frontiers | Research Topics

New insights into high-energy processes on the sun and their geospace consequences

Edited by
Xiangliang Kong, Jing Liu and Gang Li

Published in
Frontiers in Astronomy and Space Sciences

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-6030-3
DOI 10.3389/978-2-8325-6030-3

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

New insights into high-energy processes on the sun and their geospace consequences

Topic editors
Xiangliang Kong — Shandong University, Weihai, China
Jing Liu — Shandong University, China
Gang Li — Macau University of Science and Technology, SAR China

Citation
Kong, X., Liu, J., Li, G., eds. (2025). *New insights into high-energy processes on the sun and their geospace consequences*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-6030-3

Table of contents

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Editorial
PUBLISHED 10 February 2025
DOI 10.3389/fspas.2025.1560418

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Joseph E. Borovsky,
Space Science Institute (SSI), United States

*CORRESPONDENCE
Xiangliang Kong,
kongx@sdu.edu.cn

RECEIVED 14 January 2025
ACCEPTED 23 January 2025
PUBLISHED 10 February 2025

CITATION
Kong X, Liu J and Li G (2025) Editorial: New insights into high-energy processes on the Sun and their geospace consequences. *Front. Astron. Space Sci.* 12:1560418. doi: 10.3389/fspas.2025.1560418

COPYRIGHT
© 2025 Kong, Liu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: New insights into high-energy processes on the Sun and their geospace consequences

Xiangliang Kong[1,2]*, Jing Liu[1] and Gang Li[3]

[1]School of Space Science and Technology, Institute of Space Sciences, Shandong University, Weihai, Shandong, China, [2]Institute of Frontier and Interdisciplinary Science, Shandong University, Qingdao, Shandong, China, [3]State Key Laboratory of Lunar and Planetary Sciences, Macau University of Science and Technology, Taipa, Macao SAR, China

KEYWORDS

solar eruptions, solar energetic particles, x-ray and gamma-ray emissions, solar radio bursts, particle acceleration and transport, space weather, magnetosphere-ionosphere-thermosphere coupling

Editorial on the Research Topic
New insights into high-energy processes on the sun and their geospace consequences

The Sun releases an enormous amount of energy during explosive solar activities, such as solar flares and coronal mass ejections (Webb and Howard, 2012; Aschwanden et al., 2017; Benz, 2017). The solar corona can be heated up to tens of millions of degrees and a large number of charged particles can be accelerated to nearly the speed of light (Desai and Giacalone, 2016; Reames, 2017). Heated plasmas and high-energy particles increase solar radiations across the entire electromagnetic spectrum, from radio to gamma-ray wavelengths, which can have a profound effect on the Earth's upper atmosphere immediately after about 8 minutes. These create additional ionization and heating in the Earth's upper atmosphere, leading to radio blackout, GNSS signal interferences and tracking loss, increased drag on spacecraft, as well as affecting the global electric circuit (GEC), and many other phenomena (Bothmer and Daglis, 2007; Buzulukova and Tsurutani, 2022; Tacza et al., 2022). Recent studies have demonstrated that the solar flare effects can extend to the Earth's magnetosphere via electrodynamic coupling (Liu et al., 2021; Liu et al., 2024). When the high-energy particles propagate through the interplanetary medium and arrive at the vicinity of the Earth, known as solar energetic particle (SEP) events, they can pose hazardous radiation threats to astronauts and spacecraft electronics in space (Vainio et al., 2009; Shea and Smart, 2012).

This Research Topic aims to collect scientific contributions on high-energy processes on the Sun and their geospace consequences. Eight research articles and one review are contained in this electronic book, focusing on multi-wavelength observations of solar flares, acceleration and transport of energetic particles, and impacts of solar eruptions on the coupled magnetosphere–ionosphere–thermosphere system.

Qiu proposed that brightening-dimming sequences in the lower solar atmosphere can be used to identify the properties of magnetic reconnection or plasma expansion of overlying

magnetic structures in the corona. This novel method was examined by the observations of two eruptive flares.

Kuznetsov et al. presented a detailed analysis of an M-class solar flare on 2023 March 6 utilizing the combined observations in microwave and hard X-ray (HXR) from the Siberian Radioheliograph (SRH) and the HXR Imager on board the Advanced Space-based Solar Observatory (ASO-S/HXI). They further modeled the microwave emission in the 3D reconstructed flare loop using the GX simulator, which can provide spatial and spectral diagnostics of energetic electrons.

Kong et al. modeled the propagation of energetic electrons in the flare loop by numerically solving the particle transport equation. They highlighted the effects of turbulent pitch-angle scattering on the trapping/precipitation and anisotropic distribution of energetic electrons in solar flares. The simulation results help us understand the observation signatures of nonthermal HXR and microwave emissions and the excitation of coherent solar radio bursts.

Tang et al. investigated the impact of the evolution of the electron energy spectrum and velocity distribution as they propagate along the flare loop on the electron cyclotron maser instability/emission, one of the known radiation mechanisms of coherent radio bursts.

Reames reviewed how the abundance of elements can provide unique insights into the origin of SEPs in both impulsive and large gradual SEP events. The review discussed the observed properties of four sources of SEPs and the associated physical processes, including magnetic reconnection in solar jets and the seed particles for particle acceleration at CME-driven shocks.

Shalchi discussed the particle transport equation that includes the physics of pitch-angle scattering by magnetic turbulence. A new approach was developed to obtain the solutions of the equation, indicating that the two-dimensional subspace approximation is equivalent to using the telegraph equation.

Zhou et al. developed a three-component MHD model to study the evolution of the solar wind from 1 to 150 AU in the heliosphere, and the numerical simulation results were compared with the observations from New Horizons, Voyager 1 and 2. They highlighted the effects of Anomalous Cosmic Rays (ACRs) on shock-like structures of the solar wind in the outer heliosphere.

Zhang and Wang investigated a moderate geomagnetic storm in 2022 associated with a huge geohazard event of the Tonga volcano eruption. They utilized GPS total electron content (TEC) and numerical simulations from the Thermosphere Ionosphere Electrodynamic General Circulation Model (TIEGCM) to understand the plasma responses of the ionosphere-thermosphere coupled system.

Qiao et al. presented the forecast of ionospheric F2 layer (foF2) at low latitudes using deep learning models. The Informer–foF2 model demonstrates better prediction performance in predicting variations from several hours up to 48 h, compared to the widely used long short-term memory model.

Author contributions

XK: Writing–original draft, Writing–review and editing. JL: Writing–review and editing. GL: Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. XK is supported by the National Key R&D Program of China under grant 2022YFF0503002 (2022YFF0503000), the National Natural Science Foundation of China under grants 42074203.

Acknowledgments

The authors of this editorial sincerely thank all the editors, reviewers, and authors for their contributions and efforts to this Research Topic.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aschwanden, M. J., Caspi, A., Cohen, C. M. S., Holman, G., Jing, J., Kretzschmar, M., et al. (2017). Global energetics of solar flares. V. Energy closure in flares and coronal mass ejections. *V. Energy Clos. Flares Coronal Mass Ejections* 836, 17. doi:10.3847/1538-4357/836/1/17

Benz, A. O. (2017). Flare observations. *Living Rev. Sol. Phys.* 14 (2B), 2. doi:10.1007/s41116-016-0004-3

Bothmer, V., and Daglis, I. A. (2007). Space weather – physics and effects. doi:10.1007/978-3-540-34578-7.B

Buzulukova, N., and Tsurutani, B. (2022). Space weather: from solar origins to risks and hazards evolving in time. *Front. Astron. Space Sci.* 9, 429. doi:10.3389/fspas.2022.1017103

Desai, M., and Giacalone, J. (2016). Large gradual solar energetic particle events. *Living Rev. Sol. Phys.* 13, 3. doi:10.1007/s41116-016-0002-5

Liu, J., Qian, L., Wang, W., Pham, K., Kong, X., Chen, Y., et al. (2024). Energy deposition into the ionosphere during a solar flare with Extreme-ultraviolet late phase, 963. doi:10.3847/2041-8213/ad250b963L.8LL8

Liu, J., Wang, W., Qian, L., Lotko, W., Burns, A. G., Pham, K., et al. (2021). Solar flare effects in the Earth's magnetosphere. *Nat. Phys.* 17, 807–812. doi:10.1038/s41567-021-01203-5

Reames, D. V. (2017). Sol. Energ. Part. A Mod. Primer Underst. Sources, Accel. Propag 932, R. doi:10.1007/978-3-319-50871-9932

Shea, M. A., and Smart, D. F. (2012). Space weather and the ground-level solar proton events of the 23rd solar cycle, *Space Sci. Rev., Space Weather Ground-Level Sol. Prot. Events 23rd Sol. Cycle.* 171, 161–188. doi:10.1007/s11214-012-9923-z

Tacza, J., Odzimek, A., Tueros Cuadros, E., Raulin, J. P., Kubicki, M., Fernandez, G., et al. (2022). Investigating effects of solar proton events and forbush decreases on ground-level potential gradient recorded at middle and low latitudes and different altitudes. *Space weather.* 20, e2021SW002944. doi:10.1029/2021SW002944

Vainio, R., Desorgher, L., Heynderickx, D., Storini, M., Flückiger, E., Horne, R. B., et al. (2009). Dynamics of the Earth's particle radiation environment. *Earth's Part. Radiat. Environ.* 147, 187–231. doi:10.1007/s11214-009-9496-7

Webb, D. F., and Howard, T. A. (2012). Coronal mass ejections: observations. *Living Rev. Sol. Phys.* 9, 3. doi:10.12942/lrsp-2012-3

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 15 December 2023
DOI 10.3389/fspas.2023.1297632

Check for updates

OPEN ACCESS

EDITED BY
Jing Liu,
Shandong University, China

REVIEWED BY
Sampad Kumar Panda,
K L University, India
Mirela Voiculescu,
Dunarea de Jos University, Romania

*CORRESPONDENCE
Hui Wang,
h.wang@whu.edu.cn

RECEIVED 20 September 2023
ACCEPTED 05 December 2023
PUBLISHED 15 December 2023

CITATION
Zhang K and Wang H (2023), Observations and simulations of large-scale traveling ionospheric disturbances during the January 14-15, 2022 geomagnetic storm.
Front. Astron. Space Sci. 10:1297632.
doi: 10.3389/fspas.2023.1297632

COPYRIGHT
© 2023 Zhang and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Observations and simulations of large-scale traveling ionospheric disturbances during the January 14-15, 2022 geomagnetic storm

Kedeng Zhang[1,2] and Hui Wang[1]*

[1]Department of Space Physics, Hubei Luojia Laboratory, School of Electronic Information, Wuhan University, Wuhan, China, [2]State Key Laboratory of Space Weather, Chinese Academy of Sciences, Beijing, China

Using the total electron content (TEC) observations from GPS, and simulations from the Thermosphere Ionosphere Electrodynamic General Circulation Model (TIEGCM), this work investigates the large-scale traveling ionospheric disturbances (LSTIDs) and the possible involved drivers during the geomagnetic storm on January 14-15, 2022. Based on the term analysis of O^+ continuity equation in TIEGCM, it is found that the traveling atmospheric disturbances in equatorward winds are responsible for the LSTIDs, with minor contributions from plasma drifts owing to the prompt penetration electric field. A strong interhemispheric asymmetry of the LSTIDs is observed, which might be attributed to both the equatorward wind disturbances and background plasma. The stronger wind (plasma) disturbances occurs in the winter hemisphere than that in the summer hemisphere. The maximum magnitude of LSTIDs in electron density disturbances occurs at ~250 and ~270 km in the northern and southern hemispheres, respectively, owing to both the thermospheric equatorward winds and background plasma. An interesting phenomenon that tail-like LSTIDs occur at the dip equator and low latitudes might be related to the eruption of the Tonga volcano, but it is not well reproduced in TIEGCM that deserves further exploration in a future study.

KEYWORDS

large-scale traveling ionospheric disturbances, interhemispheric asymmetry, GPS-observed TEC, TIEGCM simulations, equatorward winds

1 Introduction

During the disturbance periods, the interaction between the interplanetary magnetic field (IMF) carried by the solar wind and the geomagnetic field could lead to the energy and momentum deposition from the solar wind to the Earth's upper thermosphere (e.g., Dungey, 1961). A large amount of energy and momentum deposition triggers disturbances in the thermo-sphere, traveling to middle and low latitudes from the source region. The thermospheric wind perturbations associated with the generated traveling atmospheric disturbances (TADs) could push the ionospheric plasma upward or downward along the geomagnetic field lines, causing the enhancement or depletion in plasma, referred to as large-scale traveling ionospheric disturbances (TIDs). TADs/TIDs appear as wave-like perturbations in thermospheric/ionospheric observations, i.e., thermospheric meridional winds and density, ionospheric plasma. Over the past decades, a large variety of TIDs has been observed

in a series of studies (e.g., Munro, 1958; Hocke et al., 1996; Balthazor et al., 1997; Shiokawa et al., 2007; 2013; MacDougall and Jayachandran, 2011; Yin et al., 2019; Zhang K et al., 2019; Zhang SR et al., 2019; Nishimura et al., 2020), which is a hot topic in the ionospheric research.

Based on the wavelength, phase speed, and period, TIDs could be categorized as large-scale and medium-scale (Bruinsma et al., 2009; Shiokawa et al., 2013). The large-scale TIDs (LSTIDs) have a period longer than 60 min, a horizontal velocity of 400–1,000 m/s, and a wavelength larger than 1,000 km, while the medium-scale TIDs have a period ranging from 15 min to 1 h, a horizontal velocity of 250–1,000 m/s and a wavelength of hundreds of kilometers (Afraimovich et al., 2000; Zhang SR et al., 2019). Previous studies have demonstrated that LSTIDs in the upper thermosphere could be generated by the geomagnetic activity (i.e., geomagnetic storm, and substorm) (Pi et al., 2000; Shiokawa et al., 1999; 2003; 2007; Afraimovich et al., 2008; Lei et al., 2008; Nicolls et al., 2012; Borries et al., 2016; Cherniak et al., 2018; Katamzi-Joseph et al., 2019; Zhang SR et al., 2019; Jonah et al., 2020).

In the literature, LSTIDs have been observed at high latitudes (Pi et al., 2000; Shiokawa et al., 2003; Nicolls et al., 2012). Using Fabry-Perot interferometer (FPI) observed neutral winds and incoherent scatter radar measured plasma drift at high latitudes in October 1992 for ~36 h, Pi et al. (2000) reported outstanding evidence for TADs and LSTIDs generated by the aurora heating effects. Utilizing a series of measurements at high latitudes, including neutral winds from FPI and TEC from GPS, Shiokawa et al. (2003) found prominent LSTIDs during a major storm event with a minimum DST index of −358 nT on 31 March 2001. A turning of the thermospheric meridional winds from equatorward of 94 m/s to poleward of 44 m/s was observed, indicating an intense poleward wind in the thermosphere during the LSTIDs. The generation of poleward wind in the auroral zone could be associated with an intense substorm. Analyzing the FPI-measured thermospheric wind for atmospheric gravity waves (AGWs) over Alaska on January 9-10, 2010, Nicolls et al. (2012) reported an event of AGWs with a period of 32.7 ± 0.3 min, a horizontal wavelength of 1,094 ± 408 km, the phase speed of 560 ± 210 m/s, propagation azimuth of 33.5 ± 15.8° east-of-north, which was associated with enhanced auroral activity and the potential sources might be Joule heating, Lorentz force, and body forcing of horizontal winds due to auroral activity, etc.

Apart from high-latitude LSTIDs cases, in recent decades, a large number of studies have been performed to investigate the LSTIDs at middle and low latitudes (e.g., Shiokawa et al., 2002; Afraimovich et al., 2008; Lei et al., 2008; Jonah et al., 2020; Nishimura et al., 2020). Using a comprehensive TEC observation from GPS and simulation from the Sheffield University Plasmasphere-Ionosphere Model (SUPIM), Shiokawa et al. (2002) investigated a prominent LSTIDs detected at 23-24 LT (14-15 UT) during the magnetic storm on 15 September 1999, in Japan (16°–37° magnetic latitudes, MLat). The nighttime LSTIDs might be generated by the enhancement of poleward neutral winds which propagates equatorward. The SUPIM results suggested that the equatorward movement of poleward wind pulse was linked to the auroral energy input. Afraimovich et al. (2008) compared the intensity of LSTIDs with the local electron density disturbances during the magnetic storm on October 29–31, 2003, and November 7–11, 2004. They found that LSTIDs in TEC were dominated by the auroral energy, and the TEC variations were mainly attributed to the electron density disturbances at F2-layer. A numerical simulation from the coupled magneto-sphere-ionosphere model (CMIT) was performed and TEC data from GPS in Japan were analyzed to explore prominent northward LSTIDs and two southward LSTIDs events during the magnetic storm on 15 December 2006 (Lei et al., 2008). In their results, the northward LSTIDs were generated in the southern hemisphere which propagated into the northern hemisphere, however, two southward LSTIDs were not well captured in the model. A series of LSTIDs in ionospheric TEC perturbations were generated because of the intense magnetic storm on 17 March 2015 (Borries et al., 2016), which was induced by the Joule heating in the auroral region, Lorentz force from the perturbed electric fields, and a minor particle precipitation effect. The origin, occurrence, and propagation of LSTIDs over the European on December 19–21, 2015 were investigated in Cherniak and Zakharenkova (2018). Two major sources of the LSTIDs, the quiet-time solar terminator passage, and disturbed-time auroral activity have been reported. Using ground- and space-based measurements (i.e., Global Navigation Satellite System receivers, and Swarm satellites) and TIEGCM simulations during two magnetic storm periods, Jonah et al. (2020) found that the meridional winds (background ionospheric plasma) played important roles in the propagation (amplitude) of LSTIDs at middle and low latitudes. Smaller electron density was correlated with the smaller amplitude of LSTIDs, and *vice versa*.

Recently, a moderate geomagnetic storm occurs on January 14-15, 2022. Furthermore, a huge geohazard event of the Tonga volcano eruption [−20.5° geographic latitudes (GLat), −174.5° geographic longitudes (GLon)] has great effects on the ionosphere-thermosphere coupled system. This might complicate the ionospheric plasma responses. In this work, the TEC data from GPS, and numerical simulations from TIEGCM are used to investigate the ionospheric plasma responses during the moderate geomagnetic storm, which could contribute to the understanding of the coupling between high- and low-latitudes, between ionosphere and thermosphere. In the rest of this article, Section 2 introduces the GPS data and the model; Section 3 gives the results of LSTIDs and data-model comparison; Section 4 is the discussion about potential physical mechanisms; Section 5 summaries the results of this work.

2 Data and model

The GPS TEC dataset is obtained from the International Global Navigation Satellite System Service (IGS) Working Group, which is created in 1998 (Hernández-Pajares et al., 2009; Panda et al., 2022). The individual TEC maps are developed by eight Ionospheric Associate Analysis Centers (IAACs) under IGS. For example, the University of Bern (CODE, Switzerland), the European Space Agency (ESA, Germany), the Jet Propulsion Laboratory (JPL, America), and the University Politechnical Catalonia (UPC, Spain). The TEC data from IGS is the weighted mean of the eight analysis centers. The resolution of GPS TEC data in CDF format is 15 min, 1 and 2 h. A large amount of daily TEC files since 15 January 1998, is stored in IGS. In this work, TEC data with a resolution of 15 min on January 9-16, 2022 is used to explore the ionospheric disturbances during geomagnetic activity periods. The

data could be downloaded from the link: https://cdaweb.gsfc.nasa.gov/pub/data/gps/tec15min_igs/2022/.

The Thermosphere Ionosphere Electrodynamic General Circulation Model (TIEGCM) is a three-dimensional time-dependent model of the coupled ionosphere-thermosphere system. It was developed at the High Altitude Observatory at the National Center for Atmospheric Research (NCAR/HAO). The driver of TIEGCM includes the high-latitude electric field from Heelis or Weimer models (Heelis et al., 1982; Weimer, 2005), the solar extreme ultraviolet and ultraviolet spectral fluxes that were parameterized by the $F_{10.7}$ index (Richards et al., 1994), the lower atmospheric migrating and non-migrating diurnal and semi-diurnal tides generated from the Global Scale Wave Model (GSWM) (Hagan and Forbes, 2002; 2003), or tides derived from the observations from the Sounding of the Atmosphere using Broadband Emission Radiometry and TIDI (Zhang et al., 2018; Wu et al., 2019). The horizontal resolution is 2.5° GLat × 2.5° GLon. TIEGCM has a total of 57 pressure levels in the vertical direction, with the lower and upper boundary of 97 km and ~700 km (depending on the solar activity), respectively. In this study, the migrating and non-migrating tides from GSWM were specified at the lower boundary and the electric field from the Weimer model was imposed at high latitudes.

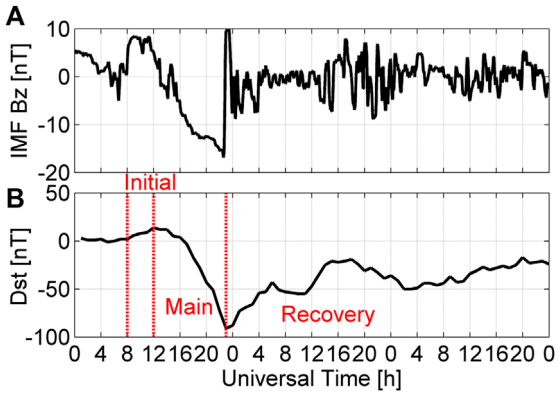

FIGURE 1
The temporal variations of IMF Bz **(A)** and Dst index **(B)** on January 14–16, 2022. The text of "Initial," "Main" and "Recovery" in the bottom panel are the initial, main, and recovery phases of the magnetic storm, respectively.

3 Results

3.1 Geomagnetic conditions

Figure 1A depicts the temporal variations of IMF Bz during the magnetic storm on January 14–16, 2022. It is found that IMF Bz is northward at 00–12 UT on January 14, with a maximum magnitude of 8.3 nT. At ~12 UT on January 14, the northward IMF Bz starts to turn southward. The southward turning of IMF Bz arrives at its minimum value of −16.8 nT at 23 UT. Then, the strong southward IMF Bz quickly turns northward for a magnitude of ~10 nT in 30 min. After that, the temporal variations of IMF Bz oscillate around 0 nT, with an absolute maximum magnitude of ~8 nT. Figure 1B shows the temporal variations of Dst index on January 14–16, 2022. Based on the Dst index, the magnetic storm can be characterized by three phases, that is, initial, main, and recovery phases. During the initial phase of 08–12 UT on January 14, the Dst index is enhanced from 2 nT to 14 nT. During the main phase of 12–23 UT, the Dst index is significantly decreased to −91 nT. Then, during the recovery phase, the Dst index gets smoothly recovery to ~−20 nT.

3.2 Data-model comparison

Figures 2A, B show the geomagnetic latitude and UT variations of TEC observed from GPS and modeled by TIEGCM at −174.5° geographic longitude (GLon) on January 14–16, 2022. Note here that Tonga volcanic eruption occurs at −174.5° GLon. The data be-fore 14 UT on January 14 is not shown here, because the Dst index in Figure 1B starts to decrease to the negative peak at around 14 UT. The prominent feature of TEC is the strong equatorial ionization anomaly (EIA), which has been reported in a series of previous studies (e.g., Lin et al., 2005; Rajesh et al., 2021). During quiet time, EIA is the region between ±10° and ±15° magnetic latitude (MLat) across the magnetic equator (center) (e.g., Panda et al., 2018; Rajesh et al., 2021; Ogwala et al., 2022). As shown in Figure 2A, during the disturbed time, EIA in both hemispheres expands to higher latitudes, even to middle and high latitudes. The poleward edge could be seen at around ±60° MLat and 00 UT/12 LT on January 15. The maximum amplitude of TEC in the EIA region could reach ~50 TECU at ±15° MLat and 00 UT/12 LT on January 15. At the dip equator, TEC is much weaker, with a value of ~30 TECU, than that in the EIA region. These are the well-known two peaks of EIA (Lin et al., 2005). At the pre-dawn sector of 14–18 UT/02–06 LT on January 14, the GPS-observed TEC has an average value of ~10 TECU, and the significant EIA has not been developed. With the onset of sunrise, the EIA begins to develop, and the maximum EIA occurs at around noon of 00–02 UT/12–14 LT on January 15. After the sunset of 08 UT/20 LT on January 15, the significant EIA disappears. During the daytime from 18 UT on January 15 to 08 UT on January 16, a similar prominent structure can be seen. An interesting phenomenon is found when the Tonga volcano eruption occurs, as indicated by the black star. A northward penetration of LSTID can be found at the post-dusk sector of 04–12 UT on January 15 after the eruption of Tonga volcano (black star), which seems like a tail-like structure following the EIA and deserves to explore. This tail-like structure disappears on January 16.

In Figure 2B, TIEGCM-modeled TEC also has a prominent character of EIA during daytime. A comparison between Figures 2A, B shows that the large-scale structure of TEC is similar to those two. The modeled EIA also expands to middle latitudes of ~±60° MLat in both hemispheres, and the peaks of TEC also occur at around 02 UT/14 LT with a magnitude of ~30 TECU. Compared to GPS-observed TEC, TEC in TIEGCM seems to be underestimated, which has been reported before (Shiokawa et al., 2007; Perlongo and Ridley, 2018) and might be attributed to the following potential reasons. For example, first, Joule heating tends to be underestimated in most large-scale models including TIEGCM due to the inability to capture small-scale features (Shiokawa et al., 2007). Second,

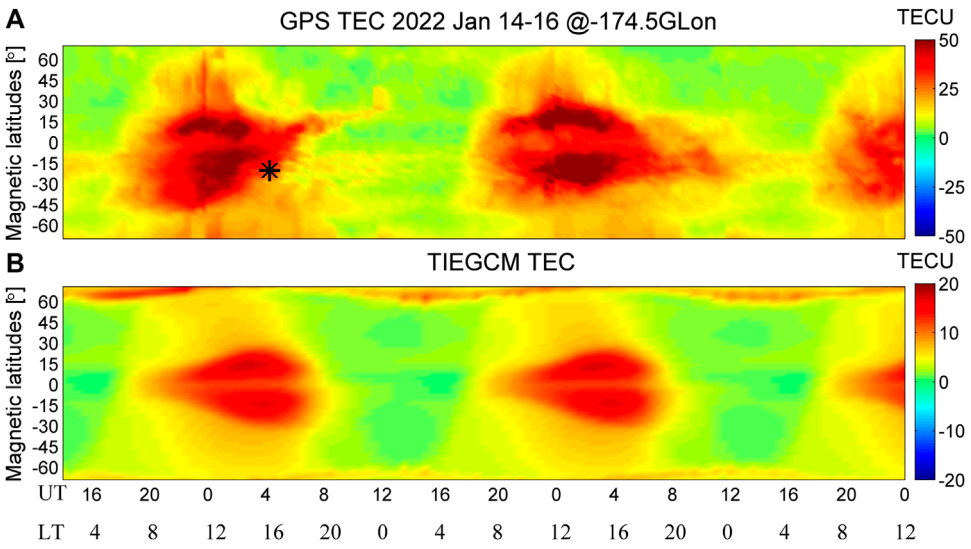

FIGURE 2
The geomagnetic latitude and UT variations of GPS observed TEC **(A)** and TIEGCM modeled TEC **(B)** at −174.5° GLon on January 14–16, 2022. TEC is given in TECU. The black star in Figure 2A is the time and location of the Tonga volcano eruption.

the supply of O^+ ions from the plasmasphere is underestimated in TIEGCM (Shiokawa et al., 2007). Third, the neutral winds are also underestimated in TIEGCM, as reported in previous studies (Perlongo and Ridley, 2018; Zhang et al., 2018). Fourth, the high-latitude electric field used in TIEGCM is an empirical model, which predicts only the state of plasma convection at high latitudes for a given 3-hourly Kp index or 1-min IMF, whereas the real high-latitude ion convection is much more complicated (Zhang et al., 2021). However, the large-scale structures of modeled TEC are similar to that of observed TEC, and a large degree of similarity between TIEGCM simulations and space-based/ground-based observations has been achieved in previous studies (Wang et al., 2012; Perlongo et al., 2018; Zhang et al., 2018; 2021). In summary, the reliability and stability of TIEGCM have been confirmed. Thus, it can be used to explore the ionospheric responses during the disturbed time in this work.

Representing the data-model comparison of the absolute vertical TEC, the ionospheric disturbances in TEC at −174.5° GLat on January 14–16 are shown in Figure 3. Note that ΔTEC in logarithm based on 10 is obtained from the ratio between the storm-time TEC on January 14–16 and background quiet-time average TEC on January 9–13. In Figure 3, at 04–16 UT/16-04 LT on January 15, an outstanding negative storm effect occurs at 15°–30° MLat in the southern hemisphere. The decrease of TEC might be caused by the changes in neutral composition owing to the thermospheric heating (Liu et al., 2014). The upwelling of molecular-rich air due to vertical advection at high latitudes would lead to a decrease in neutral composition in the ionosphere, then driven by the equatorward winds, the disturbance zone of O/N_2 would travel to lower latitudes. The decreases in O/N_2 produce the corresponding depletion in electron density. This TEC depletion follows the eruption of Tonga volcano. During the eruption, the generated gravity and lamb waves might release great energy into the ionosphere and thermosphere, causing disturbances in thermospheric winds (Harding et al., 2022; Zhang K et al., 2022; Zhang SR et al., 2022). Considering the reduction of solar radiation during nighttime, the transport effects due to disturbances in thermospheric winds might lead to the decrease in TEC, which deserves a further exploration. As indicated by two black arrows, two LSTIDs events are identified. LSTIDs in the northern (southern) hemisphere have a phase speed of ~411 (~463) m/s, consistent with previous studies (Bruinsma et al., 2009; Shiokawa et al., 2013; Zhang et al., 2019). An interesting LSTID is also observed after the onset of the Tonga volcano eruption, as indicated by the magenta arrow. This tail-like structure follows the EIA expansion during the daytime (Figure 2A). It has a phase speed of 347 m/s, which has been disclosed using GNSS TEC data (Zhang SR et al., 2022).

Previous studies had disclosed that the variations in TEC were mainly controlled by the changes of electron density at an altitude of the highest density (hmF_2) (Liu et al., 2016). The hmF_2 has an average value of ~300 km in this work (Figures not shown). Thus, the UT versus magnetic latitudes of ΔNe from TIEGCM at −174.5° GLon and ~300 km is shown in Figure 4. LSTIDs in ΔNe can be found at 22-02 UT and middle latitudes in both hemispheres, as indicated by two black arrows. Comparing Figures 3, 4, we can find that two LSTIDs in ΔTEC are also reproduced in TIEGCM. The large-scale similarity of the equatorward traveling of LSTIDs between modeled and observed results is achieved, which ensures the reliability of TIEGCM in capturing the LSTIDs. However, the tail-like LSTIDs are not well reproduced in TIEGCM, because TIEGCM does not include the effects of huge geohazard events, i.e., a violent volcano eruption.

4 Discussion

Two LSTIDs have been observed in TEC observations from IGS, and confirmed in TEC and Ne simulations from TIEGCM. Previous

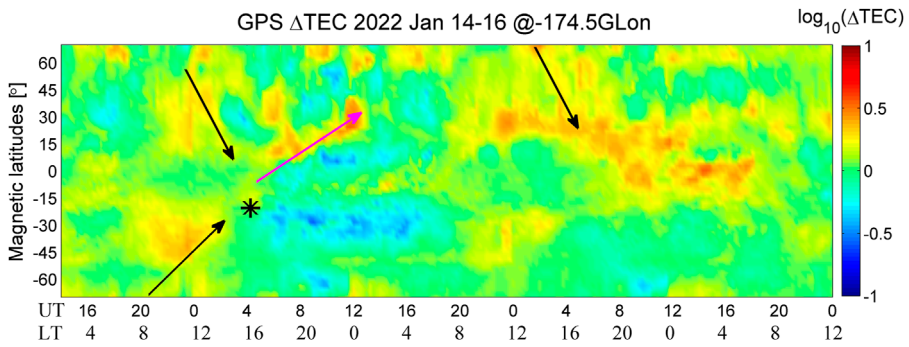

FIGURE 3
UT and MLat variations of GPS-observed residual TEC (ΔTEC). The residual TEC is obtained by a ratio between TEC on January 14–16 and the average TEC on January 9–13. ΔTEC is given in logarithm based on 10. The black and magenta arrows represent the LSTIDs.

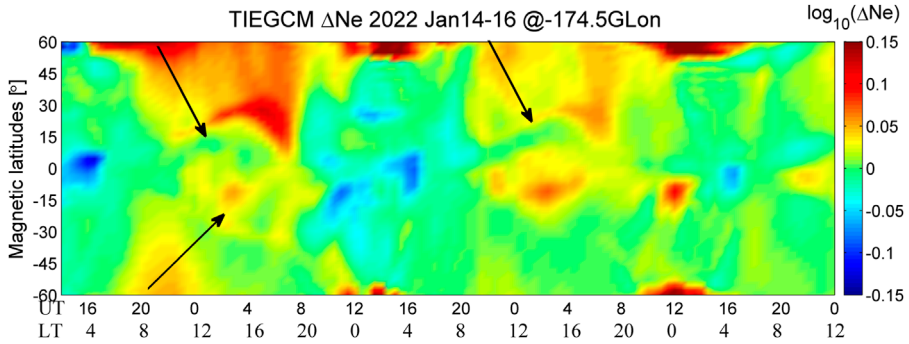

FIGURE 4
Similar to Figure 3, but for the TIEGCM modeled residual Ne (ΔNe) at 300 km.

studies have reported that large-scale ionospheric disturbances might be controlled by several forces, i.e., electric field, auroral heating, and neutral winds (e.g., Shiokawa et al., 2007; Katamzi-Joseph et al., 2019). Which one might be responsible for these two LSTIDs during storm periods? It is still unknown. Using TIEGCM, the potential drivers of two LSTIDs and their interhemispheric asymmetry have been disclosed in this work.

4.1 Term analysis of O^+ continuity

Similar to the method used in previous studies (Liu et al., 2016; Qian et al., 2016; Zhang et al., 2021), a term analysis of the ionospheric O^+ continuity equation (see Eq. 1) was performed in this work, to determine the relative contributions from neutral winds, chemical processes (including chemical production and loss rate), plasma drifts, and ambipolar diffusion.

$$\frac{\partial N}{\partial t} = q - \beta N - \nabla \bullet (NV) \quad (1)$$

where $\frac{\partial N}{\partial t}$, q, βN, and $\nabla \bullet (NV)$ are the change rate of ionospheric main ions of O^+, chemical production rate, loss rate, and effects from transport processes (including neutral winds, E × B drifts, and ambipolar diffusion), respectively. During the magnetic storm on January 14–16, 2022, the influences from chemical processes (both chemical production and loss rate) and ambipolar diffusion on O^+ changes in the formation of LSTIDs are much weaker than that from E× B drifts and neutral winds (Figures not shown). Thus, only effects from E × B drifts and neutral winds are investigated in the following.

Figure 5 shows the UT versus magnetic latitudes of total O^+ changes due to all forcing terms, E × B drifts, and neutral winds at ~300 km and −174.5° GLon. In Figure 5A, the O^+ changes due to forcing terms at middle latitudes in both hemispheres also show similar structures with LSTIDs in ΔNe at the end of January 14. The total O^+ changes in the traveling path of LSTIDs in the northern hemisphere have an average value of ~5 $cm^{-3}s^{-1}$. In the southern hemisphere, the average value of the total O^+ changes is weaker (~2 $cm^{-3}s^{-1}$) than that in the northern hemisphere. At the end of January 15, similar LSTIDs in total O^+ changes at middle latitudes in the northern hemisphere can be found, ensuring the occurrence of LSTIDs in ΔNe.

Previous studies have reported the effects of Lorenz force due to the penetration of electric field on the equatorward LSTIDs (Borries et al., 2016). Figure 5B depicts the effects of E × B drifts on the O^+ changes. It can be found that the average value of the O^+ changes is ~3 and ~2 $cm^{-3}s^{-1}$ in the LSTIDs in the northern and southern hemispheres, respectively. However, there is not a significant time delay with respect to latitudes. The disturbances of

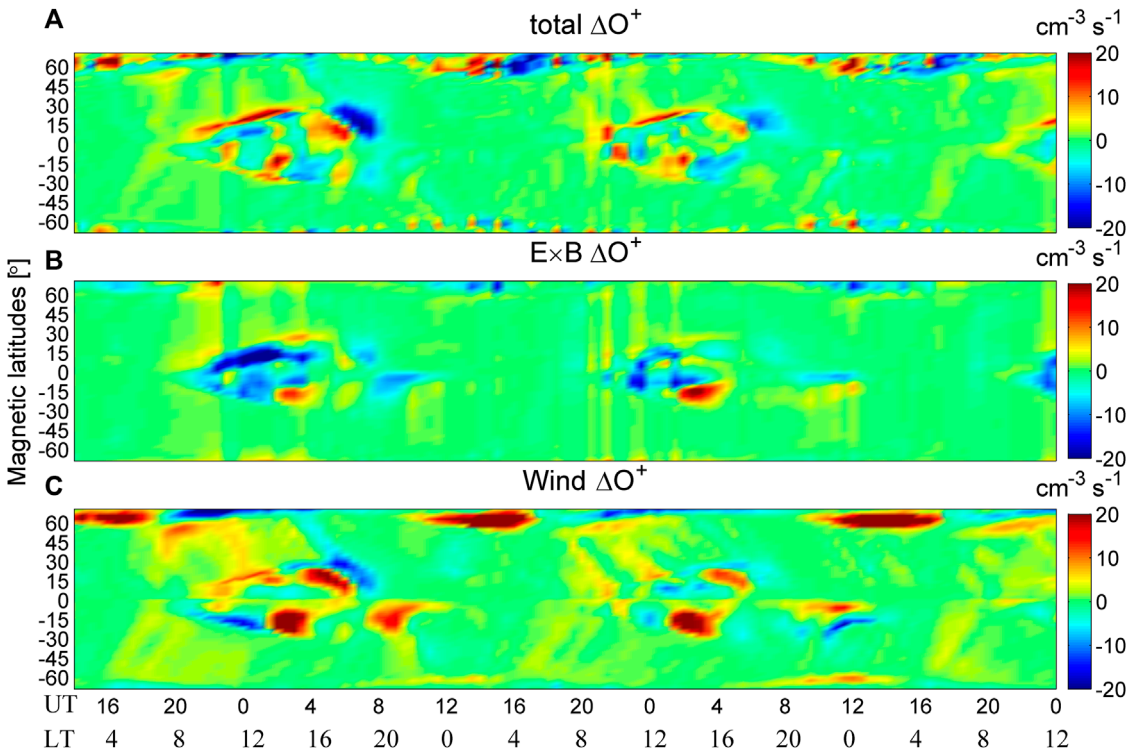

FIGURE 5
The magnetic latitudes and UT variations of total residual O^+ (ΔO^+) **(A)**, ΔO^+ due to E × B **(B)**, and ΔO^+ due to neutral winds **(C)** at ~300 km and −174.5° GLon on January 14–16, 2022. ΔO^+ is obtained by removing the background quiet-time O^+. ΔO^+ is given in $cm^{-3}s^{-1}$.

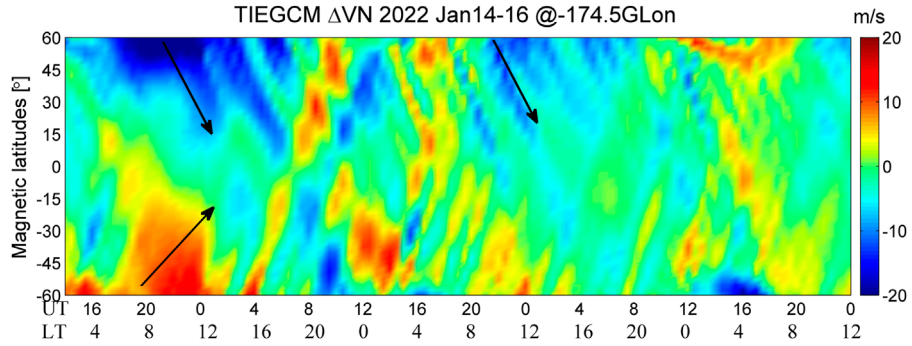

FIGURE 6
UT and MLat variations of TIEGCM-modeled meridional wind disturbances (ΔVN). ΔVN is the difference between VN during the disturbed and quiet time. ΔVN is given in m/s. Positive value stands for northward winds.

O^+ due to E × B drifts occur simultaneously at almost all latitudes (Zhang et al., 2019). Thus, we can conclude here that E × B drifts play negligible roles in the equatorward propagation of LSTIDs, but could contribute to the enhancement of ΔNe (Figure 5B). The ΔO^+ enhancement owing to the plasma transport from E × B drifts supports the occurrence of LSTIDs.

The neutral winds play important roles in the vertical movement of ionospheric plasma (Liu et al., 2016; Zhang et al., 2019). Because thermospheric winds could move the charged ions upward/downward along the geomagnetic field lines, causing the enhancement/depletion of ionospheric plasma due to the chemical recombination (Rishbeth, 1967; Zhang et al., 2012). To disclose the roles of neutral winds, Figure 5C shows the O^+ changes due to neutral winds at −174.5° GLon and ~300 km during the disturbed time. In Figures 5A, C corresponding O^+ enhancement due to neutral winds occurs at the traveling path of LSTIDs in both hemispheres. The mean value of LSTIDs in O^+ changes is approximately 6 and 3 $cm^{-3}s^{-1}$ in the northern and southern hemispheres, respectively. A comparison between Figures 5A, C indicates that the LSTIDs in ΔNe is dominated by thermospheric winds. Figure 6 illustrates the UT and magnetic latitude variations of meridional wind disturbances at −174.5° GLon and ~300 km during disturbed periods. As indicated by three black arrows, the LSTADs in the equatorward winds can

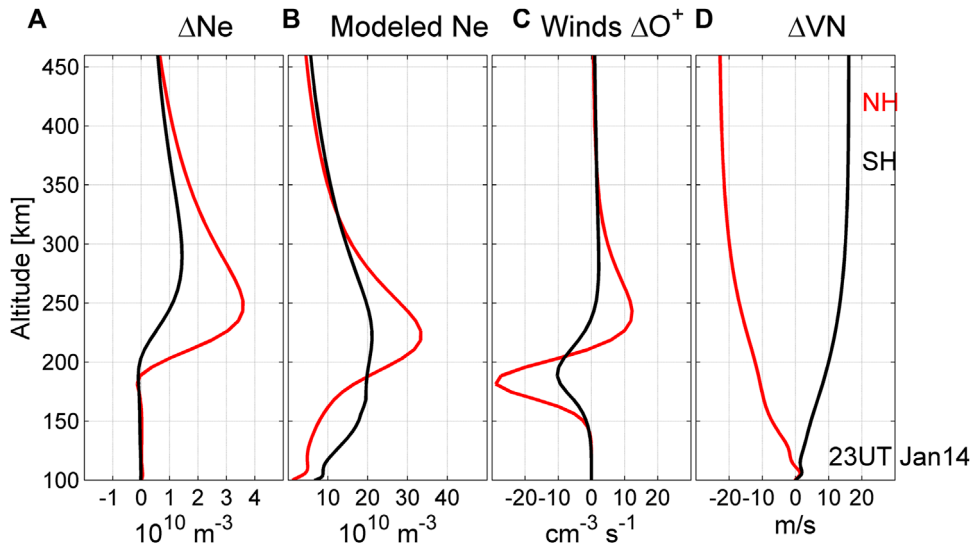

FIGURE 7
The vertical profile of ΔNe **(A)**, modeled Ne **(B)**, ΔO^+ due to neutral winds **(C)**, and meridional wind disturbances [**(D)**, ΔVN] at −174.5° GLon and 23 UT on 14 January 2022. The red and black lines represent northern and southern hemispheres, respectively.

be found. The average speed of LSTADs at the end of January 14 in the northern and southern hemispheres is ~30 and ~25 m/s, respectively. The magnitude of LSTADs at the end of January 15 in the northern hemisphere is smaller (~20 m/s) than that at the end of January 14.

The interhemispheric asymmetry of LSTIDs might be attributed to the meridional wind disturbances, which are stronger in the northern hemisphere than that in the southern hemisphere. The vertical plasma drifts due to meridional winds are expressed as follows (Eq. 2; Zhang et al., 2012):

$$VV = v_n cosDcos|I| \sin|I| \quad (2)$$

where VV, v_n, D and I are plasma drifts, thermospheric meridional winds, magnetic declination, and inclination, respectively. In general, the stronger meridional wind disturbances might mean the faster vertical plasma drifts. Thus, owing to the faster equatorward winds, more plasma at lower altitudes can be transported to higher altitudes where the chemical recombination could be slower, generating a stronger disturbance in ionospheric plasma density. Thus, the magnitude of LSTIDs at the end of January 14 is stronger in the northern hemisphere than that in the southern hemisphere. Furthermore, the disturbances in meridional winds at the end of January 15 in the southern hemisphere are not strong enough to push enough plasma up to higher altitudes, hence the LSTIDs do not occur.

4.2 Vertical profile

Previous studies have reported the features of LSTIDs observed at different altitudes, e.g., incoherent scatter radar measurements at ~325 km, Swarm-observed plasma at ~460 and ~540 km, CHAMP-observed electron density at ~400 km, and GPS-observed TEC (Shiokawa et al., 2002; Lei et al., 2008; Borries et al., 2016). As discussed before, the thermospheric meridional winds are responsible for the LSTIDs in both hemispheres via moving plasma upward along the geomagnetic field lines. In Figure 4, LSTIDs in ΔNe at ~23 UT on January 14 are located at approximately ±40°~±50° MLat. To investigate the altitudinal variations, Figure 7 gives the vertical mean profile of ΔNe, modeled Ne, ΔO^+ due to neutral winds, and meridional wind disturbances (ΔVN) at 23 UT within ±40° to ±50° MLat. In Figure 7A, ΔNe is enhanced at altitudes higher than 200 km. The maximum intensity of ΔNe is ~3.5×10^{10} m^{-3} and ~1.4×10^{10} m^{-3} in the northern and southern hemispheres, respectively. The altitude of maximum ΔNe is at ~250 (~270) km in the northern (southern) hemisphere, which might be attributed to the equatorward winds and background plasma (Afraimovich et al., 2008). As shown in Figure 7B, the background plasma density is the strongest at ~230 and ~240 km in the northern and southern hemispheres, respectively. A stronger background plasma could generate stronger LSTIDs in ΔNe (Ding et al., 2007), hence the strongest LSTIDs in ΔNe at ~250 and ~270 in the northern and southern hemispheres. During the disturbed period, ΔO^+ due to neutral winds at middle latitudes is negative at low altitudes below 220 km, and positive at high altitudes >220 km (Figure 7C). Owing to the thermospheric equatorward winds (Figure 7D), the ionospheric charged ions are moved upward along the geomagnetic field lines to a higher altitude with slower chemical recombination (Zhang SR et al., 2022), resulting in an enhancement of plasma at a higher altitude above ~220 km. Furthermore, the altitude differences between the maxima of ΔNe and background plasma is ~20 and 30 km in the northern and southern hemispheres, respectively. This interesting phenomenon can be attributed to the wind transport effects.

An interhemispheric asymmetry also occurs in the altitudinal profile of LSTIDs in ΔNe. This might be related to two potential reasons. One is the equatorward wind disturbances, which shows interhemispheric asymmetry (Figure 7D). The other might be the stronger background plasma in the northern hemisphere than that in the southern hemisphere

(Figure 7B). Owing to both the stronger ΔVN and background plasma, stronger LSTIDs occur in the northern hemisphere than that in the southern hemisphere (Zhang SR et al., 2012). The altitude differences between the maximum ΔNe and background plasma is smaller in the southern hemisphere than that in the northern hemisphere.

5 Conclusion

Using observations from GPS, and numerical simulations from TIEGCM, the LSTIDs in ΔNe and its interhemispheric asymmetry and altitudinal profile are investigated. Several interesting results are found.

1. The TIEGCM simulations show that the thermospheric equatorward winds are responsible for the LSTIDs. The interhemispheric asymmetry in plasma disturbances is related to the corresponding asymmetry in the meridional wind disturbances.
2. The vertical profile of ΔNe is also shown in this work, which is attributed to both the background plasma and transport effects from equatorward winds. The interhemispheric asymmetry also occurs in the altitudinal profile and is attributed to two factors: the background electron density and equatorward winds.
3. A tail-like LSTIDs is shown in the GPS-observed TEC after the eruption of the Tonga volcano, however, it is not reproduced by TIEGCM. The potential reason might be the huge geohazard event that is not included in the numerical physical model, and deserve further exploration in the future.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

KZ: Data curation, Writing–original draft. HW: Writing–review and editing.

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. We are grateful for the sponsor from the National Natural Science Foundation of China Basic Science Center (42188101), the Fundamental Research Funds for the Central Universities (2042023kf0099), the National Nature Science Foundation of China (No. 41974182 and 42122031). This work is supported by Hubei Provincial Natural Science Foundation of China (2023AFB616). This work is also supported by the Project Supported by the Specialized Research Fund for State Key Laboratories.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Afraimovich, E. L., Kosogorov, E. A., Leonovich, L. A., Palamartchouk, K. S., Perevalova, N. P., and Pirog, O. M. (2000). Determining parameters of large-scale traveling ionospheric disturbances of auroral origin using GPS-arrays. *J. Atmos. Solar-Terrestrial Phys.* 62 (7), 553–565. doi:10.1016/s1364-6826(00)00011-0

Afraimovich, E. L., Voeykov, S. V., Perevalova, N. P., and Ratovsky, K. G. (2008). Large-scale traveling ionospheric disturbances of auroral origin according to the data of the GPS network and ionosondes. *Adv. Space Res.* 42 (7), 1213–1217. doi:10.1016/j.asr.2007.11.023

Balthazor, R. L., and Moffett, R. J. (1997). A study of atmospheric gravity waves and travelling ionospheric disturbances at equatorial latitudes. *Ann. Geophys.* 15 (8), 1048–1056. doi:10.1007/s00585-997-1048-4

Borries, C., Mahrous, A. M., Ellahouny, N. M., and Badeke, R. (2016). Multiple ionospheric perturbations during the Saint Patrick's Day storm 2015 in the European-African sector. *J. Geophys. Res. Space Phys.* 121 (11), 11–333. doi:10.1002/2016ja023178

Bruinsma, S. L., and Forbes, J. M. (2009). Properties of traveling atmospheric disturbances (TADs) inferred from CHAMP accelerometer observations. *Adv. Space Res.* 43 (3), 369–376. doi:10.1016/j.asr.2008.10.031

Cherniak, I., and Zakharenkova, I. (2018). Large-scale traveling ionospheric disturbances origin and propagation: case study of the December 2015 geomagnetic storm. *Space weather* 16 (9), 1377–1395. doi:10.1029/2018sw001869

Ding, F., Wan, W., Ning, B., and Wang, M. (2007). Large-scale traveling ionospheric disturbances observed by GPS total electron content during the magnetic storm of 29–30 October 2003. *J. Geophys. Res. Space Phys.* 112 (6), 309. doi:10.1029/2006ja012013

Dungey, J. W. (1961). Interplanetary magnetic field and the auroral zones. *Phys. Rev. Lett.* 6 (2), 47–48. doi:10.1103/physrevlett.6.47

Hagan, M. E., and Forbes, J. M. (2002). Migrating and nonmigrating diurnal tides in the middle and upper atmosphere excited by tropospheric latent heat release. *J. Geophys. Res. Atmos.* 107 (24), 4754. doi:10.1029/2001jd001236

Hagan, M. E., and Forbes, J. M. (2003). Migrating and nonmigrating semidiurnal tides in the upper atmosphere excited by tropospheric latent heat release. *J. Geophys. Res. Space Phys.* 108 (2), 1062. doi:10.1029/2002ja009466

Heelis, R. A., Lowell, J. K., and Spiro, R. W. (1982). A model of the high-latitude ionospheric convection pattern. *J. Geophys. Res. Space Phys.* 87 (8), 6339–6345. doi:10.1029/ja087ia08p06339

Hernández-Pajares, M., Juan, J. M., Sanz, J., Orus, R., Garcia-Rigo, A., Feltens, J., et al. (2009). The IGS VTEC maps: a reliable source of ionospheric information since 1998. *J. Geodesy* 83, 263–275. doi:10.1007/s00190-008-0266-1

Hocke, K., and Schlegel, K. (1996). A review of atmospheric gravity waves and travelling ionospheric disturbances: 1982–1995. *Ann. Geophys.* 14 (9), 917. doi:10.1007/s005850050357

Jonah, O. F., Zhang, S., Coster, A. J., Goncharenko, L. P., Erickson, P. J., and Rideout, W., (2020). Understanding inter-hemispheric traveling ionospheric disturbances and their mechanisms. *Remote Sens.* 12 (2), 228. doi:10.3390/rs12020228

Katamzi-Joseph, Z. T., Aruliah, A. L., Oksavik, K., Habarulema, J. B., Kauristie, K., and Kosch, M. J. (2019). Multi-instrument observations of large-scale atmospheric gravity waves/traveling ionospheric disturbances associated with enhanced auroral activity over Svalbard. *Adv. Space Res.* 63 (1), 270–281. doi:10.1016/j.asr.2018.08.042

Lei, J., Burns, A. G., Tsugawa, T., Wang, W., Solomon, S. C., and Wiltberger, M. (2008). Observations and simulations of quasiperiodic ionospheric oscillations and large-scale traveling ionospheric disturbances during the December 2006 geomagnetic storm. *J. Geophys. Res. Space Phys.* 113 (6), 310. doi:10.1029/2008ja013090

Lin, C. H., Richmond, A. D., Heelis, R. A., Bailey, G. J., Lu, G., Liu, J. Y., et al. (2005). Theoretical study of the low-and midlatitude ionospheric electron density enhancement during the October 2003 superstorm: relative importance of the neutral wind and the electric field. *J. Geophys. Res. Space Phys.* 110 (A12), 312. doi:10.1029/2005ja011304

Liu, J., Liu, L., Nakamura, T., Zhao, B., Ning, B., and Yoshikawa, A. (2014). A case study of ionospheric storm effects during long-lasting southward IMF *Bz*-driven geomagnetic storm. *J. Geophys. Res. Space Phys.* 119 (9), 7716–7731. doi:10.1002/2014ja020273

Liu, J., Wang, W., Burns, A., Yue, X., Zhang, S., Zhang, Y., et al. (2016). Profiles of ionospheric storm-enhanced density during the 17 March 2015 great storm. *J. Geophys. Res. Space Phys.* 121 (1), 727–744. doi:10.1002/2015ja021832

MacDougall, J. W., and Jayachandran, P. T. (2011). Solar terminator and auroral sources for traveling ionospheric disturbances in the midlatitude F region. *J. Atmos. solar-terrestrial Phys.* 73 (17-18), 2437–2443. doi:10.1016/j.jastp.2011.10.009

Munro, G. H. (1958). Travelling ionospheric disturbances in the F region. *Aust. J. Phys.* 11 (1), 91–112. doi:10.1071/ph580091

Nicolls, M. J., Vadas, S. L., Meriwether, J. W., Conde, M. G., and Hampton, D. (2012). The phases and amplitudes of gravity waves propagating and dissipating in the thermosphere: application to measurements over Alaska. *J. Geophys. Res. Space Phys.* 117 (5), 323. doi:10.1029/2012ja017542

Nishimura, Y., Zhang, S. R., Lyons, L. R., Deng, Y., Coster, A. J., Moen, J. I., et al. (2020). Source region and propagation of dayside large-scale traveling ionospheric disturbances. *Geophys. Res. Lett.* 47 (19), 619. doi:10.1029/2020gl089451

Ogwala, A., Oyedokun, O. J., Ogunmodimu, O., Akala, A. O., Ali, M. A., Jamjareegulgarn, P., et al. (2022). Longitudinal variations in equatorial ionospheric TEC from GPS, global ionosphere map and international reference ionosphere-2016 during the descending and minimum phases of solar cycle 24. *Universe* 8 (11), 575. doi:10.3390/universe8110575

Panda, S. K., Haralambous, H., and Kavutarapu, V. (2018). Global longitudinal behavior of IRI bottomside profile parameters from FORMOSAT-3/COSMIC ionospheric occultations. *J. Geophys. Res. Space Phys.* 123 (8), 7011–7028. doi:10.1029/2018ja025246

Panda, S. K., Harikaa, B., Vineetha, P., Kumar Dabbakutib, J. R. K., Akhila, S., and Srujanaa, G. Validity of different global ionospheric TEC maps over Indian region. 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), Greater Noida, India, 2021, 1749–1755.

Perlongo, N. J., Ridley, A. J., Cnossen, I., and Wu, C. (2018). A year-long comparison of GPS TEC and global ionosphere-thermosphere models. *J. Geophys. Res. Space Phys.* 123 (2), 1410–1428. doi:10.1002/2017ja024411

Pi, X., Mendillo, M., Hughes, W. J., Buonsanto, M. J., Sipler, D. P., Kelly, J., et al. (2000). Dynamical effects of geomagnetic storms and substorms in the middle-latitude ionosphere: an observational campaign. *J. Geophys. Res. Space Phys.* 105 (4), 7403–7417.

Qian, L., Burns, A. G., Wang, W., Solomon, S. C., Zhang, Y., and Hsu, V. (2016). Effects of the equatorial ionosphere anomaly on the interhemispheric circulation in the thermosphere. *J. Geophys. Res. Space Phys.* 121 (3), 2522–2530. doi:10.1002/2015ja022169

Rajesh, P. K., Lin, C. H., Lin, C. Y., Chen, C. H., Liu, J. Y., and Matsuo, T., (2021). Extreme positive ionosphere storm triggered by a minor magnetic storm in deep solar minimum revealed by FORMOSAT-7/COSMIC-2 and GNSS observations. *J. Geophys. Res. Space Phys.* 126 (2), 2020JA028261. doi:10.1029/2020ja028261

Richards, P. G., Fennelly, J. A., and Torr, D. G. (1994). EUVAC: a solar EUV flux model for aeronomic calculations. *J. Geophys. Res. Space Phys.* 99 (5), 8981–8992. doi:10.1029/94ja00518

Rishbeth, H. (1967). The effect of winds on the ionospheric F2-peak. *J. Atmos. Terr. Phys.* 29 (3), 225–238. doi:10.1016/0021-9169(67)90192-4

Shiokawa, K., Lu, G., Otsuka, Y., Ogawa, T., Yamamoto, M., Nishitani, N., et al. (2007). Ground observation and AMIE-TIEGCM modeling of a storm-time traveling ionospheric disturbance. *J. Geophys. Res. Space Phys.* 112 (5), 308. doi:10.1029/2006ja011772

Shiokawa, K., Mori, M., Otsuka, Y., Oyama, S., Nozawa, S., Suzuki, S., et al. (2013). Observation of nighttime medium-scale travelling ionospheric disturbances by two 630-nm airglow imagers near the auroral zone. *J. Atmos. Solar-Terrestrial Phys.* 103, 184–194. doi:10.1016/j.jastp.2013.03.024

Shiokawa, K., Otsuka, Y., Ogawa, T., Balan, N., Igarashi, K., Ridley, A. J., et al. (2002). A large-scale traveling ionospheric disturbance during the magnetic storm of 15 September 1999. *J. Geophys. Res. Space Phys.* 107 (6), 1088. doi:10.1029/2001ja000245

Shiokawa, K., Otsuka, Y., Ogawa, T., Kawamura, S., Yamamoto, M., Fukao, S., et al. (2003). Thermospheric wind during a storm-time large-scale traveling ionospheric disturbance. *J. Geophys. Res. Space Phys.* 108 (12), 1052. doi:10.1029/2003ja010001

Wang, W., Talaat, E. R., Burns, A. G., Emery, B., Hsieh, S. Y., Lei, J., et al. (2012). Thermosphere and ionosphere response to subauroral polarization streams (SAPS): model simulations. *J. Geophys. Res. Space Phys.* 117 (7), 301. doi:10.1029/2012ja017656

Weimer, D. R. (2005). Improved ionospheric electrodynamic models and application to calculating Joule heating rates. *J. Geophys. Res. Space Phys.* 110 (5), 306. doi:10.1029/2004ja010884

Wu, Q., Sheng, C., Wang, W., Noto, J., Kerr, R., McCarthy, M., et al. (2019). The midlatitude thermospheric dynamics from an interhemispheric perspective. *J. Geophys. Res. Space Phys.* 124 (10), 7971–7983. doi:10.1029/2019ja026967

Yin, F., Lühr, H., Park, J., and Wang, L. (2019). Comprehensive analysis of the magnetic signatures of small-scale traveling ionospheric disturbances, as observed by Swarm. *J. Geophys. Res. Space Phys.* 124 (12), 10794–10815. doi:10.1029/2019ja027523

Zhang, K., Liu, J., Wang, W., and Wang, H. (2019c). The effects of IMF *B* periodic oscillations on thermospheric meridional winds. *J. Geophys. Res. Space Phys.* 124 (7), 5800–5815. doi:10.1029/2019ja026527

Zhang, K., Wang, H., Liu, J., Zheng, Z., He, Y., Gao, J., et al. (2021b). Dynamics of the tongue of ionizations during the geomagnetic storm on September 7, 2015. *J. Geophys. Res. Space Phys.* 126 (6), 2020JA029038. doi:10.1029/2020ja029038

Zhang, K., Wang, H., Yamazaki, Y., and Xiong, C. (2021a). Effects of subauroral polarization streams on the equatorial electrojet during the geomagnetic storm on June 1, 2013. *J. Geophys. Res. Space Phys.* 126 (10), 2021JA029681. doi:10.1029/2021ja029681

Zhang, K., Wang, W., Wang, H., Dang, T., Liu, J., and Wu, Q. (2018). The longitudinal variations of upper thermospheric zonal winds observed by the CHAMP satellite at low and midlatitudes. *J. Geophys. Res. Space Phys.* 123 (11), 9652–9668. doi:10.1029/2018ja025463

Zhang, S. R., Coster, A. J., Erickson, P. J., Goncharenko, L. P., Rideout, W., and Vierinen, J. (2019a). Traveling ionospheric disturbances and ionospheric perturbations associated with solar flares in September 2017. *J. Geophys. Res. Space Phys.* 124 (7), 5894–5917. doi:10.1029/2019ja026585

Zhang, S. R., Erickson, P. J., Coster, A. J., Rideout, W., Vierinen, J., Jonah, O., et al. (2019b). Subauroral and polar traveling ionospheric disturbances during the 7–9 September 2017 storms. *Space weather* 17 (12), 1748–1764. doi:10.1029/2019sw002325

Zhang, S. R., Foster, J. C., Holt, J. M., Erickson, P. J., and Coster, A. J. (2012). Magnetic declination and zonal wind effects on longitudinal differences of ionospheric electron density at midlatitudes. *J. Geophys. Res. Space Phys.* 117 (8), 329. doi:10.1029/2012ja017954

Zhang, S. R., Vierinen, J., Aa, E., Goncharenko, L. P., Erickson, P. J., Rideout, W., et al. (2022). 2022 Tonga volcanic eruption induced global propagation of ionospheric disturbances via Lamb waves. *Front. Astronomy Space Sci.* 9, 871275. doi:10.3389/fspas.2022.871275

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 27 February 2024
DOI 10.3389/fspas.2024.1350209

Check for updates

OPEN ACCESS

EDITED BY
Xiangliang Kong,
Shandong University, China

REVIEWED BY
Gang Qin,
Harbin Institute of Technology, China
Lingling Zhao,
University of Alabama in Huntsville, United States

*CORRESPONDENCE
Xiaocheng Guo,
xcguo@swl.ac.cn

RECEIVED 05 December 2023
ACCEPTED 13 February 2024
PUBLISHED 27 February 2024

CITATION
Zhou Y, Guo X and Wang C (2024), Effects of anomalous cosmic rays on the solar wind events in the outer heliosphere.
Front. Astron. Space Sci. 11:1350209.
doi: 10.3389/fspas.2024.1350209

COPYRIGHT
© 2024 Zhou, Guo and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Effects of anomalous cosmic rays on the solar wind events in the outer heliosphere

Yucheng Zhou[1,2], Xiaocheng Guo[1,2]* and Chi Wang[1,2]

[1]State Key Laboratory of Space Weather, National Space Science Center, Chinese Academy of Science, Beijing, China, [2]School of Astronomy and Space Science, University of Chinese Academy of Sciences, Chinese Academy of Science, Beijing, China

Anomalous cosmic-rays (ACRs) are thought to be originated from the acceleration of pickup ions (PUIs) at the termination shock or interplanetary shocks, and play important role for the plasma dynamics in the outer heliosphere. Due to limited observation, the effects of ACRs on the solar wind events is not well known. Under the approximation of spherical symmetry, we have developed a three-component magneto-hydrodynamic (MHD) numerical model that contains solar wind plasma, interstellar neutral atoms and ACRs, to investigate the evolution of the solar wind within a heliocentric distance from 1 to 150 astronomical units (AU). We use the solar wind observations from the OMNI database with the time from 2010.5 to 2016.0 (decimal years) at the inner boundary, and the effect of ACRs on the propagation of the solar wind events are compared with the observations from the spacecrafts of New Horizons, Voyager 1 and 2. The results show that ACRs may decrease the speed of the solar wind shocks to some extent, and the effect is positively correlated with the diffusion coefficient; a larger diffusion coefficient leads to a more pronounced effect. Moreover, the ACRs has a dissipation effect on the shock-like solar wind structures, and may play important roles on the dynamics of solar wind in the outer heliosphere.

KEYWORDS

anomalous cosmic rays, MHD, heliosphere, interplanetary shock, pickup ions

1 Introduction

The heliosphere is a huge bubble or magnetosphere-like structure formed by the interaction between the solar wind plasma and the local interstellar medium (Parker, 1961). Voyager 1 and 2 detected the termination shock (TS) about 94 and 84 Astronomical units (AU), where the solar wind is significantly decelerated and heated, and entered the heliosheath in 2004 and 2007, respectively (Stone, 2005; Richardson et al., 2008). Later, the two probes crossed the heliopause at a distance of about 120 AU and entered interstellar medium in 2012 and 2018, respectively (Stone et al., 2019a), at a time when the count rates of galactic cosmic-rays (GCRs) were rising suddenly and those of the Anomalous Cosmic-Rays (ACRs) were falling rapidly (Krimigis et al., 2013). In the heliosphere, the solar wind events such as the co-rotation interaction regions (CIRs) and interplanetary coronal mass ejections (ICMEs), have a large amount of consequences as they are propagating into the outer heliosphere. For example, a series of CIRs may form the so-called merged interaction regions (MIRs) at larger heliocentric distances (≥8 AU) (Burlaga et al., 1984). The MIRs may interact with the heliopause and generate interstellar shocks that are expected to accelerate electrons and produce radio emission that could be detected by the spacecraft in some

situations. After Voyager 1 crossed the heliopause in August 2012, two forward shocks were respectively detected by the MAG instrument in 2012.92 and 2014.66, and a suspected reverse shock was detected in 2013.35 (Burlaga and Ness, 2016); while the plasma wave instrument detected their corresponding radio emissions as well (Gurnett et al., 2015).

Attempts have been made to locate the origins of these shocks inside the heliosphere. Liu et al. (2014) investigated the origin of an interstellar shock and the radio emissions based on a one-dimensional Magneto-hydrodynamics (MHD) simulation, and suggested that the observed radio emissions and associated shock are the result of a series of ICMEs that detected in March 2012. However, due to the lack of inclusion of interstellar plasma, the termination shock and the heliopause were not included in their model; however, they used the Earth's magneto-sheath to approach the inner heliosheath and calculate the propagation time of the shock. The evolution of the solar wind events in the outer heliosphere were investigated using a more sophisticated global MHD model, in which the 1-h resolution OMNI data at the spherical inner boundary at 1 AU was used (Fermo et al., 2015); they found that the modelled termination shock and heliopause are about 30 AU farther away than the real observations. Their further work showed that the multiple co-rotating interaction regions play an important role in the formation of the first forward shock detected by Voyager 1 after the heliopause crossing, and contribute for the second forward shock after taking ICMEs into account. (Kim et al., 2017). Richardson et al. (2017) found that the pressure pulses detected by Voyager 2 in the heliosheath are correlated with the interstellar shocks by Voyager 1. Based on a spherically symmetric MHD model containing termination shocks, Guo et al. (2021) performed the numerical simulations using three solar wind sources (STEREO A and B, and OMNI) with different longitudes as inner boundary conditions, and did comparison analyses with the *in situ* observations of New Horizons, Voyager 1 and 2 in the outer heliosphere. Their results indicate the possible connection between the pressure pulse observed by Voyager 2 and the interstellar shock events observed by Voyager 1. However, the possible effects of energetic particles, such as ACRs, have not been considered in the consequences of solar wind events in the outer heliosphere.

Anomalous cosmic rays are singlely-charged ions with energies from a few MeV to 100 MeV that play an important role in the outer heliosphere (Giacalone et al., 2022). They are generally believed to originate from pick up ions (PUIs) that have the sources of interstellar neutral atoms, and are accelerated by interplanetary shocks or termination shocks through the diffuse shock acceleration mechanism (DSA) (Fisk et al., 1974; Pesses et al., 1981; Gloeckler et al., 1994; Baring et al., 1997; McComas and Schwadron, 2006). Observations have also shown that the accelerated PUIs and ACRs can couple with the solar wind plasma and have an effect on the shock structure in the heliosphere (Terasawa et al., 2006). Alexashov and Chalov (2004) considered the effect of ACRs on the outer heliosphere using numerical simulations, and found that the diffusion of the ACRs pressure leads to the formation of a smooth shock precursor at upstream of the termination shock, which in turn further increases the radial distance of the termination shock. Guo et al. (2018, 2019) also found that the escape of ACRs across the heliopause into interstellar medium leads to a contraction of the inner heliosheath, which can partly explain the observation that Voyager detected the heliopause earlier than theoretically expected (Stone et al., 2013; 2019b). In addition, the separation of the shock front from the peak of the ACRs count rate observed by Voyager 2 (Lazarus et al., 1999) may be due to an interruption in the formation of ACRs from PUIs via DSA, which leads to the peak convecting with the solar wind and separating from the shock front (Rice and Zank, 2000). After crossing the termination shock, Voyager 2 detected a short-term enhancement of the count rate of ACRs, which has been discussed through the magnetic reconnection, (e.g., Drake et al., 2010) and the related stochastic acceleration mechanism (Zank et al., 2015; Zhao et al., 2019); alternatively, it may also be related to the temporary increase in shock intensity caused by the passage of interplanetary shocks through the termination shock (Guo et al., 2022).

In this paper, we will investigate the effect of ACRs on solar wind propagation under realistic solar wind conditions by means of numerical simulations, in which the solar wind plasma, neutral atoms, and ACRs are coupled based on the MHD-neutrals-ACRs combined equations under spherical symmetry approximation. The solar wind from OMNI-database will be used as the inner boundary condition, and the effects of ACRs on the solar wind will be discussed in comparison with observations from New Horizons, Voyager 1 and 2.

2 Numerical model

In this model, the solar wind plasma is governed by the ideal MHD equations, and the ACRs are treated as mass-less fluid with pressure only and are governed by the diffusion equation in which the distribution function is isotropic and energy-averaged in phase space. The interstellar neutral atoms are fixed as the background inflow and have the charge-exchange with the solar wind protons. All variables of the solar wind are assumed to be spherically symmetric, that is, they depend exclusively on the heliocentric distance r, and the spherical derivatives ($\partial/\partial\phi = \partial/\partial\theta$) are zero in the polar coordinates (r, θ, ϕ). so that the solar wind flows along the radial direction near the ecliptic plane, being without the side impact of longitudinal and latitudinal flows. Under these approximations, the normalized equations are written as follows:

$$\frac{\partial \rho}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \rho u_r\right) = Q_N, \tag{1}$$

$$\frac{\partial \left(\rho u_r\right)}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\left(\rho u_r^2 + P_T + P_c - B_r^2\right)\right] = \frac{2P_T + 2P_c - B_\phi^2 + \rho u_\phi^2}{r} - \frac{\rho G M_s}{r^2} + Q_{Mr}, \tag{2}$$

$$\frac{\partial \left(\rho u_\phi\right)}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\left(\rho u_r u_\phi - B_r B_\phi\right)\right] = \frac{\rho u_\phi u_r - B_\phi B_r}{r} + Q_{M\phi}, \tag{3}$$

$$\frac{\partial B_\phi}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}\left[r^2\left(u_r B_\phi - B_r u_\phi\right)\right] = -\frac{u_\phi B_r - B_\phi u_r}{r}, \tag{4}$$

$$\frac{\partial}{\partial r}\left(r^2 B_r\right) = 0, \tag{5}$$

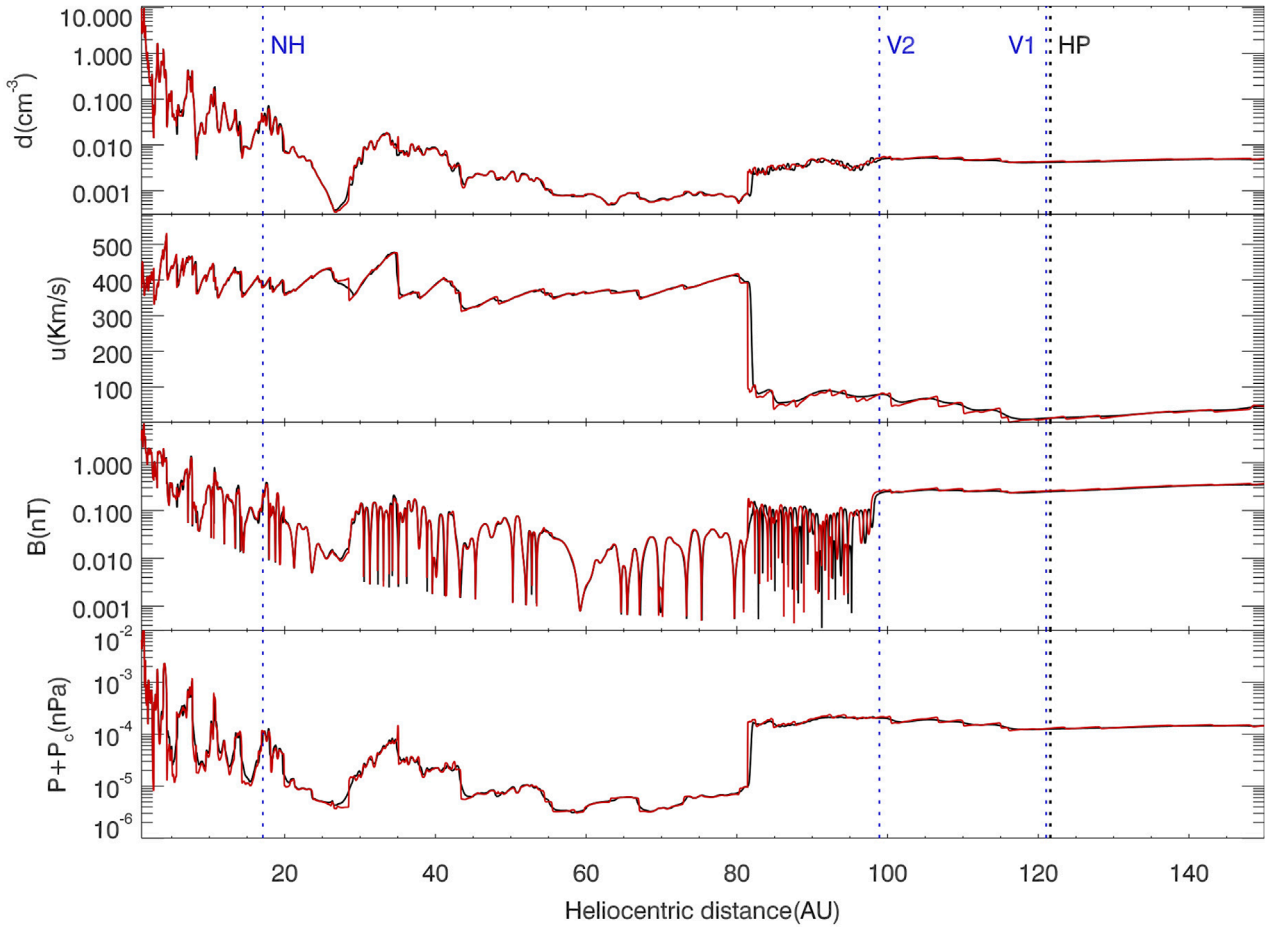

FIGURE 1
From top to bottom, the radial profiles of the solar wind density, velocity, magnetic field strength, and the total pressure of solar wind and ACRs, at the year 2012.5. The black and red solid curves are the results with ACRs and without ACRs, respectively. The positions of New Horizons, Voyager 1 and 2 are shown as the three blue vertical dashed lines. The heliopause is marked as the black vertical dash line for comparison.

$$\frac{\partial E}{\partial t}+\frac{1}{r^2}\frac{\partial}{\partial r}\left\{r^2\left[\left(E+P_T\right)u_r-B_r\left(\mathbf{u}\cdot\mathbf{B}\right)\right]\right\} = -u_r\frac{\partial P_c}{\partial r}+\frac{\alpha}{r^2}\frac{\partial\left(r^2u_r\right)}{\partial r}-\frac{\rho GM_s u_r}{r^2}+Q_E, \tag{6}$$

$$\frac{\partial E_c}{\partial t}+\frac{1}{r^2}\frac{\partial}{\partial r}\left\{r^2\left[\left(E_c+P_c\right)u_r-\kappa\frac{\partial E_c}{\partial r}\right]\right\}=u_r\frac{\partial P_c}{\partial r}-\frac{\alpha}{r^2}\frac{\partial\left(r^2u_r\right)}{\partial r}. \tag{7}$$

where$\rho,\mathbf{u},\mathbf{B},E$ represent the plasma density, velocity, magnetic field, and energy density, respectively. $E=P/(\gamma-1)+\rho u^2/2+B^2/2$, and $P_T=P+B^2/2$, where P is the thermal pressure of solar wind. The pressure of ACRs is P_c, and the corresponding energy density $E_c=P_c/(\gamma_c-1)$. The diffusion coefficient κ is spatially isotropic and set to be constant in the simulation. Namely, the simulation employs three diffusion coefficients, $\kappa_1=5\times10^{18}cm^2s^{-1}$, $\kappa_2=5\times10^{19}cm^2s^{-1}$ and $\kappa_3=5\times10^{20}cm^2s^{-1}$,and$\kappa_0$ indicates no diffusion results. The plasma adiabatic index is $\gamma=5/3$; the ACRs are considered to be non-relativistic, so the adiabatic index $\gamma_c=5/3$. Charge exchange between the plasma and interstellar neutral atoms is implemented through the three source terms (Q_N, Q_M, Q_E) (Pauls et al., 1995). Details of the charge-exchange source terms are expressed in appendix A of the previous work (Guo et al., 2019). Solar gravity is included as well, with G being the solar gravitational constant and M_s the solar mass. The above equations are implemented with the MUSCL numerical scheme (van Leer, 1979), in which the extended HLLC Riemann solver are used (Guo, 2015). Overall, the simulation code has second-order accuracy for both spatial reconstruction and time evolution.

