

DATA ASSIMILATION AND CONTROL: THEORY AND APPLICATIONS IN LIFE SCIENCES

EDITED BY: Axel Hutt, Wilhelm Stannat and Roland Potthast

PUBLISHED IN: Frontiers in Applied Mathematics and Statistics,
Frontiers in Psychology and Frontiers in Physiology



frontiers

Frontiers Copyright Statement

© Copyright 2007-2019 Frontiers Media SA. All rights reserved.

All content included on this site, such as text, graphics, logos, button icons, images, video/audio clips, downloads, data compilations and software, is the property of or is licensed to Frontiers Media SA ("Frontiers") or its licensees and/or subcontractors. The copyright in the text of individual articles is the property of their respective authors, subject to a license granted to Frontiers.

The compilation of articles constituting this e-book, wherever published, as well as the compilation of all other content on this site, is the exclusive property of Frontiers. For the conditions for downloading and copying of e-books from Frontiers' website, please see the Terms for Website Use. If purchasing Frontiers e-books from other websites or sources, the conditions of the website concerned apply.

Images and graphics not forming part of user-contributed materials may not be downloaded or copied without permission.

Individual articles may be downloaded and reproduced in accordance with the principles of the CC-BY licence subject to any copyright or other notices. They may not be re-sold as an e-book.

As author or other contributor you grant a CC-BY licence to others to reproduce your articles, including any graphics and third-party materials supplied by you, in accordance with the Conditions for Website Use and subject to any copyright notices which you include in connection with your articles and materials.

All copyright, and all rights therein, are protected by national and international copyright laws.

The above represents a summary only. For the full conditions see the Conditions for Authors and the Conditions for Website Use.

ISSN 1664-8714

ISBN 978-2-88945-985-8

DOI 10.3389/978-2-88945-985-8

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: researchtopics@frontiersin.org

DATA ASSIMILATION AND CONTROL: THEORY AND APPLICATIONS IN LIFE SCIENCES

Topic Editors:

Axel Hutt, Deutscher Wetterdienst, Germany

Wilhelm Stannat, TU Berlin, Germany

Roland Potthast, Deutscher Wetterdienst, Germany

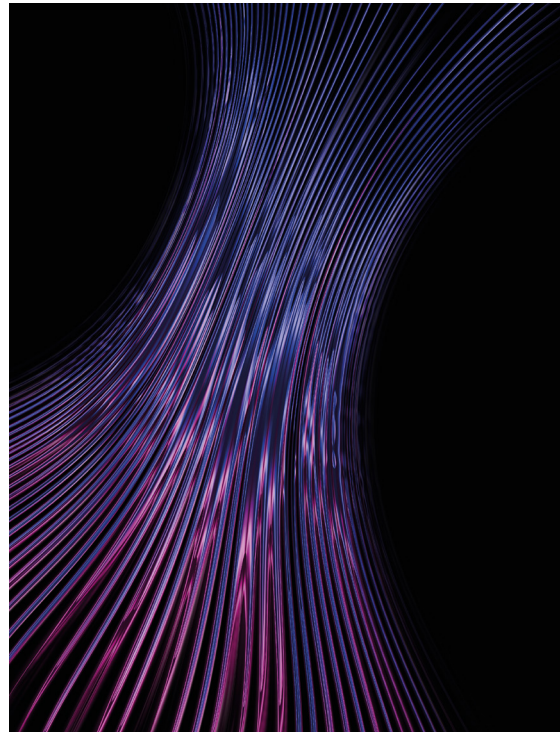


Image: SiljeAO/Shutterstock.com

The understanding of complex systems is a key element to predict and control the system's dynamics. To gain deeper insights into the underlying actions of complex systems today, more and more data of diverse types are analyzed that mirror the systems dynamics, whereas system models are still hard to derive. Data assimilation merges both data and model to an optimal description of complex systems' dynamics. The present eBook brings together both recent theoretical work in data assimilation and control and demonstrates applications in diverse research fields.

Citation: Hutt, A., Stannat, W., Potthast, R., eds. (2019). Data Assimilation and Control: Theory and Applications in Life Sciences. Lausanne: Frontiers Media. doi: 10.3389/978-2-88945-985-8

Table of Contents

- 04 Editorial: Data Assimilation and Control: Theory and Applications in Life Sciences**
Axel Hutt, Wilhelm Stannat and Roland Potthast
- 06 Double Stimulation in a Spiking Neural Network Model of the Midbrain Superior Colliculus**
Bahadır Kasap and A. John van Opstal
- 22 ExGUtils: A Python Package for Statistical Analysis With the ex-Gaussian Probability Density**
Carmen Moret-Tatay, Daniel Gamermann, Esperanza Navarro-Pardo and Pedro Fernández de Córdoba Castellá
- 33 Corrigendum: ExGUtils: A Python Package for Statistical Analysis With the ex-Gaussian Probability Density**
Carmen Moret-Tatay, Daniel Gamermann, Esperanza Navarro-Pardo and Pedro Fernández de Córdoba Castellá
- 34 Transcriptome and Proteome Alternation With Resistance to *Bacillus thuringiensis* Cry1Ah Toxin in *Ostrinia furnacalis***
Muhammad Zeeshan Shabbir, Tiantao Zhang, Zhenying Wang and Kanglai He
- 47 Extracranial Estimation of Neural Mass Model Parameters Using the Unscented Kalman Filter**
Lara Escuain-Poole, Jordi Garcia-Ojalvo and Antonio J. Pons
- 62 Statistical Data Assimilation: Formulation and Examples From Neurobiology**
Anna Miller, Dawei Li, Jason Platt, Arij Daou, Daniel Margoliash and Henry D. I. Abarbanel
- 75 Data-Driven Modeling and Prediction of Complex Spatio-Temporal Dynamics in Excitable Media**
Sebastian Herzog, Florentin Wörgötter and Ulrich Parlitz
- 85 Ultra Rapid Data Assimilation Based on Ensemble Filters**
Roland Potthast and Christian A. Welzbacher
- 97 Forecast of Spectral Features by Ensemble Data Assimilation**
Axel Hutt and Roland Potthast



Editorial: Data Assimilation and Control: Theory and Applications in Life Sciences

Axel Hutt^{1,2*}, Wilhelm Stannat³ and Roland Potthast^{1,2}

¹ Deutscher Wetterdienst, Department for Data Assimilation, Offenbach, Germany, ² Department for Mathematics and Statistics, University of Reading, Reading, United Kingdom, ³ Institute of Mathematics, TU Berlin, Berlin, Germany

Keywords: parameter estimation, superior colliculus, electroencephalography, *Ostrinia furnacalis*, excitable media, avian song system, meteorology

Editorial on the Research Topic

Data Assimilation and Control: Theory and Applications in Life Sciences

The understanding of complex systems, such as insecticides or the mammalian heart, is a key element to predict and control the system's dynamics. To gain deeper insights into the underlying actions of complex systems, today, more and more data of diverse types are analyzed that mirror the systems dynamics, whereas system models are still hard to derive. Consequently, developing techniques that permit the construction of models which are well-adapted to observed data is one of the great challenges. To match system models with diverse experimental data, data assimilation and control theory provide important techniques. They use a combination of observations and models to achieve optimal fitting of model parameters, providing optimal forecast estimations or control of the system's dynamics to make the system perform a specific task. The present Research Topic (and the corresponding e-book) brings together both recent theoretical work and applications in life sciences.

Typical research in the life science aims to understand the complex system under study involving diverse system models and observations. If a model of the system dynamics exists, it is insightful to validate the model by comparing the model's dynamical solutions with observations, either quantitatively or qualitatively. For instance, one may consider the experimental setup of a control experiment in a real-world system and simulate the experimental setup in the model framework by computing the model system's response to an equivalent external stimulation. Kasap and van Opstal have chosen this approach and simulated the control of eye saccades by electric stimulation. Their study shows good qualitative and quantitative agreement between the model dynamics and observations, validating their model. Since their effective model describes well major observation features, the successful model features can be interpreted as the major features in the brain structure.

Another approach may aim to improve or extract a model by observations. For instance, in psychology, the statistical ex-Gaussian distribution describes well the subjects' reaction times. To construct a statistical model of cognitive processes, it is important to estimate parameters of the ex-Gaussian distribution in an efficient way. Moret-Tatay et al. have developed a software library to efficiently estimate the coefficients of the ex-Gaussian distribution. Similarly, Shabbir et al. fit statistical distributions to experimental gene data to understand better why the Asian corn borer can develop resistance to genetically modified maize that is supposed to be toxic to the insect. Both latter studies aim to understand complex behavior by identifying statistical models.

OPEN ACCESS

Edited by:

Ulrich Parlitz,
Max Planck Society (MPG), Germany

Reviewed by:

Isao T. Tokuda,
Ritsumeikan University, Japan
Philip Bittihn,
Max Planck Society (MPG), Germany

*Correspondence:

Axel Hutt
digitalesbad@gmail.com

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 29 March 2019

Accepted: 29 April 2019

Published: 21 May 2019

Citation:

Hutt A, Stannat W and Potthast R
(2019) Editorial: Data Assimilation and
Control: Theory and Applications in
Life Sciences.
Front. Appl. Math. Stat. 5:25.
doi: 10.3389/fams.2019.00025

Dynamical neural models that describe mathematically the temporal evolution of neural populations play an important role in neuroscience. Escuain-Poole et al. consider a dynamical model of neural populations in the brain, that allows to explain the electroencephalogram (EEG) measured on the scalp, i.e., outside of the brain. The work shows in several theoretical studies how to estimate brain model parameters from synthetic EEG-data that are observed on the head surface. This estimation is done by the well-known Unscented Kalman filter. A similar analysis approach is *statistical data assimilation* that allows to estimate model parameters and system forecasts. Typically, statistical data assimilation provides efficient tools to estimate the posterior probability density function of model parameters. In the article of Miller et al. the authors successfully performed parameter estimations of an avian song model by statistical data assimilation and predicted the evolution of optimal model solutions.

Typical dynamical models are differential equation systems whose parameters are estimated. In the last decades, more and more of such differential equation models have been extended or even replaced by methods borrowed from artificial intelligence, such as artificial neural networks. Herzog et al. show how to estimate an underlying chaotic dynamical model by a combination of a convolutional neural network and a conditional random network. The neural network is fit to synthetic data generated from a heart tissue model. The authors show in detail that the neural network allows to faithfully reproduce the dynamics of single elements of the underlying model.

Parameter estimation is an important application of data assimilation, as demonstrated in the contributions described above. Beyond this, data assimilation techniques also provide improved forecasts. For instance, in meteorology, the solution of an atmospheric physical model represents a short-time forecast, e.g., a spatial distribution of atmospheric state variables after 1 h. A subsequent data assimilation step transforms this spatial distribution to a new spatial distribution (called analysis) that is closer to observations. One of the major aims in atmospheric data assimilation is to obtain free forecasts, i.e., long-time model solutions with the analysis as initial condition, that accurately predict the weather. Hence, in this context, data

assimilation provides optimal initial conditions for forecasts. One of the limits of standard data assimilation techniques is the condition that observations must be sufficiently dense in reasonably long fixed intervals. Potthast and Welzbacher have studied in detail a rapid data assimilation technique based on an ensemble Kalman filter that considers observations in very short time intervals. The authors show that the ultra-rapid update of the analysis significantly improves forecasts. Possible applications of the new technique range from meteorology to neuroscience.

More generally, the prediction of neural activity has attracted increasing attention over the last decade. Hutt and Potthast have proposed to forecast the spectral power of forecast time series in certain frequency bands, since it is well-known that the brain encodes and decodes information by oscillations in certain frequency ranges. To this end, the authors have applied a data assimilation cycle utilizing an ensemble Kalman filter and have computed ensemble forecasts and their time-frequency power spectral distributions. It is shown by statistical ensemble verification that these time-frequency distributions of forecasts better explain underlying oscillatory content than forecast time series.

Future research in the field may involve data assimilation of non-local observations from a theoretical perspective and more applications in biology and neuroscience.

AUTHOR CONTRIBUTIONS

AH wrote the Editorial and all authors re-read and edited the manuscript.

Conflict of Interest Statement: 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.

Copyright © 2019 Hutt, Stannat and Potthast. 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.



Double Stimulation in a Spiking Neural Network Model of the Midbrain Superior Colliculus

Bahadır Kasap and A. John van Opstal*

Department of Biophysics, Donders Institute for Brain, Cognition and Behavior, Radboud University, Nijmegen, Netherlands

OPEN ACCESS

Edited by:

Axel Hutt,
German Meteorological Service,
Germany

Reviewed by:

Meysam Hashemi,
INSERM U1106 Institut de
Neurosciences des Systèmes, France
Jorge F. Mejias,
University of Amsterdam, Netherlands

*Correspondence:

A. John van Opstal
j.vanopstal@donders.ru.nl

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 15 July 2018

Accepted: 18 September 2018

Published: 09 October 2018

Citation:

Kasap B and van Opstal AJ (2018)
Double Stimulation in a Spiking Neural
Network Model of the Midbrain
Superior Colliculus.
Front. Appl. Math. Stat. 4:47.
doi: 10.3389/fams.2018.00047

The midbrain superior colliculus (SC) is a crucial sensorimotor interface in the generation of rapid saccadic gaze shifts. For every saccade it recruits a large population of cells in its vectorial motor map. Supra-threshold electrical microstimulation in the SC reveals that the stimulated site produces the saccade vector specified by the motor map. Electrically evoked saccades (E-saccades) have kinematic properties that strongly resemble natural, visual-evoked saccades (V-saccades), with little influence of the stimulation parameters. Moreover, synchronous stimulation at two sites yields eye movements that resemble a weighted vector average of the individual stimulation effects. Single-unit recordings have indicated that the SC population acts as a vectorial pulse generator by specifying the instantaneous gaze-kinematics through dynamic summation of the movement effects of all SC spike trains. But how to reconcile the a-specific stimulation pulses with these intricate saccade properties? We recently developed a spiking neural network model of the SC, in which microstimulation initially activates a relatively small set of (~50) neurons around the electrode tip, which subsequently sets up a large population response (~5,000 neurons) through lateral synaptic interactions. Single-site microstimulation in this network thus produces the saccade properties and firing rate profiles as seen in single-unit recording experiments. We here show that this mechanism also accounts for many results of simultaneous double stimulation at different SC sites. The resulting E-saccade trajectories resemble a weighted average of the single-site effects, in which stimulus current strength of the electrode pulses serve as weighting factors. We discuss under which conditions the network produces effects that deviate from experimental results.

Keywords: saccades, motor map, spatial-temporal transformation, electrical stimulation, population coding, vector averaging

INTRODUCTION

Superior Colliculus

Because high spatial resolution is limited to the central fovea, the primate visual system needs to explore the environment through rapid and precise saccadic eye movements. Normal (human and monkey) saccades display stereotyped “main sequence” characteristics, described by linear amplitude-duration and nonlinear, saturating, amplitude-peak eye velocity relationships [1]. In addition, the horizontal and vertical velocity profiles of oblique saccades are tightly coupled, such that they are scaled versions of each other throughout the saccade, and saccade trajectories are approximately straight in all directions [2]. These properties imply that the saccadic system contains a nonlinear control stage [2–4].

Previously, these main-sequence properties had been assumed to arise at the brainstem level, possibly because of saturation of the brainstem saccadic burst neurons [3].

Recent hypotheses have suggested, however, that the saccade nonlinearity reflects a speed-accuracy trade-off, which optimally deals with spatial uncertainty in the retinal periphery and internal noise in sensorimotor pathways [5–8]. We have hypothesized that the midbrain superior colliculus (SC) would be in an excellent position to implement such a strategy [8].

The neural circuitry underlying saccade planning, selection, and execution extends from the cerebral cortex to the cerebellum, and the pons in the brainstem. The midbrain SC is the final common terminal for all cortical and subcortical outputs, and it is known to specify the vectorial eye-displacement command for the brainstem oculomotor circuitry [9–11]. The SC contains an eye-centered topographic map of visuomotor space, in which the saccade amplitude is mapped logarithmically along the rostral-caudal axis (u , in mm) and saccade direction roughly linearly along the medial-lateral direction (v , in mm; [9]). The afferent mapping (Equation 1a) and its efferent inverse (Equation 1b) are well described by Ottes et al. [12]:

$$\left. \begin{aligned} u &= B_u \ln \left(\frac{\sqrt{(x+A)^2 + y^2}}{A} \right) \\ v &= B_v \operatorname{atan} \left(\frac{y}{x+A} \right) \end{aligned} \right\} \quad (1a)$$

$$\Leftrightarrow \begin{cases} x = A \cdot \left(\exp \left(\frac{u}{B_u} \right) \cos \left(\frac{v}{B_v} \right) - 1 \right) \\ y = A \cdot \exp \left(\frac{u}{B_u} \right) \sin \left(\frac{v}{B_v} \right) \end{cases} \quad (1b)$$

with typical parameter values for the monkey SC given as $B_u \approx 1.4$ mm, $B_v \approx 1.8$ mm/rad, and $A \approx 3$ deg; see **Figure 1**. Each saccade is associated with a translation-invariant Gaussian-shaped population within this map, the center of which corresponds (through Equation 1a) to the saccade vector, (x_0, y_0) , and a width $\sigma_{\text{pop}} \approx 0.5$ mm [12, 14, 15]. Thus, the activity of neuron n in the motor map is described by:

$$F_n(u_n, v_n) = F_{\text{max}} \cdot e^{-\frac{1}{2} \cdot \left(\frac{(u_0 - u_n)^2 + (v_0 - v_n)^2}{\sigma_{\text{pop}}^2} \right)} \quad (2)$$

with F_{max} the peak activity of the population, quantified by the number of spikes in the saccade-related burst (e.g., **Figures 1, 3A**). It is generally assumed that each recruited neuron, n , in the population encodes a vectorial movement contribution to the saccade vector, which is determined by both its anatomical location within the motor map, (u_n, v_n) , and its activity, F_n [2, 11–13, 16–18].

However, the precise mechanism by which the cells contribute to the saccade is still elusive. A major hypothesis in the literature holds that the output of the population is determined by a nonlinear center-of-gravity computation [17–21]. According to this idea, the activity in the SC motor map only specifies the saccade metrics (amplitude and direction of the saccade vector) and is unrelated to the saccade kinematics. Yet, our single-unit recordings demonstrated a strong (presumably causal) relationship between the instantaneous firing patterns in the SC and associated saccade trajectories [8, 13].

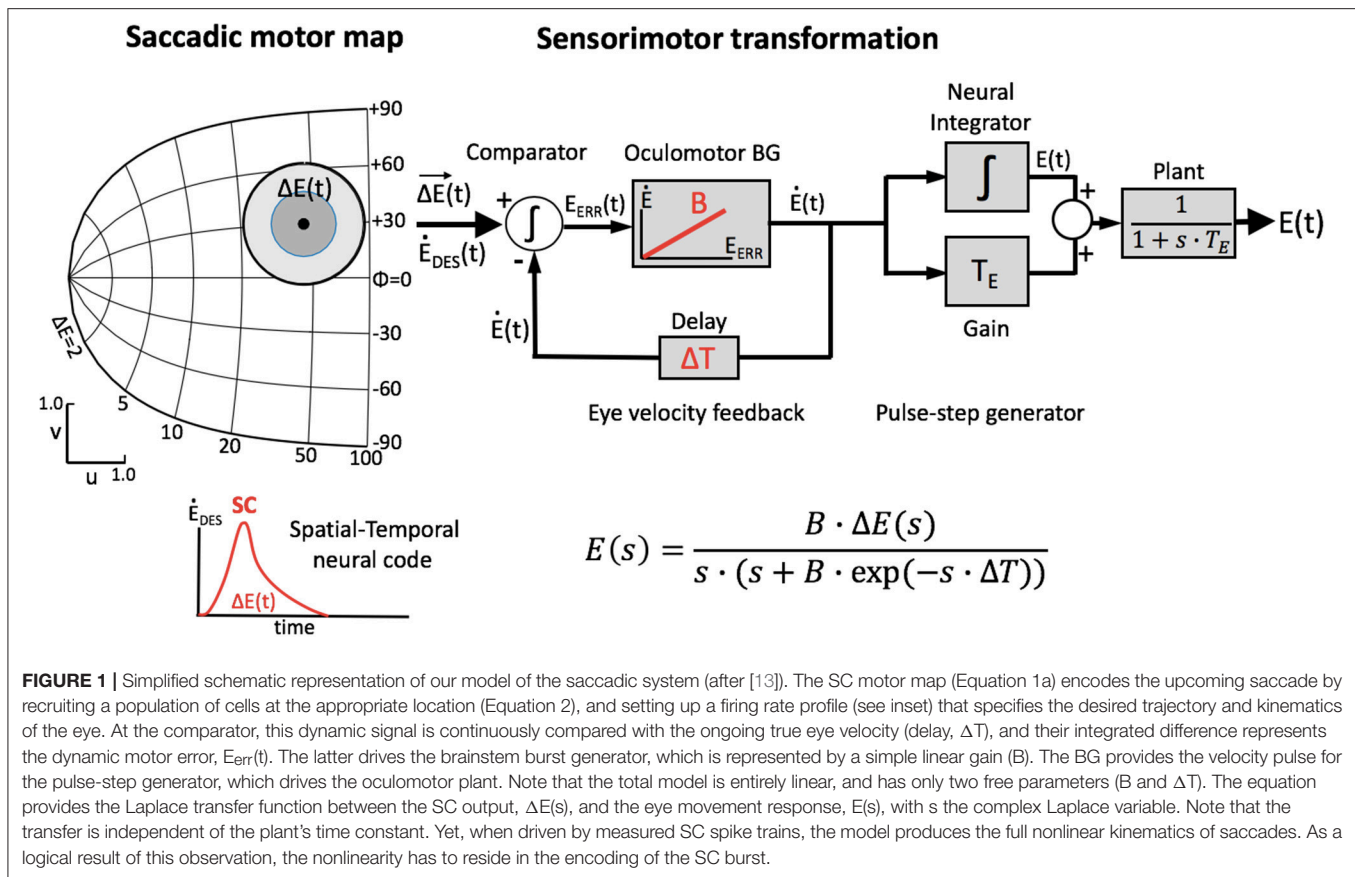
We therefore proposed and tested an extremely simple linear summation model for the recruited population that explains the encoding of spatial-temporal properties of saccade trajectories through the firing properties of SC burst cells ([8, 13]; **Figure 1**). According to this model, the saccade, $S(t)$, is generated in the following way:

$$S(t) = \sum_{n=1}^N \sum_{k=1}^{K_n < t} \delta(t - \tau_{n,k}) \cdot \mathbf{m}_n \quad (3)$$

with N the number of active cells in the population, $K_n < t$ the number of spikes in the burst of neuron n up to time t , and $\mathbf{m}_n = \zeta \cdot (x_n, y_n)$ the tiny site-specific spike vector emanating from the motor map for each spike from each cell. This spike vector is solely determined by the efferent mapping of SC site (u_n, v_n) (Equation 1b), where ζ is a fixed, small scaling constant determined by the cell density in the map and the population size, and $\delta(t - \tau_{n,k})$ is the k 'th spike fired by neuron n at time $\tau_{n,k}$.