The inner boundary is located at a heliocentric distance of 1 AU, and the simulation domain is from 1 to 150 AU, with a nonuniform grid with a total number of 10,000. The grid size ΔR is 0.004 AU at the inner boundary, 0.021 AU near the termination shock (~90 AU), and 0.034 AU at the outer boundary. Typical solar wind conditions are initially set at the inner boundary as follows: number density $10cm^{-3}$, radial velocity $u_r=400km/s$, temperature $T=5\times10^4K$, and magnetic field $B_r=2.8nT$. Assuming that the interstellar neutrals are hydrogen atoms that move along the anti-radial direction at a constant speed of 26.2 km/s and a temperature of 6,300K. Their densities follow a typical exponential distribution (Axford, 1972):

$$n_H=n_{H0}exp\left(-\frac{r_0}{r}\right) \tag{8}$$

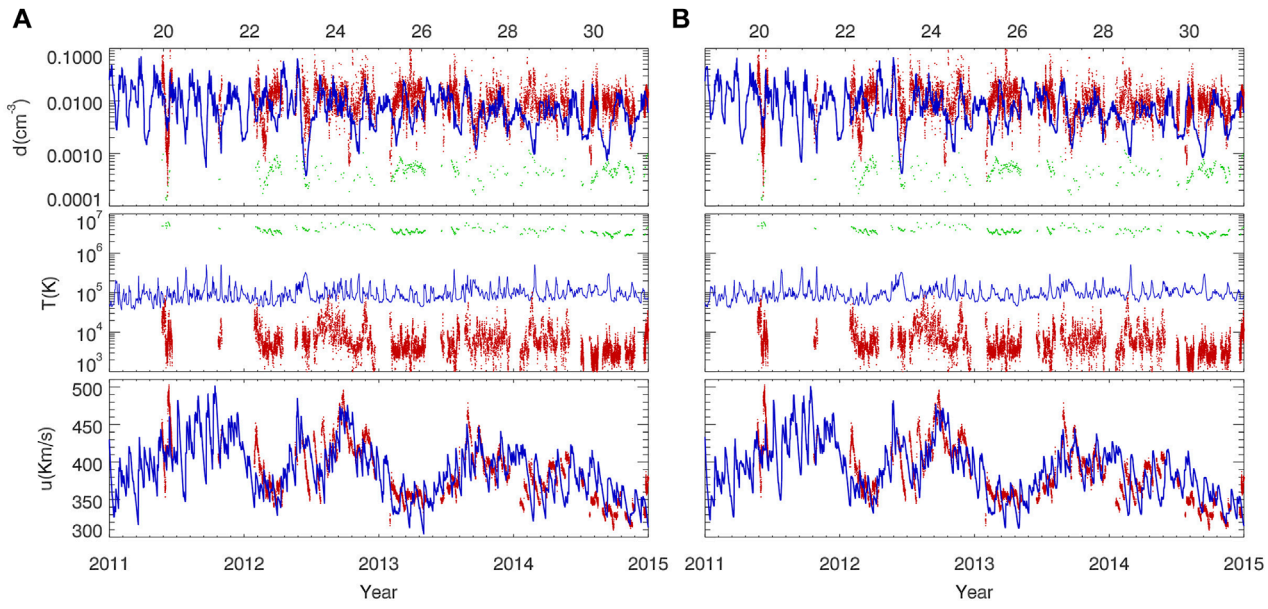

FIGURE 2
From top to bottom: Time profiles of number density, temperature and bulk flow speed from the observation of New Horizons (red: plasma; green: PUIs) and the model simulation (blue) during the years 2011.0–2015.0. **(A)** shows the simulated data without ACRs, while **(B)** shows data with ACRs.

where $n_{H0} = 0.15cm^{-3}$ is the number density of neutral atoms in interstellar space, and r_0 the penetration depth. Eq. 7 describes the transport of the cosmic-ray pressure P_c, and was derived from the Parker transport equation with the assumption of energy-averaged in phase space, (e.g., Zank et al., 1993; Chalov and Fahr, 1996). The ACRs pressure P_c is 0 everywhere initially, and the quantity α is the local injection rate of the ACRs particles from the lower-energy PUIs (Zank et al., 1993) that depends on shock strength. Similar to Rice et al. (2000), here we have an α value of:

$$\alpha = \alpha' exp\left(-\frac{r_0}{r}\right)p \tag{9}$$

where α' is a constant that measures the injection efficiency and generally takes values between 0.1 and 1.0; here it is set to be a constant of 0.8 near the shock, while zero in the uncompressed region. We assume that all particle energies can be described by a distribution function, and we consider energetic particles with energies larger than $5\,keV$ to be ACRs, so we do not distinguish between low-energy PUIs and the solar wind (Wang and Richardson, 2001), and between high-energy PUIs and ACRs (Eichler, 1979; Zank et al., 1993). The acceleration of PUIs from low-energy to high-energy (or ACRs) at the shocks are accomplished by the adiabatic heating term in the RHS of Eq. 7, during which the solar wind plasma lose energy as a compensation by the adiabatic cooling term in the RHS of Eq. 6. Once the PUIs are accelerated to the ACRs at shocks, they will convect and diffuse, being governed by the simplified Parker transport equation (Eq. 7). Galactic cosmic rays are not taken into account, because of their small pressure gradient in the inner heliosphere and their large diffusion coefficients. Therefore, they do not have a significant effect on the shock as well as on the simulation results (Rice et al., 2000).

Due to the limitations of the spherically symmetric flow approximation, the interstellar plasma is not included in the model, and therefore the heliopause cannot be reproduced. Similar to the previous approach (Florinski et al., 2004), we apply a constant pressure at the outer boundary, and a termination shock can be obtained after evolution. In order to obtain a termination shock that is stable at around 80–90 AU, we set $p = 7.2P_{ISM}$ for the outer boundary condition, where P_{ISM} is the thermal pressure of the interstellar medium. Free boundary conditions are used for other variables. Once the initial state is complete, the time-dependent solar wind data from the OMNI database is imposed at the inner boundary from the year 2010.5–2016.0, to drive the evolution of the solar wind along the radial direction. Figure 1 shows the radial profile of the solar wind quantities at the year 2012.5, with the black and red curves corresponding to those with ACRs and without ACRs, respectively. The diffusion coefficient of κ_2 is applied for the simulation. The black curve appears more smooth than the red one near shocks, and it moves into a further distance than the red one because of the existence of shock precursors (Alexashov and Chalov, 2004; Guo et al., 2019). For better comparison, the positions of New Horizons, Voyager 1 and 2 are marked in the figure. These three spacecrafts are assumed to move upwind toward the intestellar neutrals, which is roughly consistent with the reality (Guo et al., 2021). Note that Voyager 1 is approaching the heliopause which is located at 121.6 AU from the later observation, the termination shock is at about 81 AU at the time.

3 Simulation results

3.1 Comparison results with New Horizons

Figure 2 shows the simulation results and the *in situ* measurements from New Horizons, with the time ranging from 2011.0 to 2015.0, and the heliocentric distance from 18.7 to 31.3 AU. From the observations, the solar wind density, temperature

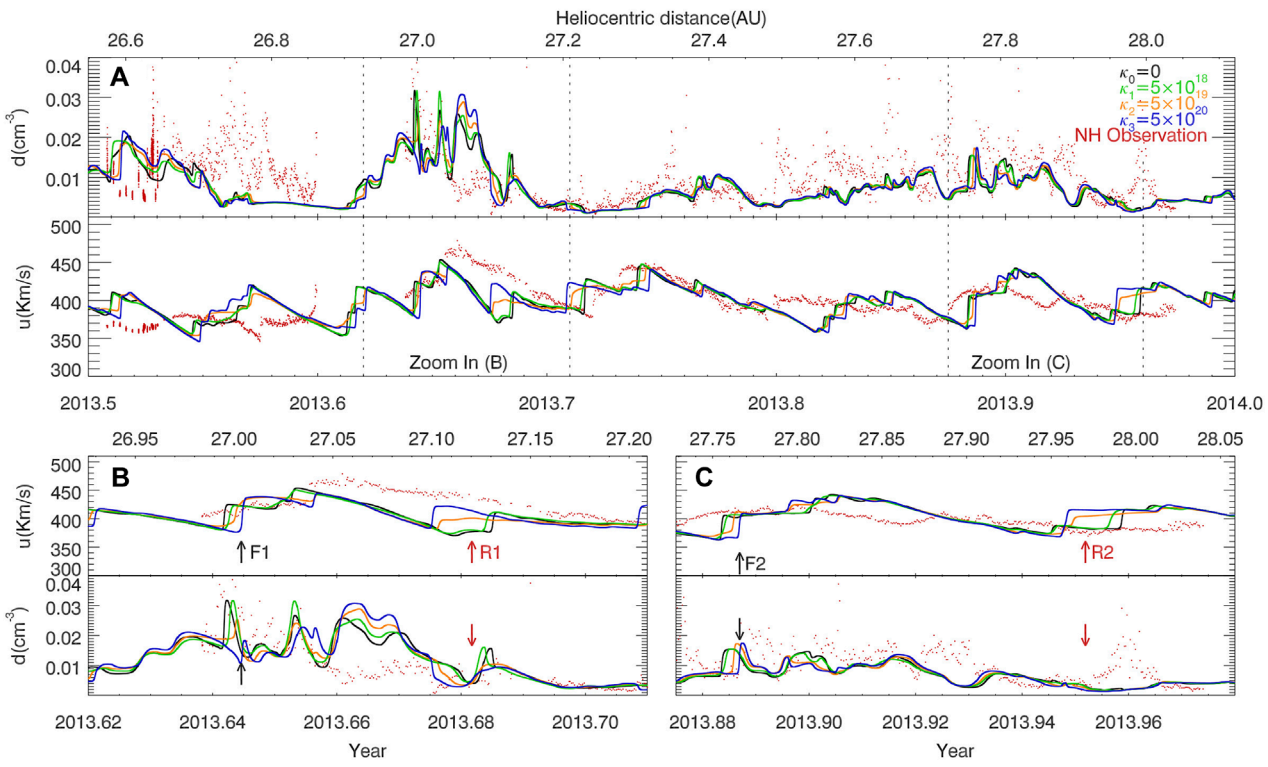

FIGURE 3
(A) Profiles of the number density and flow speed of the solar wind from the years 2013.5–2014.0. The simulation results for the four diffusion coefficients κ_0, κ_1, κ_2 and κ_3, presented by the black, green, orange and blue curves, respectively. The observations from New Horizons are plotted as red dots. The unit of diffusion coefficient is cm^2s^{-1}. Two typical views from 2013.6 to 2013.7 and 2013.87–2013.98 are shown in panels **(B, C)**, respectively.

and velocity are plotted as the red dotted curves; the density and temperature of the PUIs are represented by the green dotted curves. The blue solid curves correspond to the simulation data. Magnetic field data are not available for New Horizons due to the absence of magnetometers. We performed the simulations with the model including ACRs with the diffusion coefficient of κ_2 *versus* the model without ACRs, and the results are shown in the panels A and B, respectively. Since New Horizons is located at a lower latitude and at a distance closer to the Sun than Voyager 1 and 2, the numerical simulations match the observations of New Horizons much better than the other two as we will show below. From the comparison between A and B, we find that the simulation results with and without ACRs are basically same in large-scale solar wind structures, and the simulated plasma densities are nearly in the same order of magnitude as the observed values. Specifically, the observed densities are higher than the simulation results during most of the time in the figure. The simulated velocity variation agrees with the observed large-scale cycle variation with a period of about 1.3 years (Richardson et al., 1994). The simulated temperatures are much larger than those of the plasma and much smaller than the temperatures of PUIs, indicating that they are mixed between those of the solar wind and the PUIs. The number density of the observed thermal solar wind ions is about 95% over the total solar wind particles during the period, the calculation yields an average temperature of $2.12 \times 10^5 K$, which is a little higher than the simulated averaged temperature of $1.07 \times 10^5 K$. Note that high-energy PUIs above 5 *keV* have been classified as ACRs in our model, whereas the energetic particle instrument of New Horizons has a detection up-limit for H^+ PUIs of 7.8 *keV*(McComas et al., 2017). Therefore, in our model, most of the PUIs are considered as part of the solar wind plasma for simplicity.

For more details, we select an interval of the time 2013.5–2014.0, as shown in Figure 3. The black solid curves indicate the simulation results without ACRs; the green, orange and blue curves correspond to those with ACRs, with the diffusion coefficients of κ_1,κ_2 and κ_3, respectively; the red dots indicate the observation data from New Horizons. In panel A, each large-scale solar wind structure roughly has a corresponding simulation result, although exact coincidence in time is not achieved for the current modeling. Several shock pairs are identified, with the typical characteristics of the forward and reverse shocks, as well as density enhancement in between. Taking the event from 2013.6 to 2013.7 as an example, shown in panel B, the forward and reverse shock fronts are respectively pointed by black arrow F1 and red arrow R1. The effects of ACRs on the solar wind evolution are clearly seen for the three diffusion coefficients. For instance, at the distance of ~27 AU, ACRs significantly affect the structure of shocks with a higher diffusion coefficient corresponding to a slower shock speed, being compared with the black curves from the non-ACRs case. This phenomenon is caused by the energy loss of the shock due to the shock acceleration from the higher-energy PUIs in the solar wind to the ACRs, which are diffusive and no longer coupled with the solar wind plasma. Once the ACRs are produced at shocks and result in an increasing pressure, the thermal pressure will decrease as a response and lead to a drop of the local sonic speed

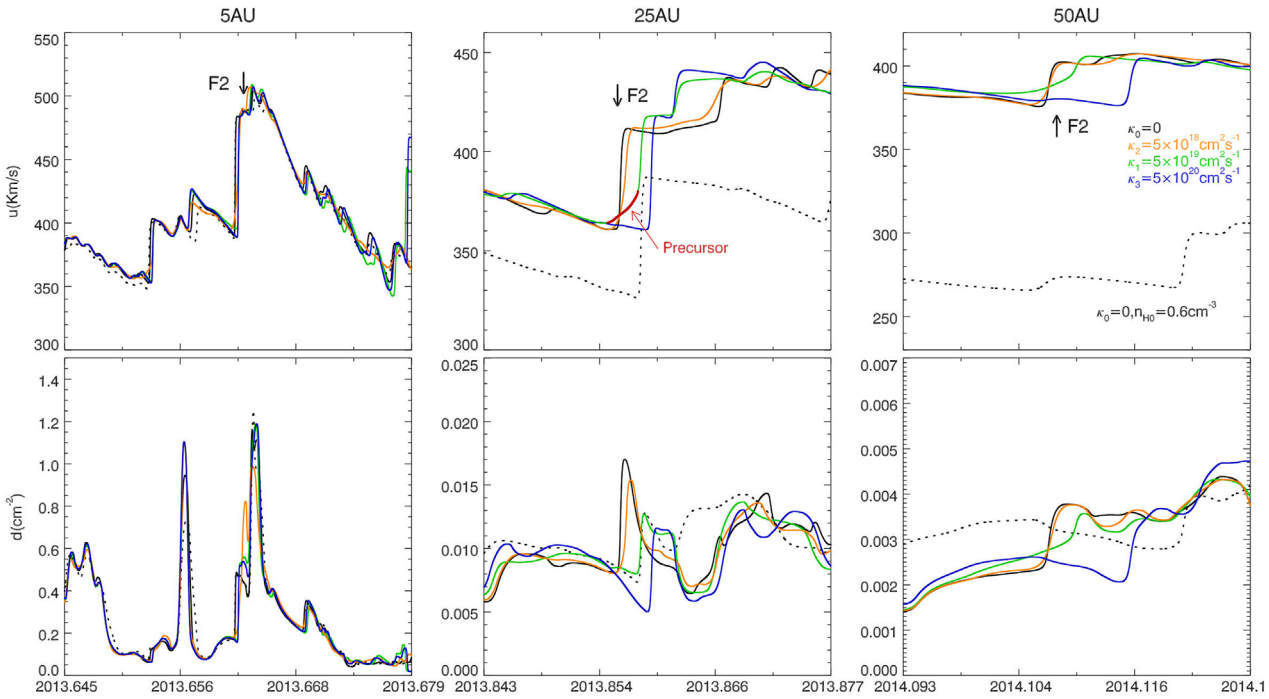

FIGURE 4
Time evolution of the solar wind structure near the 2013.88 shock (shown in Figure 3) at the three radial distances (5, 25 and 50 AU). The red arrows point to the corresponding positions of the shock front. The dotted curves show the corresponding solar wind structure for a higher neutral density of $0.6cm/m^3$, with a diffusion coefficient of $\kappa_3 = 5\times10^{20}cm^2s^{-1}$.

and consequently a slowdown in the shock speed in the solar inertial coordinate. As a result, the shock with higher diffusion case will be detected later than those from the lower diffusion or the non-ACRs cases. As the black arrows F1 and F2 indicate in panels B and C, there is about 1 day gap between the black and blue shock fronts, which respectively correspond to the two cases of κ_0 and κ_3.

On the contrary, as the red arrows R1 and R2 mark, the slower shock speed for the reverse shocks will make the shocks be detected earlier than those of the lower diffusion cases in the solar coordinate, due to its opposite propagation direction relative to the expanding solar wind flow. In panels B and C, there are about respectively three and 4 days between the two cases of κ_0 and κ_3. Notice that some shock-like structures may be identified between the above shock pairs due to the complicated interaction between the adjacent merged interaction regions. Here we do not attempt to analyze all of them since they have relatively small amplitudes.

For more details of the evolution, we trace the forward shock shown as F2 in Figure 3 at the three radial distances of 5, 25, and 50 AU, as pointed by the red arrows in Figure 4. At 5 AU, the shocks begin to depart from the others for the three different diffusion environments. As the distance increases, the separation distances between shocks become larger, with a higher diffusion corresponding to a slower shock speed and thus a latter arrival time for a same location. For example, the shock in the case of κ_3 is is detected at ~3.5 days later than in the case of no ACRs at 50 AU. There is about 182 days for the shock propagation from 1 to 50 AU in the case of no ACRs, thus the deceleration rate is ~2% for the case of κ_3.

The effect of ACRs on the shock may be characterized by the precursor that persists in the upstream of the shock in a certain diffusion environment. As the middle panel shows, the green shock fronts are obviously modulated compared to the other cases with or without ACRs. Because the PUIs are accelerated to ACRs at the shock, and further back-scattered into the upstream; the pressure gradient of ACRs then forms and decelerates the inflow solar wind, leading to the formation of precursor as a transition from the supersonic flow to the shock front (e.g., Florinski et al., 2009). The precursor is highlighted as red color from the green curve for a better identification. From a diffusion theory, the scale of the precursor λ is estimated to be κ/U, where U is the convection speed of the background flow in the shock reference frame. As the red arrow show, it is calculated that the convection flow speed in the upstream of the shock is ~150 km/s in the shock reference frame, then the scale of precursor is ~0.2 AU for a given diffusion coefficient of κ_2, which is consistent with the simulation result of ~0.35 AU. As for the lower diffusion case of κ_1, this scale is ~0.02, being too short to be reproduced well because of the limit of local grid spacing of ~0.01 AU. The large diffusion coefficient of κ_3 corresponds to a scale size of ~2 AU, which is too large to be visible because the coupling between the solar wind and ACRs is very weak.

The density distribution of different interstellar neutral atoms affects the propagation of the solar wind due to charge exchange, (e.g., Wang et al., 2000), and the shock propagation speed as well. In order to evaluate the effects of ACRs, here we make a comparison with the effects of neutrals. Based on the case of κ_0, we test the effect of interstellar neutral atoms by changing n_{H0} from 0.15 to 0.6 cm^{-3} in Eq. (8). Unlike the solo effect of ACRs on shocks, the enhancement of neutral densities greatly decrease both the solar

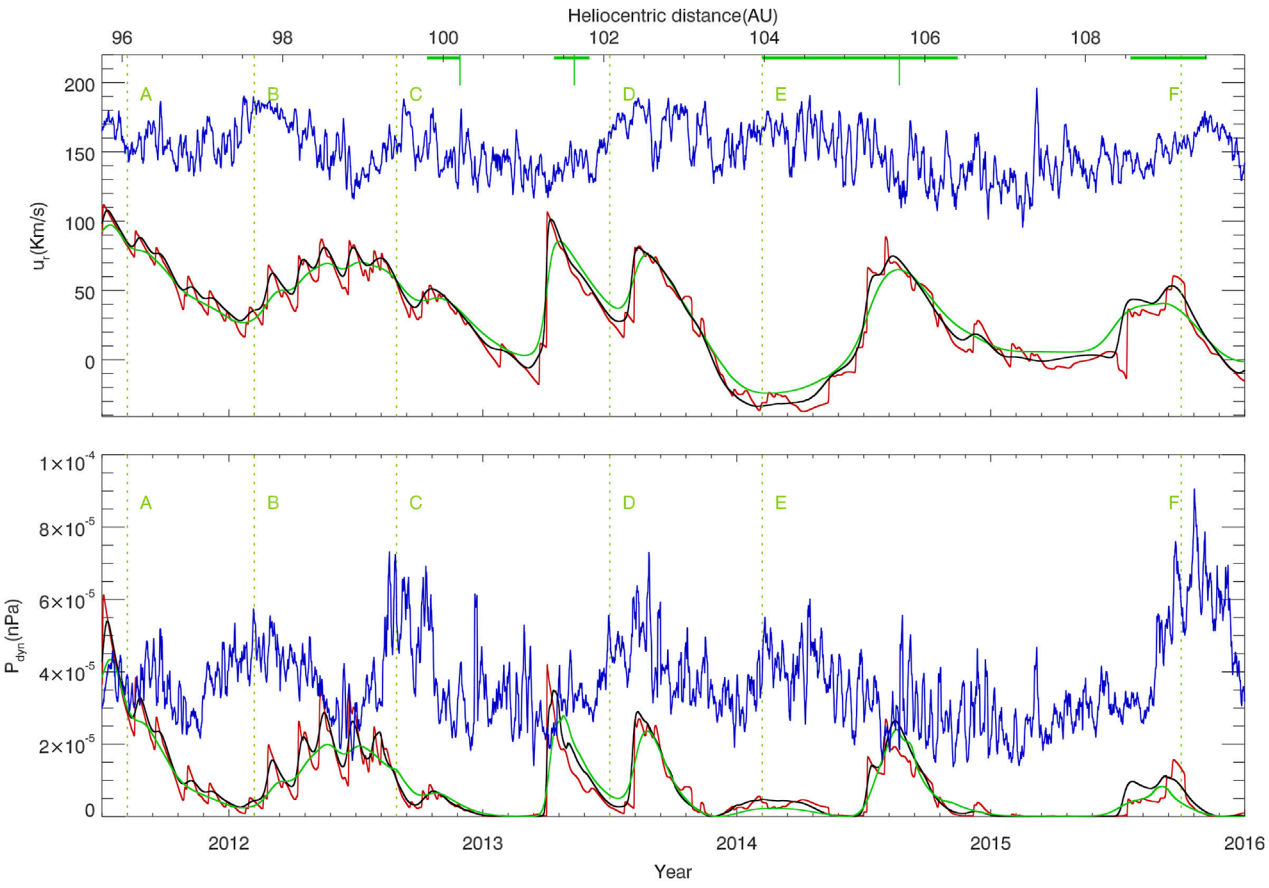

FIGURE 5
Comparison of the solar wind velocity and dynamical pressure in inner heliosheath between the *in situ* observations by Voyager 2 and the simulation results. The solid blue curves show the Voyager 2 observations, the thick green horizontal curve in the upper panel shows the plasma oscillations events detected by Voyager 1, the short green vertical line shows the three shock crossings by Voyager 1, and the green dashed line shows the six pressure pulse peaks detected by Voyager 2. The red solid curve indicates the simulation data without ACRs, the black solid and green curves are the simulation results with ACRs and correspond to the diffusion coefficients of κ_2 and κ_3, respectively. **(A–F)** mark the positions of the maximum local dynamic pressure of the MIRs recorded by Voyager 2.

wind and shock speeds because the charge exchange takes place over all the outer heliosphere. Since ACRs are generated only at the shock front, they have a relatively small effect on the solar wind plasma density. However, the shock deceleration effect by increasing 4 times of neutral density is approximately equivalent to the case of κ_3 in the simulation.

3.2 Comparison results with Voyager 2

We select the variation of solar wind during the years 2011.5–2016.0 for comparison, when Voyager 2 was still located within the heliosheath (Stone et al., 2019a). As shown in Figure 5, the blue curves indicate the observations from Voyager 2, the red, black and green curves correspond to the simulation results with the cases of κ_0, κ_2 and κ_3, respectively. The six green vertical dashed lines (labels A-F) roughly mark the positions of the maximum local dynamic pressure of the MIRs recorded by Voyager 2, which are thought to correlate with the transient events observed by Voyager 1 in the interstellar medium (Richardson et al., 2017). The thick green horizontal lines indicate the plasma oscillation events measured by Voyager 1 in the interstellar medium (Gurnett et al., 2015), and the solid green vertical lines correspond to the three shock crossing events (Burlaga and Ness, 2016). It is generally believed that the first three shock crossing events are in good agreement with the plasma oscillation events, while the potential plasma shock event being associated with the fourth plasma oscillation event was not detected by Voyager 1.

From the comparison, we can see that the simulated solar wind speed and dynamic pressure are generally lower than the observed values. For example, in 2012.7, the difference between the observed and simulated dynamic pressures is $6\times10^{-5} nPa$. For each observed pressure pulse event, we can find a corresponding simulated solar wind structure that is usually later than the observation. For example, for the observed dynamic pressure pulse B in ~2012.1, the simulations show a rough correspondence with the structure of the dynamic pressure pulse in ~2012.4. In addition, the observed pulses C, D, and E may correspond to the pulse structures of ~2013.2, ~2013.6, and ~2014.6 in the simulations, respectively. The corresponding observed solar wind speeds are larger than the simulated solar wind speeds by an average of 100 km/s. This discrepancy between simulations and observations arises from

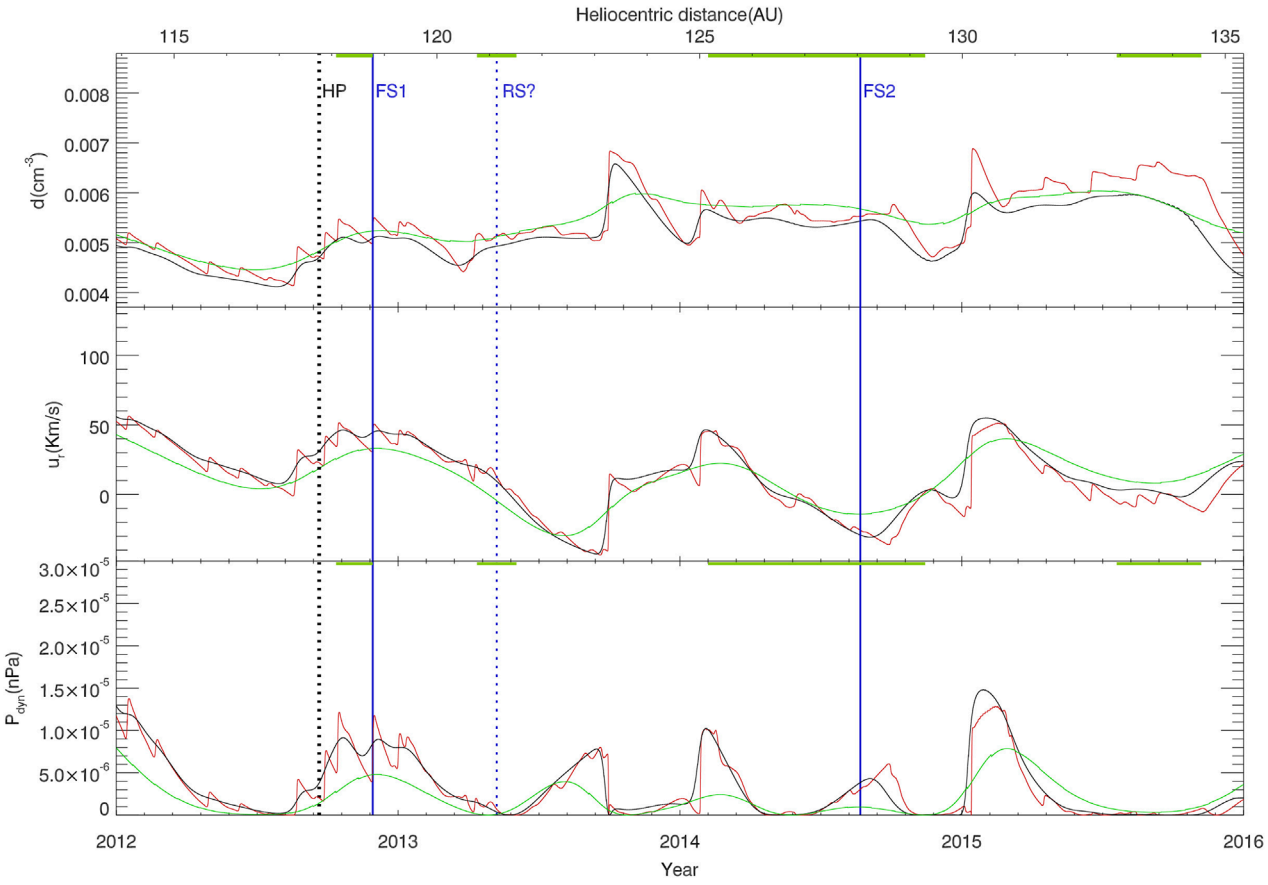

FIGURE 6
The simulated density, velocity and dynamic pressure during the period of 2012.0–2016.0 for Voyager 1. The thick green horizontal lines indicate the plasma wave events detected by Voyager 1 and the long blue vertical lines mark the time of three shock crossing events.

the complexity of the evolution of the solar wind in the inner heliosheath, one possibility is that due to the simplification of our spherically symmetric model, at this time Voyager 2 was flying at a latitude of about 34°S, which is far from the ecliptic plane where the model is best applied, and the model itself does not incorporate the non-radial action flow of the solar wind and does not take into account the evolution of the solar wind at high latitudes and its impact on the solar wind at low latitudes, so there is a large discrepancy in the correspondence between simulated times and observed times. In addition, the comparison with the plasma observations of the Voyager 2 spacecraft still remains a challenge even for the global MHD simulations (e.g., Washimi et al., 2011).

Comparing the simulation results with and without ACRs, we find that the diffusion effect of ACRs is more significant on the smoothing of the shock-like structure of solar wind for a higher diffusion coefficient. For example, the multiple structures, which are shown without ACRs during 2012.0–2012.7, merge into a large structure for the case of κ_3, because the scattering of ACRs near shocks will form the shock precursors that finally smooth the adjacent small-scale shock structures. As expected, a higher diffusion coefficient corresponds to a wider transition region for a shock. Taking the ~2013.2 shock as an example, the black curve shows that a precursor forms in the upstream of the shock for κ_2, with a size of ~0.3 AU. The green curve shows that a higher diffusion of ACRs has a more significant smoothing effect on the shock, with a ~1 AU width for the transition region. However, the scattering of ACRs near shock does not change the propagation speed of shock-like structures apparently in the inner heliosheath, the time correspondence between the simulated solar wind fluctuations and the interstellar shock signatures prescribed by Voyager 1 still remains.

3.3 Comparison results with Voyager 1

Voyager 1 became the first probe that entered the interstellar space when it crossed the heliopause at 121.6 AU in August 2013, and since then it became possible to measure the local interstellar medium. However, Voyager 1 can not measure the plasma due to the disabled plasma instrument since 1980, and it is not possible to directly compare simulated plasma parameters with observational data. Due to the limitations of the spherically symmetric model, we are unable to simulate the interstellar plasma inflow, so that heliopause is absent and we do not distinguish the inner and outer heliosheath in the simulation. In the real observations, after Voyager 1 crossed the heliopause, the detected ACRs escape rapidly and their intensity decreases to the background noise level due to the non-scattering environment for ACRs in the interstellar space (Stone et al., 2019b). Due to the limit of the model, the heliopause is absent

and the ACRs are contained all over the simulation domain. The main purpose of the comparison is to evaluate the possible time correspondence between the solar wind structures affected by ACRs and the observed interstellar shock events, since it is expected that the solar wind structures will impact with heliopause and generate the waves propagating in interstellar space (e.g., Mostafavi et al., 2022).

Figure 6 shows the simulated density, velocity and dynamic pressure results at the Voyager 1 position from the years 2012.0–2016.0. The black vertical dotted line indicates the position of heliopause at 121.6 AU observed by Voyager 1, the two blue vertical solid lines indicate the detection of two forward shocks FS1 and FS2, the vertical dashed line indicates a possible reverse shock RS1 from the magnetic field data (Burlaga and Ness, 2016), and the green horizontal solid line indicates the four radio emissions events detected by Voyager 1. As can be seen in Figure 6, there are corresponding shock-like structures near the three observed shock crossing events, and the plasma wave events as well. They are thought to be linked to the pressure pulses C, D, and E detected by Voyager 2 shown in Figure 5.

Because ACRs are not removed from the simulation in the interstellar space, the effects of ACRs on the solar wind structures may be exaggerated. However, the basic profiles of the solar wind structures look similar whether the ACRs are included or not as we see from Figures 5, 6. As expected, some local structures are smoothed by the ACRs during the evolution. For example, during the period of 2014.6–2015.0, the solar wind structures are seen in the low diffusion or no ACRs cases, but disappear in the large diffusion case. For the case of κ_2, shock precursors appear near shocks, being similar to those in Figure 5. As for the highest diffusion case of κ_3, the small shock-like structures all dissipate into a large-scale compression structure, and no precursors are found as well. It is no doubt that the high diffusion environment of κ_3 is not consistent with the observation by Voyager 1 because several shocks have been found in the interstellar space from the magnetic field measurements. In the reality, the solar wind events will encounter the heliopause and generate large-scale waves or shocks propagating into the heliosphere, it is less likely that the smoothing solar wind structures will lead to the shocks observed in interstellar medium. Note that the diffusion coefficients are set to be uniform over all the simulation domain, which apparently is a simplifed treatment because the diffusion coefficients are expected to vary with radial distance (hence turbulence) and particle rigidity, (e.g., Zhao et al., 2017; 2018). For example, the perpendicular diffusion coefficients will increase as the magnetic turbulence is enhanced near termination shocks, leading to the variation of the precursor for the shock. Similar to the work by Wang et al. (2022), a more sophiscated model with the turbulence embedded is needed for the future work, even under a context of global MHD simulation. This improved treatment will help for a better understanding with the effects of ACRs on the solar wind events in the outer heliosphere.

4 Summary

Based on the time-dependent plasma-neutrals-ACRs numerical MHD model, we investigate the effect of ACRs on the evolution of the solar wind in the outer heliosphere under the realistic solar wind conditions from the OMNI database. The solar wind data are used as the inner boundary condition at 1 AU to drive the simulation ranging from the years of 2010.5–2016.0. The simulated evolution of solar wind structures are compared with the observations from the spacecrafts of New Horizons, Voyager 1 and 2. The results show that ACRs have some effects on the evolution of the shock-like structures of solar wind in the outer heliosphere. In the supersonic solar wind region, the ACRs are able to reduce the shock propagation speeds to a certain extent because of the energy loss of solar wind caused by the diffusive ACRs transformed from the PUIs. For example, a simulation case shows that there is about 3% of the deceleration rate with a large diffusion coefficient of κ_3 for a propagating shock observed at 50 AU. In this case, the forward shocks slow down relative to the non-ACRs cases and are observed later, whereas the reverse shocks are observed earlier because of its opposite propagation direction relative to the solar wind flow. Shock precursors are commonly found at larger heliocentric distances with the diffusion coefficient of κ_2, the smoothing effects will make shock events less observable compared with the non-ACRs cases.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

YZ: Investigation, Project administration, Visualization, Writing–original draft, Writing–review and editing. XG: Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Software, Supervision, Writing–original draft, Writing–review and editing. CW: Supervision, Project administration, Funding acquisition, resources, Writing–review and editing

Funding

The author(s) declare financial support was received for the research, authorship, and/or publication of this article. The work was supported by the National Key R&D program of China No. 2021YFA0718600, NNSFC grants 42150105, 41874171, the Strategic Priority Research Program of Chinese Academy of Sciences, grant No. XDB 41000000, the Key Research Program of the Chinese Academy of Sciences (Grant No. ZDRE-KT-2021-3).

Acknowledgments

We acknowledge the use of NASA/GSFC's Space Physics Data Facility's CDAWeb service, and OMNI data.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Alexashov, D. B., Chalov, S. V., Myasnikov, A. V., Izmodenov, V. V., and Kallenbach, R. (2004). The dynamical role of anomalous cosmic rays in the outer heliosphere. *Astronomy Astrophysics* 420, 729–736. doi:10.1051/0004-6361:20041056

Axford, W. I. (1972). "The interaction of the solar wind with the interstellar medium," in *Solar wind*. Editors C. P. Sonett, P. J. Coleman, and J. M. Wilcox (Washington: Scientific and Technical Information Office, National Aeronautics and Space Administration.), 609.

Baring, M. G., Ogilvie, W., Ellison, C., and Forsyth, R. J. (1997). Acceleration of solar wind ions by nearby interplanetary shocks: comparison of Monte Carlo simulations with ulysses observations. *Astrophysical J.* 476, 889–902. doi:10.1086/303645

Burlaga, L. F., McDonald, F. B., Ness, N. F., Schwenn, R., Lazarus, A. J., and Mariani, F. (1984). Interplanetary flow systems associated with cosmic ray modulation in 1977–1980. *J. Geophys. Res. Space Phys.* 89, 6579–6587. doi:10.1029/JA089iA08p06579

Burlaga, L. F., and Ness, N. F. (2016). Observations of the interstellar magnetic field in the outer heliosheath: *VOYAGER 1*. *Astrophysical J.* 829, 134. doi:10.3847/0004-637X/829/2/134

Chalov, S. V., and Fahr, H. J. (1996). Reflection of pre-accelerated pick-up ions at the solar wind termination shock: the seed for anomalous cosmic rays. *Sol. Phys.* 168, 389–411. doi:10.1007/BF00148064

Drake, J. F., Opher, M., Swisdak, M., and Chamoun, J. N. (2010). A magnetic reconnection mechanism for the generation of anomalous cosmic rays. *Astrophysical J.* 709, 963–974. doi:10.1088/0004-637X/709/2/963

Eichler, D. (1979). Particle acceleration in collisionless shocks: regulated injection and high efficiency. *Astrophysical J.* 229, 419. doi:10.1086/156969

Fermo, R. L., Pogorelov, N. V., and Burlaga, L. F. (2015). Transient shocks beyond the heliopause. *J. Phys. Conf. Ser.* 642, 012008. doi:10.1088/1742-6596/642/1/012008

Fisk, L. A., Kozlovsky, B., and Ramaty, R. (1974). An interpretation of the observed oxygen and nitrogen enhancements in low-energy cosmic rays. *Astrophysical J.* 190, L35. doi:10.1086/181498

Florinski, V., Decker, R. B., le Roux, J. A., and Zank, G. P. (2009). An energetic-particle-mediated termination shock observed by Voyager 2. *Geophys. Res. Lett.* 36, L12101. doi:10.1029/2009GL038423

Florinski, V., Zank, G. P., Jokipii, J. R., Stone, E. C., and Cummings, A. C. (2004). Do anomalous cosmic rays modify the termination shock? *Astrophysical J.* 610, 1169–1181. doi:10.1086/421901

Giacalone, J., Fahr, H., Fichtner, H., Florinski, V., Heber, B., Hill, M. E., et al. (2022). Anomalous cosmic rays and heliospheric energetic particles. *Space Sci. Rev.* 218, 22. doi:10.1007/s11214-022-00890-7

Gloeckler, G., Geiss, J., Roelof, E. C., Fisk, L. A., Ipavich, F. M., Ogilvie, K. W., et al. (1994). Acceleration of interstellar pickup ions in the disturbed solar wind observed on ulysses. *J. Geophys. Res.* 99, 17637–17643. doi:10.1029/94JA01509

Guo, X. (2015). An extended hllc riemann solver for the magneto-hydrodynamics including strong internal magnetic field. *J. Comput. Phys.* 290, 352–363. doi:10.1016/j.jcp.2015.02.048

Guo, X., Florinski, V., and Wang, C. (2018). Effects of anomalous cosmic rays on the structure of the outer heliosphere. *Astrophysical J.* 859, 157. doi:10.3847/1538-4357/aabf42

Guo, X., Florinski, V., and Wang, C. (2019). A global MHD simulation of outer heliosphere including anomalous cosmic-rays. *Astrophysical J.* 879, 87. doi:10.3847/1538-4357/ab262b

Guo, X., Zhou, Y., Florinski, V., and Wang, C. (2022). Dynamical coupling between anomalous cosmic rays and solar wind in outer heliosphere. *Astrophysical J.* 935, 144. doi:10.3847/1538-4357/ac82ed

Guo, X., Zhou, Y., Wang, C., Ying, D., and Liu, A. (2021). PROPAGATION OF LARGE-SCALE SOLAR WIND EVENTS IN THE OUTER HELIOSPHERE FROM A NUMERICAL MHD SIMULATION. *Earth Planet. Phys.* 5, 0–231. doi:10.26464/epp2021024

Gurnett, D. A., Kurth, W. S., Stone, E. C., Cummings, A. C., Krimigis, S. M., Decker, R. B., et al. (2015). PRECURSORS TO INTERSTELLAR SHOCKS OF SOLAR ORIGIN. *Astrophysical J.* 809, 121. doi:10.1088/0004-637X/809/2/121

Kim, T. K., Pogorelov, N. V., and Burlaga, L. F. (2017). Modeling shocks detected by *voyager 1* in the local interstellar medium. *Astrophysical J.* 843, L32. doi:10.3847/2041-8213/aa7b2b

Krimigis, S. M., Decker, R. B., Roelof, E. C., Hill, M. E., Armstrong, T. P., Gloeckler, G., et al. (2013). Search for the exit: voyager 1 at heliosphere's border with the galaxy. *Science* 341, 144–147. doi:10.1126/science.1235721

Lazarus, A., Richardson, J., Decker, R., and Mcdonald, F. (1999). Voyager 2 observations of corotating interaction regions (cirs) in the outer heliosphere. *Space Sci. Rev.* 89, 53–59. doi:10.1023/A:1005252113717

Liu, Y. D., Richardson, J. D., Wang, C., and Luhmann, J. G. (2014). PROPAGATION OF THE 2012 MARCH CORONAL MASS EJECTIONS FROM THE SUN TO HELIOPAUSE. *Astrophysical J.* 788, L28. doi:10.1088/2041-8205/788/2/L28

McComas, D., and Schwadron, N. (2006). An explanation of the voyager paradox: particle acceleration at a blunt termination shock. *Geophys. Res. Lett.* 33. doi:10.1029/2005GL025437

McComas, D. J., Zirnstein, E. J., Bzowski, M., Elliott, H. A., Randol, B., Schwadron, N. A., et al. (2017). Interstellar pickup ion observations to 38 au. *Astrophysical J. Suppl. Ser.* 233, 8. doi:10.3847/1538-4365/aa91d2

Mostafavi, P., Burlaga, L. F., Cairns, I. H., Fuselier, S. A., Fraternale, F., Gurnett, D. A., et al. (2022). Shocks in the very local interstellar medium. *Space Sci. Rev.* 218, 27. doi:10.1007/s11214-022-00893-4

Parker, E. N. (1961). The stellar-wind regions. *Astrophysical J.* 134, 20. doi:10.1086/147124

Pauls, H. L., Zank, G. P., and Williams, L. L. (1995). Interaction of the solar wind with the local interstellar medium. *J. Geophys. Res.* 100, 21595–21604. doi:10.1029/95JA02023

Pesses, M. E., Jokipii, J. R., and Eichler, D. (1981). Cosmic ray drift, shock wave acceleration, and the anomalous component of cosmic rays. *Astrophysical J. Lett.* 246, L85. doi:10.1086/183559

Rice, W., and Zank, G. (2000). Shock propagation in the outer heliosphere: 3. pickup ions, mhd, cosmic rays, and energetic particles. *J. Geophys. Res. Space Phys.* 105, 5157–5166. doi:10.1029/1999JA000337

Rice, W., Zank, G., Richardson, j., and Decker, R. (2000). Ion injection and shock acceleration in the outer heliosphere. *Geophys. Res. Lett.* 27, 509–512. doi:10.1029/1999GL010731

Richardson, J. D., Kasper, J. C., Wang, C., Belcher, J. W., and Lazarus, A. J. (2008). Cool heliosheath plasma and deceleration of the upstream solar wind at the termination shock. *Nature* 454, 63–66. doi:10.1038/nature07024

Richardson, J. D., Paularena, K. I., Belcher, J. W., and Lazarus, A. J. (1994). Solar wind oscillations with a 1.3 year period. *Geophys. Res. Lett.* 21, 1559–1560. doi:10.1029/94GL01076

Richardson, J. D., Wang, C., Liu, Y. D., Šafránková, J., Němeček, Z., and Kurth, W. S. (2017). PRESSURE PULSES AT *VOYAGER 2*: DRIVERS OF INTERSTELLAR TRANSIENTS? *Astrophysical J.* 834, 190. doi:10.3847/1538-4357/834/2/190

Stone, E. C., Cummings, A. C., Heikkila, B. C., and Lal, N. (2019a). Cosmic ray measurements from Voyager 2 as it crossed into interstellar space. *Nat. Astron.* 3, 1013–1018. doi:10.1038/s41550-019-0928-3

Stone, E. C., Cummings, A. C., Heikkila, B. C., and Lal, N. (2019b). Cosmic ray measurements from voyager 2 as it crossed into interstellar space. *Nat. Astron.* 3, 1013–1018. doi:10.1038/s41550-019-0928-3

Stone, E. C., Cummings, A. C., McDonald, F. B., Heikkila, B. C., Lal, N., and Webber, W. R. (2005). Voyager 1 explores the termination shock region and the heliosheath beyond. *Science* 309, 2017–2020. doi:10.1126/science.1117684

Stone, E. C., Cummings, A. C., McDonald, F. B., Heikkila, B. C., Lal, N., and Webber, W. R. (2013). Voyager 1 observes low-energy galactic cosmic rays in a region depleted of heliospheric ions. *Science* 341, 150–153. doi:10.1126/science.1236408

Terasawa, T., Oka, M., Nakata, K., Keika, K., Nosé, M., McEntire, R. W., et al. (2006). Cosmic-ray-mediated' interplanetary shocks in 1994 and 2003. *Adv. Space Res.* 37, 1408–1412. doi:10.1016/j.asr.2006.03.012

van Leer, B. (1979). Towards the ultimate conservative difference scheme. v. a second-order sequel to godunov's method. *J. Comput. Phys.* 32, 101–136. doi:10.1016/0021-9991(79)90145-1

Wang, B.-B., Zank, G. P., Zhao, L.-L., and Adhikari, L. (2022). Turbulent cosmic ray–mediated shocks in the hot ionized interstellar medium. *Astrophysical J.* 932, 65. doi:10.3847/1538-4357/ac6ddc

Wang, C., and Richardson, J. D. (2001). Energy partition between solar wind protons and pickup ions in the distant heliosphere: a three-fluid approach. *J. Geophys. Res.* 106, 29401–29407. doi:10.1029/2001JA000190

Wang, C., Richardson, J. D., and Gosling, J. T. (2000). A numerical study of the evolution of the solar wind from Ulysses to Voyager 2. *J. Geophys. Res. Space Phys.* 105, 2337–2344. doi:10.1029/1999JA900436

Washimi, H., Zank, G. P., Hu, Q., Tanaka, T., Munakata, K., and Shinagawa, H. (2011). Realistic and time-varying outer heliospheric modelling: realistic outer heliospheric modelling. *Mon. Notices R. Astronomical Soc.* 416, 1475–1485. doi:10.1111/j.1365-2966.2011.19144.x

Zank, G. P., Hunana, P., Mostafavi, P., le Roux, J., Li, G., Webb, G., et al. (2015). Particle acceleration by combined diffusive shock acceleration and downstream multiple magnetic island acceleration. *J. Phys. Conf. Ser.* 642, 012031. doi:10.1088/1742-6596/642/1/012031

Zank, G. P., Webb, G. M., and Donohue, D. J. (1993). Particle injection and the structure of energetic-particle–modified shocks. *Astrophysical J.* 406, 67. doi:10.1086/172421

Zhao, L. L., Adhikari, L., Zank, G. P., Hu, Q., and Feng, X. S. (2017). Cosmic ray diffusion tensor throughout the heliosphere derived from a nearly incompressible magnetohydrodynamic turbulence model. *Astrophysical J.* 849, 88. doi:10.3847/1538-4357/aa932a

Zhao, L.-L., Adhikari, L., Zank, G. P., Hu, Q., and Feng, X. S. (2018). Influence of the solar cycle on turbulence properties and cosmic-ray diffusion. *Astrophysical J.* 856, 94. doi:10.3847/1538-4357/aab362

Zhao, L. L., Zank, G. P., Hu, Q., Chen, Y., Adhikari, L., leRoux, J. A., et al. (2019). Acr proton acceleration associated with reconnection processes beyond the heliospheric termination shock. *Astrophysical J.* 886, 144. doi:10.3847/1538-4357/ab4db4

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Review
PUBLISHED 08 March 2024
DOI 10.3389/fspas.2024.1368043

Check for updates

OPEN ACCESS

EDITED BY
Gang Li,
University of Alabama in Huntsville,
United States

REVIEWED BY
Subhamoy Chatterjee,
Southwest Research Institute Boulder, United States
Athanasios Papaioannou,
National Observatory of Athens, Greece
Yang Wang,
Harbin Institute of Technology, China

*CORRESPONDENCE
Donald V. Reames,
dvreames@gmail.com

RECEIVED 09 January 2024
ACCEPTED 16 February 2024
PUBLISHED 08 March 2024

CITATION
Reames DV (2024), Element abundance and the physics of solar energetic particles.
Front. Astron. Space Sci. 11:1368043.
doi: 10.3389/fspas.2024.1368043

COPYRIGHT
© 2024 Reames. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Element abundance and the physics of solar energetic particles

Donald V. Reames *

Institute for Physical Science and Technology, University of Maryland, College Park, MD, United States

The acceleration and transport of solar energetic particles (SEPs) cause their abundance, measured at a constant velocity, to be enhanced or suppressed as a function of the magnetic rigidity of each ion, and hence, of its atomic mass-to-charge ratio of *A/Q*. Ion charges, in turn, depend upon the source electron temperature. In small "impulsive" SEP events, arising from solar jets, acceleration during magnetic reconnection causes steep power-law abundance enhancements. These impulsive SEP events can have 1,000-fold enhancements of heavy elements from sources at ~2.5 MK and similar enhancements of ^{3}He/^{4}He and of streaming electrons that drive type-III radio bursts. Gamma-ray lines show that solar flares also accelerate ^{3}He-rich ions, but their electrons and ions remain trapped in magnetic loops, so they dissipate their energy as X-rays, γ-rays, heat, and light. "Gradual" SEPs accelerated at shock waves, driven by fast coronal mass ejections (CMEs), can show power-law abundance enhancements or depressions, even with seed ions from the ambient solar corona. In addition, shocks can reaccelerate seed particles from residual impulsive SEPs with their pre-existing signature heavy-ion enhancements. Different patterns of abundance often show that heavy elements are dominated by a source different from that of H and He. Nevertheless, the SEP abundance, averaged over many large events, defines the abundance of the corona itself, which differs from the solar photosphere as a function of the first ionization potential (FIP) since ions, with FIP <10 eV, are driven upward by forces of electromagnetic waves, which neutral atoms, with FIP >10 eV, cannot feel. Thus, SEPs provide a measurement of element abundance in the solar corona, distinct from solar wind, and may even better define the photosphere for some elements.

KEYWORDS

solar energetic particles, solar system abundance, solar jets, solar flares, shock waves, coronal mass ejections, heliosphere

1 Introduction

The relative abundance of chemical elements in any sample of material can be a clue to the identity and origin of that sample and to the nature of the physical processes it has undergone. Energetic particles are no exception. Abundances reveal the age of the galactic cosmic rays (GCRs) and the origin of unusual ions trapped in planetary magnetospheres. Solar energetic particles (SEPs) also display unique signature patterns of abundance that help us distinguish the physical processes that have formed them and the history they have traversed (Reames, 1988; Reames, 1999; Reames, 2013; Reames, 2021a; Reames, 2021b; Reames, 2021c).

A most unusual feature of SEP abundances is small ^{3}He-rich events with 1,000-fold enhancements of ^{3}He/^{4}He, which were later found to have enhancements of heavy elements extending as powers of the element atomic mass-to-charge ratio A/Q from C and O to up to elements as heavy as Pb, also by a factor of ~1,000 (e.g., Reames et al., 2014a). These SEP ions have been associated with magnetic reconnection in solar jets and flares (Kahler et al., 2001; Reames, 2013; Bučík, 2020; Reames, 2021a; Reames, 2021b; Reames, 2021c). In some larger SEP events, these "impulsive SEP" abundances often emerge as a signature of residual impulsive ions that have been reaccelerated by shock waves and exceed the average coronal abundances in some "gradual" SEP events (Reames, 1999; Desai and Giacalone, 2016; Reames, 2021a; Reames, 2021b). This not only divides element abundances of SEP events into haves and have-nots but also highlights variations in H and He, which can either participate or not participate in the heavy-element behavior (Reames, 2022b).

A most fundamental population that SEPs can measure is the abundance of elements in the corona itself, from which all SEPs are derived. Only in the corona are the densities low enough for ions to be accelerated without immediately losing their energy in Coulomb collisions. The corona not only provides a baseline for identifying other populations derived from it but also highlights the physical process that distinguishes it from the photosphere. We first discuss these reference coronal abundances and then the unique abundances of impulsive SEPs, followed by their presence or absence in the largest "gradual" SEP events. SEP abundances in this work are mainly derived from the Low-Energy Matrix Telescope (LEMT) on the *Wind* spacecraft (von Rosenvinge et al., 1995); ~30 years of LEMT abundance data are available at https://omniweb.gsfc.nasa.gov/.

2 Reference abundances, solar corona, and the first ionization potential

The abundances of elements C, O, and the others mentioned above, in SEPs, were first measured using nuclear emulsion detectors on sounding rockets from Ft. Churchill, Manitoba by Fichtel and Guss (1961), and those measurements were later extended up through Fe using the same technique (Bertsch et al., 1969). As satellite measurements became available (e.g., Teegarden et al., 1973), comparisons of SEPs with other abundances became more common (e.g., Webber, 1975). Meyer (1985) summarized SEP abundance measurements in large SEP events as having a common baseline, derived from abundances in the solar corona where acceleration occurs, and a second component that varied from event to event as a power law in the particle charge-to-mass ratio Q/A (Breneman and Stone, 1985). A factor in this abundance variation was the pitch-angle scattering of the ions; thus, if Fe scatters less than O, Fe/O will be enhanced early in events but will be depleted later. Such variations might average out, so it was soon possible to average ~50 or so large SEP events to remove the event-to-event variations and produce estimates of coronal abundances (Reames, 1995a; Reames, 2014); this could be compared with the solar photospheric abundances measured spectroscopically. A modern comparison of the SEPs/photospheric abundances vs. the first ionization potential (FIP) is shown in Figure 1. The photospheric abundances are a modification by Caffau et al. (2011) of the meteoritic abundances proposed by Lodders et al. (2009). A comparison using photospheric abundances proposed by Asplund et al. (2021) is shown, and the abundances are listed by Reames (2021b).

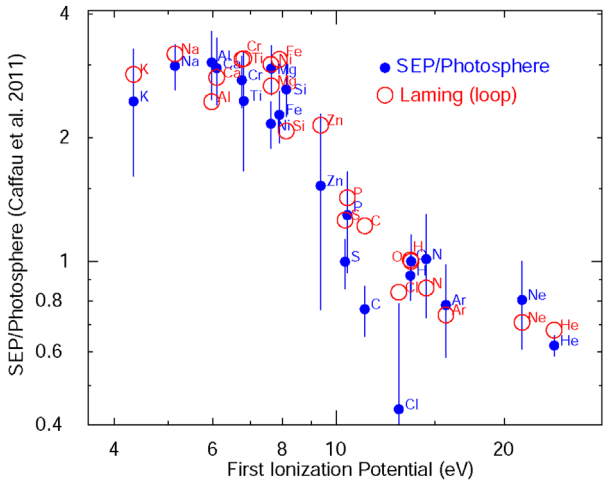

FIGURE 1
The ratio of solar energetic particles (SEPs) to photospheric abundances of elements is shown vs. the first ionization potential **(solid blue circles)** and is compared with the theory proposed by Laming et al. (2019) **(open red circles)** for element transport along closed loops in active regions.

In the theory proposed by Laming (2015) and Laming et al. (2019), shown in Figure 1, the ponderomotive force of Alfvén waves helps drive low-FIP ions up across the chromosphere into the corona, but it cannot affect un-ionized high-FIP neutral atoms. All elements become ionized in the hot ~1 MK corona. SEPs have a different FIP pattern from that of solar wind or the solar wind accelerated by shock waves at co-rotating stream interfaces (Reames et al., 1991; Reames, 1995a; Mewaldt et al., 2002; Reames, 2018a; Reames, 2021a); thus, SEP abundances do not differ when measured in fast or slow wind (Kahler et al., 2009). SEPs are *not* accelerated solar wind. Differences between the FIP patterns of SEPs and solar wind may be caused by open vs. closed field lines, where Alfvén waves resonate with the loop length of closed loops (Reames, 2018a; Laming et al., 2019). A first-order examination of Figure 1 shows a reasonable agreement when comparing elements with a similar FIP but different A/Q, e.g., Mg or Si with Fe or Ni.

To what extent does this average of abundances, over many gradual SEP events, recapture coronal abundances? It is quite possible that the increasing and decreasing power laws of abundance vs. A/Q do not perfectly average out; however, a more outstanding disagreement seems to be the single element C. How can a single element, or actually a single ratio C/O, stand out? We will return to this question after discussing the patterns of known abundance variations and their probable causes.

3 Impulsive SEP events

The idea of two fundamental mechanisms of SEP acceleration began quite early (Wild et al., 1963) with solar radio observations that distinguished the sources of type-II and type-III radio bursts. The radio emission frequency depends upon the square root of the local electron density, which decreases with distance from the Sun. Type-III radio bursts exhibit a rapid frequency decrease corresponding to the speed of 10–100 keV electrons streaming out from the Sun, while type-II bursts show a much slower frequency decrease of ~1,000 km s^{-1} shock waves driven out from the Sun. These streaming electrons propagate scatter-free because the resonant turbulence that would scatter them is absorbed by the plasma (Tan et al., 2011). Later, Lin (1970) observed beams of ~40 keV electrons associated with the impulsive type-III bursts and thought that they might involve "pure" electron events, i.e., without ions. Relativistic electrons and energetic protons only accompanied the shock-associated type-II bursts.

Soon, the SEP world was surprised by the observations of ^{3}He-rich events. While a typical solar or solar wind abundance is ^{3}He/^{4}He $\approx 5 \times 10^{-4}$, an event was soon found with ^{3}He/^{4}He = 1.5 ± 0.1 (Serlemitsos and Balasubrahmanyan, 1975; Mason, 2007). Such a high ratio could not come from the fragmentation of ^{4}He as that which occurred in GCRs since ^{3}He was not accompanied by any ^{2}H. The early idea of nuclear fragmentation was completely laid to rest by the later measurements of Be/O and B/O $<2 \times 10^{-4}$ (McGuire et al., 1979; Cook et al., 1984). Instead, this was a new mechanism involving resonant wave–particle interactions. The ^{3}He gyrofrequency, dependent upon Q/A, lay isolated at Q/A = 2/3, between those of H at Q/A = 1 and ^{4}He at Q/A = 1/2.

These two different features of impulsive SEPs, i.e., 1) "pure" electron beams producing type-III radio bursts and 2) ^{3}He-rich events, were unified by Reames et al. (1985) as different properties of the same events, and Reames and Stone (1986) explored the kilometric radio properties of ^{3}He-rich events, even tracking the flow of electrons out from the Sun. Early theories discussed selective heating by the resonant absorption of various types of plasma waves at the ^{3}He gyrofrequency, followed by the acceleration of thermal tails by some unspecified mechanism (e.g., Fisk, 1978; see other references Reames, 2021c or Reames, 2023c), but Temerin and Roth (1992) proposed electromagnetic ion cyclotron (EMIC) waves generated by streaming electron beams and added their absorption by mirroring ^{3}He for acceleration, in analogy with the production of ion conics observed in the Earth's magnetosphere.

3.1 Element abundance

Enhancements of heavy elements up to Fe in impulsive events were first reported by Mogro-Comparo and Simpson (1972). These observations were improved in subsequent generations of experiments by Mason et al. (1986) and then by Reames et al. (1994). Groups of elements were eventually resolved up to Pb at 3–10 MeV amu^{-1} (Reames, 2000; Reames and Ng, 2004; Reames et al., 2014a) and below ~1 MeV amu^{-1} (Mason et al., 2004). The average enhancement was found to be a power law in A/Q, with a power of 3.64 ± 0.15 above 1 MeV amu^{-1} and ~3.26 below, using Q-values appropriate for ~3 MK.

The early direct measurements of the ionization states of SEP elements up to Fe (Luhn et al., 1984; Luhn et al., 1987) found that Q_{Fe} = 14.1 ± 0.2 in gradual SEP events, which would correspond to a typical source plasma temperature of ~2 MK, but ^{3}He-rich events had Q_{Fe} = 20.5 ± 1.2, with elements up to Si being fully ionized, either suggesting a temperature >10 MK or the stripping of ions by passing through a small amount of material after acceleration. Reames et al. (1994) noted that impulsive ion enhancements, relative to the corona, formed three groups: 1) C, N, and O, like ^{4}He, all seemed to be un-enhanced, 2) Ne, Mg, and Si were enhanced about a similar factor of ~2.5, and 3) Fe was enhanced by a factor of ~7. The first group was probably fully ionized with A/Q = 2, while the second group would have similar abundances in their states with 2 orbital electrons, which occur at approximately 3 MK. No enhancements could occur if Ne, Mg, and Si were fully ionized, as measured. The resolution of this dilemma is that the ions in impulsive SEP events are stripped after acceleration, and it was later found that the ionization states of Fe depended upon the ion velocity (DiFabio et al., 2008), as expected from stripping. This suggested that impulsive SEP events were accelerated at ~1.5 R_S. In contrast, gradual SEP events, found to be accelerated at shock waves driven out from the Sun by coronal mass ejections (CMEs), began at 2–3 R_S (Tylka et al., 2003; Cliver et al., 2004; Reames, 2009a; Reames, 2009b).

3.2 Jets and flares

While gradual SEP events have a 96% correlation with fast, wide CMEs (Kahler et al., 1984), an early search found no meaningful association of ^{3}He-rich events with CMEs observed using the Solwind coronagraph (Kahler et al., 1985). However, with improved coronagraph sensitivity of SOHO/LASCO, Kahler et al. (2001) found narrow CMEs that were associated with the larger ^{3}He-rich events; the CME associated with the large impulsive SEP event of 1 May 2000 had a speed of 1,360 km s^{-1}, easily fast enough to drive a shock that would reaccelerate particles. These observations led Kahler et al. (2001) to associate impulsive SEP events with solar jets (e.g., Shimojo and Shibata, 2000), an association that has been extended (Nitta et al., 2006; Wang et al., 2006; Bučík et al., 2018a; Bučík et al., 2018b) and reviewed by Bučík (2020). Jets are driven by energy from magnetic reconnection, as shown in Figure 2.

A particle-in-cell simulation of a reconnection region by Drake et al. (2009) found strong A/Q-dependent enhancements in the energetic heavy ions that could match those observed. The particles were Fermi-accelerated as they were reflected (mirrored) back and forth by the approaching ends of the collapsing islands of magnetic reconnection. However, the reconnection that opens some field lines always closes others, as shown in the lower left of Figure 2. These closing field lines trap newly accelerated particles that deposit their energy as heat or in the emission of X-rays or γ-rays—a solar flare. Thus, jets would always be expected to have accompanying flares that involve the same accelerated particles. Of course, there are more realistic models of jets (e.g., Archontis and Hood, 2013; Lee et al., 2015; Pariat et al., 2015) that better describe the CME emission, but jet models do not yet include SEP acceleration.

The relationship of jets vs. flares is a close one, with similar ion acceleration on open vs. closed field lines. Similar ion enhancements

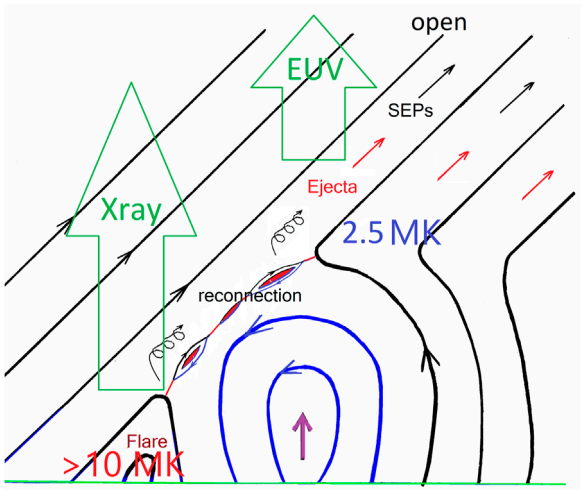

FIGURE 2
Sketches of the topology of a solar jet, where rising closed field lines of one magnetic polarity **(blue)** form islands of reconnection where they meet the oppositely directed open field lines **(black)**. SEPs are accelerated in the reconnection region and escape along the open field lines, as does the CME plasma. Newly formed closed field lines at the lower left trap energetic electrons and ions that plunge into the deeper corona and form a flare. The hot flaring region (>10 MK) emits X-rays, while the open region is observed to be an EUV-emitting region at ~2.5 MK (Reames, 2023c).

were first noted between impulsive SEPs and abundances from γ-ray line measurements in large flares by Murphy et al. (1991); then, Mandzhavidze et al. (1999) found that the energetic ions that were accelerated and trapped in all 20 available large solar flares were ^{3}He-rich. The three γ-ray lines at 0.937, 1.04, and 1.08 MeV from the de-excitation of $^{19}F^*$ were unusually strong and were produced with an especially high cross-section in the reaction ^{16}O (^{3}He, p) $^{19}F^*$. These were compared with other lines from excited ^{16}O, ^{20}Ne, and ^{56}Fe to distinguish ^{3}He from ^{4}He in the "beam." Some of the events had ^{3}He/^{4}He of ~1, while all had ^{3}He/^{4}He >0.1. Murphy et al. (2016) later found 6 key ratios of γ-ray fluxes that were dependent upon ^{3}He/^{4}He in the beam, and all of them showed an average ^{3}He/^{4}He ratio of 0.05–3.0. These studies included ~135 product de-excitation lines from ~300 proton- and He ion-induced reactions (e.g., Kozlovsky et al., 2002). These γ-ray lines are from the largest flares, not small jet-associated impulsive flares, suggesting that impulsive SEP abundances are a general consequence of the physics of magnetic reconnection.

3.3 Power-law abundance from jets with or without shocks

We always compare element abundances at the same velocity, or MeV amu^{-1}, but properties such as magnetic deflection and scattering depend upon magnetic rigidity, or the momentum per unit charge, also dependent upon A/Q, quite often depends upon a power of A/Q (e.g., Parker, 1965). Thus, it is not surprising that enhancements that depart from reference abundances vary as a power of A/Q. More specifically for impulsive SEPs, the theory proposed by Drake et al. (2009) related the power of A/Q to the power of the width distribution of reconnecting magnetic islands. However, the Q-values of the ions depend upon the source plasma temperature. Our strategy has simply been to try all temperatures in a reasonable range, determine the Q-values (using, e.g., Mazzotta et al., 1998 or Post et al., 1977), and choose the temperature and power law that yield the best least-squares fit of enhancement vs. A/Q (Reames et al., 2014b; Reames, 2018b; Reames, 2021a). Examples showing typical temperature dependence in enhancement vs. A/Q are shown in Figure 6 of the study by Reames (2022b), Figure 6 of the study by Reames (2018a), and Figure 2 of the study by Reames et al. (2014b).

Fitting 111 impulsive SEP events, Reames et al. (2014b) found 79 events at 2.5 MK and 29 at the neighboring 3.2 MK, i.e., very little variation in impulsive SEP events. Subsequently, these temperatures agree with the EUV temperatures in jets (Bučík et al., 2021). Figure 3 shows power-law fits to abundance enhancements in several impulsive SEP events. The time profiles of the events are shown in the lower panels, derived temperatures (and event durations) in the middle panels, and the best fits to the enhancements vs. A/Q in the upper panels. Event numbers marking each event onset in Figure 3 correspond to the impulsive SEP event list given by Reames et al. (2014a), along with all the events selected to have enhanced Fe/O abundances.

For events 3 and 4 in Figure 3, the power-law fits, obtained for ions with $Z \geq 6$, extend to include proton measurements at $A/Q = 1$. This is taken to mean that all of these ions come from the same population, i.e., the magnetic reconnection in the associated jet. Reames (2020a) defined these "pure" reconnection events as SEP1 events; they had either no visible CMEs or CME speeds <500 km s^{-1}, i.e., no shock acceleration was likely. Event 5 is ambiguous; the excess protons do not fit the power law, but the theory proposed by Drake et al. (2009) allows for enhancements that start above $Z = 2$, and H and ^{4}He could both be un-enhanced, i.e., at the same level, as in this event, which also shows no CMEs.

Earlier, there were thought to be only two types of SEP events, ^{3}He-rich or "impulsive" SEP events with unique element abundances and shock-accelerated "gradual" SEP events that accelerated coronal ions. Abundances in the gradual events varied primarily because of the differences in element transport; since Fe scatters less than O, Fe/O will be enhanced early and depressed later, following a power law in A/Q. This was observed by Breneman and Stone (1985) and implied in the discussions by Meyer (1985). This simplicity ended when Mason et al. (1999) found the enhancement of ^{3}He in a large gradual SEP event. Clearly, shocks could reaccelerate residual ions from small impulsive and ambient coronal ions, and these two seed populations became widely discussed (Tylka et al., 2001; Tylka et al., 2005; Desai et al., 2003; Tylka and Lee, 2006). Eventually, Reames (2020a) suggested organizing the combination of acceleration mechanisms and seed populations into four physical categories, as shown in Table 1.

However, event 92, as shown in Figure 3F, is an event with large proton excess and even an excess of ^{4}He; it suggests power-law contributions that are shock-accelerated from two seed populations, ambient coronal ions for H and ^{4}He and residual impulsive suprathermal ions for $Z \geq 6$. The event would be classified as SEP2 if all the SEP1 ions came from a single impulsive jet event and as SEP3 if the pre-accelerated impulsive seed ions had been collected from many previous SEP1 or SEP2 events before shock acceleration. In

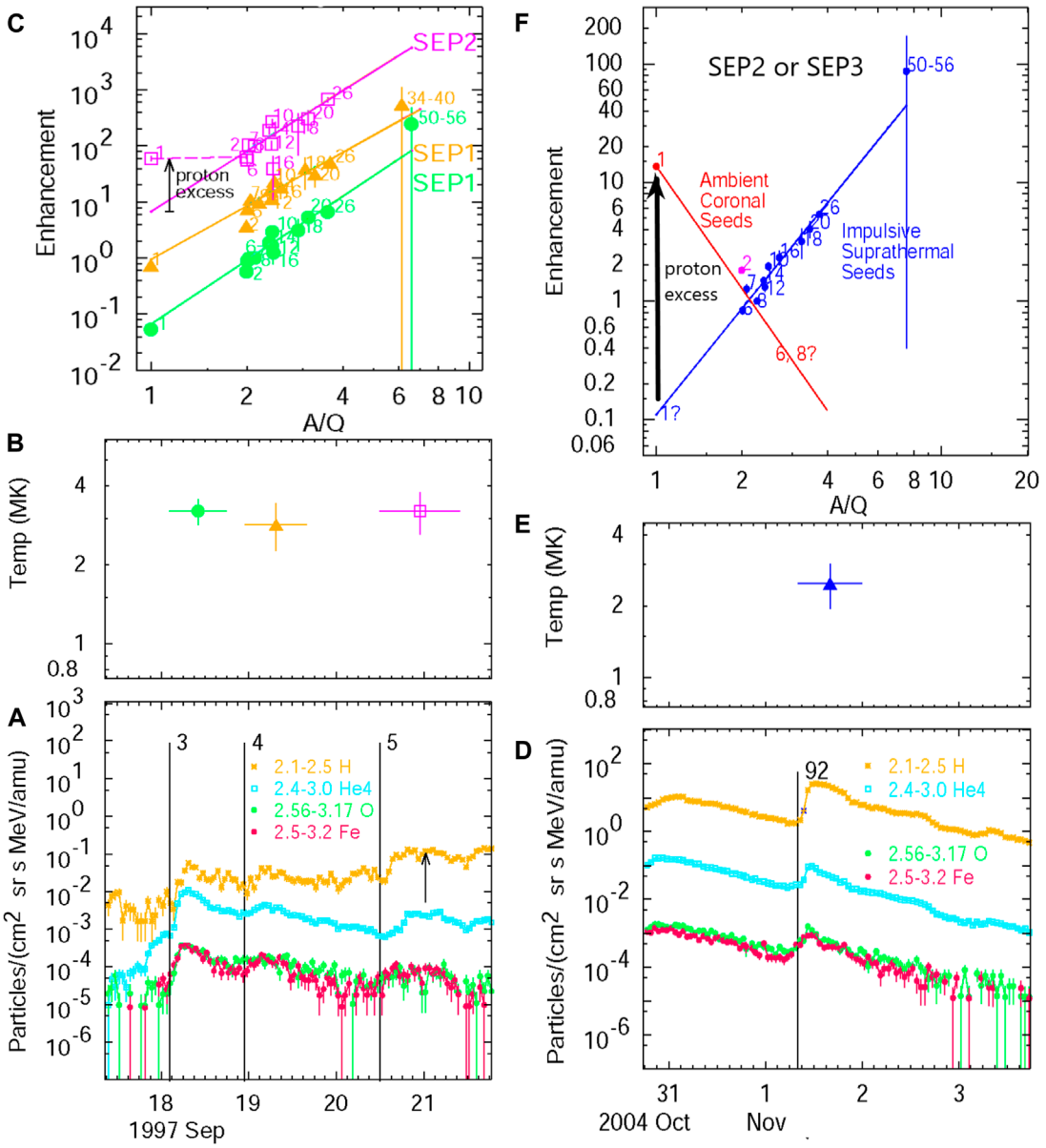

FIGURE 3
Panels **(A** and **D)** show intensities of H, ^{4}He, O, and Fe at the listed energies in MeV amu^{-1} for **(A)**, which is a sequence of three small impulsive SEP events, and **(D)**, a larger event. Panels **(B** and **E)** show the derived best-fit temperatures (and durations) for each event below; panels **(C** and **F)** show the corresponding best-fit power-law abundance enhancements (normalized at O and shifted ×10 for each event), with the measurements for each element labeled by atomic numbers Z. Only elements with $Z \geq 6$ are included in the fits. Panel **(F)** distinguished two possible seed populations, contributing to acceleration in the shock wave driven by the 925-km s^{-1}-associated CME (see text).

fact, this seems to be a SEP3 event, despite its short duration, since its intensity is high, and an earlier event in the same location is similar in character. It should also be noted that the scatter of $Z \geq 6$ points about the fit line, shown in Figure 3F, is quite small compared with those in Figure 3C, suggesting that the output of many small jets have been averaged to reduce abundance variations in SEP3 events (see Figure 8 in the study by Reames, (2020a)). We consider other SEP2 and SEP3 examples below.

The abundance of ^{4}He in SEP events can vary because it can be dominated by either coronal or impulsive seed components. However, there are also other variations in ^{4}He that have been summarized in greater detail by Reames (2022b).

4 Gradual SEP events

The first recognized SEP events (Forbush, 1946) were the immense ground-level enhancement (GLE) events, where GeV protons initiate nuclear cascades through the atmosphere that exceed those of GCRs. While these events were erroneously

TABLE 1 Properties of four sources of SEPs.

	Observed properties	Physical association
SEP1	Fe-rich power-law enhancement vs. A/Q at all Z; T ≈ 2.5 MK	Magnetic reconnection in solar jets with no fast shock
SEP2	Fe-rich power-law enhancement vs. A/Q at Z >2; T ≈ 2.5 MK. Proton excess ~×10. CME speed >500 km/s	Jets with fast, narrow CMEs drive shocks that reaccelerate local SEP1 seeds to dominate high Z and ambient plasma to dominate H (and ^{4}He)
SEP3	Fe-rich power-law enhancement vs. A/Q at Z >2; T ≈ 2.5 MK. Proton excess ~×10. CME speed >>500 km/s	Fast, wide CME-driven shocks accelerate the SEP1 residue left by many jets in active-region pools, plus H (and ^{4}He) from ambient plasma at low Z
SEP4	Power-law or flat vs. A/Q for all ions with 0.8 <T <1.8 MK. Fast, wide CMEs	Very fast, wide CME-driven shocks accelerate all dominant ions as seeds from the ambient plasma

attributed to solar flares for many years (Gosling, 1993; Gosling, 1994), Kahler et al. (1984) had found that large SEP events had a 96% association with fast, wide shock waves, driven out from the Sun by CMEs, reaffirming the finding of shock acceleration in radio type-II bursts by Wild et al. (1963) two decades earlier. Mason et al. (1984) concluded that only "large-scale shock acceleration" could explain the extensive rigidity-independent spread of SEPs, and recent findings from missions like STEREO now show how shock waves and SEPs wrap around the Sun (e.g., Reames, 2023a; Reames, 2023b). Evidence for shock acceleration has grown (e.g., Reames, 1995b; Reames, 1999; Zank et al., 2000; Zank et al., 2007; Lee et al., 2012; Reames, 2013; Desai and Giacalone, 2016; Kouloumvakos et al., 2019; Reames, 2021b), especially for GLEs (Tylka and Dietrich, 2009; Gopalswamy et al., 2012; Mewaldt et al., 2012; Gopalswamy et al., 2013a; Raukunen et al., 2018).

One line of evidence has been the onset timing or solar particle release (SPR) time of the SEPs compared with X-ray or γ-ray onset times of associated flares. Ions with lower velocities have increasingly delayed onsets that extrapolate back to a single SPR time, with delay = distance along the observer's field line divided by velocity. For impulsive events, the SPR times and X-ray onsets agree closely (Tylka et al., 2003), but for gradual events like GLEs, the SPR time can lag the X-ray and γ-ray onset by as much as half an hour (Tylka et al., 2003; Reames, 2009a; Reames, 2009b), sometimes even after the associated flare is completely over. The SPR time corresponds to the time the shock at the leading edge of the CME reaches 2–3 solar radii (Reames, 2009a; Reames, 2009b; Cliver et al., 2004), presumably when the shock emerges above closed magnetic loops and its speed exceeds the declining Alfvén speed. Type-II radio emission shows that shocks can begin at ~1.5 AU (Gopalswamy et al., 2013b), but the SPR time depends upon the observer's longitude, which often differs from that of the ewarliest source of radio emission (Reames, 2009b). Variations in these parameters or delays in the shock interception of the observer's field line (Reames, 2023a; Reames, 2023b) cause variations in SPR delay.

Abundance enhancements or suppressions in gradual SEP events relative to the reference coronal abundances have been classed as SEP3 or SEP4 by Reames (2020a), depending upon whether the source of seed particles for shock acceleration is purely ambient coronal abundances (SEP4) or whether the residual impulsive suprathermal ions dominate the heavy elements (SEP3).

4.1 Moderate-sized SEP4 events

Figure 4 shows abundance data for two typical SEP4 events. These gradual events are intense enough to measure significant enhancements of most elements in several time intervals, e.g., every 8 h. The derived temperatures are less than those in impulsive events, and the abundance patterns vary little during the events, suggesting that there is too little variation in scattering to separate different elements in time. Most importantly, the power-law fits, obtained for elements C and above, can be extended to fit H and He reasonably well, suggesting that all the elements have originated from a single population. This, the declining enhancement, and the lower temperature suggest that the population is ambient coronal ions, un-enhanced by any impulsive pre-accelerated population. Similar plots can be observed in some events from three spacecraft (e.g., *Wind* and STEREO A and B) spaced at ~120^0 around the Sun (e.g., see Figure 7 in the study by Reames, (2022b)). The main reason for the systematic decline with A/Q may be that higher-rigidity elements leak away faster. The energy spectra of ions are correlated with variations in A/Q in these events (Reames, 2021d; Reames, 2022a) and are also relatively steep.

Typically, enhancements decrease with A/Q in many gradual events, as shown in Figure 4, but there is also a class of events where A/Q dependence is flat, i.e., the abundances are nearly coronal. However, they cannot be used to determine a temperature well since the enhancements are independent of A/Q. Thus, these events are unremarkable and often overlooked. However, one is shown in Figure 5 as an example of an event where abundance enhancements (Figure 5C) begin as quite flat, i.e., independent of A/Q, but then steepen as the higher-rigidity ions preferentially leak away. Here, source temperatures are poorly determined when enhancements are flat but improve as they steepen.

4.2 GLEs that are SEP4 events

As intensities in gradual SEP events increase, the high-energy protons stream ahead to amplify Alfvén waves that scatter subsequent ions (Stix, 1992; Ng et al., 1999; Ng et al. 2003; Ng et al. 2012). Scattering varies as a power law in magnetic rigidity so that Fe can propagate away from the shock more easily than O; for example, Figure 6 shows abundance fits for two GLEs that are SEP4 events. The fitted temperatures are near ~1 MK, and the fits for elements with $Z \geq 6$ are in reasonable agreement with H and He for both positive and negative power-law slopes. Heavy elements tend to be enhanced ahead of the shock and, hence, depressed behind. These enhancements primarily originate from strengthened preferential scattering during transport in these events and *not* from impulsive

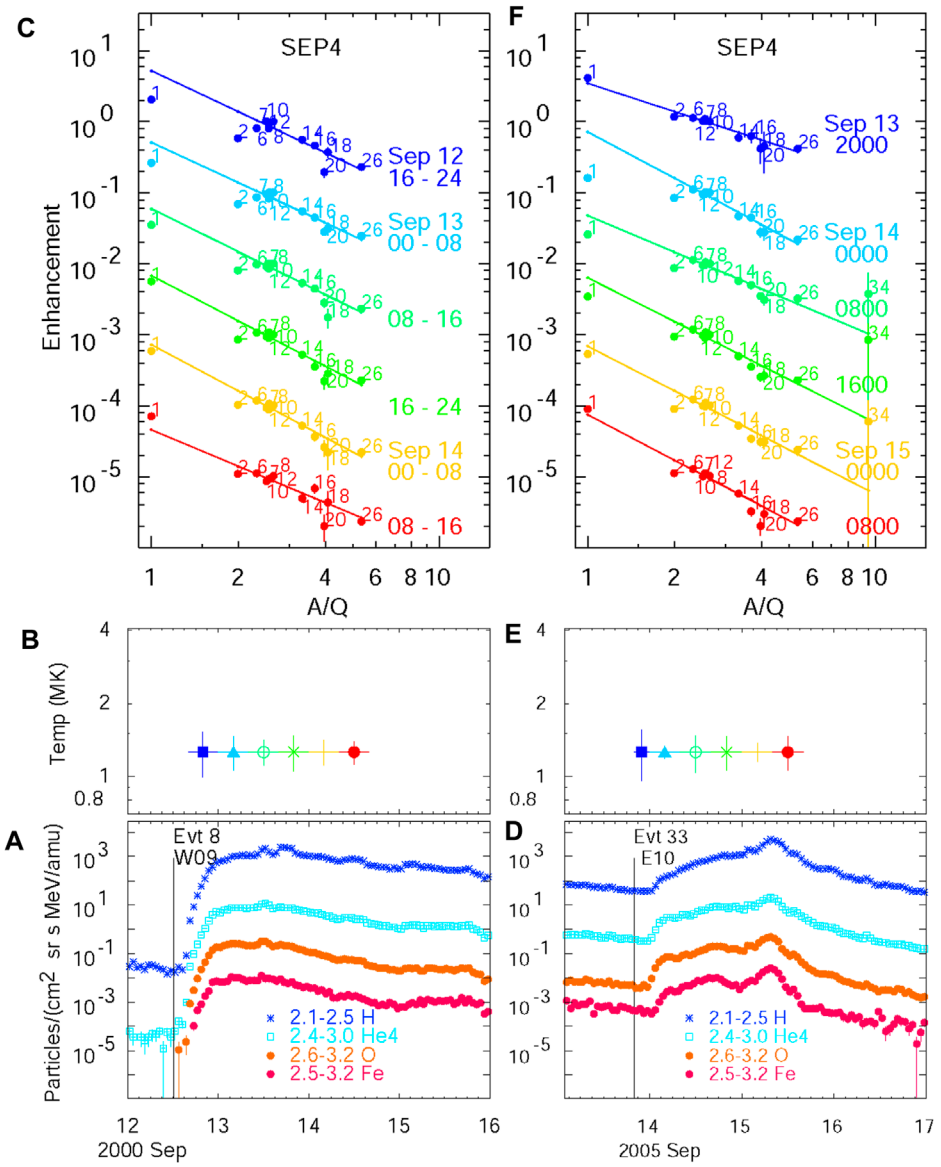

FIGURE 4
Typical small gradual SEP4 events. Panels **(A and D)** show the time evolution of the elements H, ^{4}He, O, and Fe and the listed energies (in MeV amu^{-1}) in two gradual SEP events numbered 8 and 33 (in reference to the list given by Reames, (2016)). Panels **(B** and **E)** show the derived source plasma temperatures in a series of intervals **(colors)** during each event. Panels **(C** and **F)** show abundance enhancements vs. *A/Q* for elements (normalized at O and shifted ×0.1 for each interval), with atomic numbers *Z* shown for each time interval **(color)** and the best fits of elements with $Z \geq 6$ extended to $A/Q = 1$. All the elements, including H, tend to fit each power law.

seed particles, as indicated by the lower temperature, the inclusion of protons in a single seed population, and the return to heavy-element suppression behind the shock. Where abundances show an increase in *A/Q*, the energy spectra become quite flat in the plateau region (Reames and Ng, 2010; Ng et al., 2012; Reames and Ng, 2014; Reames, 2021a) because of the underlying correlation between abundances and spectra (Reames, 2021d; Reames, 2022a).

For these large events, transport out to 1 AU becomes much more important. The strong transport-induced abundance increases shown in Figure 6 are a consequence of the high SEP intensities. The high intensities of protons streaming away from the shock amplify resonant Alfvén waves (Melrose, 1980; Stix, 1992) that scatter subsequent ions. The wave number of resonant waves is $k \approx B/\mu P$, where B is the magnetic field intensity, P is the particle rigidity or momentum per unit charge, and μ is the cosine of the particle pitch angle relative to B. The spectrum of waves not only traps particles near the shock but extends far out into space, bounding intensities at the "streaming limit" (Reames and Ng, 1998; Reames and Ng, 2010; Reames and Ng, 2014), thus driving more acceleration (Lee, 1983; Lee, 2005; Zank et al., 2000; Ng and Reames, 2008; Afanasiev et al., 2015; Afanasiev et al., 2023), and the transport strongly favors the escape of Fe vs. O (Parker, 1965; Ng et al., 1999; Tylka et al., 2001; Ng et al., 2003; Tylka et al., 2005; Tylka and Lee, 2006; Ng et al., 2012). In very large events, intensities

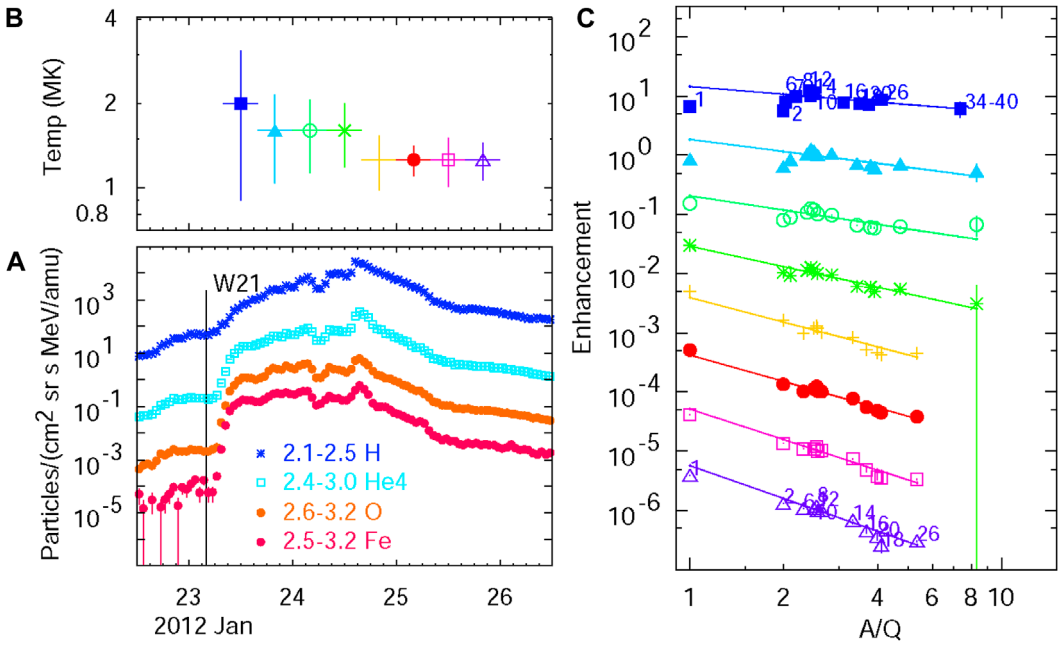

FIGURE 5
Typical gradual SEP4 event with initially flat (coronal) abundances. Panel **(A)** shows the time evolution of elements H, ^{4}He, O, and Fe and the listed energies (in MeV amu^{-1}) in the gradual SEPs. Panel **(B)** shows the derived source plasma temperatures in a series of intervals **(colors** and **symbols)** during the event. Panel **(C)** shows abundance enhancements vs. *A/Q* for elements in each time interval **(color** and **symbols)** ×0.1 and the best fits of elements with $Z \geq 6$ extended to $A/Q = 1$. Element atomic numbers *Z* are listed for the first and last intervals.

of ions below ~1 MeV amu^{-1} (and their abundances) remain in the pre-event background until the shock comes very near to the observer.

4.3 GLEs that are SEP3 events

However, large gradual SEP events, even GLEs, can pick up pre-accelerated residual impulsive seed particles from multi-jet collections, often observed to accumulate near active regions (Desai et al., 2003; Wiedenbeck et al., 2008; Bučík et al., 2014; Bučík et al., 2015; Chen et al., 2015; Reames, 2022a). Recently, Kouloumvakos et al. (2023) found an average connection time to ^{3}He-rich active regions of 4.1 ± 1.8 days, suggesting a width of ~52^0 in longitude. These seed particles contribute with their characteristic enhancement pattern at high *Z* and its source temperature, but ambient coronal ions still dominate H and, possibly, He. Figure 7 shows three events sequentially, from a single region rotating across the Sun, which show the characteristic behavior of SEP2 and then SEP3 events: temperature >2 MK (like impulsive events) and the enhanced high-*Z* fit line that fails to include the proton intensities. Figures 7D, E show He being enhanced as well, and the high-*Z* enhancements flatten with time, probably from preferential leakage at high *Z*, making temperature measurement difficult. The source rotates from W71 to W84 to W120 at the rate of ~13^0 day^{-1}. CME speeds for the three events are 830, 1,199, and 2,465 km s^{-1}, and the last two events are both GLEs. These "double-dipping" SEP3 events are not uncommon (Reames, 2022a) and often include GLEs. Of course, the GLE is determined by the protons, not the high-*Z* ions, but the location or configuration of these events could be a factor.

While it is tempting to think that the single SEP2 event "feeds" the SEP3 events shown in Figure 7, details of the abundances differ. The SEP2 event has an unusually high Ne amount, as shown in Figure 7C, but the later SEP3 events shown in Figures 7D, E do not. Presumably, the seed population for the SEP3 events is fed by many subsequent smaller impulsive events on 15 April that do not contribute energetic ions at 1 AU. SEP3 abundances always show smaller variations than SEP2 events (e.g., Figure 8 in the study by Reames (2021b)). Otherwise, searches for abundance features, like the Ne enhancement here, sometimes implicate spectral fluctuations (Reames, 2019). However, below 1 MeV amu^{-1}, Mason et al. (2016) found extreme spikes in S in 16 events in 16 years. S may be a second-harmonic resonance at $A/Q = 3$ related to the ^{3}He resonance at $A/Q = 1.5$ since S and ^{3}He have similar spectra. In contrast, S ($Z = 16$) is actually suppressed in Figure 3C.

It is common that the high-*Z* enhancement in SEP3 events is less than that in any preceding SEP2 events that may feed the impulsive pool, as shown in Figure 7. Shock acceleration probably reduces the enhancement from seed impulsive ions just as modest SEP4 events depress the abundances of the ambient coronal seed ions, as shown in Figure 4.

Figure 8 directly compares a SEP3 event with two SEP4 events. While the temperatures of the seed populations of the two event types differ, the most notable difference occurs in the proton enhancement. The protons fit the power-law extrapolation from

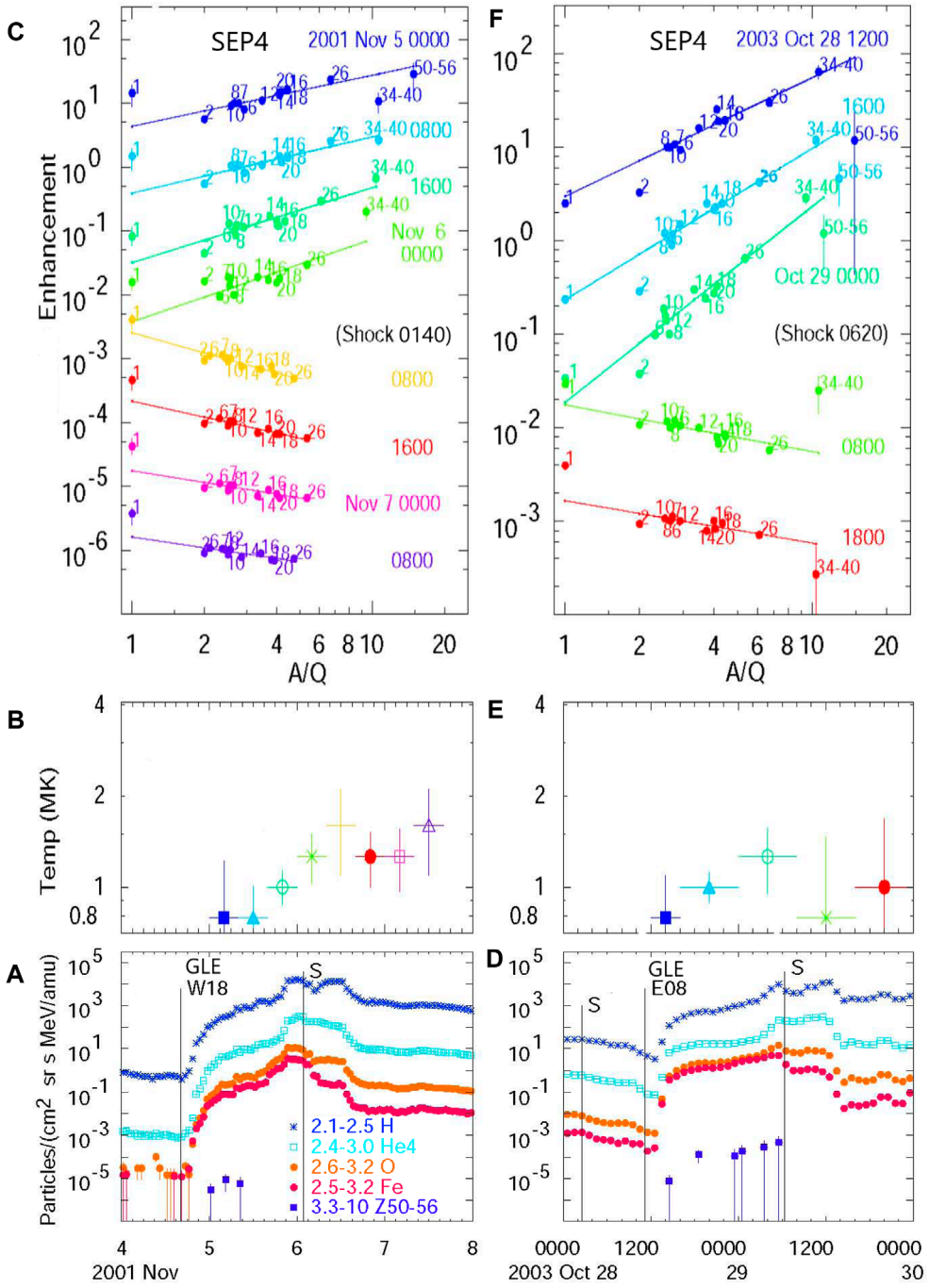

FIGURE 6
Large ground-level enhancements (GLEs) as gradual SEP4 events. Panels **(A** and **D)** show the time evolution of the elements H, ^{4}He, O, and Fe, and elements with 50 $\leq Z \leq$56, along with the listed energies (in MeV amu^{-1}) in 2 SEP4 events. Panels **(B** and **E)** show the derived source plasma temperatures in a series of intervals **(colors)** during each event. Panels **(C** and **F)** show abundance enhancements vs. *A/Q* for elements (normalized at O and shifted ×0.1 for each interval) with *Z* shown, for each time interval **(color)** and the best least-squares fits of elements with $Z \geq 6$ extended to *A/Q* = 1. The high-Z enhancements go away after the shocks pass.

high *Z* in Figure 8F, but they are clearly enhanced in Figure 8C, as they were in Figure 7. All three events shown in Figure 8 show enhancements of high-*Z* elements, at least initially, but it is the departure of the protons from the fit that suggests the presence of two seed populations in SEP3 (and SEP2) events.

It is important to realize that we cannot exclude the possibility of some impulsive suprathermal seeds in *any* gradual event, including SEP4 events. Mason et al. (1999) found a modest increase in ^{3}He in a large gradual event. We cannot distinguish the presence of impulsive heavy elements unless they actually dominate the event

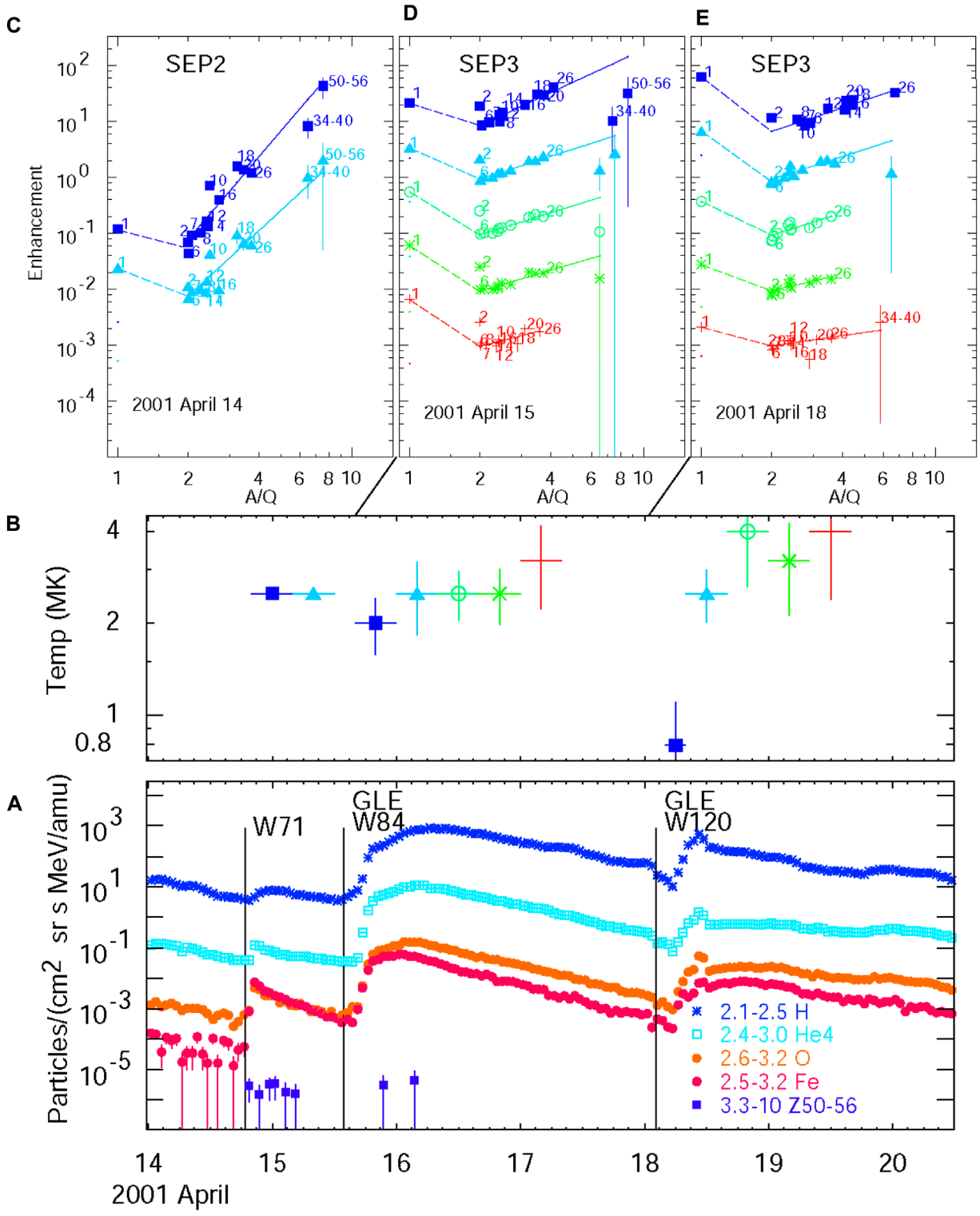

FIGURE 7
Large GLEs as gradual SEP3 events. Panel **(A)** shows the time evolution of elements H, ^{4}He, O, and Fe, and elements with 50 ≤Z ≤56, along with the listed energies (in MeV amu^{-1}) in 2 SEP3 events. Panel **(B)** shows the derived source plasma temperatures in a series of time intervals **(colors)** during the events. Panels **(C, D,** and **E)** show abundance enhancements vs. A/Q for elements (normalized at O and shifted ×0.1 for each interval), with representative atomic numbers Z shown, with the best least-squares fits of elements with Z ≥6, for each time interval **(color)** mapped to the events below. Two different seed populations dominate the high and low Z in SEP2 **(C)** and SEP3 **(D** and **E)** events.

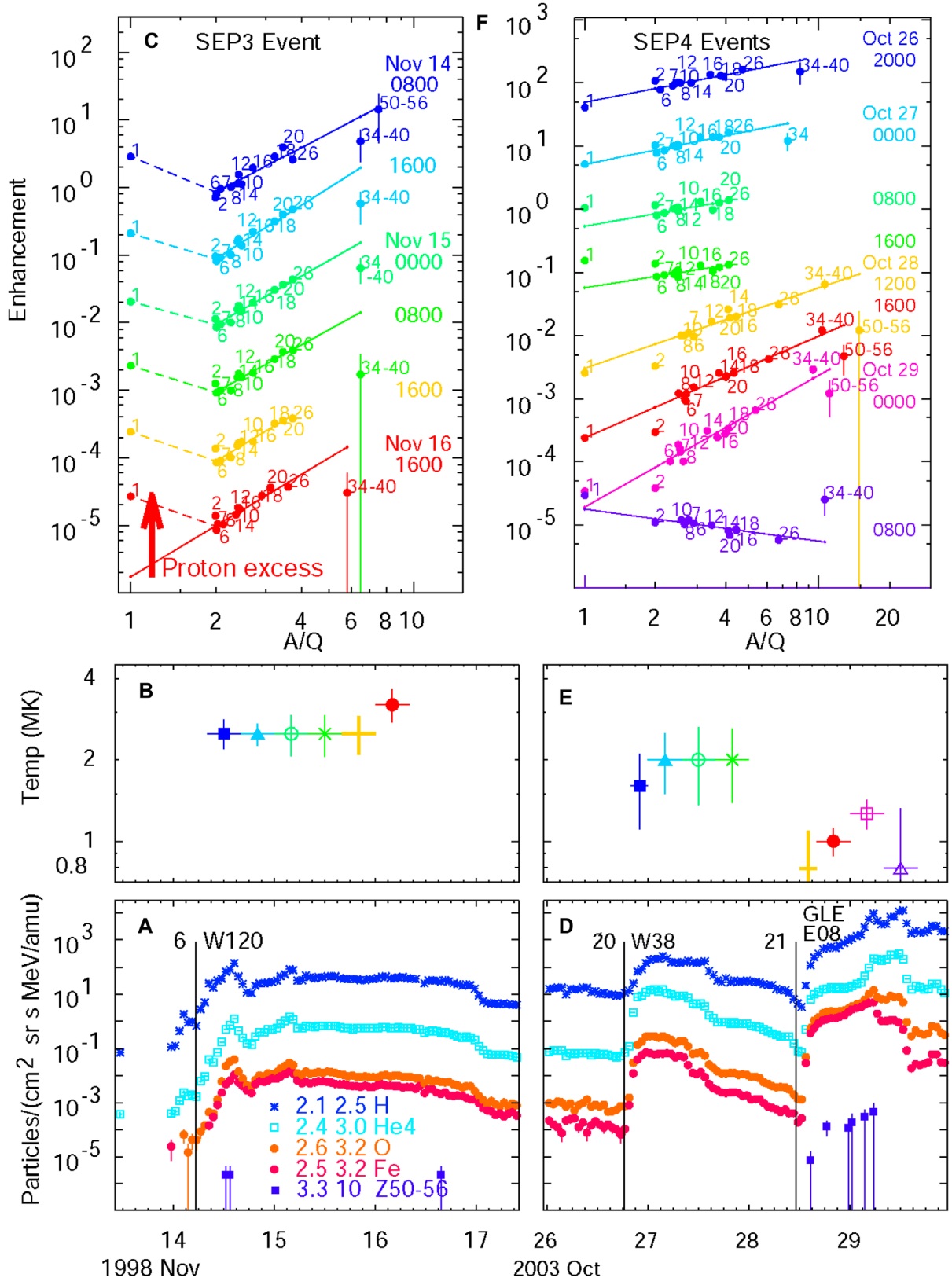

FIGURE 8
Compares a large SEP3 **(left)** event with a pair of large SEP4 events **(right)**; one is a GLE. Panels **(A** and **D)** show the time evolution of the elements H, ^{4}He, O, and Fe, and elements with 50 $\leq Z \leq$56, along with the listed energies (in MeV amu^{-1}). Panels **(B** and **E)** show the derived source plasma temperatures in a series of intervals **(colors)** during each event. Panels **(C** and **F)** show abundance enhancements vs. *A/Q* for elements (normalized at O and shifted ×0.1 for each interval) with *Z* listed for each time interval **(color)**, along with the best least-squares fits of elements with $Z \geq 6$. Systematic proton excess suggests that the SEP3 event in **(C)** has two seed populations for the shock, while for SEP4 events in **(F)**, all elements tend to fit a single power law.

so we see their characteristic A/Q pattern and higher temperature. If only ~10% of Fe was from residual impulsive ions, we would never know it. It is also quite possible that very large events become SEP4 events because strong shocks sweep up enough ambient plasma to swamp any residual impulsive ions that are also available. It is also possible that SEP4 events occur because there are no residual SEP1 ions available for the shock to reaccelerate.

5 Conditions for SEP3 events

GLEs are determined by proton intensities, and protons are accelerated from the ambient plasma in both SEP3 and SEP4 events; so, GLE existence is independent of the dominant source of high-Z ions. In solar cycle 23, 6 of the 15 GLEs were SEP3 events and 9 were SEP4 events. Solar cycle 24 is much weaker with only 2 GLEs, both of which were SEP4 events. During solar cycle 24, STEREO plus Earth provided 3 approximately equally spaced locations around the Sun. Cohen et al. (2017) observed H, He, O, and Fe for gradual SEP events observed by two or three spacecraft. Of 41 events, 10 were measured on all three spacecraft. All of the 10 were SEP4 events, and only 1 of the 2-spacecraft events (4 August 2011) was a SEP3 event (see Figure 10 in the study by Reames (2020b)).

Why are there so few SEP3 events in cycle 24? These events require a stream of residual impulsive SEP1 or SEP2 ions flowing out from an active region. These streams or persistent ^{3}He-rich regions are often observed during the solar maxima (Richardson et al., 1990; Desai et al., 2003; Wiedenbeck et al., 2008; Bučík et al., 2014; Bučík et al., 2015; Chen et al., 2015; Reames, 2022a). Then, that same region emits a fast, wide CME-driven shock that accelerates these residual impulsive ions, along with ambient ions that dominate the protons. In a weak solar cycle, 1) the number of sufficiently fast CMEs is reduced and 2) the probability of ^{3}He-rich streams is reduced, so perhaps the number of SEP3 events is reduced by the product of the two factors. Perhaps a very strong cycle would mostly have SEP3 events.

What are the conditions for producing a SEP3 event? Gopalswamy et al. (2022) asked an important question: "Can type-III radio storms be a source of seed particles to shock acceleration?" These authors identify a large shock event that follows from the same active region as a storm of type-III bursts. Should we expect a SEP3 event? Figure 9 shows the time variation of SEP species during this period. He intensities show ^{3}He/^{4}He of ~1 during the type-III storm, but the intensities are too small to show measurable Fe and O. Intensities of ^{3}He are not reliable during the large event because of the background from ^{4}He, but most likely, ^{3}He/^{4}He <0.1. However, Figure 9B and 9C quite clearly show that the large event is a SEP4 event and definitely not a SEP3 event. Source temperatures are quite low, the protons fit the same population as the high-Z ions, and the power-law fits actually tend to decrease with A/Q. Impulsive suprathermal seed ions do not dominate this event.

Unfortunately, the intensities shown in Figure 9 are insufficient for a firm conclusion since the type-III storm is too weak to show high-Z ions, but the large events are clearly not dominated by impulsive seeds, so it is unlikely that ions from the storm have contributed enough seed particles. There are other cases like this where large shock waves do not seem to reaccelerate the available impulsive suprathermal ions; many large SEP4 events are preceded by several type-III bursts. Are the intensities of impulsive seeds just too low? Yet, there are also many cases, like that shown in Figure 7, where consecutive large shocks dip into a single persistent impulsive population (Reames, 2022a). Worse, we have no cases where a single large event is SEP3 at one longitude and SEP4 at another; such cases might guide us to the location of the impulsive seeds, but there are too few of them in cycle 24. Perhaps, some shocks are driven in the direction of the impulsive seeds, while others are driven away from them, but we do not know for sure. The answer to the question posed by Gopalswamy et al. (2022) is not yet clear; impulsive seed ions are inadequate, or there may be other required conditions. However, the association of ^{3}He with the type-III emission remains strong.

In fact, the measurements of ^{3}He/^{4}He during a type-III storm, as shown in Figure 9A, address another important question: are all the electron events that produce type-III bursts ^{3}He-rich? A value of ^{3}He/^{4}He ≈ 1 persists during the entire type-III storm, suggesting that even these small events are ^{3}He-rich. This may be the first reported measurement of ^{3}He/^{4}He associated with the small events in a type-III storm. At the other extreme, γ-ray lines show that some of the largest flares are ^{3}He-rich (Mandzhavidze et al., 1999; Murphy et al., 2016). Are ^{3}He-rich events a persistent consequence of magnetic reconnection?

6 C/O and the element abundance of the photosphere

After averaging over all the smooth power-law abundance variations, how could the variation of a single abundance C/O stand out so significantly in Figure 1? SEPs have C/O = 0.42 ± 0.01, while the recent photospheric values are 0.550 ± 0.76 (Caffau et al., 2011) and 0.589 ± 0.063 (Asplund et al., 2021). Earlier photospheric measurements by Anders and Grevesse (1989) of C/O = 0.489 were lower and in better agreement. Since the FIP of C is lower than that of O, it seems extremely unlikely that C/O could be suppressed in transit to the corona. The mixing of various seed populations, as described above, can surely cause doubt in the relative abundances of H and ^{4}He, but the power-law behavior of both transport and acceleration most likely preserves the relationship of closely spaced C, N, and O. How could C/O in SEPs possibly be suppressed below that in the photosphere? We previously suggested that the problem might lie with the coronal abundance of C (Reames, 2021b).

It is well-known that the increasing photospheric abundances of "heavy elements," like C, N, and O, have also caused problems for stellar models and helioseismology (e.g., Basu and Antica, 2008). These abundances determine the opacity of the stellar material, and the new (since 1990) lower abundances disagree with the helioseismic constraints.

To correct the C/O problem shown in Figure 1, rather than decrease photospheric C, as previously considered, suppose we increase photospheric O instead. While this is less convenient, since O is our reference, reducing the photospheric C/O to 0.42 amounts to a reduction in SEPs of O by 31%, as shown in Figure 1. At this point, C, O, Ne, and Ar all line up. That is, their SEPs and

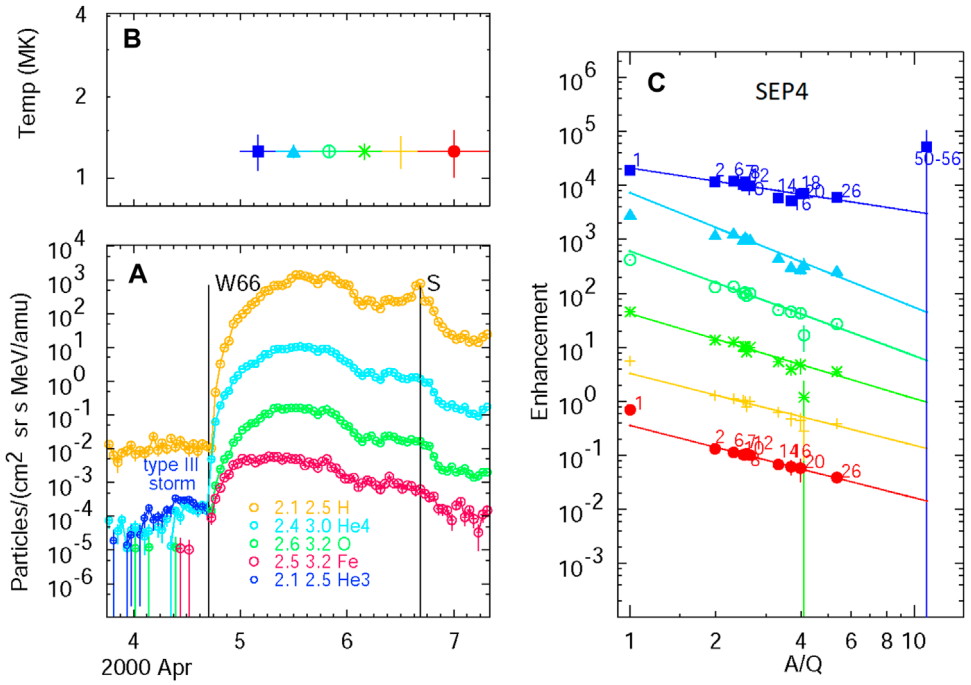

FIGURE 9
Panel **(A)** shows the intensities of the listed SEP ions that are shown during a type-III radio storm and the following large SEP4 event from the same active region studied by Gopalswamy et al. (2022). **(B)** shows the derived temperatures for the selected intervals. Panel **(C)** shows abundance enhancements vs. *A/Q* for elements with *Z* listed for each time interval **(color and symbol)**, along with best least-squares fits of elements with $Z \geq 6$ extended down to protons. Ions during the type-III storm are ^{3}He-rich, but Fe/O is not measurable. ^{3}He cannot be reliably measured during a large event because of the high-^{4}He background. The large event is SEP4 and shows no evidence of impulsive seed ions from the type-III storm.

TABLE 2 SEP-based high-first ionization potential photosphere.

	Z	First ionization potential (FIP) [eV]	SEPs	SEP photosphere [dex]
C	6	11.3	420 ± 10	8.50
N	7	14.5	128 ± 8	7.98
O	8	13.6	1,000 ± 10	8.88
Ne	10	21.6	157 ± 10	8.07
Ar	18	15.8	4.3 ± 0.4	6.51

photospheric abundances are the same. If we assume that the SEP-derived photosphere has the observed SEP abundance ratios of all high-FIP elements (other than H and He), normalized to the abundance of C observed by Caffau et al. (2011), we find the values in Table 2. The higher abundance of O returns to the values observed by Anders and Grevesse (1989).

7 Discussion

Our analysis of source temperatures has assumed Maxwellian electron velocity distributions controlling the relationship between plasma temperatures and *Q*-values of ions (e.g., Mazzotta et al., 1998). Recently, Lee et al. (2022) and Lee et al. (2024) tested this assumption using more realistic kappa distributions, which include high-energy tails of the electron distribution, to fit average impulsive event enhancements. These authors find that the derived source temperatures are not significantly affected and that *A/Q* values differ at most by 10%–20% in extreme cases. This is an important confirmation of the temperatures deduced from SEP abundances. These authors also note that the derived source temperatures agree with active-region temperatures, i.e., they are unaffected by electron heating. Thus, the SEPs must leave the acceleration region on a time scale shorter than their ionization time

scales. However, at low densities, these ionization times could be quite long.

Laming and Kuroda (2023) suggested that heavy ions in impulsive SEP events are enhanced as part of the FIP process. Of course, the FIP process occurs in the dense chromosphere, while ion acceleration must occur at a much lower density in jets (Bučík, 2020), decoupling enhancement and acceleration; presumably, the enhancement could affect a region that would later release a jet. However, this would suggest that both SEPs and the CME from a jet would have strong *A/Q*-dependent enhancements, which have never been observed for CMEs. It is also true that the high-FIP elements He, C, N, O, Ne, S, and Ar fit the same pattern of enhancements as the low-FIP elements Mg, Si, Ca, and Fe (Figure 2 in the study by Reames, 2023c; Figure 8 in the study by Reames et al., 2014a). Independent of the impulsive abundance enhancements, we have separate evidence to show that acceleration occurs at magnetic reconnection sites in solar jets. Developing the power-law enhancement during this acceleration (Drake et al., 2009) seems to be most likely since only the SEPs would be affected.

However, concerning other mechanisms, we should again note that there is evidence that second-harmonic resonant processes related to the enhancement of ^{3}He may contribute to low-energy abundance enhancements, especially of S in relatively rare, small impulsive events (Mason et al., 2016). These rare S-rich events have steep spectra that are only observed below ~1 MeV amu^{-1} (Mason et al., 2016), probably where $A/Q \approx 3$ for S, which occurs at ~2 MK. These higher-order resonant enhancements are associated with the ^{3}He resonance at $A/Q \approx 1.5$. Resonant element enhancement at higher temperatures was studied by Roth and Temerin (1997). More recently, at even lower energies, i.e., <300 keV amu^{-1}, Mason et al. (2023) found extreme enhancements in heavy elements, e.g., Fe/O, presumably because O has $A/Q \approx 2$, and these resonant waves have been absorbed by ^{4}He. This is reminiscent of the hot-plasma study of "the He valley" (Steinacker et al., 1997), where the wave absorption band of ^{4}He can shape the abundances of other ions. Yet, all of the resonant modifications of $Z > 6$ abundances seem to be confined to steep spectra below ~1 MeV amu^{-1}, while higher-energy abundances are dominated by power laws in *A/Q*.

There has been growing interest in the importance of streamers in SEP intensities and the production of GLEs. In streamers, higher densities and lower Alfvén speeds produce higher Alfvénic Mach numbers (Liu et al., 2023); regions of higher θ_{Bn} (e.g., Kong et al., 2017; Kong et al., 2019) may also be a factor in shock acceleration. We have yet to explore any relationship between streamers and the streams of impulsive suprathermal ions above active regions that distinguish between SEP3 and SEP4 events.

8 Summary

SEP element abundances relative to O are compared at the same velocity (i.e., MeV amu^{-1}), so that any dependence upon magnetic rigidity (often a power-law) appears as a dependence upon *A/Q*, assuming that a power law allows the best-fit determination of *Q*-values and, hence, the source temperature.

Impulsive events are accelerated in magnetic reconnection regions in solar jets and escape on open-field lines from sources at ~2.5 MK. They produce ^{3}He-rich events, heavy-ion enhancements, and electron beams that drive type-III radio bursts. The smaller SEP1 events have no fast shocks available for additional acceleration. Steep power-law abundance enhancements vs. *A/Q* (e.g., Figure 3C) include *all* elements, including H (events 3 and 4) or, in some cases, begin above ^{4}He (event 5).

Gamma-ray line measurements show us that solar flares involve the same ^{3}He-rich acceleration mechanisms as solar jets and show significant nuclear fragmentation, but the high (>10 MK) temperatures and nuclear fragments of bright, hot flares are not found in SEPs, indicating that the energetic ions in flares are efficiently trapped magnetically and are not observed to "leak out" into space. Thus, flares show no contribution to either impulsive or gradual SEPs.

In large gradual SEP4 events, fast, wide shock waves, driven by CMEs, are completely dominated by ions accelerated from ambient coronal seed material at 0.8–1.8 MK, beginning at 2–3 R_S. All elements, including H and ^{4}He, again tend to fit on a single power law vs. *A/Q*. Slopes of power-law abundance enhancements or depressions vs. *A/Q* mainly depend upon scattering during transport. 1) Intensities in smaller SEP4 events produce minimal amplification of Alfvén waves, allowing high-rigidity ions to leak away preferentially so abundances decrease vs. *A/Q* (Figures 4C, F). 2) Intensities in larger SEP4 events produce some amplification of Alfvén waves, producing balanced trapping and leakage over the *A/Q* range; so, abundances are initially quite flat vs. *A/Q* (Figure 5C). 3) Intensities of protons streaming out from very large SEP4 events (e.g., GLEs) produce significant amplification of Alfvén waves that scatter the subsequent ions during transport out from the shock but allow high-rigidity ions to preferentially reach the observer ahead the shock, so abundances increase vs. *A/Q* (Figures 6C, F); depleted high-rigidity ions are seen as a depression vs. *A/Q* downstream of the shock.

SEP events can include both shock-accelerated coronal seed ions and shock-reaccelerated impulsive residual seed ions. Impulsive SEP2 events are intended to involve both seeds and shock from a single jet, while gradual SEP3 events describe a large shock traversing an active region with a collection of impulsive ions from multiple previous jets. Both classes are dominated by H (and possibly ^{4}He) from the coronal seeds and by $Z \geq 6$ ions from the previously enhanced impulsive seeds. Figure 7C shows the enhancement pattern of an impulsive SEP2 event, Figures 7D, E and Figure 8C show clear SEP3 events, and Figure 3F shows an ambiguous single-jet event (SEP2) that occurs in a pre-existing impulsively seeded region (SEP3). SEP3 events tend to have smaller abundance fluctuations since they average over many individual jets.

SEP3 events are very rare in solar cycle 24, which is probably a combined effect of both fewer impulsive seeds and fewer large shocks to encounter them. Furthermore, while type-III storms clearly produce ^{3}He-rich events, the subsequent strong shocks from that region do not necessarily find dominant high-*Z* impulsive seed particles.

The SEP value of C/O = 0.42 conflicts with much higher photospheric values of up to 0.59 and lacks explanation. Could the

photospheric value of O actually be 30%–40% higher, as previously found, and as helioseismology independently suggests?

Author contributions

DR: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing–original draft, and writing–review and editing.

Funding

The author declares that no financial support was received for the research, authorship, and/or publication of this article.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Afanasiev, A., Battarbee, M., and Vainio, R. (2015). Self-consistent Monte Carlo simulations of proton acceleration in coronal shocks: effect of anisotropic pitch-angle scattering of particles. *Astron. Astrophys.* 584, 81. doi:10.1051/0004-6361/201526750

Afanasiev, A., Vainio, R., Trotta, D., Nyberg, S., Talebpour Sheshvan, N., Hietala, H., et al. (2023). Self-consistent modeling of the energetic storm particle event of November 10, 2012. *Astron. Astrophys.* 679 (2023), A111. doi:10.1051/0004-6361/202346220

Anders, E., and Grevesse, N. (1989). Abundances of the elements: meteoritic and solar. *Geochim. Cosmochim. Acta* 53, 197–214. doi:10.1016/0016-7037(89)90286-X

Archontis, V., and Hood, A. W. (2013). A numerical model of standard to blowout jets. *Astrophys. J. Lett.* 769, L21. doi:10.1088/2041-8205/769/2/l21

Asplund, M., Amarsi, A. M., and Grevesse, N. (2021). The chemical make-up of the Sun: a 2020 vision. *Astron. Astrophys.* 653, A141, 01661. arXiv:2105. doi:10.1051/0004-6361/202140445

Basu, S., and Antica, H. M. (2008). Helioseismology and solar abundances. *Phys. Rep.* 457, 217–283. doi:10.1016/j.physrep.2007.12.002

Bertsch, D. L., Fichtel, C. E., and Reames, D. V. (1969). Relative abundance of iron-group nuclei in solar cosmic rays. *Astrophys. J. Lett.* 157, L53. doi:10.1086/180383

Breneman, H. H., and Stone, E. C. (1985). Solar coronal and photospheric abundances from solar energetic particle measurements. *Astrophys. J. Lett.* 299, L57. doi:10.1086/184580

Bučík, R. (2020). ^{3}He-rich solar energetic particles: solar sources. *Space Sci. Rev.* 216, 24. doi:10.1007/s11214-020-00650-5

Bučík, R., Innes, D. E., Chen, N. H., Mason, G. M., Gómez-Herrero, R., and Wiedenbeck, M. E. (2015). Long-lived energetic particle source regions on the Sun. *J. Phys. Conf. Ser.* 642, 012002. doi:10.1088/1742-6596/642/1/012002

Bučík, R., Innes, D. E., Mall, U., Korth, A., Mason, G. M., and Gómez-Herrero, R. (2014). Multi-spacecraft observations of recurrent ^{3}He-rich solar energetic particles. *Astrophys. J.* 786, 71. doi:10.1088/0004-637X/786/1/71

Bučík, R., Innes, D. E., Mason, G. M., Wiedenbeck, M. E., Gómez-Herrero, R., and Nitta, N. V. (2018a). ^{3}He-rich solar energetic particles in helical jets on the Sun. *Astrophys. J.* 852, 76. doi:10.3847/1538-4357/aa9d8f

Bučík, R., Mulay, S. M., Mason, G. M., Nitta, N. V., Desai, M. I., and Dayeh, M. A. (2021). Temperature in solar sources of ^{3}He-rich solar energetic particles and relation to ion abundances. *Astrophys. J.* 908, 243. doi:10.3847/1538-4357/abd62d

Bučík, R., Wiedenbeck, M. E., Mason, G. M., Gómez-Herrero, R., Nitta, N. V., and Wang, L. (2018b). ^{3}He-rich solar energetic particles from sunspot jets. *Astrophys. J. Lett.* 869, L21. doi:10.3847/2041-8213/aaf37f

Caffau, E., Ludwig, H.-G., Steffen, M., Freytag, B., and Bonofacio, P. (2011). Solar chemical abundances determined with a CO5BOLD 3D model atmosphere. *Sol. Phys.* 268, 255–269. doi:10.1007/s11207-010-9541-4

Chen, N. H., Bučík, R., Innes, D. E., and Mason, G. M. (2015). Case studies of multi-day ^{3}He-rich solar energetic particle periods. *Astron. Astrophys.* 580, 16. doi:10.1051/0004-6361/201525618

Cliver, E. W., Kahler, S. W., and Reames, D. V. (2004). Coronal shocks and solar energetic proton events. *Astrophys. J.* 605, 902–910. doi:10.1086/382651

Cohen, C. M. S., Mason, G. M., and Mewaldt, R. A. (2017). Characteristics of solar energetic ions as a function of longitude. *Astrophys. J.* 843, 132. doi:10.3847/1538-4357/aa7513

Cook, W. R., Stone, E. C., and Vogt, R. E. (1984). Elemental composition of solar energetic particles. *Astrophys. J.* 279, 827. doi:10.1086/161953

Desai, M. I., and Giacalone, J. (2016). Large gradual solar energetic particle events. *Living Rev. Sol. Phys.* 13, 3. doi:10.1007/s41116-016-0002-5

Desai, M. I., Mason, G. M., Dwyer, J. R., Mazur, J. E., Gold, R. E., Krimigis, S. M., et al. (2003). Evidence for a suprathermal seed population of heavy ions accelerated by interplanetary shocks near 1 AU. *Astrophys. J.* 588, 1149–1162. doi:10.1086/374310

DiFabio, R., Guo, Z., Möbius, E., Klecker, B., Kucharek, H., Mason, G. M., et al. (2008). Energy-dependent charge states and their connection with ion abundances in impulsive solar energetic particle events. *Astrophys. J.* 687, 623–634. doi:10.1086/591833

Drake, J. F., Cassak, P. A., Shay, M. A., Swisdak, M., and Quataert, E. (2009). A magnetic reconnection mechanism for ion acceleration and abundance enhancements in impulsive flares. *Astrophys. J. Lett.* 700, L16–L20. doi:10.1088/0004-637X/700/1/L16

Fichtel, C. E., and Guss, D. E. (1961). Heavy nuclei in solar cosmic rays. *Phys. Rev. Lett.* 6, 495–497. doi:10.1103/PhysRevLett.6.495

Fisk, L. A. (1978). He-3-rich flares - a possible explanation. *Astrophys. J.* 224, 1048. doi:10.1086/156456

Forbush, S. E. (1946). Three unusual cosmic ray increases possibly due to charged particles from the Sun. *Phys. Rev.* 70, 771–772. doi:10.1103/physrev.70.771

Gopalswamy, N., Akiyama, S., Mäkelä, P., Yashiro, S., and Hong Xie, H. (2022). *Can type III radio storms be a source of seed particles to shock acceleration? 3rd URSI AT-AP-RASC*. Gran Canaria. arXiv:2205.15852.

Gopalswamy, N., Xie, H., Akiyama, S., Yashiro, S., Usoskin, I. G., and Davila, J. M. (2013a). The first ground level enhancement event of solar cycle 24: direct observation of shock formation and particle release heights. *Astrophys. J. Lett.* 765, L30. doi:10.1088/2041-8205/765/2/L30

Gopalswamy, N., Xie, H., Mäkelä, P., Yashiro, S., Akiyama, S., Uddin, W., et al. (2013b). Height of shock formation in the solar corona inferred from observations of type II radio bursts and coronal mass ejections. *Adv. Space Res.* 51, 1981–1989. doi:10.1016/j.asr.2013.01.006

Gopalswamy, N., Xie, H., Yashiro, S., Akiyama, S., Mäkelä, P., and Usoskin, I. G. (2012). Properties of Ground level enhancement events and the associated solar eruptions during solar cycle 23. *Space Sci. Rev.* 171, 23–60. doi:10.1007/s11214-012-9890-4

Gosling, J. T. (1993). The solar flare myth. *J. Geophys. Res.* 98, 18937–18949. doi:10.1029/93JA01896

Gosling, J. T. (1994). Corrections to "The solar flare myth.". *J. Geophys. Res.* 99, 4259. doi:10.1029/94JA00015

Kahler, S. W., Reames, D. V., and Sheeley, N. R., Jr. (2001). Coronal mass ejections associated with impulsive solar energetic particle events. *Astrophys. J.* 562, 558–565. doi:10.1086/323847

Kahler, S. W., Reames, D. V., Sheeley, N. R., Jr., Howard, R. A., Kooman, M. J., and Michels, D. J. (1985). A comparison of solar ^{3}He-rich events with type II bursts and coronal mass ejections. *Astrophys. J.* 290, 742. doi:10.1086/163032

Kahler, S. W., Sheeley, N. R., Jr., Howard, R. A., Koomen, M. J., Michels, D. J., McGuire, R. E., et al. (1984). Associations between coronal mass ejections and solar energetic proton events. *J. Geophys. Res.* 89, 9683–9693. doi:10.1029/JA089iA11p09683

Kahler, S. W., Tylka, A. J., and Reames, D. V. (2009). A comparison of elemental abundance ratios in SEP events in fast and slow solar wind regions. *Astrophys. J.* 701, 561–570. doi:10.1088/0004-637X/701/1/561

Kong, X., Guo, F., Chen, Y., and Giacalone, J. (2019). The acceleration of energetic particles at coronal shocks and emergence of a double power-law feature in particle energy spectra. *Astrophys. J.* 883, 49. doi:10.3847/1538-4357/ab3848

Kong, X., Guo, F., Giacalone, J., Li, H., and Chen, Y. (2017). The acceleration of high-energy protons at coronal shocks: the effect of large-scale streamer-like magnetic field structures. *Astrophys. J.* 851, 38. doi:10.3847/1538-4357/aa97d7

Kouloumvakos, A., Mason, G. M., Ho, G. C., Allen, R. C., Wimmer-Schweingruber, R. F., Rouillard, A. P., et al. (2023). Extended ^{3}He-rich time periods observed by solar orbiter: magnetic connectivity and sources. *Astrophys. J.* 956, 123. doi:10.3847/1538-4357/acf44e

Kouloumvakos, A., Rouillard, A. P., Wu, Y., Vainio, R., Vourlidas, A., Plotnikov, I., et al. (2019). Connecting the properties of coronal shock waves with those of solar energetic particles. *Astrophys. J.* 876, 80. doi:10.3847/1538-4357/ab15d7

Kozlovsky, B., Murphy, R. J., and Ramaty, R. (2002). Nuclear deexcitation gamma-ray lines from accelerated particle interactions. *Astrophys. J. Suppl.* 141, 523–541. doi:10.1086/340545

Laming, J. M. (2015). The FIP and inverse FIP effects in solar and stellar coronae. *Living Rev. Sol. Phys.* 12, 2. doi:10.1007/lrsp-2015-2

Laming, J. M., and Kuroda, N. (2023). Element abundances in impulsive solar energetic particle events. *Astrophys. J.* 951, 86. doi:10.3847/1538-4357/acd69a

Laming, J. M., Vourlidas, A., Korendyke, C., Chua, D., Cranmer, S. R., Ko, Y. K., et al. (2019). Element abundances: a new diagnostic for the solar wind. *Astrophys. J.* 879, 124. doi:10.3847/1538-4357/ab23f1

Lee, E. J., Archontis, V., and Hood, A. W. (2015). Plasma jets and eruptions in solar coronal holes: a three-dimensional flux emergence experiment. *Astrophys. J. Lett.* 798, L10. doi:10.1088/2041-8205/798/1/L10

Lee, J. Y., Kahler, S., Ko, Y. K., and Raymond, J. C. (2022). "A study of mass-to-charge ratio with various kappa values in impulsive SEP events," in AGU Fall Meeting, Chicago, IL, Bibcode. 2022AGUFMSH42A.04L.

Lee, J. Y., Kahler, S., Raymond, J. C., and Ko, Y. K. (2024). Solar energetic particle charge states and abundances with nonthermal electrons. *Astrophys. J.* in press, arXiv: 2024arXiv240101604L.

Lee, M. A. (1983). Coupled hydromagnetic wave excitation and ion acceleration at interplanetary traveling shocks. *J. Geophys. Res.* 88, 6109–6119. doi:10.1029/JA088iA08p06109

Lee, M. A. (2005). Coupled hydromagnetic wave excitation and ion acceleration at an evolving coronal/interplanetary shock. *Astrophys. J. Suppl.* 158, 38–67. doi:10.1086/428753

Lee, M. A., Mewaldt, R. A., and Giacalone, J. (2012). Shock acceleration of ions in the heliosphere. *Space Sci. Rev.* 173, 247–281. doi:10.1007/s11214-012-9932-y

Lin, R. P. (1970). The emission and propagation of 40 keV solar flare electrons. I: the relationship of 40 keV electron to energetic proton and relativistic electron emission by the sun. *Sol. Phys.* 12, 266. doi:10.1007/BF00227122

Liu, W., Kong, X., Guo, F., Zhao, L., Feng, S., Yu, F., et al. (2023). Effects of coronal magnetic field configuration on particle acceleration and release during the ground level enhancement events in solar cycle 24. *Astrophys. J.* 954, 203. doi:10.3847/1538-4357/ace9d2

Lodders, K., Palme, H., and Gail, H.-P. (2009). "Abundances of the elements in the solar system," in *Landolt-börnstein, new series VI/4B*. Editor J. E. Trümper (Berlin: Springer), 560. Chap. 4.4. doi:10.1007/978-3-540-88055-4_34

Luhn, A., Klecker, B., Hovestadt, D., Gloeckler, G., Ipavich, F. M., Scholer, M., et al. (1984). Ionic charge states of N, Ne, Mg, Si and S in solar energetic particle events. *Adv. Space Res.* 4, 161. doi:10.1016/0273-1177(84)90307-7

Luhn, A., Klecker, B., Hovestadt, D., and Möbius, E. (1987). The mean ionic charge of silicon in He-3-rich solar flares. *Astrophys. J.* 317, 951. doi:10.1086/165343

Mandzhavidze, N., Ramaty, R., and Kozlovsky, B. (1999). Determination of the abundances of subcoronal ^{4}He and of solar flare-accelerated ^{3}He and ^{4}He from gamma-ray spectroscopy. *Astrophys. J.* 518, 918–925. doi:10.1086/307321

Mason, G. M. (2007). ^{3}He-rich solar energetic particle events. *Space Sci. Rev.* 130, 231–242. doi:10.1007/s11214-007-9156-8

Mason, G. M., Gloeckler, G., and Hovestadt, D. (1984). Temporal variations of nucleonic abundances in solar flare energetic particle events. II - evidence for large-scale shock acceleration. *Astrophys. J.* 280, 902. doi:10.1086/162066

Mason, G. M., Mazur, J. E., and Dwyer, J. R. (1999). [TSUP]3[/TSUP]H[CLC]e[/CLC] Enhancements in Large Solar Energetic Particle Events. *Astrophys. J. Lett.* 525, L133–L136. doi:10.1086/312349

Mason, G. M., Mazur, J. E., Dwyer, J. R., Jokippi, J. R., Gold, R. E., and Krimigis, S. M. (2004). Abundances of heavy and ultraheavy ions in ^{3}He-rich solar flares. *Astrophys. J.* 606, 555–564. doi:10.1086/382864

Mason, G. M., Nitta, N. V., Wiedenbeck, M. E., and Innes, D. E. (2016). Evidence for a common acceleration mechanism for enrichments of 3He and heavy ions in impulsive SEP events. *Astrophys. J.* 823, 138. doi:10.3847/0004-637X/823/2/138

Mason, G. M., Reames, D. V., Klecker, B., Hovestadt, D., and von Rosenvinge, T. T. (1986). The heavy-ion compositional signature in He-3-rich solar particle events. *Astrophys. J.* 303, 849. doi:10.1086/164133

Mason, G. M., Roth, I., Nitta, N. V., Bučík, R., Lario, D., Ho, G. C., et al. (2023). Heavy-ion acceleration in ^{3}He-rich solar energetic particle events: new insights from Solar Orbiter. *Astrophys. J.* 957, 112. doi:10.3847/1538-4357/acf31b

Mazzotta, P., Mazzitelli, G., Colafrancesco, S., and Vittorio, N. (1998). Ionization balance for optically thin plasmas: rate coefficients for all atoms and ions of the elements H to Ni. *Astron. Astrophys. Suppl.* 133, 403–409. doi:10.1051/aas:1998330

McGuire, R. E., von Rosenvinge, T. T., and McDonald, F. B. (1979). "A survey of solar cosmic ray composition," in Proc. 16th Int. Cosmic Ray Conf., Tokyo, 61.

Melrose, D. B. (1980). *Plasma astrophysics*, 1. New York: Gordon & Breach).

Mewaldt, R. A., Cohen, C. M. S., Leske, R. A., Christian, E. R., Cummings, A. C., Stone, E. C., et al. (2002). Fractionation of solar energetic particles and solar wind according to first ionization potential. *Advan. Space Res.* 30, 79–84. doi:10.1016/S0273-1177(02)00263-6

Mewaldt, R. A., Looper, M. D., Cohen, C. M. S., Haggerty, D. K., Labrador, A. W., Leske, R. A., et al. (2012). Energy spectra, composition, and other properties of ground-level events during solar cycle 23. *Space Sci. Rev.* 171, 97–120. doi:10.1007/s11214-012-9884-2

Meyer, J. P. (1985). The baseline composition of solar energetic particles. *Astrophys. J. Suppl.* 57, 151. doi:10.1086/191000

Mogro-Campero, A., and Simpson, J. A. (1972). Enrichment of very heavy nuclei in the composition of solar accelerated particles. *Astrophys. J. Lett.* 171, L5. doi:10.1086/180856

Murphy, R. J., Kozlovsky, B., and Share, G. H. (2016). Evidence for enhanced ^{3}He in flare-accelerated particles based on new calculations of the gamma-ray line spectrum. *Astrophys. J.* 833, 196. doi:10.3847/1538-4357/833/2/196

Murphy, R. J., Ramaty, R., Kozlovsky, B., and Reames, D. V. (1991). Solar abundances from gamma-ray spectroscopy: comparisons with energetic particle, photospheric, and coronal abundances. *Astrophys. J.* 371, 793. doi:10.1086/169944

Ng, C. K., and Reames, D. V. (2008). Shock acceleration of solar energetic protons: the first 10 minutes. *Astrophys. J. Lett.* 686, L123–L126. doi:10.1086/592996

Ng, C. K., Reames, D. V., and Tylka, A. J. (1999). Effect of proton-amplified waves on the evolution of solar energetic particle composition in gradual events. *Geophys. Res. Lett.* 26, 2145–2148. doi:10.1029/1999GL900459

Ng, C. K., Reames, D. V., and Tylka, A. J. (2003). Modeling shock-accelerated solar energetic particles coupled to interplanetary Alfvén waves. *Astrophys. J.* 591, 461–485. doi:10.1086/375293

Ng, C. K., Reames, D. V., and Tylka, A. J. (2012). Solar energetic particles: shock acceleration and transport through self-amplified waves. *AIP Conf. Proc.* 1436, 212. doi:10.1063/1.4723610

Nitta, N. V., Reames, D. V., DeRosa, M. L., Yashiro, S., and Gopalswamy, N. (2006). Solar sources of impulsive solar energetic particle events and their magnetic field connection to the earth. *Astrophys. J.* 650, 438–450. doi:10.1086/507442

Pariat, E., Dalmasse, K., DeVore, C. R., Antiochos, S. K., and Karpen, J. T. (2015). Model for straight and helical solar jets. I. Parametric studies of the magnetic field geometry. *Astron. Astrophys.* 573, A130. doi:10.1051/0004-6361/201424209

Parker, E. N. (1965). The passage of energetic charged particles through interplanetary space. *Planet. Space Sci.* 13, 9–49. doi:10.1016/0032-0633(65)90131-5

Post, D. E., Jensen, R. V., Tarter, C. B., Grasberger, W. H., and Lokke, W. A. (1977). Steady-state radiative cooling rates for low-density, high temperature plasmas. *At. Data Nucl. Data tables.* 20, 397–439. doi:10.1016/0092-640X(77)90026-2

Raukunen, O., Vainio, R., Tylka, A. J., Dietrich, W. F., Jiggens, P., Heynderickx, D., et al. (2018). Two solar proton fluence models based on ground level enhancement observations. *J. Spa. Wea. Spa. Clim.* 8, A04. doi:10.1051/swsc/2017031

Reames, D. V. (1988). Bimodal abundances in the energetic particles of solar and interplanetary origin. *J. Lett.* 330, L71. doi:10.1086/185207

Reames, D. V. (1995a). Coronal Abundances determined from energetic particles. *Adv. Space Res.* 15 (7), 41–51. doi:10.1016/0273-1177(94)00018-v

Reames, D. V. (1995b). Solar energetic particles: a paradigm shift. *Revs. Geophys. Suppl.* 33, 585–589. doi:10.1029/95RG00188

Reames, D. V. (1999). Particle acceleration at the Sun and in the heliosphere. *Space Sci. Rev.* 90, 413–491. doi:10.1023/A:1005105831781

Reames, D. V. (2000). Abundances of trans-iron elements in solar energetic particle events. *Astrophys. J. Lett.* 540, L111–L114. doi:10.1086/312886

Reames, D. V. (2009a). Solar release times of energetic particles in ground-level events. *Astrophys. J.* 693, 812–821. doi:10.1088/0004-637X/693/1/812

Reames, D. V. (2009b). Solar energetic-particle release times in historic ground-level events. *Astrophys. J.* 706, 844–850. doi:10.1088/0004-637X/706/1/844

Reames, D. V. (2013). The two sources of solar energetic particles. *Space Sci. Rev.* 175, 53–92. doi:10.1007/s11214-013-9958-9

Reames, D. V. (2014). Element abundances in solar energetic particles and the solar corona. *Sol. Phys.* 289, 977–993. doi:10.1007/s11207-013-0350-4

Reames, D. V. (2016). Temperature of the source plasma in gradual solar energetic particle events. *Sol. Phys.* 291, 911–930. doi:10.1007/s11207-016-0854-9

Reames, D. V. (2018a). The "FIP effect" and the origins of solar energetic particles and of the solar wind. *Sol. Phys.* 293, 47. doi:10.1007/s11207-018-1267-8

Reames, D. V. (2018b). Abundances, ionization states, temperatures, and FIP in solar energetic particles. *Space Sci. Rev.* 214, 61. doi:10.1007/s11214-018-0495-4

Reames, D. V. (2019). Helium suppression in impulsive solar energetic-particle events. *Sol. Phys.* 294, 32. (arXiv: 1812.01635). doi:10.1007/s11207-019-1422-x

Reames, D. V. (2020a). Four distinct pathways to the element abundances in solar energetic particles. *Space Sci. Rev.* 216, 20. doi:10.1007/s11214-020-0643-5

Reames, D. V. (2020b). Distinguishing the rigidity dependences of acceleration and transport in solar energetic particles. *Sol. Phys.* 295, 113. arXiv 2006.11338. doi:10.1007/s11207-020-01680-6

Reames, D. V. (2021a). "Solar energetic particles," in *Lec. Notes phys. 978*. Second Edition (Cham, Switzerland: Springer Nature). open access. doi:10.1007/978-3-030-66402-2

Reames, D. V. (2021b). Sixty years of element abundance measurements in solar energetic particles. *Space Sci. Rev.* 217, 72. doi:10.1007/s11214-021-00845-4

Reames, D. V. (2021c). Fifty years of ^{3}He-rich events. *Front. Astron. Space Sci.* 8, 164. doi:10.3389/fspas.2021.760261

Reames, D. V. (2021d). On the correlation between energy spectra and element abundances in solar energetic particles. *Sol. Phys.* 296, 24. doi:10.1007/s11207-021-01762-z

Reames, D. V. (2022a). Energy spectra vs. element abundances in solar energetic particles and the roles of magnetic reconnection and shock acceleration. *Sol. Phys.* 297, 32. doi:10.1007/s11207-022-01961-2

Reames, D. V. (2022b). Solar energetic particles: spatial extent and implications of the H and He abundances. *Space Sci. Rev.* 218, 48. doi:10.1007/s11214-022-00917-z

Reames, D. V. (2023a). How do shock waves define the space-time structure of gradual solar energetic-particle events? *Space Sci. Rev.* 219, 14. doi:10.1007/s11214-023-00959-x

Reames, D. V. (2023b). Review and outlook of solar-energetic-particle measurements on multispacecraft missions. *Front. Astron. Space Sci.* 10. doi:10.3389/fspas.2023.1254266

Reames, D. V. (2023c). Element abundances in impulsive solar energetic-particle events. *Universe* 9, 466. doi:10.3390/universe9110466

Reames, D. V., Cliver, E. W., and Kahler, S. W. (2014a). Abundance enhancements in impulsive solar energetic-particle events with associated coronal mass ejections. *Sol. Phys.* 289, 3817–3841. doi:10.1007/s11207-014-0547-1

Reames, D. V., Cliver, E. W., and Kahler, S. W. (2014b). Variations in abundance enhancements in impulsive solar energetic-particle events and related CMEs and flares. *Sol. Phys.* 289, 4675–4689. doi:10.1007/s11207-014-0589-4

Reames, D. V., Meyer, J. P., and von Rosenvinge, T. T. (1994). Energetic-particle abundances in impulsive solar flare events. *Astrophys. J. Suppl.* 90, 649. doi:10.1086/191887

Reames, D. V., and Ng, C. K. (1998). Streaming-limited intensities of solar energetic particles. *Astrophys. J.* 504, 1002–1005. doi:10.1086/306124

Reames, D. V., and Ng, C. K. (2004). Heavy-element abundances in solar energetic particle events. *Astrophys. J.* 610, 510–522. doi:10.1086/421518

Reames, D. V., and Ng, C. K. (2010). Streaming-limited intensities of solar energetic particles on the intensity plateau. *Astrophys. J.* 723, 1286–1293. doi:10.1088/0004-637X/723/2/1286

Reames, D. V., and Ng, C. K. (2014). *The streaming limit of solar energetic-particle intensities, living with a star workshop on extreme space weather events*. Boulder, Co. June 9-11 (2014) arXiv 1412.2279.