Our linear dynamic ensemble-coding model is illustrated in **Figure 1**. The SC provides a feedforward motor command by the temporal integration of all spike trains of the total population. The integrated signal represents the cumulative desired displacement of the eye, whereas the population firing rate represents the desired eye velocity (inset). The SC output thus represents both a spatial (by the location of the population) and a temporal (the instantaneous firing rates) neural code of the eye movement. The SC signal is continuously compared with an efference copy of the true eye velocity (with delay, ΔT), which is generated by the brainstem saccadic burst generator (BG). Note that in our model the BG is taken as a simple no-memory linear system (gain, B). The BG output is subsequently fed through a parallel circuit, consisting of the eye-position integrator and a static gain (T_E). These signals combine at the oculomotor neurons to produce the pulse-step innervation for the oculomotor plant. The latter is usually modeled by a simple first-order low-pass filter with time constant T_E . We showed that this entirely linear model resulted to account for the full nonlinear kinematics of saccades. We therefore proposed that the main-sequence properties should originate at the level of the SC motor map [8, 13]. The neural mechanism underlying this property was identified as a precise tuning of the peak firing rates and burst durations in the SC as a function of their location in the map, while keeping the number of spikes in the population fixed. As a result, the instantaneous firing rates of the neurons together encode all measured properties of saccadic velocity profiles [22].

Recently, we implemented a simple spiking neural network model for the SC that can generate realistic saccades to visual targets [23]. This minimalistic (one-dimensional) model with lateral excitatory-inhibitory interactions among the SC cells accounts for most of the experimentally observed firing properties of saccade-related neurons in the motor map [8, 13], and yields saccades with normal main-sequence properties. The model takes a fixed Gaussian input from upstream sources (e.g., the cortical frontal eye fields, or FEF), and assumes precisely-tuned biophysical properties of the SC network neurons, and their interconnections.



Microstimulation

Electrical stimulation at a particular site in the motor map produces a saccadic gaze shift with metrics that correspond well to the efferent mapping function (Equation 1b), and with normal main-sequence kinematics [9, 15, 24, 25]. These studies have also shown that the properties of electrically evoked (E-)saccades are largely invariant to a wide range of stimulation parameters, which might appear problematic for the linear ensemble-coding model of Equation 3.

Note that two factors contribute to the neural responses to electrical microstimulation: (1) direct (feedforward) current activation of cell bodies and axons by the electric field of the electrode, and (2) synaptic activation through lateral (feedback) interactions among the neurons in the motor map [26].

We recently argued that as current strength falls off rapidly with distance from the electrode tip, only a small number of SC neurons will be directly stimulated by the electrode's electric field (e.g., [27]). Thus, the major factor determining the microstimulation effects would be synaptic transmission. Indeed, several studies have suggested the existence of a functional organization of lateral excitatory-inhibitory interactions within the SC (anatomy: [28, 29]; electrophysiology: [30–32], and pharmacology: [33]).

We thus extended our spiking model to account for single-site microstimulation results over a wide range of stimulation

parameters [26]. The network was tuned such that, above a threshold, the E-saccades were insensitive to changes in the stimulation parameters. This result supports the idea that the excitatory-inhibitory interactions effectively normalize the total SC output. Under microstimulation, the network thus creates a population that is virtually identical to the one elicited by a visual stimulus. It may be expected that such intrinsic normalization could ensure a behavior that resembles (nonlinear) weighted-averaging without the need for a nonlinear, activation-dependent weighting scheme that is implemented downstream from the motor map.

Double Stimulation

In this paper, we further explored the predictions of our model for synchronous and asynchronous electrical stimulation at two different sites. Robinson [9] and Nota and Gnadt [34] demonstrated that double stimulation in the SC produced eye movements that resemble the weighted average of the individual stimulation effects, with the stimulation current strengths and relative timings acting as weighting factors. Similar weighting effects occur when an electrical stimulus is combined with a behaviorally relevant visual stimulus [35]. Results such as these have prompted computational modelers to propose a downstream vector-averaging mechanism that acts on the SC output by explicitly calculating the center of gravity of the population (see above; [17–21]; review in [36]). The neural

mechanism that would implement such a neural computation, however, remains unspecified.

Figure 2 illustrates two extreme outcomes for mechanisms that would both calculate the center of gravity (CoG) of the effects of the total activity: averaging at the level of the motor map (Equation 4a), vs. averaging at the level of the brainstem (Equation 4b), i.e.,:

$$\vec{S}_{CoG}^{SC} = \frac{\sum_{n=1}^{N_{POP}} F_n \cdot \vec{w}_n}{\sum_{n=1}^{N_{POP}} F_n} \quad \text{with } \vec{w}_n = (u_n, v_n) \quad (4a)$$

$$\text{vs. } \vec{S}_{CoG}^{DOWN} = \frac{\sum_{n=1}^{N_{POP}} F_n \cdot \vec{m}_n}{\sum_{n=1}^{N_{POP}} F_n} \quad \text{with } \vec{m}_n = (x_n, y_n) \quad (4b)$$

Note that in the former case (**Figure 2A**), the resulting saccade is horizontal with a constant amplitude of 20 deg, regardless the direction of the single-site responses. In the case of Equation (4b), however, response amplitude varies with the angle, Φ , of the single-site stimulation response as $R_{CoG} = R_{SITE} \cdot \cos \Phi_{SITE}$ (**Figure 2B**).

In an earlier modeling study we had shown that lateral excitatory/inhibitory synaptic interactions within the SC motor map, in combination with the linear ensemble-coding scheme of Van Gisbergen et al. [14], could account for saccade-averaging effects to (synchronous) double stimulation [37, 38]. However, the model's output of that study only focused on the saccade-vector endpoints, as it was not equipped to generate saccade trajectories and their kinematics.

Here we employ the dynamic ensemble-coding scheme of Equation (3) to our spiking collicular network to simulate two-dimensional saccade trajectories under a variety of electrical double-stimulation conditions. We show that linear dynamic ensemble-coding with lateral excitatory-inhibitory interactions in the motor map can account for most of the experimental vector-averaging results to double stimulation [9, 20, 35], without the need for additional computational nonlinearities, such as a downstream population center-of-gravity computation [20, 21, 34], or a spike-counting cut-off threshold [13, 39, 40]. The results of our model simulations suggest several interesting limiting cases to the averaging behavior, which, to our knowledge, have so far not been investigated in experimental studies. We also discuss to what extent the model's responses deviate from experimental findings, and suggest some further refinements to the model.

METHODS

The Log-Polar Mapping

Without loss of generality, we simplified the afferent motor map of Equation (1a) to the isotropic complex logarithmic function, by setting $B_u = B_v = 1$, and $A = 0$:

$$u(R) = \ln(R) \quad \text{and} \quad v(\phi) = \phi, \quad \text{with } R = \sqrt{x^2 + y^2} \quad \text{and} \quad \phi = \text{atan}\left(\frac{y}{x}\right) \quad (5a)$$

Thus, a single spike's movement contribution to the saccade from a cell at site (u, v) is determined by the simplified efferent mapping relations:

$$m_x(u, v) = \zeta \cdot \exp(u) \cdot \cos(v) \quad \text{and} \quad m_y(u, v) = \zeta \cdot \exp(u) \cdot \sin(v) \quad (5b)$$

We modeled the spiking neural network by a rectangular grid of 201×201 neurons, representing the gaze motor-map of the right hemifield with $0 < u < 5$ mm (i.e., up to $R = 148$ deg), and $-\frac{\pi}{2} < v < \frac{\pi}{2}$ mm. Under single-site stimulation, the center location of the recruited population determines the direction and amplitude of the saccade, whereas the temporal activity profile encodes the eye-movement kinematics through Equation (3). As described in our previous studies [23, 26], and briefly summarized below (Equations 13 and 14), the eye-movement main-sequence kinematics result from the location-dependent biophysical properties of the neurons, and their lateral excitatory-inhibitory connectivity profiles.

The Adex Neuron Model

We studied the dynamics of the network through simulations developed in C++/CUDA [41], by custom code that implemented dynamic parallelism on a GPU [42], developed and tested on a Tesla K40 with CUDA Toolkit 7.0, Linux Ubuntu 16.04 LTS. Simulations ran with a time resolution of 0.01 ms. Brute-force search and genetic algorithms were used for parameter identification and network tuning since there exists no analytical solutions for the system [23, 26]. Sample simulation and analysis code can be found under https://bitbucket.org/bkasap/sc_doublestimulation/.

Neurons were described by the adaptive exponential integrate-and-fire (AdEx) model [43, 44], which is a conductance-based model with an exponential membrane potential dependence. The nonlinear temporal dynamics of neuron n are described by two coupled differential equations that determine the two state variables: the cell's membrane potential, V , and the adaptation current, q :

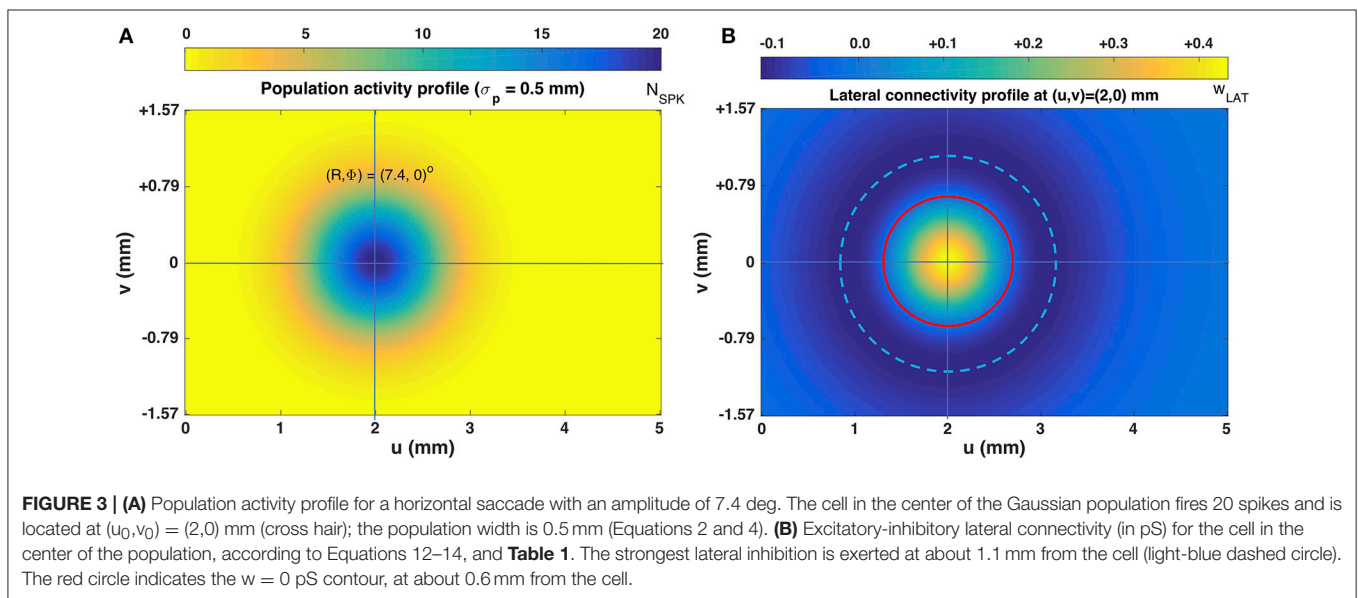
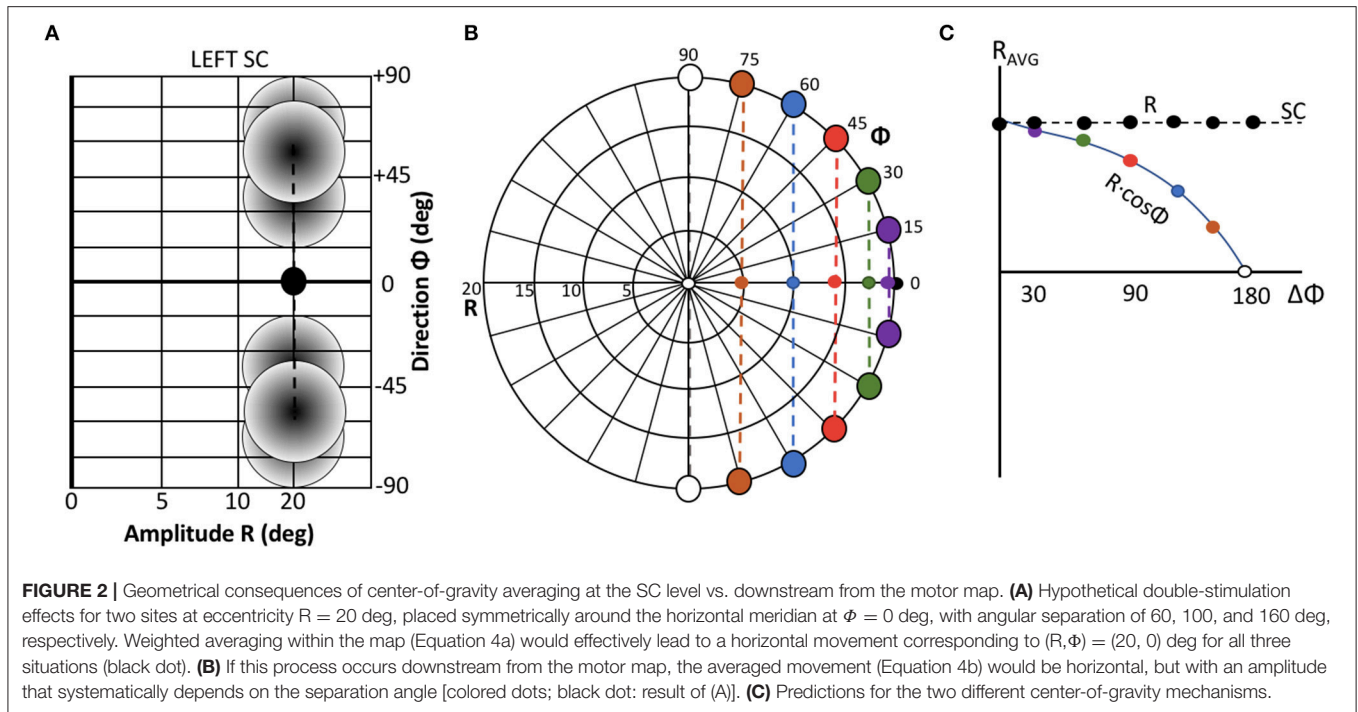
$$C \frac{dV_n}{dt} = -g_L(V_n - E_L) + g_L \eta \exp\left(\frac{V_n - V_T}{\eta}\right) - q_n + I_{inp,n}(t) \quad (6a)$$

$$\tau_{q,n} \frac{dq_n}{dt} = a(V_n - E_L) - q_n \quad (6b)$$

C is the membrane capacitance, g_L is the leak conductance, E_L is the leak reversal potential, η is a slope factor, V_T determines the neural spiking threshold, $\tau_{q,n}$ is the adaptation time constant, a is the sub-threshold adaptation constant, and $I_{inp,n}$ is the cell's total synaptic input current.

Once the membrane potential crosses V_T , the exponential term in Equation (6a) starts to dominate. To limit the membrane potential, we incorporated a ceiling threshold at $V_{peak} = -30$ mV for spike generation. For each spiking event at time τ , the membrane potential is reset to its resting potential, V_{rst} , and the adaptation current, q_n , is increased by b to implement the spike-triggered neural adaptation:

$$V_n(\tau) \rightarrow V_{rst} \quad \text{and} \quad q_n(\tau) \rightarrow q_n(\tau) + b \quad (7)$$



In our model, two biophysical parameters specify the firing properties of the SC neurons: the adaptation time constant, $\tau_{q,n}$ (taken to be location dependent; [23]), and the synaptic input current, $I_{inp,n}$, which is partly determined by the intra-collicular connections (see below). In our model, both depend systematically on the rostral-causal location (u) of the cells within the network. The remaining parameters, C , g_L , E_L , η , V_T , and a , were fixed and tuned such that the cells showed neural bursting behavior (see **Table 1** for the list and values of all parameters used in the simulations, and [26], for example responses and phase plots).

Current Spread

We applied electrical stimulation by the input current, centered around site $[u_E, v_E]$. We assumed an exponential spatial decay of the electric field from the tip of each stimulation electrode. For stimulation at a single site at time t_1 :

$$I_E(u, v, t) = I_0 \cdot \exp\left(-\lambda \cdot \sqrt{(u - u_E)^2 + (v - v_E)^2}\right) \cdot P(t - t_1) \quad (8)$$

with λ (mm^{-1}) a spatial decay constant, I_0 the current intensity at site (u_E, v_E) (in pA), and a rectangular stimulation pulse given

TABLE 1 | List of all parameters used in the simulations.

MICROSTIMULATION PARAMETERS		
λ	10 mm ⁻¹	Spatial decay constant
I_0	150 (40–280) pA	Intracellular current intensity
$P(t)$	I_0 (for $0 < t < D_S$)	Rectangular stimulus pulse
D_S	100 (25–250) ms	Stimulation duration
NEURAL PARAMETERS		
C	600 pF	Membrane capacitance
g_L	20 nS	Leak conductance
E_L	–53 mV	Leak reversal potential
η	2 mV	Spike slope factor
V_T	–50 mV	Exponential threshold
V_{peak}	–30 mV	Spiking threshold
V_{rst}	–45 mV	Reset potential
a	0 nS	Sub-threshold adaptation
b	120 pA	Spike-triggered adaptation
τ_q	100–30 ms	Location-dependent adaptation time constant; varies with (u_n) (Equation 13)
ζ	$5.087 \cdot 10^{-5}$	Spike-vector scaling
SYNAPTIC PARAMETERS		
E_{exc}	0 mV	Excitatory reversal potential
E_{inh}	–80 mV	Inhibitory reversal potential
τ_{exc}	5 ms	Excitatory conductance decay
τ_{inh}	10 ms	Inhibitory conductance decay
LATERAL CONNECTIVITY PARAMETERS		
w_{exc}	45 pS	Excitatory scaling factor
σ_{exc}	0.4 mm	Range of excitatory synapses
w_{inh}	14 pS	Inhibitory scaling factor
σ_{inh}	1.2 mm	Range of inhibitory synapses
s_n	0.0113–0.0148	Location-dependent synaptic scaling parameter; varies with (u_n) , Equation 14).

by $P(t) = 1$ for $0 < t - t_1 < D_S$, and 0 elsewhere. Thus, only a small set of neurons around the stimulation site will be directly activated with this input current (see [26]). In double-stimulation trials, two stimuli were applied at different sites. The total current is then given by:

$$I_E(u, v, t) = \sum_{n=1}^2 I_{0,n} \cdot \exp \left(-\lambda \cdot \sqrt{(u - u_{E,n})^2 + (v - v_{E,n})^2} \right) \cdot P_n(t - t_n) \quad (9)$$

In these simulations, stimulus amplitudes, sites, durations, and their relative timings were systematically varied.

Synapse Dynamics and Lateral Connections

The total input current for neuron n depends on the spiking activity of its surrounding neurons through conductance-based synaptic transmission, and external electric current inputs (Equations 8 or 9):

$$I_{inp, n}(t) = g_n^{exc}(t)(E_e - V_n(t)) + g_n^{inh}(t)(E_i - V_n(t)) + I_E(u_n, v_n, t) \quad (10)$$

where g_n^{exc} and g_n^{inh} are excitatory and inhibitory synaptic conductances acting upon neuron n , E_e , and E_i are excitatory and inhibitory reversal potentials, respectively. These conductances increase instantaneously for each presynaptic spike by a factor that is determined by the synaptic connection strength between neurons, and they subsequently decay over time in an exponential way:

$$\tau_{exc} \frac{dg_n^{exc}}{dt} = -g_n^{exc} + \tau_{exc} \sum_i^{N_{pop}} w_{i,n}^{exc} \sum_s^{N_{spks}^i} \delta(t - \tau_{i,s}) \quad (11a)$$

$$\tau_{inh} \frac{dg_n^{inh}}{dt} = -g_n^{inh} + \tau_{inh} \sum_i^{N_{pop}} w_{i,n}^{inh} \sum_s^{N_{spks}^i} \delta(t - \tau_{i,s}) \quad (11b)$$

with τ_{exc} and τ_{inh} , the excitatory and inhibitory time constants; $w_{i,n}^{exc}$ and $w_{i,n}^{inh}$ are the intracollicular excitatory and inhibitory connection strengths between neurons i and n , respectively (Equations 12a,b) and $\tau_{i,s}$ are the spike timings of all presynaptic SC neurons projecting to neuron n .