Reames, D. V., Richardson, I. G., and Barbier, L. M. (1991). On the differences in element abundances of energetic ions from corotating events and from large solar events. *Astrophys. J. Lett.* 382, L43. doi:10.1086/186209

Reames, D. V., and Stone, R. G. (1986). The identification of solar ^{3}He-rich events and the study of particle acceleration at the sun. *Astrophys. J.* 308, 902. doi:10.1086/164560

Reames, D. V., von Rosenvinge, T. T., and Lin, R. P. (1985). Solar ^{3}He-rich events and nonrelativistic electron events - a new association. *Astrophys. J.* 292, 716. doi:10.1086/163203

Richardson, I. G., Reames, D. V., Wenzel, K. P., and Rodriguez-Pacheco, J. (1990). Quiet-time properties of low-energy (less than 10 MeV per nucleon) interplanetary ions during solar maximum and solar minimum. *Astrophys. J. Lett.* 363, L9. doi:10.1086/185853

Roth, I., and Temerin, M. (1997). Enrichment of ^{3}He and heavy ions in impulsive solar flares. *Astrophys. J.* 477, 940–957. doi:10.1086/303731

Serlemitsos, A. T., and Balasubrahmanyan, V. K. (1975). Solar particle events with anomalously large relative abundance of ^{3}He. *Astrophys. J.* 198, 195. doi:10.1086/153592

Shimojo, M., and Shibata, K. (2000). Physical parameters of solar X-ray jets. *Astrophys. J.* 542, 1100–1108. doi:10.1086/317024

Steinacker, J., Meyer, J.-P., Steinacker, A., and Reames, D. V. (1997). The helium valley: comparison of impulsive solar flare ion abundances and gyroresonant acceleration with oblique turbulence in a hot multi-ion plasma. *Astrophys. J.* 476, 403–427. doi:10.1086/303589

Stix, T. H. (1992). *Waves in plasmas*. New York: AIP.

Tan, L. C., Reames, D. V., Ng, C. K., Shao, X., and Wang, L. (2011). What causes scatter-free transport of non-relativistic solar electrons? *Astrophys. J.* 728, 133. doi:10.1088/0004-637X/728/2/133

Teegarden, B. J., von Rosenvinge, T. T., and McDonald, F. B. (1973). Satellite measurements of the charge composition of solar cosmic rays in the $6 \leq Z \leq 26$ interval. *Astrophys. J.* 180, 571. doi:10.1086/151985

Temerin, M., and Roth, I. (1992). The production of ^{3}He and heavy ion enrichment in ^{3}He-rich flares by electromagnetic hydrogen cyclotron waves. *Astrophys. J. Lett.* 391, L105. doi:10.1086/186408

Tylka, A. J., Cohen, C. M. S., Dietrich, W. F., Krucker, S., McGuire, R. E., Mewaldt, R. A., et al. (2003). in Proc. 28th Int. Cosmic Ray Conf., Tsukuba, 3305.

Tylka, A. J., Cohen, C. M. S., Dietrich, W. F., Lee, M. A., Maclennan, C. G., Mewaldt, R. A., et al. (2005). Shock geometry, seed populations, and the origin of variable elemental composition at high energies in large gradual solar particle events. *Astrophys. J.* 625, 474–495. doi:10.1086/429384

Tylka, A. J., Cohen, C. M. S., Dietrich, W. F., Maclennan, C. G., McGuire, R. E., Ng, C. K., et al. (2001). Evidence for remnant flare suprathermals in the source population of solar energetic particles in the 2000 bastille day event. *Astrophys. J. Lett.* 558, L59–L63. doi:10.1086/323344

Tylka, A. J., and Dietrich, W. F. (2009). "A new and comprehensive analysis of proton spectra in ground-level enhanced (GLE) solar particle events," in Proceedings of 31st International Cosmic Ray Conference Lodz. http://icrc2009.uni.lodz.pl/proc/pdf/icrc0273.pdf.

Tylka, A. J., and Lee, M. A. (2006). A model for spectral and compositional variability at high energies in large, gradual solar particle events. *Astrophys. J.* 646, 1319–1334. doi:10.1086/505106

von Rosenvinge, T. T., Barbier, L. M., Karsch, J., Liberman, R., Madden, M. P., Nolan, T., et al. (1995). The energetic particles: acceleration, composition, and transport (EPACT) investigation on the wind spacecraft. *Space Sci. Rev.* 71, 155–206. doi:10.1007/BF00751329

Wang, Y.-M., Pick, M., and Mason, G. M. (2006). Coronal holes, jets, and the origin of ^{3}He-rich particle events. *Astrophys. J.* 639, 495–509. doi:10.1086/499355

Webber, W. R. (1975). Solar and galactic cosmic ray abundances – a comparison and some comments. Proc. 14th Int. Cos, Ray Conf., Municxh. 5, 1597.

Wiedenbeck, M. E., Cohen, C. M. S., Cummings, A. C., de Nolfo, G. A., Leske, R. A., Mewaldt, R. A., et al. (2008). Persistent energetic ^{3}He in the inner heliosphere. *Proc. 30th Int. Cosm. Ray Conf. (Mérida)* 1, 91.

Wild, J. P., Smerd, S. F., and Weiss, A. A. (1963). Solar bursts. *Annu. Rev. Astron. Astrophys.* 1, 291–366. doi:10.1146/annurev.aa.01.090163.001451

Zank, G. P., Li, G., and Verkhoglyadova, O. (2007). Particle acceleration at interplanetary shocks. *Space Sci. Rev.* 130, 255–272. doi:10.1007/s11214-007-9214-2

Zank, G. P., Rice, W. K. M., and Wu, C. C. (2000). Particle acceleration and coronal mass ejection driven shocks: a theoretical model. *J. Geophys. Res.* 105, 25079–25095. doi:10.1029/1999JA000455

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 09 May 2024
DOI 10.3389/fspas.2024.1385820

Check for updates

OPEN ACCESS

EDITED BY
Gang Li,
University of Alabama in Huntsville,
United States

REVIEWED BY
Patricio A. Muñoz,
Technical University of Berlin, Germany
Christian L. Vásconez,
National Polytechnic School, Ecuador

*CORRESPONDENCE
Andreas Shalchi,
andreasm4@yahoo.com

RECEIVED 13 February 2024
ACCEPTED 05 April 2024
PUBLISHED 09 May 2024

CITATION
Shalchi A (2024), Transport of energetic particles in turbulent space plasmas: pitch-angle scattering, telegraph, and diffusion equations.
Front. Astron. Space Sci. 11:1385820.
doi: 10.3389/fspas.2024.1385820

COPYRIGHT
© 2024 Shalchi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Transport of energetic particles in turbulent space plasmas: pitch-angle scattering, telegraph, and diffusion equations

Andreas Shalchi*

Department of Physics and Astronomy, University of Manitoba, Winnipeg, Canada

Introduction: In this article, we revisit the pitch-angle scattering equation describing the propagation of energetic particles through magnetized plasma. In this case, solar energetic particles and cosmic rays interact with magnetohydrodynamic turbulence and experience stochastic changes in the pitch-angle. Since this happens over an extended period of time, a pitch-angle isotropization process occurs, leading to parallel spatial diffusion. This process is described well by the pitch-angle scattering equation. However, the latter equation is difficult to solve analytically even when considering special cases for the scattering coefficient.

Methods: In the past, a so-called subspace approximation was proposed, which has important applications in the theory of perpendicular diffusion. Alternatively, an approach based on the telegraph equation (also known as telegrapher's equation) has been developed. We show that two-dimensional subspace approximation and the description based on the telegraph equation are equivalent. However, it is also shown that the obtained distribution functions contain artifacts and inaccuracies that cannot be found in the numerical solution to the problem. Therefore, an N-dimensional subspace approximation is proposed corresponding to a semi-analytical/semi-numerical approach. This is a useful alternative compared to standard numerical solvers.

Results and Discussion: Depending on the application, the N-dimensional subspace approximation can be orders of magnitude faster. Furthermore, the method can easily be modified so that it can be used for any pitch-angle scattering equation.

KEYWORDS

cosmic rays, magnetic fields, turbulence, diffusion, transport

1 Introduction

The motion of energetic particles such as cosmic rays through plasma is a complicated stochastic process. It is described via transport equations containing different diffusion parameters. The simplest form of a transport equation which is used in this field is the pitch-angle scattering equation (Shalchi, 2009; Zank, 2014)

$$\frac{\partial f}{\partial t}+v\mu\frac{\partial f}{\partial z}=\frac{\partial}{\partial\mu}\left[D_{\mu\mu}(\mu)\frac{\partial f}{\partial\mu}\right], \tag{1}$$

where we have used time t, particle position along the mean magnetic field z, pitch-angle cosine μ, particle speed v, and pitch-angle scattering coefficient $D_{\mu\mu}$. The analytical

form of the latter parameter is difficult to determine since it contains information about the interaction between magnetohydrodynamic turbulence and energetic and electrically charged particles. Very originally, a quasi-linear approach was developed to determine the coefficient $D_{\mu\mu}$ (Jokipii 1966). However, this approach is inaccurate, and it fails to describe correctly the scattering of particles at 90° corresponding to $\mu = 0$ (Shalchi 2009). Therefore, the so-called second-order quasi-linear theory (SOQLT) was developed by Shalchi (2005), which provides non-vanishing scattering at $\mu = 0$, resolving the 90°-problem. This theory was further explored analytically in Shalchi et al. (2009), and the so-called *isotropic form*

$$D_{\mu\mu} = (1-\mu^2)\, D \tag{2}$$

was derived in the limit of a stronger turbulent magnetic field. In Eq. 2, the parameter D does not depend on μ, but it is a complicated function of turbulence and particle properties (Shalchi et al., 2009).

In addition to the question of what the correct analytical form of $D_{\mu\mu}$ is, one desires to find solutions to Eq. 1. However, so far, no exact solution to the pitch-angle scattering equation has been found, and one has to rely on either a numerical approach or approximations. However, one can show that in the late-time limit, the pitch-angle-averaged distribution function

$$M(z,t) = \frac{1}{2}\int_{-1}^{+1} d\mu\; f(\mu,z,t) \tag{3}$$

satisfies a diffusion or heat transfer equation of the form

$$\frac{\partial M}{\partial t} = \kappa_{\parallel}\frac{\partial^2 M}{\partial t^2}, \tag{4}$$

where the parallel spatial diffusion coefficient is related to the pitch-angle scattering coefficient via (Earl, 1974)

$$\kappa_{\parallel} = \frac{v^2}{8}\int_{-1}^{+1} d\mu\; \frac{(1-\mu^2)^2}{D_{\mu\mu}(\mu)}. \tag{5}$$

The heat transfer equation shown above can easily be solved. For sharp initial conditions, for instance, the solution is simply a normalized Gaussian distribution

$$M(z,t) = \frac{1}{\sqrt{4\pi\kappa_{\parallel} t}}\, e^{-\frac{z^2}{4\kappa_{\parallel} t}} \tag{6}$$

centered at z = 0 and having the second moment $\langle z^2\rangle = 2\kappa_{\parallel} t$. One can also write down the more general solution

$$M(z,t) = \frac{1}{\sqrt{4\pi\kappa_{\parallel} t}}\int_{-\infty}^{+\infty} dz'\; M(z',t=0)\, e^{-\frac{(z-z')^2}{4\kappa_{\parallel} t}} \tag{7}$$

which depends on the initial distribution $M(z', t = 0)$ and has a Gaussian integral kernel.

More recently (Tautz and Lerche, 2016 and references therein), it was argued that the diffusive solution does not always provide a good approximation, and one should instead use a *telegraph equation* of the form

$$\tau\ddot{M} + \dot{M} = \kappa_{\parallel}\frac{\partial^2 M}{\partial z^2}, \tag{8}$$

where we have used the *telegraph time scale* τ. It should be noted that using the telegraph equation instead of the diffusion equation was, in particular, suggested in the context of *adiabatic focusing* (Litvinenko and Schlickeiser, 2013; Effenberger and Litvinenko, 2014), but this effect is omitted in this paper.

Independently, a two-dimensional subspace approximation to the solution of Eq. 1 has been developed (see Shalchi et al. (2011) for the original description of this approach and Shalchi (2020) for a review). Although this approach provides only an approximation to the solution of Eq. 1 for the isotropic case, it provides a pitch-angle-dependent solution. The two-dimensional subspace approximation was successfully applied in the theory of perpendicular transport and contributed significantly to the development of advanced particle transport theories (Shalchi, 2020; Shalchi, 2021).

In this paper, we revisit pitch-angle scattering and parallel spatial diffusion as well as the corresponding transport equations. Through this study, we aim to perform the following tasks:

1. We review the two-dimensional subspace approximation and summarize the corresponding results.
2. We show the equivalence of the two-dimensional subspace approximation and the telegraph equation.
3. We derive an approximation for the Fourier-transformed distribution function corresponding to the correctly normalized solution of the telegraph equation.
4. We propose an N-dimensional subspace approximation to numerically solve the pitch-angle scattering equation. This approach can be several orders of magnitude faster than standard solvers.
5. All numerical and analytical approaches are compared with each other. This will help us understanding the respective advantages and disadvantages of the different techniques.

Those tasks will be performed in Sections. 2–4, and in Section 5, we provide the summary and and conclusions. This article has several appendices containing mathematical details.

2 The two-dimensional subspace approximation

The two-dimensional subspace approximation was originally developed by Shalchi et al. (2011) to solve pitch-angle scattering Equation 1. We summarize the corresponding results, rewrite previously found solutions, and discuss the relation to the telegraph equation as follows. The following three subsections were mostly taken from Shalchi (2020) but have been modified significantly.

2.1 The isotropic scattering coefficient

For the isotropic scattering coefficient, as given by Eq. 2, the parallel spatial diffusion coefficient is obtained via Eq. 5. Alternatively, we can compute the parallel mean free path that is defined via $\lambda_{\parallel} = 3\kappa_{\parallel}/v$. For the isotropic case, those parameters are given by

$$\kappa_{\parallel} = \frac{v^2}{6D} \qquad \text{and} \qquad \lambda_{\parallel} = \frac{v}{2D}. \tag{9}$$

Eq. 1 corresponds to a partial differential equation with the variables t, z, and μ. As a first step toward a solution, we use the Fourier transform

$$f(z,\mu,t) = \int_{-\infty}^{+\infty} dk\, F_k(\mu,t)\, e^{ikz} \tag{10}$$

so that the pitch-angle scattering equation becomes

$$\frac{\partial F_k}{\partial t} + iv\mu k F_k = \frac{\partial}{\partial \mu}\left[D_{\mu\mu}\frac{\partial F_k}{\partial \mu}\right]. \tag{11}$$

The inverse Fourier transform is then given by the following equation:

$$F_k(\mu,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dz\, f(z,\mu,t)\, e^{-ikz}. \tag{12}$$

For the isotropic scattering coefficient, Eq. 11 is simplified to

$$\frac{\partial F_k}{\partial t} + iv\mu k F_k = D\frac{\partial}{\partial \mu}\left[(1-\mu^2)\frac{\partial F_k}{\partial \mu}\right]. \tag{13}$$

To continue, we expand the solution of Eq. 13 in a series of *Legendre polynomials*

$$F_k(\mu,t) = \sum_{n=0}^{\infty} C_n(t)\, P_n(\mu), \tag{14}$$

where the coefficients C_n are functions of time, though they also depend on k. This dependence is not explicitly written down during the following investigations. Using Eq. 14 in the differential Equation 13 yields

$$\begin{aligned}&\sum_n \dot{C}_n P_n + iv\mu k \sum_n C_n P_n \\ =&D\sum_n C_n \frac{\partial}{\partial \mu}\left[(1-\mu^2)\frac{\partial P_n}{\partial \mu}\right],\end{aligned} \tag{15}$$

where $\dot{C}_n$ denotes the time derivative of the coefficient C_n. In order to further valuate Eq. 15, we use the following two relations for Legendre polynomials (Abramowitz and Stegun, 1974)

$$\frac{\partial}{\partial \mu}\left[(1-\mu^2)\frac{\partial P_n}{\partial \mu}\right] = -n(n+1)P_n \tag{16}$$

and

$$\mu P_n = \frac{n+1}{2n+1}P_{n+1} + \frac{n}{2n+1}P_{n-1}. \tag{17}$$

With those two relations, Eq. 15 can be written as follows:

$$\begin{aligned}&\sum_n \dot{C}_n P_n + ivk\sum_n C_n\left(\frac{n+1}{2n+1}P_{n+1} + \frac{n}{2n+1}P_{n-1}\right) \\ &\quad = -D\sum_n C_n n(n+1)P_n.\end{aligned} \tag{18}$$

To continue, we multiply this equation by the Legendre polynomial P_m, integrate over μ, and use the orthogonality relation of Legendre polynomials (Abramowitz and Stegun, 1974)

$$\int_{-1}^{+1} d\mu\, P_n P_m = \frac{2}{2m+1}\delta_{nm}. \tag{19}$$

After performing those steps, we derive the *recurrence relation*

$$\begin{aligned}\dot{C}_m = &-Dm(m+1)C_m - ivk\frac{m}{2m-1}C_{m-1} \\ &- ivk\frac{m+1}{2m+3}C_{m+1}.\end{aligned} \tag{20}$$

Alternatively, one can use the coefficient Q_m defined via

$$C_m = (2m+1)(-i)^m Q_m. \tag{21}$$

With this, the *recurrence relation* can be written as follows

$$\begin{aligned}(2m+1)\dot{Q}_m = &-Dm(m+1)(2m+1)Q_m \\ &+ vkmQ_{m-1} - vk(m+1)Q_{m+1}.\end{aligned} \tag{22}$$

For the case of no scattering $D = 0$, we can compare this with the relation (see Equation of Abramowitz and Stegun (1974))

$$(2n+1)j_n' = nj_{n-1} - (n+1)j_{n+1}, \tag{23}$$

where we have used *spherical Bessel functions*. Thus, we find $Q_m = j_m(vkt)$ for the scatter-free case and

$$C_m = (2m+1)(-i)^m j_m(vkt). \tag{24}$$

Using this in Eq. 14 yields

$$\begin{aligned}F_k(\mu,t) &= \sum_{n=0}^{\infty}(2n+1)(-i)^n j_n(vkt)P_n(\mu) \\ &= e^{-iv\mu kt},\end{aligned} \tag{25}$$

where we have used Equation 92 from Shalchi et al. (2011). This is also known as *plane wave expansion* widely used in quantum mechanics. It should be noted that Eq. 25 corresponds to the unperturbed or scatter-free solution. It can be easily obtained directly from Eq. 11 for the case $D = 0$.

We use Eq. 20 which corresponds to an infinite set of coupled ordinary differential equations. For $m = 0$, for instance, we find

$$\dot{C}_0 = -\frac{1}{3}ivkC_1 \tag{26}$$

and for $m = 1$, we obtain

$$\dot{C}_1 = -2DC_1 - ivkC_0 - \frac{2}{5}ivkC_2. \tag{27}$$

It is problematic here that the coefficients C_0 and C_1 are coupled to C_2. Therefore, it is not possible to derive an exact solution for the coefficients C_n.

2.2 The two-dimensional approximation

Since an exact solution to Eq. 20 seems impossible to be found, one needs to rely on approximations. In the following, we discuss the two-dimensional (2D) subspace approximation originally developed by Shalchi et al. (2011), meaning we set

$$C_m = 0 \quad \text{for} \quad m \geq 2 \tag{28}$$

so that only the coefficients C_0 and C_1 are used. In Lasuik and Shalchi (2019), one can find the solution obtained by using a three-dimensional subspace approximation. It is shown that the three-dimensional (3D) solution is too complicated for most applications.

Within the two-dimensional subspace approximation, the expansion (14) is reduced to

$$F(\mu,t) = C_0 + \mu C_1. \tag{29}$$

In this case, Eqs 26, 27 can be combined to eliminate C_1. Since we set $C_2 = 0$, we found the second-order differential equation

$$\ddot{C}_0 = -2D\dot{C}_0 - \frac{1}{3}v^2k^2C_0. \tag{30}$$

Using the *ansatz*

$$C_0 = be^{\omega t} \tag{31}$$

leads to the quadratic equation

$$\omega^2 + 2D\omega + \frac{1}{3}v^2k^2 = 0. \tag{32}$$

Alternatively, for $C_2 = 0$, Eqs 26, 27 can be written as the matrix equation

$$\begin{pmatrix}\dot{C}_0\\ \dot{C}_1\end{pmatrix} = \begin{pmatrix}0 & -ivk/3\\ -ivk & -2D\end{pmatrix}\begin{pmatrix}C_0\\ C_1\end{pmatrix}. \tag{33}$$

After using Eq. 31 for both functions $C_0(t)$ and $C_1(t)$, the problem of finding the two ω is expressed as

$$\det\begin{pmatrix}\omega & ivk/3\\ ivk & \omega + 2D\end{pmatrix} = 0, \tag{34}$$

leading to the same quadratic equation as given by Eq. 32. The latter equation can easily be solved by the following equation:

$$\omega_\pm = -D \pm \sqrt{D^2 - \frac{1}{3}v^2k^2}. \tag{35}$$

We conclude that the eigenvalues can be complex depending on the wave number k. With this, the coefficient C_0 can be written as the linear combination

$$C_0 = b_+e^{\omega_+ t} + b_-e^{\omega_- t} \tag{36}$$

with the two unknown coefficients $b_\pm$. It follows from Eq. 26 that

$$C_1 = -\frac{3}{ivk}\left(b_+\omega_+e^{\omega_+ t} + b_-\omega_-e^{\omega_- t}\right). \tag{37}$$

The coefficients $b_\pm$ will be determined below. Before we perform this task, we write down the solution for $F_k(\mu,t)$. We need to combine Eq. 29 with Eqs 36, 37 to derive

$$\begin{aligned}F_k(\mu,t) = {} & b_+e^{\omega_+ t} + b_-e^{\omega_- t}\\ & - \frac{3\mu}{ivk}\left(b_+\omega_+e^{\omega_+ t} + b_-\omega_-e^{\omega_- t}\right).\end{aligned} \tag{38}$$

In order to find the coefficients $b_\pm$, we can use the initial condition

$$f(z,\mu,t=0) = 2\delta(z)\,\delta(\mu-\mu_0), \tag{39}$$

meaning that the particle has its initial position at $z = 0$ and the initial pitch-angle cosine μ_0. Using this in the inverse Fourier transform given by Eq. 12, yields after some straightforward algebra

$$F_k(\mu,t=0) = \frac{1}{\pi}\delta(\mu-\mu_0). \tag{40}$$

The latter initial condition used in expansion (Eq. 14) allows us to write

$$\sum_n C_n(t=0)P_n(\mu) = \frac{1}{\pi}\delta(\mu-\mu_0). \tag{41}$$

In order to determine the coefficients $C_n(t=0)$, we multiply this by P_m and integrate over μ to get

$$C_m(t=0) = \frac{2m+1}{2\pi}P_m(\mu_0). \tag{42}$$

To perform this task, we have used again the orthogonality relation (Eq. 19). For $m = 0$ and $m = 1$, this yields[1]

$$C_0(t=0) = \frac{1}{2\pi} \tag{43}$$

and

$$C_1(t=0) = \frac{3\mu_0}{2\pi}. \tag{44}$$

To determine the coefficients $b_\pm$, we write down Eqs 36, 37 for $t = 0$ and use Eqs 43, 44 to deduce

$$\begin{aligned}b_+ + b_- &= \frac{1}{2\pi},\\ b_+\omega_+ + b_-\omega_- &= -\frac{ivk\mu_0}{2\pi}.\end{aligned} \tag{45}$$

This system of two equations is solved by the following equation

$$b_\pm = \mp\frac{ivk\mu_0 + \omega_\mp}{2\pi(\omega_+ - \omega_-)}. \tag{46}$$

Using this result and Eq. 35 in Eq. 38 provides the two-dimensional subspace approximation to the solution $F_k(\mu,t)$. In Section 2.4, we provide a more detailed discussion of this solution.

Our solution is based on the expansion given by Eq. (29). One can easily demonstrate using Eq. 3 and

$$J(z,t) = \frac{v}{2}\int_{-1}^{+1}d\mu\;\mu f(\mu,z,t), \tag{47}$$

together with the orthogonality relation (Eq. 19), that the function $C_0(t)$ corresponds to the Fourier transform of the pitch-angle-averaged distribution function $M(z,t)$, and $C_1(t)$ corresponds to the Fourier transform of the current density or diffusion flux $J(z,t)$. Those two quantities are related to each other via the one-dimensional continuity equation

$$\frac{\partial M}{\partial t} + \frac{\partial J}{\partial z} = 0 \tag{48}$$

which is obtained by averaging Eq. 1 over all μ and using Eqs 3, 47. The exact relations to the coefficients are

$$M(z,t) = \int_{-\infty}^{+\infty}dk\;C_0(t)\,e^{ikz} \tag{49}$$

and

$$J(z,t) = \frac{v}{3}\int_{-\infty}^{+\infty}dk\;C_1(t)\,e^{ikz}. \tag{50}$$

It should be noted that the latter relation is obtained by combining Eq. 47 with Eq. 10 and the expansion given by Eq. 14. After combining these three relations and using the orthogonality relation (Eq. 19), one can obtain Eq. 50. As demonstrated, the coefficients $C_0(t)$ and $C_1(t)$ are directly linked to physical quantities. In particular, the coefficient $C_0(t)$ is very important because it is simply the Fourier transform of the pitch-angle-averaged distribution function $M(z,t)$.

1 Note: there is a typo in Shalchi (2020) where one can find the incorrect formula $C_0(t=0) = 1/(3\pi)$.

2.3 Further physical quantities

An important quantity in particle transport theory is the characteristic function $\langle e^{\pm ikz}\rangle$. We define the ensemble average via

$$\langle A\rangle = \frac{1}{4}\int_{-1}^{+1} d\mu \int_{-1}^{+1} d\mu_0 \int_{-\infty}^{+\infty} dz\, A(z,\mu,t)\, f(z,\mu,t). \tag{51}$$

It should be noted that in some cases, one could aim for a result that depends on μ_0. Then, the corresponding average is omitted.

To determine the characteristic function, we average over all quantities, and thus, we have

$$\begin{aligned}\langle e^{-ikz}\rangle = \frac{1}{4}\int_{-1}^{+1} d\mu_0 \int_{-1}^{+1} d\mu \\ \times \int_{-\infty}^{+\infty} dz\, e^{-ikz} f(z,\mu,t).\end{aligned} \tag{52}$$

Replacing $f(z,\mu,t)$ therein by using Eq. 12 leads to

$$\langle e^{-ikz}\rangle = \frac{\pi}{2}\int_{-1}^{+1} d\mu_0 \int_{-1}^{+1} d\mu\, F_k(\mu,t). \tag{53}$$

We now replace $F_k(\mu,t)$ by using Eq. 14 and use $P_0(\mu)=1$ to get

$$\langle e^{-ikz}\rangle = \frac{\pi}{2}\int_{-1}^{+1} d\mu_0 \int_{-1}^{+1} d\mu \sum_{n=0}^{\infty} C_n P_n P_0. \tag{54}$$

Due to the orthogonality of Legendre polynomials (Eq. 19), this is reduced to

$$\langle e^{-ikz}\rangle = \pi\int_{-1}^{+1} d\mu_0\, C_0. \tag{55}$$

To solve the remaining integral, we use Eq. 36 to write

$$\langle e^{-ikz}\rangle = \pi\int_{-1}^{+1} d\mu_0 \left(b_+ e^{\omega_+ t} + b_- e^{\omega_- t}\right). \tag{56}$$

In order to replace $b_\pm$, we use Eq. 46. The integrals over the terms containing μ_0 vanish, and we finally find

$$\langle e^{\pm ikz}\rangle = \frac{\omega_+ e^{\omega_- t} - \omega_- e^{\omega_+ t}}{\omega_+ - \omega_-}. \tag{57}$$

It should be noted that the parameters $\omega_\pm$ are given by Eq. 35. For the case that the $\omega_\pm$ are real, the characteristic function given by Eq. 57 is real as well. For the case that the $\omega_\pm$ are complex, it follows from Eq. 35 that $\omega_+^* = \omega_-$. Therefore, the characteristic function is always real, and we have $\langle e^{+ikz}\rangle = \langle e^{-ikz}\rangle$.

Based on Eq. 35, it can be shown that Eq. 57 contains two asymptotic limits, namely, (see Shalchi (2020) for more details)

$$\langle e^{\pm ikz}\rangle \approx \begin{cases} e^{-\kappa_\parallel k^2 t} & \text{for } v^2k^2 \ll 3D^2 \\ \cos\left(\dfrac{vkt}{\sqrt{3}}\right) e^{-Dt} & \text{for } v^2k^2 \gg 3D^2. \end{cases} \tag{58}$$

For small wave numbers, we find the characteristic function of diffusion Equation 4. The result obtained for large wave numbers can be understood as a damped unperturbed orbit.

By comparing Eq. 53 with Eq. 10 and using Eq. 3, we can relate the characteristic function to the μ- and μ_0-averaged functions $M(z,t)$. This relation is given by the following equation:

$$M(z,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, \langle e^{\pm ikz}\rangle e^{ikz}. \tag{59}$$

Furthermore, we can compare this with Eq. 49 to find

$$\langle e^{\pm ikz}\rangle = 2\pi C_0(t). \tag{60}$$

As an example, we consider the limit $D\to\infty$ so that we can use the first line of Eq. 58 in Eq. 59. We can easily derive

$$\begin{aligned}M(z,t) &= \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, \cos(kz)\, e^{-\kappa_\parallel k^2 t} \\ &= \frac{1}{\sqrt{4\pi\kappa_\parallel t}} e^{-\frac{z^2}{4\kappa_\parallel t}}\end{aligned} \tag{61}$$

corresponding to a Gaussian solution. The result obtained here is the diffusive solution that one would expect in this case (see Eq. 6 in this paper).

Other physical quantities can be derived by using the subspace approximation, alternative approximations, or even exact calculations (Shalchi, 2006; Shalchi, 2011).

2.4 Rewriting the solution

Eq. 29 corresponds to an integral representation of the solution of the Fourier-transformed pitch-angle scattering equation. This result is based on the 2D subspace approximation. Using therein Eqs 36 and 37, as well as (Eq. 46) yields

$$\begin{aligned}F_k(\mu,t) = \frac{3\mu_0\mu\omega_+ - \omega_- - ivk(\mu_0+\mu)}{2\pi(\omega_+-\omega_-)} e^{\omega_+ t} \\ - \frac{3\mu_0\mu\omega_- - \omega_+ - ivk(\mu_0+\mu)}{2\pi(\omega_+-\omega_-)} e^{\omega_- t},\end{aligned} \tag{62}$$

where the functions $\omega_\pm$ are given by Eq. 35.

The μ- and μ_0-averaged solution is then

$$\begin{aligned}M_k(t) &= \frac{1}{4}\int_{-1}^{+1} d\mu_0 \int_{-1}^{+1} d\mu\, F_k(\mu,t) \\ &= \frac{1}{2\pi(\omega_+-\omega_-)}\left[\omega_+ e^{\omega_- t} - \omega_- e^{\omega_+ t}\right]\end{aligned} \tag{63}$$

in Fourier space. To find the solution in the configuration space, we use Eq. 10 to derive

$$M(z,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk \left[\frac{\omega_+}{\omega_+-\omega_-} e^{\omega_- t} - \frac{\omega_-}{\omega_+-\omega_-} e^{\omega_+ t}\right] e^{-ikz}. \tag{64}$$

Alternatively, this result can also be obtained by combining Eqs 57, 59.

It is convenient to define the parameter

$$\xi := \sqrt{D^2 - \frac{1}{3}v^2k^2}, \tag{65}$$

and it follows from Eq. 35 that

$$\omega_\pm = -D \pm \xi. \tag{66}$$

From this, we can easily deduce

$$\omega_+ - \omega_- = 2\xi. \tag{67}$$

Therewith, the solution in the configuration space is given as the following Fourier transform

$$\begin{aligned}M(z,t) = \frac{1}{2\pi} e^{-Dt}\int_0^{\infty} dk \left[\left(1+\frac{D}{\xi}\right)e^{\xi t} + \left(1-\frac{D}{\xi}\right)e^{-\xi t}\right] \\ \times \cos(kz),\end{aligned} \tag{68}$$

where we have used the integrand, which is an even function of k. With the help of hyperbolic functions, this can be written in a more compact form

$$M(z,t) = \frac{1}{\pi} e^{-Dt} \int_0^\infty dk \left[\cosh(\xi t) + \frac{D}{\xi} \sinh(\xi t) \right] \cos(zk). \quad (69)$$

It should be noted that the quantity ξ, given by Eq. 65, can either be real or imaginary depending on the value of k. Eq. 69 provides an integral representation of the μ- and μ_0-averaged distribution function based on the 2D subspace approximation. Alternative forms are presented in Supplementary Appendix S1 of this paper. In Supplementary Appendix S2, we provide an approximative solution of the remaining integral.

2.5 Relation to the Telegrapher's equation

We have derived an ordinary differential equation for the function $C_0(t)$ previously (Eq. 30), which can be written as follows

$$\ddot{C}_0 + 2D\dot{C}_0 = -\frac{1}{3} v^2 k^2 C_0. \quad (70)$$

As shown via Eq. 49, $C_0(t)$ is the Fourier transform of the distribution function $M(z,t)$. Thus, working in the configuration space instead of the Fourier space allows us to write Eq. 70 as

$$\ddot{M} + 2D\dot{M} = \frac{1}{3} v^2 \frac{\partial^2 M}{\partial z^2}. \quad (71)$$

The latter equation has the same form as Eq. 8, and, thus, it corresponds to a telegraph equation. A quick alternative derivation of the latter equation can be found in Supplementary Appendix S3. It should be noted that the coefficient $C_0(t)$ used here depends also on the initial pitch-angle cosine μ_0. If one averages over the latter quantity, the two-dimensional subspace approximation provides Eq. 69. In Supplementary Appendix S4, we demonstrate that the latter form indeed solves Eq. 71. Using therein Eq. 9 and the scattering time $\tau = 1/(2D)$ yields the telegraph equation, as given by Eq. 8. Therefore, we have shown the complete equivalence of the two-dimensional subspace approximation and the telegraph equation. The solution given by Eq. 69 is correctly normalized. In order to demonstrate this, we consider

$$\int_{-\infty}^{+\infty} dz\, M(z,t) = \frac{1}{\pi} e^{-Dt} \int_0^\infty dk \left[\cosh(\xi t) + \frac{D}{\xi} \sinh(\xi t) \right] \times \int_{-\infty}^{+\infty} dz\, \cos(zk). \quad (72)$$

Therein, we use (Zwillinger, 2012)

$$\int_{-\infty}^{+\infty} dz\, e^{i(k'-k)z} = 2\pi\delta(k'-k) \quad (73)$$

to write this as

$$\int_{-\infty}^{+\infty} dz\, M(z,t) = e^{-Dt} \int_{-\infty}^{+\infty} dk \left[\cosh(\xi t) + \frac{D}{\xi} \sinh(\xi t) \right] \delta(k) = e^{-Dt} \left[\cosh(\xi t) + \frac{D}{\xi} \sinh(\xi t) \right]_{k=0}. \quad (74)$$

From Eq. 65, it follows that $\xi_{k=0} = D$, and, thus, we find

$$\int_{-\infty}^{+\infty} dz\, M(z,t) = e^{-Dt} [\cosh(Dt) + \sinh(Dt)] = 1. \quad (75)$$

As already pointed out in Tautz and Lerche (2016), one can use the transformation

$$M(z,t) = \Psi(z,t)\, e^{-Dt}, \quad (76)$$

and Eq. 8 becomes

$$\tau\ddot{\Psi} - \kappa_\parallel \frac{\partial^2 \Psi}{\partial z^2} = \frac{1}{4\tau} \Psi. \quad (77)$$

This corresponds to the *Klein–Gordon equation* but with imaginary mass. After comparing Eqs 69, 76 with each other, we can easily read off the function $\Psi(z,t)$.

We have focused on the function $C_0(t)$ previously. We can also derive an ordinary differential equation for $C_1(t)$. By combining Eqs 26, 27, we derive

$$\ddot{C}_1 + 2D\dot{C}_1 = -\frac{1}{3} v^2 k^2 C_1, \quad (78)$$

where we have set $C_2 = 0$ corresponding to the 2D subspace approximation. Eq. 78 is the same equation as we have derived above for C_0. The function $C_1(t)$ corresponds to the Fourier-transformed current density, as shown by Eq. 50. Therefore, the telegraph and Klein–Gordon equations can also be derived for the current density function. In order to obtain the current density, as a further solution to the telegraph equation, we can combine Eq. 69 with the continuity Equation 48. We can easily derive

$$J(z,t) = \frac{v^2}{3\pi} e^{-Dt} \int_0^\infty dk\, \frac{k}{\xi} \sinh(\xi t) \sin(zk), \quad (79)$$

where ξ is given by Eq. 65. Of course, integrating the obtained $J(z,t)$ over all z yields 0, meaning that the found solution to the telegraph equation is not normalized to 1.

3 The N-dimensional subspace approximation

Previously, we have used the expansion in the Legendre polynomials (see Eq. 14 of this paper). The functions $C_n(t)$ therein are given by the recurrence relation (Eq. 20). This relation is still exact and can be written as the following matrix equation

$$\frac{d}{dt} \vec{C} = \mathbf{A} \vec{C} \quad (80)$$

with the tridiagonal matrix $\mathbf{A}$ having the components

$$\begin{aligned} A_{n,n-1} &= -ivk \frac{n}{2n-1}, \\ A_{n,n} &= -n(n+1)D, \\ A_{n,n+1} &= -ivk \frac{n+1}{2n+3}. \end{aligned} \quad (81)$$

The vector $\vec{C}$ in Eq. 80 contains the functions $C_n(t)$ needed in the expansion given by Eq. 14. The formal solution of Eq. 80 can be written as follows:

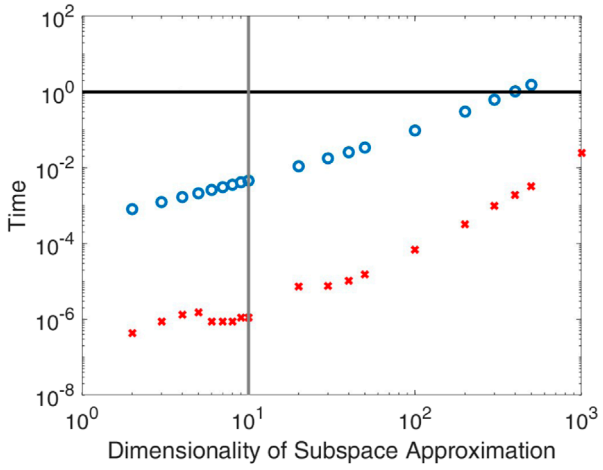

FIGURE 1
Shown are runtimes of codes used to solve the pitch-angle scattering equation based on different techniques. The black horizontal line represents the pure numerical solution, providing a result which depends on the pitch-angle cosine μ and the initial pitch-angle cosine μ_0. This pure numerical method is described in Supplementary Appendix S5 and corresponds to an *implicit Euler method*. The blue circles represent the N-dimensional subspace approximation described in Section 3 also providing a pitch-angle-dependent result. The red crosses represent the N-dimensional subspace approximation for the μ- and μ_0-averaged case. For a small dimensionality (small N), the runtimes are insignificant. It should also be noted that one obtains an accurate result for $N = 10$ (vertical gray line), meaning that the subspace approximation is several orders of magnitude faster than standard numerical solvers. It should be noted that all results are normalized with respect to the runtime of the pure numerical method and have been obtained by using MATLAB running on the same computer.

$$\vec{C}(t) = e^{\mathbf{A}t}\vec{C}(t=0), \tag{82}$$

where we have used the *matrix exponential*. The initial conditions $C_n(t=0)$ are given by Eq. 42. Eq. 82 can be easily evaluated with software such as MATLAB. However, it is required to work with a finite matrix $\mathbf{A}$. This corresponds to the subspace approximation outlined above. Let us assume that we work with an $N \times N$-matrix. This then corresponds to an *N-dimensional subspace approximation*. The method described here corresponds to a semi-numerical/semi-analytical approach that solves the pitch-angle scattering equation, but this method can be faster if one needs the solution only for a specific time t. Standard numerical solvers (see Supplementary Appendix S5 of this paper) require a high time resolution to be accurate. Therefore, one typically needs to work with roughly thousand time-steps so that the solution converges to the true solution of the differential equation. The N-dimensional subspace approximation can be applied to a single time value. As shown via Figures 2–9, an accurate solution is obtained for $N = 10$.

For certain applications, one could be interested in the μ- and μ_0-averaged solution only. Analytical solutions of diffusion and telegraph equations are incomplete and inaccurate depending on the considered application. For the case of pitch-angle-averaged solutions, the N-dimensional subspace approximation is particularly powerful, as outlined below. First, we define the matrix exponential used already above via

$$\mathbf{E} \coloneqq e^{\mathbf{A}t}. \tag{83}$$

Then, Eq. 82 can be written as follows

$$\vec{C}(t) = \mathbf{E}\vec{C}(t=0) \tag{84}$$

or in component notation,

$$C_n(t) = \sum_{m=0}^{N-1} E_{nm} C_m(t=0). \tag{85}$$

At the initial time, the components of the vector $\vec{C}(t=0)$ are given by Eq. 42. If those coefficients are averaged over μ_0, we can easily derive

$$C_m(t=0) = \frac{1}{2\pi}\delta_{m0}, \tag{86}$$

meaning that all coefficients are 0, except $C_0(t=0)$. Therefore, we can write the time-dependent coefficients as

$$C_n(t) = E_{n0}C_0(t=0) \equiv \frac{1}{2\pi}E_{n0}. \tag{87}$$

Furthermore, the μ-dependent solution is given by Eq. 14. After μ-averaging of the latter expansion, we find

$$M_k(t) = C_0(t) = \frac{1}{2\pi}E_{00}, \tag{88}$$

where $M_k(t)$ is the Fourier-transformed distribution function as observed by Eq. 49. It should be noted that the function $C_0(t)$ discussed here is also μ_0-averaged. Furthermore, the characteristic function is easily obtained via

$$\left\langle e^{\pm ikz} \right\rangle = E_{00}, \tag{89}$$

meaning that the 00-component is simply the characteristic function. Thus, it follows from Eq. 59 that

$$M(z,t) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} dk\, E_{00}(k,t)\, e^{ikz}, \tag{90}$$

which is an integral and matrix exponential representation of the μ- and μ_0-averaged distribution function. Therefore, in order to obtain the distribution function $M(z,t)$ for given z and t, we need to numerically solve the k-integral in Eq. 90. For each value of k, we set up the matrix $\mathbf{A}$ defined via Eq. 81, numerically compute the matrix exponential $\mathbf{E}$, and use the component E_{00} in the numerically evaluated k-integral. The distribution functions shown in Figures 6, 7, based on the 10D subspace approximation, for instance, can be computed with a regular computer within a few seconds. Figure 1 shows a comparison in speed between different numerical methods. This comparison includes the N-dimensional subspace approximation described above and the pure numerical approach described in Supplementary Appendix S5 of this paper, which corresponds to an *implicit Euler method*.

The μ- and μ_0-dependent Fourier-transformed solution is given by the following equation:

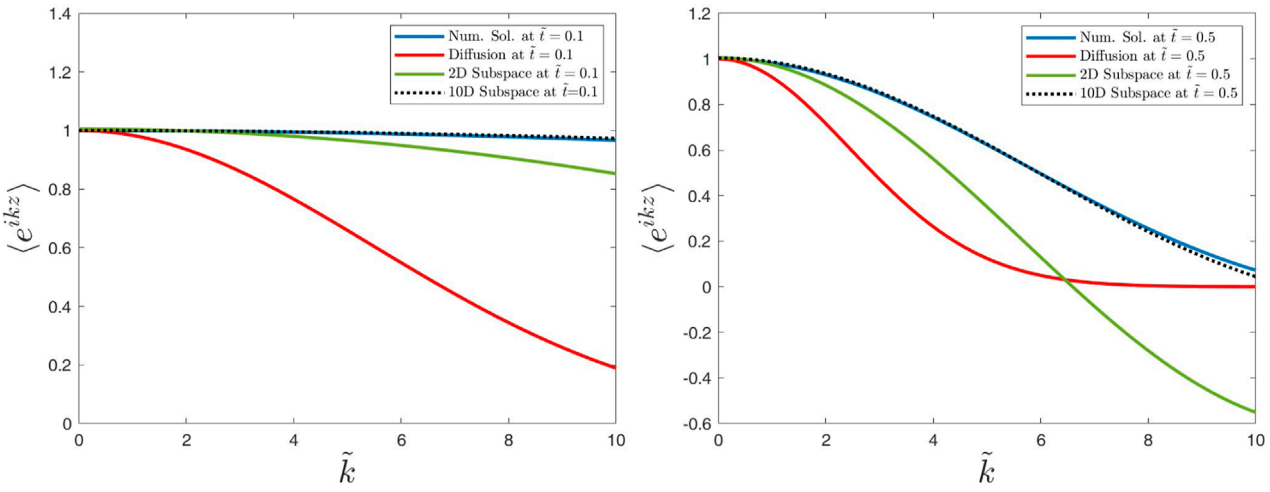

FIGURE 2
Numerical and analytical solutions obtained for the characteristic function $\langle e^{ikz} \rangle$ *versus* the dimensionless wave number $\tilde{k} = vk/D$. The numerical solution refers to the *implicit Euler method* described in Supplementary Appendix S5, and the N-dimensional subspace approximation, which is a semi-analytical/semi-numerical method, is described in Section 3. Shown are plots for $\tilde{t} = Dt = 0.1$ (left panel) and $\tilde{t} = 0.5$ (right panel). For the initial pitch-angle cosine, we have used $\mu_0 = 0$. It should be noted that the characteristic function is μ-averaged.

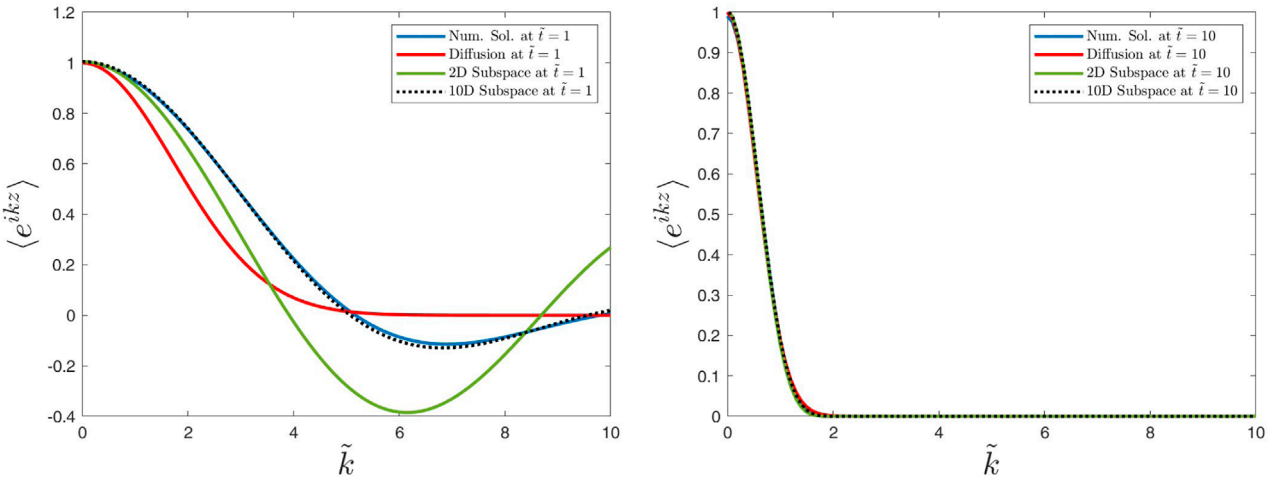

FIGURE 3
Caption is as in Figure 2, but we have considered the times $\tilde{t} = 1$ (left panel) and $\tilde{t} = 10$ (right panel). It should be noted that for the latter case, all four results are in coincidence.

$$F_k(\mu,t) = \frac{1}{2\pi} \sum_{n,m} (2m+1)\, E_{nm}(k,t)\, P_n(\mu)\, P_m(\mu_0), \tag{91}$$

where we have combined Eqs 14, 42, and 85. The Fourier transform can be performed using Eq. 10 and solving the k-integral numerically. Of course, obtaining and plotting the pitch-angle-dependent result is more time-consuming when compared to the pitch-angle-averaged solution.

In certain analytical theories developed for describing the perpendicular transport of energetic particles, one needs to know the function (Shalchi, 2010; Shalchi, 2017; Shalchi, 2020; Shalchi, 2021)

$$\Gamma_k(t) \coloneqq \langle \mu_0 \mu e^{-ikz} \rangle \tag{92}$$

that is somewhat similar but not identical compared to the characteristic function discussed above. In order to express $\Gamma_k(t)$ as before, we perform the same mathematical steps. The pitch-angle-dependent solution is given by Eq. 14. In order to obtain $\Gamma_k(t)$, we need

$$\Gamma_k(t) = \frac{1}{4} \int_{-1}^{+1} d\mu \int_{-1}^{+1} d\mu_0\, \mu\mu_0 \times \int_{-\infty}^{+\infty} dz\, f(z,\mu,t)\, e^{-ikz}. \tag{93}$$

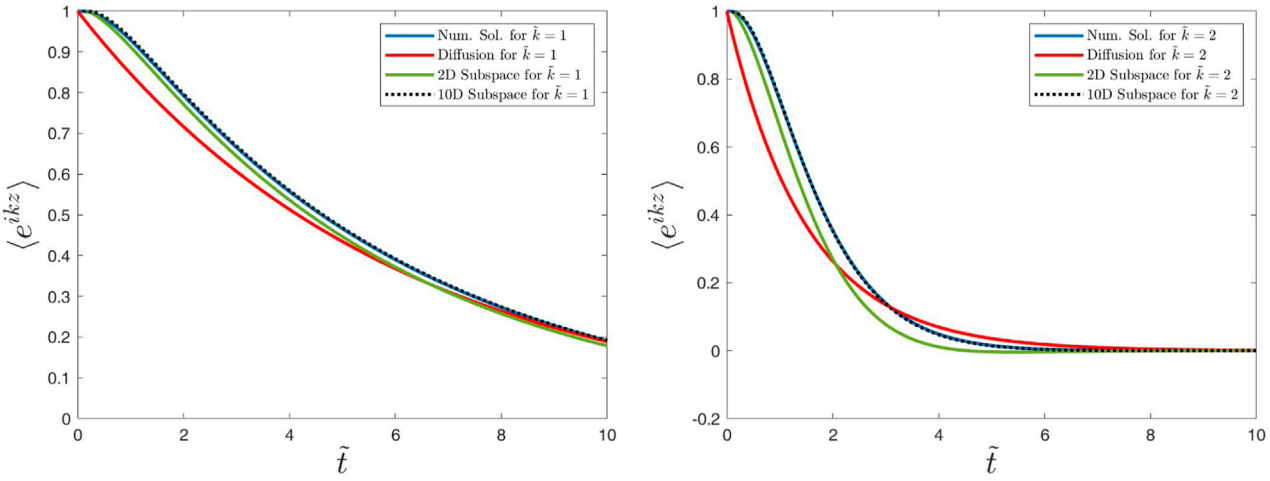

FIGURE 4
Numerical and analytical solutions obtained for the characteristic function $\langle e^{ikz}\rangle$ *versus* dimensionless time $\tilde{t} = Dt$. Shown are plots for $\tilde{k} = 1$ (left panel) and $\tilde{k} = 2$ (right panel). For the initial pitch-angle cosine, we have used $\mu_0 = 0$. It should be noted that the characteristic function is μ-averaged.

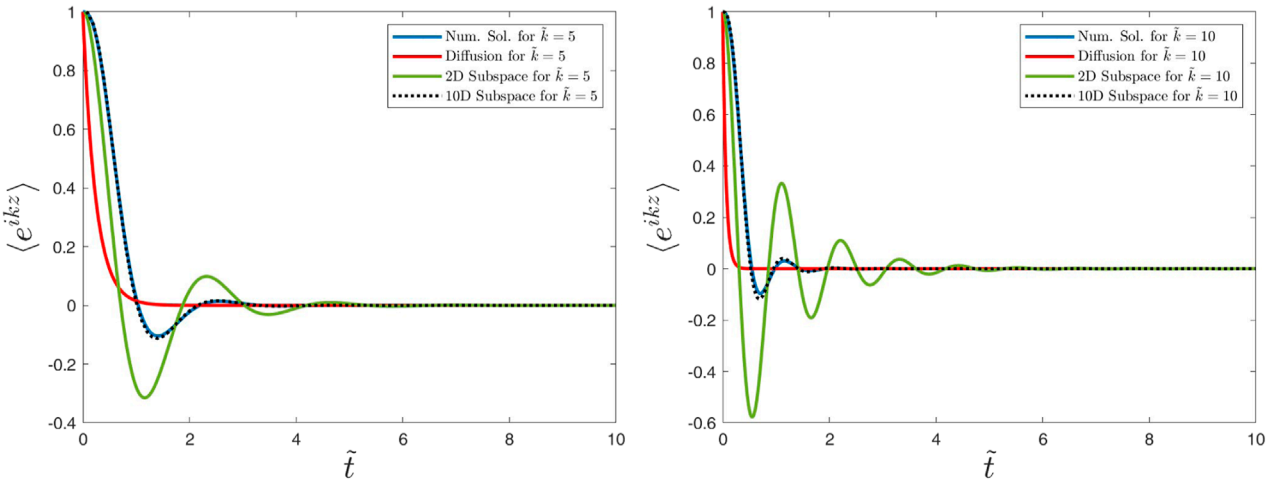

FIGURE 5
Caption is as in Figure 4, but we have considered the values $\tilde{k} = 5$ (left panel) and $\tilde{k} = 10$ (right panel).

To evaluate this further, we use Eqs 12, 14. After using those two relations, we derive

$$\begin{aligned}\Gamma_k(t) &= \frac{\pi}{2}\int_{-1}^{+1} d\mu \int_{-1}^{+1} d\mu_0\, \mu\mu_0 F_k(\mu,t)\\ &= 2\pi\sum_{n=0}^{\infty}\frac{1}{2}\int_{-1}^{+1} d\mu_0\, \mu_0 C_n(t)\,\frac{1}{2}\int_{-1}^{+1} d\mu\, \mu P_n(\mu).\end{aligned} \tag{94}$$

For the μ-integral, we can use the orthogonality relation (Eq. 19) to find

$$\frac{1}{2}\int_{-1}^{+1} d\mu\, \mu P_n(\mu) = \frac{1}{3}\delta_{n1}. \tag{95}$$

Using the above relation allows us to perform the following steps:

$$\begin{aligned}\Gamma_k(t) &= \frac{\pi}{3}\int_{-1}^{+1} d\mu_0\, \mu_0 C_1(t)\\ &= \frac{\pi}{3}\sum_{m=0}^{N-1} E_{1m}\int_{-1}^{+1} d\mu_0\, \mu_0 C_m(t=0)\\ &= \frac{1}{6}\sum_{m=0}^{N-1}(2m+1)\,E_{1m}\int_{-1}^{+1} d\mu_0\, P_1(\mu_0)\,P_m(\mu_0)\\ &= \frac{1}{3}\sum_{m=0}^{N-1} E_{1m}\delta_{m1}\\ &= \frac{1}{3}E_{11},\end{aligned} \tag{96}$$

where we have used Eqs 19, 42, and 85. Therefore, the derived function $\Gamma_k(t)$ corresponds to the matrix element E_{11} which can be computed quickly based on the N-dimensional subspace

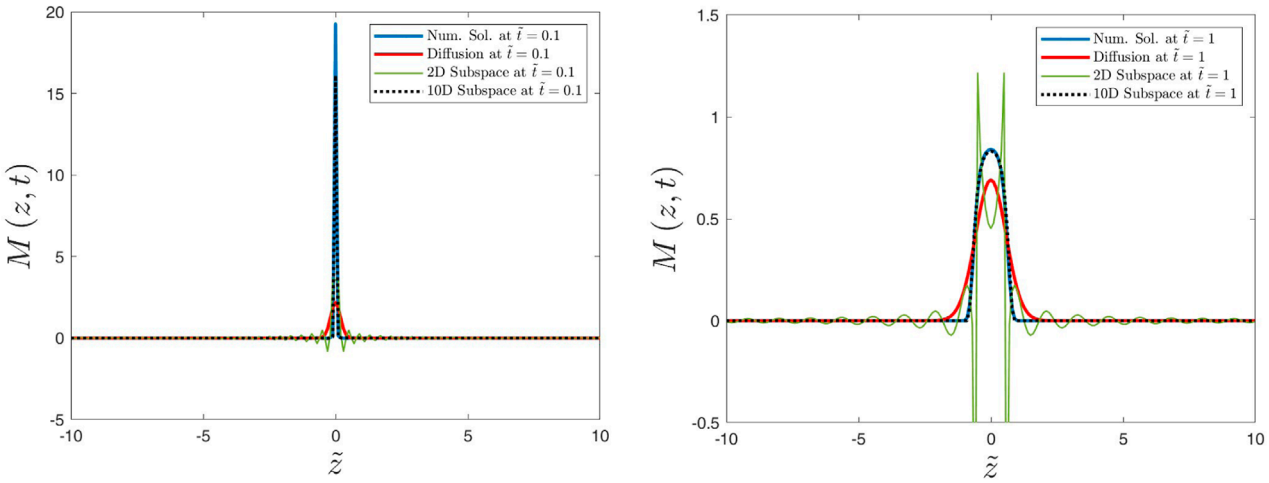

FIGURE 6
Numerical and analytical solutions for the μ-averaged distribution function $M(z,t)$ *versus* the parallel position $\tilde{z} = Dz/v$. Shown are plots for $\tilde{t} = 0.1$ (left panel) and $\tilde{t} = 1$ (right panel). For the initial pitch-angle cosine, we have used $\mu_0 = 0$.

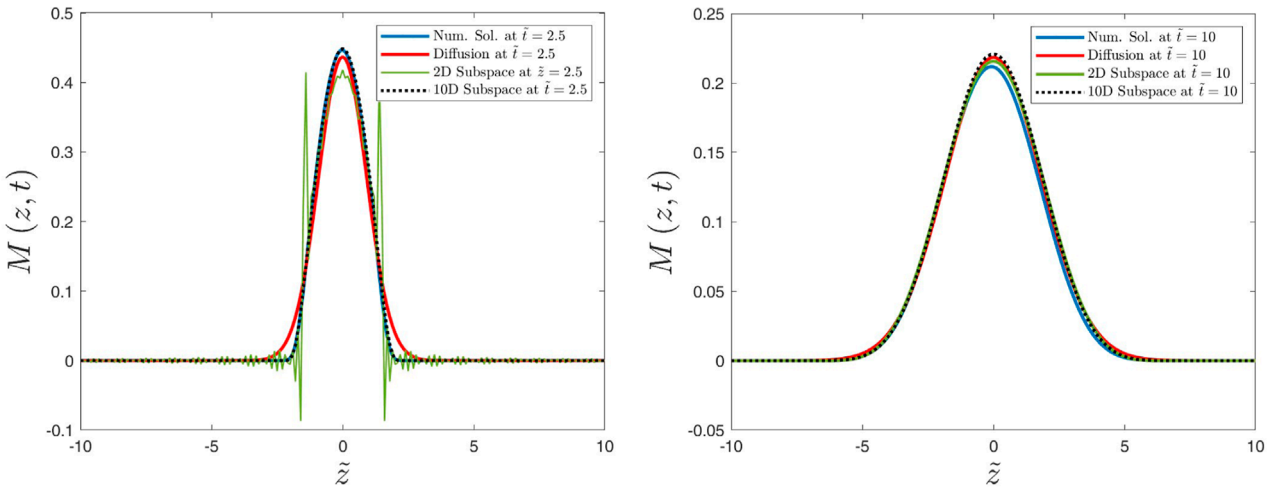

FIGURE 7
Caption is as in Figure 6, but we have considered the times $\tilde{t} = 2.5$ (left panel) and $\tilde{t} = 10$ (right panel).

approximation. Figure 12 shows some example plots for the quantity $\Gamma_k(t)$.

4 Comparison of results

We have solved the pitch-angle scattering equation numerically using an *implicit Euler method* (Supplementary Appendix S5) and the N-dimensional subspace approximation outlined in the previous section. We have considered two cases, namely, $N = 2$ (corresponding to the pure analytical case discussed above) and $N = 10$ (which provides an accurate result). In most cases, we have only considered the μ-averaged solution to reduce the number of plots. Some results are also averaged over the initial pitch-angle cosine μ_0.

Figures 2–5 show the characteristic function $\langle e^{ikz} \rangle \equiv \langle e^{-ikz} \rangle$, which corresponds to the Fourier transform of the distribution function $M(z,t)$. In Figures 2, 3, the characteristic function is plotted *versus* the dimensionless wave number $\tilde{k} = vk/D$ for different values of the dimensionless time $\tilde{t} = Dt$. We have also shown the solution of the diffusion equation as given by the first line of Eq. 58. We can easily see that all solutions agree with each other at later times. This is not the case for early times where the 2D subspace approximation and the diffusive solution differ significantly from the numerical solution. The 10D subspace approximation agrees very well with the numerical solution in all considered cases.

Figures 4, 5 show the characteristic function *versus* time $\tilde{t}$ for different values of $\tilde{k}$. We can easily see agreement for smaller values of $\tilde{k}$ but disagreement for larger values. However, the 10D subspace

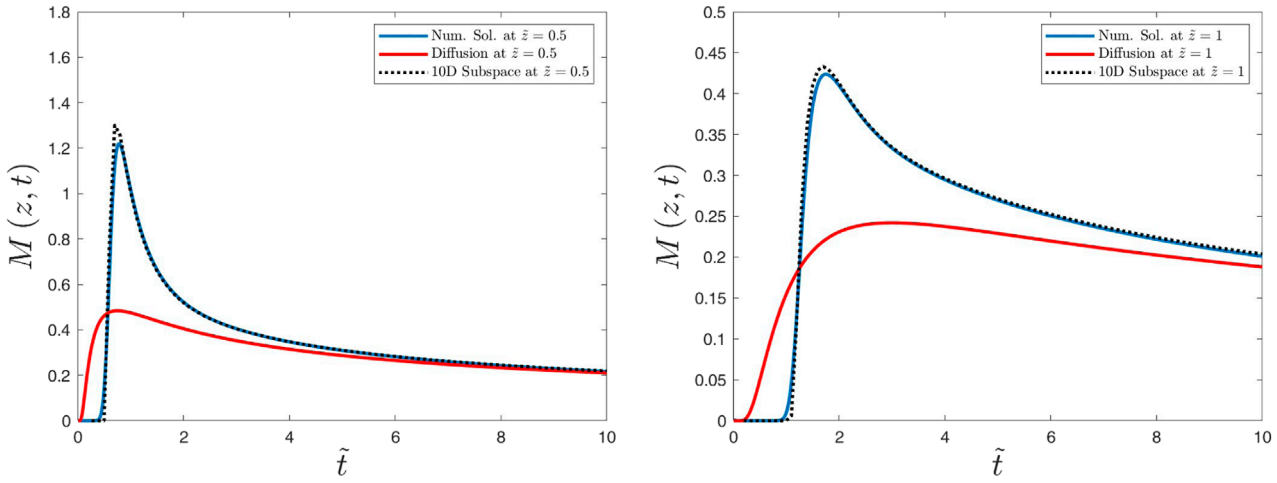

FIGURE 8
Numerical and analytical solutions for the μ-averaged distribution function $M(z,t)$ *versus* time $\tilde{t}$. Shown are plots for $\tilde{z} = 0.5$ (left panel) and $\tilde{z} = 1$ (right panel). For the initial pitch-angle cosine, we have used $\mu_0 = 0.5$.

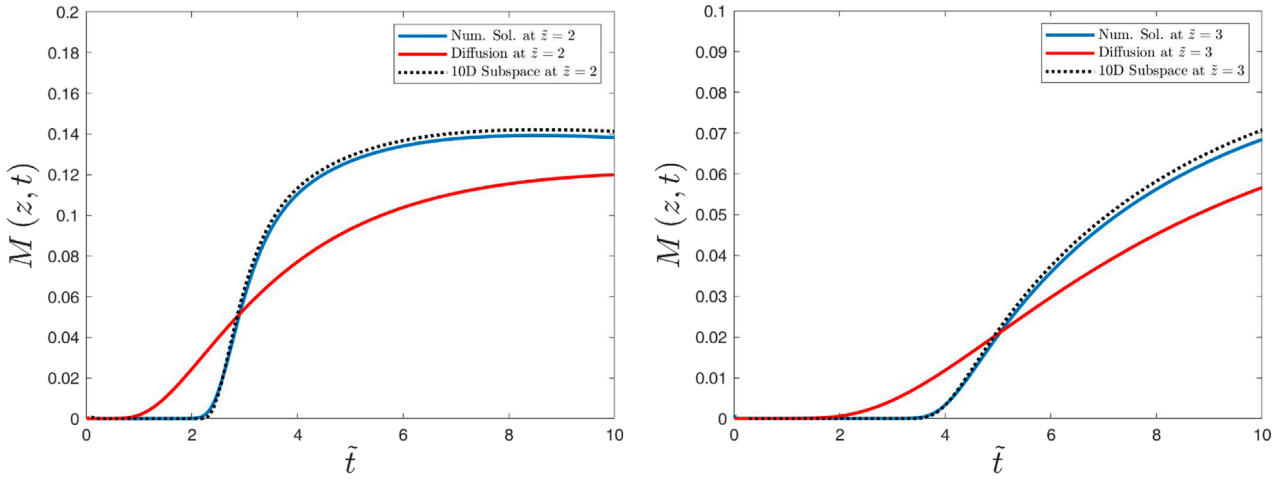

FIGURE 9
Caption is as in Figure 8, but we have considered $\tilde{z} = 2$ (left panel) and $\tilde{z} = 3$ (right panel).

approximation and the numerical solution agree very well with each other. It has to be emphasized that the 10D subspace approximation solution, which can be seen as a semi-analytical/semi-numerical technique, is several orders of magnitude faster than standard numerical solvers.

Figures 6–9 show the distribution function $M(z,t)$. Figures 6, 7 show this function *versus* the dimensionless position $\tilde{z} = Dz/v$ for different times. For late times, all considered results agree with each other, as expected. The corresponding distributions are well-described by the Gaussian given by Eq. 6. For early times, however, diffusive and 2D subspace results do not agree well with the pure numerical solution. The 2D subspace solution contains spikes that are a consequence of the Dirac delta (Supplementary Appendix S2). The diffusive solution is non-zero everywhere in space. Numerical and 10D subspace solutions correctly describe that the distribution function is exactly 0 for $|z| > vt$ due to the finite propagation speed of the particles. The latter effect can be observed much better by plotting the distribution function *versus* time $\tilde{t}$ for different values of $\tilde{z}$. This is done via Figures 8, 9.

Figure 10 shows the comparison of the time evolution of $M(z,t)$ based on diffusion equation and the 10D subspace approximation. We can clearly see the similarity for later times. For early times, on the other hand, we observe significant differences. In particular, the 10D solution provides $M(|z| > vt) = 0$ as needed. Alternatively, we have plotted $M(z,t)$ *versus* time for different positions (Figure 11) where the aforementioned effect can be observed more clearly.

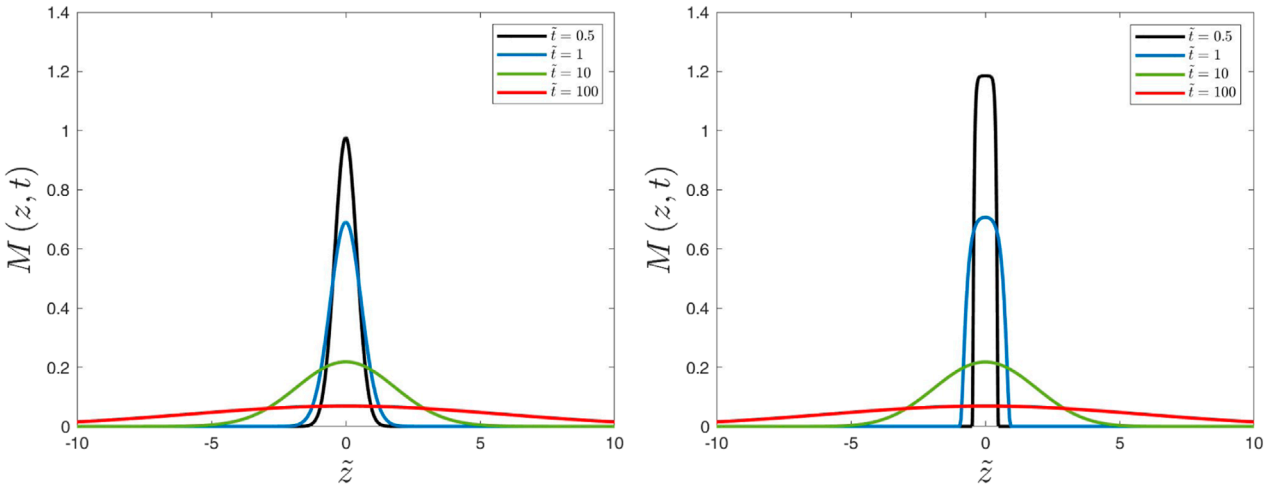

FIGURE 10
Time evolution of distribution functions. The left panel shows the solution of the diffusion equation as given by Eq. 6 for different times, and the right panel shows the μ_0-and μ-averaged solutions of the pitch-angle scattering equation based on the 10D dimensional subspace approximation.

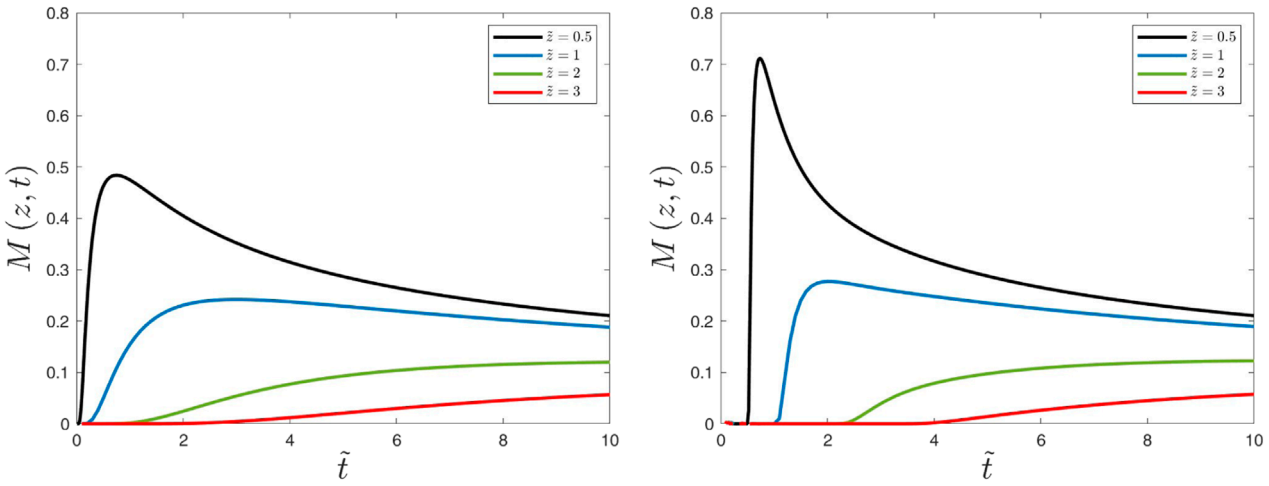

FIGURE 11
Distribution functions *versus* time $\tilde{t}$ at given positions $\tilde{z}$. The left panel shows the solution of the diffusion equation as given by Eq. 6 for different positions, and the right panel shows the μ_0-and μ-averaged solutions of the pitch-angle scattering equation based on the 10-dimensional subspace approximation.

Last but not the least, we have computed the function $\Gamma_k(t)$ defined via Eq. 92. The latter function enters certain analytical theories for perpendicular diffusion. According to Figure 12, the 2D subspace approximation works overall well for computing this quantity. This explains why analytical theories for perpendicular diffusion, in which the 2D subspace approximation was used, agree well with performed test-particle simulations (Shalchi, 2020; Shalchi, 2021).

5 Summary and conclusion

In this paper, we have focused on the most basic transport equation, namely, the pitch-angle scattering equation, as given by Eq. 1. Analytical and numerical investigations of pitch-angle-dependent transport equations have been the subject of several papers published during recent years. In addition to studies of the basic pitch-angle scattering equation (Shalchi et al., 2011; Tautz and Lerche, 2016; Lasuik and Shalchi, 2017; Lasuik and Shalchi, 2019), authors have explored the impact of so-called focusing, an effect which is related to a non-constant mean magnetic field (Danos et al., 2013; Litvinenko and Schlickeiser, 2013; Effenberger and Litvinenko, 2014; Lasuik et al., 2017; Wang and Qin, 2020; Wang and Qin, 2021; Wang and Qin, 2023). Even more complicated cases, including perpendicular particle transport, have been investigated by Wang and Qin (2024).

In this article, we have reviewed the two-dimensional subspace approximation originally developed by Shalchi et al. (2011)

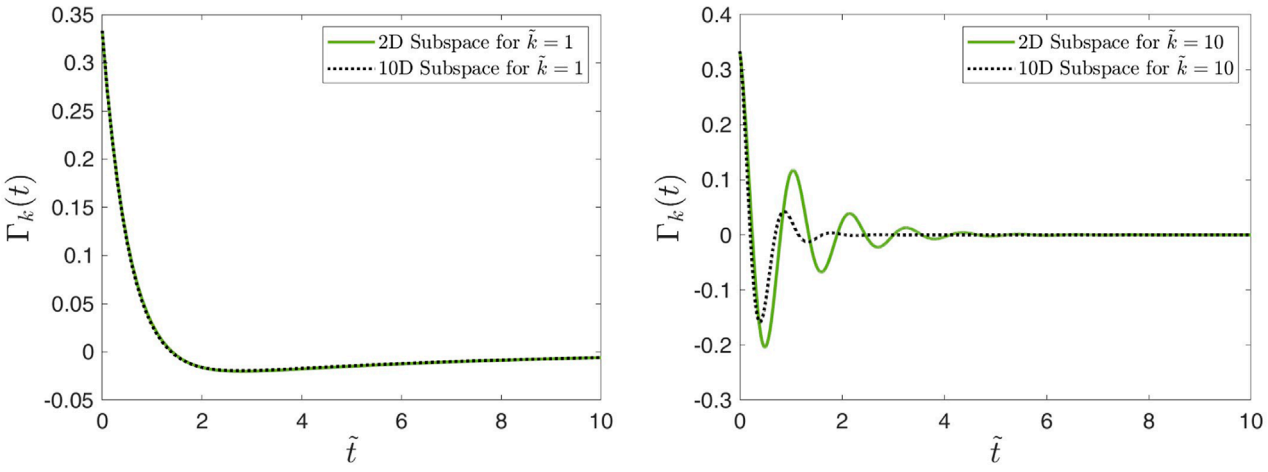

FIGURE 12
Numerical results for the function $\Gamma_k(t)$ as defined via Eq. 92 based on 2D and 10D subspace approximations. The left panel shows results for $\tilde{k}=1$ and the right panel for $\tilde{k}=10$. It should be noted that it follows from the definition of this function that $\Gamma_k(t=0)=1/3$. The results for smaller values of $\tilde{k}$ are not shown here, but the agreement between 2D and 10D subspace approximations would be almost perfect in such cases.

and discussed the provided solutions in configuration and Fourier spaces. We have also demonstrated that the two-dimensional subspace approximation is equivalent to using a telegraph equation. Normalized solutions in configuration and Fourier spaces are also discussed. However, we also argue that such solutions do not often provide appropriate results even if the pitch-angle average is considered. Although it was often argued that the telegraph equation is more complete than the usual diffusion approach (Tautz and Lerche, 2016), the solution discussed here contains artifacts that are not realistic. In particular, we observe spikes at $z=\pm vt$ (Figures 6, 7 as well as Supplementary Appendix S2).

Therefore, it is important to solve the pitch-angle scattering equation numerically. However, standard approaches such as implicit Euler or *Crank–Nicolson* solvers are time-consuming to use. In this paper, we have, thus, developed an N-dimensional subspace approach. This method can be seen as a semi-analytical/semi-numerical method. It has the advantage of being is several orders of magnitude faster than standard solvers (see Figure 1 of this paper). This is in particular the case if one is only interested in pitch-angle-averaged solutions at a given time. Standard solvers require a high time resolution in order to provide an accurate result. The N-dimensional subspace technique can be applied for a single time value if this is everything what is needed. It should also be emphasized that the N-dimensional subspace method can be easily parallelized since for a given k and t, one can compute the matrix exponentials independently of other values. This is also valid if one is looking for a μ- and μ_0-dependent result.

In this paper, we have computed several quantities such as distribution and characteristic functions as well as the function $\Gamma_k(t)$ which is defined via Eq. 92 of this paper. We have compared numerical solutions obtained by using a standard solver with results obtained by using the N-dimensional subspace approximation for different values of N and the diffusive solution. The main difference is that pure numerical and 10D solutions provide $M(|z|>vt)=0$, meaning that the particles have a finite propagation speed. All results are visualized in Figures 2–12. One can clearly see that for $N=10$, we obtain an accurate result that agrees well with the pure numerical solution of the pitch-angle scattering equation.

It has to be noted that the N-dimensional subspace method presented in this paper was specifically developed for the basic pitch-angle scattering equation and an isotropic scattering coefficient. However, this approach can be easily modified so that it can be used for more general transport equations including focused transport equations and other forms of the pitch-angle scattering coefficient.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding author.

Author contributions

AS: writing–original draft and writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The authors also acknowledge the support by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fspas.2024.1385820/full#supplementary-material

References

Abramowitz, M., and Stegun, I. A. (1974). *Handbook of mathematical functions*. New York: Dover Publications.

Crank, J., and Nicolson, P. (1947). A practical method for numerical evaluation of solutions of partial differential equations of the heat conduction type. *Proc. Camb. Phil. Soc.* 43, 50–67. doi:10.1017/s0305004100023197

Danos, R. J., Fiege, J. D., and Shalchi, A. (2013). Numerical analysis of the fokker-planck equation with adiabatic focusing: isotropic pitch-angle scattering. *Astrophysical J.* 772, 35. doi:10.1088/0004-637x/772/1/35

Earl, J. A. (1974). The diffusive idealization of charged-particle transport in random magnetic fields. *Astrophys. J.* 193, 231. doi:10.1086/153152

Effenberger, F., and Litvinenko, Y. E. (2014). The diffusion approximation versus the telegraph equation for modeling solar energetic particle transport with adiabatic focusing, I. Isotropic pitch-angle scattering. *Astrophysical J.* 783, 15. doi:10.1088/0004-637x/783/1/15

Gradshteyn, I. S., and Ryzhik, I. M. (2000). *Table of integrals, series, and products*. New York: Academic Press.

Jokipii, J. R. (1966). Cosmic-Ray propagation. I. Charged particles in a random magnetic field. *Astrophys. J.* 146, 480. doi:10.1086/148912

Lasuik, J., Fiege, J. D., and Shalchi, A. (2017). Numerical analysis of the Fokker-Planck equation with adiabatic focusing: realistic pitch-angle scattering. *Adv. Space Res.* 59, 722–735. doi:10.1016/j.asr.2016.10.027

Lasuik, J., and Shalchi, A. (2017). Solutions of the cosmic ray velocity diffusion equation. *Adv. Space Res.* 60, 1532–1546. doi:10.1016/j.asr.2017.06.035

Lasuik, J., and Shalchi, A. (2019). Subspace approximations to the cosmic ray Fokker-Planck equation. *Mon. Notices R. Astronomical Soc.* 485, 1635–1650. doi:10.1093/mnras/stz474

Litvinenko, Y. E., and Schlickeiser, R. (2013). The telegraph equation for cosmic-ray transport with weak adiabatic focusing. *A A* 554, A59. doi:10.1051/0004-6361/201321327

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (2007). *Numerical Recipes*. Cambridge University Press.

Shalchi, A. (2005). Second-order quasilinear theory of cosmic ray transport. *Phys. Plasmas* 12, 052905. doi:10.1063/1.1895805

Shalchi, A. (2006). Analytical investigation of the two-dimensional cosmic ray Fokker-Planck equation. *Astronomy Astrophysics* 448, 809–816. doi:10.1051/0004-6361:20053664

Shalchi, A. (2009). Nonlinear cosmic ray diffusion theories. *Astrophysics Space Sci. Libr.* 362. Springer, Berlin. doi:10.1007/978-3-642-00309-7

Shalchi, A. (2010). A unified particle diffusion theory for cross-field scattering: subdiffusion, recovery of diffusion, and diffusion in 3D turbulence. *Astrophys. J.* 720, L127–L130. doi:10.1088/2041-8205/720/2/l127

Shalchi, A. (2011). Velocity correlation functions of charged particles derived from the fokker-planck equation. *Adv. Space Res.* 47, 1147–1164. doi:10.1016/j.asr.2010.12.002

Shalchi, A. (2017). Time-dependent perpendicular transport of energetic particles in magnetic turbulence with transverse complexity. *Phys. Plasmas* 24, 050702. doi:10.1063/1.4982805

Shalchi, A. (2020). Perpendicular transport of energetic particles in magnetic turbulence. *Space Sci. Rev.* 216, 23. doi:10.1007/s11214-020-0644-4

Shalchi, A. (2021). Perpendicular diffusion of energetic particles: a complete analytical theory. *Astrophys. J.* 923, 209. doi:10.3847/1538-4357/ac2363

Shalchi, A., Škoda, T., Tautz, R. C., and Schlickeiser, R. (2009). Analytical description of nonlinear cosmic ray scattering: isotropic and quasilinear regimes of pitch-angle diffusion. *Astron. Astrophys.* 507, 589–597. doi:10.1051/0004-6361/200912755

Shalchi, A., Tautz, R. C., and Rempel, T. J. (2011). Test-particle transport: higher-order correlations and time-dependent diffusion. *Plasma Phys. control. Fusion* 53, 105016. doi:10.1088/0741-3335/53/10/105016

Tautz, R. C., and Lerche, I. (2016). Application of the three-dimensional telegraph equation to cosmic-ray transport. *Res. Astronomy Astrophysics* 16, 162. doi:10.1088/1674-4527/16/10/162

Wang, J. F., and Qin, G. (2020). The invariance of the diffusion coefficient with iterative operations of the charged particle transport equation. *Astrophysical J.* 899, 39. doi:10.3847/1538-4357/aba3c8

Wang, J. F., and Qin, G. (2021). Study of momentum diffusion with the effect of adiabatic focusing. *Astrophysical J. Suppl. Ser.* 257, 44. doi:10.3847/1538-4365/ac1bb3

Wang, J. F., and Qin, G. (2023). Relationship of transport coefficients with statistical quantities of charged particles. *Astrophysical J.* 954, 213. doi:10.3847/1538-4357/ace9d3

Wang, J. F., and Qin, G. (2024). The effect of solar wind on charged particles' diffusion coefficients. *Astrophysical J.* 961, 6. doi:10.3847/1538-4357/ad09b7

Watson, G. N. (2011). *A treatise on the theory of Bessel functions*. Cambridge: Cambridge Mathematical Library.

Zank, G. P. (2014). Transport processes in space Physics and astrophysics. *Lect. Notes Phys.* 877. Springer, New York. doi:10.1007/978-1-4614-8480-6

Zwillinger, D. (2012). *Standard mathematical tables and formulae*. Boca Raton: CRC Press.

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 21 May 2024
DOI 10.3389/fspas.2024.1404145

Check for updates

OPEN ACCESS

EDITED BY
Xiangliang Kong,
Shandong University, China

REVIEWED BY
Rudolf A. Treumann,
Ludwig Maximilian University of Munich, Germany
Hao Ning,
Shandong University, China

*CORRESPONDENCE
J. F. Tang,
tjf1027@163.com

RECEIVED 20 March 2024
ACCEPTED 30 April 2024
PUBLISHED 21 May 2024

CITATION
Tang JF, Wu DJ, Chen L, Tan CM and Wang JB (2024), Electron cyclotron maser instability by evolving fast electron beams in the flare loops. *Front. Astron. Space Sci.* 11:1404145. doi: 10.3389/fspas.2024.1404145

COPYRIGHT
© 2024 Tang, Wu, Chen, Tan and Wang. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Electron cyclotron maser instability by evolving fast electron beams in the flare loops

J. F. Tang[1,2]*, D. J. Wu[3], L. Chen[3], C. M. Tan[2] and J. B. Wang[1]

[1]College of Engineering and Design, Lishui University, Lishui, China, [2]State Key Laboratory of Space Weather, Chinese Academy of Sciences, Beijing, China, [3]Key Laboratory of Planetary Sciences, Purple Mountain Observatory, Chinese Academy of Sciences, Nanjing, China

The electron cyclotron maser instability (ECMI) stands as a pivotal coherent radio emission mechanism widely implicated in various astrophysical phenomena. In the context of solar activity, ECMI is primarily instigated by energetic electrons generated during solar eruptions, notably flares. These electrons, upon leaving the acceleration region, traverse the solar atmosphere, forming fast electron beams (FEBs) along magnetic field lines. It is widely accepted that as these FEBs interact with the ambient plasma and magnetic fields, they give rise to radio and hard X-ray emission. Throughout their journey in the plasma, FEBs undergo modifications in their energy spectrum and velocity spatial distribution due to diverse energy loss mechanisms and changes in ambient plasma parameters. In this study, we delve into the impact of the evolving energy spectrum and velocity anisotropic distribution of FEBs on ECMI during their propagation in flare loops. Our findings indicate that if we solely consider the progressively flattened lower energy cutoff behavior as FEBs descend along flare loops, the growth rates of ECMI decrease accordingly. However, when accounting for the evolution of ambient magnetic plasma parameters, the growth rates of ECMI increase as FEBs delve into denser atmosphere. This underscores the significant influence of the energy spectrum and velocity anisotropy distribution evolution of FEBs on ECMI. Our study sheds light on a more comprehensive understanding of the dynamic spectra of solar radio emissions.

KEYWORDS

electron cyclotron maser instability, energetic electrons, radio radiation, evolution, flare loop

1 Introduction

The Sun is a highly active star, and flares represent some of the most powerful eruptions it experiences. During a flare event, an excess of 10^{32} erg of stored magnetic energy can be rapidly released (Fletcher et al., 2011). This released energy is transferred into the thermal energy of local plasma, the enhanced emission of the entire electromagnetic spectrum, plasma motions such as jets and coronal mass ejections (CME), and the acceleration of particles. While the mechanism by which flares convert magnetic energy into energetic particles remains an unanswered question, magnetic reconnection is generally accepted as a triggering mechanism (Masuda et al., 1994; Yokoyama et al., 2001; Imada et al., 2013). For instance, energetic particles can be accelerated by the electric fields within large-scale reconnecting current sheets during a large two-ribbon flare (Benka and Holman, 1994; Zharkova and Gordovskyy, 2004; Sharykin et al., 2014). As these flare-accelerated fast

e and interact with the ambient plasma, they produce radio and hard X-ray radiation. Consequently, FEBs can be inferred indirectly from solar radio bursts and hard X-ray bursts (Dulk et al., 1992; Holman et al., 2011), or directly detected by particle detectors as they travel into near-Earth interplanetary space (Lin, 2011). The most direct radio observational evidence of FEBs is type III solar radio bursts, which are generally accepted to be produced by FEBs during flare events.

Type III solar radio bursts are one of the most extensively studied during the past few decades. The key issue is how the FEBs generate electromagnetic waves near the local plasma frequency. A so-called plasma emission theory was put forth by (Ginzburg and Zhelezniakov, 1958) to explain the observed characteristics of type III bursts, in which the beam-generated Langmuir waves play a key role. This plasma theory was developed by many authors (Cairns, 1987a; Cairns, 1987b; Cairns, 1987c; Wu et al., 1994; Yoon, 1995; Yoon, 1997; Yoon, 1998) and they all assume that the ambient magnetic field in the source region of type III radio bursts is very weak. For the most solar radio bursts, including type III bursts, they usually produced near an active region in the low corona, so the weak magnetic field approximation is not appropriate because the plasma in the source region is strongly magnetized.

Another important coherent radiation process, the electron cyclotron maser instability (ECMI), which emits radiation near the electron cyclotron frequency and its harmonics via wave-particle interaction was first suggested by Twiss (1958) and Schneider (1959). This early maser instability theory has also been applied to various radio emissions, such as solar radio type I bursts (Twiss and Roberts, 1958; Mangeney and Veltri, 1976), and Jovian decametric radiation (Hirshfield, 1963; Goldstein and Eviatar, 1972; Melrose, 1976). The ECMI theory achieved a major breakthrough when Wu and Lee (1979) accounted for weak relativistic effects and applied this theory to auroral kilometer radiation (AKR). Since then, the ECMI theory has been extensively studied and discussed, being applied to radio emissions from magnetized planets (Hewitt et al., 1981; Zarka, 1998; Hess et al., 2007), various solar radio bursts (Melrose and Dulk, 1982; Aschwanden and Benz, 1988; Aschwanden, 1990; Vlasov et al., 2002; Wu et al., 2002; Yoon et al., 2002; Wang, 2004; Treumann et al., 2011; Zhao et al., 2013), and radio emissions from various stellar environments (Stepanov et al., 2001; Begelman et al., 2005; Hallinan et al., 2008; Callingham et al., 2021). Most earlier theories suggest that ECMI is excited by various velocity anisotropic distributions of energetic electrons, which exhibit an inverted perpendicular population as $\partial f_b/\partial v_\perp > 0$, providing free energy for ECMI. Hard X-ray observations demonstrate that flare-accelerated FEBs usually have an approximate power-law energy spectrum (Lin, 1974; Hudson and Ryan, 1995; Aschwanden, 2002). These power-law FEBs traveling in the solar atmosphere are the main source of solar radio bursts (Aschwanden, 2002; Fleishman, 2004; Wu et al., 2007). A series of recent works (Wu and Tang, 2008; Tang and Wu, 2009; Tang et al., 2011; Tang et al., 2012) have shown that the low-energy cutoff behavior of power-law electrons is another effective driving source of ECMI.

As mentioned above, it is the FEBs traveling in the solar atmosphere and interacting with the ambient plasma that produce radio and hard X-ray radiation. The strong correlation between SRBs and HXRs in some statistical works (Kosugi et al., 1988; White et al., 2011) also implies that they have the same source, i.e., the FEBs (Raoult et al., 1990; Aschwanden, 2002; White et al., 2011). High-resolution imaging observations of HXRs show that the HXR source of a flare often consists of a coronal looptop source and two or more footpoint sources (Hudson, 1978; Hoyng et al., 1981). The standard flare HXR emission model, which involves thin- and thick-target bremsstrahlung emission at looptop and footpoints, respectively, indicates a spectral index difference between the looptop source and footpoint sources, $\Delta\gamma = 2$. However, the majority of HXR observations show that the spectral index difference $\Delta\gamma \neq 2$, implying that the energy spectra of FEBs change during their propagation in the flare loops from the top to foots. The collisional energy loss and deceleration by induced electric fields of energetic electrons are the main reasons for the change of the energy spectra of FEBs. Considering the collisional and noncollisional energy losses of electrons, Tang et al. (2020) investigated the parametric evolution of power-law spectra of flare-accelerated FEBs when propagating in the flare loops. The variations of the energy spectral parameters of power-law FEBs, such as the spectral index α, the steepness index δ, and lower cutoff energy E_c, all have important effects on the maser instability. Tang et al. (2016) investigated the ECMI driven by evolving FEBs in the coronal loop preliminarily. In their model, the magnetic field configuration of the source region is a giant expansion arch structure, and the CME-driven shock front is the acceleration site for FEBs. Since the specific models of density and magnetic field for the coronal loop are not considered, the energy loss of energetic electrons described by the loss factor Q and the magnetic mirror ratio parameter σ are empirically determined. Therefore, further research is needed to fully understand the ECMI theory excited by evolving FEBs and other related issues.

In this paper, we focus on the influence of the evolution of FEBs when propagating in the flare loop on ECMI. We propose that the flare-accelerated FEBs initially exhibit a power-law energy distribution and travel along the flare loop from the acceleration region to the footpoints. Utilizing the coronal magnetic field and density models of active region by Zhao (1995), the semiempirical homogeneous plane-parallel flare atmospheric model by Machado et al. (1980), and the evolution of energy spectral parameters of FEBs by Tang et al. (2020), we investigate the characteristics of ECMI excited by evolving FEBs propagating along the flare loops. The structure of this paper is organized as follows: Section 2 introduces the density and magnetic field models of the flare loop. In Section 3, we present the calculating results of maser instability. Finally, Section 4 provides the discussion and conclusions.

2 The physical model

2.1 Density and magnetic field models of flare loops

It is generally believed that the acceleration of charged particles by magnetic reconnection in the impulsive phase may be the most

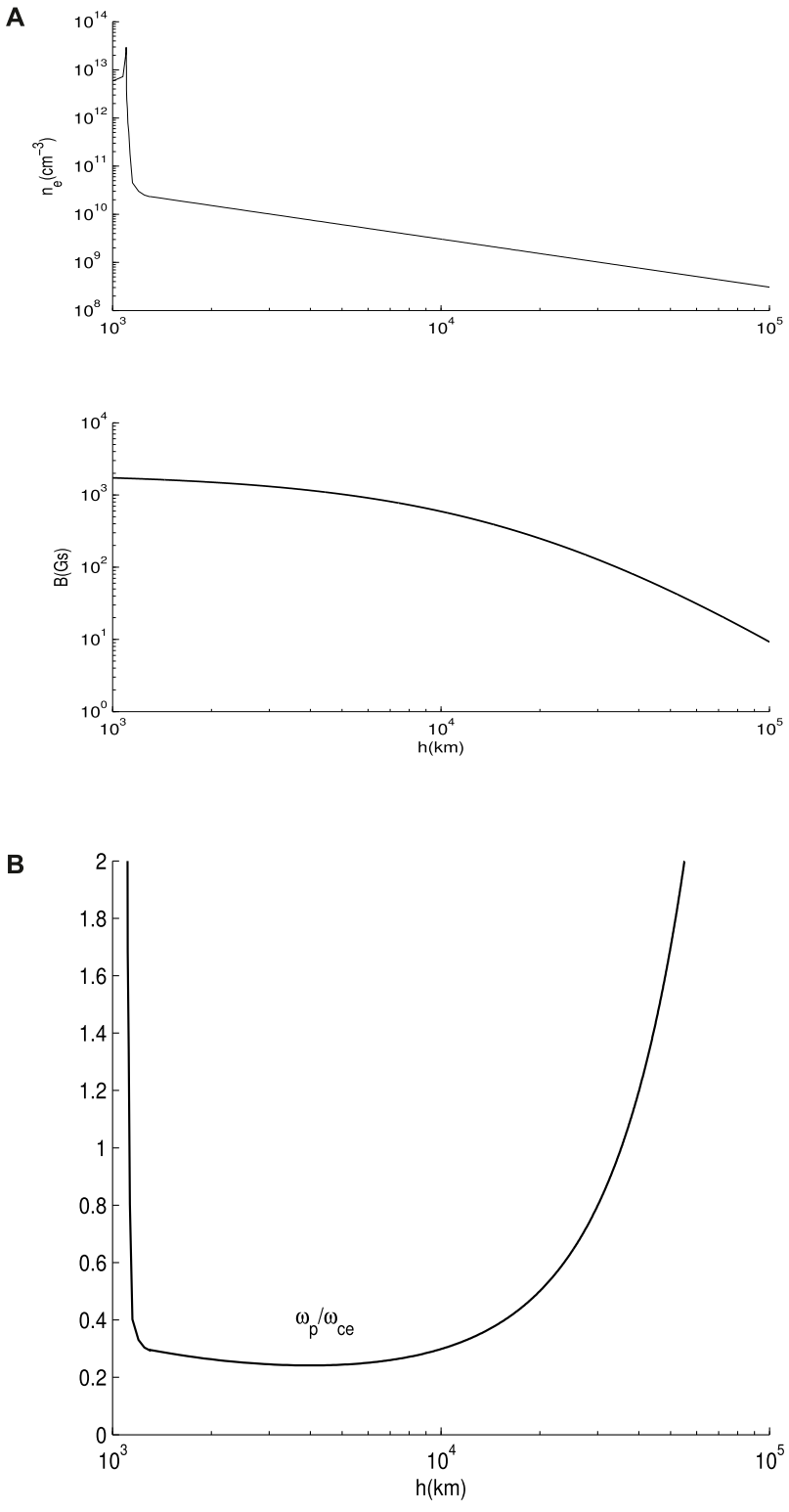

FIGURE 1
(A): Density and magnetic field strength in the flare loop vs. height. **(B)**: Radial distribution of the ratio of plasma frequency ω_p to electron cyclotron frequency ω_{ce} along the flare loop.

prominent feature of a solar flare. In this paper, we assume that the flare-accelerated FEBs have an initial power-law distribution with a lower energy cutoff, denoted as E_c, when they leave the acceleration region. Following the results of Tang et al. (2020), we analyze the influence of the evolving FEBs on ECMI when traveling in the flare loops. The height of the electron acceleration site,

approximately corresponding to the height of the flare loops, is about 20–50 Mm (Aschwanden et al., 1998; Reid et al., 2011). Therefore, Tang et al. (2020) utilized the semiempirical homogeneous plane-parallel flare atmospheric model (F2 model) (Machado et al., 1980) and extended the loop to a height of $h = 40$ Mm. Observations and studies show that flares tend to occur strongly in magnetic active regions. Hence, we approximately model the magnetic field of the flare loop and extend the density distribution from the transition region to the looptop using the solar active regions model proposed by Zhao (1995):

$$B(r,h) = \frac{\left[r^2 + 4(h+d)^2\right]^{1/2}}{2\left[r^2 + (h+d)^2\right]^2} d^3 B_0, \tag{1}$$

and

$$N(r,h) = 7.272 \times 10^{17} \mu(r)\, h^{-0.9978}. \tag{2}$$

In Eqs 1, 2, r represents the distance to the central axis of the active region, h denotes the height above the photosphere, and d indicates the depth of a vertical dipole sunspot below the photosphere. B_0 represents the magnetic field strength at the point on the axis of the sunspot and at the photospheric level, i.e., $(r = 0, h = 0)$.

Figure 1A illustrates the radial distribution of plasma density and magnetic field along the flare loops. The density curve n_e is plotted based on the flare atmospheric model (Machado et al., 1980) and coronal density model of active regions (Eq. (2)). Here, we set $\mu(r) = 4$, $r = 0$, $d = 2 \times 10^4$ km, and magnetic field $B_0 = 2000$ G, respectively. In Figure 1B, the radial distribution of the ratio of the electron plasma frequency (ω_p) to the cyclotron frequency (ω_{ce}) is presented. It is evident from Figure 1B that the plasma frequency ω_p is smaller than the electron cyclotron frequency ω_{ce} in the height range from the upper part of the transition region to near the top of the flare loop. This suggests that ECM is a viable emission mechanism in almost the entire flare loops. Morosan et al. (2016) and Régnier (2015) similarly reported that the condition for ECMI ($\omega_p/\omega_{ce} < 1$) is possible at heights $< 1.1R_\odot$ and $< 1.2R_\odot$ within the active region. Figure 2 displays the radial distribution of the magnetic mirror ratio parameter σ with the height of flare loops.

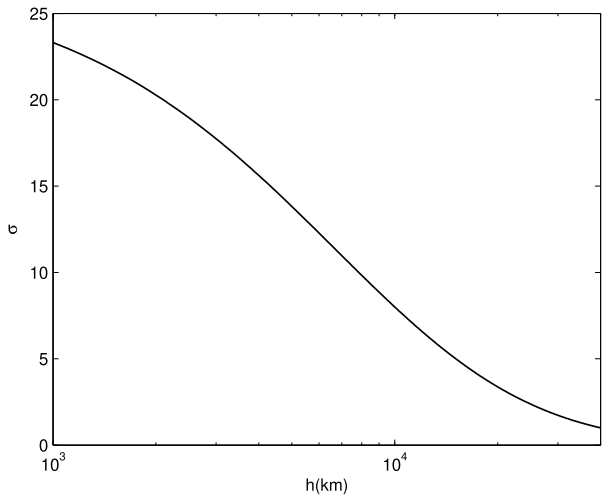

FIGURE 2
The magnetic mirror ratio σ vs. height of flare loop.

2.2 The evolving FEBs

As we know, FEBs propagating along the magnetic field in the solar atmosphere are believed to be the sources of solar radio bursts and HXR bursts. However, due to various reasons such as energy loss, diffusion caused by wave-particle interactions, etc., the distribution of FEBs is highly unstable. Due to energy loss during their long journey, the properties of FEBs, such as the energy spectral parameters, can be significantly modified (Tang et al., 2020). Melrose and Wheatland (2016) and Ning et al. (2021) proposed that a horseshoe distribution can be formed when beam electrons traveling inside flare loops. It is generally believed that when a parallel electric field exists, the acceleration effect of the parallel electric field causes electrons to escape, forming a horseshoe distribution. The corresponding relaxation includes the time for establishing the parallel electric field, electron acceleration, and then escape. In the case of AKR, it is indeed observed that there are parallel electric fields and horseshoe distributions in the auroral electron acceleration region. Here, we have not considered the issue of parallel electric fields at the moment, so we have also not considered horseshoe distributions. For loss cone distribution, since it directly forms from electron leakage at small injection angles without undergoing diffusion, acceleration, or other kinetic processes, there is no relaxation time. If we must define a relaxation time, it would be the escape time of quasi-parallel (small injection angle) electrons, which should be extremely short and can be considered instantaneous. Considering the initial beam-like velocity distribution of accelerated electron beams (Wang, 2004) and the power-law energy distribution, we employ the distribution function of FEBs when they leave the acceleration site as follows:

$$F_0(u,\mu) = A_0 \tanh\left(\frac{u}{u_c}\right)^{2\delta}\left(\frac{u}{u_c}\right)^{-2\alpha} \times \exp\left[-\frac{(u\mu - u_s)^2}{\beta^2} - \frac{u^2(1-\mu^2)}{\beta^2}\right]. \tag{3}$$

Here, $u^2 = u_\perp^2 + u_\parallel^2$, where u denotes the momentum per unit mass, $u_\parallel$ and $u_\perp$ represent the components of u parallel and perpendicular to the ambient magnetic field, respectively. $\mu = u_\parallel/u$ is the pitch angle cosine. A_0 is the normalization coefficient. The hyperbolic tangent function $\tanh(u/u_c)^{2\delta}$ describes the general cutoff behavior, with parameters δ and α representing the steepness index and spectrum index of FEBs, respectively. $E_c = \frac{1}{2}mu_c^2$ denotes the cutoff energy. u_s represents the beam velocity, and β is the momentum dispersion in parallel and perpendicular directions.

The results of Tang et al. (2020) show that as the FEBs precipitate in the flare loops, the steepness cutoff behavior will be flatten, the lower cutoff energy E_c decreases, and

the initial single power-law spectrum can evolve into a broken power-law spectrum. Taking into account the magnetic mirror effect caused by the convergence of magnetic fields in the lower atmosphere and disregarding the influence of the double power-law spectrum on the ECM, we can approximate the distribution function of FEBs at a certain height as follows:

$$F(u,\mu) = A_0 \tanh\left(\frac{u}{u_{ch}}\right)^{2\delta_h}\left(\frac{u}{u_{ch}}\right)^{-2\alpha_h} \times\left[1-\exp\left[\left(1-\sigma_h\right)\frac{1-\mu^2}{\mu^2}\right]\right] \times\exp\left[-\frac{\left(u\mu-u_s\right)^2}{\beta^2}-\frac{u^2\left(1-\mu^2\right)}{\beta^2}\right], \tag{4}$$

here, $E_{ch}=\frac{1}{2}mu_{ch}^2$, δ_h, α_h, and σ_h represent the cutoff energy, the steepness index, the spectrum index, and the magnetic mirror ratio parameter of FEBs at a certain height h.

3 Numerical results

3.1 General formulation of ECMI

ECMI is a coherent mechanism that directly amplifies electromagnetic radiation near the electron cyclotron frequency and its harmonic frequencies. Due to its high efficiency, ECMI has been extensively investigated and applied to various high-power radio bursts in magnetized plasma. Electromagnetic waves can be amplified when the resonance condition between energetic electrons and waves is met:

$$\gamma - s\Omega/\omega_q - N_q\mu\cos\theta u/c = 0. \tag{5}$$

Here, γ and s denote the Lorentz factor and harmonic number, respectively. Parameters Ω represent the electron cyclotron frequency, and ω_q is the frequency of the excited wave, with all frequencies normalized by the plasma frequency ω_p. N_q denotes the refractive index of the excited wave propagating with a phase angle θ to the magnetic field. The subscript $q=\pm$ indicates the wave modes for the ordinary mode (O mode, $q=+$) and the extraordinary mode (X mode, $q=-$), respectively.

We consider the density of FEBs to be much lower than that of the background plasma. Therefore, when discussing the dispersion relation of waves, the FEBs can be neglected. However, when discussing the growth rate, the effect of the FEBs predominates. So, the dispersion relation of the excited wave can be approximately by the cold-plasma theory (Chen et al., 2002; Wu et al., 2002):

$$N_q^2 = 1-\frac{1}{\omega_q\left(\omega_q+\tau_q\Omega\right)}, \tag{6}$$

and

$$\tau_q = -s_q + q\sqrt{s_q^2+\cos^2\theta}$$
$$s_q = \omega_q\Omega\sin^2\theta/2\left(\omega_q^2-1\right). \tag{7}$$

When the frequency of the excited wave $\omega\simeq s\Omega$, the temporal growth rate can be given by the following (Wu et al., 2002):

$$\frac{\omega_{qi}}{\omega_{ce}} = \frac{\pi}{2}\frac{n_b}{n_0}\iiint d^3\mathbf{u}\frac{\gamma\left(1-\mu^2\right)}{\Omega\omega_q\left(1+T_q^2\right)R_q}\delta\left(\gamma-\frac{s\Omega}{\omega_q}-\frac{N_q\mu u}{c}\cos\theta\right) \times\left\{\frac{\omega_q}{\Omega}\left[\gamma K_q\sin\theta+T_q\left(\gamma\cos\theta-\frac{N_q\mu u}{c}\right)\right]\times\frac{J_s\left(b_q\right)}{b_q}+J_s'\left(b_q\right)\right\}^2 \times\left[u\frac{\partial}{\partial u}+\left(\frac{N_q u\cos\theta}{c\gamma}-\mu\right)\frac{\partial}{\partial\mu}\right]F_b(u,\mu), \tag{8}$$

and

$$b_q = N_q\frac{\omega_q}{\Omega}\frac{u}{c}\sqrt{1-\mu^2}\sin\theta,$$
$$R_q = 1-\frac{\Omega\tau_q}{2\omega_q\left(\omega_q+\tau_q\Omega\right)^2}\left(1-\frac{qs_q}{\sqrt{s_q^2+\cos^2\theta}}\frac{\omega_q^2+1}{\omega_q^2-1}\right),$$
$$K_q = \frac{\Omega\sin\theta}{\left(\omega_q^2-1\right)\left(\omega_q+\tau_q\Omega\right)},$$
$$T_q = -\frac{\cos\theta}{\tau_q}. \tag{9}$$

Here, n_b and n_0 denote the number densities of energetic electrons and ambient thermal electrons, respectively. $J_s(b_q)$ represents the first kind of Bessel function.

3.2 ECMI by the evolving FEBs

As mentioned above, the energy spectra of FEBs can significantly vary due to their interaction with the ambient plasma as they travel down along the flare loop (Tang et al., 2020). With the evolving distribution function $F(u,\mu)$ given by Eq. 4, the growth rates of ECMI by FEBs when they reach a certain height can be calculated based on Eq. 8. For the given initial parameters (α, δ, u_c, u_s, β) of FEBs, the evolving energy spectral parameters, such as the cutoff energy E_{ch} (u_{ch}), the power-law spectral index α_h, and the steepness index δ_h at a certain height are based on the calculation results of Tang et al. (2020). The frequency ratio Ω and the magnetic mirror ratio σ are deduced from the density and magnetic field models of the flare loops. Figures 3–5 illustrate the effect of the parametric evolution of FEBs energy spectrum on ECMI when propagating from the loop top to a certain height. The growth rate depends on the frequency and propagation angle of the excited wave, i.e., on parameters ω_q and θ. The peak growth rates ω_i are normalized by $\omega_{ce}n_b/n_0$. The parameter n_b/n_0 represents the density ratio between energetic electrons and background electrons, and it actually varies as the FEBs move through the solar atmosphere. Different density ratios have an impact on both the dispersion relation and the growth rate of the excited wave (Winglee and Dulk, 1986; Yasnov and Karlický, 2004; Li et al., 2019; Ning et al., 2023). Here, the density ratio $n_b/n_0 \ll 1$. Therefore, the state of the background plasma does not directly affect the growth rate, but only indirectly affects it through the dispersion relation. Of course, for relatively

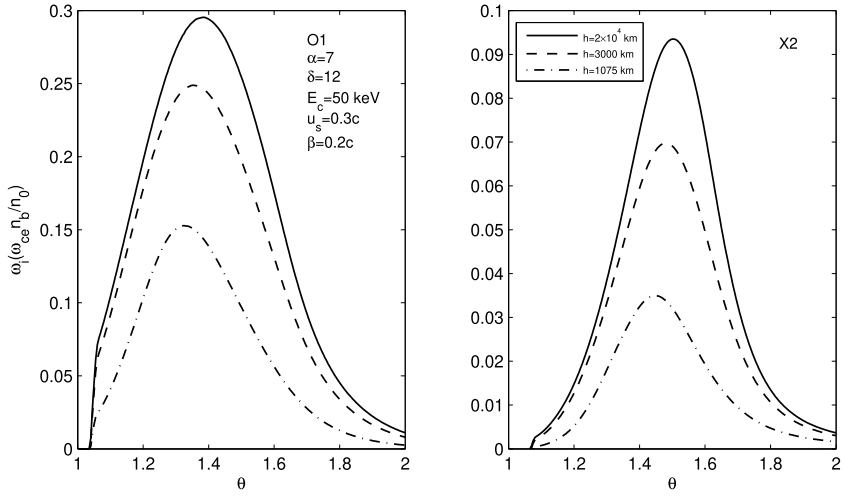

FIGURE 3
Peak growth rates of the fundamental waves for O mode and harmonic waves (s = 2) for X mode excited by the FEBs traveling down along the flare loop. The initial parameters of FEBs, such as spectrum index $\alpha = 7$, steepness index $\delta = 12$, $E_c = 50$ keV, and $u_s = 0.3c$, and $\beta = 0.2c$ have been used.

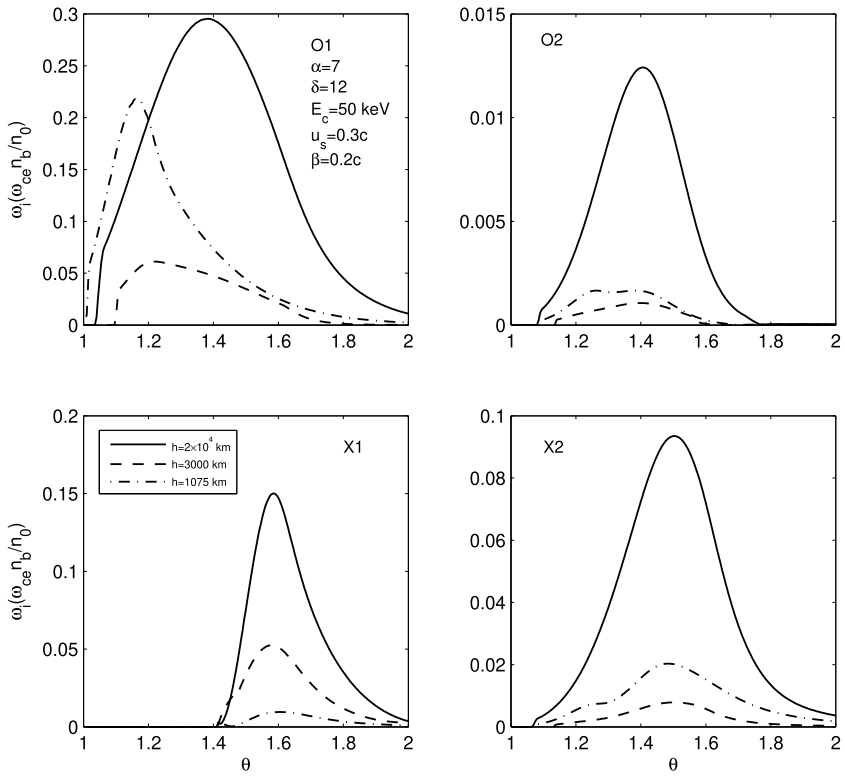

FIGURE 4
Peak growth rates of the fundamental waves and harmonic waves (s = 2) for O and X modes excited by the FEBs when traveling down along flare loop. Here, the frequency ratio Ω and the magnetic mirror ratio σ vary with the magnetic field and density models of the flare loop, other parameters are same as in Figure 3.

large ratios of beam/background electron densities, the state of background electrons may indeed have a direct impact on the calculation of the growth rate.

Figure 3 depict the peak growth rates calculated by varying the frequency ω_q for a given wave phase angle θ, where O1 and X2 are the fundamental waves (s = 1) for the O mode and

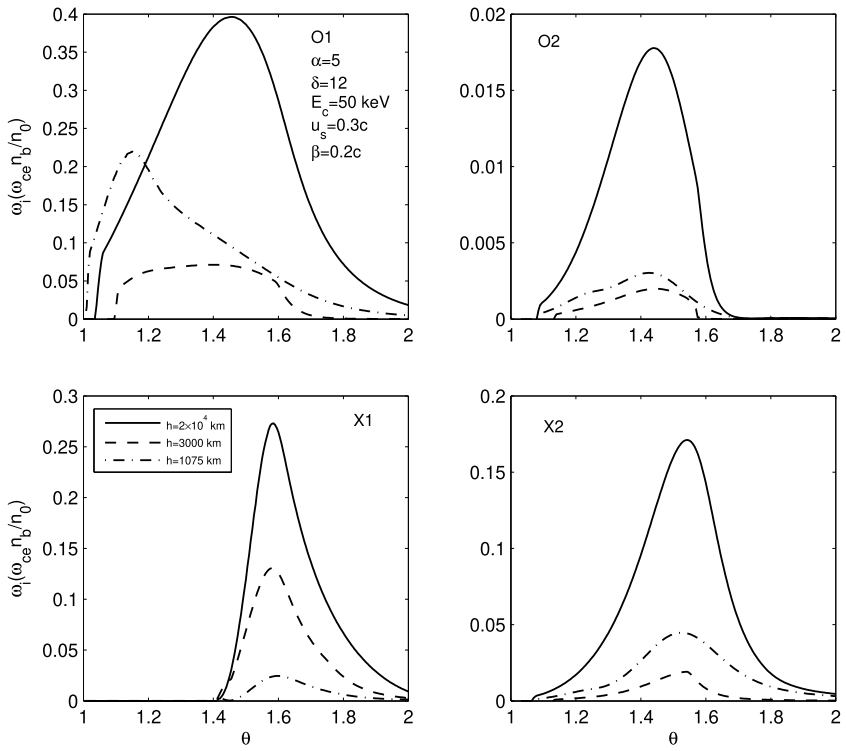

FIGURE 5
Peak growth rates of the fundamental waves and harmonic waves (s = 2) for O and X modes excited by the FEBs when traveling down along flare loop. Here, the frequency ratio Ω and the magnetic mirror ratio σ vary with the magnetic field and density models of the flare loop, the initial spectrum index $\alpha = 5$, other parameters are same as in Figure 3.

the harmonic waves (s = 2) for the X mode, respectively. In this figure, the frequency ratio Ω and the magnetic mirror ratio parameter σ are fixed, and we only consider the influence of the energy spectrum parameters on ECMI. The initial parameters of FEBs when they leave the acceleration site, such as the spectrum index $\alpha = 7$, steepness index $\delta = 12$, $E_c = 50$ keV, $u_s = 0.3c$, and $\beta = 0.2c$ have been used. The energetic electrons of FEBs are accelerated at the top of a flare loop with the height $h = 4 \times 10^4$ km. Here, the solid lines, dashed lines, and dot-dashed lines denote the peak growth rates of ECMI excited by the evolving FEBs when they reach a height of $h = 2 \times 10^4$ km, $h = 3000$ km, and $h = 1075$ km, respectively. According to the calculation results of Tang et al. (2020), we have the steepness index $\delta_h = 10.9$, spectral index $\alpha_h = 6.6$, cutoff energy $E_{ch} = 47.5$ keV for height $h = 3000$ km, and $\delta_h = 7.3$, $\alpha_h = 4.8$, $E_{ch} = 33$ keV for height $h = 1075$ km. Since the energy loss of energetic electrons is extremely small in the upper corona, we take the spectral parameters at $h = 2 \times 10^4$ km as the initial valves approximately. As seen from Figure 3, the peak growth rates of ECMI decrease rapidly as FEBs precipitate down along the flare loops. This indicates that the variations of the energy spectral parameters of FEBs do have a significant effect on the ECMI.

Figure 4 illustrates the peak growth rates of the fundamental waves and harmonic waves (s = 2) for the O and X modes by the evolving FEBs. In this Figure, except for the frequency ratio Ω and magnetic mirror ratio σ, the initial parameters of FEBs (α, δ, E_c, u_s and β) and the evolving energy spectral parameters (α_h, δ_h and E_{ch}) are the same as in Figure 3. Based on the magnetic field and density models of the flare loop as shown in Figure 1A, we can determine the frequency radio $\Omega = \omega_{ce}/\omega_p$ and magnetic mirror ratio σ at a certain height. The solid lines, dashed lines, and dot-dashed lines also represent the ECMI growth rates by the evolving FEBs at height $h = 2 \times 10^4$ km, $h = 3000$ km, and $h = 1075$ km, respectively.

The results from Figure 4 show that O1, O2, and X2 modes exhibit similar variations in growth rates. The peak growth rates of these three modes at height $h = 2 \times 10^4$ km are nearly an order of magnitude greater than those at $h = 3000$ km. However, the growth rates increase as FEBs precipitate downward into deeper height (such as $h = 1075$ km), especially for the O1 mode. This suggests that the energy spectral parameters of FEBs have a significant effect on ECMI. Moreover, compared with Figure 3, the growth rate increases at lower altitudes, indicating that the velocity anisotropy of FEBs, such as loss-cone distribution due to the convergence of magnetic fields, mainly affects the ECMI in the transition region. For mode X1, as the FEBs descend from the loop top to a height of $h = 1075$ km, the peak growth rate decreases monotonically. The peak growth rate at $h = 2 \times 10^4$ km is approximately three times higher than that at $h = 3000$ km. From Figure 4, it can be seen that the O1 mode has the highest growth rate, while the growth rate of the O2 mode is more than one order of magnitude lower than that of the O1 mode.

In Figures 5, 6, we also present the peak growth rates of the fundamental waves and harmonic waves (s = 2) for O and X

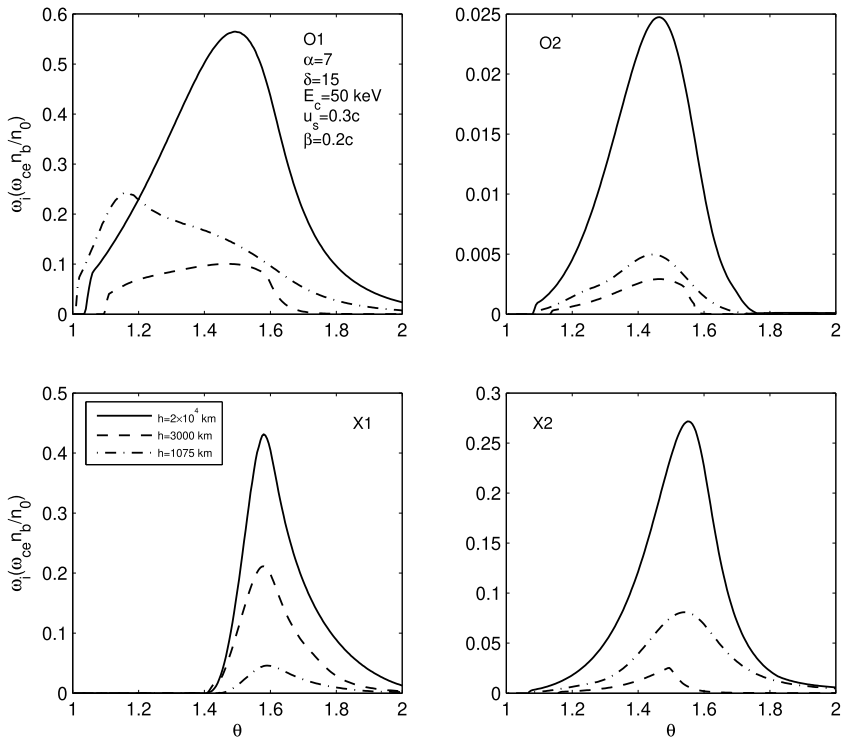

FIGURE 6
Peak growth rates of the fundamental waves and harmonic waves (s = 2) for O and X modes excited by the FEBs when traveling down along flare loop. Here, the frequency ratio Ω and the magnetic mirror ratio σ vary with the magnetic field and density models of the flare loop, the initial steepness index $\delta = 15$, other parameters are same as in Figure 3.

modes. For Figure 5, we have used the initial spectrum index $\alpha = 5$ and steepness index $\delta = 12$, while for Figure 6, we have used $\alpha = 7$ and $\delta = 15$. Other initial parameters of FEBs, such as $E_c = 50$ keV, $u_s = 0.3c$, and $\beta = 0.2c$, as well as the frequency ratio Ω and mirror ratio σ, remain the same as in Figure 4. The different curves, solid lines, dashed lines, and dot-dashed lines, also represent the peak growth rates of ECMI when FEBs reach heights of $h = 2 \times 10^4$ km, $h = 3000$ km, and $h = 1075$ km, respectively. The evolving energy spectral parameters $\alpha_h = 4.8$, $\delta_h = 11$, $E_{ch} = 47.5$ keV and $\alpha_h = 3.6$, $\delta_h = 7.1$, $E_{ch} = 33$ keV for height $h = 3000$ km and $h = 1075$ km in Figure 5, and $\alpha_h = 6.6$, $\delta_h = 13.5$, $E_{ch} = 47.5$ keV and $\alpha_h = 4.8$, $\delta_h = 8.9$, $E_{ch} = 33$ keV for $h = 3000$ km and $h = 1075$ km in Figure 6 have been used, respectively. Figures 5, 6 show that, as the FEBs descend, the characteristics of the peak growth rates change are similar to those in Figure 4, except for the specific numerical values of the growth rates. For the O1, O2, and X2 modes, the peak growth rates decrease initially in corona and then increase as FEBs travel down into denser atmosphere. And for the X1 mode, the peak growth rates decreases monotonically as FEBs precipitate along the loop. Similarly, the O1 mode consistently exhibits the highest growth rates, while the growth rates of O2 mode is more than one order of magnitude lower.

The calculations presented in Figures 3–6 demonstrate that the loss-cone velocity anisotropy of FEBs, resulting from the convergence of magnetic fields, has a more significant effect on the O1 mode than other three modes.

4 Summary and conclusion

Particle acceleration stands as one of the most crucial and widespread processes observed in various burst activities within space and cosmic plasma environments. Concerning the Sun, it is widely accepted that the energetic electrons are generated through the magnetic reconnection process during solar flare (Masuda et al., 1994; Yokoyama et al., 2001; Imada et al., 2013), or via shock acceleration near the corona shock wave (Drury, 1983; Blandford and Eichler, 1987; Park et al., 2012; Guo and Giacalone, 2015). These accelerated electrons propagate along magnetic fields as FEBs, which in turn produce solar radio bursts and HXR bursts in the solar atmosphere. One of the most direct pieces of evidence for FEBs is provided by solar radio type III bursts, which are excited as FEBs propagate upward along open magnetic field structures. As FEBs propagate downward along magnetic loop, both radio and hard X-ray emissions can be generated. The pivotal question lies in understanding how FEBs contribute to the generation of radio radiation.

ECMI represent a coherent radio radiation mechanism that directly amplify electromagnetic waves near the electron gyrofrequency and its harmonics within magnetized plasma. ECMI has been extensively studied and applied to various radio phenomena, including AKR from Earth and other magnetized planets (Wu and Lee, 1979; Zarka, 1998), various solar radio bursts (Melrose and Dulk, 1982; Wu et al., 2002; Yoon et al.,

2002; Treumann et al., 2011), and radio emission from flare stars (Stepanov et al., 2001), M dwarf stars (Callingham et al., 2021), and Blazar jets (Begelman et al., 2005).

Various energy loss processes, including collisions with ambient plasma and deceleration by induced electric fields, can lead to significant changes in the energy spectrum distribution of FEBs. Additionally, the spatial velocity anisotropic distributions of FEBs can arise due to variations of ambient magnetic field as they travel through the solar atmosphere. The lower energy cutoff behavior of power-law electrons and various velocity anisotropic distributions, such as loss-cone and temperature anisotropy, serve as effective driving sources for ECMI. Therefore, it is crucial to investigate ECMI produced by evolving FEBs.

In this paper, the beam electrons are accelerated near the looptop by the magnetic reconnection process during flares. The parametric evolution of power-law spectra of FEBs is based on the calculations of Tang et al. (2020). Taking into account the density and magnetic field models of active regions (Zhao, 1995) and the semiempirical homogeneous plane-parallel flare atmospheric model (Machado et al., 1980), we investigate the characteristics of ECMI by beam electrons traveling downward along a flare loop. Our results demonstrate that the evolution of the energy spectrum and velocity anisotropic distributions significantly influence ECMI as FEBs propagate in the flare loop.

When the ambient magnetic plasma parameters, frequency ratio Ω, and the magnetic mirror ratio σ are fixed, it is evident from Figure 3 that the growth rates of ECMI decrease as FEBs precipitate from the top of the flare loop to the footpoint. This decrease in growth rates is caused by the evolution of the lower energy cutoff behavior, which becomes increasingly flat as FEBs precipitate down along flare loops. Considering the magnetic field and density models of flare loops, the results of Figures 4–6 show that the growth rates of O1, O2, and X2 modes decrease initially as FEBs propagate in the corona and then increase as FEBs precipitate into the denser transition region. This suggests that the velocity anisotropy distribution of FEBs, caused by the convergence of magnetic fields, is an important driving source of ECMI in the denser transition region, and this new driver is more efficient for the O1 mode. For the X mode, it is strongly excited at the fundamental as Melrose and Dulk (1982) and (Treumann, 2006) pointed out. And the growth rate of X1 mode decreases continuously as the FEBs descend. It is absolutely clear that the O mode can easily leave the plasma as there is no restriction on its propagation. However, the X1 mode cannot escape directly. It may undergo resonance mode conversion or nonlinear wave-wave coupling to become an escapable radiation wave. Baumjohann and Treumann (2023) investigated the excitation of X mode, involving nonlinear wave-wave interaction and favored for its large expected growth rates. However, the issue of how the X1 mode can escape from the plasma is quite a different subject which deserves a completely separate investigation, and this is beyond the scope of the current article's main focus. It could be a subject for further separate discussion in the future. This evolutionary trend of the growth rates is similar to the results presented in Figures 2, 3 of Tang et al. (2016). Emission excited by such evolving FEBs in flare loops will form three separate radio sources. Additionally, our results indicate that the anisotropic driving source also affects the phase angles θ of the maximum growth rates of O1 mode, causing them to deviate more from the vertical direction. These findings contribute to a more comprehensive understanding of the dynamic spectra of solar radio bursts.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

JT: Conceptualization, Data curation, Formal Analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing–original draft. DW: Writing–review and editing. LC: Writing–review and editing. CT: Writing–review and editing. JW: Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work acknowledges the support from the National Natural Science Foundation of China (NSFC) under grants 12173076 and 42174195, from the Lishui University Initial Funding under grant QD2182, from the National Key R&D Program of China under grant 2021YFA1600503, as well as from the Strategic Priority Research Program of the Chinese Academy of Sciences under grant XDB0560000.

Acknowledgments

We also acknowledge the Project Supported by the Specialized Research Fund for State Key Laboratories.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aschwanden, M. J. (1990). The saturation of the electron-cyclotron maser instability and the interpretation of solar millisecond spikes. *Astronomy Astrophysics* 237, 512.

Aschwanden, M. J. (2002). Particle acceleration and kinematics in solar flares – a synthesis of recent observations and theoretical concepts (invited review). *Space Sci. Rev.* 101, 1–227. doi:10.1023/A:1019712124366

Aschwanden, M. J., and Benz, A. O. (1988). On the electron-cyclotron maser instability. II. Pulsations in the quasi-stationary state. *ApJ* 332, 466. doi:10.1086/166670

Aschwanden, M. J., Schwartz, R. A., and Dennis, B. R. (1998). Deconvolution of directly precipitating and trap-precipitating electrons in solar flare hard X-rays. II.*Compton gamma ray Observatory*Data Analysis. *ApJ* 502, 468–480. doi:10.1086/305891

Baumjohann, W., and Treumann, R. A. (2023). ECMI resonance in AKR revisited: hyperbolic resonance, harmonics, and wave–wave interaction. *Front. Phys.* 11, 1174557. doi:10.3389/fphy.2023.1174557

Begelman, M. C., Ergun, R. E., and Rees, M. J. (2005). *ApJ* 625, 51–59. doi:10.1086/429550

Benka, S. G., and Holman, G. D. (1994). A thermal/nonthermal model for solar hard X-ray bursts. *ApJ* 435, 469. doi:10.1086/174829

Blandford, R., and Eichler, D. (1987). Particle acceleration at astrophysical shocks: a theory of cosmic ray origin. *Phys. Rep.* 154, 1–75. doi:10.1016/0370-1573(87)90134-7

Cairns, I. H. (1987a). Fundamental plasma emission involving ion sound waves. *J. Plasma Phys.* 38, 169–178. doi:10.1017/S0022377800012496

Cairns, I. H. (1987b). Second harmonic plasma emission involving ion sound waves. *J. Plasma Phys.* 38, 179–198. doi:10.1017/S0022377800012502

Cairns, I. H. (1987c). Third and higher harmonic plasma emission due to Raman scattering. *J. Plasma Phys.* 38, 199–208. doi:10.1017/S0022377800012514

Callingham, J. R., Pope, B. J. S., Feinstein, A. D., Vedantham, H. K., Shimwell, T. W., Zarka, P., et al. (2021). Low-frequency monitoring of flare star binary CR Draconis: long-term electron-cyclotron maser emission. *A&A* 648, A13. doi:10.1051/0004-6361/202039144

Chen, Y. P., Zhou, G. C., Yoon, P. H., and Wu, C. S. (2002). A beam-maser instability: direct amplification of radiation. *Phys. Plasmas* 9, 2816–2821. doi:10.1063/1.1481748

Drury, L. O. (1983). An introduction to the theory of diffusive shock acceleration of energetic particles in tenuous plasmas. *Rep. Prog. Phys.* 46, 973–1027. doi:10.1088/0034-4885/46/8/002

Dulk, G. A., Kiplinger, A. L., and Winglee, R. M. (1992). Characteristics of hard X-ray spectra of impulsive solar flares. *ApJ* 389, 756. doi:10.1086/171248

Fleishman, G. D. (2004). Natural spectral bandwidth of electron cyclotron maser emission. *Astron. Lett.* 30, 603–614. doi:10.1134/1.1795949

Fletcher, L., Dennis, B. R., Hudson, H. S., Krucker, S., Phillips, K., Veronig, A., et al. (2011). An observational overview of solar flares. *SSRv* 159, 19–106. doi:10.1007/s11214-010-9701-8

Ginzburg, V. L., and Zheleznіakov, V. V. (1958). On the possible mechanisms of sporadic solar radio emission (radiation in an isotropic plasma). *Sov. Ast.* 2, 653.

Goldstein, M. L., and Eviatar, A. (1972). The plasma Physics of the jovian decameter radiation. *ApJ* 175, 275. doi:10.1086/151555

Guo, F., and Giacalone, J. (2015). The acceleration of electrons at collisionless shocks moving through A turbulent magnetic field. *ApJ* 802, 97. doi:10.1088/0004-637X/802/2/97

Hallinan, G., Antonova, A., Doyle, J. G., Bourke, S., Lane, C., and Golden, A. (2008). Confirmation of the electron cyclotron maser instability as the dominant source of radio emission from very low mass stars and Brown dwarfs. *ApJ* 684, 644–653. doi:10.1086/590360

Hess, S., Mottez, F., and Zarka, P. (2007). Jovian *S* burst generation by Alfvén waves. *J. Geophys. Res. (Space Phys.)* 112, A11212. doi:10.1029/2006JA012191

Hewitt, R. G., Melrose, D. B., and Ruennmark, K. G. (1981). A cyclotron theory for the beaming pattern of jupiter's decametric radio emission. *PASA* 4, 221–226. doi:10.1017/S132335800001643X

Hirshfield, J. L., and Bekefi, G. (1963). Decameter radiation from jupiter. *Nature* 198, 20–22. doi:10.1038/198020a0

Holman, G. D., Aschwanden, M. J., Aurass, H., Battaglia, M., Grigis, P. C., Kontar, E. P., et al. (2011). Implications of X-ray observations for electron acceleration and propagation in solar flares. *SSRv* 159, 107–166. doi:10.1007/s11214-010-9680-9

Hoyng, P., Duijveman, A., Machado, M. E., Rust, D. M., Svestka, Z., Boelee, A., et al. (1981). Origin and location of the hard X-ray emission in a two-ribbon flare. *ApJL* 246, L155. doi:10.1086/183574

Hudson, H., and Ryan, J. (1995). High-energy particles in solar flares. *ARA&A* 33, 239–282. doi:10.1146/annurev.aa.33.090195.001323

Hudson, H. S. (1978). A purely coronal hard X-ray event. *ApJ* 224, 235. doi:10.1086/156370

Imada, S., Aoki, K., Hara, H., Watanabe, T., Harra, L. K., and Shimizu, T. (2013). *ApJL* 776, L11. doi:10.1088/2041-8205/776/1/L11

Kosugi, T., Dennis, B. R., and Kai, K. (1988). Energetic electrons in impulsive and extended solar flares as deduced from flux correlations between hard X-rays and microwaves. *ApJ* 324, 1118. doi:10.1086/165967

Li, C., Chen, Y., Kong, X., Hosseinpour, M., and Wang, B. (2019). Effect of the temperature of background plasma and the energy of energetic electrons on Z-mode excitation. *ApJ* 880, 31. doi:10.3847/1538-4357/ab270f

Lin, R. P. (1974). Non-relativistic solar electrons. *SSRv* 16, 189. doi:10.1007/BF00240886

Lin, R. P. (2011). Energy release and particle acceleration in flares: summary and future prospects. *SSRv* 159, 421–445. doi:10.1007/s11214-011-9801-0

Machado, M. E., Avrett, E. H., Vernazza, J. E., and Noyes, R. W. (1980). Semiempirical models of chromospheric flare regions. *ApJ* 242, 336. doi:10.1086/158467

Mangeney, A., and Veltri, P. (1976). On the theory of type I solar radio bursts. I - beam plasma instabilities in a turbulent magnetized plasma. II - a model for the source. *A&A* 47, 165.

Masuda, S., Kosugi, T., Hara, H., Tsuneta, S., and Ogawara, Y. (1994). A loop-top hard X-ray source in a compact solar flare as evidence for magnetic reconnection. *Nature* 371, 495–497. doi:10.1038/371495a0

Melrose, D. B. (1976). An interpretation of Jupiter's decametric radiation and the terrestrial kilometric radiation as direct amplified gyroemission. *ApJ* 207, 651. doi:10.1086/154532

Melrose, D. B., and Dulk, G. A. (1982). Electron-cyclotron masers as the source of certain solar and stellar radio bursts. *ApJ* 259, 844. doi:10.1086/160219

Melrose, D. B., and Wheatland, M. S. (2016). *SoPh* 291, 3637–3658. doi:10.1007/s11207-016-1006-y

Morosan, D. E., Zucca, P., Bloomfield, D. S., and Gallagher, P. T. (2016). Conditions for electron-cyclotron maser emission in the solar corona. *A&A* 589, L8. doi:10.1051/0004-6361/201628392

Ning, H., Chen, Y., Li, C., Ye, S., Kuznetsov, A., and Wu, S. (2023). Excitation of extraordinary modes inside the source of Saturn's kilometric radiation. *A&A* 678, A94. doi:10.1051/0004-6361/202347149

Ning, H., Chen, Y., Ni, S., Li, C., Zhang, Z., Kong, X., et al. (2021). Harmonic electron-cyclotron maser emissions driven by energetic electrons of the horseshoe distribution with application to solar radio spikes. *A&A* 651, A118. doi:10.1051/0004-6361/202140427

Park, J., Workman, J. C., Blackman, E. G., Ren, C., and Siller, R. (2012). Particle-in-cell simulations of particle energization from low Mach number fast mode shocks. *Phys. Plasmas* 19, 062904. doi:10.1063/1.4729913

Raoult, A., Mangeney, A., and Vlahos, L. (1990). *A&A* 233, 229.

Régnier, S. (2015). A new approach to the maser emission in the solar corona. *A&A* 581, A9. doi:10.1051/0004-6361/201425346

Reid, H. A. S., Vilmer, N., and Kontar, E. P. (2011). Characteristics of the flare acceleration region derived from simultaneous hard X-ray and radio observations. *A&A* 529, A66. doi:10.1051/0004-6361/201016181

Schneider, J. (1959). Stimulated emission of radiation by relativistic electrons in a magnetic field. *PhRvL* 2, 504–505. doi:10.1103/PhysRevLett.2.504

Sharykin, I., Liu, S., and Fletcher, L. (2014). Onset of electron acceleration in A flare loop. *ApJ* 793, 25. doi:10.1088/0004-637X/793/1/25

Stepanov, A. V., Kliem, B., Zaitsev, V. V., Fürst, E., Jessner, A., Krüger, A., et al. (2001). Microwave plasma emission of a flare on AD Leo. *A&A* 374, 1072–1084. doi:10.1051/0004-6361:20010518

Tang, J., Wu, D., and Yan, Y. (2012). Effects of power-law spectrum behaviors of nonthermal electrons on a ring-beam maser instability. *Mech. Astronomy* 55, 744–750. doi:10.1007/s11433-012-4683-z

Tang, J. F., and Wu, D. J. (2009). Electron-cyclotron maser emission by power-law electrons in coronal loops. *A&A* 493, 623–628. doi:10.1051/0004-6361:200810792

Tang, J. F., Wu, D. J., Chen, L., Xu, L., and Tan, B. L. (2020). Parametric evolution of power-law energy spectra of flare accelerated electrons in the solar atmosphere. *ApJ* 904, 1. doi:10.3847/1538-4357/abc2ca

Tang, J. F., Wu, D. J., Chen, L., Zhao, G. Q., and Tan, C. M. (2016). Electron cyclotron maser emissions from evolving fast electron beams. *ApJ* 823, 8. doi:10.3847/0004-637X/823/1/8

Tang, J. F., Wu, D. J., and Yan, Y. H. (2011). Effects of the temperature anisotropy on the maser instability excited by lower energy cutoff behavior. *ApJ* 727, 70. doi:10.1088/0004-637X/727/2/70

Treumann, R. A. (2006). The electron–cyclotron maser for astrophysical application. *A&A Rv* 13, 229–315. doi:10.1007/s00159-006-0001-y

Treumann, R. A., Nakamura, R., and Baumjohann, W. (2011). A model of so-called Zebra" emissions in solar flare radio burst continua. *Ann. Geophys.* 29, 1673–1682. doi:10.5194/angeo-29-1673-2011

Twiss, R. Q. (1958). Radiation transfer and the possibility of negative absorption in radio Astronomy. *Aust. J. Phys.* 11, 564. doi:10.1071/PH580564

Twiss, R. Q., and Roberts, J. A. (1958). Electromagnetic radiation from electrons rotating in an ionized medium under the action of a uniform magnetic field. *Aust. J. Phys.* 11, 424. doi:10.1071/PH580424

Vlasov, V. G., Kuznetsov, A. A., and Altyntsev, A. T. (2002). The maser mechanism for solar millisecond spikegeneration in inhomogeneous plasma. *A&A* 382, 1061–1069. doi:10.1051/0004-6361:20011601

Wang, D.-y. (2004). A model for the coherent emission of solar radio moving type IV bursts. *Chin. Astronomy Astrophysics* 28, 404–411. doi:10.1016/j.chinastron.2004.09.004

White, S. M., Benz, A. O., Christe, S., Fárník, F., Kundu, M. R., Mann, G., et al. (2011). The relationship between solar radio and hard X-ray emission. *Space Sci. Rev.* 159, 225–261. doi:10.1007/s11214-010-9708-1

Winglee, R. M., and Dulk, G. A. (1986). The electron-cyclotron maser instability as a source of plasma radiation. *Astrophysical J.* 307, 808. doi:10.1086/164467

Wu, C. S., and Lee, L. C. (1979). A theory of the terrestrial kilometric radiation. *ApJ* 230, 621. doi:10.1086/157120

Wu, C. S., Wang, C. B., Yoon, P. H., Zheng, H. N., and Wang, S. (2002). Generation of type III solar radio bursts in the low corona by direct amplification. *ApJ* 575, 1094–1103. doi:10.1086/341468

Wu, C. S., Yoon, P. H., and Zhou, G. C. (1994). Generation of radiation in solar corona and interplanetary space by energetic electrons. *ApJ* 429, 406. doi:10.1086/174331

Wu, D. J., Huang, J., Tang, J. F., and Yan, Y. H. (2007). Solar microwave drifting spikes and solitary kinetic alfvén waves. *ApJL* 665, L171–L174. doi:10.1086/521360

Wu, D. J., and Tang, J. F. (2008). Effects of the lower energy cutoff behavior of power-law electrons on the electron-cyclotron maser instability. *ApJL* 677, L125–L128. doi:10.1086/588025

Yasnov, L. V., and Karlický, M. (2004). The Growth rate of upper-hybrid waves and dynamics of microwave zebra structures. *SoPh* 219, 289–299. doi:10.1023/B:SOLA.0000022942.17621.88

Yokoyama, T., Akita, K., Morimoto, T., Inoue, K., and Newmark, J. (2001). Clear evidence of reconnection inflow of a solar flare. *ApJL* 546, L69–L72. doi:10.1086/318053

Yoon, P. H. (1995). Plasma emission by a nonlinear beam instability. *Phys. Plasmas* 2, 537–548. doi:10.1063/1.870979

Yoon, P. H. (1997). Plasma emission by a nonlinear beam instability in a weakly magnetized plasma. *Phys. Plasmas* 4, 3863–3881. doi:10.1063/1.872508

Yoon, P. H. (1998). On the higher-order nonlinear corrections to the theory of plasma emission by a nonlinear beam instability. *Phys. Plasmas* 5, 2590–2595. doi:10.1063/1.872945

Yoon, P. H., Wu, C. S., and Wang, C. B. (2002). Generation of type III solar radio bursts in the low corona by direct amplification. II. Further numerical study. *ApJ* 576, 552–560. doi:10.1086/341634

Zarka, P. (1998). Auroral radio emissions at the outer planets: observations and theories. *J. Geophys. Res.* 103, 20159–20194. doi:10.1029/98JE01323

Zhao, G. Q., Chen, L., and Wu, D. J. (2013). Solar type iii radio bursts modulated by homochromous alfvén waves. *ApJ* 779, 31. doi:10.1088/0004-637X/779/1/31

Zhao, R.-Y. (1995). A model of solar (Radio) active regions. *Ap&SS* 234, 125–137. doi:10.1007/BF00627287

Zharkova, V. V., and Gordovskyy, M. (2004). Particle acceleration asymmetry in a reconnecting nonneutral current sheet. *ApJ* 604, 884–891. doi:10.1086/381966

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 17 June 2024
DOI 10.3389/fspas.2024.1407955

Check for updates

OPEN ACCESS

EDITED BY
Gang Li,
University of Alabama in Huntsville,
United States

REVIEWED BY
Xiaocheng Guo,
National Space Science Center, Chinese Academy of Sciences (CAS), China
Subhamoy Chatterjee,
Southwest Research Institute Boulder,
United States

*CORRESPONDENCE
Alexey Kuznetsov,
a_kuzn@iszf.irk.ru

RECEIVED 27 March 2024
ACCEPTED 28 May 2024
PUBLISHED 17 June 2024

CITATION
Kuznetsov A, Wu Z, Anfinogentov S, Su Y and Chen Y (2024), Electron acceleration and transport in the 2023-03-06 solar flare. *Front. Astron. Space Sci.* 11:1407955. doi: 10.3389/fspas.2024.1407955

COPYRIGHT
© 2024 Kuznetsov, Wu, Anfinogentov, Su and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Electron acceleration and transport in the 2023-03-06 solar flare

Alexey Kuznetsov[1,2]*, Zhao Wu[3,4], Sergey Anfinogentov[1], Yang Su[5] and Yao Chen[3,4,6]

[1]Department of Radioastrophysics, Institute of Solar-Terrestrial Physics, Irkutsk, Russia, [2]Department of Geography, Irkutsk State University, Irkutsk, Russia, [3]School of Space Science and Physics, Shandong University, Weihai, China, [4]Laboratory for Electromagnetic Detection, Institute of Space Sciences, Shandong University, Weihai, China, [5]Purple Mountain Observatory, Chinese Academy of Sciences (CAS), Nanjing, China, [6]Institute of Frontier and Interdisciplinary Science, Shandong University, Qingdao, China

We investigated in detail the M5.8 class solar flare that occurred on 2023-03-06. This flare was one of the first strong flares observed by the Siberian Radioheliograph in the microwave range and the Advanced Space-based Solar Observatory in the X-ray range. The flare consisted of two separate flaring events (a "thermal" and a "cooler" ones), and was associated with (and probably triggered by) a filament eruption. During the first part of the flare, the microwave emission was produced in an arcade of relatively short and low flaring loops. During the second part of the flare, the microwave emission was produced by energetic electrons trapped near the top of a large-scale flaring loop; the evolution of the trapped electrons was mostly affected by the Coulomb collisions. Using the available observations and the GX Simulator tool, we created a 3D model of the flare, and estimated the parameters of the energetic electrons in it.

KEYWORDS

solar flares, solar microwave emission, solar X-ray emission, particle acceleration, remote astrophysical diagnostics

1 Introduction

Solar flares are complicated phenomena that cover a broad range of heights in the solar atmosphere and produce electromagnetic emission in a broad range of wavelengths. The flares occur basically due to the sudden magnetic reconnection processes in the solar corona, which result in plasma heating, acceleration of charged particles, etc. (e.g., Benz and Güdel, 2010; Emslie et al., 2012).

To obtain a comprehensive picture of a flare, we need observations in different spectral ranges: e.g., the hard X-rays and white-light and ultraviolet (UV) continuum emissions are produced by non-thermal electrons mainly in the chromosphere at the footpoints of the coronal flaring loops; in the corona, the same electrons produce the microwave continuum emission due to the gyrosynchrotron mechanism; the soft X-rays and extreme ultraviolet (EUV) emission reflect the dynamics of the hot thermal plasma in the corona. Both the evolution of the spatially resolved images and the delays between the emissions at different wavelengths can reflect the dynamics of acceleration and transport of the non-thermal particles (e.g., Aschwanden, 2002). The recent commissioning of such solar-oriented astronomical instruments as the Siberian

Radioheliograph (SRH, Altyntsev et al., 2020) and the Hard X-Ray Imager on board the Advanced Space-based Solar Observatory (ASO-S/HXI, Su et al., 2019; Gan et al., 2023), which provide imaging spectroscopy observations in the microwave and hard X-ray ranges, respectively, offers new opportunities to study the solar flares.

In addition to the multiwavelength observations, understanding the nature of solar flares requires data-constrained modeling, which enables us to estimate the physical parameters in the flaring regions and to link the observed phenomena with the underlying processes of energy release and particle acceleration and transport. The recent advances in this field include, e.g., the case studies by Kuznetsov and Kontar (2015), Kuroda et al. (2018), Fleishman et al. (2018), Fleishman et al. (2021b), Fleishman et al. (2023), where the 3D structures of flares were reconstructed, and the spatial and energy distributions of energetic electrons and their dynamics were determined.

Here we investigate the GOES M5.8 class solar flare that occurred on 2023-03-06, at ~02:15–03:30 UT; it was one of the first strong flares observed by the SRH and ASO-S/HXI, as well as by other instruments. We present the results of observations and 3D modeling, and analyze the factors affecting the transport of non-thermal electrons.

2 Instruments and data

The microwave images of the flare were obtained using the Siberian Radioheliograph (SRH, Altyntsev et al., 2020). This instrument consists of three independent antenna arrays, two of which (for the frequency bands of 2.8–5.8 and 5.8–11.8 GHz) were operable at the considered date, thus providing imaging observations with the spatial resolutions of $15'' - 30''$and $12'' - 24''$, respectively. The observations were performed at 16 equidistant frequencies in each frequency band, i.e., at 32 frequencies in total, with the time resolution of ~3 s. The flux calibration was performed using the estimated microwave flux from the quiet-Sun regions (Zirin et al., 1991). To obtain a better alignment of the microwave images with magnetograms and images in other spectral ranges, we also performed simulations of the thermal gyroresonance emission from a non-flaring active region (AR 13245) just before the considered flare (at 02:10 UT) using the GX Simulator code (Nita et al., 2018; Nita et al., 2023; see also Fleishman et al., 2021a), and determined the position deviations between the observed and synthetic microwave images; the shifts needed to remove those deviations were then applied to all observed images throughout the flare. In addition to the imaging observations, we used the spatially unresolved measurements by the Nobeyama Radiopolarimeters (NoRP, Shimojo and Iwai, 2023), Palehua station of the Radio Solar Telescope Network (RSTN[1]), and Chashan Broadband Solar millimeter spectrometer (CBS, Shang et al., 2022; Shang et al., 2023) in the microwave range at a number of frequencies from 1 to 40 GHz.

The initial stage of the considered flare (until ~ 02:32 UT) was observed also by the Hard X-Ray Imager on board the Advanced Space-based Solar Observatory (ASO-S/HXI, Su et al., 2019; Gan et al., 2023). This instrument provides imaging spectroscopy observations of the solar X-ray emission in the energy range of ~10 – 400 keV with a spatial resolution down to about $3.2''$ at 30 keV. After ~ 02:32 UT, the X-ray data from the ASO-S/HXI became unreliable due to a strong parasite signal caused by the radiation belt particles. More continuous (but spatially unresolved) hard X-ray data were provided by the Konus-Wind spectrometer on board the Wind spacecraft (KW, Lysenko et al., 2022). In the considered event, this instrument operated in the waiting mode and recorded the X-ray count rates in the energy ranges of 19–78, 78–323, and 323–1,280 keV with the time resolution of ~3 s. The spatially unresolved soft X-ray data were provided by the Geostationary Operational Environmental Satellite (GOES[2]).

In addition to the above observations, we used the data from the instruments on board the Solar Dynamic Observatory: UV and EUV images from the Atmospheric Imaging Assembly (SDO/AIA, Lemen et al., 2012) and magnetograms from the Helioseismic and Magnetic Imager (SDO/HMI, Scherrer et al., 2012). All instruments used in this study are summarized in Table 1.

3 Observations

The considered GOES M5.8 class solar flare occurred on 2023-03-06 in the active region AR 13243 near the western solar limb, at N18W64. Figure 1 demonstrates the light curves of the flare at several selected microwave frequencies and X-ray energy ranges. The SRH light curves represent the microwave fluxes integrated over the $200'' \times 200''$ area centered at the flare. As has been said above, the reliable ASO-S/HXI data are available only before ~ 02:32 UT; no flare-related X-ray signal above 50 keV has been detected during that time interval. No flare-related X-ray signal in the KW 323–1,280 keV channel has been detected as well.

From the light curves, one can notice that the flare actually consisted of two separate (but closely related) flaring events, separated by the vertical thick dashed grey line in Figure 1. The first part of the flare (before ~02:33 UT) was mostly "thermal" (cf. Fleishman et al., 2015, and references therein): the X-ray spectrum was relatively soft, with no significant flux above ~50 keV, but relatively high fluxes at lower energies. The microwave emission demonstrated a good correlation with the hard X-rays above ~20 keV (which indicates its non-thermal origin), but was relatively weak. The GOES soft X-ray flux was sufficiently high, too, and demonstrated a noticeable delay with respect to the non-thermal emissions.

In the second, "cooler" part of the event (after ~ 02:33 UT), the X-ray spectrum became considerably harder, with the KW 78–323 keV flux considerably higher, but the KW 19–78 keV and GOES 1–8 Å fluxes lower than during the first part of the flare. The non-thermal microwave emission, too, reached much higher intensities than during the first part of the flare. One can distinguish the impulsive phase of the flare (~02:34–02:43), which was characterized by a prominent hard X-ray emission with multiple local peaks, corresponding likely to separate acts of magnetic reconnection.

1 https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-radio/rstn-1-second/

2 https://www.ngdc.noaa.gov/stp/satellite/goes-r.html

TABLE 1 Instruments used in this study.