We incorporated a Mexican hat-type lateral connection scheme [45]:

$$w_{i,n} = s_n \cdot (w_{i,n}^{exc} - w_{i,n}^{inh}), \quad \text{with} \quad (12)$$

$$w_{i,n}^{exc} = \bar{w}_{exc} \exp \left(-\frac{\|u_i - u_n\|^2}{2\sigma_{exc}^2} \right) \quad (12a)$$

$$w_{i,n}^{inh} = \bar{w}_{inh} \exp \left(-\frac{\|u_i - u_n\|^2}{2\sigma_{inh}^2} \right) \quad (12b)$$

where $\bar{w}_{exc} > \bar{w}_{inh}$ and $\sigma_{inh} > \sigma_{exc}$, and s_n is a location-dependent synaptic scaling parameter, which accounts for the location-dependent change in neuronal sensitivity that is related to the variation in their adaptation time constants. Note, that in our model each SC neuron exerts both excitatory and inhibitory effects on the other neurons in the map, depending on inter-neuron distance. Thus, for simplicity, the inhibitory connections were not mediated by a separate class of inhibitory interneurons.

Figure 1B exemplifies the connectivity profile for a single site.

The strong short-range excitatory and weak long-range inhibitory synapses act as a dynamic soft winner-take-all (WTA) mechanism: not just one neuron remains active, but the “winner” affects the temporal activity patterns of the other active neurons too. The central neuron thus governs the population activity, since it usually is the most active one (but note that under double-stimulation conditions this may change; see section Results). As a result, all recruited neurons exhibit similarly-shaped bursting profiles as the most active neuron, leading to spike-train synchronization within the population [8, 23, 26].

Network Tuning

The intrinsic biophysical properties of the neurons were enforced by systematically varying the adaptation time constant, $\tau_{q,n}$, and the synaptic weight-scaling parameter, s_n . Changes in the

adaptive properties result in a varying susceptibility to synaptic input, while the synaptic scaling corrects for the total input activity. Following the brute-force genetic algorithm from our recent paper [23, 26], the optimal location-dependent $[\tau_{q,n}, s_n]$ value pairs for the neurons were fitted to ensure a systematic negative rostral-caudal gradient of the peak firing rates ($f_{peak} \propto \frac{1}{\sqrt{R}}$) and a fixed number of spikes per neuron for its preferred saccade ($N_{SPK} = 20$) under a single-site microstimulation condition with $I_0 = 150$ pA and $D_S = 100$ ms.

In short, the algorithm optimized the network “fitness,” by incorporating the scaled contributions of the cells’ peak firing rates, their total spike counts, and an inter-cellular synchronization index within the recruited population. As a result, the adaptive time constant, $\tau_{q,n}$, decreased linearly from 100 to 30 ms with the anatomical rostral-caudal location of the neuron, u_n , according to:

$$\tau_{q,n} = 100 - 14 \cdot u_n \text{ ms, with } u_n \in [0, 5] \text{ mm} \quad (13)$$

The optimal synaptic scaling factor for the lateral excitatory/inhibitory connections (Equation 12) could be fitted by a monotonically decreasing 5th-order polynomial in u_n (sin mm; [26]):

$$s(u_n) = 0.0148 + (-2.52 \cdot u_n + 1.6856 \cdot u_n^2 - 1.49 \cdot u_n^3 + 0.4318 \cdot u_n^4 - 0.04737 \cdot u_n^5) \cdot 10^{-4} \quad (14)$$

Table 1 provides the model’s full parameter list.

Figure 3B illustrates the lateral connectivity profile for one of the cells [at $(u, v) = (2.0, 0.0)$ mm] in the motor map, together with the Gaussian population activity around that cell, associated with a small horizontal V-saccade of $[R, \Phi] = [7.4, 0]$ deg (**Figure 2A**). Note that the lateral interaction profiles are similar in shape and extent across all cells in the motor map, but the absolute values of the excitatory peak and inhibitory trough decrease in a systematic way with the rostral-caudal coordinate, u , as $s_{(0)} = 0.0148$ and $s_{(5)} = 0.0113$, from Equation (14).

RESULTS

Single-Site Stimulation

Figures 4A–C shows the recruited neural population at a rostral stimulation site ($R = 2$ deg, $\Phi = 0$ deg) for stimulation with an amplitude of $I_0 = 150$ pA and duration $D_S = 100$ ms. The diameter of the circular population extends to about 1 mm in the motor map, with the cumulative spike count of the central cells reaching ~ 20 spikes. **Figure 4B** provides the neuronal bursts (top spike patterns) from 12 selected cells, together with their calculated spike-density functions. The peak firing rate of the central cells was close to 700 spikes/s and dropped in a regular fashion with distance from the population center. Note also that the cells near the edge of the population were recruited slightly later than the central cells, but that their peak firing rates were reached nearly simultaneously. Moreover, the bursts all appeared to have the same shape. **Figure 4C** presents the saccade of 2 deg

(top: as function of time; bottom: as a spatial trajectory) encoded by this population through Equation (3).

Figures 4D–F shows the results for stimulation at a more caudal location in the motor map, yielding an oblique saccade with $R = 21$ deg, $\Phi = 30$ deg. The size of the evoked population activity is very similar to that of the rostral population, and also the number of spikes elicited by the cells is the same. The peak firing rates of the neurons, however, were markedly lower at the caudal site, reaching a maximum of about 450 spikes/s. As a result, the burst durations increased accordingly, from about 35 ms at the rostral site, to more than 70 ms at the caudal site. Note also that the horizontal and vertical position and velocity temporal profiles are scaled versions of each other, leading to a straight oblique saccade trajectory (**Figure 4F**, lower panel).

Synchronous Stimulation at Nearby Rostral-Caudal Sites

Figure 5 shows the network response to synchronous double stimulation for two nearby sites, at $R = 10$ and $R = 20$ deg (i.e., $u = 2.3$ and 3.0 mm; Equation 5a) on the horizontal meridian [i.e., $\Phi = 0$ ($v = 0$ mm), for both sites]. The microstimulation parameters were taken the same at both locations ($I_0 = 150$ pA for $D_S = 100$ ms). After about 30 ms following population activity onset, the highest merged population activity is observed, in which the most active neurons are found between the two stimulation sites (**Figures 5A,B**). The firing rates of the two neurons closest to the stimulation electrodes are highlighted in **Figure 5B**. Note that the resulting firing rates at these stimulation sites are markedly lower than at the center of the total population. Note also that these firing rates are highly similar. For single-site stimulation, these firing rates would have been different, due to the tuning properties of the neurons within the motor map (Equation 13). These interesting equilibrating population dynamics result from the mutual excitatory/inhibitory interactions among the neurons, as given by Equations (12, 14) (cf. with **Figure 3B**).

Synchronous Stimulation at Widely Separated Rostral-Caudal Sites

Figure 6 illustrates the network response to synchronous double stimulation with the same intensity and duration as in **Figure 5**, at two sites on the horizontal meridian that are separated by nearly 3 mm: $R = 2$ deg and $R = 35$ deg, respectively (at $u = 0.7$ and 3.6 mm). About 30 ms after activity onset, two separated populations can be observed, in which the most active neurons now coincide with the two stimulation sites (**Figure 6A**). The firing rates of the two neurons closest to the stimulation electrodes are again highlighted in **Figure 6B**. Note that the peak firing rate at the small-amplitude stimulation site (green line) is markedly lower (by almost 50%) and has a much longer duration than for the single-site stimulation result (cf. **Figure 4B**). Both populations appear to result in comparable firing dynamics, which again is due to the mutual interactions among the neurons across the motor map (cf. with **Figure 3B**). However, because

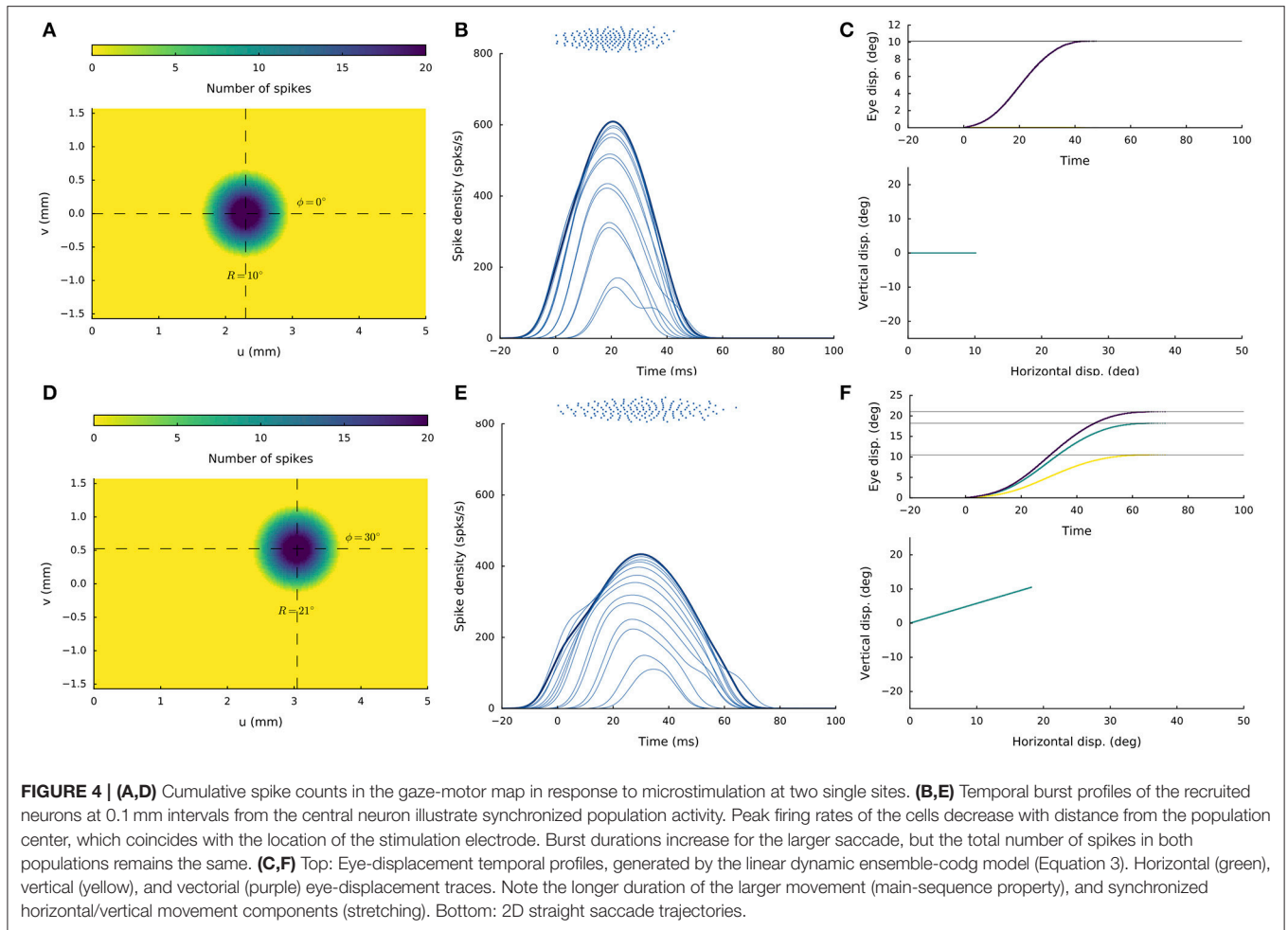


FIGURE 4 | (A,D) Cumulative spike counts in the gaze-motor map in response to microstimulation at two single sites. **(B,E)** Temporal burst profiles of the recruited neurons at 0.1 mm intervals from the central neuron illustrate synchronized population activity. Peak firing rates of the cells decrease with distance from the population center, which coincides with the location of the stimulation electrode. Burst durations increase for the larger saccade, but the total number of spikes in both populations remains the same. **(C,F)** Top: Eye-displacement temporal profiles, generated by the linear dynamic ensemble-codg model (Equation 3). Horizontal (green), vertical (yellow), and vectorial (purple) eye-displacement traces. Note the longer duration of the larger movement (main-sequence property), and synchronized horizontal/vertical movement components (stretching). Bottom: 2D straight saccade trajectories.

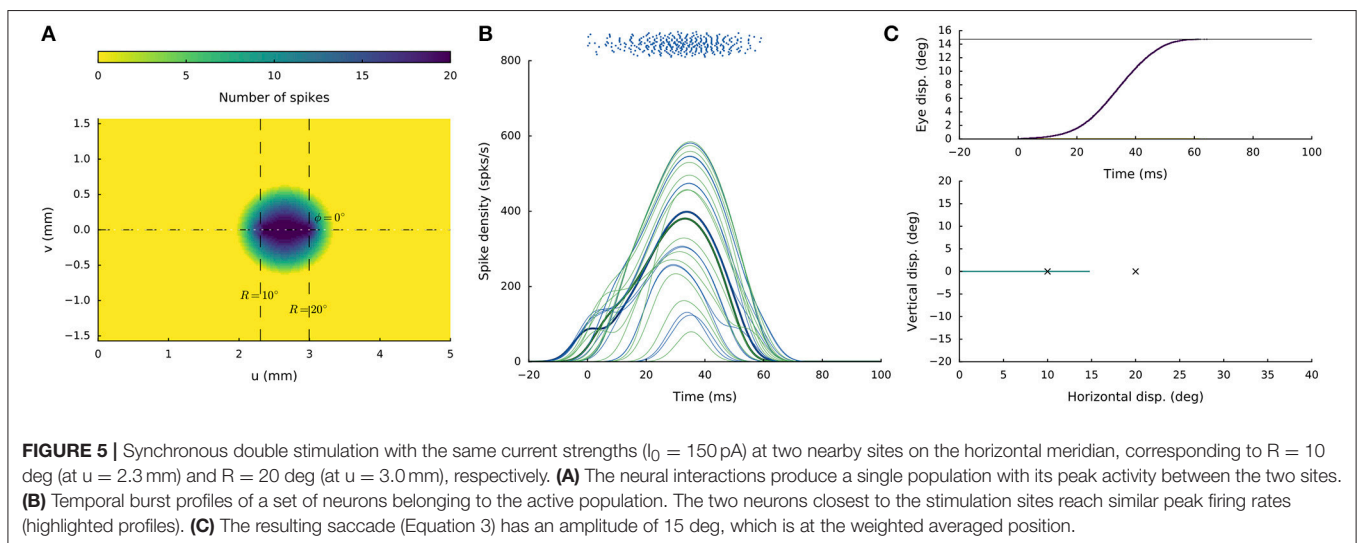
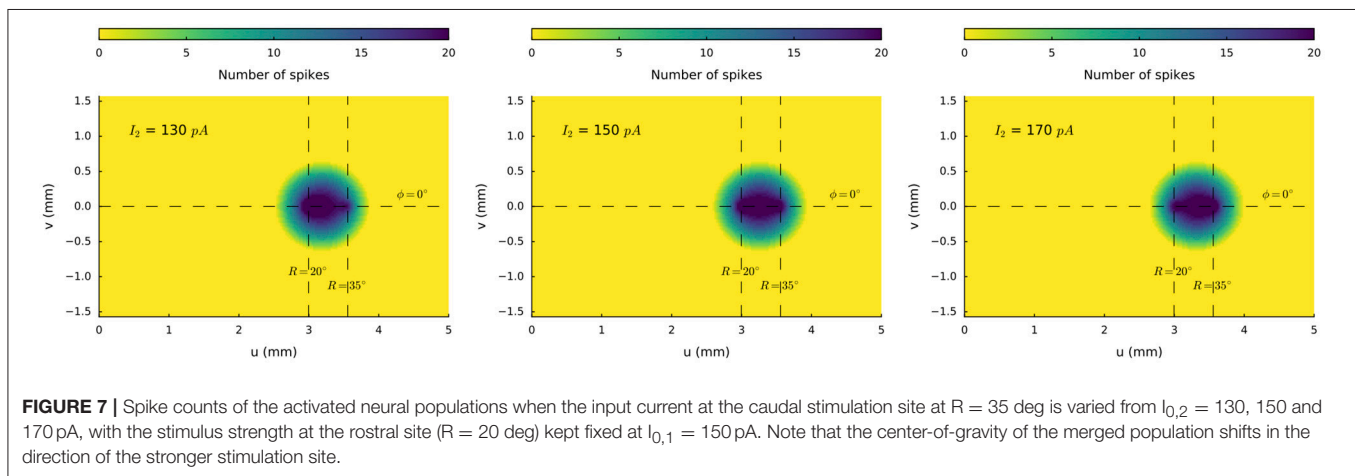
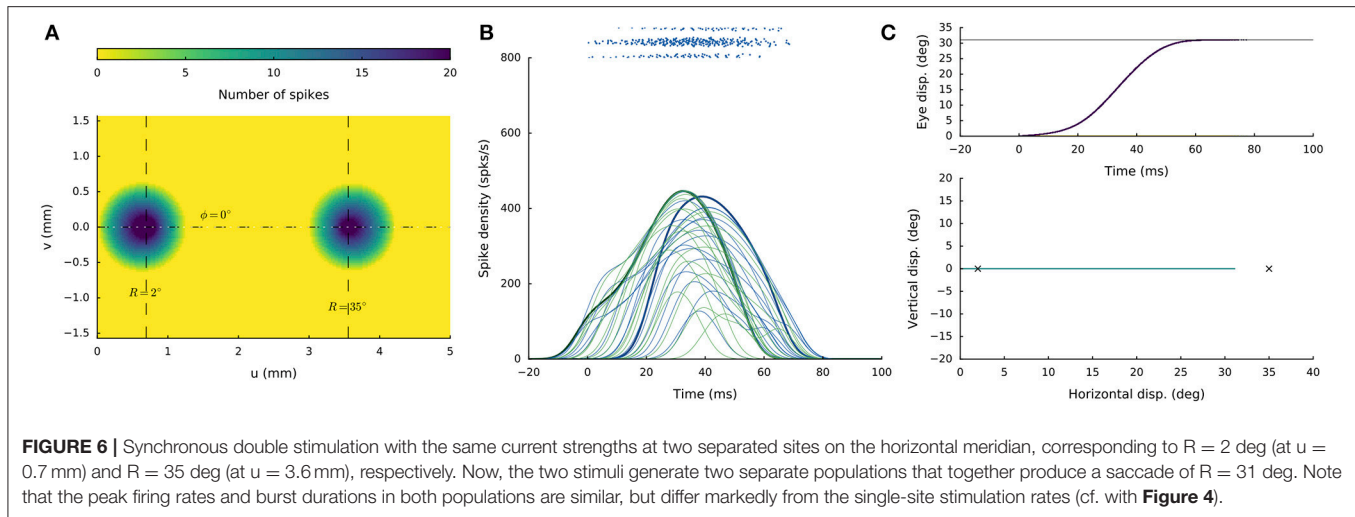


FIGURE 5 | Synchronous double stimulation with the same current strengths ($I_0 = 150$ pA) at two nearby sites on the horizontal meridian, corresponding to $R = 10$ deg (at $u = 2.3$ mm) and $R = 20$ deg (at $u = 3.0$ mm), respectively. **(A)** The neural interactions produce a single population with its peak activity between the two sites. **(B)** Temporal burst profiles of a set of neurons belonging to the active population. The two neurons closest to the stimulation sites reach similar peak firing rates (highlighted profiles). **(C)** The resulting saccade (Equation 3) has an amplitude of 15 deg, which is at the weighted averaged position.

the strength of the interaction profiles is site-specific (Equations 12-14), the populations show different onset dynamics, with the caudal site starting later than the rostral site.

The resulting horizontal saccade has an amplitude of 31 deg, which differs from the linear summation of the two stimulation effects ($R_{\text{SUM}} = 37$ deg).



Weighted Averaging for Rostral-Caudal Sites

We next illustrate the effect of varying the relative current strengths at two stimulation sites on the horizontal meridian (at $R = 20^\circ$ and $R = 35^\circ$, respectively) for synchronous double stimulation. The stimulation amplitude at the rostral electrode was kept constant at $I_{0,1} = 150$ pA, whereas the stimulus intensity at the caudal site was varied systematically between $I_{0,2} = 100$ and 200 pA in 10 pA steps. **Figure 7** illustrates three stimulus situations: $I_{0,2} = 130$ pA, $I_{0,2} = 150$ pA, and $I_{0,2} = 170$ pA. In all three cases a merged population is seen, in which the center-of-gravity of the activity gradually shifts from the rostral to the more caudal site.

Figure 8 shows the result of systematically varying the relative stimulus intensities on the evoked saccade amplitudes (all saccades were horizontal, like in **Figures 4, 5**). The individual stimulation sites produced saccades of $R = 20$ and $R = 35^\circ$, respectively (red symbols). Synchronous stimulation at the two sites, with $I_{1,0} = 150$ pA (fixed), resulted in eye-movements with amplitudes that systematically varied as a function of $I_{2,0}$ between 22.4 and 30° .

Double Stimulation at Medial-Lateral Sites

We next illustrate the effects of synchronous stimulation at two sites that encode the same saccade amplitude ($u = \text{constant}$), but different saccade directions (different v coordinates). In **Figure 9** the two stimulation electrodes were placed at $R = 20^\circ$ and were separated by $\Delta\Phi = 60^\circ$ around the horizontal meridian (cf. **Figure 2A**). The resulting activity shows a merged population with its most intensely firing cells located on the horizontal meridian at $R = 20^\circ$ ($u = 3$ mm). In **Figure 9B** we show the SC bursts for a group of selected cells, with the two sites corresponding to the up and down electrode highlighted by the bold green and blue lines, respectively. Note that the stimulation sites are markedly less active than the cells near the horizontal meridian, and also that their firing rates are much reduced (by more than 40%) with respect to the single-site stimulation effect (cf. **Figure 4D**). The sites near the horizontal meridian, on the other hand, display firing rates (> 500 spikes/s) that significantly exceed the peak firing rate (~ 450 spikes/s) of the single-site stimulation effect at the coordinate for a comparable saccade amplitude.