Instrument	Spectral range	Spatial resolution	Time resolution
SRH	3–6 GHz (16 channels)	15″ – 30″	3.5 s
	6–12 GHz (16 channels)	12″ – 24″	3.5 s
RSTN	0.245, 0.41, 0.61, 1.415, 2.695, 4.995, 8.8, 15.4 GHz	—	1 s
NoRP	1.0, 2.0, 3.75, 9.4, 17.0, 34.0 GHz	—	1 s
CBS	35–40 GHz (10 channels)	—	0.54 s
SDO/HMI	Magnetograms	0.5″	12 min
SDO/AIA	1,600, 1,700 Å	0.6″	24 s
	94, 131, 171, 193, 211, 304, 335 Å	0.6″	12 s
GOES	1–8 Å	—	1 s
ASO-S/HXI*	10–400 keV	6.5″**	4 s/1 min**
Konus-Wind	19–1,280 keV (3 channels)	—	2.95 s

*The reliable ASO-S/HXI data for the considered flare were only available before ~02:32 UT; the data from other instruments were available for the entire duration of the flare.
**To produce the hard X-ray images of the considered flare (in the 20–40 keV range), the integration time of 1 min was used; 6.5″ is an effective spatial resolution of the reconstructed images in this study.

The microwave emission demonstrated firstly a similar dynamics with multiple peaks (especially at the frequencies of ~10 GHz) corresponding to the hard X-ray peaks; however, in contrast to the hard and soft X-rays, the microwave emission demonstrated also an overall increasing trend likely caused by a gradual accumulation of energetic particles in the flaring loop(s). The microwave emission reached a maximum at ~ 02:43 UT. After that, the hard X-ray emission dropped rapidly to the background level, and the microwave and soft X-ray emissions demonstrated a gradual decay that lasted for up to ~50 min.

A notable feature of the considered event was a filament eruption that occurred immediately before the flare. Figure 2 demonstrates a sequence of the SDO/AIA 335 Å EUV images of the flaring region. The eruption started at ~ 02:11 UT, i.e., well before the brightenings in the microwave and X-ray ranges. At 02:19:01 UT, when the first microwave and hard X-ray peak was observed, the filament had already risen up to a height of about 20,000 km. We have found no correlation between the filament parameters and the microwave and X-ray emissions. Therefore, although the filament eruption could trigger the magnetic reconnection and thus initiate the flare, at later stages (after the trigger) the evolutions of the flare and the filament likely diversified and became independent of each other.

3.1 Source structure and evolution, part I

Figure 3 demonstrates the images of the 2023-03-06 flare (during its first part) at several selected wavelengths, at three different times corresponding to the hard X-ray peaks, which are also representative of the flare structure and evolution. The ASO-S/HXI images in the 20–40 keV range were reconstructed by HXI_Clean with the preliminarily calibrated sub-collimator groups G3–G10, which generated a spatial resolution of ~6.5″. In the SDO/AIA 1600 Å UV images, one can identify two parallel flare ribbons. The hard X-ray emission, as observed by the ASO-S/HXI, initially (at ~02:18–02:20 UT, including the first emission peak) originated from an elongated region near the south-western edge of the flare ribbons, being likely produced in a flaring loop (or loops) connecting the ribbons. Then, at ~02:21–02:23 UT (i.e., including the major emission peak), the hard X-ray source extended noticeably to the north-east, forming an elongated structure that followed the flare ribbons. Finally, after ~02:23 UT, the hard X-ray brightening near the north-eastern edge of the flare ribbons disappeared, and the emission was again (until the end of the ASO-S/HXI observations) dominated by a relatively compact south-western source associated with the tops of the flaring loops visible in the EUV 335 and 131 Å channels; meanwhile, the total hard X-ray flux (above 20 keV) decreased with time more-or-less gradually, with a weaker peak at ~02:28 UT, as seen in Figure 1.

In the microwave range, at high frequencies (~11.80 GHz), there was a distinctive compact source located near the south-western edge of the flare ribbons, which barely changed its shape and position throughout the considered time interval; an additional weaker source appeared near the north-eastern edge of the flare ribbons at ~02:21–02:23 UT, i.e., simultaneously with a hard X-ray brightening at the same location. At lower frequencies (~5.60 GHz), the microwave source was more elongated; its peak firstly (at ~ 02:19 UT) nearly coincided with the 11.80 GHz peak, then (at ~02:21–02:23) shifted a bit to north-east, and finally (after ~02:25 UT) returned back to its initial position.

To explore the evolution of the microwave sources in more detail, we plotted the locations of the source peaks vs. time (see Figure 4); the peak locations were determined by fitting the microwave maps by an elliptical Gaussian. One can see from the figure that the source motions were rather complicated and frequency-dependent. At low frequencies (2.80–4.20 GHz), the source was firstly located close to the north-eastern edge of the flare ribbons, then shifted to south-west along the ribbons, and finally returned back to nearly the initial position. At higher frequencies (4.40–11.80 GHz), the picture was opposite: the source was firstly located near the south-western edge of the flare ribbons, then shifted to north-east along the ribbons, and finally returned back to nearly the initial position. At high frequencies (~11.80 GHz), the source displacement with time was relatively small, while at the middle frequencies (~5.60 GHz), the displacement was much larger and the source reached the middle of the flare ribbons. The maximum displacement of the microwave sources from their initial/final positions towards the middle of the flare ribbons occurred at around 02:23 UT, i.e., at the time when an additional hard X-ray brightening appeared at that location.

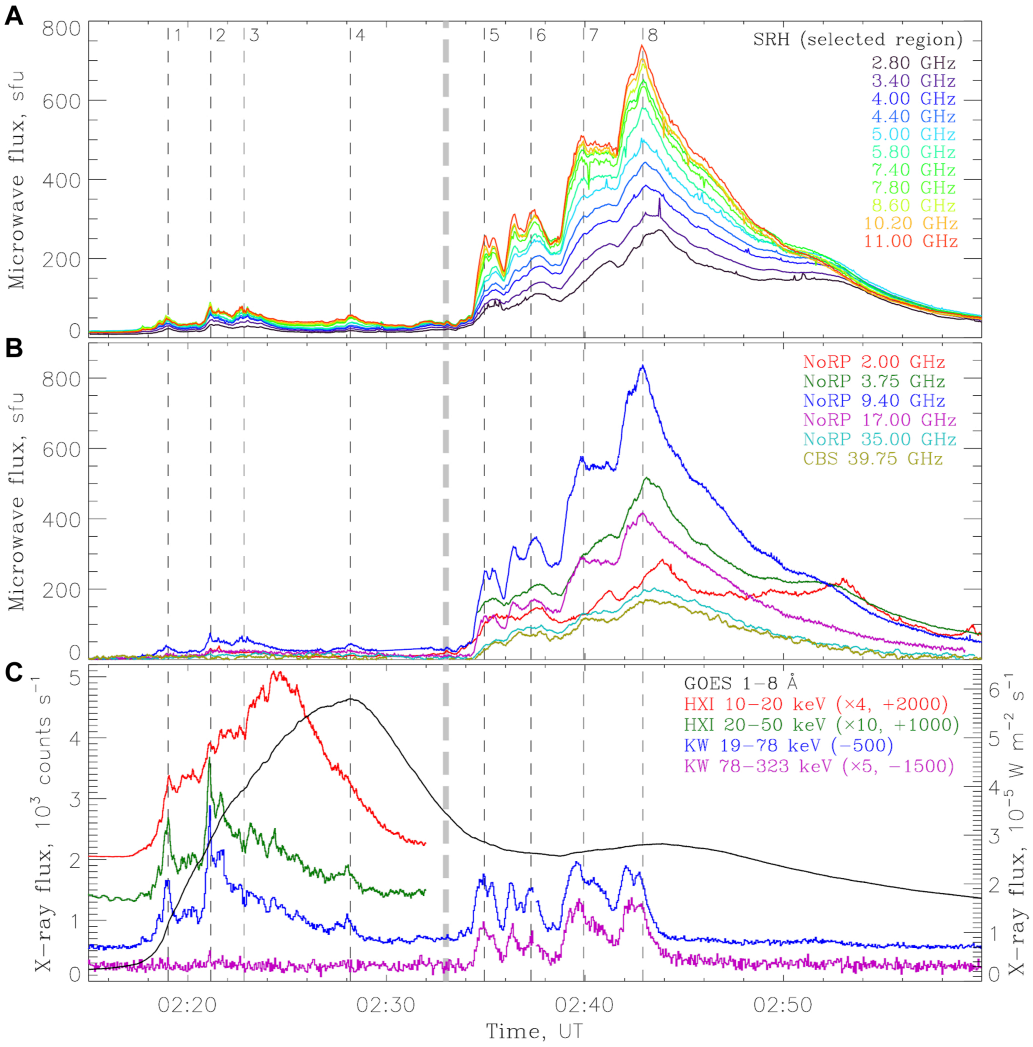

FIGURE 1
Light curves of the 2023-03-06 solar flare in the microwave **(A, B)** and X-ray **(C)** spectral ranges. The SRH light curves in the **(A)** represent the microwave fluxes from the flaring region; the NoRP and CBS light curves in the **(B)** are background-subtracted. In the **(C)**, the HXI and KW fluxes are in counts s^{-1}, while the GOES flux is in W $m^{-2}s^{-1}$.

Summarizing the presented observations, we conclude that during the first part of the 2023-03-06 flare, the microwave and hard X-ray emissions were likely produced in a sheared arcade of relatively short and low flaring loops connecting the flare ribbons; this arcade (at least, a part of it) can be seen, e.g., in the 335 Å EUV image at 02:28:11 UT in Figure 3. The hard X-ray emission was of non-thermal thin-target origin. The arcade was located below the erupted filament, and the magnetic reconnection in it was likely triggered by the eruption. The energy release and particle acceleration occurred along the entire arcade, but were not evenly distributed in space and time: the south-western part of the arcade usually dominated, but during a certain time interval (~02:21–02:23 UT) an intensive particle acceleration occurred near the middle of the arcade as well; the dynamics of the microwave and hard X-ray sources reflected the described dynamics of the energy release process.

3.2 Source structure and evolution, part II

Figure 5 demonstrates the images of the 2023-03-06 flare (during its second part, the impulsive phase) at several selected wavelengths, at three different times corresponding to the microwave emission peaks. Unfortunately, as has been said above, we have no imaging X-ray data for this time interval. The flare retained its two-ribbon structure (as seen in the 1,600 Å UV images), although the ribbons changed their configuration and expanded somewhat in the north-eastern direction in comparison with the first part of the flare. In the 131 Å EUV images, one can identify a loop-like structure that connected the flare ribbons, with the footpoints corresponding to the regions of the strongest magnetic fields of opposite polarities; this structure broadened gradually with time.

A similar loop-like structure is visible in the microwave images at high frequencies (~11.80 GHz): initially (at ~02:35 UT), the

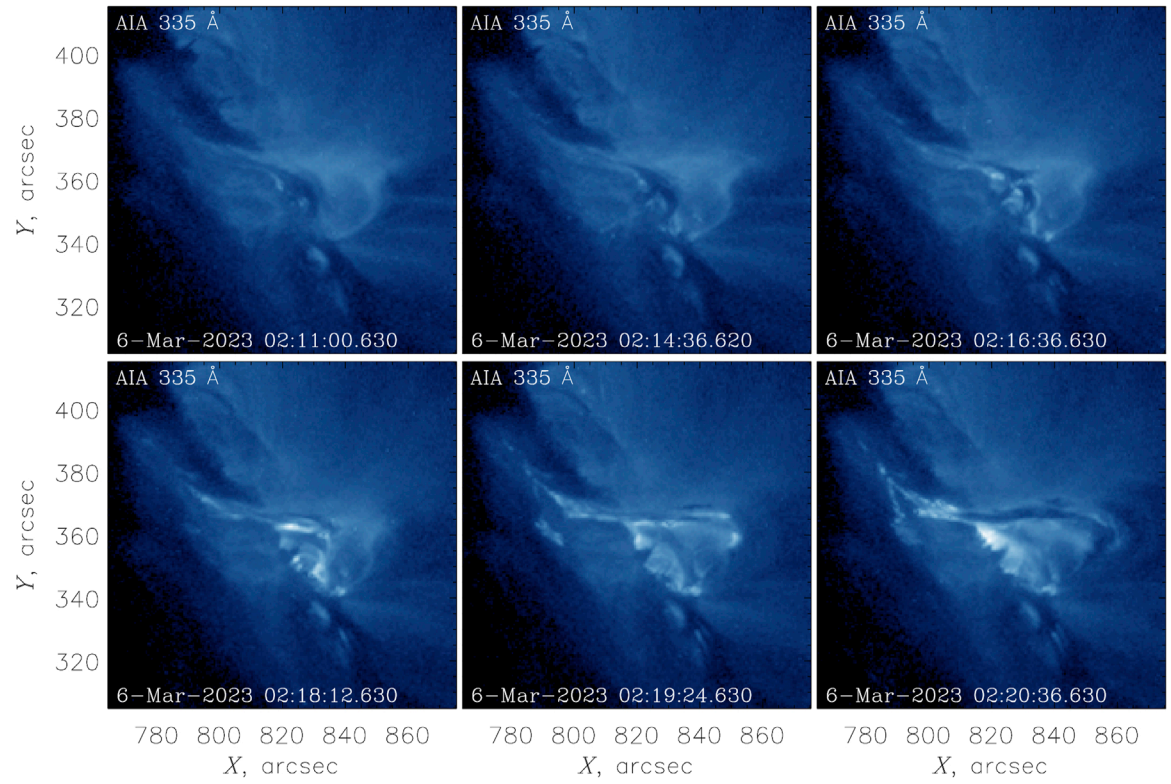

FIGURE 2
SDO/AIA 335 Å images of the flaring region at six representative times at the beginning of the flare, demonstrating the filament eruption.

north-eastern footpoint of that loop dominated; at later times (until ~02:43 UT), the south-western footpoint became gradually more pronounced, and the source peak shifted towards the loop top. At lower frequencies (~5.6 GHz and below), the microwave source demonstrated no definite structure, since its size was comparable with the SRH beam size; nevertheless, a gradual shift with time in the western direction can be noticed as well.

Figure 6 shows the motions of the microwave source peaks within the considered time interval. At all frequencies, the emission sources moved gradually along the loop visible in the 131 Å EUV images, from its north-eastern footpoint towards the loop top. This gradual motion demonstrated no visible correlation with the variations of the emission intensity (i.e., with the local peaks in the light curves, see Figure 1). A small departure of the 11.80 GHz source peaks in the southern direction during the time interval marked as 7–8 (02:39:56–02:42:55 UT) likely had an instrumental origin related to an insufficient spatial resolution, when in the presence of two nearby actual emission sources (near the loop top and at the south-western footpoint) the resulting observed source centroid was shifted towards the footpoint. At the flare decay phase (after 02:43 UT), the microwave source peaks at all frequencies were located at the loop top.

Summarizing the presented observations, we conclude that during the second part of the 2023-03-06 flare, the microwave emission likely originated from a single large-scale flaring loop (or a tightly packed bundle of such loops). This flaring loop was located above the loop arcade formed at the previous stage of the considered flare. The magnetic reconnection in this large-scale loop, again, could be triggered by the rising filament, although we cannot determine reliably the location of the reconnection site. Initially, the microwave emission was produced mainly in a strong magnetic field near the north-eastern footpoint; the subsequent evolution of the microwave emission sources reflected the process of gradual accumulation of energetic electrons within the loop (mainly near its top), which resulted in the respective shift of the dominant emission source towards the loop top (see also Section 4).

3.3 Particle dynamics

We now analyze the parameters and evolution of the energetic electrons in the considered event. During the first ("thermal") part of the flare, we have found no significant delays between the microwave and hard X-ray emissions (between the peaks in the light curves, see Figure 1), which indicates that the particle trapping and accumulation in the flaring loops were negligible. Also, the lack of reliable high-energy and high-frequency data (the fluxes in the KW 78–123 keV channel, NoRP 35 GHz channel, and CBS 35–40 GHz channels were too low) does not allow us to infer the parameters of the energetic electron spectrum during this time interval; we can only conclude that the spectrum was sufficiently soft.

Figure 7 demonstrates a zoomed-in fragment of the flare light curves at several selected microwave and hard X-ray channels for the second ("non-thermal") part of the 2023-03-06 flare (we consider

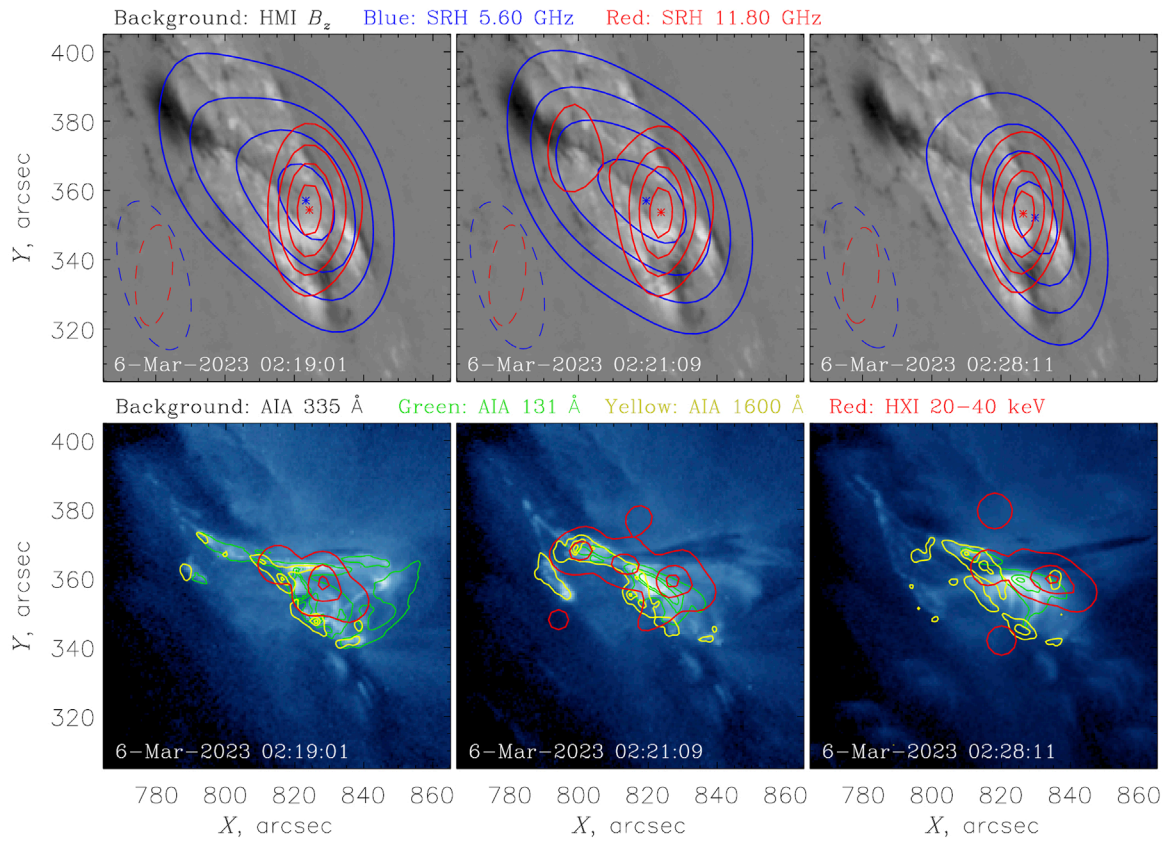

FIGURE 3
Images of the flaring region at three representative times during the first part of the flare. The top row demonstrates the SRH microwave intensity contours at 5.60 and 11.80 GHz (as solid lines, at 30, 50, 70, and 90% of the respective maximum intensities) overlaid on the SDO/HMI line-of-sight magnetograms; the dashed lines are the corresponding SRH beam contours at 1/2 level. The bottom row demonstrates the SDO/AIA EUV and UV contours at 131 and 1,600 Å, and the ASO-S/HXI X-ray contours in the 20–40 keV range (as solid lines, at 10, 50, and 90% of the respective maximum intensities) overlaid on the SDO/AIA 335 Å images.

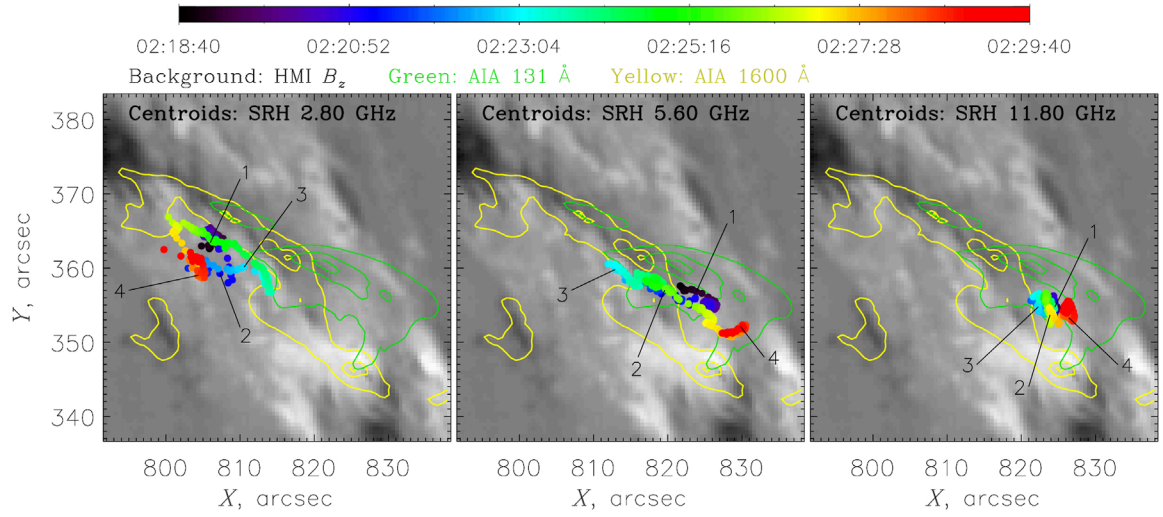

FIGURE 4
Motion of the microwave emission sources (at the frequencies of 2.80, 5.60, and 11.80 GHz) during the first part of the flare. The colored dots represent the locations of the smoothed maxima (centroids) of the microwave sources at different times; the numbers 1–4 mark the times indicated by vertical dashed lines in Figure 1. The centroid locations are overlaid on the SDO/HMI line-of-sight magnetogram; the solid lines show the SDO/AIA EUV and UV contours at 131 and 1,600 Å (at 10, 50, and 90% of the respective maximum intensities) at the representative time of 02:24:00 UT.

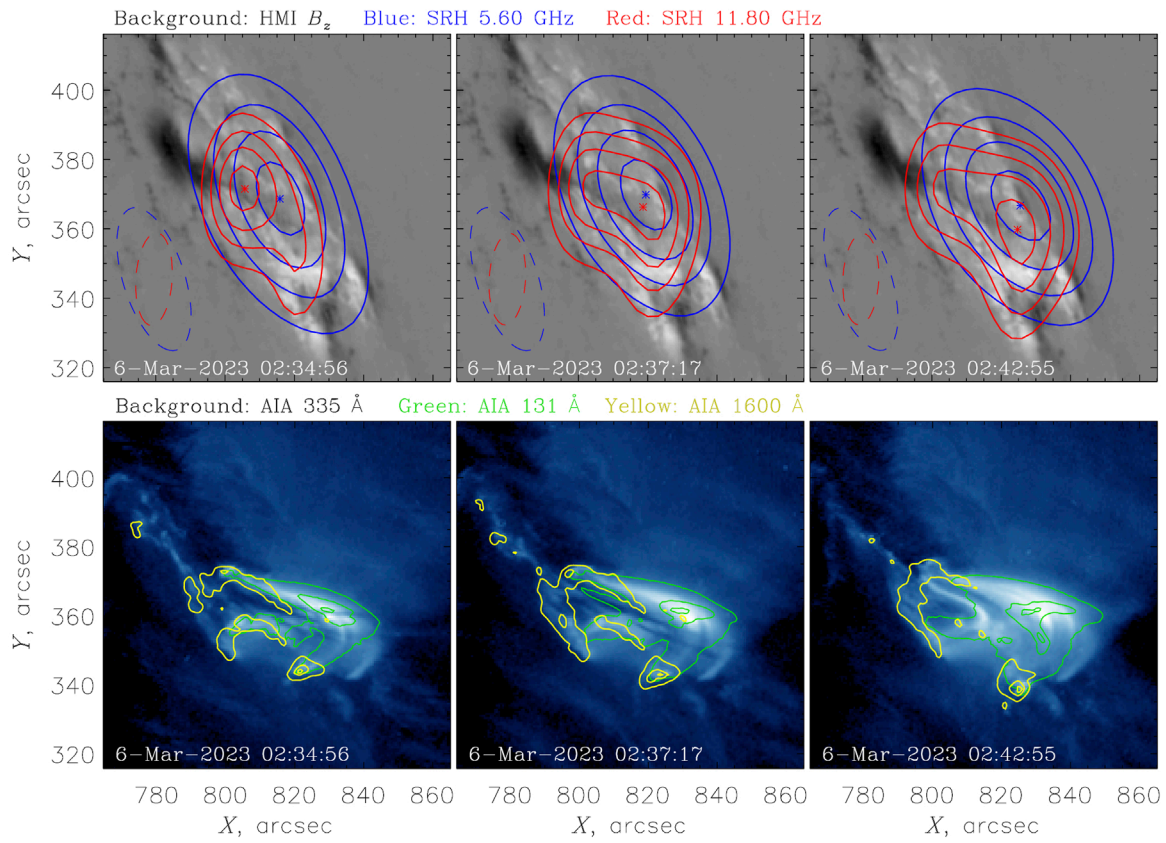

FIGURE 5
Images of the flaring region at three representative times during the second part of the flare. The top row demonstrates the SRH microwave intensity contours at 5.60 and 11.80 GHz (as solid lines, at 30, 50, 70, and 90% of the respective maximum intensities) overlaid on the SDO/HMI line-of-sight magnetograms; the dashed lines are the corresponding SRH beam contours at 1/2 level. The bottom row demonstrates the SDO/AIA EUV and UV contours at 131 and 1,600 Å (as solid lines, at 10, 50, and 90% of the respective maximum intensities) overlaid on the SDO/AIA 335 Å images.

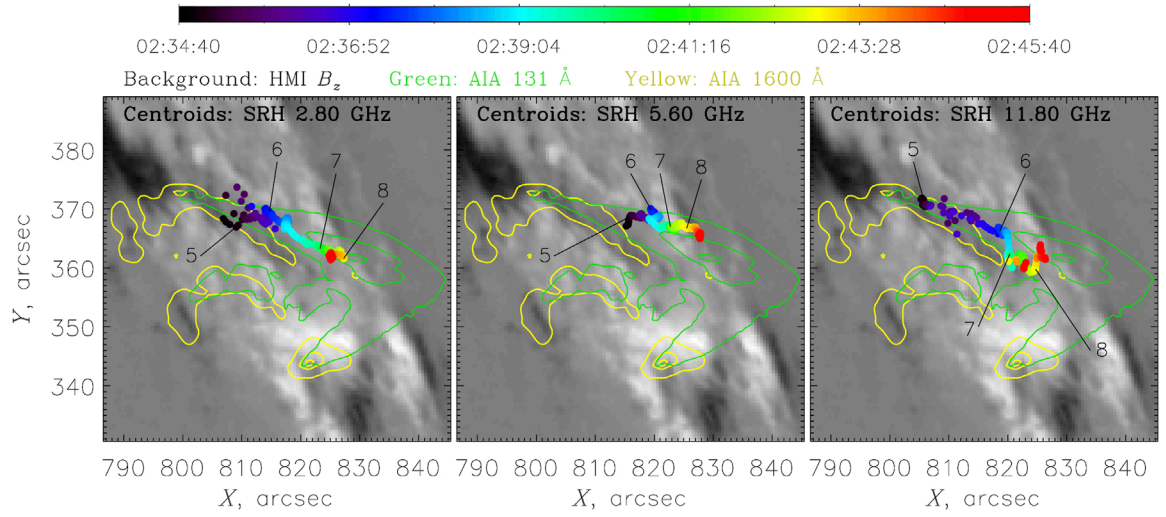

FIGURE 6
Motion of the microwave emission sources (at the frequencies of 2.80, 5.60, and 11.80 GHz) during the second part of the flare. The colored dots represent the locations of the smoothed maxima (centroids) of the microwave sources at different times; the numbers 5–8 mark the times indicated by vertical dashed lines in Figure 1. The centroid locations are overlaid on the SDO/HMI line-of-sight magnetogram; the solid lines show the SDO/AIA EUV and UV contours at 131 and 1,600 Å (at 10, 50, and 90% of the respective maximum intensities) at the representative time of 02:34:58 UT.

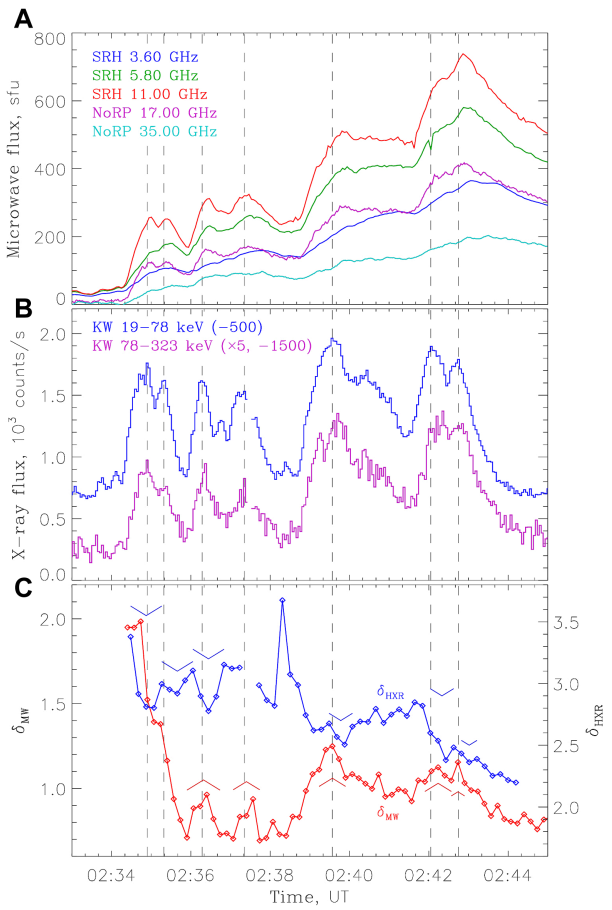

FIGURE 7
Zoomed-in light curves of the 2023–03–06 solar flare in the microwave **(A)** and hard X-ray **(B)** spectral ranges. The power-law spectral indices of the optically thin microwave emission (δ_{MW}, derived from the NoRP fluxes at 17 and 34 GHz) and the hard X-ray emission (δ_{HXR}, derived from the KW fluxes in the two presented energy ranges) are shown in the **(C)**, with the upward and downward angle brackets marking the time intervals with the hard-soft-hard and soft-hard-soft spectral evolution patterns, respectively. The vertical dashed lines correspond to the X-ray emission peaks at 19–78 keV.

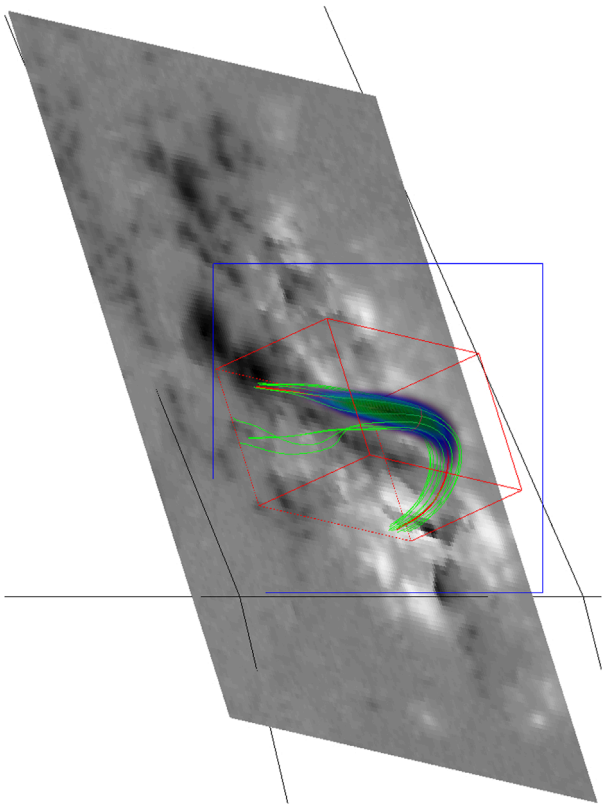

FIGURE 8
The model of the flaring region (screenshot from GX Simulator), corresponding to 02:34:56 UT: the selected coronal flux tube overlaid on the SDO/HMI magnetogram. The light green lines show the representative magnetic field lines bounding the flux tube, while the green-blue cloud shows the distribution of the energetic electrons.

here the impulsive phase only). The figure also shows the optically thin microwave spectral index δ_{MW}, defined as $I_{MW}(f) \propto f^{-\delta_{MW}}$, where I_{MW} is the microwave flux and f is the emission frequency, and the hard X-ray spectral index δ_{HXR}, defined as $I_{HXR}(E) \propto E^{-\delta_{HXR}}$, where I_{HXR} is the X-ray flux and E is the X-ray photon energy; the indices were derived respectively from the NoRP data at 17 and 35 GHz, and the KW data in the 19–78 and 78–323 keV channels. The spectral indices of the observed emissions are related to the spectral index of the emitting electrons δ as $\delta_{MW} \simeq 0.90\delta - 1.22$ for the optically thin gyrosynchrotron emission (Dulk and Marsh, 1982), and $\delta_{HXR} = \delta - 1$ for the thick-target bremsstrahlung X-ray emission (Brown, 1971).

One can see from the figure that the microwave emission was delayed with respect to the hard X-ray one, which represents a signature of the particle transport processes (including trapping). The delays were frequency-dependent and reached ~30 s at ~3–4 GHz and ~10 s at ~10 GHz and higher. The optically thin microwave spectral index δ_{MW} demonstrated a correlation with the hard X-ray light curves: the spectral index increased (softened) during the hard X-ray pulses (i.e., when the energetic particles injection occurred), and then gradually decreased (hardened) in the absence of the injection; i.e., around each microwave and hard X-ray emission peak, the microwave emission and hence the energetic electrons producing the emission demonstrated the "hard-soft-hard" pattern (cf. Ning, 2008; Huang and Nakajima, 2009; Yan et al., 2023; Wu et al., 2024, etc.).

From the hard X-ray light curves, one can notice that most of the emission peaks at higher energies (78–323 keV) were slightly delayed with respects to the peaks at lower energies (19–78 keV); we have estimated the delays as $\lesssim$ 5–7s. As a result, the hard X-ray spectral index δ_{HXR} decreased (hardened) slightly during each emission peak, and then increased (softened) again; i.e., the hard X-ray emission and hence the electrons producing the emission demonstrated the "soft-hard-soft" pattern. In addition to those rapid variations, the hard X-ray emission demonstrated an overall hardening trend throughout the impulsive phase of the flare. We also note that the spectral index of the energetic electrons δ derived from the microwave observations (~2.1 – 2.7) was systematically lower (i.e., harder) than the same index derived from the hard X-ray observations (~3.3 – 4.1); this difference is typical of solar flares

TABLE 2 Parameters of the energetic electron distributions in the GX Simulator models used to simulate the microwave emission of the 2023-03-06 solar flare at different times: characteristic spatial scales in the directions across (σ_{r0}) and along (σ_s) the magnetic field, shifts relative to the loop top (s_0), maximum densities (n_{b0}), spectral indices (δ), and total numbers of the energetic electrons within the flaring loop in the 1–10 MeV energy range ($N_{1-10MeV}$).

Time, UT	σ_{r0}, km	σ_s, km	s_0, km	n_{b0}, cm^{-3}	δ	$N_{1-10MeV}$
02:34:56	1,285	8,145	3,710	3.30×10^{7}	2.90	8.76×10^{29}
02:37:17	1,800	9,860	2,385	1.55×10^{6}	2.30	1.42×10^{30}
02:39:56	2,315	10,125	1,060	5.80×10^{6}	2.60	2.33×10^{30}
02:42:55	2,575	10,700	0	3.40×10^{6}	2.45	3.46×10^{30}

(e.g., White et al., 2011) and reflects the fact that the microwave and hard X-ray emissions are produced respectively by the trapped electrons in the solar corona and by the precipitating electrons in the chromosphere and/or transition region. Other physical implications of the above-described features are discussed in Section 5.

4 Modeling

To model the microwave emission of the considered flare, we used the GX Simulator code (Nita et al., 2015; 2023). This code allows one to create a 3D magnetic field model of the flaring region using the nonlinear force-free field extrapolation, to select a flaring loop (or loops), to fill the flaring loop(s) with thermal and non-thermal electrons, and to compute the corresponding gyrosynchrotron and free-free microwave emission using the "fast gyrosynchrotron codes" by Fleishman and Kuznetsov (2010); Kuznetsov and Fleishman (2021). For comparison with the observations, the computed microwave emission maps were then convolved with the SRH beam.

As has been said above, during the first part of the flare, the emission was likely produced in an arcade consisting of multiple flaring loops. A model of such a structure would have too many free parameters. In addition, the lack of microwave data in the optically thin frequency range does not allow us to constrain reliably the spectrum of the emitting electrons; therefore, we do not consider that time interval here. In contrast, during the second part of the flare, the observed structure of the emission sources could be described reasonably well by a single-loop model. Based on the available images in the microwave, UV, and EUV ranges, we selected the flaring loop shown in Figure 8 that provided the best agreement with the observations. The loop had the length of 53,000 km and rose up to the height of 17,000 km; the magnetic field strength (at the loop axis) varied from 170 G at the loop top up to 1,600 and 1,380 G in the north-eastern and south-western footpoints, respectively. By analogy with a number of previous simulations (e.g., Kuznetsov and Kontar, 2015; Kuroda et al., 2018; Fleishman et al., 2021b; 2023; Wu et al., 2024), in order to reduce the number of free parameters, the magnetic structure of the loop (determined by the selected axial magnetic field line) was assumed to be the same at all times throughout the impulsive phase of the flare, and only the parameters of the energetic electrons varied.

For the energetic electrons, we adopted a single power-law energy distribution function in the form of $f(E) \propto E^{-\delta}$, with the electron energy E in the range from 0.01 to 10 MeV, and the electron number density equal to n_b; the pitch-angle distribution was assumed to be isotropic. The spatial distribution of the energetic electrons within the flaring region was described by the model function in the form of

$$n_b(r,s) = n_{b0} \exp\left[-\frac{r^2}{2\sigma_r^2(s)}\right] \exp\left[-\frac{(s-s_0)^2}{2\sigma_s^2}\right],$$

where s and r are the coordinates along and across the selected flaring loop, respectively, with the coordinate s measured relative to the loop top and positive in the direction towards the north-eastern footpoint, and the coordinate r measured relative to the loop axis; σ_s and σ_r are the characteristic scales of the distribution in the respective directions, and n_{b0} is the peak electron number density. Following the magnetic flux conservation, the transverse scale σ_{r0} varied along the loop as $\sigma_r(s)/\sigma_{r0} = \sqrt{B_0/B(s)}$, where B_0 and σ_{r0} are the magnetic field strength and the transverse scale σ_r at the loop top, and $B(s)$ is a local magnetic field strength.

We note that the above model is oversimplified and accounts for only the basic characteristics of the energetic electrons in the flaring region. Therefore, our aim was to reproduce: a) the total (spatially integrated) microwave emission spectra of the flare, primarily in the optically thin frequency range (above ~10 GHz), and b) the 2D locations of the microwave source peaks, as well as the microwave brightness distributions along the flaring loop at high frequencies (namely, at 11.80 GHz). The model parameters that provided the best agreement with the observations at four different times (corresponding to the microwave emission peaks) are presented in Table 2, while the corresponding synthetic and observed images and spectra are shown in Figure 9.

From Table 2 and Figure 9, one can notice that the energetic electrons in the considered event (during the second part of the flare) were likely concentrated near the top of the flaring loop. Even at 02:34:56 UT, the estimated displacement of the electron distribution from the loop top s_0 was relatively small, while the observed emission was concentrated near a footpoint due to a stronger magnetic field there. With time, the electron distribution peak approached gradually the loop top (s_0 decreased). The energetic electron distribution along the flaring loop broadened gradually with time, and the effective thickness of the loop increased as well (both σ_s and σ_r increased). The spectral index of the energetic electrons δ varied with time, in agreement with the estimations based on the observed microwave spectral index δ_{MW} (see Figure 7)

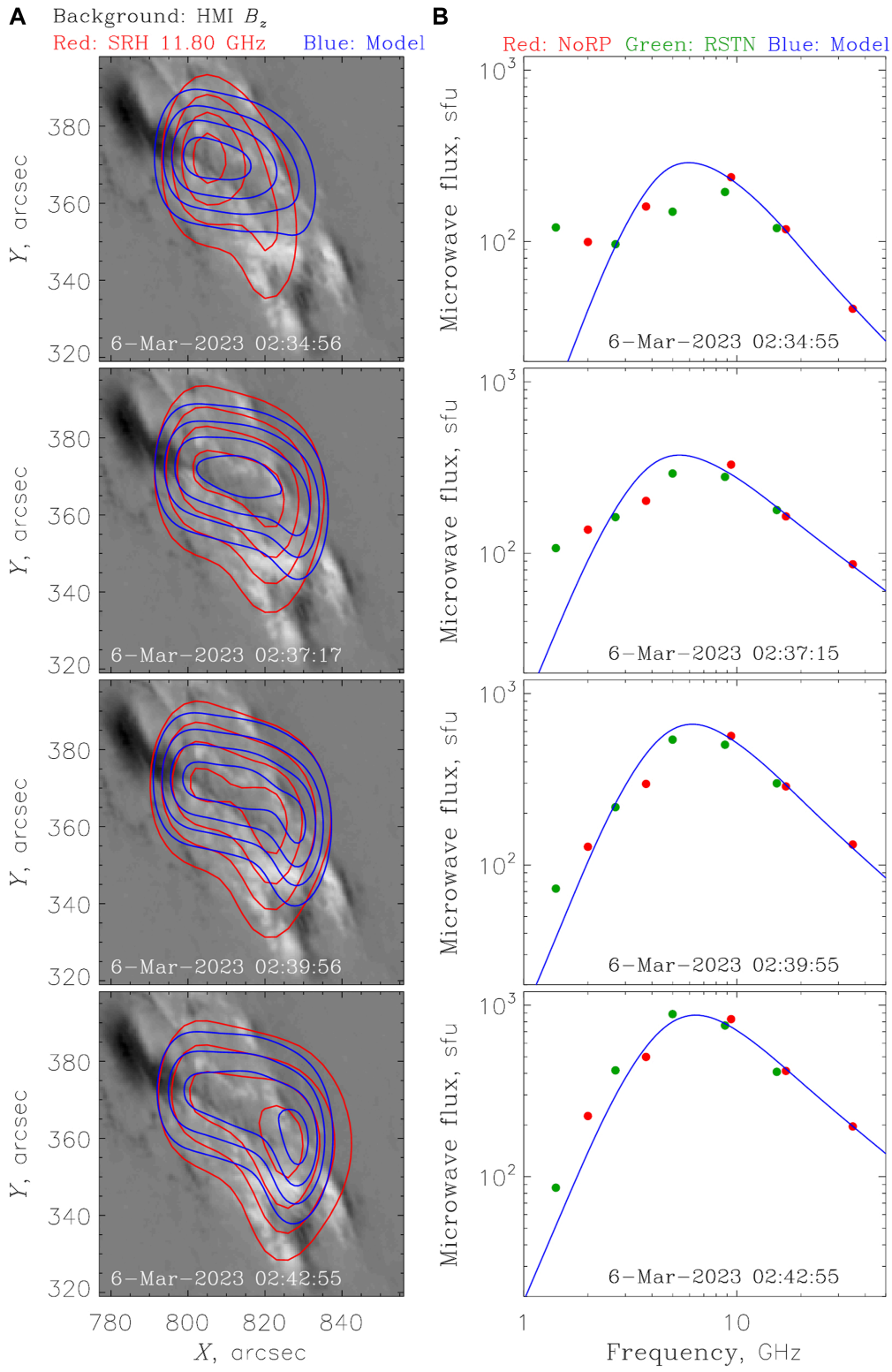

FIGURE 9
Comparison of the observed and simulated microwave emission parameters at four different times. **(A)** Observed and simulated microwave intensity contours at 11.80 GHz (at 30, 50, 70, and 90% of the respective maximum intensities) overlaid on the SDO/HMI line-of-sight magnetograms. **(B)** Observed (by NoRP and RSTN) and simulated total emission spectra; the error bars of the observations are smaller than or comparable to the symbol size.

and the empirical formula by Dulk and Marsh (1982). The parameter n_{b0} is not representative, because it depends strongly on the low-energy cutoff of the electron distribution, which was not reliably known and was chosen arbitrarily. A more reliable characteristic is the total number of energetic electrons at high energies, say, above 1 MeV. According to Table 2, this number ($N_{1-10\mathrm{MeV}}$) increased gradually with time during the impulsive phase of the flare, which reflected the process of accumulation of the energetic electrons in the flaring loop.

5 Discussion

The presented observations offer some clues into the non-thermal particle transport processes in the considered flare. During the first part of the flare, we have found no significant delays between the microwave and hard X-ray emissions, which indicates that the energetic electrons injected into the flaring loops then precipitated into the chromosphere nearly immediately, without experiencing a noticeable trapping. Such a behaviour is quite consistent with the above-described scenario suggesting that during the first part of the flare, the energy release processes occurred in an arcade of relatively short flaring loops with low mirror ratios.

On the other hand, during the second part of the flare, the microwave emission peaks were delayed with respect to the X-ray ones by up to ~10 – 30 s. These delays were much longer than any expected time-of-flight delays, and could be naturally attributed to the particle trapping and accumulation processes, given that the energy release likely occurred in a relatively long flaring loop with the mirror ratio of up to ~10. The evolution of the trapped energetic particles in coronal magnetic tubes is governed primarily by their pitch-angle scattering due to Coulomb collisions or/and interaction with a magnetohydrodynamic turbulence, when the particles scattered into the loss cone escape from the trap (are precipitated in the chromosphere). Thus the observed hardening of the microwave emission spectrum (and hence of the energy spectrum of the trapped energetic electrons in the flaring loop) with time in the absence of the particle injection, see Figure 7, indicates that the lower-energy electrons were scattered more efficiently in the considered event, which favours the Coulomb collisions as the dominant scattering factor. In turn, the intermittent acts of energy release (highlighted by the hard X-ray emission pulses) resulted in injection of additional portions of electrons with a softer energy spectrum into the flaring loop, and this interplay between the particle injection and escape processes formed the observed hard-soft-hard pattern in the optically thin spectral index of the microwave emission. At the same time, the total number of the high-energy electrons increased more-or-less steadily with time until the end of the impulsive phase of the flare.

The observed delays between the hard X-ray emission peaks at higher and lower energies (see Figure 7) are consistent with the above scenario, because the hard X-ray emission is produced by the precipitating electrons that escaped from the magnetic trap. As has been said above, in the considered event, the lower-energy trapped electrons were scattered into the loss cone more efficiently (due to the Coulomb collisions), and therefore escaped from the trap and reached the loop footpoints first, then followed by the higher-energy electrons. This explanation implies that the energetic electrons were injected somewhere in the coronal part of the flaring loop, with an isotropic pitch-angle distribution or preferably in the direction across the local magnetic field, so that a significant fraction of the particles became trapped immediately after the injection. The suitable particle acceleration mechanisms include, e.g., stochastic acceleration or acceleration in a collapsing magnetic trap (e.g., Zharkova et al., 2011).

The delay between the hard X-ray emission pulses at the energies of E_2 and E_1 can be estimated as $\Delta t \simeq \tau(E_2) - \tau(E_1)$, where τ is the characteristic electron scattering time due to the Coulomb collisions, given, e.g., by Eq. 12.5.11 in the monograph of Aschwanden (2005). For the energies of $E_1 = 20$ keV and $E_2 = 80$ keV, the observed delays of $\Delta t \simeq 5$–7 s would occur due to scattering in a thermal plasma with the density of about $3 \times 10^{10}\mathrm{cm}^{-3}$ and the temperature of a few MK. The obtained plasma density value seems to be typical of the coronal flaring loops.

6 Conclusion

We presented the results of observations and simulations of a M5.8 class solar flare that occurred on 2023-03-06 near the north-eastern solar limb. The flare was observed by a number of instruments, including the Siberian Radioheliograph, Nobeyama Radiopolarimeters, and Radio Solar Telescope Network in the microwave range, Hard X-ray Imager on board the Advanced Space-based Solar Observatory and Konus-Wind on board the Wind spacecraft in the hard X-ray range, and the Solar Dynamic Observatory in the optical, ultraviolet, and extreme ultraviolet ranges. The main results can be summarized as follows:

- The flare consisted of two separate flaring events. The first part of the flare was mostly "thermal," with a relatively soft spectrum of energetic particles and weak microwave emission. During the second part of the flare, the spectrum of energetic particles was much harder and the microwave emission was much stronger than during the first part.
- A filament eruption occurred at the location of the future flare ~6 minutes before the flare onset. This eruption likely triggered the magnetic reconnection process and thus initiated the flare.
- During the first part of the flare, the microwave and X-ray emissions were produced in an arcade of relatively short and low flaring loops. Trapping and accumulation of the energetic particles in the flaring loops were negligible. The evolution of the microwave and hard X-ray sources reflected the dynamics of the energy release processes in the arcade.
- During the second part of the flare, the microwave emission was produced in a single large-scale flaring loop. The energetic particles were concentrated near the loop top. The evolution of the microwave source reflected the process of gradual accumulation of energetic electrons in the flaring loop. Around the individual emission pulses, the dynamics of the trapped and precipitating energetic electrons demonstrated the hard-soft-hard and soft-hard-soft patterns, respectively. The evolution of the trapped energetic electrons was mostly determined by the Coulomb collisions.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: https://ftp.rao.istp.ac.ru/SRH/ (SRH), https://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-features/solar-radio/ (RSTN), https://solar.nro.nao.ac.jp/norp/html/daily/ (NoRP), http://47.104.87.104/SRData/ (CBS), http://jsoc.stanford.edu/ajax/lookdata.html (SDO), https://www.ngdc.noaa.gov/stp/satellite/goes-r.html (GOES), http://aso-s.pmo.ac.cn/sodc/dataArchive.jsp (ASO-S), https://www.ioffe.ru/LEA/kwsun/index.html (Konus-Wind).

Author contributions

AK: Investigation, Writing–original draft. ZW: Investigation, Writing–review and editing. SA: Data curation, Software, Writing–review and editing. YS: Data curation, Software, Writing–review and editing. YC: Conceptualization, Writing–review and editing.

Funding

The authors declare that financial support was received for the research, authorship, and/or publication of this article. This work was supported by the Ministry of Science and Technology of the People's Republic of China (grant No. 110000206220220025). AK and SA acknowledge financial support from the Ministry of Science and Higher Education of the Russian Federation. ASO-S mission is supported by the Strategic Priority Research Program on Space Science, the Chinese Academy of Sciences, Grant No. XDA15320000.

Acknowledgments

We thank the teams of the Siberian Radioheliograph and Advanced Space-based Solar Observatory for maintaining these instruments and for the help with analyzing the observations.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Altyntsev, A., Lesovoi, S., Globa, M., Gubin, A., Kochanov, A., Grechnev, V., et al. (2020). Multiwave siberian Radioheliograph. *Solar-Terrestrial Phys.* 6, 30–40. doi:10.12737/stp-62202003

Aschwanden, M. J. (2002). Particle acceleration and kinematics in solar flares - a synthesis of recent observations and theoretical concepts (invited review). *Space Sci. Rev.* 101, 1–227. doi:10.1023/A:1019712124366

Aschwanden, M. J. (2005). "Physics of the solar corona," in *An introduction with problems and solutions*. 2nd edition (Chichester, UK: Praxis Publishing Ltd).

Benz, A. O., and Güdel, M. (2010). Physical processes in magnetically driven flares on the sun, stars, and young stellar objects. *Annu. Rev. Astron. Astrophys.* 48, 241–287. doi:10.1146/annurev-astro-082708-101757

Brown, J. C. (1971). The deduction of energy spectra of non-thermal electrons in flares from the observed dynamic spectra of hard X-ray bursts. *Sol. Phys.* 18, 489–502. doi:10.1007/BF00149070

Dulk, G. A., and Marsh, K. A. (1982). Simplified expressions for the gyrosynchrotron radiation from mildly relativistic, nonthermal and thermal electrons. *Astrophys. J.* 259, 350–358. doi:10.1086/160171

Emslie, A. G., Dennis, B. R., Shih, A. Y., Chamberlin, P. C., Mewaldt, R. A., Moore, C. S., et al. (2012). Global energetics of thirty-eight large solar eruptive events. *Astrophys. J.* 759, 71. doi:10.1088/0004-637X/759/1/71

Fleishman, G. D., Anfinogentov, S. A., Stupishin, A. G., Kuznetsov, A. A., and Nita, G. M. (2021a). Coronal heating law constrained by microwave gyroresonant emission. *Astrophys. J.* 909, 89. doi:10.3847/1538-4357/abdab1

Fleishman, G. D., Kleint, L., Motorina, G. G., Nita, G. M., and Kontar, E. P. (2021b). Energy budget of plasma motions, heating, and electron acceleration in a three-loop solar flare. *Astrophys. J.* 913, 97. doi:10.3847/1538-4357/abf495

Fleishman, G. D., and Kuznetsov, A. A. (2010). Fast gyrosynchrotron codes. *Astrophys. J.* 721, 1127–1141. doi:10.1088/0004-637X/721/2/1127

Fleishman, G. D., Nita, G. M., and Gary, D. E. (2015). Energy partitions and evolution in a purely thermal solar flare. *Astrophys. J.* 802, 122. doi:10.1088/0004-637X/802/2/122

Fleishman, G. D., Nita, G. M., Kuroda, N., Jia, S., Tong, K., Wen, R. R., et al. (2018). Revealing the evolution of non-thermal electrons in solar flares using 3D modeling. *Astrophys. J.* 859, 17. doi:10.3847/1538-4357/aabae9

Fleishman, G. D., Nita, G. M., and Motorina, G. G. (2023). Data-constrained 3D modeling of a solar flare evolution: acceleration, transport, heating, and energy budget. *Astrophys. J.* 953, 174. doi:10.3847/1538-4357/ace1f4

Gan, W., Zhu, C., Deng, Y., Zhang, Z., Chen, B., Huang, Y., et al. (2023). The advanced space-based solar observatory (ASO-S). *Sol. Phys.* 298, 68. doi:10.1007/s11207-023-02166-x

Huang, G., and Nakajima, H. (2009). Statistics of flaring loops observed by Nobeyama Radioheliograph. II. Spectral evolution. *Astrophys. J.* 702, 19–26. doi:10.1088/0004-637X/702/1/19

Kuroda, N., Gary, D. E., Wang, H., Fleishman, G. D., Nita, G. M., and Jing, J. (2018). Three-dimensional forward-fit modeling of the hard X-ray and microwave emissions of the 2015 june 22 M6.5 flare. *Astrophys. J.* 852, 32. doi:10.3847/1538-4357/aa9d98

Kuznetsov, A. A., and Fleishman, G. D. (2021). Ultimate fast gyrosynchrotron codes. *Astrophys. J.* 922, 103. doi:10.3847/1538-4357/ac29c0

Kuznetsov, A. A., and Kontar, E. P. (2015). Spatially resolved energetic electron properties for the 21 may 2004 flare from Radio observations and 3D simulations. *Sol. Phys.* 290, 79–93. doi:10.1007/s11207-014-0530-x

Lemen, J. R., Title, A. M., Akin, D. J., Boerner, P. F., Chou, C., Drake, J. F., et al. (2012). The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO). *Sol. Phys.* 275, 17–40. doi:10.1007/s11207-011-9776-8

Lysenko, A. L., Ulanov, M. V., Kuznetsov, A. A., Fleishman, G. D., Frederiks, D. D., Kashapova, L. K., et al. (2022). KW-sun: the konus-wind solar flare database in hard X-ray and soft gamma-ray ranges. *Astrophys. J. Suppl. Ser.* 262, 32. doi:10.3847/1538-4365/ac8b87

Ning, Z. (2008). Microwave and hard X-ray spectral evolution for the 13 december 2006 solar flare. *Sol. Phys.* 247, 53–62. doi:10.1007/s11207-007-9101-8

Nita, G. M., Fleishman, G. D., Kuznetsov, A. A., Anfinogentov, S. A., Stupishin, A. G., Kontar, E. P., et al. (2023). Data-constrained solar modeling with GX simulator. *Astrophys. J. Suppl. Ser.* 267, 6. doi:10.3847/1538-4365/acd343

Nita, G. M., Fleishman, G. D., Kuznetsov, A. A., Kontar, E. P., and Gary, D. E. (2015). Three-dimensional Radio and X-ray modeling and data analysis software: revealing flare complexity. *Astrophys. J.* 799, 236. doi:10.1088/0004-637X/799/2/236

Nita, G. M., Viall, N. M., Klimchuk, J. A., Loukitcheva, M. A., Gary, D. E., Kuznetsov, A. A., et al. (2018). Dressing the coronal magnetic extrapolations of active regions with a parameterized thermal structure. *Astrophys. J.* 853, 66. doi:10.3847/1538-4357/aaa4bf

Scherrer, P. H., Schou, J., Bush, R. I., Kosovichev, A. G., Bogart, R. S., Hoeksema, J. T., et al. (2012). The helioseismic and magnetic imager (HMI) investigation for the solar dynamics observatory (SDO). *Sol. Phys.* 275, 207–227. doi:10.1007/s11207-011-9834-2

Shang, Z., Wu, Z., Liu, Y., Bai, Y., Lu, G., Zhang, Y., et al. (2023). The calibration of the 35-40 GHz solar Radio spectrometer with the new moon and a noise source. *Astrophys. J. Suppl. Ser.* 268, 45. doi:10.3847/1538-4365/acee00

Shang, Z., Xu, K., Liu, Y., Wu, Z., Lu, G., Zhang, Y., et al. (2022). A Broadband solar Radio dynamic spectrometer working in the millimeter-wave band. *Astrophys. J. Suppl. Ser.* 258, 25. doi:10.3847/1538-4365/ac4257

Shimojo, M., and Iwai, K. (2023). Over seven decades of solar microwave data obtained with Toyokawa and Nobeyama Radio Polarimeters. *Geosci. Data J.* 10, 114–129. doi:10.1002/gdj3.165

Su, Y., Liu, W., Li, Y.-P., Zhang, Z., Hurford, G. J., Chen, W., et al. (2019). Simulations and software development for the Hard X-ray Imager onboard ASO-S. *Res. Astron. Astrophys.* 19, 163. doi:10.1088/1674-4527/19/11/163

White, S. M., Benz, A. O., Christe, S., Fárník, F., Kundu, M. R., Mann, G., et al. (2011). The relationship between solar Radio and hard X-ray emission. *Space Sci. Rev.* 159, 225–261. doi:10.1007/s11214-010-9708-1

Wu, Z., Kuznetsov, A., Anfinogentov, S., Melnikov, V., Sych, R., Wang, B., et al. (2024). A multipeak solar flare with a high turnover frequency of the gyrosynchrotron spectra from the loop-top source. *Astrophys. J.* 968, 5. doi:10.3847/1538-4357/ad46ff

Yan, F., Wu, Z., Shang, Z., Wang, B., Zhang, L., and Chen, Y. (2023). The first flare observation with a new solar microwave spectrometer working in 35-40 GHz. *Astrophys. J. Lett.* 942, L11. doi:10.3847/2041-8213/acad02

Zharkova, V. V., Arzner, K., Benz, A. O., Browning, P., Dauphin, C., Emslie, A. G., et al. (2011). Recent advances in understanding particle acceleration processes in solar flares. *Space Sci. Rev.* 159, 357–420. doi:10.1007/s11214-011-9803-y

Zirin, H., Baumert, B. M., and Hurford, G. J. (1991). The microwave brightness temperature spectrum of the quiet sun. *Astrophys. J.* 370, 779. doi:10.1086/169861

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 22 July 2024
DOI 10.3389/fspas.2024.1401846

Check for updates

OPEN ACCESS

EDITED BY
Xiangliang Kong,
Shandong University, Weihai, China

REVIEWED BY
Sijie Yu,
New Jersey Institute of Technology, United States
Tingyu Gou,
Harvard University, United States
Yulei Wang,
Nanjing University, China

*CORRESPONDENCE
Jiong Qiu,
qiu@montana.edu

RECEIVED 16 March 2024
ACCEPTED 29 May 2024
PUBLISHED 22 July 2024

CITATION
Qiu J (2024), Tracing field lines that are reconnecting, or expanding, or both. *Front. Astron. Space Sci.* 11:1401846. doi: 10.3389/fspas.2024.1401846

COPYRIGHT
© 2024 Qiu. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Tracing field lines that are reconnecting, or expanding, or both

Jiong Qiu*

Department of Physics, Montana State University, Bozeman, MT, United States

The explosive release of energy in the solar atmosphere is driven magnetically, but the mechanisms that trigger the onset of the eruption remain controversial. In the case of flares and coronal mass ejections (CMEs), ideal or non-ideal instabilities usually occur in the corona, but it is difficult to obtain direct observations and diagnostics there. To overcome this difficulty, we analyze observational signatures in the upper chromosphere or transition region, particularly brightening and dimming at the base of coronal magnetic structures. In this paper, we examine the time evolution of spatially resolved light curves in two eruptive flares and identify a variety of tempo-spatial sequences of brightening and dimming, such as dimming followed by brightening and dimming preceded by brightening. These brightening–dimming sequences are indicative of the configuration of energy release in the form of plasma heating or bulk motion. We demonstrate the potential of using these analyses to diagnose the properties of magnetic reconnection and plasma expansion in the corona during the early stages of the eruption.

KEYWORDS

magnetic reconnection, solar flares, coronal mass ejections, solar eruptions, ultraviolet radiation

1 Introduction

It is well known that explosive energy releases in the form of solar flares and coronal mass ejections (CMEs) are driven magnetically. In this process, magnetic energy is converted to the kinetic energy of particles and the heat and bulk motion of plasmas (Thompson et al., 2021). CMEs open up a portion of the solar corona, along which energetic particles are released into interplanetary space and can impact the space weather. At present, routine measurements of the full disk magnetic field are available only in the lower solar atmosphere, the photosphere. Therefore, the capability to identify the solar surface signatures of an open magnetic structure and track magnetic field evolution prior to eruption will help us explore mechanisms governing the onset of the eruption and predict space weather.

Traditionally, coronal holes (Cranmer, 2009), or regions of persistent lack of emission at soft X-ray (SXR) and extreme ultraviolet (EUV) wavelengths, are considered to map the feet of open field lines on the solar surface. The lifetimes of coronal holes range from many days to multiple weeks (Lowder et al., 2017, and references therein). The temporary, and often abrupt, opening up of the magnetic structure associated with a CME produces transient coronal holes, also called coronal dimmings (Sterling and Hudson, 1997). A comprehensive review of observational signatures and an interpretation of coronal dimmings has been

provided by Veronig et al. (2024, in press). Mechanisms causing the explosive "opening up" of the corona include the onset of an ideal instability, in which the global force balance is lost during the quasi-equilibrium evolution of the system (Forbes and Isenberg, 1991; Török and Kliem, 2005; Kliem and Török, 2006; Isenberg and Forbes, 2007), or the onset of a non-ideal instability, often referred to as magnetic reconnection, which abruptly changes the connectivity of the field lines and, in this way, removes or weakens the constraints that would maintain the force balance of a coronal structure. The "tether-cutting" (Moore et al., 2001) and "break-out" (Antiochos et al., 1999) configurations have been the most well-known reconnection geometry, which occurs either below or above the erupting structure, be it a magnetic flux rope or a sheared arcade (Patsourakos et al., 2020).

In the case of major solar eruptive events, the onset of the eruption, either due to ideal or non-ideal instabilities, takes place in the corona, but perturbations propagate along magnetic field lines to reach the lower atmosphere on Alfvénic timescales. In this paper, we focus on dimming and brightening signatures observed at the foot or base of magnetic structures that are undergoing dynamic evolution, such as magnetic reconnection or plasma expansion/eruption. As a direct consequence of magnetic reconnection, energy flux, via particle beams (Fisher et al., 1985), thermal conduction (Longcope, 2014), or Alfvén waves (Fletcher and Hudson, 2008), is transported along newly reconnected field lines to be deposited in the denser lower atmosphere, producing enhanced brightening there. On the other hand, dimming at the base of the corona is primarily an effect of plasma rarefaction due to the expansion of the overlying coronal structures such as a CME, which reduces the pressure of the overlying corona. The expansion or eruption of a coronal structure may occur before or after the onset of magnetic reconnection; therefore, dimming may be observed either before or after brightening in the lower atmosphere.

A solar eruptive event involves a rather complex magnetic configuration, and different parts of the system undergo different dynamics and also interact with each other, such as through reconnection. The spatially resolved, full-disk observations, like those provided by the Solar Dynamics Observatory (SDO; Pesnell et al., 2012), allow us to infer the evolution of different parts of the coronal structures during or before the eruption by examining the behavior at their base in the lower atmosphere. In this paper, we conduct such an experiment on two eruptive flares, using the tempo-spatial sequence of brightening and dimming to reconstruct the evolution of coronal structures in the early phase of their eruption. In the following section, we describe the strategy to identify dimming signatures at the base of the corona (S2). We apply the analysis to observations of an X-class eruptive event SOL20120712, which shows post-eruption dimming (S3), and of a C-class eruptive event SOL20110621, which also exhibits pre-eruption dimming (S4), and use these signatures to infer the properties of the overlying coronal structure. We summarize what is learned from this experiment and discuss the potential to reconstruct the evolution of overlying coronal structures from the brightening and dimming signatures at their base (S5).

2 Tempo-spatial sequence of brightening and dimming

In this paper, we analyze the brightening and dimming signatures in the lower atmosphere of two eruptive events. They exhibit a variety of dimming–brightening sequences that can be used to diagnose the dynamic evolution of overlying coronal structures in the early phase of the eruption.

2.1 Impulsive and prolonged brightening

We identify brightening in the lower atmosphere, as these locations usually map the feet of reconnecting field lines along which energy flux is transported and deposited in the denser lower atmosphere. It has been a well-adopted practice to integrate magnetic flux in these areas of brightening in the lower atmosphere as an estimate of the amount of reconnection flux ψ_{rec} and its time derivative, the reconnection rate $\dot{\psi}_{rec}$ (Fletcher and Hudson, 2001; Asai et al., 2004; Qiu et al., 2004; 2010; Kazachenko et al., 2017). Typically, observations in the optical (such as the H_{alpha} line) and ultraviolet (such as the UV 1,600 Å passband) wavelengths are analyzed to identify brightening signatures. In this study, we analyze either the 1,600 Å or the EUV 304 Å observations from the Atmospheric Imaging Assembly (AIA; Lemen et al., 2012). The contribution to the 1,600 Å broadband emission includes the continuum formed in the temperature minimum region and, in particular, during the flare, optically thin lines like C IV formed at a transition-region temperature of 100,000 K (Simões et al., 2019). The contribution to the EUV 304 Å broadband is more complex, from the upper chromosphere, transition region, and low corona, at a temperature of 0.01–1 MK (O'Dwyer et al., 2010). The EUV 304 Å passband is more sensitive to weak brightening in the upper chromosphere and transition region; on the other hand, structures in the corona, such as filaments, active region loops, and flare loops, are often observed in this passband and are sometimes hard to distinguish from the brightening at the base of the coronal structures.

When magnetic reconnection occurs in the corona, energy flux along newly reconnected field lines travels to the lower atmosphere on Alfvénic timescales and produces a rapid or *impulsive brightening* there. If reconnection forms closed field lines (post-reconnection flare loops), the impulsive energy release in these loops drives chromospheric evaporation, which significantly increases the density (and temperature) of the plasmas trapped in the loops. As a result, *prolonged brightening* is often observed at the base of the post-reconnection flare loops for an extended duration of more than 10 min, before the brightness is attenuated to a pre-flare level, reflecting the timescales of chromospheric evaporation, cooling of the heated corona, and often, continuous gradual heating (Qiu et al., 2013; Qiu and Longcope, 2016). Therefore, light curves similar to those shown in Figure 1A are indicative of reconnection forming closed field lines or flare loops. On the other hand, reconnection leading to open field lines, which do not trap plasmas, would produce only impulsive brightening at its base, as shown in Figure 1B. The pixel light curves are, therefore, indicative of the reconnection geometry.

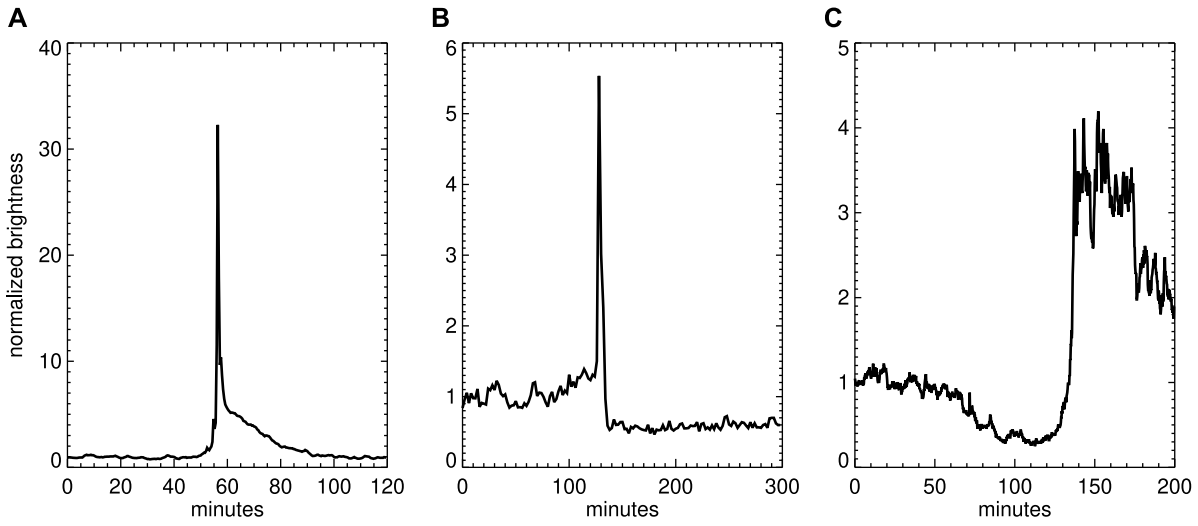

FIGURE 1
Light curves of the brightness in the UV 1,600 or EUV 304 Å passbands in several pixels from various flares observed by the AIA, showing **(A)** impulsive brightening followed by prolonged brightening, **(B)** rapid dimming following impulsive brightening, or **(C)** gradual dimming before brightening.

In this study, we identify brightening pixels in the UV 1,600 Å passband or EUV 304 Å passband when the pixel brightness I is enhanced to be more than N times of its base brightness I_0 for more than τ min; I_0 is the average of the brightness over 20 min during the quiescent (pre-flare) period. N and τ are empirically selected to pick out as many brightening pixels as possible in the lower atmosphere while minimizing the contribution by brightening coronal loops.

2.2 Rapid and gradual dimming

The majority of coronal dimming is observed in the soft X-ray and EUV wavelengths (Sterling and Hudson, 1997; Thompson et al., 1998; Mandrini et al., 2007; Qiu et al., 2007; Temmer et al., 2017; Dissauer et al., 2018; Wang et al., 2019). In this study, we examine signatures in the upper chromosphere, or transition region, where coronal structures undergoing dynamic evolution are anchored. We identify dimming signatures at the base of the corona primarily using observations in the He II 304 passband on the AIA, complemented by analysis of more conventional observations in the EUV passbands, including the EUV 171, 193, and 211 Å passbands. The other EUV passbands are sensitive to temperatures ≥ 1 MK (O'Dwyer et al., 2010) of plasmas in the corona. As such, dimming identified in the EUV 304 Å passband tends to occupy smaller areas than those identified in other bands, which often include dimming signatures due to the **removal** or re-orientation of coronal loops along the line of sight (Harvey and Recely, 2002; Harra et al., 2007; Qiu et al., 2007; Downs et al., 2015). In other words, detecting dimming in the transition region lines helps minimize the projection effect.

We examine the time evolution of the brightness at each pixel, normalized to its base brightness. The base brightness is the mean brightness over 20–30 min prior to the eruption. The dimming pixels are identified if the brightness is reduced to ≤ 80% of the base brightness continuously for ≥10 min. The choice of the minimum dimming depth at 80% and the minimum dimming duration of 10 min is mostly empirical, justified by the statistical performance of the pixel brightness. For example, the fluctuations in the quiescent brightness of individual pixels are found to be about 10%; therefore, a persistent decrease in the brightness at 20% below the base brightness is considered to reflect genuine dimming.

We assume that these dimming pixels map the feet of the magnetic structures that are expanding or erupting in the corona. Here, we define "expansion" and "eruption" as the global or average bulk motion of the plasma in a magnetic structure, the former referring to motion at subsonic speeds and the latter at supersonic speeds. If dimming is preceded by brightening at the same or adjacent locations, this is likely an indication that reconnection is opening overlying field lines, and the magnetic flux integrated into the brightening area provides an estimate of the amount of flux removed from above. Furthermore, the dynamic properties, such as the mean speed of the expansion of the overlying corona, can be also estimated from the evolution of the dimming depth.

The pixel light curves given in Figure 1 show various dimming signatures, such as *rapid dimming* following impulsive brightening (b), indicative of reconnection of the opening up of field lines along which plasmas rapidly expand, or *gradual dimming* over tens of minutes, followed by brightening or rapid dimming (panel c), suggesting the quasi-equilibrium expansion of the overlying coronal structure before the onset of reconnection or eruption. The tempo-spatial sequence of dimming and brightening therefore provides clues to the dynamic evolution of the corona.

3 Reconnection-driven post-eruption dimming

An X-class eruptive flare occurred on 12 July 2012 in NOAA-11520. The flare was accompanied by a fast CME with its early-phase motion best captured by STEREO-A (Cheng et al.,

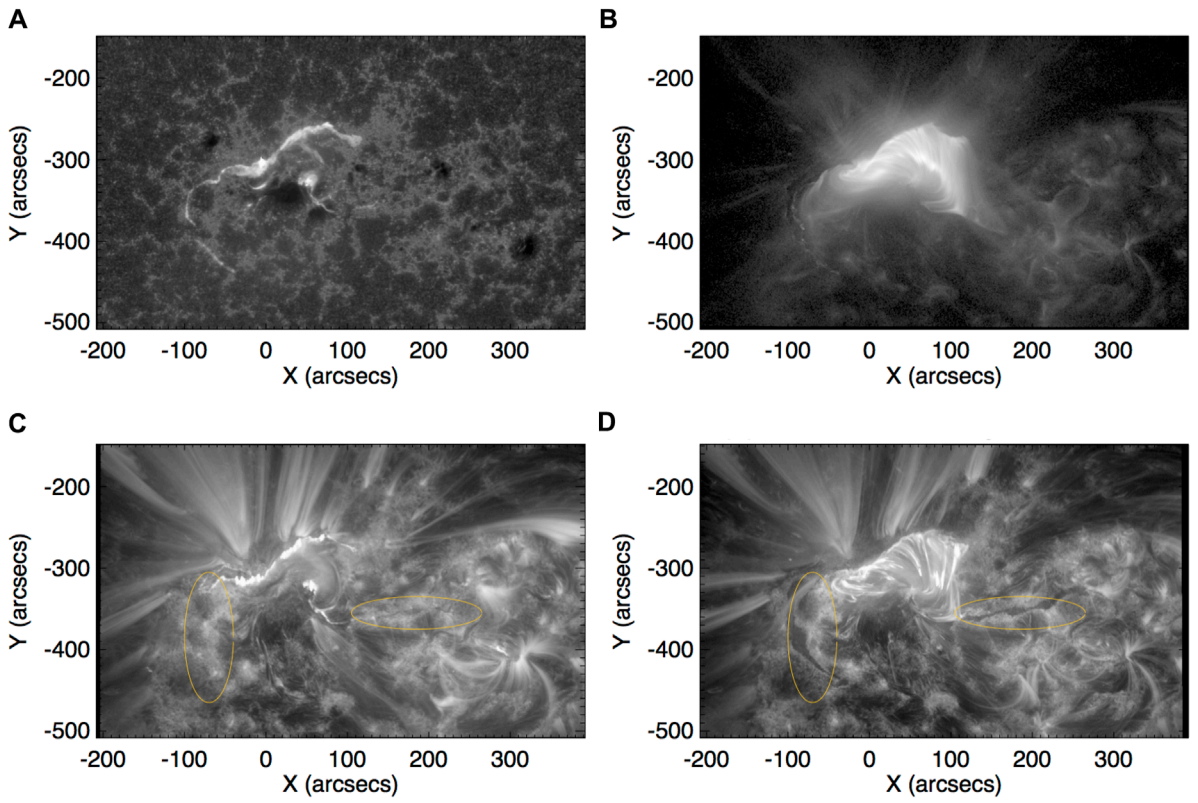

FIGURE 2
Overview of the SOL20120712 eruptive flare observed by the AIA. Top: flare ribbons observed in the UV 1,600 Åpassband **(A)** and flare loops observed in the EUV 131 Åpassband **(B)**. Bottom: flare ribbons and loops observed in the EUV 171 Åpassband at two times **(C,D)**, showing a pair of post-eruption dimming ribbons (their locations indicated by the two orange ovals) extending from the ends of the two flare ribbons. All images are co-aligned to 17:00 UT.