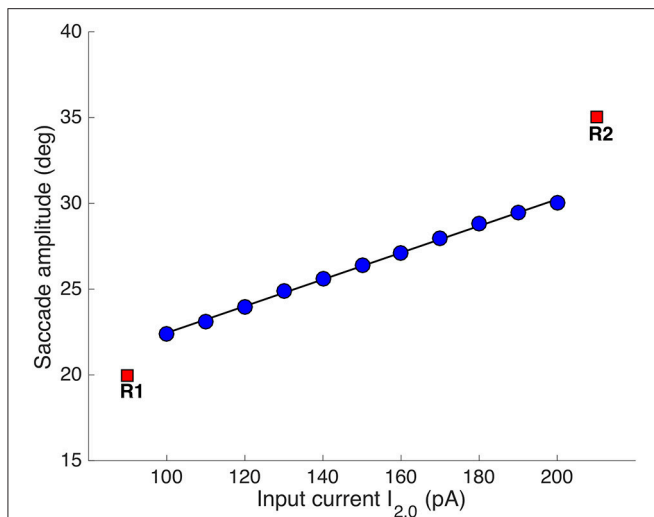


FIGURE 8 | Synchronous double stimulation with varying current strengths at the caudal stimulation site. The input current at $[R, \Phi] = [35, 0]$ deg varied between 100–200 pA, while it was fixed to 150 pA at $[R, \Phi] = [20, 0]$ deg (same stimulus durations of 100 ms). Varying the stimulation strengths shifted the merged population's center-of-gravity as in **Figure 6**. The resulting eye-displacement vectors varied from 22.4 to 30 deg (slope of the linear regression line: 7.8 deg/100 pA).

The resulting saccade is horizontal and has an amplitude of $R = 13$ deg. In other words, the amplitude is much smaller than the saccade corresponding to the site of maximal activity, which would be $R = 20$ deg. It is also somewhat smaller than the projection of the saccade vectors onto the horizontal meridian, which would correspond to an amplitude of $R_{CoG} = 20 \cdot \cos(30) = 17.3$ deg (cf. **Figure 2C**).

Double Stimulation: Evoked Saccade Amplitude Depends on Medial-Lateral Separation

To appreciate the complex interactions between the neural populations along the medial-lateral (ν) axis in the motor map, **Figure 10** shows the results for the evoked saccade amplitude (blue symbols) as function of the medial-lateral separation, $\Delta\nu$, or, equivalently, as function of the angular separation between the two single-site movements. The figure also indicates the simple predictions from the pure center-of-gravity calculations that would result from the motor map ($R = 20$ deg for all sites), and from downstream averaging (the red line). It is clear that the evoked saccades follow neither prediction. Although the averaging effects are clearly due to the neural interactions with the SC motor map (as we have not incorporated a downstream center-of-gravity mechanism in our model, see Equation 3), they clearly differ from the simple scheme of center-of-gravity computation. Instead, the results reflect the intricate neural dynamics as well as the influence of the lateral excitatory-inhibitory interactions (see **Figure 3B**).

For example, for small spatial separations (up to about 0.7 mm), the two populations strongly overlap (as in **Figure 9**).

As a result, they are partly dominated by the mutual excitatory interactions, leading to a slight increase in the saccade amplitude by about one deg. When the sites are separated by about 1 mm, both populations undergo mostly inhibitory influences, leading to a reduced saccade amplitude. This effect increases up to about $\Delta\nu = 1.4$ mm, where the evoked saccade (at these current levels) reaches a minimum of 7.0 deg. In this region the inhibitory interactions are the strongest (see **Figure 3B**). As the electrodes are positioned further apart, the saccade amplitude is still small, but slightly increases up to about 9 deg, because of the slightly lower strength of the lateral inhibition.

Lateral-Medial Double Stimulation at Different Current Strengths

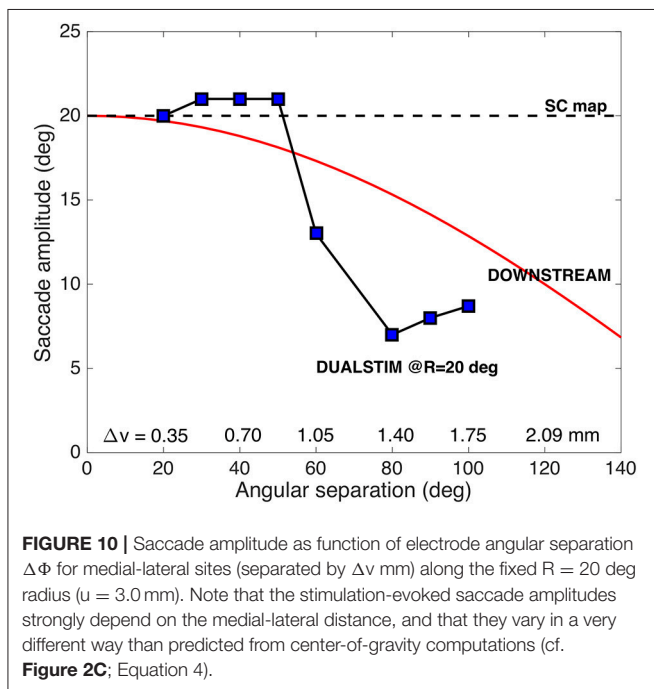
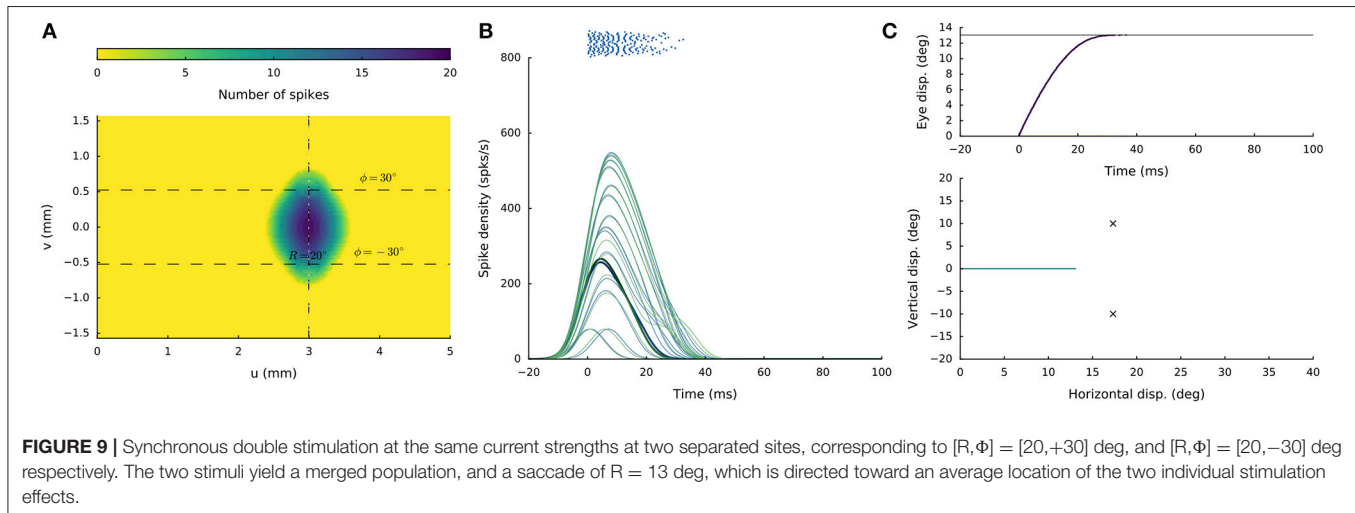
Weighted saccade averaging can also occur when the electrodes are positioned along the medial-lateral axis, but the effects resulted to depend strongly on both the electrode separation and on the strengths of the two currents. For example, when one electrode was kept fixed at the supra-threshold stimulation intensity of $I_{0,1} = 150$ pA, and the other electrode was varied between $I_{0,2} = 100$ –200 pA, the following pattern emerged for all angular separation conditions:

- (i) For currents below $I_{0,2} = 150$ pA, site 1 always fully dominated, and all saccades were directed toward the first site.
- (ii) Above $I_{0,2} = 150$ pA, site 2 dominated and saccades were directed to the second site.
- (iii) Only when the currents were equal, $I_{0,1} = I_{0,2} = 150$ pA, averaging was obtained according to the relationship seen in **Figure 9**. In other words, in these double-stimulation conditions the saccade direction behaved as a bistable variable. This response behavior is illustrated in **Figure 11** for an angular separation of 30 deg ($\Delta\nu = 0.52$ mm; black symbols).

True averaging of the saccade direction was only obtained when (i) the fixed stimulation current at site 1 was lowered to slightly above the threshold for evoking a saccade (e.g., to $I_{0,1} = 120$ pA), and (ii) the two sites were close together. **Figure 11** shows the results of such weighted stimulation effects for the same sites (blue symbols). The figure shows that from $I_{0,2} = 130$ pA onwards, a clear weighted averaging pattern was obtained, in which the saccade direction varied systematically with the difference in current strength. Note that for currents below about $I_{0,2} = 130$ pA, also the saccade amplitude started to decrease, as for these cases both currents were getting close to their saccade-evoking thresholds.

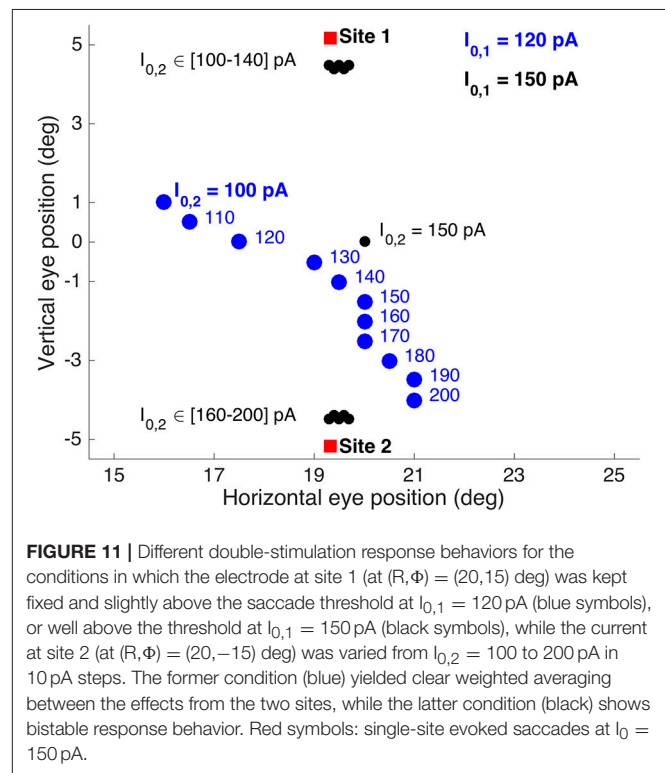
Double Stimulation With Delay

In a similar way as observed for the interactions along the medial-lateral coordinate (see sections Double Stimulation: Evoked Saccade Amplitude Depends on Medial-Lateral Separation and Lateral-Medial Double Stimulation at Different Current Strengths), imposing a temporal delay between the two supra-threshold electrode currents (when both at 150 pA) produced different response behaviors, depending on the electrode separations and current strengths. For supra-threshold stimulation at both sites, a curved saccade trajectory would only emerge when the delay was very short (typically, below 6 ms),



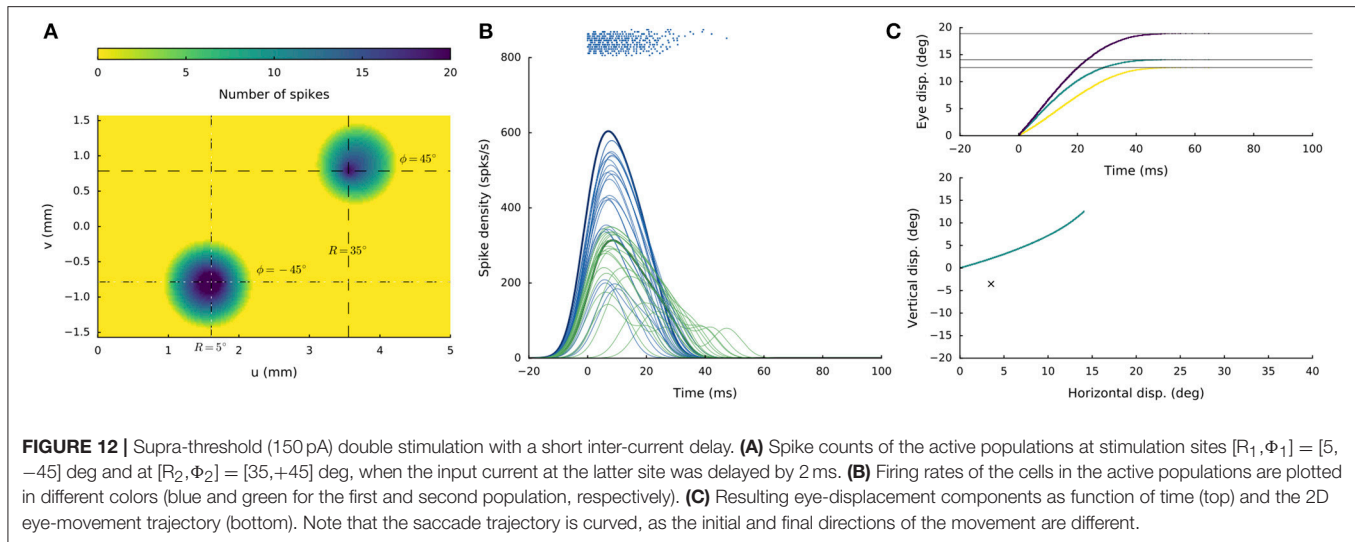
and the stimulation sites are separated in both the medial-lateral and rostral-caudal dimensions of the motor map. An example of such a stimulation condition is shown in **Figure 12**. The two sites were at $[R, \Phi] = [5, -45]$ and $[35, +45]$ deg, respectively, and the current strengths were 150 pA at both sites, whereby the stimulation pulse at the second site was delayed by 2 ms. Both electrodes set up a population response, leading to a curved saccade trajectory with an overall amplitude of $R = 19$ deg and a direction of about $\Phi = 40$ deg, which is a weighted average of the individual stimulus effects. When the delay was increased to 4 ms the initial direction of the saccade was horizontal curving toward the final site location in midflight of the response (not shown).

At delays above 5 ms, the saccade was invariably directed at the endpoint of the first site, as the second site would be strongly



inhibited by the activated first population. As a result, the second site would not be able to set up an appropriate population response to produce a colliding saccadic on its own.

When the stimulation sites and current strengths, as well as the delays were systematically varied, the occurrence of curved saccade trajectories resulted to be quite rare. Instead, we often obtained a bistable response behavior, in which a small change in one of the stimulation parameters (e.g., the current strength at the first electrode) could fully change the saccadic response from being directed to the first site, toward the second site.



An example of this bistable behavior on the stimulation conditions is shown in **Figure 13**, where the two sites were at $[R_1, \Phi_1] = [20, +30]$ deg and $[R_2, \Phi_2] = [40, -30]$ deg, respectively, and the delay was 10 ms. The stimulation current, $I_{0,2}$, was 150 pA in both cases, whereas $I_{0,1}$ was either 140 pA, or 130 pA. In the former condition, a straight saccade is directed toward site 1, whereas in the latter case, a straight saccade is made in the direction of site 2.

We systematically varied the inter-stimulus delay t_2 from (2, 5, 10, 20, 50) ms and $I_{0,1}$ from (200, 190, ..., 80) pA ($I_{0,2}$ fixed at 150 pA), and obtained similar bistable results for many cases. Note, however, that these two sites are separated by about 1.26 mm, which falls in the strongest inhibitory range of the lateral connectivity profile. In the situation of **Figure 12** the two sites are further apart, given weaker mutual inhibition and allowing more excitatory interactions (see **Figure 3B** and section Discussion).

DISCUSSION

Summary

Synchronous double stimulation in a spiking neural network model of the SC with Gaussian excitatory-inhibitory interactions results in saccade responses that display many of the features that have been reported in electrophysiological studies [9, 25, 34]: when the electrodes were located on an iso-direction line ($v = \text{constant}$) the resulting saccade amplitudes were a weighted average of the individual stimulus effects, with the current strengths acting as weighting parameters (**Figures 5–8**). When the electrodes were positioned along iso-eccentricity lines ($u = \text{constant}$), however, the response patterns appeared to be more complex: weighted averaging was obtained for low stimulation currents at nearby stimulation sites, but when the electrodes were moved further apart and/or the current levels increased, we obtained bistable response behavior (**Figures 9–11**). When a delay was introduced between the first and second stimulus pulse, the averaged saccade trajectories

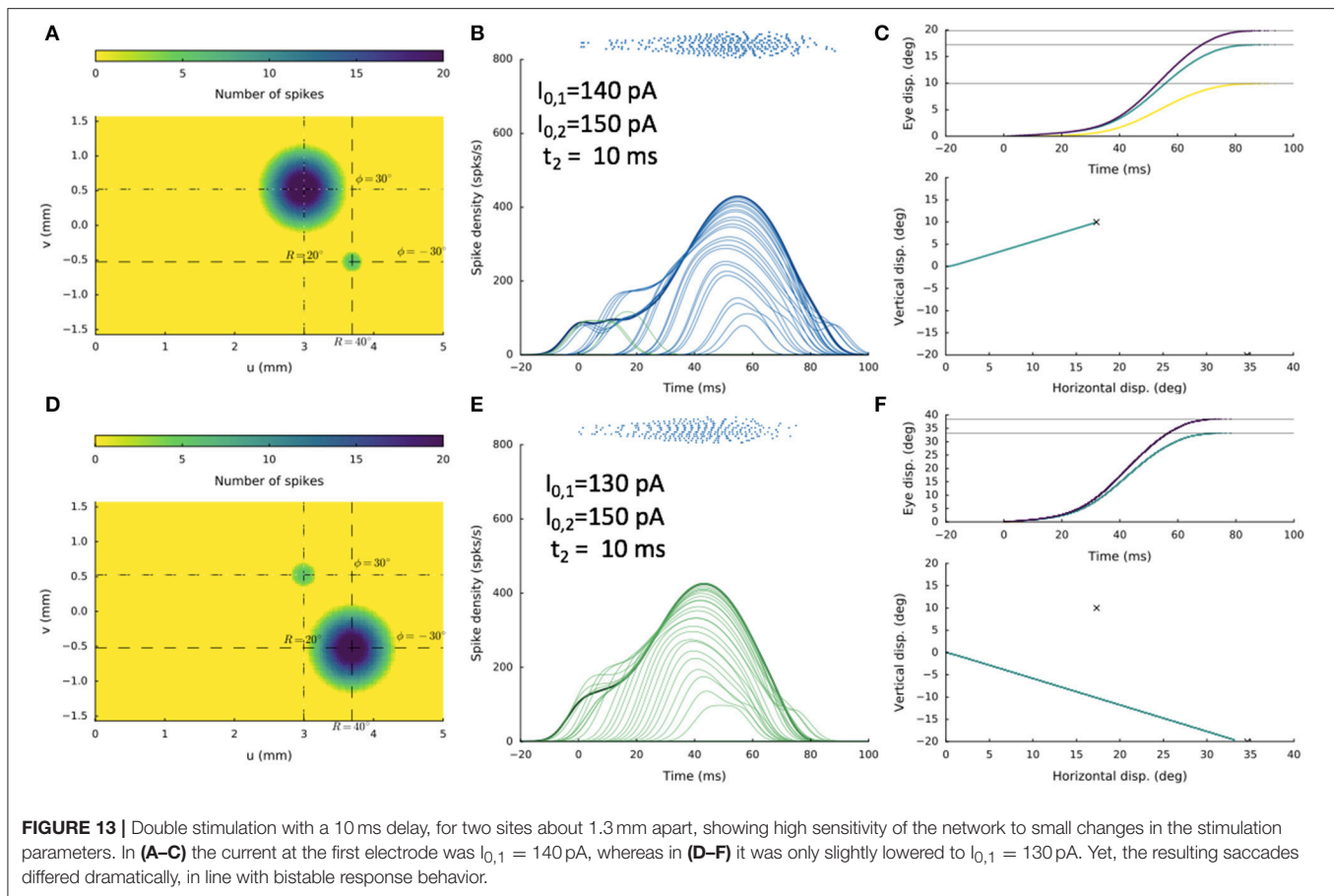
could become curved, provided the delay was short (< 6 ms; **Figure 12**). For longer delays, saccades were invariably directed toward the site evoked by the first electrode when its current intensity was above the normal saccade-initiation threshold (150 pA). In other cases, we obtained bistable response behavior, in which the saccade was directed either to the first site, or to the second site, without averaging (**Figure 13**).

The weighted averaging effects, which betray a nonlinearity in the system, are entirely due to the neural dynamics (Equations 6–7) and synaptic connectivity patterns (Equations 12–14) within the SC motor map, as the downstream motor circuitry in our model was taken entirely linear (Equation 3). Yet, the averaging results of our simulations do not correspond at all to the simple prediction of a center of gravity calculation at the level of the motor map either (Equation 4a; **Figure 2B**), as for iso-eccentricity stimulation the evoked saccade amplitudes varied strongly with the electrode separation (**Figure 10**), in a pattern that somewhat resemble the effect of downstream averaging. Whether these predictions truly deviate from observed experimental data on synchronous double stimulation is hard to tell, as precise measurements and quantification of this phenomenon are rare (e.g., 25, 34). The same may hold for the exact paths followed by curved trajectories evoked by delayed electrical double stimulation [25, 34, 39].

In what follows, we discuss these apparent discrepancies with the experimental data.

Model Structure

The subtle different behaviors observed for iso-direction vs. iso-eccentricity stimulation are likely caused by the differences in neural organization for the u - and v -coordinates in our model. The tuning parameters of the neuronal dynamics (the adaptive time constant, Equation 13) and the lateral synaptic projection strengths (the scaling parameter, Equation 14) both only vary with the rostral-caudal coordinate (u), and are assumed constant along iso-eccentricity lines.



These biophysical neural tunings were required to explain the firing behavior of collicular neurons under single-site visual stimulation conditions [8, 13, 23], and the nonlinear saccadic main sequence kinematics (see Introduction). From our single-unit recordings we noted that the peak firing rates of SC neurons in the center of the population decreased systematically with the saccade amplitude, meanwhile increasing their burst durations to keep the number of spikes in the saccade-related burst invariant across the motor map for slow, fast, small and large saccades. As single-site microstimulation produces normal saccadic eye movements, we argued that the same population activity would emerge during electrical stimulation and for natural visual stimulation. The neural population dynamics are then explained by synaptic lateral interactions, and are hardly influenced by the externally applied electrical stimulation current. We assumed that the stimulation current directly activates only a small subset of the neurons around the electrode. Indeed, under these assumptions, most single-site microstimulation results could be accounted for as well [26].

One discrepancy with experimental observations concerned the near-threshold behavior of the network: around the stimulation threshold, the network's saccades become much slower than main sequence (as evoked firing rates decrease), but their size (determined by the total number of spikes in the burst) remained unaffected. However, experiments have revealed

that near the threshold, saccades become both slower than main sequence *and* smaller [15, 35]. This would suggest that near threshold not only the firing rates are reduced, but also the number of spikes. The current model does not incorporate this possibility.

We here conjecture that the failure to produce different numbers of spikes for near-threshold conditions may also underlie the bistable character of our model to some of the double-stimulation conditions, and its reluctance to readily produce curved saccades. In double stimulation, the two electrodes exert a mutual inhibitory influence, which brings the weaker stimulation site to near- or below-threshold levels under many conditions. Indeed, when the stimulation sites fall in each other's strongest inhibitory zones, the bistable effects are nearly impossible to overcome (e.g., **Figures 11, 13**). On the other hand, when the stimulation electrodes are placed along the u -direction in the map, bi-stability is less common. This is probably due to the decreasing strength of the lateral connectivity patterns along this dimension, as dictated by Equation 14 (the most caudal sites exert nearly 25% less influence than the most rostral sites).

One possibility to overcome this discrepancy is to introduce variability (noise) in the neural population, e.g., at the level of the synaptic conductances (Equation 11), and at the adaptive time constants (Equation 13), that relies on the total input strength to the neuron (multiplicative noise; [8]). This will affect the total

number of spikes of the neuron, and therefore could potentially lead to smaller saccades for effectively weak inputs.

Untested Predictions

The neural interactions, imposed by the two separated electrodes, cause some interesting (and somewhat unexpected) behaviors of the neural firing properties, which so far have not been tested experimentally. Under single-site stimulation, the activity of the central cell, which encodes the ensuing saccade amplitude and direction, fully determines the firing-rate profile of all other cells, as well as the saccade kinematics (neural synchronization; e.g., **Figure 4**). Under double-stimulation at different nearby sites, however, the most active cells are no longer found at the stimulation electrodes, but at a location in between. The firing rates of these most active cells now determine the full saccade kinematics and the firing profiles of the other cells (e.g., **Figures 6, 7, 9**). Interestingly, the kinematics of the resulting saccades (which are slower) and the firing rates of these most active cells (which are higher) differ from the effects of single stimulation at that most active site. Unfortunately, it is difficult to test this prediction experimentally for the firing rates under electrical double stimulation, because of the strong electrical artifacts produced by the electrodes.

However, the effects of double stimulation on the emerging eye-movement kinematics can be readily assessed. As far as the main-sequence properties are concerned, averaging saccades under double visual stimulation appear to be slower than saccades of the same amplitude to a single visual stimulus, and the associated firing rates in the SC are lower (e.g., [46]). To our knowledge, the detailed velocity profiles under electrical double-stimulation have so far not been quantified in experimental studies.

Lateral Interactions

The simulations of electrical double stimulation made clear that the shape of the Mexican-hat profile affects the activity profiles of both active neuron populations and of the resulting saccades (e.g., **Figure 11**). The presence of lateral interactions within the SC has been well established by both anatomical and physiological evidence [28, 30, 33]. Modeling studies have suggested different synaptic interaction profiles, such as local excitation and global constant inhibition [37], or Mexican-hat type Gaussian profiles [45]. In the present study, we fixed the ranges of the excitatory and inhibitory interactions (σ_{exc} and σ_{inh}) for all cells and tuned their synaptic strengths in line with the proposal of Trappenberg et al. ([45]; Equation 14). Although it is conceivable that different profiles with shorter ranges could generate similar population activities (see below), anatomical studies so far do not allow to quantify the connectivity profiles and ranges, except for recent *in-vitro* studies [31, 32].

In contrast to the model of Van Opstal and Van Gisbergen [38], in the present model the effective range of the electrical current was assumed to be small (Equation 10; [26]). This assumption was inspired by recent findings from stimulation experiments with simultaneous calcium imaging in frontal

cortical tissue [27, 47]. In our model, the stimulation profile is subsequently combined with the Mexican-hat interaction function of Equations 12–14. We have shown earlier, using a static population model of the SC, that a weak global constant inhibition in combination with a delta function for the excitatory profile (i.e., only self-excitation) could yield saccade-averaging results if the current-spread function was a Gaussian with a much broader extent as in the present study, and whereby its width depended in a nonlinear way on the applied current strength [38].

Note that for network models such as these, including our own, the overall spatial effect of the stimulation (ignoring time) is in fact given by the convolution of the electrical stimulation profile with the weighting kernel of the excitatory-inhibitory interactions. Each cell's membrane potential is thus described by:

$$V_n(u, v) = \iint_{(u,v)_{\min}}^{(u,v)_{\max}} w_n(\sigma, \tau) \cdot I_{\text{INP}}(u - \sigma, v - \tau) \cdot d\sigma d\tau \quad (15)$$

which constitutes one equation for the membrane potential of neuron n , as a multiplicative combination of two functions. It is therefore conceivable that many potential functions could fulfill Equation 15. However, the nonlinear dynamics of the current model (Equations 6–7) makes a simple analytical approach to find the optimal solution that satisfies all experimental constraints not feasible. Further study is therefore required to analyze the effects of different profiles on the total network behavior across a wide range of sensory and electrical stimulation conditions.

As a final note, the electrical stimulation inputs were simply taken as constant rectangular pulses, instead of trains of short-duration stimulation pulses. In the latter case, which is physiologically more realistic, also the pulse intervals (stimulation frequency), pulse durations (stimulus train lengths), pulse heights, pulse interleave times, and pulse polarity may all play a role in the evoked E-saccades under single and double stimulation paradigms [24, 25, 34]. Incorporating these different stimulation parameter settings in our spiking neural-network model will require some tedious retuning of the network parameters, but may be worth the effort for its potential to generate novel neural dynamics.

AUTHOR CONTRIBUTIONS

AvO, BK: Writing manuscript, preparation of figures; BK: Model implementation, model simulations; AvO: Conceptualization.

ACKNOWLEDGMENTS

This work was supported by the European Commission through FP7 Marie Curie PEOPLE-2012-ITN, project NETT (grant 289146; BK), and by a Horizon 2020 ERC Advanced Grant, project ORIENT (grant 693400; AvO; BK). The Tesla K40 used for this research was donated by the NVIDIA Corporation.

REFERENCES

- Bahill AT, Clark MR, Stark L. The main sequence, a tool for studying human eye movements. *Math Biosci.* (1975) **204**:191–20. doi: 10.1016/0025-5564(75)90075-9
- Van Gisbergen JA, Van Opstal AJ, Schoenmakers JJM. Experimental test of two models for the generation of oblique saccades. *Exp Brain Res.* (1985) **57**:321–36. doi: 10.1007/BF00236538
- Van Gisbergen JA, Robinson DA, Gielen S. A quantitative analysis of generation of saccadic eye movements by burst neurons. *J Neurophysiol.* (1981) **45**:417–42. doi: 10.1152/jn.1981.45.3.417
- Smit AC, Van Opstal AJ, Van Gisbergen JAM. Component stretching in fast and slow oblique saccades in the human. *Exp Brain Res.* (1990) **81**:325–34. doi: 10.1007/BF00228123
- Harris CM, Wolpert DM. The main sequence of saccades optimizes speed-accuracy trade-off. *Biol Cybern.* (2006) **95**:21–9. doi: 10.1007/s00422-006-0064-x
- Tanaka H, Krakauer JW, Qian N. An optimization principle for determining movement duration. *J Neurophysiol.* (2006) **95**:3875–86. doi: 10.1152/jn.00751.2005
- Van Beers RJ. Saccadic eye movements minimize the consequences of motor noise. *PLoS ONE* 3:e2070. doi: 10.1371/journal.pone.0002070
- Goossens HH, Van Opstal AJ. (2012) Optimal control of saccades by spatial-temporal activity patterns in monkey Superior Colliculus. *PLoS Comput Biol.* **8**:e1002508. doi: 10.1371/journal.pcbi.1002508
- Robinson DA. Eye movements evoked by collicular stimulation in the alert monkey. *Vision Res.* (1972) **12**:1795–808. doi: 10.1016/0042-6989(72)90070-3
- Scudder A. A new local feedback model of the saccadic burst generator. *J Neurophysiol.* (1988) **59**:1455–75. doi: 10.1152/jn.1988.59.5.1455
- Moschovakis AK, Kitama T, Dalezio Y, Petit J, Brandi AM, Grantyn AA. An anatomical substrate for the spatiotemporal transformation. *J Neurosci.* (1998) **18**:10219–29. doi: 10.1523/JNEUROSCI.18-23-10219.1998
- Ottens FP, Van Gisbergen JAM, Eggermont JJ. Visuomotor fields of the superior colliculus: a quantitative model. *Vision Res.* (1986) **26**:857–73. doi: 10.1016/0042-6989(86)90144-6
- Goossens HH, Van Opstal AJ. Dynamic ensemble coding of saccades in the monkey Superior Colliculus. *J Neurophysiol.* (2006) **95**:2326–41. doi: 10.1152/jn.00889.2005
- Van Gisbergen JA, Van Opstal AJ, Tax AAM. Collicular ensemble coding of saccades based on vector summation. *Neuroscience* (1987) **21**:541–55. doi: 10.1016/0306-4522(87)90140-0
- Van Opstal AJ, Van Gisbergen JA, Smit AC. Comparison of saccades evoked by visual stimulation and collicular electrical stimulation in the alert monkey. *Exp Brain Res.* (1990) **79**:299–312. doi: 10.1007/BF00608239
- McIlwain JT. Lateral spread of neural excitation during microstimulation in intermediate gray layer of cat's superior colliculus. *J Neurophysiol.* (1982) **47**:167–78.
- Lee C, Rohrer WH, Sparks DL. Population coding of saccadic eye movements by neurons in the superior colliculus. *Nature* (1988) **332**:357–60. doi: 10.1038/332357a0
- Groh JM. Converting neural signals from place codes to rate codes. *Biol Cybern.* (2001) **85**:159–65. doi: 10.1007/s004220100249
- Quaia C, Lefèvre P, Optican LM. Model of the control of saccades by superior colliculus and cerebellum. *J Neurophysiol.* (1999) **82**:999–1018. doi: 10.1152/jn.1999.82.2.999
- Port NL, Wurtz RH. Sequential activity of simultaneously recorded neurons in the superior colliculus during curved saccades. *J Neurophysiol.* (2003) **90**:1887–903. doi: 10.1152/jn.01151.2002
- Walton MMG, Sparks DL, Gandhi NJ. Simulations of saccade curvature by models that place Superior Colliculus upstream from the local feedback loop. *J Neurophysiol.* (2005) **93**:2354–8. doi: 10.1152/jn.01199.2004
- Van Opstal AJ, Goossens HH. Linear ensemble coding in midbrain Superior Colliculus specifies the saccade kinematics. *Biol Cybern.* (2008) **98**:561–77. doi: 10.1007/s00422-008-0219-z
- Kasap B, Van Opstal AJ. A spiking neural network model of the midbrain superior colliculus that generates saccadic motor commands. *Biol Cybern.* (2017) **111**:249–68. doi: 10.1007/s00422-017-0719-9
- Stanford TR, Freedman EG, Sparks DL. Site and parameters of microstimulation: evidence for independent effects on the properties of saccades evoked from the primate superior colliculus. *J Neurophysiol.* (1996) **76**:3360–81. doi: 10.1152/jn.1996.76.5.3360
- Katnani HA, Gandhi NJ. The relative impact of microstimulation parameters on movement generation. *J Neurophysiol.* (2012) **108**:528–38. doi: 10.1152/jn.00257.2012
- Kasap B, Van Opstal AJ. Microstimulation in a spiking neural network model of the midbrain superior colliculus. *bioRxiv* [preprint] (2018). doi: 10.1101/424127
- Histed MH, Bonin V, Reid RC. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron* (2009) **63**:508–522. doi: 10.1016/j.neuron.2009.07.016
- Behan M, Kime NM. Intrinsic circuitry in the deep layers of the cat superior colliculus. *Vis Neurosci.* (1996) **13**:1031–42. doi: 10.1017/S0952523800007689
- Olivier E, Porter JD, May PJ. Comparison of the distribution and somato-dendritic morphology of tecto-tectal neurons in the cat and monkey. *Vis Neurosci.* (1998) **15**:903–22. doi: 10.1017/S095252389815513X
- Munoz DP, Istvan PJ. Lateral inhibitory interactions in the intermediate layers of the monkey superior colliculus. *J Neurophysiol.* (1998) **79**:1193–209. doi: 10.1152/jn.1998.79.3.1193
- Phongphanphanee P, Mizuno F, Lee PH, Yanagawa Y, Isa T, Hall WC. A circuit model for saccadic suppression in the superior colliculus. *J Neurosci.* (2011) **31**:1949–54. doi: 10.1523/JNEUROSCI.2305-10.2011
- Phongphanphanee P, Marino RA, Kaneda K, Yanagawa Y, Munoz DP, Isa T. Distinct local circuit properties of the superficial and intermediate layers of the rodent superior colliculus. *Eur J Neurosci.* (2014) **40**:2329–40. doi: 10.1111/ejn.12579
- Meredith MA, Ramoa AS. Intrinsic circuitry of the superior colliculus: pharmacophysiological identification of horizontally oriented inhibitory interneurons. *J Neurophysiol.* (1998) **79**:1597–602. doi: 10.1152/jn.1998.79.3.1597
- Noto CT, Gnadt JW. Saccade trajectories evoked by sequential and colliding stimulation of the monkey superior colliculus. *Brain Res.* (2009) **1295**:99–118. doi: 10.1016/j.brainres.2009.07.069
- Katnani HA, Van Opstal AJ, Gandhi NJ. A test of spatial-temporal decoding mechanisms in the superior colliculus. *J Neurophysiol.* (2012) **107**:2442–52. doi: 10.1152/jn.00992.2011
- Gandhi NJ, Katnani HA. Motor functions of the superior colliculus. *Annu Rev Neurosci.* (2011) **34**:205–31. doi: 10.1146/annurev-neuro-061010-113728
- Van Opstal AJ, Van Gisbergen JA. A nonlinear model for collicular spatial interactions underlying the metrical properties of electrically elicited saccades. *Biol Cybern.* (1989) **60**:171–83. doi: 10.1007/BF00207285
- Van Opstal AJ, Van Gisbergen JAM. A model for collicular efferent mechanisms underlying the generation of saccades. *Brain Behav Evol.* (1989) **33**:90–4. doi: 10.1159/000115906
- Van der Willigen RF, Goossens HH, Van Opstal AJ. Linear visuomotor transformations in midbrain superior colliculus control saccadic eye movements. *J Integr Neurosci.* (2011) **10**:277–301. doi: 10.1142/S0219635211002750
- Van Opstal AJ. *The Auditory System and Human Sound Localization Behavior*. Amsterdam: Elsevier Publishers, Academic Press (2016).
- Nickolls J, Buck I, Garland M, Skadron K. Scalable parallel programming with CUDA. *AMC Queue* (2008) **6**:40–53. doi: 10.1145/1365490.1365500
- Kasap B, Van Opstal AJ. Dynamic parallelism for synaptic updating in GPU-accelerated spiking neural network simulations. *Neurocomputing* (2018) **302**:55–65. doi: 10.1016/j.neucom.2018.04.007
- Brette R, Gerstner W. Adaptive exponential integrate-and-fire model as an effective description of neuronal activity. *J Neurophysiol.* (2005) **94**:3637–42. doi: 10.1152/jn.00686.2005
- Touboul J, Brette R. Dynamics and bifurcations of the adaptive exponential integrate-and-fire model. *Biol Cybern.* (2008) **99**:319–34. doi: 10.1007/s00422-008-0267-4

45. Trappenberg TP, Dorris MC, Munoz DP, Klein RM. A model of saccade initiation based on the competitive integration of exogenous and endogenous signals in the Superior Colliculus. *J Cogn Neurosci.* (2001) **13**:256–71. doi: 10.1162/089892901564306
46. Van Opstal AJ, Van Gisbergen JA. Role of monkey superior colliculus in saccade averaging. *Exp Brain Res.* (1990) **79**:143–9. doi: 10.1007/BF00228883
47. Histed MH, Ni AM, Maunsell JH. Insights into cortical mechanisms of behavior from microstimulation experiments. *Prog Neurobiol.* (2013) **103**:115–30. doi: 10.1016/j.pneurobio.2012.01.006

Conflict of Interest Statement: 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.

Copyright © 2018 Kasap and van Opstal. 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.



ExGUtills: A Python Package for Statistical Analysis With the ex-Gaussian Probability Density

Carmen Moret-Tatay^{1*}, Daniel Gamermann², Esperanza Navarro-Pardo³ and Pedro Fernández de Córdoba Castellá⁴

¹ Department of Neuropsychology, Methodology, Basic and Social Psychology, Faculty of Psychology, Universidad Católica de Valencia San Vicente Mártir, Valencia, Spain, ² Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ³ Department of Developmental and Educational Psychology, Faculty of Psychology, Universitat de Valencia, Valencia, Spain, ⁴ Grupo de Modelización Interdisciplinar, Instituto Universitario de Matemática Pura y Aplicada, InterTech, Universitat Politècnica de València, Valencia, Spain

OPEN ACCESS

Edited by:

Axel Hutt,
German Meteorological Service,
Germany

Reviewed by:

Denis Cousineau,
University of Ottawa, Canada
Miguel A. Vadillo,
Universidad Autónoma de Madrid,
Spain

*Correspondence:

Carmen Moret-Tatay
mariaacarmen.moret@ucv.es

Specialty section:

This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 29 December 2017

Accepted: 11 April 2018

Published: 01 May 2018

Citation:

Moret-Tatay C, Gamermann D,
Navarro-Pardo E and Fernández de
Córdoba Castellá P (2018) ExGUtills: A
Python Package for Statistical Analysis
With the ex-Gaussian Probability
Density. *Front. Psychol.* 9:612.
doi: 10.3389/fpsyg.2018.00612

The study of reaction times and their underlying cognitive processes is an important field in Psychology. Reaction times are often modeled through the ex-Gaussian distribution, because it provides a good fit to multiple empirical data. The complexity of this distribution makes the use of computational tools an essential element. Therefore, there is a strong need for efficient and versatile computational tools for the research in this area. In this manuscript we discuss some mathematical details of the ex-Gaussian distribution and apply the ExGUtills package, a set of functions and numerical tools, programmed for python, developed for numerical analysis of data involving the ex-Gaussian probability density. In order to validate the package, we present an extensive analysis of fits obtained with it, discuss advantages and differences between the least squares and maximum likelihood methods and quantitatively evaluate the goodness of the obtained fits (which is usually an overlooked point in most literature in the area). The analysis done allows one to identify outliers in the empirical datasets and criteriously determine if there is a need for data trimming and at which points it should be done.

Keywords: response times, response components, python, ex-Gaussian fit, significance testing

1. INTRODUCTION

The reaction time (RT) has become one of the most popular dependent variables in cognitive psychology. Over the last few decades, much research has been carried out on problems focusing exclusively on success or fail in trials during the performance of a task, emphasizing the importance of RT variables and their relationship to underlying cognitive processes (Sternberg, 1966; Wickelgren, 1977; McVay and Kane, 2012; Ratcliff et al., 2012). However, RT has a potential disadvantage: its skewed distribution. One should keep in mind that in order to perform data analysis, it is preferable that the data follow a known distribution. If the distribution is not symmetrical, it is possible to carry out some data transformation techniques (e.g., the Tukey scale for correcting skewness distribution), or to apply some trimming techniques, but with these techniques, statistics may be altered (in other words a high concentration of cases in a given range

may be favored and as a result, statistics can appear biased). Moreover, transformations can affect the absolute value of the data or modify the relative distances between data. When conducting trimming it is not easy to distinguish noisy data from valid information, or in other words, to set the limits between outliers and extreme data (Heathcote et al., 1991). Whether we include or exclude outliers often depends on the reason why they might occur, dealing with the decision to classify them as variability in the measurement or as an experimental error. Another option, for the analysis of skewed data, is to characterize them with a known skewed distribution. This procedure allows one to determine the probability of an event based on the statistical model used to fit the data. A common problem with this approach is to estimate the parameters that characterize the distribution. In practice, when one wants to find out the probability for an event numerically, a quantified probability distribution is required.

Going back to the point on characterizing data with a specific distribution, there is one distribution that has been widely employed in the literature when fitting RT data: the exponentially modified Gaussian distribution (West, 1999; Leth-Steensen et al., 2000; West and Alain, 2000; Balota et al., 2004; Hervey et al., 2006; Epstein et al., 2011; Gooch et al., 2012; Navarro-Pardo et al., 2013). This distribution is characterized by three parameters, μ , σ and τ . The first and second parameters (μ and σ), correspond to the average and standard deviation of the Gaussian component, while the third parameter (τ) is the decay rate of the exponential component. This distribution provides good fits to multiple empirical RT distributions (Luce, 1986; Lacouture and Cousineau, 2008; Ratcliff and McKoon, 2008), however there are currently no published statistical tables available for significance testing with this distribution, though there are softwares like S-PLUS (Heathcote, 2004) or PASTIS (Cousineau and Larochelle, 1997) and programming language packages available for R, MatLab or Mathematica.

In this article we present a package, developed in Python, for performing statistical and numerical analysis of data involving the ex-Gaussian function. Python is a high-level interpreted language. Python and R are undoubtedly two of the most widespread languages, as both are practical options for building data models with a lot of community support. However, the literature seems to be rather scarce in terms of computations with the ex-Gaussian function in Python. The package presented here is called ExGUtils (from ex-Gaussian Utilities), it comprises functions for different numerical analysis, many of them specific for the ex-Gaussian probability density.

The article is organized as follows: in the next section we present the ex-Gaussian distribution, its parameters and a different way in which the distribution can be parameterized. Following this, we discuss two fitting procedures usually adopted to fit probability distributions: the least squares and the maximum likelihood. In the third section we present the ExGUtils module and we apply it in order to fit experimental data, evaluate the goodness of the fits and discuss the main differences in the two fitting methods. In the last section we present a brief overview.

2. THE ex-GAUSSIAN DISTRIBUTION AND ITS PROBABILITY DENSITY

Given a randomly distributed X variable that can assume values between minus infinity and plus infinity with probability density given by the gaussian distribution,

$$g(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right), \quad (1)$$

and a second random Y variable that can assume values between zero and plus infinity with probability density given by an exponential distribution,

$$h(x) = \frac{1}{\tau} e^{-\frac{x}{\tau}}, \quad (2)$$

let's define the Z variable as the sum of the two previous random variables: $Z = X + Y$.

The gaussian distribution has average μ and standard deviation σ , while the average and standard deviation of the Y variable will be both equal to τ . The Z variable will also be a random variable, whose average will be given by the sum of the averages of X and Y and whose variance will be equal to the sum of the variances of X and Y :

$$M = \mu + \tau \quad (3)$$

$$S^2 = \sigma^2 + \tau^2 \quad (4)$$

Defined as such, the variable Z has a probability density with the form (Grushka, 1972):

$$f(x) = \frac{1}{2\tau} \exp\left(\frac{1}{2\tau}\left(2\mu + \frac{\sigma^2}{\tau} - 2x\right)\right) \operatorname{erfc}\left(\frac{\mu + \frac{\sigma^2}{\tau} - x}{\sqrt{2}\sigma}\right) \quad (5)$$

which receives the name of ex-Gaussian distribution (from exponential modified gaussian distribution). The erfc function is the complementary error function. One must be careful, for μ and σ are NOT the average and standard deviation for the ex-Gaussian distribution, instead the average and variance of the ex-Gaussian distribution is given by Equations (3)–(4): $M = \mu + \tau$ and $S^2 = \sigma^2 + \tau^2$. On the other hand, a calculation of the skewness of this distribution results in:

$$K = \int_{-\infty}^{\infty} \left(\frac{x-M}{S}\right)^3 f(x) dx = \frac{2\tau^3}{(\sigma^2 + \tau^2)^{\frac{3}{2}}}, \quad (6)$$

While the gaussian distribution has null skewness, the skewness of the exponential distribution is exactly equal to two. As a result the skewness of the ex-Gaussian has an upper bound equal to two in the limit $\sigma \ll \tau$ (when the exponential component dominates) and a lower bound equal to zero in the limit $\sigma \gg \tau$ (when the gaussian component dominates).

Let's parameterize the ex-Gaussian distribution in terms of its average M , standard deviation S and a new skewness parameter

$\lambda = \sqrt[3]{\frac{K}{2}}$. Defined in this way, the λ parameter can have values between 0 and 1. Now, defining the standard coordinate z ($z = \frac{x-M}{S}$) one can have the ex-Gaussian distribution normalized for average 0 and standard deviation 1 in terms of a single parameter, its asymmetry λ :

$$f_{\lambda}(z) = \frac{1}{2\lambda} \exp\left(\frac{1}{2\lambda^2}(-2z\lambda - 3\lambda^2 + 1)\right) \operatorname{erfc}\left(\frac{-z + \frac{1}{\lambda} - 2\lambda}{\sqrt{2}\sqrt{1-\lambda^2}}\right). \quad (7)$$

in this case, in terms of λ , the parameters μ , σ and τ are given by:

$$\mu = -\lambda \quad (8)$$

$$\sigma = \sqrt{1-\lambda^2} \quad (9)$$

$$\tau = \lambda. \quad (10)$$

Thus, the ex-gaussian represents a family of distributions that can be parametrized in terms of their asymmetry. Ranging from the

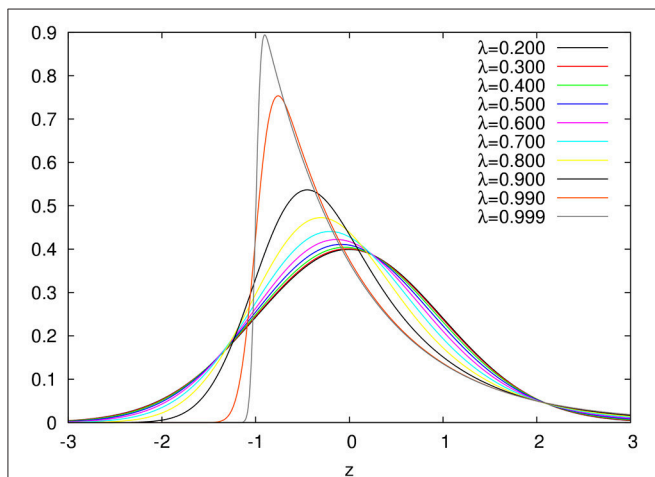


FIGURE 1 | ex-Gaussian distributions for different values of the λ asymmetry parameter.

exponential (maximum asymmetry in the limit when $\lambda = 1$) to a gaussian (symmetrical distribution in the limit when $\lambda = 0$).

In **Figure 1**, we show plots for the ex-Gaussian function for different values of the parameter λ . We should note that for very small values of λ (less than around 0.2), the ex-Gaussian is almost identical to the gaussian function (see **Figure 2**)¹.

Given a probability density, an important function that can be calculated from it is its cumulative distribution (its left tail), which is the result of the integral

$$F(z) = \int_{-\infty}^z f(x) dx. \quad (11)$$

The importance of this function is that given the cumulative distribution one is able to calculate the probability of an event. For the ex-gaussian, the expression for its cumulative distribution is given by:

$$F(x) = \frac{1}{2} \operatorname{erfc}\left(-\frac{x-\mu}{\sqrt{2}\sigma}\right) - \frac{1}{2} \exp\left(\frac{\sigma^2}{\tau^2} - \frac{x-\mu}{\tau}\right) \operatorname{erfc}\left(-\frac{\frac{x-\mu}{\sigma} - \frac{\sigma}{\tau}}{\sqrt{2}}\right) \quad (12)$$

Let's also define z_{α} , the value of z for which the right tail of the distribution has an area equal to α :

$$\alpha = \int_{z_{\alpha}}^{\infty} f(x) dx. \quad (13)$$

$$1 - F(z_{\alpha}) = \alpha \quad (14)$$

so, solving the Equation (14), one is able to obtain the value of z_{α} for any given α .

¹In this cases, the numerical evaluation of the ex-Gaussian distribution in Equation (5) becomes unstable and one can without loss (to a precision of around one part in a million) approximate the ex-Gaussian by a gaussian distribution.

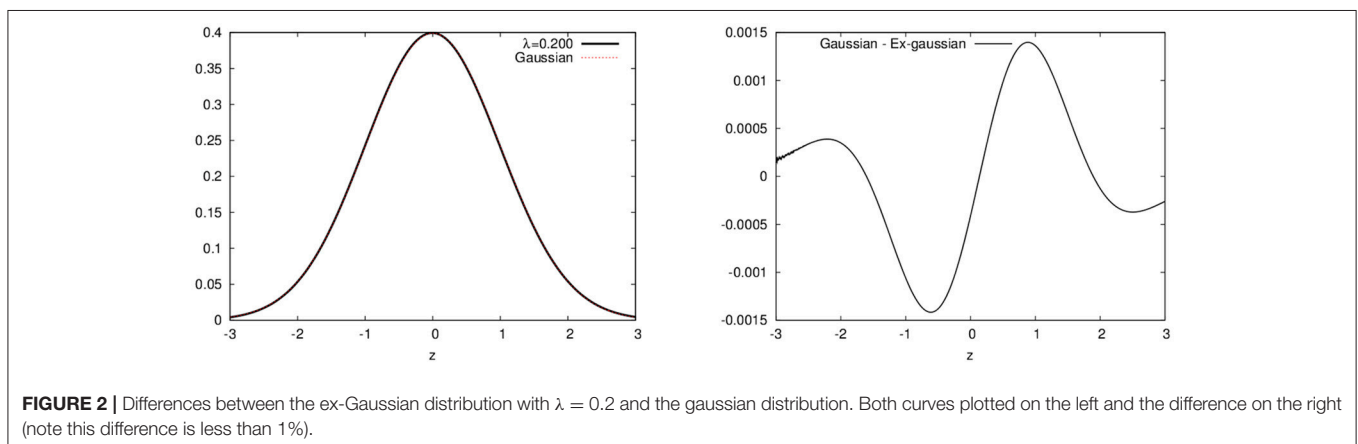


FIGURE 2 | Differences between the ex-Gaussian distribution with $\lambda = 0.2$ and the gaussian distribution. Both curves plotted on the left and the difference on the right (note this difference is less than 1%).

3. FITTING THE PROBABILITY DISTRIBUTION

We are interested in the following problem: given a dataset, to estimate the parameters μ , σ and τ that, plugged into Equation (5), best fit the data.

We must now define what it means to best fit the data. Different approaches here will result in different values for the parameters. The most trivial approach would be to say that the best parameters are those that result in the fitted ex-Gaussian distribution with the same statistical parameters: average (M), standard deviation (S) and asymmetry (K or λ). So, one can take the dataset, calculate M , S , and K and use the relations between them and the parameters μ , σ and τ :

$$M = \mu + \tau \quad (15)$$

$$S = \sqrt{\sigma^2 + \tau^2} \quad (16)$$

$$\lambda = \sqrt[3]{\frac{K}{2}} = \frac{\tau}{\sqrt{\sigma^2 + \tau^2}} \quad (17)$$

$$\mu = M - S\lambda \quad (18)$$

$$\sigma = S\sqrt{1 - \lambda^2} \quad (19)$$

$$\tau = S\lambda \quad (20)$$

This method of evaluating the parameters from the statistic (*momenta*) is known as the method of the moments as is usually the worst possible approach given the resulting bias. For instance, in some experiments, one finds the K parameter bigger than 2 (or $\lambda > 1$) and from Equation (17) one sees that, in order to have $K > 2$, σ cannot be a real number.

Another approach is to find the parameters that minimize the sum of the squared differences between the observed distribution and the theoretical one (least squares). In order to do that, one must, from the dataset, construct its distribution (a histogram), which requires some parametrization (dividing the whole range of observations in fixed intervals). Since a potentially arbitrary choice is made here, the results might be dependent on this choice. When analyzing data, we will study this dependency and come back to this point.

The last approach we will study is the maximum likelihood method. The function in Equation (5) is a continuous probability distribution for a random variable, which means that $f(x)dx$ can be interpreted as the probability that a observation of the random variable will have the x value (with the infinitesimal uncertainty dx). So, given a set of N observations of the random variable, $\{x_i\}$, with $i = 1, 2, \dots, N$, the likelihood \mathcal{L} is defined as the probability of such a set, given by:

$$\mathcal{L} = \prod_{i=1}^N f(x_i; \mu, \sigma, \tau) \quad (21)$$

$$\ln \mathcal{L} = \sum_{i=1}^N \ln (f(x_i; \mu, \sigma, \tau)) \quad (22)$$

The maximum likelihood method consists in finding the parameters μ , σ and τ that maximize the likelihood \mathcal{L} (or its

logarithm² $\ln \mathcal{L}$). Note that in this approach, one directly uses the observations (data) without the need of any parametrization (histogram).

In both approaches, least squares and maximum likelihood, one has to find the extreme (maximum or minimum) of a function. The numerical algorithm implemented for this purpose is the steepest descent/ascent (descent for the minimum and ascent for the maximum). The algorithm consists in interactively changing the parameters of the function by amounts given by the gradient of the function in the parameter space until the gradient falls to zero (to a certain precision). There are other optimization methods, like the simplex (Van Zandt, 2000; Cousineau et al., 2004), which also iteratively updates the parameters (in the case of the simplex without the need to compute the gradients). We chose to implement steepest ascent in order to gain in efficiency: since one is able to evaluate the gradients, this greedy algorithm should converge faster than the sample techniques used by simplex. But in any case, both algorithms (steepest descent and simplex) should give the same results, since both search the same maximum or minimum.

4. THE EXGUTILS MODULE

ExGUtils is a python package with two modules in its 3.0 version: one purely programmed in python (`pyexg`) and the other programmed in C (`uts`). The advantage of having the functions programmed in C is speed, stability and numerical precision.

As mentioned, the package has two modules: `pyexg` and `uts`. The first one comprises all functions with source code programmed in python, some of which depend on the `numpy`, `scipy` and `random` python packages. On the other hand, the module `uts` contains functions with source code programmed in C. In Table 1 one can find a complete list of all functions contained in both modules and the ones particular to each one. The source distribution of the ExGUtils module comes with a manual which explains in more detail and with examples the functions.

5. APPLICATIONS

We use here the ExGUtils package in order to analyze data from the experiment in Navarro-Pardo et al. (2013). From this work, we analyse the datasets obtained for the reaction times of different groups of people in recognizing different sets of words in two possible experiments (yes/no and go/nogo). In the Appendix B we briefly explain the datasets analyzed here (which are provided as Supplementary Material for download).

In our analysis, first each dataset is fitted to the ex-Gaussian distribution through the three different approaches aforementioned:

- moments → Estimating the parameters through the sample statistics Equations (18–20).

²Note that, since the logarithm is an monotonically increasing function, the maximal argument will result in the maximum value of the function as well.

TABLE 1 | Functions present in the package modules.

Module	Function	Brief description
Present	drand	Returns a random number with homogeneous distribution between 0 and 1
in	exp_rvs	Returns a random number with exponential distribution between 0 and infinity
both	gauss_rvs	Returns a random number with gaussian distribution between minus infinity and infinity
modules	exg_rvs	Returns a random number with ex-Gaussian distribution between minus infinity and infinity
	gauss_pdf	Evaluates the gaussian distribution at a given point
	gauss_cdf	Evaluates the gaussian cumulative distribution at a given point
	exg_pdf	Evaluates the ex-Gaussian distribution at a given point
	exg_cdf	Evaluates the ex-Gaussian cumulative distribution at a given point
	exg_lamb_pdf	Evaluates the ex-Gaussian distribution parameterized in terms of its asymmetry at a given point
	exg_lamb_cdf	Evaluates the ex-Gaussian cumulative distribution parameterized in terms of its asymmetry at a given point
	pars_to_stats	Given the parameters μ , σ and τ , evaluates the corresponding statistics M , S , and K
	stats_to_pars	Given the statistics M , S and K , evaluates the corresponding parameters μ , σ and τ
	histogram	Given a set of observations, produces an histogram
	stats	Given a set of observations, returns the statistics M , S , and K
	stats_his	Given a set of observations, presented as a histogram, returns the statistics M , S , and K
	correlation	Given a set of points, returns the linear correlation coefficient for the points
	minsquare	Given a set of points, fits a polynomial to the data using the least square method
	exgLKHD	Evaluates the likelihood and its gradient in the parameter space for a dataset in a given point of the parameter space
	maxLKHD	Evaluates the parameters μ , σ and τ that maximize the likelihood for a given dataset
	exgSQR	Evaluates the sum of squared differences and its gradient in the parameter space for an histogram in a given point of the parameter space
	minSQR	Evaluates the parameters μ , σ and τ that minimize the sum of squared differences for a given histogram
Only	int_points_gauss	Creates a point partition of an interval for evaluating a
in		gaussian integral
uts	intsum	Evaluates the gaussian integral for a function calculated at the points in a gaussian partition
Only	zero	Finds the zero of an equation
in	ANOVA	Performs an ANOVA test
pyexg	integral	Evaluates an integral

In python type `help(FUNC)` (where `FUNC` should be the name of a given function), in order to obtain the list of arguments that each function should receive and in which order.

- minSQR → Estimation through least square method, using as initial point in the steepest descent algorithm the μ , σ and τ obtained from the method of moments above³.
- maxLKHD → Estimation through maximum likelihood method, using as initial point in the steepest ascent algorithm the μ , σ and τ obtained from the method of moments³.

In **Table 2**, one can see the estimated parameters and the corresponding statistics for the different experiments. From the table, one sees that in the case of the experiments performed with young people, the value of the skewness, K , is bigger than two. This happens because of a few atypical measurements far beyond the bulk of the distribution. In fact, many researches opt for trimming extreme data, by “arbitrarily” choosing a cutoff and removing data points beyond this cutoff. One must, though, be careful for the ex-Gaussian distribution does have a long right tail, so we suggest a more criterious procedure:

Having the tools developed in `EXGUTILS`, one can use the parameters obtained in the fitting procedures (either minSQR or maxLKHD) in order to estimate a point beyond which one should find no more than, let's say, 0.1% of the distribution. In the Appendix A (Supplementary Material), the Listing 1 shows a quick python command line in order to estimate this point in the case of the young_gng experiment. The result informs us that, in principle, one should not expect to have more than 0.1% measurements of reaction times bigger than 1472.84 ms if the parameters of the distribution are the ones adjusted by maxLKHD for the young_gng empirical data. In fact, in this experiment, one has 2396 measurements of reaction times, from those, 8 are bigger than 1472.8 ms (0.33%). If one now calculates the statistics for the data, removing these 8 outliers, one obtains:

moments: $M = 593.80$ $S = 154.30$ $K = 1.91$ $\mu = 441.82$
 $\sigma = 26.67$ $\tau = 151.98$

minSQR: $M = 590.11$ $S = 142.44$ $K = 1.67$ $\mu = 455.96$
 $\sigma = 47.89$ $\tau = 134.14$

maxLKHD: $M = 593.80$ $S = 148.44$ $K = 1.69$ $\mu = 453.52$
 $\sigma = 48.52$ $\tau = 140.29$

In **Figure 3** one can see the histogram of data plotted along with three ex-Gaussians resulting from the above parameters.

Now, one might ask, having these different fits for the same experiment, how to decide which one is the best? Accepting the parameters of a fit is the same as accepting the null hypothesis that the data measurements come from a population with an ex-Gaussian distribution with the parameters given by the ones obtained from the fit. In Clauset et al. (2009) the authors suggest a procedure in order to estimate a p -value for this hypothesis when the distribution is a power-law. One can generalize the procedure for any probability distribution, like the ex-Gaussian, for example:

³In the cases where K was bigger than 2, the initial parameters were calculated as if $K = 1.9$. Note that the final result of the search should not depend on the initial search point if it starts close to the local maximum/minimum.

TABLE 2 | Parameters and statistics obtained with the three fitting methods.

Experiment	Moments					minSQR					maxLKHD							
	M	S	K	μ	σ	τ	M	S	K	μ	σ	τ	M	S	K	μ	σ	τ
elder_gng	831.14	318.95	1.75	526.06	93.02	305.08	841.45	334.97	1.85	515.16	75.76	326.29	831.14	318.39	1.79	524.49	85.67	306.65
elder_hfng	798.55	310.00	1.94	491.52	42.78	307.03	803.42	308.32	1.82	504.66	76.18	298.76	796.63	312.32	1.90	489.39	56.11	307.24
elder_hfyn	826.15	278.61	1.62	566.56	101.16	259.59	810.03	256.41	1.68	568.03	84.74	242.00	826.15	275.10	1.75	563.06	80.40	263.08
elder_lfng	863.73	324.53	1.60	562.65	121.14	301.08	887.69	371.88	1.90	522.22	68.76	365.46	863.73	339.88	1.87	531.43	71.38	332.30
elder_lfyn	884.53	315.93	1.59	591.97	119.27	292.55	882.16	310.04	1.77	584.33	86.12	297.84	884.53	308.13	1.62	597.54	112.18	286.98
elder_pseudo	1189.64	416.92	0.88	872.59	270.73	317.05	1233.90	518.22	1.79	734.21	137.33	499.69	1189.64	447.89	1.60	773.74	166.24	415.90
elder_yn	854.88	298.93	1.63	575.78	107.04	279.11	846.94	286.75	1.75	572.77	84.01	274.17	854.88	292.72	1.68	578.59	96.69	276.29
young_gng	597.90	169.90	2.71	-	-	-	590.36	142.94	1.66	455.98	48.71	134.38	597.90	154.25	1.72	451.09	47.33	146.81
young_hfng	562.94	141.88	3.04	-	-	-	555.47	115.19	1.54	449.84	45.93	105.63	562.94	126.30	1.65	444.47	43.77	118.47
young_hfyn	621.16	176.99	3.88	-	-	-	610.51	128.15	1.53	493.32	51.86	117.19	621.16	148.38	1.65	482.08	51.70	139.08
young_lf	632.96	187.61	2.46	-	-	-	625.36	161.43	1.65	473.80	55.61	151.55	632.96	173.23	1.71	468.51	54.43	164.45
young_lfng	632.96	187.61	2.46	-	-	-	625.36	161.43	1.65	473.80	55.61	151.55	632.96	173.23	1.71	468.51	54.43	164.45
young_lfyn	668.57	184.88	2.10	-	-	-	660.23	165.16	1.60	507.06	61.77	153.18	668.58	176.56	1.70	501.24	56.31	167.34
young_pseudo	722.53	190.36	2.37	-	-	-	718.60	175.25	1.64	554.66	61.92	163.94	722.53	180.58	1.68	552.03	59.47	170.50
young_yn	644.37	182.41	2.90	-	-	-	635.20	148.99	1.62	496.41	54.18	138.78	644.37	164.31	1.69	488.96	53.34	155.42

For each one of the three procedures, we present the obtained parameters and the correspondent statistics. Note that for the method of the moments, the statistics is directly obtained from the dataset and the parameters are evaluated based on this statistic from Equations (18–20), while the other methods obtain the parameters from the dataset and the statistics presented in the table is evaluated based on these parameters from the expressions in eqs. Some conditions were not estimated because of a few atypical measurements far beyond the bulk of the distribution (K larger than 2), (15–17).

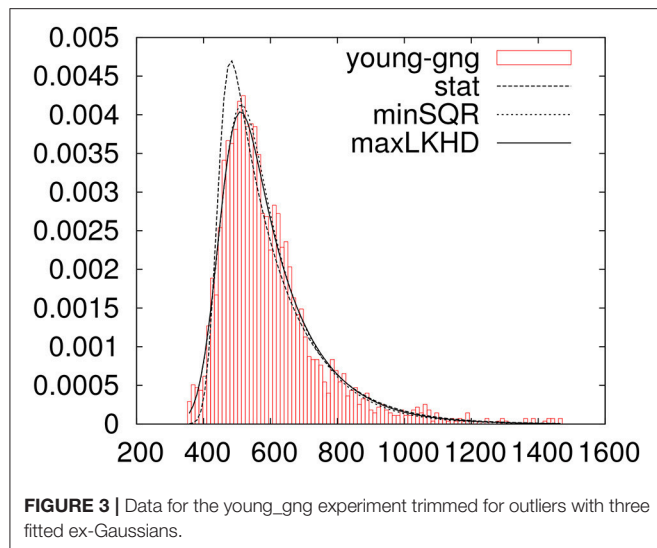


FIGURE 3 | Data for the young_gng experiment trimmed for outliers with three fitted ex-Gaussians.

- Take a measure that quantifies the distance between the data and the fitted theoretical distribution. One could use $\ln \mathcal{L}$ or χ^2 , but, as our fitting procedures maximize or minimize these measures, as the authors in Clauset et al. (2009) suggest, in order to avoid any possible bias, we evaluate the Kolmogorov-Smirnov statistic, which can be calculated for reaction-time data without the need of any parametrization.
- Randomly generate many data samples of the same size as the empirical one using the theoretical distribution with the parameters obtained from the fit to the empirical data.
- Fit each randomly generated data sample to the theoretical distribution using the same fit procedure used in the case of the empirical data.
- Evaluate the Kolmogorov-Smirnov statistic between the random sample and its fitted theoretical distribution.

Following this procedure, one can evaluate the probability that a random data sample, obtained from the fitted distribution, has a bigger distance to the theoretical curve than the distance between the empirical data and its fitted distribution. If this probability is higher than the confidence level one is willing to work with, one can accept the null hypothesis knowing that the probability that one is committing a type I error if one rejects the null hypothesis is p .

In the Appendix A (Supplementary Material) we provide listings with the implementation, in python via the ExGUtils package, of the functions that evaluate this p probability and the Kolmogorov-Smirnov statistic. In Table 3 we provide the values of p obtained for the experiments, using minSQR and maxLKHD approaches (p_1 and p_2 , respectively).

We can see that there are some discrepancies in Table 3. Sometimes minSQR seems to perform better, sometimes maxLKHD. One might now remember that the minSQR method depends on a parametrization of the data. In order to perform the fit, one needs to construct a histogram of the data, and there is an arbitrary choice in the number of intervals one divides the data into. In the fits performed till now, this number is set to be the

TABLE 3 | Probabilities p_1 and p_2 for the fits.

Experiment	minSQR		maxLKHD	
	KS	p_2 ($\bar{KS} \pm sd$)	KS	p_1 ($\bar{KS} \pm sd$)
elder_gng	64.52	0.001 (29.47 \pm 8.12)	38.89	0.096 (29.96 \pm 12.54)
elder_hfgng	44.32	0.001 (20.85 \pm 5.73)	49.61	0.003 (21.33 \pm 5.86)
elder_hfyn	34.10	0.019 (20.10 \pm 5.35)	35.30	0.021 (20.44 \pm 7.49)
elder_lfgng	42.83	0.005 (21.73 \pm 5.98)	31.70	0.043 (20.96 \pm 5.94)
elder_lfyn	17.25	0.634 (19.76 \pm 5.18)	29.00	0.028 (19.15 \pm 5.63)
elder_pseudo	62.79	0.000 (26.12 \pm 6.81)	53.10	0.009 (25.69 \pm 10.41)
elder_yn	32.87	0.258 (28.77 \pm 7.42)	62.72	0.012 (29.00 \pm 14.16)
young_gng	35.92	0.136 (28.60 \pm 7.39)	69.38	0.003 (28.66 \pm 8.36)
young_hfgng	21.33	0.305 (19.70 \pm 4.99)	34.11	0.016 (20.13 \pm 6.16)
young_hfyn	29.75	0.049 (19.59 \pm 5.04)	45.20	0.009 (19.83 \pm 7.03)
young_lf	22.06	0.318 (20.39 \pm 5.81)	37.78	0.015 (20.67 \pm 7.82)
young_lfgng	22.06	0.299 (20.08 \pm 5.25)	37.78	0.012 (20.27 \pm 6.52)
young_lfyn	23.62	0.182 (19.66 \pm 5.03)	17.66	0.542 (19.56 \pm 7.43)
young_pseudo	20.35	0.867 (27.86 \pm 7.20)	28.48	0.386 (28.44 \pm 10.87)
young_yn	38.34	0.097 (28.07 \pm 7.03)	54.20	0.003 (28.13 \pm 8.66)

KS is the Kolmogorov-Smirnov statistic calculated between the data and its fitted ex-Gaussian. In columns p_1 and p_2 , one finds the probabilities that a randomly generated dataset has a bigger KS statistic than the empirical data. In parenthesis, the average KS statistic and standard deviation for the generated random samples.

default in the histogram function of the ExGUtils package, namely two times the square root of the number of measurements in the data.

In order to study the effect of the number of intervals in the values for the parameters and of p_2 , we performed the procedure of fitting the data through minSQR after constructing the histogram with different number of intervals. In Figure 4 we show the evolution of the p_2 probability, along with the values for μ , σ , and τ obtained by minSQR for the histograms constructed with a different number of intervals for the young_hfgng experiment.

From the figure one sees that while the number of intervals is unreasonably small compared to the size of the empirical dataset, the values for the fitted ex-Gaussian parameters fluctuate, while the p probability is very small, but, once the number of intervals reaches a reasonable value, around 40, the values for the parameters stabilize and the value of p also gets more stable. So the question remains, why the values for the probability obtained with maxLKHD method is so small in the case of this experiment? The fact is that the likelihood of the dataset is very sensible to outliers. For the value of the probability [$f(x)$ in Equation 5] gets very small for the extreme values. Therefore, in these cases, it might be reasonable to make some criterious data trimming. So we proceed as follows: Given a dataset, we first perform a pre-fitting by maxLKHD. Using the parameters obtained in this fit, we estimate the points where the distribution has a left and right tails of 0.1% and remove measurements beyond these points. With the trimmed dataset, removed of outliers, we perform fits again and evaluate the p_1 and p_2 probabilities. In Table 4, we show the results for this new round of fitting and probability evaluations. In more than half of the experiments where one could see a big

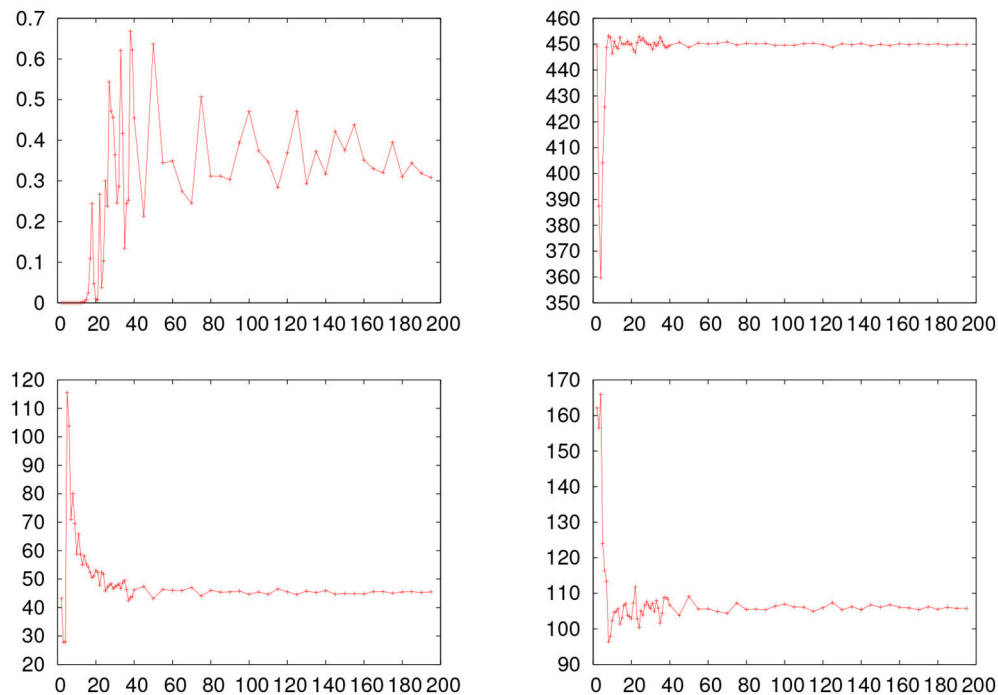


FIGURE 4 | Fitting results obtained with the minSQR method varying the number of intervals in the histogram for the young_hfgng experiment. The horizontal line shows the value obtained with the maxLKHD method. **(Upper left)** Evolution of the p probability. **(Upper right)** Evolution of μ . **(Bottom left)** Evolution of σ . **(Bottom right)** Evolution of τ .

TABLE 4 | The $p1$ and $p2$ probabilities for the fits.

Experiment	N	N' (%)	minSQR		maxLKHD	
			KS	$p2$ ($\bar{KS} \pm sd$)	KS	$p1$ ($\bar{KS} \pm sd$)
elder_gng	2,348	2 (0.09)	66.58	0.000 (28.92 \pm 7.32)	50.24	0.040 (30.98 \pm 17.55)
elder_hfgng	1,174	8 (0.68)	34.20	0.040 (20.67 \pm 5.70)	32.64	0.010 (20.66 \pm 5.83)
elder_hfyn	1,175	2 (0.17)	32.09	0.040 (20.01 \pm 4.86)	24.76	0.090 (19.22 \pm 6.69)
elder_lfgng	1,174	1 (0.09)	43.49	0.000 (21.47 \pm 5.83)	33.22	0.030 (20.57 \pm 6.90)
elder_lfyn	1,139	4 (0.35)	19.97	0.550 (20.55 \pm 6.37)	19.71	0.620 (19.97 \pm 6.11)
elder_pseudo	1,910	5 (0.26)	57.26	0.000 (26.91 \pm 6.64)	57.11	0.010 (26.61 \pm 10.06)
elder_yn	2,314	5 (0.22)	36.83	0.240 (28.57 \pm 7.46)	29.72	0.230 (30.54 \pm 14.33)
young_gng	2,396	10 (0.42)	38.93	0.250 (27.82 \pm 6.32)	43.11	0.020 (30.19 \pm 17.07)
young_hfgng	1,200	8 (0.67)	23.28	0.780 (19.25 \pm 4.39)	17.82	0.430 (18.07 \pm 4.13)
young_hfyn	1,180	9 (0.76)	27.97	0.050 (19.68 \pm 4.91)	28.93	0.010 (20.74 \pm 7.71)
young_lf	1,196	5 (0.42)	25.11	0.310 (20.09 \pm 5.21)	25.32	0.020 (19.69 \pm 4.29)
young_lfgng	1,196	5 (0.42)	25.11	0.280 (20.51 \pm 5.08)	25.32	0.080 (20.55 \pm 5.05)
young_lfyn	1,132	3 (0.27)	25.20	0.230 (19.42 \pm 5.40)	16.60	0.780 (20.72 \pm 8.53)
young_pseudo	2,326	10 (0.43)	23.33	0.940 (27.59 \pm 7.05)	25.85	0.870 (28.45 \pm 12.48)
young_yn	2,312	12 (0.52)	46.10	0.130 (27.80 \pm 7.87)	28.58	0.210 (31.21 \pm 19.74)

KS is the Kolmogorov-Smirnov statistic calculated between the data and its fitted ex-Gaussian. N is the number of data points in each empirical dataset, N' in the number of points removed by the trimming and in brackets next to it its proportion in relation to the total data. In columns $p1$ and $p2$, one finds the probabilities that a randomly generated dataset has a bigger KS statistic than the empirical data. In parenthesis, the average KS statistic and standard deviation for the generated random samples.

discrepancy between $p1$ and $p2$ in Table 3, the trimmed data do show better results. For some datasets, the trimming had no impact on the discrepancy. In any case, one might wonder

about the impact of the trimming in the obtained parameters. Therefore, in Table 5, we show the results obtained with different trimming criteria.

TABLE 5 | Results for different trimming on the data.

Experiment	%	minSQR				maxLKHD			
		μ	σ	τ	p2	μ	σ	τ	p1
elder_gng	0.1	513.52	73.00	329.54	0.001	518.71	75.02	313.04	0.026
elder_gng	0.5	516.62	76.61	319.50	0.002	521.83	70.31	299.00	0.011
elder_gng	1.0	516.04	76.80	317.93	0.000	523.84	66.32	291.17	0.014
elder_hfgng	0.1	509.10	84.96	285.05	0.043	504.96	65.26	297.06	0.012
elder_hfgng	0.5	509.39	89.51	277.28	0.020	511.19	65.09	277.33	0.020
elder_hfgng	1.0	508.40	83.49	279.35	0.016	512.79	59.89	272.67	0.005
elder_hfyn	0.1	564.82	82.19	246.63	0.052	558.93	71.17	266.45	0.148
elder_hfyn	0.5	565.70	83.88	242.73	0.036	559.98	68.60	261.73	0.143
elder_hfyn	1.0	566.73	87.05	235.38	0.006	561.88	65.77	255.95	0.094
elder_lfgng	0.1	521.64	62.39	368.34	0.006	530.64	68.95	333.51	0.041
elder_lfgng	0.5	523.29	67.46	359.50	0.006	530.25	60.81	329.35	0.011
elder_lfgng	1.0	523.37	67.70	356.20	0.002	533.09	59.45	318.33	0.008
elder_lfyn	0.1	583.03	84.58	301.15	0.562	581.72	76.56	305.56	0.577
elder_lfyn	0.5	584.32	86.07	296.15	0.524	584.60	78.19	296.28	0.329
elder_lfyn	1.0	586.72	85.93	287.48	0.470	589.73	77.85	278.47	0.027
elder_pseudo	0.1	735.04	133.55	498.90	0.001	755.81	134.79	436.48	0.012
elder_pseudo	0.5	733.65	135.57	499.00	0.001	754.68	132.25	438.02	0.017
elder_pseudo	1.0	732.54	135.87	498.14	0.000	752.31	124.65	442.19	0.014
elder_yn	0.1	572.16	81.99	275.26	0.251	567.87	73.30	288.63	0.280
elder_yn	0.5	573.64	84.34	270.01	0.373	570.72	72.30	278.01	0.378
elder_yn	1.0	573.82	84.87	266.60	0.246	573.48	72.59	268.80	0.159
young_gng	0.1	456.35	48.59	133.40	0.292	453.37	47.60	140.66	0.013
young_gng	0.5	456.95	47.02	132.15	0.177	456.29	43.54	134.00	0.167
young_gng	1.0	457.70	46.28	130.55	0.096	457.63	40.37	131.00	0.013
young_hfgng	0.1	449.79	45.31	105.15	0.707	448.42	44.89	109.02	0.565
young_hfgng	0.5	450.77	44.72	103.91	0.500	449.62	40.74	107.45	0.704
young_hfgng	1.0	451.94	44.75	101.09	0.208	451.50	37.51	103.23	0.226
young_hfyn	0.1	493.66	50.92	116.16	0.032	487.17	51.93	126.49	0.009
young_hfyn	0.5	494.62	50.74	114.27	0.054	488.97	51.00	122.73	0.025
young_hfyn	1.0	495.77	50.10	111.55	0.083	493.08	49.40	114.69	0.170
young_lf	0.1	473.36	54.44	151.84	0.287	471.09	54.85	157.76	0.037
young_lf	0.5	474.18	55.22	148.96	0.207	474.72	51.93	148.93	0.117
young_lf	1.0	475.03	54.10	147.35	0.067	475.22	45.69	148.46	0.019
young_lfgng	0.1	473.36	54.44	151.84	0.290	471.09	54.85	157.76	0.054
young_lfgng	0.5	474.18	55.22	148.96	0.201	474.72	51.93	148.93	0.119
young_lfgng	1.0	475.03	54.10	147.35	0.068	475.22	45.69	148.46	0.021
young_lfyn	0.1	508.16	61.53	151.83	0.228	503.17	57.27	162.27	0.776
young_lfyn	0.5	508.79	62.11	148.82	0.306	506.82	56.33	153.58	0.713
young_lfyn	1.0	508.92	59.52	148.67	0.278	508.72	51.89	151.43	0.545
young_pseudo	0.1	555.42	63.03	161.81	0.951	555.36	60.57	162.27	0.858
young_pseudo	0.5	556.11	63.54	159.16	0.364	556.92	57.17	158.77	0.194
young_pseudo	1.0	557.18	62.50	157.25	0.096	559.57	54.06	153.59	0.021
young_yn	0.1	497.56	54.59	136.65	0.141	492.23	53.69	146.70	0.144
young_yn	0.5	498.05	54.18	135.23	0.374	495.25	52.33	139.85	0.605
young_yn	1.0	498.17	53.86	134.10	0.556	496.97	50.70	136.71	0.494

The column % indicates the amount of tail trimmed to the left and right of the data.

Now, having the full picture, one can realize that some values of p are indeed small, indicating that either the ex-Gaussian distribution is not that good a model in order to fit the empirical results, or there is still some systematic error in the analysis of the experiments. Most of these empirical datasets where one sees very low values of p are with elderly people. These have the τ parameter much bigger than the σ which indicates a very asymmetric distribution with a long right tail. Indeed, a careful analysis of the histograms will show that the tail in these empirical distributions seems to be cut short at the extreme of the plots, so that the limit time in the experiment should be bigger than 2,500 ms in order to get the full distribution. One might argue that the trimming actually was removing data, but most of the removed points in the trimming of elderly data, was from the left tail and not from the right. This issue will result in the wrong evaluation of the KS statistics, since it assumes that one is dealing with the full distribution. This kind of analysis might guide better experimental designs.

6. OVERVIEW

The ex-Gaussian fit has turned into one of the preferable options when dealing with positive skewed distributions. This technique provides a good fit to multiple empirical data, such as reaction times (a popular variable in Psychology due to its sensibility to underlying cognitive processes). Thus, in this work we present a python package for statistical analysis of data involving this distribution.

This tool allows one to easily work with alternative strategies (fitting procedures) to some traditional analysis like trimming. This is an advantage given that an ex-Gaussian fit includes all data while trimming may result in biased statistics because of the cuts.

REFERENCES

- Balota, D. A., Cortese, M. J., Sergent-Marshall, S. D., Spieler, D. H., and Yap, M. (2004). Visual word recognition of single-syllable words. *J. Exp. Psychol. Gen.* 133:283. doi: 10.1037/0096-3445.133.2.283
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law distributions in empirical data. *SIAM Rev.* 51, 661–703. doi: 10.1137/070710111
- Cousineau, D., Brown, S., and Heathcote, A. (2004). Fitting distributions using maximum likelihood: methods and packages. *Behav. Res. Methods Instrum. Comput.* 36, 742–756. doi: 10.3758/BF03206555
- Cousineau, D., and Larochelle, S. (1997). Pastis: a program for curve and distribution analyses. *Behav. Res. Methods Instrum. Comput.* 29, 542–548. doi: 10.3758/BF03210606
- Epstein, J. N., Langberg, J. M., Rosen, P. J., Graham, A., Narad, M. E., Antonini, T. N., et al. (2011). Evidence for higher reaction time variability for children with adhd on a range of cognitive tasks including reward and event rate manipulations. *Neuropsychology* 25:427. doi: 10.1037/a0022155
- Gooch, D., Snowling, M. J., and Hulme, C. (2012). Reaction time variability in children with adhd symptoms and/or dyslexia. *Dev. Neuropsychol.* 37, 453–472. doi: 10.1080/87565641.2011.650809
- Grushka, E. (1972). Characterization of exponentially modified gaussian peaks in chromatography. *Anal. Chem.* 44, 1733–1738. doi: 10.1021/ac60319a011
- Heathcote, A. (2004). Fitting wald and ex-wald distributions to response time data: an example using functions for the s-plus package. *Behav. Res. Methods Instrum. Comput.* 36, 678–694. doi: 10.3758/BF03206550

Moreover, this tool is programmed as Python modules, which allow the researcher to integrate them with any other Python resource available. They are also open-source and free software which allows one to develop new tools using these as building blocks.

7. AVAILABILITY

ExGUtils may be downloaded from the Python Package index (<https://pypi.python.org/pypi/ExGUtils/3.0>) for free along with the source files and the manual with extended explanations on the functions and examples.

AUTHOR CONTRIBUTIONS

CM-T participated in the conception, design, and interpretation of data, and in drafting the manuscript. DG participated in the design, and analysis and interpretation of data, and in drafting the manuscript. EN-P and PF participated in revising the manuscript.

FUNDING

This work has been financed under the Generalitat Valenciana research project GV/2016/188 (Prof. Carmen Moret-Tatay) and the Universidad Católica de Valencia, San Vicente Mártir.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpsyg.2018.00612/full#supplementary-material>

- Heathcote, A., Popiel, S. J., and Mewhort, D. (1991). Analysis of response time distributions: an example using the stroop task. *Psychol. Bull.* 109:340.
- Hervey, A. S., Epstein, J. N., Curry, J. F., Tonev, S., Eugene Arnold, L., Keith Conners, C., et al. (2006). Reaction time distribution analysis of neuropsychological performance in an adhd sample. *Child Neuropsychol.* 12, 125–140. doi: 10.1080/09297040500499081
- Lacouture, Y., and Cousineau, D. (2008). How to use matlab to fit the ex-gaussian and other probability functions to a distribution of response times. *Tutor. Quant. Methods Psychol.* 4, 35–45. doi: 10.20982/tqmp.04.1.p035
- Leth-Steensen, C., King Elbaz, Z., and Douglas, V. I. (2000). Mean response times, variability, and skew in the responding of adhd children: a response time distributional approach. *Acta Psychol.* 104, 167–190. doi: 10.1016/S0001-6918(00)00019-6
- Luce, R. D. (1986). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford psychology series. New York, NY; Oxford: Oxford University Press; Clarendon Press.
- McVay, J. C., and Kane, M. J. (2012). Drifting from slow to “D’oh!”: working memory capacity and mind wandering predict extreme reaction times and executive control errors. *J. Exp. Psychol. Learn. Mem. Cogn.* 38:525. doi: 10.1037/a0025896
- Navarro-Pardo, E., Navarro-Prados, A. B., Gamermann, D., and Moret-Tatay, C. (2013). Differences between young and old university students on a lexical decision task: evidence through an ex-Gaussian approach. *J. Gen. Psychol.* 140, 251–268. doi: 10.1080/00221309.2013.817964

- Ratcliff, R., Love, J., Thompson, C. A., and Opfer, J. E. (2012). Children are not like older adults: a diffusion model analysis of developmental changes in speeded responses. *Child Dev.* 83, 367–381. doi: 10.1111/j.1467-8624.2011.01683.x
- Ratcliff, R., and McKoon, G. (2008). The diffusion decision model: theory and data for two-choice decision tasks. *Neural Comput.* 20, 873–922. doi: 10.1162/neco.2008.12-06-420
- Sternberg, S. (1966). High-speed scanning in human memory. *Science* 153, 652–654. doi: 10.1126/science.153.3736.652
- Van Zandt, T. (2000). How to fit a response time distribution. *Psychon. Bull. Rev.* 7, 424–465. doi: 10.3758/BF03214357
- West, R. (1999). Age differences in lapses of attention in the stroop task. *J. Gerontol. Ser. B Psychol. Sci. Soc. Sci.* 54, P34–P43. doi: 10.1093/geronb/54B.1.P34
- West, R., and Alain, C. (2000). Age-related decline in inhibitory control contributes to the increased stroop effect observed in older adults. *Psychophysiology* 37, 179–189. doi: 10.1111/1469-8986.3720179
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychol.* 41, 67–85. doi: 10.1016/0001-6918(77)90012-9

Conflict of Interest Statement: 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.

Copyright © 2018 Moret-Tatay, Gamermann, Navarro-Pardo and Fernández de Córdoba Castellá. 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 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.



Corrigendum: ExGUtills: A Python Package for Statistical Analysis With the ex-Gaussian Probability Density

Carmen Moret-Tatay^{1*}, Daniel Gamermann², Esperanza Navarro-Pardo³ and Pedro Fernández de Córdoba Castellá⁴

OPEN ACCESS

Edited and reviewed by:
Frontiers in Psychology,
Frontiers Media SA, Switzerland

***Correspondence:**
Carmen Moret-Tatay
mariacarmen.moret@ucv.es

Specialty section:
This article was submitted to
Cognition,
a section of the journal
Frontiers in Psychology

Received: 08 June 2018
Accepted: 11 June 2018
Published: 26 June 2018

Citation:
Moret-Tatay C, Gamermann D,
Navarro-Pardo E and Fernández de
Córdoba Castellá P (2018)
Corrigendum: ExGUtills: A Python
Package for Statistical Analysis With
the ex-Gaussian Probability Density.
Front. Psychol. 9:1108.
doi: 10.3389/fpsyg.2018.01108

¹ Department of Neuropsychology, Methodology, Basic and Social Psychology, Faculty of Psychology, Universidad Católica de Valencia San Vicente Mártir, Valencia, Spain, ² Instituto de Física, Universidade Federal do Rio Grande do Sul (UFRGS), Porto Alegre, Brazil, ³ Department of Developmental and Educational Psychology, Faculty of Psychology, Universitat de Valencia, Valencia, Spain, ⁴ Grupo de Modelización Interdisciplinar, Instituto Universitario de Matemática Pura y Aplicada, InterTech, Universitat Politècnica de València, Valencia, Spain

Keywords: response times, response components, python, ex-Gaussian fit, significance testing

A corrigendum on

ExGUtills: A Python Package for Statistical Analysis With the ex-Gaussian Probability Density
by Moret-Tatay, C., Gamermann, D., Navarro-Pardo, E., and Fernández de Córdoba Castellá, P.
(2018). *Front. Psychol.* 9:612. doi: 10.3389/fpsyg.2018.00612

We hereby inform that there was an error in the email of the corresponding author. This is how the author should be contacted: mariacarmen.moret@ucv.es.

The original article has been updated.

Conflict of Interest Statement: 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.

Copyright © 2018 Moret-Tatay, Gamermann, Navarro-Pardo, Fernández de Córdoba Castellá. 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 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.



Transcriptome and Proteome Alternation With Resistance to *Bacillus thuringiensis* Cry1Ah Toxin in *Ostrinia furnacalis*

Muhammad Zeeshan Shabbir[†], Tiantao Zhang[†], Zhenying Wang and Kanglai He^{*}

State Key Laboratory for Biology of Plant Diseases and Insect Pests, Institute of Plant Protection, Chinese Academy of Agricultural Sciences, Beijing, China

OPEN ACCESS

Edited by:

Roland Potthast,
University of Reading,
United Kingdom

Reviewed by:

Oksana Sorokina,
University of Edinburgh,
United Kingdom
Priyanka Baloni,
Institute for Systems Biology (ISB),
United States

*Correspondence:

Kanglai He
hekanglai@caas.cn

[†]These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Systems Biology,
a section of the journal
Frontiers in Physiology

Received: 14 September 2018

Accepted: 11 January 2019

Published: 01 February 2019

Citation:

Shabbir MZ, Zhang T, Wang Z
and He K (2019) Transcriptome
and Proteome Alternation With
Resistance to *Bacillus thuringiensis*
Cry1Ah Toxin in *Ostrinia furnacalis*.
Front. Physiol. 10:27.
doi: 10.3389/fphys.2019.00027

Background: Asian corn borer (ACB), *Ostrinia furnacalis* can develop resistance to transgenic *Bacillus thuringiensis* (Bt) maize expressing Cry1Ah-toxin. However, the mechanisms that regulate the resistance of ACB to Cry1Ah-toxin are unknown.

Objective: In order to understand the molecular basis of the Cry1Ah-toxin resistance in ACB, “omics” analyses were performed to examine the difference between Cry1Ah-resistant (ACB-AhR) and susceptible (ACB-BtS) strains of ACB at both transcriptional and translational levels.

Results: A total of 7,007 differentially expressed genes (DEGs) and 182 differentially expressed proteins (DEPs) were identified between ACB-AhR and ACB-BtS and 90 genes had simultaneous transcription and translation profiles. Down-regulated genes associated with Cry1Ah resistance included aminopeptidase N, ABCC3, DIMBOA-induced cytochrome P450, alkaline phosphatase, glutathione S-transferase, cadherin-like protein, and V-ATPase. Whereas, anti-stress genes, such as heat shock protein 70 and carboxylesterase were up-regulated in ACB-AhR, displaying that a higher proportion of genes/proteins related to resistance was down-regulated compared to up-regulated. The Kyoto encyclopedia of genes and genomes (KEGG) analysis mapped 578 and 29 DEGs and DEPs, to 27 and 10 pathways, respectively ($P < 0.05$). Furthermore, real-time quantitative (qRT-PCR) results based on relative expression levels of randomly selected genes confirmed the “omics” response.

Conclusion: Despite the previous studies, this is the first combination of a study using RNA-Seq and iTRAQ approaches on Cry1Ah-toxin binding, which led to the identification of longer length of unigenes in ACB. The DEGs and DEPs results are valuable for further clarifying Cry1Ah-mediated resistance.

Keywords: *Ostrinia furnacalis*, *Bacillus thuringiensis*, Cry1Ah toxin, qRT-PCR, RNA-Seq, iTRAQ

INTRODUCTION

Maize (*Zea mays* L.) is the main crop in terms of production and planting area (Wang et al., 2014). Asian corn borer (ACB), *Ostrinia furnacalis* (Guenée), is an economically important pest of maize causing 20–80% yield losses (Nicolas et al., 2013) by attacking fresh whorl leaves, silks, ears, and cobs, finally leading to devastation by boring into the stalk, ear shanks and cobs of corn

(He et al., 2003). The potential of ACB having adaptation to many host crops and higher fecundity are the key factors in developing Bt resistance (Zhang et al., 2014). Genetically modified crops produced by Bt are effective in controlling this endemic pest of maize, likewise, the transgenic Bt maize, Cry1Ac, Cry1Ab, Cry1Ie, and Cry1Ah which express transgenic insecticidal proteins are assumed to show effectiveness against infestation of ACB (Zhang et al., 2013; Shabbir et al., 2018). Although Bt transgenic crops are likely to hold great promise to improve insect pest management, the efficacy of Bt maize can be reduced by the evolution of target insect resistance. The increased occurrence of functional resistance in the pest populations is causing hazardous loss to the continuing success of Cry proteins (Tabashnik et al., 2003). Previously, evolution of potential resistance to various Bt toxins Cry1Ac, Cry1Ie, and Cry1F has been observed in ACB in laboratory selection (He et al., 2003; Wang et al., 2016), and now one ACB-AhR strain had developed resistance to Cry1Ah, and readily consumed Cry1Ah-Bt maize (Shabbir et al., 2018).

However, complete recognition of the mechanism of Bt resistance is essential to delay the resistance evolution in target insect pest. Currently, two different hypotheses for modes of actions are directed for Cry toxin: the pore formation model and signal transduction model (Soberón et al., 2009). The pore formation model has been reported to propose that reduction of Bt toxins in toxin binding sites in brush border membrane vesicles (BBMVs) of insect midgut is the major factor of the evolution of resistance in target insect pests (Daniel et al., 2002). After the crystalline inclusions, toxins are ingested and solubilized in the gut to the protoxin, which is cleaved by midgut proteases and binds to activated toxins (Bravo and Soberón, 2008; Soberón et al., 2009). The interaction of toxins with cadherin enables additional proteolytic cleavages that prompt the toxin oligomerization. Subsequently, these oligomers bind to secondary receptors, aminopeptidase N (APN) and alkaline phosphatase (ALP), as they have a larger affinity to bind these proteins as compared to the monomeric toxin. After binding, these oligomers insert into the membrane and create pores which make it more permeable. Finally, these pores cause osmotic shock in the membrane, ultimately leading to the death of cells (Soberón et al., 2009). According to the signal transduction model, the binding of Cry1A to cadherin is supposed to activate a cascade pathway involved in the stimulation of a G protein and adenylate cyclase to increase cAMP, causing activation of protein kinase A, and finally death of the cell (Zhang et al., 2006). Previously, several studies have reported binding receptors, including cadherin protein (Xu et al., 2005), APN (Tiewisiri and Wang, 2011), ALP (Jurat-Fuentes and Adang, 2007), membrane glycolipids (Griffitts and Aroian, 2005), and ABCC2 of ABC transporters (Gahan et al., 2010; Baxter et al., 2011). The differences in the sequences of amino acids and expression of mRNA of four APN genes have been observed between ACB-AhR and ACB-BtS strains (Xu et al., 2014). In addition, V-type ATPase and HSP 70 kDa proteins had been documented as Bt binding proteins in ACB using a proteomic approach (Xu et al., 2013). However, the studies describing the Bt resistance mechanism are still limited in ACB.

Gene expression analysis is extensively used for studying regulatory mechanisms that control cellular processes in plants, animal, and microbes. Recent advancement in high-throughput RNA sequencing (RNA-seq) technology and isobaric tags for relative and absolute quantification (iTRAQ) gene expression based on next generation sequencing technology significantly has upgraded transcriptome analysis (Wang et al., 2010; Chen et al., 2011). In the present study, we compared midgut tissues of ACB-AhR and ACB-BtS strains at transcriptome (RNA-seq) and proteome (iTRAQ) level to determine the molecular mechanism of Bt Cry1Ah resistance in ACB. The differentially expressed genes (DEGs) and differentially expressed proteins (DEPs) were further validated by quantitative real-time qRT-PCR analysis. These approaches are valuable for the understanding of systemic differences between susceptible and Bt resistant genotypes, and to identify the genes/proteins that might be involved in conferring resistance to Cry1Ah-toxin.

MATERIALS AND METHODS

Insects

The susceptible strain (ACB-BtS) and the Cry1Ah resistant strain (ACB-AhR), as reported previously (Shabbir et al., 2018), were used in the study. In our previous study, ACB-AhR had developed 200-fold resistance to Cry1Ah after 48 generations of selection (Shabbir et al., 2018). However, in the present study, the ACB-AhR was selected to detect the Cry1Ah resistance-relative genes in ACB. Four to five individual larvae from fifth instar larvae were collected as one biological replicate for both ACB-BtS and ACB-AhR. Three biological replicates for each sample were collected and processed independently. Three replicates were used in gene expression profile analysis, and Illumina sequencing, as well as three biological replicates which were used for the qRT-PCR analysis. All samples were stored at -80°C until assayed.

Library Preparation for Transcriptome Sequencing

A total amount of 1.5 μg RNA from the fifth instars larvae was used as input material for RNA sample preparation for each of the ACB-AhR and ACB-BtS strains. Sequences libraries were generated using NEBNext[®] Ultra[™] RNA Library Prep Kit for Illumina (NEB, United States) according to manufacturer's instructions and index codes were added to attribute sequences to each sample. The mRNA was purified from total RNA using poly-T oligo-attached magnetic beads and broken into short fragments using divalent cations under elevated temperature in NEBNext First Strand Synthesis Reaction Buffer (5X). First-strand cDNA was synthesized using random hexamer primer and M-MuLV Reverse Transcriptase (RNase H⁻). Second-strand cDNA was subsequently performed using DNA polymerase I and RNase H. Remaining overhangs were converted into blunt ends via exonuclease/polymerase activities. After adenylation of 3' ends of DNA fragments, NEBNext Adaptor with a hairpin loop structure was ligated to prepare for hybridization.

In order to select cDNA fragments of preferentially 250–300 bp in length, the library fragments were purified with AMPure XP system (Beckman Coulter, Beverly, MA, United States). Then before PCR, 3 μ l USER Enzyme (NEB, United States) was used with size-selected, adapter-ligand cDNA at 37°C for 15 min followed by 5 min at 95°C. PCR was performed with Phusion High-Fidelity DNA polymerase, Universal PCR primers, and Index (X) Primer. Finally, PCR products were purified (AMPure XP system) and library quality was assessed on the Agilent Bioanalyzer 2100 system. The RNA-seq data has been submitted to SRA database and the accession ID is PRJNA508227.

Assembly and Functional Gene Annotation

The reads containing poly-N (<10%), and low quality reads ($q < 20$) were removed from raw data. Q20, Q30, GC-content and sequence duplication level of the clean data were also assessed based on high quality clean data. Subsequently, the clean reads were accomplished using Trinity software (Grabherr et al., 2013). Gene functional annotation sequences were searched using BLAST against NCBI NR database was searched using BLAST against NCBI NR database¹ with a cut-off E-value of 10^{-5} . Functional gene annotations were collected for transcript sequences ≥ 150 bp using Blast2GO (Conesa et al., 2005). DEGs were calculated in FPKM (fragments per kilobase pair of exon model per million fragments mapped) for comparing the expression of up- or down-regulated transcripts in two groups. BLASTx algorithm was used to assign gene ontology (GO) terms from the GO database² and the DEGs were assigned into different pathways by the Kyoto encyclopedia of genes and genomes (KEGG) pathways databases.

Screening of Differentially Expressed Genes Between ACB-AhR and ACB-BtS

The mapped reads of ACB-AhR and ACB-BtS groups were assembled using the DESeq (2010) R package (1.10.1). DESeq fetches statistical routines to regulate differential expression in digital gene expression data using a model based on the negative binomial distribution. The resulting *P*-values were adjusted using the *q*-value. Genes with an adjusted *P*-value <0.05 found by DESeq were assigned as differentially expressed. Then, the FPKM value between the biological replications was analyzed for each gene. The significance of digital gene expression profiles was analyzed as described previously (Audic and Claverie, 1997). The fold change of each gene was then calculated by the formula of \log_2 (ACB-AhR_FPKM/ACB-BtS_FPKM). False discovery rate (FDR) method was used to determine the threshold of *P*-value in differential gene expression tests. “FDR” ≤ 0.05 and the absolute value of \log_2 -ratio ≥ 1 was the threshold to evaluate the significance level of differentiated gene expression for comparing the gene expression between two strains of ACB.

¹<https://www.ncbi.nlm.nih.gov/refseq/about/nonredundantproteins/>

²<http://www.geneontology.org/>

Protein Quantification and Database Search Using iTRAQ Labeling

The midgut tissues of ACB-AhR and ACB-BtS samples were individually milled in liquid nitrogen then put into 1 ml of lysis buffer (50 mM Tris buffer, 8 M urea, 1% SDS, pH 8), and ultrasonic was used to extract the protein. The lysis solution was centrifuged at 4°C, $12,000 \times g$ for 15 min to collect the supernatant, then four volumes of precooled acetone (include 10 mM DTT) was added to a sample extract, and samples were placed at 20°C for 2 h. It was centrifuged again, and the pellet was collected to wash twice with cold acetone. Finally, the precipitation was dissolved by the dissolution buffer containing Tris-base (pH 8) 8M Urea solution. The protein was determined by using the Bradford method and analyzed on the SDS-PAGE gel. After 100 μ l protein from each sample was digested with trypsin gold (Promega, Madison, WI, United States) at 37°C for 16 h, and the resultant peptides were dried by vacuum centrifugation. The peptides were reconstituted in 20 μ l of 0.5 M TEAB (pH 8.5) and processed according to the manufacturer's protocol for 8-plex iTRAQ (AB Sciex, Foster City, CA, United States) (Noirel et al., 2011). Then, pooled mixtures of iTRAQ-labeled peptides were fractionated by XBridge BEHC18 column BEH C18 4.6 \times 250 mm, 5 μ m, (Waters, Milford, MA, United States) on a Rigol L3000 HPLC operating at 1 ml/min. Mobile phases A (2% acetonitrile, 20 mM NH₄FA, adjusted pH to 10.0 using NH₃·H₂O) and B (98% acetonitrile, 20 mM NH₄FA, adjusted pH to 10.0 using NH₃·H₂O) were used to develop a gradient elution. Collected fractions were pooled into 15 final fractions and analyzed by Q-Exactive HF-X mass spectrometer (Matrix Science Limited, Washington, DC, United States).

Peptides were identified separately by searching against a specified database Proteome Discoverer 2.2 (PD 2.2, Thermo). A peptide mass tolerance of 10 ppm and fragment mass tolerance of 0.02 Da were acceptable for product ion scans. When the Proteome Discoverer 2.2 software was used to search the database, 5,900 proteins were identified at FDR less than 1.0%. Proteins comprising of similar peptides and could not be distinguished based on MS/MS analysis were grouped separately as protein groups. To analyze the differential expression ratios, all identified peptides from a protein were used to find an average protein ratio relative to the control label (i.e., fold change). Mann–Whitney test was used to analyze the differential expression of proteins between ACB-AhR and ACB-BtS larvae midgut and the significant ratios, defined as $P < 0.05$ and $|\log_2 \text{FC}| > *(\text{ratio} > * \text{ or ratio} < *[\text{fold change, FC}])$, were used to screen the DEPs.

GO Classification of Differentially Expressed Genes and Proteins Pathway Enrichment Analysis

Functional annotation of the genes and proteins which were identified in ACB midgut sample was implemented using Goseq R packages based Wallenius non-central hyper-geometric distribution (Young et al., 2010), an integrated GO annotation and mining tool that assigns gene ontology through BLAST searches against nucleotide and protein databases. GO functional

significance enrichment analysis gives GO functional entries that are significantly enriched in DEGs compared to the genomic background. The analysis first maps DEGs to each term in the Gene Ontology database (see footnote 2), calculate the number of genes for each term, and then find differences in expression compared to the entire genomic background and then used a hypergeometric test to find significantly enriched GO terms for DEGs compared to the ACB transcriptome/proteome background. In order to better study the function of differential genes, we not only performed enrichment analysis (GO enrichment, KEGG enrichment) for all the differential genes in each combination but also separated differential genes in each combination according to up- or down-regulation. The differential expression of the genes was determined by performing independent alignments of short reads count obtained from analysis of gene expression levels. For samples with biological replicates, the analysis was performed using DESeq (Anders and Huber, 2010), and the screening threshold was $p_{adj} < 0.05$. The P -value was checked by using the following formula:

$$P=1 - \sum_{i=0}^{m-1} \frac{\binom{M}{i} \binom{N-M}{n-i}}{\binom{N}{n}}$$

(1) N is the number of genes with pathway annotation in all genes. (2) n is the number of DEGs in N . (3) M is the number of genes annotated as a particular pathway in all genes. (4) m is the number of DEGs annotated as a specific pathway. Pathway with $FDR \leq 0.05$ was defined as a pathway that was significantly enriched in DEGs or proteins. All identified transcripts and proteins were mapped to a pathway in the KEGG database. Significantly enriched metabolic pathways containing DEGs and DEPs were determined using the same formula as in GO analysis. Here N means the number of all the genes/proteins with KEGG annotation, n represents the number of DEGs or DEPs in N , M is the number of all genes or proteins annotated to specific pathways, and m is the number of DEGs or DEPs in M .

Relationship Between RNA-Seq and iTRAQ

To evaluate the expression level of genes and proteins in ACB-AhR and ACB-BtS, the relationship between transcriptomic and proteomics levels was evaluated. The mRNA information obtained from the transcriptome was integrated with the DEPs information identified by the proteome and was searched for

the expression patterns of corresponding genes ($P < 0.05$). The significance of the overlapping between the identified transcripts and proteins was determined using Pearson's chi-square test with Yates' continuity correction (Song et al., 2012).

RT-qPCR for Expression Analysis

The genes related to resistance selected from transcriptomic and proteomic analysis were verified using qRT-PCR. Total RNA was prepared from different tissues of ACB-AhR and ACB-BtS strains, with three technical replicates performed for each of three biological replicates. cDNAs were synthesized using the One-Step gDNA Removal and cDNA Synthesis SuperMix (TransGen Biotech Co., Ltd., Beijing, China) following the kit manual. β -actin was used as a reference gene (accession number-EU585777.1), and it was used to select the cDNA templates on the PCR equipment. Primers (Supplementary Table S9) were designed manually or using the Primer 5 tool³. Individual qRT-PCR reactions were repeated four times; water was used as the negative control. Before gene quantification, the amplification efficiency between the target gene and the reference gene were checked. qRT-PCR reactions were performed on the Applied Bio System 7500 Real-Time PCR System (Applied Biosystems, Foster City, CA, United