2014; Dudík et al., 2014; Zhu et al., 2020). Figure 2 shows several snapshots of the event observed by the AIA. The eruption occurred in the core active region, forming two bright ribbons in magnetic fields of opposite polarities, outlining the feet of post-reconnection flare loops or closed field lines produced by reconnection. We refer to these two ribbons as *flare ribbons*. Shortly afterward, another set of two long ribbons extended from the flare ribbons in the core region, first briefly brightening and then quickly dimming. We call these two ribbons that became dark *dimming ribbons*, and they were also located in magnetic fields of opposite polarities.

Figure 3A shows, in red, the time sequence of the brightening in the *flare ribbons* in the core region and also the brightening adjacent to the dimming ribbons away from the core region. This reconnection mask is derived using the UV 1,600 Å observations, with the method described by Qiu et al. (2007, 2010). In the same figure, blue indicates the time sequence of the dimming as the brightness is reduced to 80% of the base level (see S2 for identification of the dimming pixels). The dimming (and the preceding brightening) appears to spread rapidly along the dimming ribbons at an apparent speed of more than 100 km s $^{-1}$, which was noted as a signature of "slipping reconnection" by Dudík et al. (2014).

The magnetic flux integrated in the brightening pixels and in the dimming pixels is measured and shown in Figure 3B, showing that, in this event, the onset of flare reconnection starts at approximately 16 UT, prior to the rapid increase in the dimming. The flare reconnection rate peaks at 16:20 UT and stops at approximately 17 UT, when a total amount of 7×10^{21} Mx has been reconnected. The dimming flux is also measured, which increases quickly after 16:20 UT, and the dimming rate peaks at approximately 17 UT, with the total amount of dimming flux reaching 5×10^{20} Mx encompassed in the dimming ribbons. The associated CME is observed by STEREO; it shows a rapid acceleration starting from 16:10 UT and reaching a maximum at 16:20 UT (Zhu et al., 2020), when the rate of flare reconnection peaks.

The dimming ribbons are most prominently observed in the EUV 304, 171, and 193 Å passbands. Figure 4 shows the time–distance diagrams of the normalized brightness produced along five slits across each *dimming ribbon*; the locations of the slits are indicated by horizontal or vertical bars in Figure 3A. In particular, the diagrams of the 304 Å passband clearly demonstrate dimming after impulsive brightening. Figure 5 shows the epoch plot of the pixel light curves in the two dimming ribbons, or the evolution of the brightness (normalized to the base brightness) with respect to the time of the peak brightness. The figure illustrates the timescales of impulsive brightening, followed by rapid dimming, both within a couple of minutes. In most of these places, dimming does not recover for more than a few hours.

The tempo–spatial sequence of the brightening and dimming in this event suggests the scenario of magnetic reconnection between

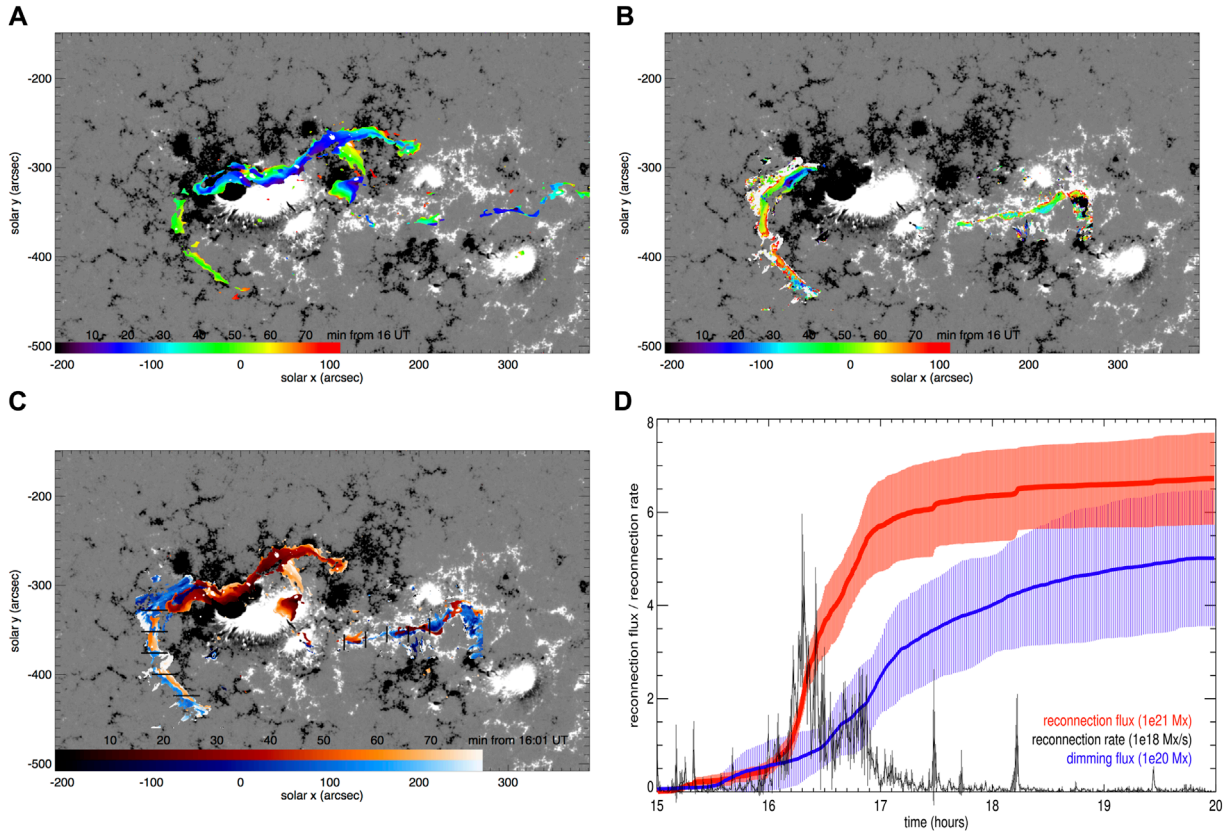

FIGURE 3
Evolution of brightening and dimming. **(A)** Mapping of the brightening evolution (from AIA 1,600 Å) on a photospheric magnetogram (grayscale) of the longitudinal magnetic field component B_{los}. **(B)** Mapping of the dimming evolution (from AIA 304 Å) on the photospheric magnetogram. **(C)** Mapping of brightening (orange) and dimming (blue) on the photospheric magnetogram. **(D)** Reconnection flux ψ_{rec} and reconnection rate $\dot{\psi}_{rec}$ measured from the brightening map and the dimming flux ψ_{dim} measured from the dimming map. In **(A–C)**, the grayscale of the B_{los} map is saturated at ±300 (G). In **(A)** and **(B)**, the rainbow color indicates the onset time of the brightening or the dimming, respectively. In **(C)**, the orange (blue) color scheme indicates the onset time of brightening (dimming) at a given location. It should be noted that in the dimming ribbons, the brightening is covered by the subsequent dimming. The five horizontal bars across the dimming ribbon in the east and the five vertical bars across the dimming ribbon in the west denote the locations of the time–distance diagrams of the normalized brightness in Figure 4.

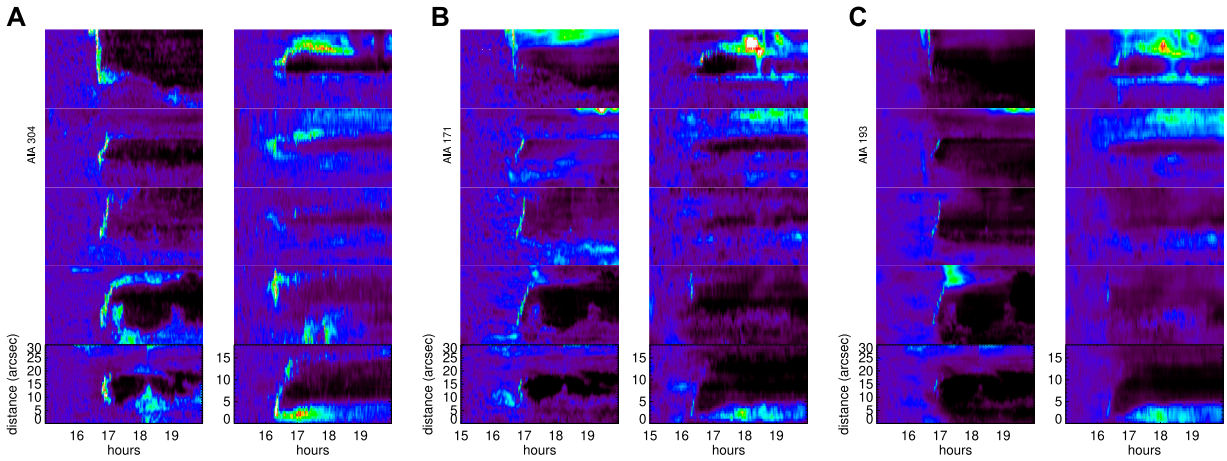

FIGURE 4
Time–distance diagrams of the normalized brightness along the five slits from top to bottom across the dimming ribbon in the east (left) and along the five slits from left to right across the dimming ribbon in the west (right), in the 304 Å **(A)**, 171 Å **(B)**, and 193 Å **(C)** passbands, respectively, all showing rapid dimming after impulsive brightening.

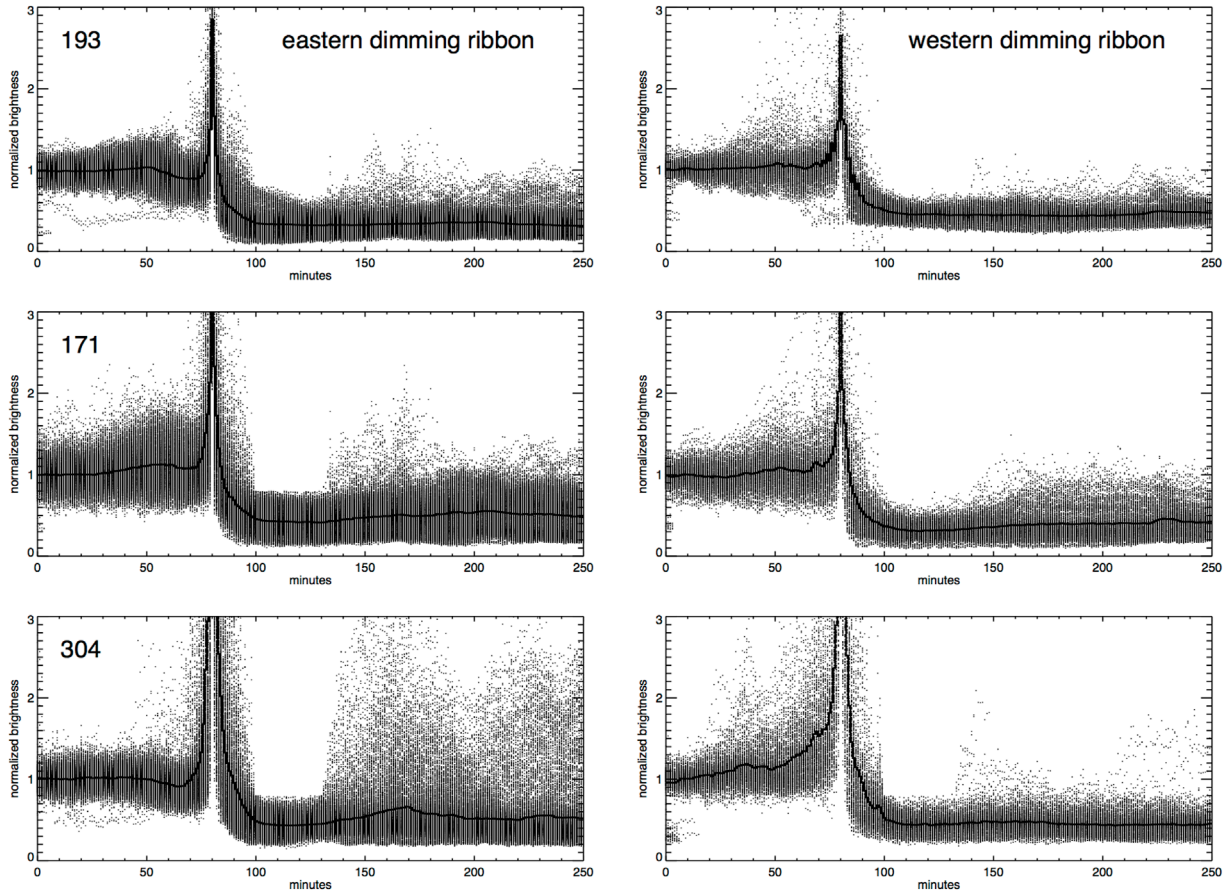

FIGURE 5
Epoch plots of the normalized brightness light curves of approximately 1,000 pixels in the eastern dimming ribbon (left) and of approximately 300 pixels in the western dimming ribbon (right), observed in the EUV 193 Å, 171 Å, and 304 Å passbands, respectively.

an erupting structure from the core region and the overlying arcades or the strapping field. The two dimming ribbons outline the photospheric intersection of separatrices in the complex multi-polar magnetic field where reconnection tends to occur (Dudík et al., 2014). In particular, a potential field extrapolation suggests that the dimming ribbon in the negative magnetic field to the east outlines the feet of an arcade connecting the dimming ribbon to the outer edge of the flare ribbon in the positive magnetic field, and the dimming ribbon in the positive magnetic field to the west maps the feet of another set of arcades with their conjugate feet in the negative magnetic field north of the flare ribbon (Cooper Downs, personal communication). These overlying arcades have to open up for the underlying structure to escape the solar corona.

In summary, in this eruptive event, the dimming analysis does not reveal any significant pre-eruption dynamics of the coronal magnetic structures. Dimming occurs after the onset of flare reconnection in the core region. The pair of dimming ribbons extending outward from the flare ribbons most likely map the feet of several overlying arcades that open up by reconnecting with the erupting structure. In this course, an overlying flux of 5×10^{20} Mx, which is 10% of the flare reconnection flux, is removed from the path of the eruption.

4 Pre-eruption dimming indicative of gradual expansion

A C-class eruptive flare occurred on 21 June 2011. It has been studied by Zhu et al. (2020) and Vievering et al. (2023). The associated CME is best observed by STEREO-A from the limb during its early phase; following the trajectory of the CME, its onset is determined to be at 2:15 UT (Zhu et al., 2020). Figure 6 shows the evolution of the event observed in the EUV 304 Å (left) and 211 Å (right) passbands. A filament is visible prior to the eruption and is erupted around the time of the CME onset. During the eruption, two flare ribbons are brightened as depicted in a standard model; the eruption also causes coronal dimming in a large area to the south of the source region. In this study, we focus on the core region where the eruption originates, attempting to understand the early-phase evolution.

This event exhibits a variety of brightening and dimming signatures different from those of the SOL20120712 event. As shown in Figure 6, prominent dimming signatures are observed in both passbands and, in particular, in the EUV 304 Å base ratio images (middle panels). The dimming primarily occurs in two regions demarcated by the orange boxes that are located at the two ends

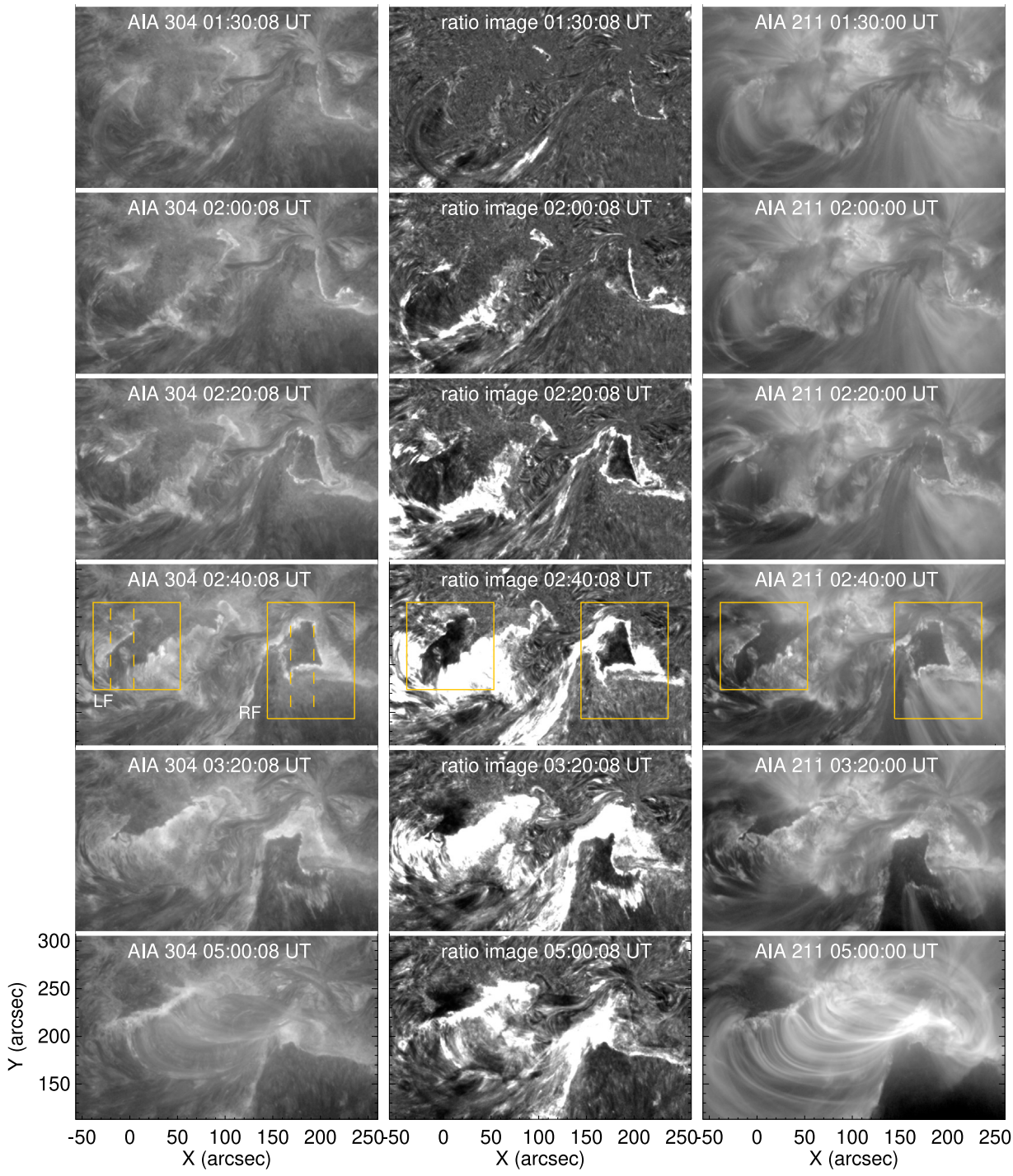

FIGURE 6
Overview of the SOL20110621 eruptive flare observed by the AIA, in the EUV 304 Å passband (left) and 211 Å passband (right). Also shown are base ratio images in the 304 Å passband (middle), with the images normalized to the base image, which is the average of the images from 1:00 UT to 1:20 UT. The two orange boxes denote regions of prominent dimming at the left (LF) and right (RF) feet of a filament visible in the EUV images. The four vertical dashed bars in the left panel mark the slits, along which the time–distance diagrams are produced in Figure 8.

of the filament. The dimming morphology resembles the twin-dimming or core-dimming geometry, similar to the events reported by Webb et al. (2000), Cheng and Qiu (2016), and Wang et al. (2019), suggesting that the twin-dimming regions may map the feet of a flux rope. It is also noted that along the edge of the dimming cores at the far ends of the two flare ribbons, brightening occurs as early as 1:30 UT, well before the onset of the eruption and flare reconnection that produced the two ribbons.

In this event, the brightening is relatively weak compared with the X-class flare studied in S3; therefore, we use EUV 304 Å passband

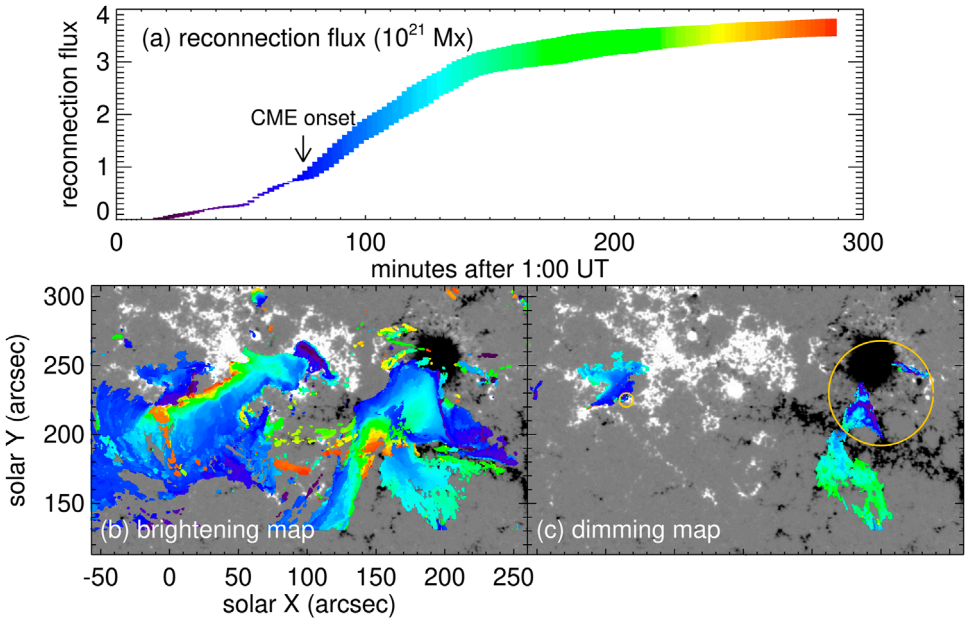

FIGURE 7
(A) Magnetic flux, or reconnection flux, measured in brightening pixels. **(B–C)** Time evolution of the brightening and dimming mapped on a photospheric magnetogram of the line-of-sight magnetic field component. At each location, the color indicates the time of the onset of the brightening (B) or dimming (C) as defined in the text. The time of the color scheme is shown in (A).

observations, which are more sensitive to weak brightening than the 1,600 Å passband, to identify brightening signatures at the base of the corona. Two kinds of brightening signatures are observed: flare ribbons underlying closed flare loops exhibit impulsive brightening, followed by a prolonged brightening that gradually decays for over 10 min; on the other hand, several pixels of weak and brief brightening are also identified. Figure 7B shows the timing of brightening[1] superimposed on a photospheric magnetogram of B_{los}. The magnetic flux integrated into the brightening pixels is shown in Figure 7A. It should be noted that due to some mixture of the brightening of coronal features, especially during the filament eruption between 2:15 and 2:40 UT, the measured flux is an overestimate of the reconnection flux ψ_{rec} (see the measurement obtained by Zhu et al. (2020) using 1,600 Å passband images, which likely underestimates the total ψ_{rec}). It is noted that, in this event, the initial brightening occurs adjacent to the twin dimmings at the far ends of the later formed by two flare ribbons, suggesting that pre-eruption reconnection might take place between the hypothetic flux rope and ambient fields. The amount of flux estimated from the early brightening signatures prior to the eruption (2:15 UT) is about one-fifth of the total reconnection flux.

We identify dimming using EUV 304 Å imaging observations following the method outlined in S2. We form the base image as the average of the images between 1:00 and 1:20 UT and define dimming as persistent attenuation at ≤80% of the base brightness I_0 at the same location for more than 10 min. The onset of dimming, which is shown in Figure 7C, is identified as the time when the brightness starts to decrease below 80% of I_0. The analysis reveals two groups of dimming signatures: one group exhibits gradual dimming over tens of minutes prior to the eruption, and in the other group, dimming occurs after the onset of the eruption (at 2:15 UT), often preceded by impulsive brightening.

Figure 8 shows the time–distance diagrams of the normalized brightness along a few slits shown in Figure 6 and the light curves of a few selected pixels along the slits exhibiting various dimming–brightening sequences similar to those shown in Figure 1. The epoch plots of the two groups of dimming light curves are given in Figure 9, showing the evolution of the pixel brightness with respect to their peak time. The first group exhibits gradual dimming (Figure 9A), with the observed onset of the dimming ranging between 10 and 100 min (average at 54 ± 19 min) before they brightened. The other group of dimming pixels is characterized by rapid dimming following impulsive brightening (Figure 9B), and the peak dimming depth is approximately 42% ± 12%. In most places, dimming did not recover after more than 3 hours.

The locations of the two groups of dimmings are shown in Figure 7C. The pre-eruption gradual dimmings are within the orange circles; they are mostly clustered in the right region in negative magnetic fields and are bounded by early brightening that outlines a triangle-shaped hook. In the left region in positive magnetic fields, only a small area right next to the flare ribbon exhibits pre-eruption gradual dimming. The post-eruption dimming occupies a larger area in both the left and right feet regions. It appears that, at the onset of the eruption, the pre-eruption dimming spread out into the post-eruption dimming, with a rapid brightening during the transition.

1 The onset time of the brightening depends on the empirical threshold used to identify the brightening, and it is typically within 0–3 min of the time of peak brightness.

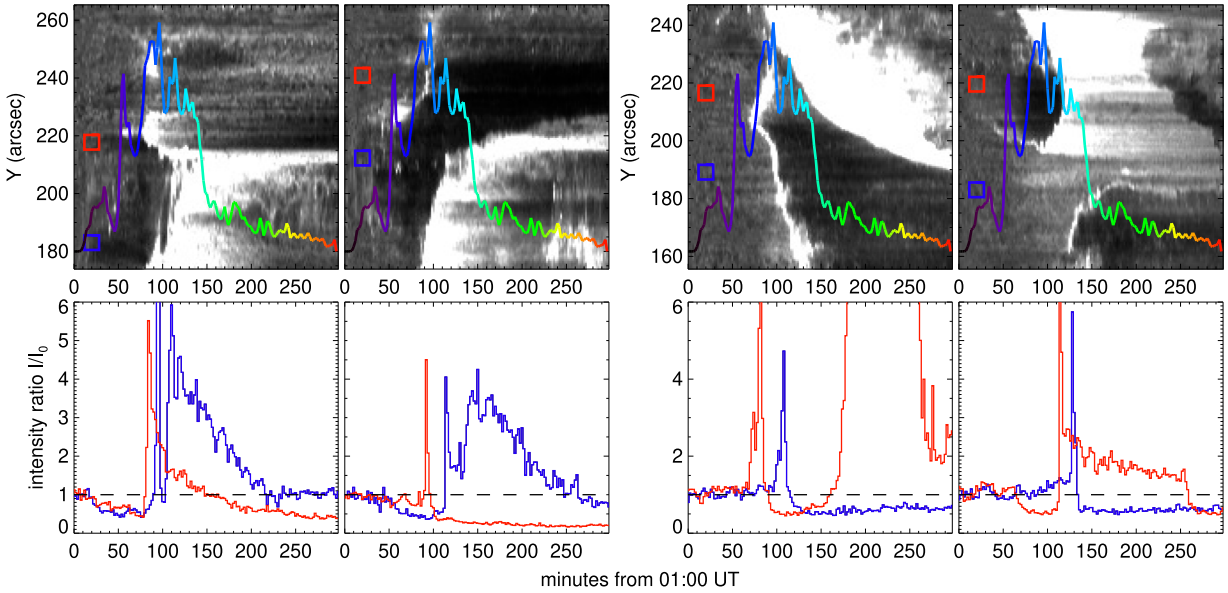

FIGURE 8
Top: time–distance diagrams of the normalized brightness along the four slits shown in Figure 6, two in the left foot region (left) and two in the right foot region (right), superimposed with the reconnection rate $\dot{\psi}_{rec}$. The color code of $\dot{\psi}_{rec}$ is the same as in Figure 7. Bottom: light curves of the normalized brightness at a few locations indicated by the red or blue symbols in the top panels.

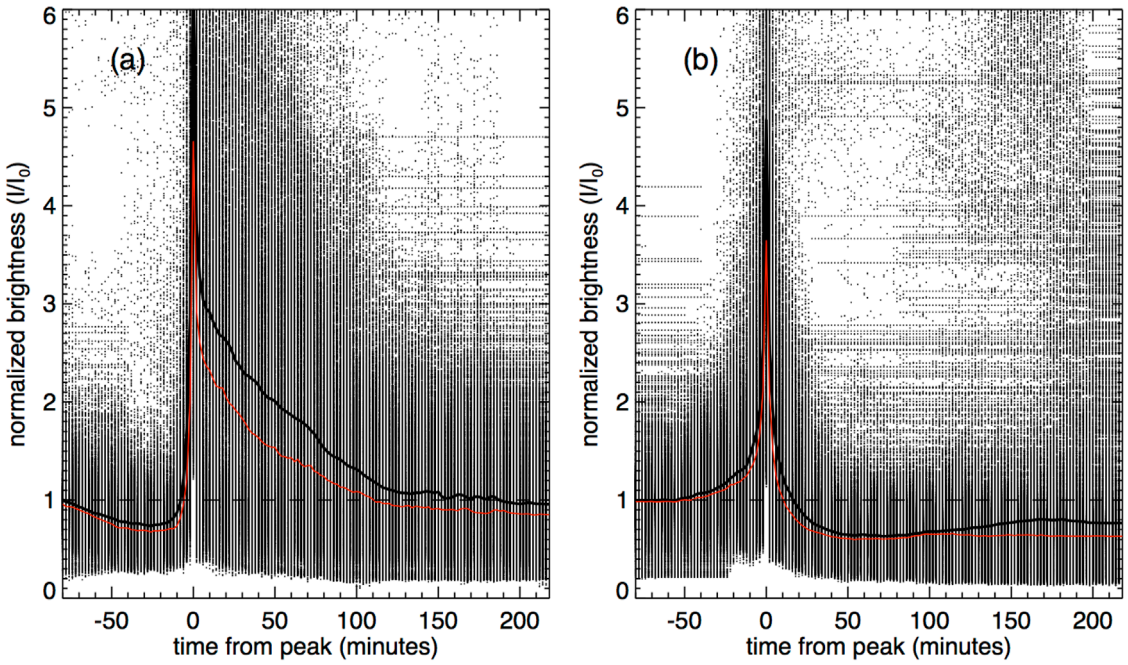

FIGURE 9
Epoch plots of dimming light curves in two groups showing the pre-eruption gradual dimming (left) and the post-eruption impulsive dimming (right). Solid black and red curves indicate the average and median light curves, respectively, of all light curves in a group.

The post-eruption dimming light curves in the second group are similar to those in the SOL20120712 event, suggesting that the erupting structure reconnects with overlying fields and escapes from the corona. However, the geometry of the dimming is different; in this event, the post-eruption dimming appears to map the feet of the erupting structure that has shifted outward after the onset of the flare reconnection, rather than the feet of overlying arcades. This should be confirmed with future data-driven MHD modeling.

The pre-eruption gradual dimming at the feet of the filament is likely a signature of the quasi-equilibrium expansion of a coronal

structure. We can estimate the mean speed of the expansion from the slope of the dimming depth (Qiu and Cheng, 2017). At the base of the corona, the optically thin transition-region emission can be modeled as the "pressure gauge" (Qiu et al., 2013, and references therein) so that the brightness is proportional to the mean pressure of the overlying corona. As the overlying corona gradually expands, the dimming in terms of the ratio of the brightness to the base brightness, $R \equiv I/I_0$, roughly varies as $\dot{R} \approx \alpha(L_0/L)^{\alpha}(\langle v \rangle/L)$, where $\alpha \geq 1$ is a factor close to unity dependent on the gas expansion model, L is the equivalent height along the line of sight, and L_0 is the height before the expansion. For slow (subsonic) expansion, approximating $L \approx L_0$ and $\alpha \approx 1$, the expansion velocity is roughly $\langle v \rangle \approx \dot{R}L_0$. Fitting the dimming light curves by $R(t) \equiv I(t)/I_0 = R_0 - \dot{R}t$, we derive the dimming slope $\dot{R} \approx 0.001 - 0.01$ min^{-1}, the average being 0.003 min^{-1}, or about 0.3% decrease in brightness per minute. For L_0 of a few tens of Mm, $\langle v \rangle$ is approximately a few kilometers per second. The estimated subsonic expansion speed is consistent with the expansion speed directly observed from limb observations of a different event, which also exhibits persistent gradual dimming, when viewed from the disk, prior to the eruption (Wang et al., 2019).

In summary, the eruptive event SOL20110621 displays different brightening–dimming geometries and sequences from those of the SOL20120712 event. Both pre-eruption and post-eruption dimming occur at the feet of a filament adjacent to flare ribbons, and its geometry resembles the core twin dimming. The twin dimming likely maps the conjugate feet of a coronal structure that gradually expands and then erupts, and this structure carries the total (axial) flux of up to a few times 10^{19} Mx, estimated by integrating the flux in the dimming area. The early brightening surrounding the gradually dimming cores may indicate the coronal structure expanding through (and interacting with) the ambient field.

5 Summary

In this paper, we present an experiment that analyzes the brightening and dimming signatures in the lower atmosphere and uses the tempo-spatial sequence of these signatures to identify overlying magnetic structures that are undergoing dynamic evolution in the early phase of their eruption, such as reconnection or expansion. The experiment is applied to two eruptive events, in which pixel light curves exhibit a variety of brightening–dimming sequences.

For the SOL20120712 event, we did not find signatures indicative of dynamic evolution prior to flare reconnection, which occurs nearly simultaneously with the eruption. The erupting structure then reconnects with the overlying arcades, producing impulsive brightening and rapid dimming at the feet of the arcades as the overlying field lines are opened up and removed from their path. We did not find any signatures likely to map the feet of the erupting structure in this event (e.g., Gou et al., 2023). This is possibly due to the stringent criteria for selecting dimming pixels in this study, which require persistent dimming (for more than 10 min) observed with low-temperature lines characterizing signatures at the base of the corona, such as in the transition region. These requirements are reinforced to help minimize projection effects and noise fluctuations; on the other hand, dimming at these lines, particularly in the EUV 304 Å passband, is relatively weak. Therefore, dynamic dimming variations on short timescales may not be detected with the method used in this paper, and the estimated dimming flux is likely to be the lower limit.

The SOL20110621 event exhibits gradual pre-eruption dimming at the feet of a filament that erupts later, and the pre-eruption dimming is accompanied by early brightening that outlines the boundary of the gradual dimming at the far ends of flare ribbons formed later. These are likely signatures of the quasi-equilibrium expansion of a coronal structure through the ambient field, which persists for tens of minutes before the explosive loss of equilibrium, leading to eruption. If this structure embodies a pre-existing flux rope, then its gradual expansion and accompanying reconnection with the overlying fields (e.g., Longcope and Forbes, 2014) would bring it to a larger height in favor of the onset of the ideal instabilities, such as the torus instability (Kliem et al., 2014).

Comprehensive modeling (e.g., Rempel et al., 2023) is needed to reconstruct the three-dimensional magnetic configuration and the evolution toward eruption in both events. The identified tempo-spatial sequence of dimming and brightening provides additional observational constraints, or the observed boundary conditions, for the successful modeling of real eruptive events.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

JQ: conceptualization, data curation, formal analysis, funding acquisition, investigation, methodology, software, and writing–original draft.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. This work was motivated by discussions during the coronal dimming workshop sponsored by the International Space Science Institute (ISSI) at Bern, Switzerland, and during the magnetic flux rope workshop sponsored by the Institute for Space–Earth Environmental (ISEE) Research in Nagoya University of Japan. JQ was supported by NASA grant nos 80NSSC22K0519 and 80NSSC23K0414. *SDO* is a mission of NASA's Living With a Star program.

Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Antiochos, S. K., DeVore, C. R., and Klimchuk, J. A. (1999). A model for solar coronal mass ejections. *ApJ* 510, 485–493. doi:10.1086/306563

Asai, A., Yokoyama, T., Shimojo, M., Masuda, S., Kurokawa, H., and Shibata, K. (2004). Flare ribbon expansion and energy release rate. *ApJ* 611, 557–567. doi:10.1086/422159

Cheng, J. X., and Qiu, J. (2016). *ApJ* 825, 37. doi:10.3847/0004-637X/825/1/37

Cheng, X., Ding, M. D., Zhang, J., Sun, X. D., Guo, Y., Wang, Y. M., et al. (2014). *ApJ* 789, 93. doi:10.1088/0004-637X/789/2/93

Cranmer, S. R. (2009). Coronal holes. *Living Rev. Sol. Phys.* 6, 3. doi:10.12942/lrsp-2009-3

Dissauer, K., Veronig, A. M., Temmer, M., Podladchikova, T., and Vanninathan, K. (2018). On the detection of coronal dimmings and the extraction of their characteristic properties. *ApJ* 855, 137. doi:10.3847/1538-4357/aaadb5

Downs, C., Török, T., Titov, V., Liu, W., Linker, J., and Mikic, Z. (2015). *AAS/AGU triennial earth-sun summit*, 1. Indianapolis, IN: AAS/AGU Triennial Earth-Sun Summit. 304.01.

Dudík, J., Janvier, M., Aulanier, G., Del Zanna, G., Karlický, M., Mason, H. E., et al. (2014). SLIPPING MAGNETIC RECONNECTION DURING AN X-CLASS SOLAR FLARE OBSERVED BY*SDO*/AIA. *ApJ* 784, 144. doi:10.1088/0004-637X/784/2/144

Fisher, G. H., Canfield, R. C., and McClymont, A. N. (1985). Flare loop radiative hydrodynamics - Part Six - chromospheric evaporation due to heating by nonthermal electrons. *ApJ* 289, 425. doi:10.1086/162902

Fletcher, L., and Hudson, H. (2001). *SoPh* 204, 69–89. doi:10.1023/A:1014275821318

Fletcher, L., and Hudson, H. S. (2008). Impulsive phase flare energy transport by large-scale Alfvén waves and the electron acceleration problem. *ApJ* 675, 1645–1655. doi:10.1086/527044

Forbes, T. G., and Isenberg, P. A. (1991). A catastrophe mechanism for coronal mass ejections. *ApJ* 373, 294. doi:10.1086/170051

Gou, T., Liu, R., Veronig, A. M., Zhuang, B., Li, T., Wang, W., et al. (2023). Complete replacement of magnetic flux in a flux rope during a coronal mass ejection. *Nat. Astron.* 7, 815–824. doi:10.1038/s41550-023-01966-2

Harra, L. K., Hara, H., Imada, S., Young, P., Williams, D. R., Sterling, A. C., et al. (2007). *PASJ* 59–S801. doi:10.1093/pasj/59.sp3.S801

Harvey, K. L., and Recely, F. (2002). *SoPh* 211, 31–52. doi:10.1023/A:1022469023581

Isenberg, P. A., and Forbes, T. G. (2007). A three-dimensional line-tied magnetic field model for solar eruptions. *ApJ* 670, 1453–1466. doi:10.1086/522025

Kazachenko, M. D., Lynch, B. J., Welsch, B. T., and Sun, X. (2017). A database of flare ribbon properties from the solar dynamics observatory. I. Reconnection flux. *ApJ* 845, 49. doi:10.3847/1538-4357/aa7ed6

Kliem, B., Lin, J., Forbes, T. G., Priest, E. R., and Török, T. (2014). *ApJ* 789, 46. doi:10.1088/0004-637X/789/1/46

Kliem, B., and Török, T. (2006). Torus instability. *PhRvL* 96, 255002. doi:10.1103/PhysRevLett.96.255002

Lemen, J. R., Title, A. M., Akin, D. J., Boerner, P. F., Chou, C., Drake, J. F., et al. (2012). The atmospheric imaging assembly (AIA) on the solar dynamics observatory (SDO). *SoPh* 275, 17–40. doi:10.1007/s11207-011-9776-8

Longcope, D. W. (2014). *ApJ* 795, 10. doi:10.1088/0004-637X/795/1/10

Longcope, D. W., and Forbes, T. G. (2014). Breakout and tether-cutting eruption models are both catastrophic (sometimes). *SoPh* 289, 2091–2122. doi:10.1007/s11207-013-0464-8

Lowder, C., Qiu, J., and Leamon, R. (2017). Coronal holes and open magnetic flux over cycles 23 and 24. *SoPh* 292, 18. doi:10.1007/s11207-016-1041-8

Mandrini, C. H., Nakwacki, M. S., Attrill, G., van Driel-Gesztelyi, L., Démoulin, P., Dasso, S., et al. (2007). *SoPh* 244, 25–43. doi:10.1007/s11207-007-9020-8

Moore, R. L., Sterling, A. C., Hudson, H. S., and Lemen, J. R. (2001). Onset of the magnetic explosion in solar flares and coronal mass ejections. *ApJ* 552, 833–848. doi:10.1086/320559

O'Dwyer, B., Del Zanna, G., Mason, H. E., Weber, M. A., and Tripathi, D. (2010). SDO/AIA response to coronal hole, quiet Sun, active region, and flare plasma. *A&A* 521, A21. doi:10.1051/0004-6361/201014872

Patsourakos, S., Vourlidas, A., Török, T., Kliem, B., Antiochos, S. K., Archontis, V., et al. (2020). Decoding the pre-eruptive magnetic field configurations of coronal mass ejections. *SSRv* 216, 131. doi:10.1007/s11214-020-00757-9

Pesnell, W. D., Thompson, B. J., and Chamberlin, P. C. (2012). The solar dynamics observatory (SDO). *SoPh* 275, 3–15. doi:10.1007/s11207-011-9841-3

Qiu, J., and Cheng, J. (2017). Gradual solar coronal dimming and evolution of coronal mass ejectionin the early phase. *ApJL* 838, L6. doi:10.3847/2041-8213/aa6798

Qiu, J., Hu, Q., Howard, T. A., and Yurchyshyn, V. B. (2007). On the magnetic flux budget in low-corona magnetic reconnection and interplanetary coronal mass ejections. *ApJ* 659, 758–772. doi:10.1086/512060

Qiu, J., Liu, W., Hill, N., and Kazachenko, M. (2010). *ApJ* 725, 319–330. doi:10.1088/0004-637X/725/1/319

Qiu, J., and Longcope, D. W. (2016). *ApJ* 820, 14. doi:10.3847/0004-637X/820/1/14

Qiu, J., Sturrock, Z., Longcope, D. W., Klimchuk, J. A., and Liu, W.-J. (2013). *ApJ* 774, 14. doi:10.1088/0004-637X/774/1/14

Qiu, J., Wang, H., Cheng, C. Z., and Gary, D. E. (2004). Magnetic reconnection and mass acceleration in flare–coronal mass ejection events. *ApJ* 604, 900–905. doi:10.1086/382122

Rempel, M., Chintzoglou, G., Cheung, M. C. M., Fan, Y., and Kleint, L. (2023). Comprehensive radiative MHD simulations of eruptive flares above collisional polarity inversion lines. *ApJ* 955, 105. doi:10.3847/1538-4357/aced4d

Simões, P. J. A., Reid, H. A. S., Milligan, R. O., and Fletcher, L. (2019). The spectral content of SDO/AIA 1600 and 1700 Å filters from flare and plage observations. *ApJ* 870, 114. doi:10.3847/1538-4357/aaf28d

Sterling, A. C., and Hudson, H. S. (1997). [ITAL]Yohkoh[/ITAL] SXT observations of X-ray "dimming" associated with a halo coronal mass ejection. *ApJL* 491, L55–L58. doi:10.1086/311043

Temmer, M., Thalmann, J. K., Dissauer, K., Veronig, A. M., Tschernitz, J., Hinterreiter, J., et al. (2017). On flare-CME characteristics from sun to earth combining remote-sensing image data with *in situ* measurements supported by modeling. *SoPh* 292, 93. doi:10.1007/s11207-017-1112-5

Thompson, B. J., Plunkett, S. P., Gurman, J. B., Newmark, J. S., St. Cyr, O. C., and Michels, D. J. (1998). *SOHO*/EIT observations of an Earth-directed coronal mass ejection on May 12, 1997. *Geophys. Res. Lett.* 25, 2465–2468. doi:10.1029/98GL50429

Thompson, B. J., Qiu, J., Lugaz, N., and Webb, D. F. (2021). "Solar flares and coronal mass ejections," in *Solar physics and solar wind*. Editors N. E. Raouafi, and A. Vourlidas, 1, 179–220. doi:10.1002/9781119815600.ch5

Török, T., and Kliem, B. (2005). Confined and ejective eruptions of kink-unstable flux ropes. *ApJL* 630, L97–L100. doi:10.1086/462412

Vievering, J. T., Vourlidas, A., Zhu, C., Qiu, J., and Glesener, L. (2023). Evolution of solar eruptive events: investigating the relationships among magnetic reconnection, flare energy release, and coronal mass ejections. *ApJ* 946, 81. doi:10.3847/1538-4357/acbe3d

Wang, W., Zhu, C., Qiu, J., Liu, R., Yang, K. E., and Hu, Q. (2019). Evolution of a magnetic flux rope toward eruption. *ApJ* 871, 25. doi:10.3847/1538-4357/aaf3ba

Webb, D. F., Lepping, R. P., Burlaga, L. F., DeForest, C. E., Larson, D. E., Martin, S. F., et al. (2000). The origin and development of the May 1997 magnetic cloud. *J. Geophys. Res.* 105, 27251–27259. doi:10.1029/2000JA000021

Zhu, C., Qiu, J., Liewer, P., Vourlidas, A., Spiegel, M., and Hu, Q. (2020). *ApJ* 893, 141. doi:10.3847/1538-4357/ab838a

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 04 September 2024
DOI 10.3389/fspas.2024.1418918

Check for updates

OPEN ACCESS

EDITED BY
Daniel Okoh,
The National Space Research and Development Agency (NASRDA), Nigeria

REVIEWED BY
Sampad Kumar Panda,
KL University, India
Yun Gong,
Wuhan University, China

*CORRESPONDENCE
Zan-Yang Xing,
xingzanyang@sdu.edu.cn
Qing-He Zhang,
zhangqinghe@sdu.edu.cn

RECEIVED 17 April 2024
ACCEPTED 09 August 2024
PUBLISHED 04 September 2024

CITATION
Qiao F, Xing Z-Y, Zhang Q-H, Zhang H-B, Zhang S-R, Wang Y, Ma Y-Z, Zhang D, Lu S and Varghese M (2024) A long time-series forecasting informer architecture-based ionospheric foF2 model in the low-latitude region.
Front. Astron. Space Sci. 11:1418918.
doi: 10.3389/fspas.2024.1418918

COPYRIGHT
© 2024 Qiao, Xing, Zhang, Zhang, Zhang, Wang, Ma, Zhang, Lu and Varghese. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

A long time-series forecasting informer architecture-based ionospheric foF2 model in the low-latitude region

Feng Qiao[1,2], Zan-Yang Xing[2]*, Qing-He Zhang[2,3]*, Hong-Bo Zhang[4], Shun-Rong Zhang[5], Yong Wang[2], Yu-Zhang Ma[2], Duan Zhang[2,3], Sheng Lu[2] and Manu Varghese[2]

[1]School of Information Science and Engineering, Shandong University, Qingdao, China, [2]Shandong Provincial Key Laboratory of Optical Astronomy and Solar-Terrestrial Environment, Institute of Space Sciences, Shandong University, Weihai, China, [3]State Key Laboratory of Space Weather, Center for Space Science and Applied Research, Chinese Academy of Sciences, Beijing, China, [4]China Research Institute of Radiowave Propagation, Qingdao, China, [5]MIT Haystack Observatory, Westford, MA, United States

Deep learning models have made great accomplishments in space weather forecasting. The critical frequency of the ionospheric F2 layer (foF2) is a key ionospheric parameter, which can be understood and predicted by some advanced new deep learning technologies. In this paper, we utilized an Informer architecture model to predict foF2 for several hours up to 48 h and analyzed its variations during periods of quiet, moderate, and intense geomagnetic conditions. The Informer method forecasts the temporal variations of foF2 by processing and training the past and present foF2 data from the Haikou station, China, during 2006–2014. It is evident that the Informer–foF2 model achieves better prediction performance than the widely used long short-term memory model. The Informer–foF2 model captures the correlation features within the foF2 time series and better predicts the variations ranging for hours up to days during different geomagnetic activities.

KEYWORDS

Informer, foF2, ionosphere, long short-term memory, long sequence time-series forecasting

1 Key Points

- An Informer architecture-based model is used to forecast ionospheric foF2 at low latitudes.
- The Informer–foF2 model has advantages in predicting variations from several hours up to 48 h.
- The Informer–foF2 model forecast for geomagnetic storms is in good agreement with the observations.

2 Introduction

The F2 layer of the ionosphere has the highest degree of ionization, which is closely linked to the Global Positioning System (GPS) and other navigation systems as well as

long-range, high-frequency (HF) communications. The critical frequency of F2 layer (foF2), as one of the most important parameters of radio science, controls the electromagnetic wave propagation through the ionosphere. Therefore, accurate prediction of foF2 is important and a difficult problem in space weather forecasting, especially for long time series at low latitudes (Cander et al., 1998; Rao et al., 2018; Zhang et al., 2022; Fan et al., 2019). There are many precedents for researchers to combine ionospheric phenomena with deep learning (Chen et al., 2019; Zhou et al., 2021; Wang et al., 2022; Yang et al., 2024; Hu and Zhang, 2018). Although many models could simulate and predict the ionospheric F2 layer variations, the complex changes in solar activities make the spatial and temporal variability of foF2 difficult to be predicted and even more difficult to achieve the expected performance at low latitudes. Other ionosphere models such as the International Reference Ionosphere (IRI) (empirical model) (Wichaipanich et al., 2017), random forest (machine learning model), and autoregressive integrated moving average model (ARIMA, time-series model) cannot predict the sudden changes in foF2 caused by a geomagnetic storm. These models do not perform well in long-term forecasting since they are restricted by long-range dependencies. In order to achieve improvement in the forecasting accuracy over a longer time span, deep learning techniques are implemented to forecast foF2 for a sufficiently long duration. Long sequence time-series forecasting means the forecast of high-resolution ionospheric parameters continuously ahead of a few days.

There are several studies on ionospheric modeling using the basic deep learning approach of artificial neural networks (ANNs), which began in the mid-1990s (Willscroft and Poole, 1996; Altinay et al., 1997; Sai Gowtam and Tulasi Ram, 2017; Wang et al., 2013; Kim et al., 2020; Moon et al., 2020; Li et al., 2021; Zhou et al., 2021). Williscroft and Poole (1996) used 10 years of foF2 data observed from the Grahamstown ionosonde and the sunspot number and Ap index to train a simple ANN. Altinay et al. (1997) developed a 1-h prediction model for foF2 using 10 years of foF2 data observed from the Poitiers ionosonde in Central Europe and the Kp index to train a multilayer perceptron model. Since these ANN-based models do not consider past data for the current specific period, the predictive performance cannot be applicable to the phenomena affected by the prior states that are older than the specified period and are naturally not competent for long time-series prediction tasks.

To overcome the shortcomings of the above algorithms, the long short-term memory (LSTM) algorithm, a technique that remembers past data and reflects it in predictions, was developed (Ergen and Kozat, 2018). Earlier, foF2 and hmF2 parameters were predicted by Moon et al. (2020) using LSTM models. Kim et al. (2020) used a physics-based model, which could predict up to 24 h of mid-latitude ionospheric data by inputting the parameters predicted by Moon et al. (2020). Although the LSTM-based model showed reasonably good prediction performance on geomagnetically quiet days, the model predictions are not correlated well during geomagnetically disturbed periods (Kim et al., 2021). The current LSTM model significantly improves the prediction accuracy and time span of the long time series, but its response to abrupt changes was not sufficiently rapid due to the retention of memory states (Lissa et al., 2020).

The LSTM models incorporating the attention module have been applied in long-term time-series prediction studies. The attention mechanism offers the possibility to focus on the response triggered by a certain factor Xia et al., 2022, such as the foF2 anomalies caused by geomagnetic perturbations (Liu and Guo, 2019; Rao et al., 2021). Tang et al. (2023) combined BiLSTM with Attention to predict foF2. With the widespread use of the transformer model (Vasmani et al., 2017), the fully attention-based model is applied to more and more fields, and its excellent performance has led to its use for solving long sequence time-series forecasting (LSTF) problems, but the computational complexity limitation of the transformer itself has led to unsatisfactory results. The transformer-like model Informer proposed by Zhou et al. (2021), which uses the *Probsparse* self-attention mechanism, has shown excellent capability in the long-term time-series prediction problem, and at the same time, the high computational speed, low complexity, and high accuracy demonstrated by the Informer are very attractive. The long-term time-series prediction of the ionospheric foF2 has been implemented in this study based on the foF2 data from the Haikou station at a low-latitude region. We achieved better long-term forecasting performance and avoided self-correction due to geomagnetic activities (Gao et al., 2020). We present the prediction results of foF2 at low latitudes using the Informer model and discuss its advantages and limitations. The Informer-based model shows great potential in solving LSTF problems and saving computing resources. From an experimental standpoint, the Informer-based model has excellent performance in time-series prediction and has great value for further research and application. In this study, we refer to the 48-h prediction as a long-term time-series prediction and attempt to find the most suitable model for long-term time-series prediction on foF2. The models constructed in this paper are applicable to the quiet, moderate, and intense geomagnetic periods, and the comparisons of these models in different periods are discussed.

3 Informer model architecture

Due to the significant correlation with the foF2 time series, recurrent neural networks (RNNs) have been extensively used in foF2 prediction. However, typical RNNs are limited in their long-term dependence due to gradient issues, which do not represent abrupt events. The transformer module by using self-attention can solve those problems to some extent, but the module is limited by computation complexity and error accumulation in the decoding process directly in long sequence time-series forecasting. So the Informer–foF2 model is developed to perform foF2 forecasting using *ProbSparse* self-attention to simplify the calculation and the generative decoder to output the forecast results directly (Bi et al., 2022).

The Informer model was proposed by Zhou et al., in 2021, and we fine-tuned the model to make it better at predictive tasks of foF2. Figure 1 shows the Informer–foF2 architecture. It has two critical components denoted as encoder and decoder modules. The encoder converts input information into a dense vector of fixed dimensionality and extracts features from elements to generate feature mapping. Inversely, the

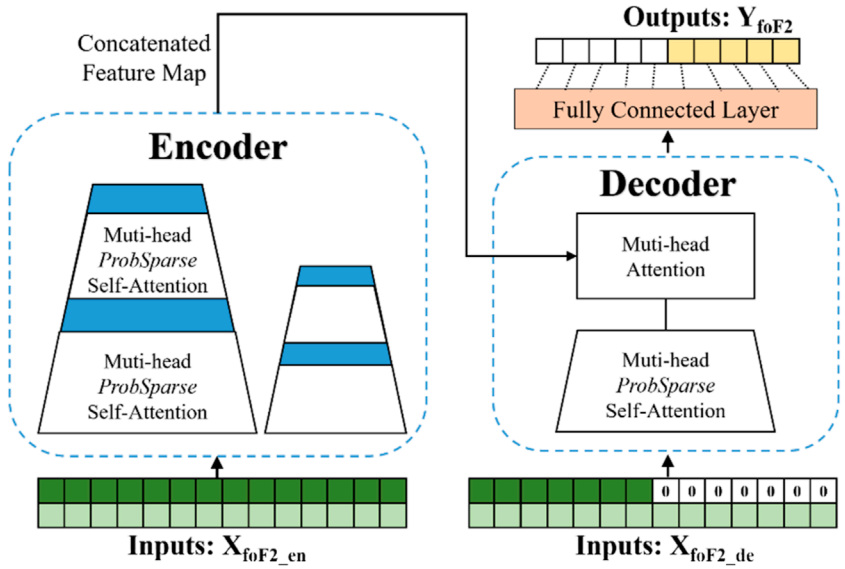

FIGURE 1
Informer–foF2 architecture.

decoder combines the information and feature mapping to the predicted outputs.

ProbSparse self-attention differs in finding a more active query to simplify corresponding attention calculation from traditional multi-head self-attention. The sparsity measurement of *ProbSparse* self-attention is shown below Equation 1.

$$\overline{M}(q_i, K) = \max j\left\{\frac{q_i k_j^T}{\sqrt{d}}\right\} - \frac{1}{L_K}\sum_{j=1}^{L_K}\frac{q_i k_j^T}{\sqrt{d}}, \tag{1}$$

where L_k is 1/q $(k_j \mid q_i)$, the i times query of sparsity measurement is defined as q_i; and q_i and k_j from Q and K, generated in the encoder training process, represent query and key vectors, respectively. *ProbSparse* self-attention calculates the corresponding attention between the active queries and keys, replacing the other by a uniform distribution. Attention calculation is defined as follows Equation 2, where $\overline{Q}$ is a sparse matrix and it contains the sparsity measurement:

$$\overline{A}(Q, K, V) = Softmax\left(\frac{\overline{Q}K^T}{\sqrt{d}}\right)V. \tag{2}$$

The distilling block is using 1-d max pooling operation to accomplish feature downsampling in order to make encoder feature extraction quick and simple. The *ProbSparse* self-attention and distilling operations alternately stack. The encoder output feature mapping is acquired by a two-channel stack.

For the decoder architecture, the generative decoder generates all predicted outputs at once to replace the transformer's decoder, in which it avoids the time-consuming dynamic decoding process in the encoder–decoder architecture. A decoder is composed of two decoder layers, each with a self-attention, a cross-attention, a three-layer norm, and a dropout. The decoder input is a truncation of the later part of the encoder input and a matrix with the same shape as the predicted target. After passing through the decoder, each placeholder (position to be predicted) has a vector, which is then input into a fully connected layer to obtain the predicted results.

A detailed explanation of the Informer architecture is given by Zhou et al. (2021). The specific model composition of the Informer for the foF2 prediction is shown in Table 1.

The LSTM-based model and IRI will appear as contrast models to demonstrate the advantages of the Informer-based model. The LSTM model is a widely used time-series forecasting model in various fields. In previous studies, LSTM has been shown to significantly improve model prediction RMSE (root mean squared error) over empirical models such as IRI and pure time-series models such as ARIMA. LSTM is an excellent model for time-series problems with multivariate effects, especially for time-series data prediction tasks such as foF2, where the fluctuations are severe and perturbations are diverse. The principle of the BiLSTM model is similar to that of the LSTM model, and the improvement is less significant than that of the Informer. This work mainly discusses the effect of the self-attention model on the long-term time-series prediction of foF2, so LSTM is chosen as the representative model to verify the effect of the Informer. The LSTM used in this comparison has been tuned for several experiments and can basically reach the average level of the LSTM prediction models for low-latitude regions, as found in previously published results by several researchers. In order to better demonstrate the model enhancement, we also compare the model outputs with the commonly used empirical model, IRI-2016. We trained 1-h, 5-h, 12-h, 24-h, and 48-h models based on LSTM and Informer. In addition, we perform the same procedure for the corresponding IRI output for foF2 values at the Haikou station (20.0° N and 110.1° E), China, and use these outputs for comparisons with the actual foF2 measurements from Haikou.

TABLE 1 Specific model composition of the Informer.

Encoder			N
Inputs	1×3Conv1d	Embedding (d = 512)	4
ProbSparse Self-attention Block	Multi-head *ProbSparse* attention (h = 8)		
	Add, LayerNorm, dropout (p = 0.05)		
	Pos-wise FNN (d_{inner} = 2048), GELU		
	Add, LayerNorm, Dropout (p = 0.05)		
Distilling	1×3Conv1d, ELU		
	Max pooling (stride = 2)		
Decoder			**N**
Inputs	1×3Conv1d	Embedding (d = 512)	2
Masked PSB	Add mask on attention block		
ProbSparse Self-attention Block	Multi-head *ProbSparse* attention (h = 8)		
	Add, LayerNorm, Dropout (p = 0.05)		
	Pos-wise FNN (d_{inner} = 2048), GELU		
	Add, LayerNorm, Dropout (p = 0.05)		
Batch-Size	32		
Epochs	50		

4 Data and model inputs

The ionospheric foF2 measurements are manually scaled from the ionograms of the Haikou station (20.0° N and 110.1° E) in 1-h resolution. The solar radio flux F10.7 is available from the Space Weather website https://www.spaceweather.gc.ca/index-en.php in 1-day resolution. The geomagnetic indices such as Dst and Kp are available from the Kyoto World Data Center, https://wdc.kugi.kyoto-u.ac.jp/dstae/index.html, having resolutions of 1 h and 3 h, respectively.

The foF2 data from 1 January 2006 to 31 May 2014 are the input and target of the Informer–foF2 model. In order to ensure the input data continuity of the time-series model, the missing data of foF2 have been interpolated by the linear interpolation method. The model output at the time of the geomagnetic disturbances is generated considering F10.7, Dst, and Kp indices. The data have been processed at 1-h intervals and are shown in Figure 2. Due to foF2 exhibiting outstanding time-auto correlation, the 24-h prediction window was chosen for the study. We use the foF2 of the past 96 h to predict the foF2 of the future 48 h, which are all multiples of 24.

In order to make the Informer–foF2 model fit better with the prepared training datasets, we changed the numbers of the encoder and decoder layers, multi-head attention, dropout, minimum batch size, and embedding to optimize the configuration of the proposed model. The total number of foF2 samples is split as 70% for training, 10% for validation, and 20% for testing and prediction.

To quantitatively evaluate the performance of our model, we calculated the RMSE (root mean square error), MAE (mean absolute error), MSE (mean square error), MAPE (mean absolute percentage error), and MSPE (mean squared percentage error) as evaluating indicators. The calculation methods are as follows: $RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2}$, $MSE = \frac{1}{n}\sum_{i=1}^{n}(\widehat{y}_i - y_i)^2$, $MAE = \frac{1}{n}\sum_{i=1}^{n}|\widehat{y}_i - y_i|$, $MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\widehat{y}_i - y_i}{y_i}\right|$, and $MSPE = \frac{100\%}{n}\sum_{i=1}^{n}\left(\frac{\widehat{y}_i - y_i}{y_i}\right)^2$. Here, y_i is the ground truth and $\widehat{y}_i$ is the model prediction. The primary reference metric is RMSE.

5 Results

Table 2 summarizes the performance of the models used in the study. The Informer–foF2 model achieves the best performance among different models. The results have shown that the Informer–foF2 model shows better predictive performance for lower RMSE by approximately 17%, 47%, 29%, 35%, and 27% compared to the LSTM–foF2 model. Moreover, the Informer model significantly outperforms the best level of LSTM models, as shown in Table 2. Figure 3 shows the RMSE of LSTM–foF2 and Informer–foF2 models. Informer has a significant improvement in prediction accuracy and is more accurate for long-term (5–48 h) forecasting. It is more suitable for long-term forecasting than LSTM and is far superior to the IRI-2016 model. Therefore the IRI-2016 model will not be referenced for comparison in the latter picture.

Figure 4 shows the RMSE of Informer–foF2 for each hour during the entire 48-h prediction. The forecast errors corresponding to the 24th hour and 48th hour are marked in red, with a significant decrease compared to other points in the forecast. It is obvious that as time goes on, the errors increase gradually but not linearly; after approximately 15 h, the increase in RMSE becomes gentle. It shows that the Informer–foF2 model has enormous advantages in long-term time-series prediction, and we believe that the correlation among foF2 data also plays an important role in this work. The results have also shown clear 24-h diurnal periodicities at the Haikou station.

Figure 5 shows that the diurnal variations of foF2 are contained within 2-day forecasts for the low-latitude region, where the predictive output of the Informer and LSTM models is benchmarked against actual measurement values. The blue solid line denotes the forecasts by the Informer model, the black dashed line indicates LSTM outputs, and the red dotted line is the actual foF2 measurements. During the spring equinox, the prediction curve of the Informer-48 model aligns more closely with the actual values. At the summer solstice, the LSTM-48 model prediction deviates significantly at approximately 10 h, and in the latter half of the prediction window, both model prediction curves diverge from the actual values to some extent; however, the difference curve shows that the Informer-48 model performs relatively better. Around the autumn equinox, both models perform

10.3389/fspas.2024.1418918

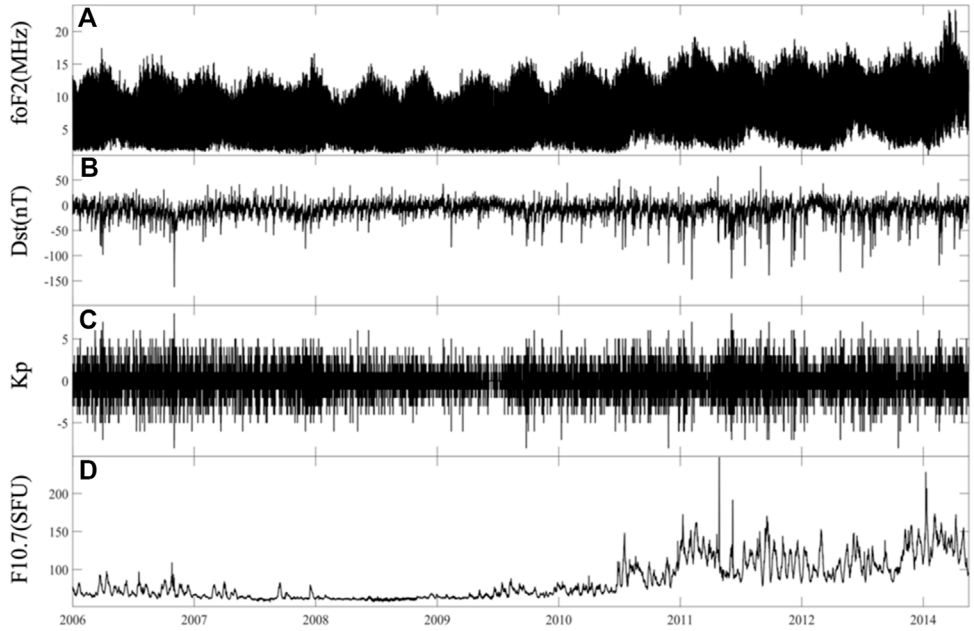

FIGURE 2
Actual values of **(A)** foF2, **(B)** Dst, **(C)** Kp, and **(D)** F10.7 to the Informer–foF2 model input during January 2006–May 2014.

TABLE 2 Accuracy of foF2 forecasting compared to the IRI-2016 model and other LSTM observations.

Model	Predict time	MAE	MSE	RMSE	MAPE	MSPE
IRI-2016	1 h	1.21	2.62	1.62	0.16	0.045
Informer–foF2	1 h	0.61	0.71	0.84	0.07	0.011
	5 h	0.91	1.54	1.24	0.10	0.023
	12 h	1.01	1.86	1.36	0.11	0.032
	24 h	1.04	1.98	1.41	0.12	0.035
	48 h	1.09	2.12	1.45	0.13	0.038
LSTM–foF2	1 h	0.75	1.06	1.03	0.09	0.022
	5 h	1.23	2.81	1.67	0.14	0.047
	12 h	1.29	3.01	1.73	0.15	0.052
	24 h	1.24	2.79	1.67	0.14	0.053
	48 h	1.32	3.17	1.78	0.16	0.064

similarly for the initial 36 h, but beyond this point, the LSTM-48 model prediction curve noticeably deviates from the actual data, demonstrating the Informer-48 model's advantage in longer-term predictions. During the winter solstice, it is observed that the model performance deteriorates, starting approximately 36 h toward the end of the prediction window, with the LSTM-48 model exhibiting significantly more fluctuations than the Informer-48 model. Thus, the Informer-based ionospheric F2 layer critical frequency prediction model exhibits superior performance in low-latitude ionospheric foF2 forecasting, showing substantial advantages over the LSTM model.

In order to show the performance of the Informer–foF2 model in the storm period, two storm events have been selected, which occurred on 17 March 2013 and 28 March 2013 (Figures 6, 7). The storm on 17 March had a Dst minimum of −131 nT (03 LT, 19 March), while the storm on 28 March had −59 nT (20 LT, 28

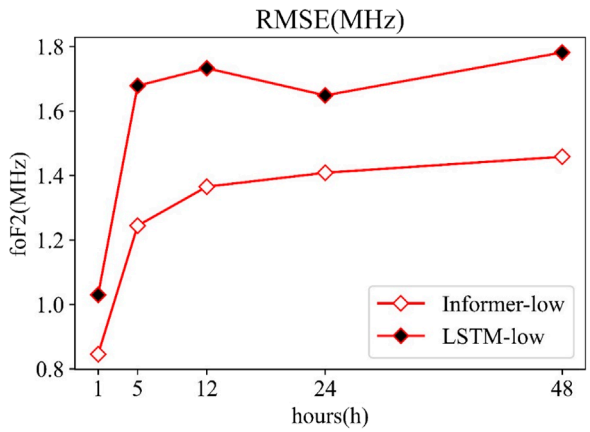

FIGURE 3
RMSE score of each model. The white dots represent different time span prediction models based on Informer, while the black dots represent LSTM-based models.

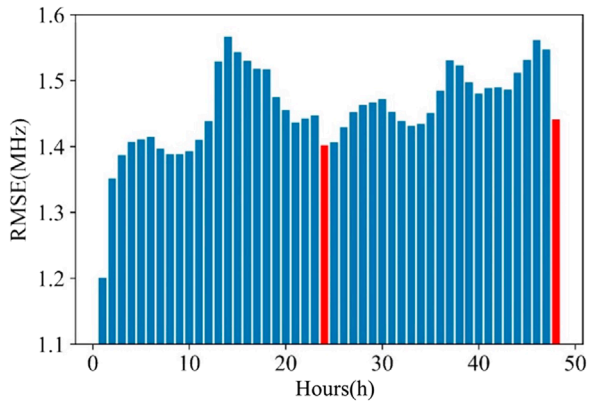

FIGURE 4
Informer–foF2 RMSE of each predicts step. 1 to 48 represent each predicted value for the next 48 h, marked in red at 24 and 48 h.

March) (Shim et al., 2018). The blue line represents the results of the Informer 48-h prediction model of the event (the labels in the figures is Informer-48); the (Rao et al., 2021) black line represents the results of the LSTM 48-h prediction model (the labels in the figures is LSTM-48); the pink and green lines represent the results of the Informer and LSTM 1-h prediction models (the labels in the figures are Informer-1 and LSTM-1), respectively; and the red line is the real foF2 value (the labels in the figures is REAL). The gray region needs special attention as it can more clearly distinguish the performance differences of the models.

As shown in Figure 6, the measured foF2 values at the Haikou station are presented alongside the predicted foF2 values by both the Informer–foF2 and LSTM–foF2 models. During the intense geomagnetic storm, the Dst index plummeted to approximately −120 nT and persisted for an extended period, inducing fluctuations in foF2, which in turn affect the performance of predictive models, particularly those spanning longer durations. It is evident that prior to the storm, the actual measurements closely align with the predictions from the Informer–foF2 model. As the storm commences, both models continue to approximate the actual foF2 values well, and upon examining the differences between the predicted and measured values, the results from the Informer-48 and LSTM-48 models are remarkably similar, suggesting that both the LSTM and Informer models provide a good fit for the data to be forecasted in the initial phase of long sequence time-series predictions.

Notably, as the Dst index hovers at approximately −100 nT for approximately 10 h before continuing its descent, a significant deviation is observed in the LSTM 48-h prediction model from the actual values during this period, as shown in the gray-shaded area 1, whereas the Informer-48 model maintains a good performance. A shortcoming of the 48-h models is their inability to accurately grasp fluctuations that oscillate within smaller time intervals; the attention-based Informer model exhibits a pronounced advantage in unpredictable fluctuation events such as geomagnetic storms. As the storm gradually abates, a brief resurgence in the predictive accuracy of both models is observed. However, it is clear that the predictions from the Informer-48 and LSTM-48 models begin to diverge from the actual ionospheric foF2 values, particularly toward the end of the forecasting window, as depicted in gray-shaded area 2, where both models' predictions substantially veer off the true foF2 curve, with the discrepancies significantly increasing. In addition to the objective influence of time and the cumulative effect of predictive errors, the persistent impact of the storm's incomplete recovery should be considered, with all three factors contributing to the substantial divergence in the predictive models. The aforementioned temporal objective influence refers to the noticeable decline in model performance during nighttime (Feng et al., 2021). Both gray areas 1 and 2 occur during the local time interval of approximately 22:00 to 4:00, where a considerable drop in model performance is evident. Overall, the Informer-48 ionospheric foF2 prediction model exhibits a certain superiority over the LSTM-48 model, clearly closer to predicting the true values.

Figure 7 shows that, compared with the characterization of the strong magnetic storm described in Figure 6, the prediction accuracy of the model for the moderate magnetic storm is significantly closer to the measured value. In the 10 h preceding the forecast, both the Informer-48 and LSTM-48 models exhibit deviations from the actual foF2 measurements, with the Informer-48 predictions more closely mirroring the actual data. At the minimum of the Dst index within gray-shaded area 1, the Informer-48's forecasted values are seen to align almost congruently with the actual foF2 measurements, whereas the LSTM-48's forecasts display a discernible discrepancy, highlighting the LSTM model's difficulties in capturing the abrupt shifts associated with geomagnetic activity. In the waning hours of this event, as denoted by gray-shaded area 2, the Informer-48 model's performance remains superior to that of the LSTM-48 model. This advantage in the predictive capability at the end of the forecast window further underscores the Informer-48 model's proficiency in accurately undertaking long-duration foF2 forecasting tasks.

Figure 8 describes the performance of the low-latitude forecasting model during a geomagnetic quiet period, providing a lucid illustration of the model's capabilities when the perturbations of magnetic storms are absent. At the inception of the forecast, as

10.3389/fspas.2024.1418918

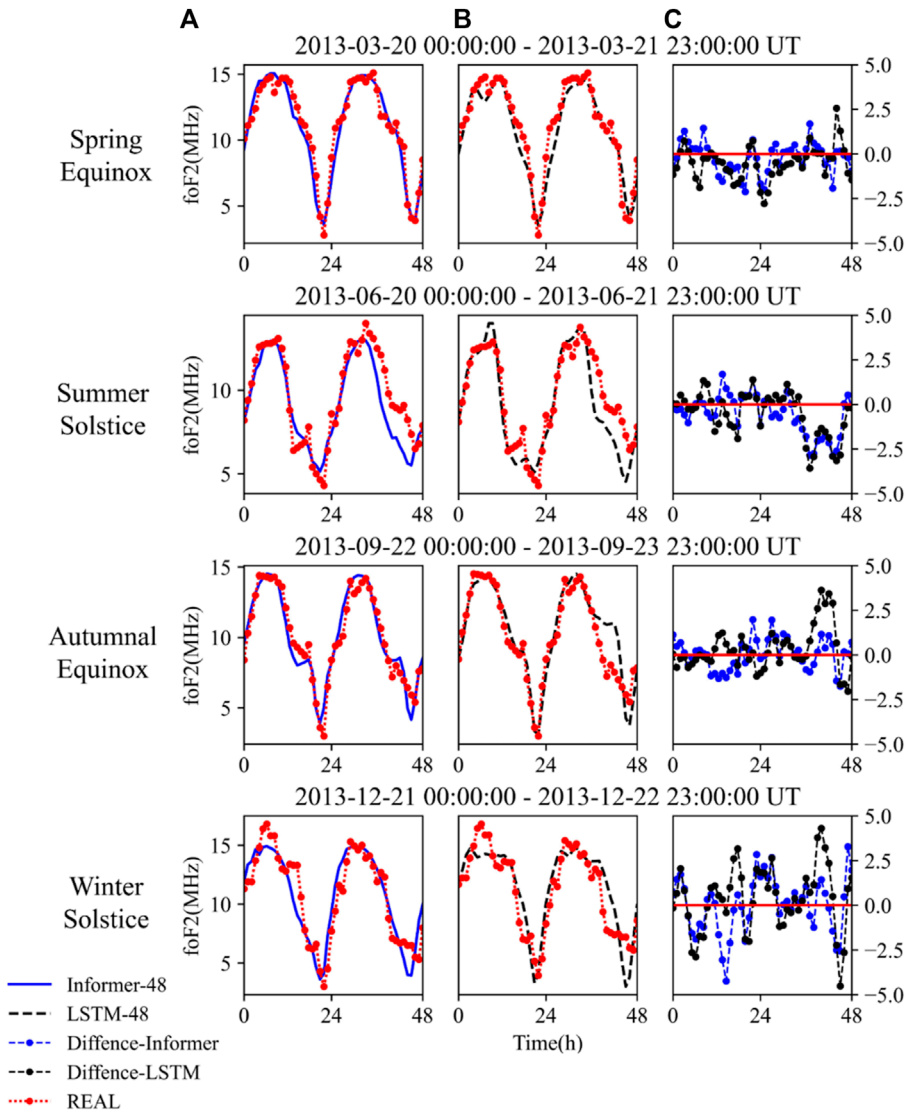

FIGURE 5
Figure shows the comparison between the Informer-48 and LSTM-48 model predictions against the actual values around the spring equinox, summer solstice, autumnal equinox, and winter solstice. **(A)** shows the comparison of the Informer model with the actual observations of foF2. **(B)** shows the comparison between the predicted values of the LSTM model and the measured values of foF2. **(C)** shows the difference between the predicted and measured values of the two models.

highlighted by the gray-shaded area 1 in the diagram, the LSTM-48 model's predictions manifest a notable deviation from the actual measurements. Similarly, within gray-shaded area 2, the LSTM-48 model continues to demonstrate suboptimal performance, starkly contrasted by the Informer-48 model's precise mirroring of the actual foF2 trajectory. These specific intervals correspond to the periods of gradual decline from the day's peak foF2 values, where the LSTM's retention of historical temporal data may be the culprit of the observed predictive inaccuracies. This underscores the advantage of utilizing the one-step generative decoder inherent to the Informer-48 model.

We calculate the RMSE of models, as shown in Figures 6–8, to more visually display the performance of the mode during geomagnetic storm events and quiet period. The results are shown in Table 3. The calmer the geomagnetic activity, the smaller the model prediction error will be. This is because the occurrence of geomagnetic storms is very sudden, and their effects are difficult to predict, and even introducing many parameters to assist in prediction will inevitably have a certain lag, affecting the model prediction results.

Combining various factors, the Informer model demonstrates great performance in the foF2 prediction task, while showing good stability and prediction accuracy in the long duration time series, and it captures general trends of variations during the storm.

6 Conclusion

In this paper, the Informer long-term time-series architecture is implemented to forecast the ionospheric foF2 variation in the

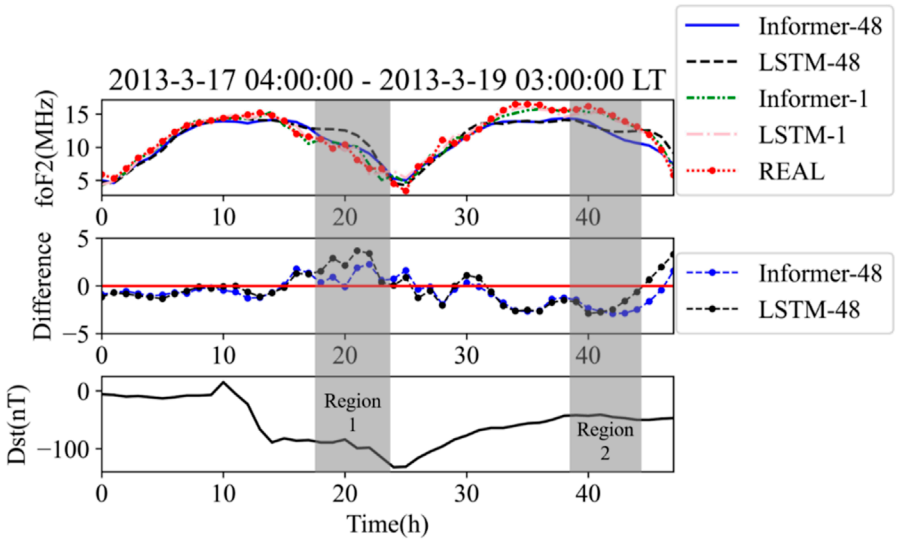

FIGURE 6
Comparison between the observed values and predicted foF2 values during the intense geomagnetic storm on 17 March 2013.

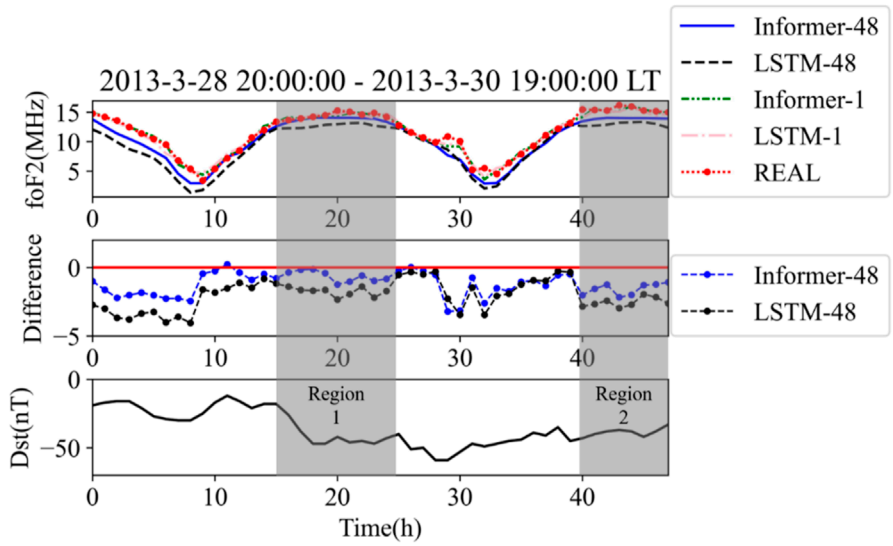

FIGURE 7
Comparison between the observed values and predicted foF2 values during the moderate geomagnetic storm on 28 March 2013.

low latitude. The model input is the ionospheric foF2 data at the Haikou station, China, along with space weather parameters Kp, F10.7, and Dst from 2006 to 2014. Prediction results are compared with the output values of traditional LSTM models and are analyzed in detail during moderate and intense geomagnetic storm events, separately. The overall performance of the model for all the events was discussed, and the conclusion was drawn that the prediction performance of the model is degraded at night. The analysis of events also demonstrated the superior advantage of the informer model in long-term (from hours to within 2 days) sequence prediction tasks. In the measurement examples used for verification, the Informer–foF2–48 h model predicts foF2 parameters with the RMSE of 1.245 MHz and shows that the prediction accuracy maintains stability as the prediction window expands. The RMSE demonstrates that the proposed method performed well in long-term time-series prediction compared with other models and captured some variations in the ionospheric foF2. The prediction results of the proposed Informer–foF2 model provide insights into ionospheric foF2 prediction at low latitudes and long-term time-series prediction from hours up to 2 days. The Informer-based ionospheric foF2 prediction model proved that it can forecast the continuous change in ionospheric foF2 more accurately and more reliably than existing methods.

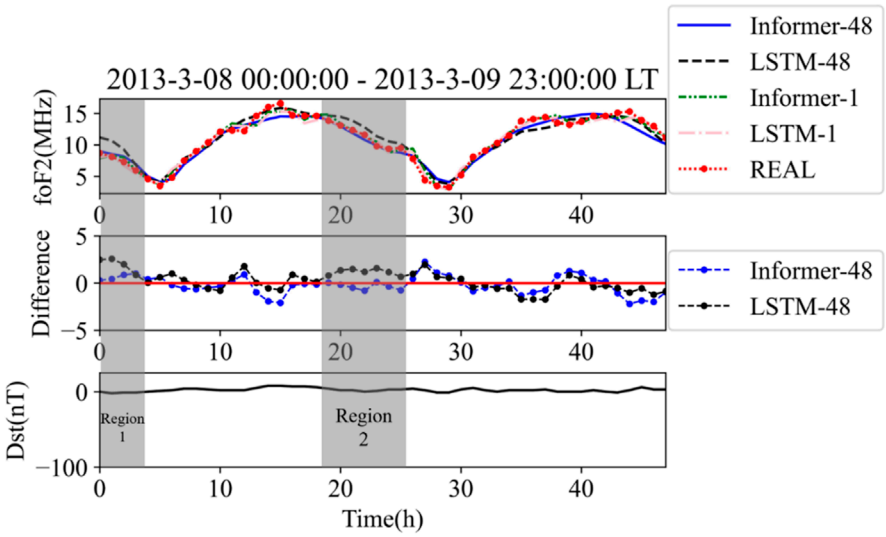

FIGURE 8
Comparison between the observed values and predicted foF2 values during the geomagnetic quiet period.

TABLE 3 RMSE score of 48-h models in events.

Model	RMSE—strong storm	RMSE—medium storm	RMSE—quiet period
LSTM-48	1.68	2.28	1.08
Informer-48	1.46	1.45	0.96
LSTM-1	0.81	0.63	0.71
Informer-1	0.80	0.58	0.69

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

FQ: writing–original draft and writing–review and editing. Z-YX: writing–review and editing. Q-HZ: writing–review and editing. H-BZ: writing–review and editing. S-RZ: writing–review and editing. YW: writing–review and editing. Y-ZM: writing–review and editing. DZ: writing–review and editing. SL: writing–review and editing. VM: writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. The work was supported by the National Natural Science Foundation of China (Grants 42325404, 42120104003, and 42204164), the Stable-Support Scientific Project of China Research Institute of Radiowave Propagation (Grant A132312191), the Chinese Meridian Project, the China Postdoctoral Science Foundation (Grant 2021M701974), and the foundation of National Key Laboratory of Electromagnetic Environment (Grant No. 6142403230102).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Altinay, O., Tulunay, E., and Tulunay, Y. (1997). Forecasting of ionospheric critical frequency using neural networks. *Geophys. Res. Lett.* 24 (12), 1467–1470. doi:10.1029/97gl01381

Bi, C., Ren, P., Yin, T., Zhang, Y., Li, B., and Xiang, Z. (2022). An informer architecture-based ionospheric foF2 model in the middle latitude region. *IEEE Geoscience Remote Sens. Lett.* 19, 1–5. doi:10.1109/lgrs.2022.3160422

Cander, L. R., Milosavljevic, M. M., Stankovic, S. S., and Tomasevic, S. (1998). Ionospheric forecasting technique by artificial neural network. *Electron. Lett.* 34, 1573–1574. doi:10.1049/el:19981113

Chen, Z., Jin, M., Deng, Y., Wang, J., Huang, H., Deng, X., et al. (2019). Improvement of a deep learning algorithm for total electron content maps: image completion. *J. Geophys. Res.* 124 (1), 790–800. doi:10.1029/2018ja026167

Ergen, T., and Kozat, S. S. (2018). Online training of LSTM networks in distributed systems for variable length data sequences. *IEEE Trans. Neural Netw. Learn. Syst.* 29 (10), 5159–5165. doi:10.1109/tnnls.2017.2770179

Fan, J., Liu, C., Lv, Y., Han, J., and Wang, J. (2019). A short-term forecast model of foF2 based on Elman neural network. *Appl. Sci.* 9 (14), 2782. doi:10.3390/app9142782

Feng, J., Zhou, Y., Zhou, Y., Gao, S., Zhou, C., Tang, Q., et al. (2021). Ionospheric response to the 17 March and 22 June 2015 geomagnetic storms over Wuhan region using GNSS-based tomographic technique. *Adv. Space Res.* 67 (1), 111–121. doi:10.1016/j.asr.2020.10.008

Gao, C., Jin, S., and Yuan, L. (2020). Ionospheric responses to the June 2015 geomagnetic storm from ground and LEO GNSS observations. *Remote Sens.* 12 (14), 2200. doi:10.3390/rs12142200

Hu, A., and Zhang, K. (2018). Using bidirectional long short-term memory method for the height of F2 peak forecasting from ionosonde measurements in the Australian Region. *Remote Sens.* 10 (10), 1658. doi:10.3390/rs10101658

Kim, J.-H., Kwak, Y.-S., Kim, Y., Moon, S.-I., Jeong, S.-H., and Yun, J. (2021). Potential of regional ionosphere prediction using a long short-term memory deep-learning algorithm specialized for geomagnetic storm period. *Space weather.* 19, e2021SW002741. doi:10.1029/2021sw002741

Kim, J. H., Kwak, Y. S., Kim, Y. H., Moon, S. I., Jeong, S. H., and Yun, J. Y. (2020). Regional ionospheric parameter estimation by assimilating the LSTM trained results into the Sami2 model. *Space weather.* 18 (10). doi:10.1029/2020sw002590

Li, X., Zhou, C., Tang, Q., Zhao, J., Zhang, F., Xia, G., et al. (2021). Forecasting ionospheric foF2 based on deep learning method. *Remote Sens.* 13 (19), 3849. doi:10.3390/rs13193849

Lissa, D., Srinivasu, V. K. D., Prasad, D. S. V. V. D., and Niranjan, K. (2020). Ionospheric response to the 26 August 2018 geomagnetic storm using GPS-TEC observations along 80° E and 120° E longitudes in the Asian sector. *Adv. Space Res.* 66 (6), 1427–1440. doi:10.1016/j.asr.2020.05.025

Liu, G., and Guo, J. (2019). Bidirectional LSTM with attention mechanism and convolutional layer for text classification. *Neurocomputing* 337, 325–338. doi:10.1016/j.neucom.2019.01.078

Moon, S., Kim, Y. H., Kim, J. H., Kwak, Y. S., and Yoon, J. Y. (2020). Forecasting the ionospheric F2 parameters over jeju station (33.43°N, 126.30°E) by using long short-term memory. *J. Korean Phys. Soc.* 77 (12), 1265–1273. doi:10.3938/jkps.77.1265

Rao, S. S., Chakraborty, M., and Pandey, R. (2018). Ionospheric variations over Chinese EIA region using foF2 and comparison with IRI-2016 model. *Adv. Space Res.* 62 (1), 84–93. doi:10.1016/j.asr.2018.04.009

Rao, T. V., Sridhar, M., Ratnam, D. V., Harsha, P. B. S., and Srivani, I. (2021). A bidirectional long short-term memory-based ionospheric foF2 and hmF2 models for a single station in the low latitude region. *IEEE Geoscience Remote Sens. Lett.* 19 (99), 1–5. doi:10.1109/LGRS.2020.3045702

Sai Gowtam, V., and Tulasi Ram, S. (2017). An artificial neural network-based ionospheric model to predict N_mF_2 and h_mF_2 using long-term data set of FORMOSAT-3/COSMIC radio occultation observations: preliminary results. *J. Geophys. Res. Space Phys.* 122 (11), 11743–11755. doi:10.1002/2017ja024795

Shim, J. S., Tsagouri, I., Goncharenko, L., Rastaetter, L., Kuznetsova, M., Bilitza, D., et al. (2018). Validation of ionospheric specifications during geomagnetic storms: TEC and foF2 during the 2013 March storm event. *Space weather.* 16 (11), 1686–1701. doi:10.1029/2018sw002034

Tang, J., Yang, D., and Ding, M. (2023). Forecasting ionospheric foF2 using bidirectional LSTM and attention mechanism. *Space weather.* 21, e2023SW003508. doi:10.1029/2023SW003508

Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., et al. (2017). Attention is all you need. *Adv. neural Inf. Process. Syst.* doi:10.48550/arXiv.1706.03762

Wang, R., Zhou, C., Deng, Z., Ni, B., and Zhao, Z. (2013). Predicting foF2 in the China region using the neural networks improved by the genetic algorithm. *J. Atmos. Solar-Terr. Phys.* 92, 7–17. doi:10.1016/j.jastp.2012.09.010

Wang, P., Li, J., Han, B., ZeJun, H. U., Gao, X.Bo, Liu, J. J., et al. (2022). Modeling the polar ionospheric convection velocity vectors using shadow neural networks. *Chin. J. Geophys. (in Chinese)* 65 (4), 1197–1213. doi:10.6038/cjg2022P0255

Wichaipanich, N., Hozumi, K., Supnithi, P., and Tsugawa, T. (2017). A comparison of neural network-based predictions of foF2 with the IRI-2012 model at conjugate points in Southeast Asia. *Advances in Space Research* 59 (12), 2934–2950. doi:10.1016/j.asr.2017.03.023

Williscroft, L. A., and Poole, A. W. V. (1996). Neural networks, foF2, sunspot number and magnetic activity. *Geophysical Research Letters* 23 (24), 3659–3662. doi:10.1029/96gl03472

Xia, G., Liu, M., Zhang, F., and Zhou, C. (2022). CAiTST: conv-attentional image time sequence transformer for ionospheric TEC maps forecast. *Remote Sensing* 14 (17), 4223. doi:10.3390/rs14174223

Yang, Q., Xie, M., Su, H., Han, D., and He, Q. (2024). Statistical characteristics of multi-scale auroral arc width based on machine learning. *Journal of Geophysical Research Space Physics* 129, e2023JA031954. doi:10.1029/2023ja031954

Zhang, B., Wang, Z., Shen, Y., Li, W., Xu, F., and Li, X. (2022). Evaluation of foF2 and hmF2 parameters of IRI-2016 model in different latitudes over China under high and low solar activity years. *Remote Sensing* 14 (4), 860. doi:10.3390/rs14040860

Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., et al. (2021). Informer: beyond efficient transformer for long sequence time-series forecasting. *Proc. 35th AAAI conf. artif. intell.* 35, 11106–11115. doi:10.1609/aaai.v35i12.17325

frontiers | Frontiers in Astronomy and Space Sciences

TYPE Original Research
PUBLISHED 14 January 2025
DOI 10.3389/fspas.2024.1510579

Check for updates

OPEN ACCESS

EDITED BY
Rudolf A. Treumann,
Ludwig Maximilian University of Munich, Germany

REVIEWED BY
Masahiro Hoshino,
The University of Tokyo, Japan
Silvia Perri,
University of Calabria, Italy

*CORRESPONDENCE
Xiangliang Kong,
kongx@sdu.edu.cn

RECEIVED 13 October 2024
ACCEPTED 26 December 2024
PUBLISHED 14 January 2025

CITATION
Kong X, Ning H and Chen Y (2025) Modeling the transport and anisotropy of energetic electrons in solar flares.
Front. Astron. Space Sci. 11:1510579.
doi: 10.3389/fspas.2024.1510579

COPYRIGHT
© 2025 Kong, Ning and Chen. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Modeling the transport and anisotropy of energetic electrons in solar flares

Xiangliang Kong[1,2,3]*, Hao Ning[2] and Yao Chen[1,2]

[1]School of Space Science and Physics, Institute of Space Sciences, Shandong University, Weihai, Shandong, China, [2]Institute of Frontier and Interdisciplinary Science, Shandong University, Qingdao, Shandong, China, [3]Yunnan Key Laboratory of Solar Physics and Space Science, Kunming, China

Transport of energetic electrons in the flare loop is important to understanding nonthermal emissions in solar flares. In this work, we model the propagation of electrons by numerically solving the particle transport equation which includes the physics of magnetic mirroring and turbulent pitch-angle diffusion. We find that both the fractions of electrons trapped in the looptop and precipitating into the solar surface display a non-monotonic behavior with increasing scattering rate. In the moderate diffusion regime, the precipitation fraction is highest and we expect intense nonthermal HXR and microwave emissions at the footpoints. With no or weak pitch-angle scattering, the velocity space distribution can be highly anisotropic both in the looptop and loopleg regions. Different patterns of stripes with positive gradients in the perpendicular direction can drive the electron cyclotron maser instability with higher efficiency than the classical loss-cone distribution, facilitating the excitation of coherent solar radio bursts. Our simulation results highlight the effects of turbulent pitch-angle scattering on electron trap/precipitation and anisotropic distribution in solar flares, which may help us understand the precipitation of magnetospheric electrons accounting for the aurora as well.

KEYWORDS

solar flares, energetic electrons, particle transport, solar X-ray emission, solar radio emission

1 Introduction

Solar flares are the most powerful energy-release phenomena on the Sun (e.g., Fletcher et al., 2011; Benz, 2017). A large number of charged particles are accelerated to high energies, including electrons, protons, and heavy ions, which can further excite nonthermal emissions from radio to gamma-rays via different radiation mechanisms. Although the primary acceleration mechanism remains unclear (Miller et al., 1997; Zharkova et al., 2011; Kong et al., 2019; Li et al., 2021), it is generally believed that electrons are accelerated in the coronal region above flare loops. After being injected at the top of flare loops, accelerated electrons travel to the loop footpoints and deposit energy in the high-density chromosphere, resulting in chromospheric evaporation. In some strong flares, energetic particles can deliver energy to the deeper atmosphere and have impact on the photosphere, suggested as the driver of white-light flares and sunquakes (e.g., Wu et al., 2023). Therefore, the acceleration and transport of energetic electrons plays a central role in the solar flare dynamics.

Energetic electrons are not free-streaming and subject to various effects during their transport from the looptop to the footpoints. The transport effects include magnetic mirroring due to the convergence in magnetic field, pitch-angle scattering by magnetic

turbulence, energy loss and pitch-angle scattering via Coulomb collisions, return current, etc (e.g., Fletcher and Martens, 1998; Minoshima et al., 2011; Jeffrey et al., 2014; Kontar et al., 2014; Bian et al., 2017; Effenberger and Petrosian, 2018; Musset et al., 2018; Allred et al., 2020; Tang et al., 2020; Kong et al., 2022). Therefore, the pitch-angle distribution of energetic electrons is time dependent and should not be isotropic. Anisotropy in the velocity space is of critical importance to nonthermal emissions. It plays a fundamental role in plasma wave excitation in coherent emission mechanisms of solar radio bursts (Melrose, 2017). For example, electron-cyclotron maser emission requires a positive gradient of perpendicular direction, such as loss cone and horseshoe distributions (e.g., Melrose andWheatland, 2016; Zhao G. Q. et al., 2016; Ning et al., 2021a; Ning et al., 2021b; Tang et al., 2024). Anisotropic distribution can also affect the intensity, spectrum, and polarization of incoherent emissions, e.g., in microwave (e.g., Kuznetsov and Fleishman, 2021) and X-rays (e.g., Kuznetsov and Fleishman, 2021) and X-rays (e.g., Charikov et al., 2012; Melnikov et al., 2013).

Magnetic turbulence is an essential element both in particle acceleration (e.g., stochastic or shock acceleration) and transport processes in solar flares. Recent observations from nonthermal broadening of spectral lines by Hinode/EIS (e.g., Stores et al., 2021) revealed the presence of turbulence throughout the flare loop, although the strongest is at the looptop. In MHD simulations of magnetic reconnection in solar flares, the impact of reconnection outflows on the flare loop can trigger various instabilities and cause a highly turbulent plasma environment (e.g., Ruan et al., 2023; Wang et al., 2023; Ye et al., 2023). Recently, Effenberger and Petrosian (2018) studied the particle escape time for different initial pitch-angle distributions by solving the Fokker-Planck transport equation and assuming isotropic pitch-angle scattering by magnetic turbulence. Melnikov and Filatov (2020) investigated the conditions for the generation of whistler turbulence in the flare loop, which can resonate with energetic electrons and significantly affect their spectral and pitch-angle distributions (Melnikov and Filatov, 2021).

In this work, we numerically model the propagation of energetic electrons after being injected into the flare loop, and focus on the effects of magnetic mirror and turbulent scattering on the transport and anisotropic distribution of electrons. The paper is organized as follows. Section 2 describes our numerical model and Section 3 presents the simulation results. Summary and discussion are given in Section 4.

2 Numerical model

For the flare loop, we use an analytical two-dimensional magnetic field model in the $x-y$ plane (Lin et al., 1995; Minoshima et al., 2010),

$$A_z(x,y) = -\frac{y+d}{x^2+(y+d)^2} - \frac{y}{(a+d)^2},$$

$$B_x(x,y) = \frac{\partial A_z}{\partial y} = -\left[\frac{x^2-(y+d)^2}{\{x^2+(y+d)^2\}^2} + \frac{1}{(a+d)^2}\right],$$

$$B_y(x,y) = -\frac{\partial A_z}{\partial x} = -\frac{2x(y+d)}{\{x^2+(y+d)^2\}^2},$$

where A_z is the flux function, B_x and B_y are two components of the magnetic field, a is the height of X-type neutral line at the top of flare loops, and d is the depth of the dipole below the photosphere. In Figures 1A, 2, the thin curves are contours of A_z and illustrate the magnetic field lines of the flare loop model. Here we assume a = 80 Mm, d = 50 Mm.

Following our previous work Kong et al. (2022), we model the transport of energetic electrons in the flare loop by numerically solving the focused transport equation (Roelof, 1969; Skilling, 1971; van den Berg et al., 2020). The equation includes various transport effects, such as streaming along the magnetic field, advection with the solar wind, pitch-angle scattering, magnetic focusing/mirroring, and adiabatic cooling. Therefore, it has been widely applied to study the acceleration and transport of solar energetic particles (e.g., Qin et al., 2006; Zhang et al., 2009; Dröge et al., 2010; Zuo et al., 2011; Wang et al., 2012; Zhao L. et al., 2016; Hu et al., 2017; Zhang and Zhao, 2017; Wei et al., 2019; Wijsen et al., 2019). A similar Fokker-Planck transport equation has also been used in modeling energetic electrons in solar flares, in which the effects of magnetic mirroring, Coulomb collisions, and pitch-angle scattering are often included (e.g., Hamilton and Petrosian, 1990; Fletcher, 1995; Kontar et al., 2014; Effenberger and Petrosian, 2018; Melnikov and Filatov, 2021).

In this work, we focus on the effect of pitch-angle scattering on electron trapping/precipitation and the anisotropic distribution of energetic electrons. We neglect the advection term and the energy change due to Coulomb collisions, compression and shear in plasma flow (Kong et al., 2022). Test-particle simulations in synthetic turbulence suggested that the perpendicular diffusion coefficient is a few percent of the parallel diffusion coefficient (Giacalone and Jokipii, 1999). Cross-field diffusion may affect both the size and energy dependence of nonthermal emissions (Kontar et al., 2011), and the escape of electrons to the open field line. Here it is neglected for simplicity. The reduced particle transport equation can be written as (Roelof, 1969; Effenberger and Petrosian, 2018),

$$\frac{\partial f}{\partial t} = -v\mu\hat{\boldsymbol{b}}\cdot\nabla f - \frac{v(1-\mu^2)}{2L_B}\frac{\partial f}{\partial \mu} + \frac{\partial}{\partial \mu}D_{\mu\mu}\frac{\partial f}{\partial \mu},$$

where f is the distribution function of charged particles, v is the particle speed, μ is the pitch-angle cosine, and t is the time. The terms on the right-hand side describe the electron streaming along the direction of magnetic field $\hat{\boldsymbol{b}}$, the magnetic mirroring effect with the focusing length $L_B = (\hat{\boldsymbol{b}}\cdot\nabla lnB)^{-1}$, and the pitch-angle diffusion with a coefficient $D_{\mu\mu}$.

The pitch-angle diffusion coefficient $D_{\mu\mu}$ describes the resonant interaction between the particle and the turbulent magnetic field. In the quasi-linear theory, it is given by (Jokipii, 1971),

$$D_{\mu\mu} = \frac{\pi}{4}\Omega_0(1-\mu^2)\frac{k_r P(k_r)}{B_0^2},$$

where $\Omega_0 = qB_0/m$ is the particle gyrofrequency with the mass m and the charge q, $P(k)$ is the turbulence power spectrum, and $k_r = \Omega_0/(v|\mu|)$ is the resonant wavenumber. We assume the form of Kolmogorov turbulence spectrum with the spectral index Γ = 5/3. In the non-relativistic limit, the pitch-angle diffusion coefficient can be expressed as (Beeck and Wibberenz, 1986),

$$D_{\mu\mu} = D_{\mu\mu 0}\left(\frac{p}{p_0}\right)^{\Gamma-1}(1-\mu^2)\left(|\mu|^{\Gamma-1}+h_0\right).$$

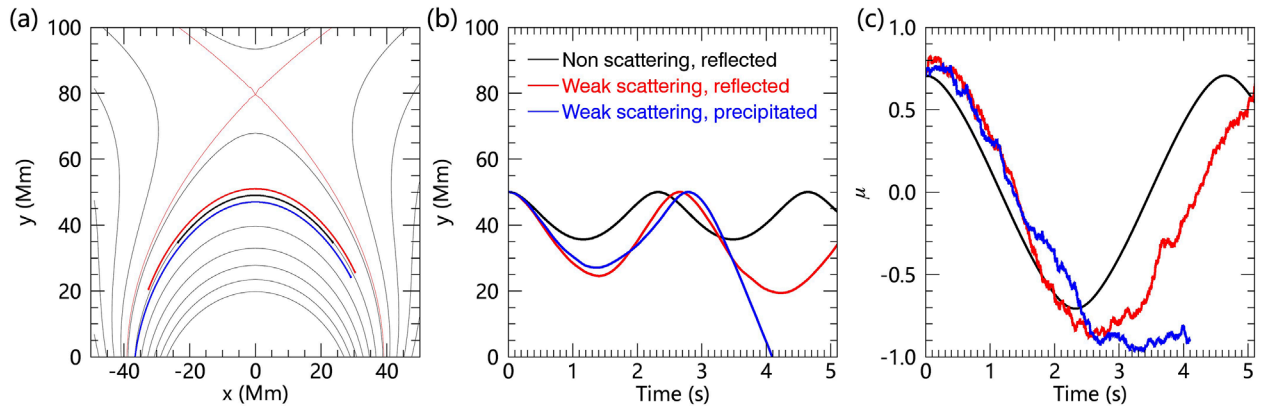

FIGURE 1
Simulation results for three representative electrons. **(A)**: electron trajectories plotted over the magnetic field lines, **(B, C)**: temporal variations of y and μ. The electron without scattering is plotted in black, while the two electrons with weak scattering are plotted in red and blue. Note that in panel **(A)** the red and blue curves are shifted to avoid overlapping.

$D_{\mu\mu 0}$ is a constant describing the scattering rate and depends on the level of magnetic field fluctuation. p_0 is the particle momentum at the energy E_0 = 10 keV. The parameter h_0 is added to describe the finite scattering through $\mu = 0$ and here we set $h_0 = 0.05$.

Because the transport equation is essentially a Fokker-Planck equation, it can be recast into a set of stochastic differential equations (SDEs) (e.g., Zhang, 1999; Strauss and Effenberger, 2017). Here we use the following time-forward SDEs to trace the particle's position and pitch-angle (Kong et al., 2022).

$$d\boldsymbol{X} = v\mu\hat{\boldsymbol{b}}dt,$$

$$d\mu = \left[-\frac{v(1-\mu^2)}{2L_B} + \frac{\partial D_{\mu\mu}}{\partial \mu}\right]dt + \sqrt{2D_{\mu\mu}}dW_\mu(t),$$

where dW_μ is a Wiener process.

In the simulations, we assume that electrons have been accelerated near the top of the flare loop and only consider the transport process in the loop. Energetic electrons with a power-law energy spectrum, $f(E) \sim E^{-\delta}$, are impulsively injected in the looptop region, given by x = [-2, 2] Mm and y = [48, 52] Mm. Here we set the electron energy spectral index $\delta = 3$, and the energy range is between 0.7 and 153 keV (electron velocity between 0.05 c and 0.64 c, c is the speed of light). The initial electron pitch-angle distribution is assumed to be isotropic. In each simulation, a total of 6 million pseudo-particles are injected. We note that the injection of accelerated electrons is not necessarily at the top of the loop in a realistic solar flare. For example, when the reconnection takes place between a closed loop with other loops or an open field line. This may give rise to asymmetric distribution in space and anisotropy of energetic electrons.

To study the effect of turbulent scattering on electron transport and anisotropic distribution, we conduct five simulation runs with different levels of magnetic fluctuations by changing the value of $D_{\mu\mu 0}$. We take $D_{\mu\mu 0} = 0$ (*RunA*), 0.0272 s^{-1} (*RunB*), 0.272 s^{-1} (*RunC*), 2.72 s^{-1} (*RunD*), and 27.2 s^{-1} (*RunE*). Then, the time scale of turbulent scattering is approximately $\tau_d = 1/D_{\mu\mu 0}$, varying between 36.8 s and 0.0368 s from *RunB* to *RunE*. Note that for electrons with the energy E_0 = 10 keV, $v_{e0} = 0.195\ c = 5.85 \times 10^7$ m s^{-1}, and a loop length of L_0 = 100 Mm, the crossing time scale in the loop where they are injected, $\tau_c = L_0/v_{e0} = 1.71$ s.

Three regimes of turbulent pitch-angle diffusion was defined in Bespalov et al. (1987), weak ($\tau_d > \sigma\tau_c$), moderate ($\tau_c < \tau_d < \sigma\tau_c$), and strong ($\tau_d < \tau_c$), where σ is the mirror ratio of the flare loop. For the field lines where electrons are injected, the magnitudes of magnetic field in the looptop and at the footpoint are 40.8 G and 241 G, respectively. Then, the mirror ratio is $\sigma = B_{FP}/B_{LT} = 5.9$, and the critical pitch angle is $\theta_c = arcsin\sqrt{1/\sigma} = 24.3°$. Therefore, *RunB* corresponds to the weak diffusion regime, *RunC* in the moderate diffusion regime, and *RunD* and *RunE* in the strong diffusion regime. We can also calculate the particle mean free path for 10 keV electrons, $\lambda_{\parallel 0} = 3\kappa_{\parallel 0}/v_{e0}$, where $\kappa_{\parallel 0}$ is the spatial diffusion coefficient along the direction of the magnetic field and related to the pitch-angle diffusion coefficient $D_{\mu\mu}$ (Kong et al., 2022). Then, we can get $\lambda_{\parallel 0}$ = 4,100 Mm in *RunB*, much larger than the loop length; $\lambda_{\parallel 0}$ = 410 Mm in *RunC*, comparable to the loop length; $\lambda_{\parallel 0}$ = 41 Mm and 4.1 Mm in *RunD* and *RunE*, smaller than the loop length.

3 Simulation results

To test the validity of the simulation, particularly the pitch-angle scattering through $\mu = 0$, we first examine the trajectory of a single electron. Figure 1 shows the simulation results for three representative electrons, trajectories plotted over the magnetic field lines, variations of y position and μ as a function of time. The three electrons are injected in the same position at the looptop with the same energy of 10 keV and initial pitch-angle of 45°. For the electron without turbulent scattering (as in *RunA*), the curves are plotted in black. It is reflected at y_R = 35.7 Mm, where the magnetic field strength is B_R = 81.7 G. We then can get the critical pitch-angle at the reflection point is $arcsin\sqrt{B_{LT}/B_R} = 45°$, consistent with the initial condition. For the two electrons with weak turbulent scattering (as in *RunB*), the curves are plotted in red and

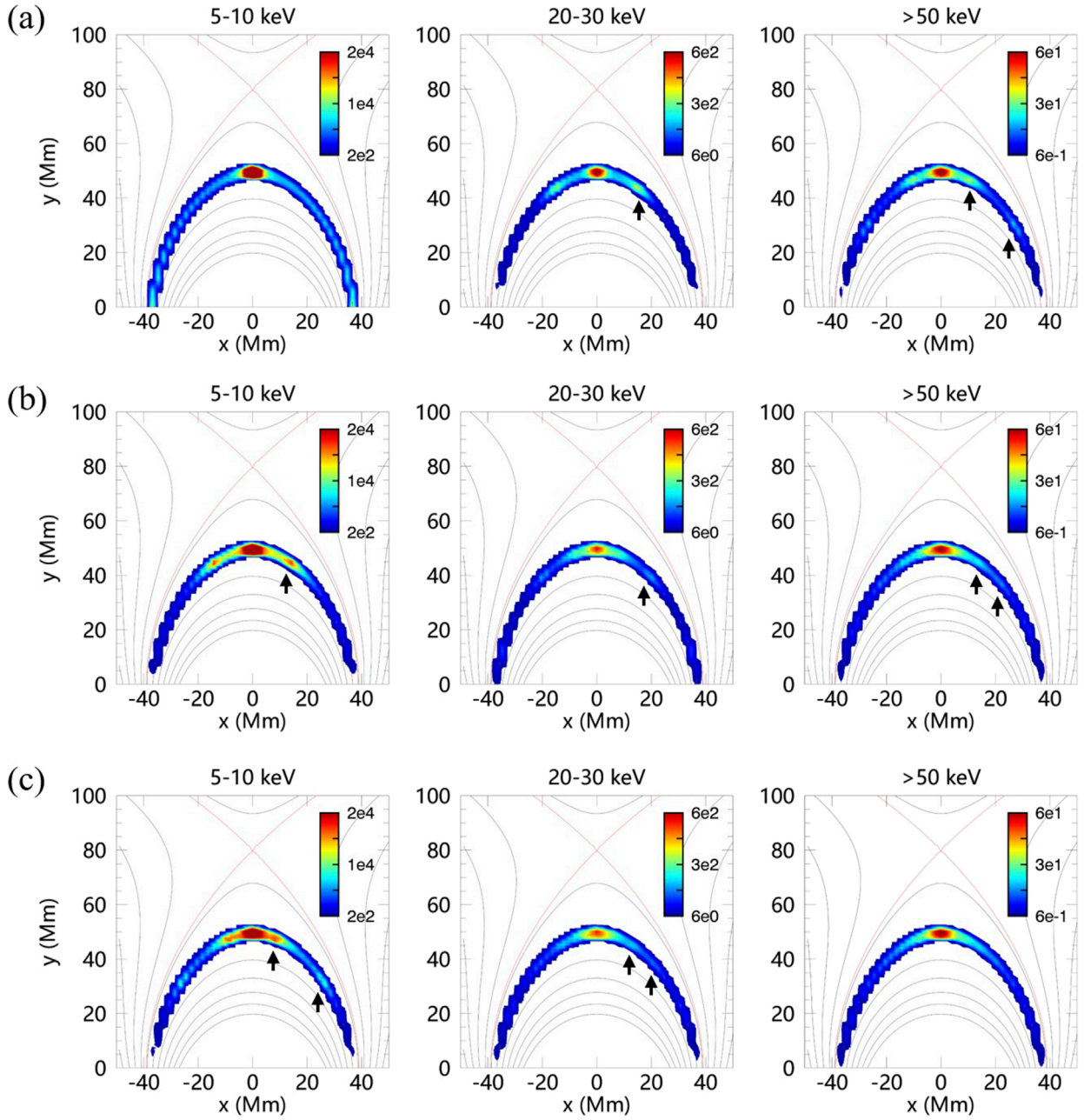

FIGURE 2
Spatial distributions of electrons at three different energies, 5–10 keV, 20–30 keV, > 50 keV, in *RunA*. Panels **(A–C)** are at three simulation times, τ_c **(A)**, $2\tau_c$ **(B)**, and $3\tau_c$ **(C)**, respectively. Black arrows denote streams of electrons bouncing back and forth in the right side of the loop.

blue, respectively. Due to the pitch-angle diffusion, electrons can be scattered into the loss cone. Therefore, the two electrons can go deeper than the expected reflection position y_R. For the electron plotted in blue, it is not reflected while moving to the left footpoint and finally precipitates into the solar surface. As seen from the evolution of μ in panel (c), the electrons can be scattered smoothly through $\mu = 0$.

Figure 2 displays the spatial distributions of energetic electrons at three energy ranges, 5–10 keV, 20–30 keV, and > 50 keV, in *RunA*. The simulation times in panels (a)-(c) are τ_c = 1.71 s, $2\tau_c$ = 3.42 s, and $3\tau_c$ = 5.13 s, respectively. Due to the trapping effect of magnetic mirror, most electrons are concentrated around the top of the flare loop. Since the initial pitch-angle distribution of injected electrons is isotropic, electrons with larger pitch-angles take much more time as they move from the loop top to lower altitudes. Therefore, we can see multiple streams of electrons bouncing back and forth in the loop, as denoted by the black arrows (only the right side is marked). The number density of streaming electrons is smaller than that trapped at the looptop and the pattern varies with energy. As shown below, the streams of electrons are the reason for the presence of stripes in the

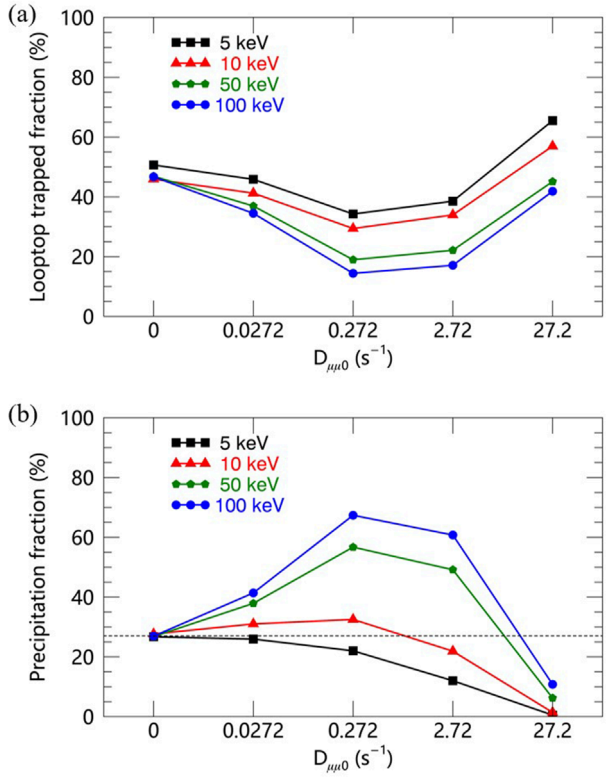

FIGURE 3
Fractions of electrons at various energies trapped in the looptop **(A)** and precipitating to the solar surface **(B)** at time $3\tau_c$ as a function of the scattering rates in the five simulation runs. The dashed line in panel **(B)** denotes the expected precipitating fraction F_{pc} = 27% in the magnetic mirror without turbulent scattering.

velocity space distribution. For different simulation runs, the spatial distribution is generally similar. With increasing scattering rate, the distribution gets smoother and streaming electrons are harder to be distinguished.

We now analyze the effect of different scattering rates (as described by $D_{\mu\mu 0}$) on the trapping and precipitation of electrons in the flare loop. Figure 3 shows the fractions of electrons trapped in the looptop and precipitating to the solar surface for various energies, 5 keV, 10 keV, 50 keV, and 100 keV, respectively, at the end of the simulation ($3\tau_c$). For each energy, the trapped fraction is defined as $F_t = N_{y>45}/N_{inject}$, where $N_{y>45}$ is the number of electrons that remain trapped at $y >$ 45 Mm in the simulation domain and N_{inject} is the injected population. For the precipitating fraction, it is defined as $F_p = 1 - N_{loop}/N_{inject}$, where N_{loop} is the number of electrons that remain bouncing in the loop and have not reached the bottom boundary. As noted above, for the field line where electrons are initially injected, the mirror ratio σ = 5.9 and the critical pitch angle θ_c = 24.3°. The electrons with pitch angle smaller than θ_c fall into the loss cone and can escape. Therefore, it results in an expected precipitating fraction F_{pc} = 27%, as denoted by the dashed line in Figure 3B. In the simulation of *RunA* without turbulent scattering, the precipitating fractions at different energies agree well with the theoretical predication.

As shown in Figure 3, with increasing scattering rate, the variations of both the trapped fraction (F_t) and precipitation fraction (F_p) display a non-monotonic pattern. From the non-scattering case in *RunA* to the weak and intermediate scattering cases in *RunB* and *RunC*, the trapped fraction decreases and the precipitation fraction increases. Due to pitch-angle scattering, more and more electrons with initial pitch-angle larger than the critical value θ_c = 24.3° are scattered into the loss cone and escape. However, for the low-energy electrons of 5 keV, although the trapped fraction decreases, there is no rise in the precipitation fraction. It suggests that the electrons have not reached the solar surface while they have left the looptop, possibly due to their low speed. For the strong scattering cases in *RunD* and *RunE*, the pitch-angle scattering is so frequent that electrons should stay at the looptop for much longer time before moving to lower altitudes. Thus, from moderate to strong scattering, the trapped fraction increases and the precipitation fraction decreases. This indicates that the precipitation fraction is highest in the moderate diffusion regime, therefore, in favor of high intensity of nonthermal HXR and microwave emissions in the footpoints. In contrast, to reproduce a bright nonthermal source in the looptop, either weak or strong scattering is required. We also find that the magnitude of variation is energy dependent. The trapped fraction decreases at higher energies, while the precipitation fraction increases with energy.

Due to the effects of magnetic mirror and turbulent scattering, the particle distribution in the velocity space varies along the flare loop and with time. Figures 4, 5 show the velocity space distribution in the looptop and loopleg regions, respectively. $v_\parallel$ and $v_\perp$ are velocity components in the parallel and perpendicular directions. Here the looptop is integrated over x = [-5, 5] Mm and y = [45, 55] Mm, and the loopleg on the right side is integrated over x = [20, 40] Mm and y = [20, 30] Mm. The simulation results in *RunA*, *RunB*, and *RunE* are displayed in panels (a), (b), and (c), respectively. The left and right columns are at two different times, τ_c and $3\tau_c$, respectively. The dashed line in each panel illustrates the critical pitch angle θ_c = 24.3° for the magnetic field lines where electrons are injected.

In the non-scattering (*RunA*) and weak scattering (*RunB*) runs, the velocity space distributions are obviously anisotropic, but exhibit different patterns in the looptop and loopleg regions. As shown in panel (a) in Figure 4, at the looptop, multiple narrow bands (which resemble branches or fishbone) stretch out from the vertical axis and present positive gradients in the perpendicular direction, i.e., $\partial f/\partial v_\perp > 0$. The number of bands increases with time and they gather towards the origin of the coordinate system. If we continue to run the simulation, the gap between stripes gets smaller and the distribution will evolve into a double-sided loss cone. Those electrons at the looptop are mainly reflected and trapped electrons, therefore most electrons are distributed in the perpendicular direction. Since electrons with larger $v_\parallel$ can leave the looptop faster or be reflected faster, multiple streams of electrons can be observed as shown in Figure 2, leading to fishbone-like multiple bands as time goes on. As shown in panel (b) in Figure 4, with weak scattering, similar stripes can be seen at the early time, which also exhibits $\partial f/\partial v_\perp > 0$. Due to turbulent scattering, the width of stripes increases and some electrons fill in the gaps between stripes. At later time as shown on the right, the gaps between stripes are nearly smoothed out and the distribution resemble a

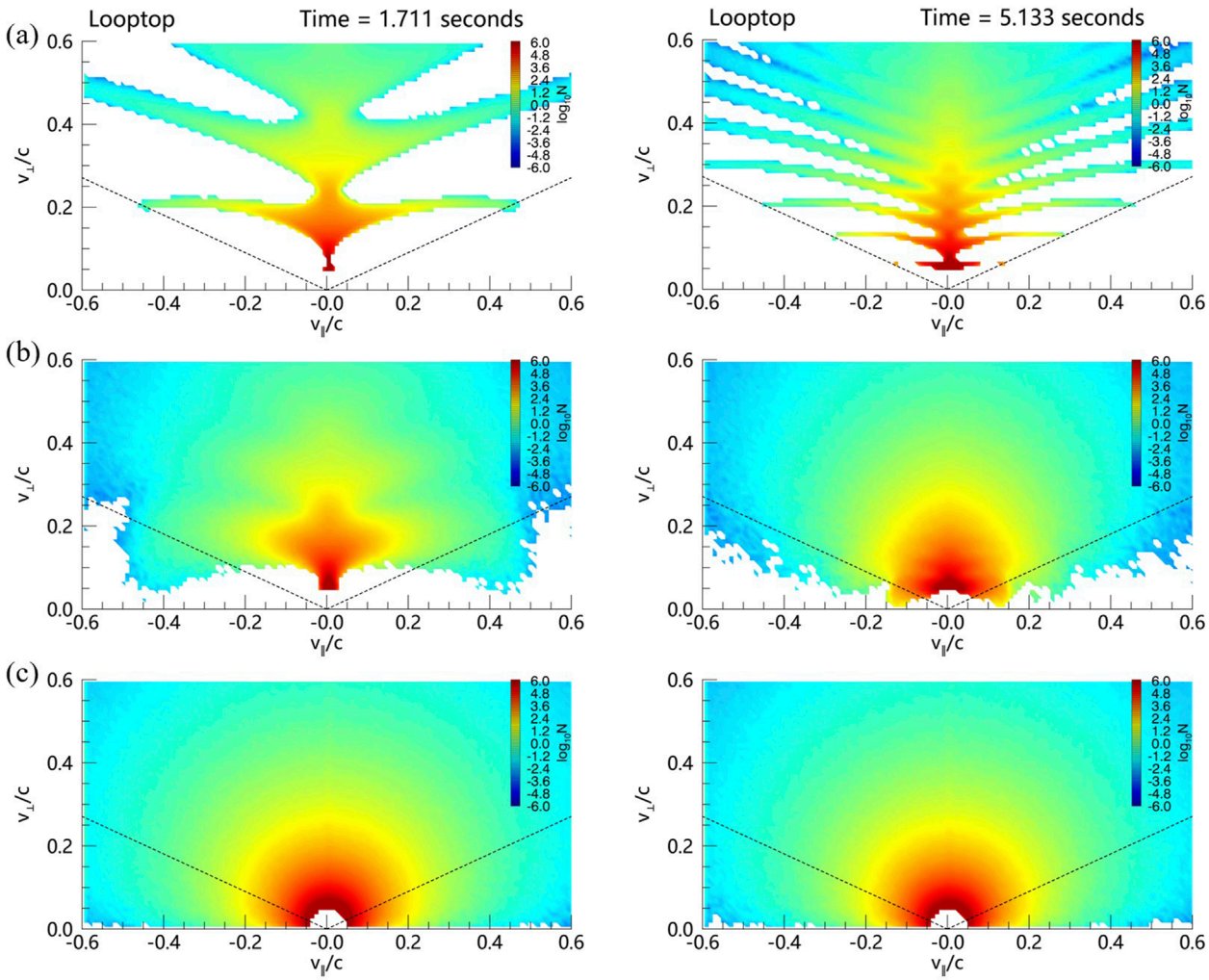

FIGURE 4
Velocity space distribution in the looptop region at two different times, τ_c (left) and $3\tau_c$ (right). Panels **(A–C)** are results in *RunA*, *RunB*, and *RunE*, respectively. The dashed line in each panel denotes the critical pitch angle $\theta_c = 24.3°$.

double-sided loss cone. As shown in panel (c), for the simulation with strong scattering (*RunE*), the distribution has already become nearly isotropic at the early time. In addition, compared with the non-scattering case in panel (a), electrons are scattered into the loss cone (below the dashed line) at various energies as a result of pitch-angle scattering, consistent with the results as shown in Figure 3.

As shown in Figure 5, in the loopleg region, the velocity space distributions look different from that in the looptop. In panels (a) and (b), at the early time, the first stripe (close to the origin of the coordinate system) is circular-shaped and represents the contribution from the beam-like electrons before getting reflected. The distribution resembles the so-called horseshoe distribution as observed in the source of auroral kilometric radiation (AKR, see, e.g., Ergun et al., 2000; Treumann, 2006). It contains positive gradients in both parallel and perpendicular directions, i.e., $\partial f/\partial v_\parallel > 0$ and $\partial f/\partial v_\perp > 0$. At later time, after being reflected, $v_\parallel$ evolves from positive to negative values and the distribution presents mainly $\partial f/\partial v_\perp > 0$. Similarly, the other stripes appearing later also have $\partial f/\partial v_\perp > 0$. As in the looptop, if we continue to run the simulation, the distribution will eventually evolve into a double-sided loss cone, but it is asymmetric. For the case with strong scattering, as shown in panel (c), the distribution is nearly isotropic.

4 Summary and discussion

In this work, we numerically model the transport of energetic electrons in the flare loop after being injected around the top of the loop. We examine the effect of turbulent pitch-angle scattering on the trap/precipitation fraction and velocity space distribution by introducing different levels of scattering rates. We find that both the fractions of electrons trapped in the looptop and precipitating into the solar surface vary in a non-monotonic way with increasing scattering rate. From non-/weak to intermediate scattering, the

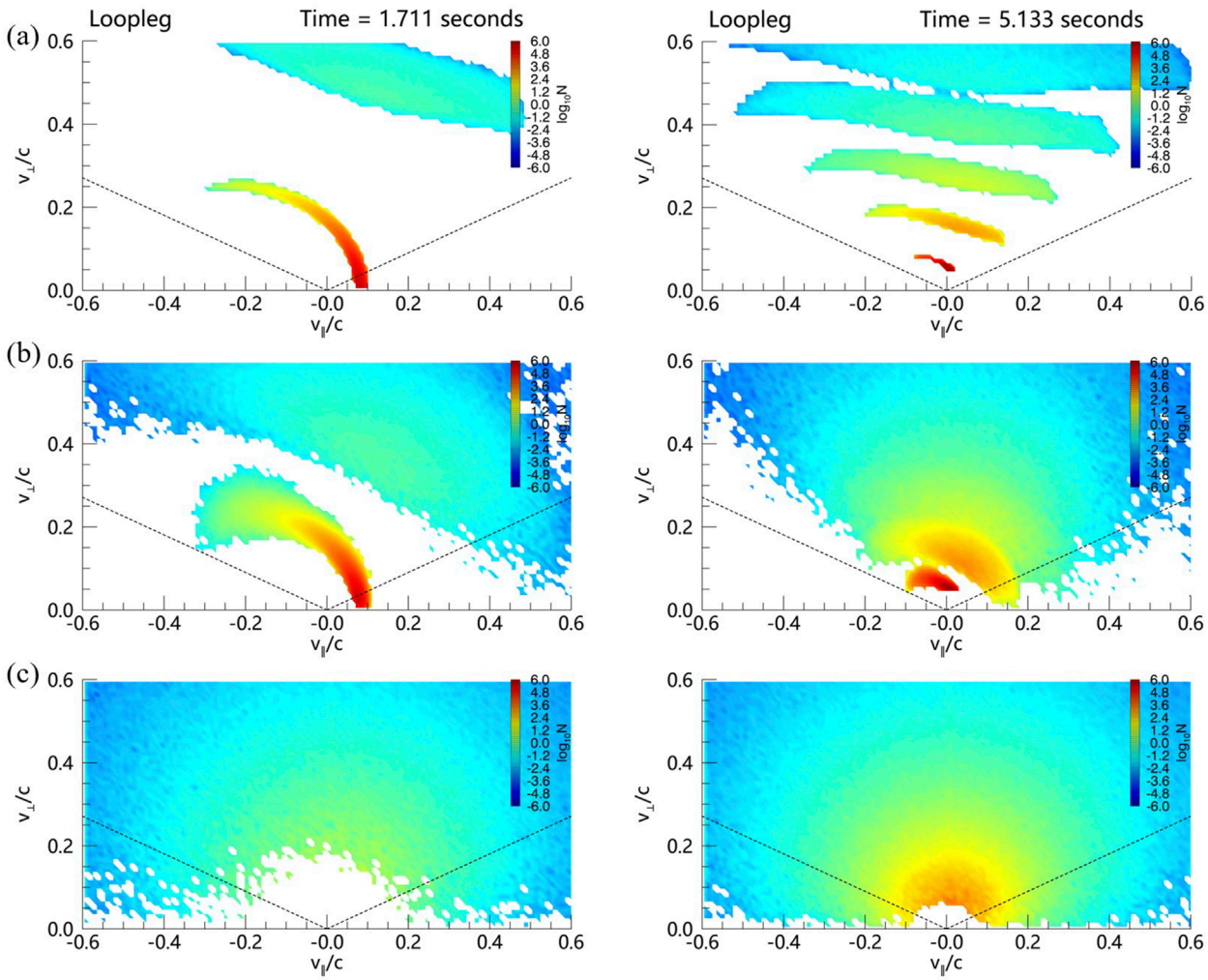

FIGURE 5
Same as plotted in Figure 4, but for velocity space distribution in the right loopleg region.

trapped fraction decreases and the precipitation fraction increases, while from intermediate to strong scattering, the trapped fraction increases and the precipitation fraction decreases. Therefore, in the moderate diffusion regime, we expect intense nonthermal HXR and microwave emissions in the footpoints due to the highest precipitation fraction. In addition, the trap/precipitation fraction apparently shows energy dependence, which will affect the electron energy spectra both in the looptop and loopleg regions. We also find that the velocity space distribution varies both along the loop and with time. With non-/weak turbulent scattering, it presents different patterns of stripes and is highly anisotropic both in the looptop and loopleg, and gradually evolves into a double-sided loss-cone as the simulation continues. In the case of enhanced turbulent scattering, the distribution becomes nearly isotropic because a large number of electrons can be scattered into the loss-cone.

In flare regions with strong magnetic fields, the plasma frequency can be smaller compared to the electron gyro-frequency, i.e., $\omega_{pe}/\Omega_{ce} < 1$. Different anisotropic features in the velocity space can drive the electron cyclotron maser instability (ECMI) in different manners. In the looptop region, most electrons are distributed in the perpendicular direction, with horizontal branches stretching outward. Positive gradients along the loss-cone boundaries mainly generate fundamental X-mode emissions via ECMI, propagating along the parallel and oblique directions (see, e.g., Yoon and Ziebell, 1995; Ning et al., 2021b). According to the plasma kinetic theory, the linear growth rates of ECMI can be approximated with the integral of the velocity distribution function gradient ($\partial f/\partial v_\perp$) along the resonance curve in the phase space (Wu and Lee, 1979; Wu, 1985). We note that in the non-scattering case, the distribution presents branch features with sharp gradients where the resonance curve could pass through. This could drive ECMI with higher efficiency, compared to the classical loss-cone distribution. In the loopleg region, the distribution resembles the horseshoe distribution in the source of planetary AKR. Recently, the horseshoe-driven ECMI has been applied to explain the solar spikes (e.g., Melrose and Wheatland, 2016; Ning et al., 2021a). Multi-stripe distribution has been demonstrated in earlier studies (e.g., White et al., 1983). Yousefzadeh et al. (2021) carried out kinetic simulations and found that such electrons mainly generate second

harmonic X-mode emissions (X2), which could solve the escaping difficulty of fundamental emission in solar corona. For the strong scattering case, the distributions in both regions are nearly isotropic, making it hard to drive the ECMI.

We used the reduced transport equation that includes magnetic mirroring and turbulent pitch-angle scattering in this study. Other effects such as Coulomb collisions and cross-filed diffusion have been neglected and will be discussed in future work. We considered different regimes of turbulent pitch-angle diffusion as defined in Bespalov et al. (1987). However, the level of magnetic turbulence in realistic flares remains unclear. Recently, some studies (Kontar et al., 2017; Stores et al., 2021) investigated the spatial and temporal distributions of turbulence in one solar flare from the observations of nonthermal broadening of spectral lines. They calculated the turbulent kinetic energy density from the nonthermal broadening velocity (v_{nth}), which approximates the energy density associated with the magnetic field fluctuations. Then, one can estimate the level of turbulent magnetic fluctuation $\delta B/B \sim v_{nth}/v_A$ (Kontar et al., 2017). Taking the Alfven speed v_A ~2000 km s^{-1}, while v_{nth} ranging between ~10–100 km s^{-1}, we can get $\delta B/B$ is about 0.05%–5%. This indicates that the turbulence is relatively weak on average, at least for this flare event, and may provide the required condition for anisotropic distribution.

A similar electron trap and precipitation process occurs in the Earth's magnetosphere, where the resonant interaction between energetic electrons and plasma waves such as chorus waves has been applied to explain the characteristics of aurora (e.g., Thorne et al., 2010; Zhang et al., 2022). Our simulation results may provide helpful insights to the dynamics of energetic particles in the radiation belts of magnetosphere.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

XK: Conceptualization, Investigation, Methodology, Project administration, Writing–original draft, Writing–review and editing. HN: Writing–review and editing. YC: Writing–review and editing.

Funding

The author(s) declare that financial support was received for the research, authorship, and/or publication of this article. XK is supported by the National Key R&D Program of China under grant 2022YFF0503002 (2022YFF0503000), the National Natural Science Foundation of China under grants 42074203, and Yunnan Key Laboratory of Solar Physics and Space Science under grant YNSPCC202218. HN is supported by NSFC 12203031 and the China Postdoctoral Science Foundation (2022TQ0189). The work was carried out at National Supercomputer Center in Tianjin (TH-3F) and Guangzhou (TianHe-2).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Allred, J. C., Alaoui, M., Kowalski, A. F., and Kerr, G. S. (2020). Modeling the transport of nonthermal particles in flares using fokker-planck kinetic theory. *Astrophys. J.* 902, 16. doi:10.3847/1538-4357/abb239

Beeck, J., and Wibberenz, G. (1986). Pitch angle distributions of solar energetic particles and the local scattering properties of the interplanetary medium. *Astrophys. J.* 311, 437. doi:10.1086/164784

Benz, A. O. (2017). Flare observations. *Living Rev. Sol. Phys.* 14 (2B), 2. doi:10.1007/s41116-016-0004-3

Bespalov, P. A., Zaitsev, V. V., and Stepanov, A. V. (1987). On the origin of time delays in hard X-ray and gamma-ray emission of solar flares. *Sol. Phys.* 114, 127–140. doi:10.1007/bf00193073

Bian, N. H., Emslie, A. G., and Kontar, E. P. (2017). The role of diffusion in the transport of energetic electrons during solar flares. *Astrophys. J.* 835, 262. doi:10.3847/1538-4357/835/2/262

Charikov, Y. E., Mel'nikov, V. F., and Kudryavtsev, I. V. (2012). Intensity and polarization of the hard X-ray radiation of solar flares at the top and footpoints of a magnetic loop. *Geomagnetism Aeronomy* 52, 1021–1031. doi:10.1134/S0016793212080051

Dröge, W., Kartavykh, Y. Y., Klecker, B., and Kovaltsov, G. A. (2010). Anisotropic three-dimensional focused transport of solar energetic particles in the inner heliosphere. *Astrophys. J.* 709, 912–919. doi:10.1088/0004-637x/709/2/912

Effenberger, F., and Petrosian, V. (2018). The relation between escape and scattering times of energetic particles in a turbulent magnetized plasma: application to solar flares. *Appl. Sol. Flares* 868, L28. doi:10.3847/2041-8213/aaedb3

Ergun, R. E., Carlson, C. W., McFadden, J. P., Delory, G. T., Strangeway, R. J., and Pritchett, P. L. (2000). Electron-cyclotron maser driven by charged-particle acceleration from magnetic field-aligned electric fields. *Astrophys. J.* 538, 456–466. doi:10.1086/309094

Fletcher, L. (1995). On the generation of loop-top impulsive hard X-ray sources. 303, L9–303L.

Fletcher, L., Dennis, B. R., Hudson, H. S., Krucker, S., Phillips, K., Veronig, A., et al. (2011). An observational overview of solar flares 159, 19–106. doi:10.1007/978-1-4614-3073-5_3

Fletcher, L., and Martens, P. C. H. (1998). A model for hard X-ray emission from the top of flaring loops. *Astrophys. J.* 505, 418–431. doi:10.1086/306137

Giacalone, J., and Jokipii, J. R. (1999). The transport of cosmic rays across a turbulent magnetic field. *Astrophys. J.* 520, 204–214. doi:10.1086/307452

Hamilton, R. J., and Petrosian, V. (1990). Effects of Coulomb collisions on cyclotron maser and plasma wave growth in magnetic loops. *Astrophys. J.* 365, 778. doi:10.1086/169531

Hu, J., Li, G., Ao, X., Zank, G. P., and Verkhoglyadova, O. (2017). Modeling particle acceleration and transport at a 2-D CME-driven shock. *J. Geophys. Res. (Space Physics)* 122, 10,938–10,963. doi:10.1002/2017JA024077

Jeffrey, N. L. S., Kontar, E. P., Bian, N. H., and Emslie, A. G. (2014). On the variation of solar flare coronal X-ray source sizes with energy. *Astrophys. J.* 787, 86. doi:10.1088/0004-637x/787/1/86

Jokipii, J. R. (1971). Propagation of cosmic rays in the solar wind. *Rev. Geophys. Space Phys.* 9, 27–87. doi:10.1029/rg009i001p00027

Kong, X., Chen, B., Guo, F., Shen, C., Li, X., Ye, J., et al. (2022). Numerical modeling of energetic electron acceleration, transport, and emission in solar flares: connecting loop-top and footpoint hard X-ray sources. *Astrophys. J. Lett.* 941, L22. doi:10.3847/2041-8213/aca65c

Kong, X., Guo, F., Shen, C., Chen, B., Chen, Y., Musset, S., et al. (2019). The acceleration and confinement of energetic electrons by a termination shock in a magnetic trap: an explanation for nonthermal loop-top sources during solar flares. 887, L37. doi:10.3847/2041-8213/ab5f67887L

Kontar, E. P., Bian, N. H., Emslie, A. G., and Vilmer, N. (2014). Turbulent pitch-angle scattering and diffusive transport of hard X-ray-producing electrons in flaring coronal loops. *Astrophys. J.* 780, 176. doi:10.1088/0004-637x/780/2/176

Kontar, E. P., Hannah, I. G., and Bian, N. H. (2011). Acceleration, magnetic fluctuations, and cross-field transport of energetic electrons in a solar. *Flare Loop* 730, L22. doi:10.1088/2041-8205/730/2/l22

Kontar, E. P., Perez, J. E., Harra, L. K., Kuznetsov, A. A., Emslie, A. G., Jeffrey, N. L. S., et al. (2017). Turbulent kinetic energy in the energy balance of a solar flare. *Phys. Rev. Lett.* 118, 155101. doi:10.1103/PhysRevLett.118.155101

Kuznetsov, A. A., and Fleishman, G. D. (2021). Ultimate fast gyrosynchrotron codes. *Astrophys. J.* 922, 103. doi:10.3847/1538-4357/ac29c0

Li, X., Guo, F., and Liu, Y.-H. (2021). The acceleration of charged particles and formation of power-law energy spectra in nonrelativistic magnetic reconnection. *Phys. Plasmas* 28, 052905. doi:10.1063/5.0047644

Lin, J., Forbes, T. G., Priest, E. R., and Bungey, T. N. (1995). Models for the motions of flare loops and ribbons. *Sol. Phys.* 159, 275–299. doi:10.1007/BF00686534

Melnikov, V. F., Charikov, Y. E., and Kudryavtsev, I. V. (2013). Spatial brightness distribution of hard X-Ray emission along flare loops. *Geomagnetism Aeronomy* 53, 863–866. doi:10.1134/S0016793213070153

Melnikov, V. F., and Filatov, L. V. (2020). Conditions for whistler generation by nonthermal electrons in flare loops. *Geomagnetism Aeronomy* 60, 1126–1131. doi:10.1134/S0016793220080150

Melnikov, V. F., and Filatov, L. V. (2021). Nonthermal electron diffusion modes in whistler turbulence in flare loops. *Geomagnetism Aeronomy* 61, 1189–1196. doi:10.1134/S0016793221080144

Melrose, D. B. (2017). Coherent emission mechanisms in astrophysical plasmas. *Rev. Mod. Plasma Phys.* 1, 5. doi:10.1007/s41614-017-0007-0

Melrose, D. B., and Wheatland, M. S. (2016). Is cyclotron maser emission in solar flares driven by a horseshoe distribution? *Sol. Phys.* 291, 3637–3658. doi:10.1007/s11207-016-1006-y

Miller, J. A., Cargill, P. J., Emslie, A. G., Holman, G. D., Dennis, B. R., LaRosa, T. N., et al. (1997). Critical issues for understanding particle acceleration in impulsive solar flares. *J. Geophys. Res.* 102, 14631–14659. doi:10.1029/97ja00976

Minoshima, T., Masuda, S., and Miyoshi, Y. (2010). Drift-kinetic modeling of particle acceleration and transport in solar flares. *Astrophys. J.* 714, 332–342. doi:10.1088/0004-637x/714/1/332

Minoshima, T., Masuda, S., Miyoshi, Y., and Kusano, K. (2011). Coronal electron distribution in solar flares: drift-kinetic model. *Astrophys. J.* 732, 111. doi:10.1088/0004-637x/732/2/111

Musset, S., Kontar, E. P., and Vilmer, N. (2018). Diffusive transport of energetic electrons in the solar corona: X-ray and radio diagnostics. *diagnostics* 610, A6. doi:10.1051/0004-6361/201731514

Ning, H., Chen, Y., Ni, S., Li, C., Zhang, Z., Kong, X., et al. (2021a). Harmonic electron-cyclotron maser emissions driven by energetic electrons of the horseshoe distribution with application to solar radio spikes. 651, A118. doi:10.1051/0004-6361/202140427

Ning, H., Chen, Y., Ni, S., Li, C., Zhang, Z., Kong, X., et al. (2021b). Harmonic maser emissions from electrons with loss-cone distribution in solar active regions. 920, L40. doi:10.3847/2041-8213/ac2cc6920L

Qin, G., Zhang, M., and Dwyer, J. R. (2006). Effect of adiabatic cooling on the fitted parallel mean free path of solar energetic particles. *J. Geophys. Res. (Space Physics)* 111, A08101. doi:10.1029/2005JA011512

Roelof, E. C. (1969). "Propagation of solar cosmic rays in the interplanetary magnetic field," in *Lectures in high-energy astrophysics*. Editors H. Ögelman, and J. R. Wayland, 111.

Ruan, W., Yan, L., and Keppens, R. (2023). Magnetohydrodynamic turbulence formation in solar flares: 3D simulation and synthetic observations. *Astrophys. J.* 947, 67. doi:10.3847/1538-4357/ac9b4e

Skilling, J. (1971). Cosmic rays in the galaxy: convection or diffusion? *Astrophys. J.* 170, 265. doi:10.1086/151210

Stores, M., Jeffrey, N. L. S., and Kontar, E. P. (2021). The spatial and temporal variations of turbulence in a solar flare. *Astrophys. J.* 923, 40. doi:10.3847/1538-4357/ac2c65

Strauss, R. D. T., and Effenberger, F. (2017). A hitch-hiker's guide to stochastic differential equations. *Solut. Methods Energetic Part. Transp. Space Phys. Astrophysics* 212, 151–192. doi:10.1007/s11214-017-0351-y

Tang, J. F., Wu, D. J., Chen, L., Tan, C. M., and Wang, J. B. (2024). Electron cyclotron maser instability by evolving fast electron beams in the flare loops. *Front. Astronomy Space Sci.* 11, 1404145. doi:10.3389/fspas.2024.1404145

Tang, J. F., Wu, D. J., Chen, L., Xu, L., and Tan, B. L. (2020). Parametric evolution of power-law energy spectra of flare accelerated electrons in the solar atmosphere. *Atmosphere* 904, 1. doi:10.3847/1538-4357/abc2ca

Thorne, R. M., Ni, B., Tao, X., Horne, R. B., and Meredith, N. P. (2010). Scattering by chorus waves as the dominant cause of diffuse auroral precipitation 467, 943–946. doi:10.1038/nature09467

Treumann, R. A. (2006). The electron-cyclotron maser for astrophysical application. *Astron. Astrophys. Rev.* 13, 229–315. doi:10.1007/s00159-006-0001-y

van den Berg, J., Strauss, D. T., and Effenberger, F. (2020). A primer on focused solar energetic particle transport. *Space Sci. Rev.* 216, 146. doi:10.1007/s11214-020-00771-x

Wang, Y., Cheng, X., Ding, M., Liu, Z., Liu, J., and Zhu, X. (2023). Three-dimensional turbulent reconnection within the solar flare current sheet. *Curr. Sheet* 954, L36. doi:10.3847/2041-8213/acf19d

Wang, Y., Qin, G., and Zhang, M. (2012). Effects of perpendicular diffusion on energetic particles accelerated by the interplanetary coronal mass ejection shock. *Astrophys. J.* 752, 37. doi:10.1088/0004-637x/752/1/37

Wei, W., Shen, F., Yang, Z., Zhao, L., Wang, Y., Zuo, P., et al. (2019). Modeling solar energetic particle transport in 3D background solar wind: influences of the compression regions. *J. Atmos. Solar-Terrestrial Phys.* 182, 155–164. doi:10.1016/j.jastp.2018.11.012

White, S. M., Melrose, D. B., and Dulk, G. A. (1983). Electron cyclotron masers during solar flares. *Publ. Astron. Soc. Aust.* 5, 188–191. doi:10.1017/s1323358000016829

Wijsen, N., Aran, A., Pomoell, J., and Poedts, S. (2019). Modelling three-dimensional transport of solar energetic protons in a corotating interaction region generated with EUHFORIA 622, A28. doi:10.1051/0004-6361/201833958

Wu, C. S. (1985). Kinetic cyclotron and synchrotron maser instabilities: radio emission processes by direct amplification of radiation. *Space Sci. Rev.* 41, 215–298. doi:10.1007/bf00190653

Wu, C. S., and Lee, L. C. (1979). A theory of the terrestrial kilometric radiation. *Astrophys. J.* 230, 621–626. doi:10.1086/157120

Wu, H., Dai, Y., and Ding, M. D. (2023). Highly energetic electrons accelerated in strong solar flares as a preferred driver of sunquakes. *Astrophys. J. Lett.* 943, L6. doi:10.3847/2041-8213/acb0d1

Ye, J., Raymond, J. C., Mei, Z., Cai, Q., Chen, Y., Li, Y., et al. (2023). Three-dimensional simulation of thermodynamics on confined turbulence in a large-scale CME-flare current sheet. *Astrophys. J.* 955, 88. doi:10.3847/1538-4357/acf129

Yoon, P. H., and Ziebell, L. F. (1995). Quasilinear evolution of cyclotron maser instability. *Phys. Rev. E* 51, 4908–4916. doi:10.1103/PhysRevE.51.4908

Yousefzadeh, M., Ning, H., and Chen, Y. (2021). Harmonic electron cyclotron maser emission excited by energetic electrons traveling inside a coronal loop. *Astrophys. J.* 909, 3. doi:10.3847/1538-4357/abd8d5

Zhang, M. (1999). A markov stochastic process theory of cosmic-ray modulation. *Astrophys. J.* 513, 409–420. doi:10.1086/306857

Zhang, M., Qin, G., and Rassoul, H. (2009). Propagation of solar energetic particles in three-dimensional interplanetary magnetic fields. *Astrophys. J.* 692, 109–132. doi:10.1088/0004-637x/692/1/109

Zhang, M., and Zhao, L. (2017). Precipitation and release of solar energetic particles from the solar coronal magnetic field. *Astrophys. J.* 846, 107. doi:10.3847/1538-4357/aa86a8

Zhang, X.-J., Artemyev, A., Angelopoulos, V., Tsai, E., Wilkins, C., Kasahara, S., et al. (2022). Superfast precipitation of energetic electrons in the radiation belts of the Earth. *Nat. Commun.* 13, 1611. doi:10.1038/s41467-022-29291-8

Zhao, G. Q., Feng, H. Q., Wu, D. J., Chen, L., Tang, J. F., and Liu, Q. (2016a). Cyclotron maser emission from power-law electrons with strong pitch-angle anisotropy. *Astrophys. J.* 822, 58. doi:10.3847/0004-637x/822/2/58

Zhao, L., Zhang, M., and Rassoul, H. K. (2016b). Double power laws in the event-integrated solar energetic particle spectrum. *Astrophys. J.* 821, 62. doi:10.3847/0004-637x/821/1/62

Zharkova, V. V., Arzner, K., Benz, A. O., Browning, P., Dauphin, C., Emslie, A. G., et al. (2011). Recent advances in understanding particle acceleration processes in solar flares. *Space Sci. Rev.* 159, 357–420. doi:10.1007/s11214-011-9803-y

Zuo, P., Zhang, M., Gamayunov, K., Rassoul, H., and Luo, X. (2011). Energy spectrum of energetic particles accelerated by shock waves: from focused transport to diffusive acceleration. *Astrophys. J.* 738, 168. doi:10.1088/0004-637x/738/2/168

Frontiers in
Astronomy and Space Sciences

Explores planetary science and extragalactic astronomy in all wavelengths

Advances the understanding of our universe - from planetary science to extragalactic astronomy, to high-energy and astroparticle physics.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact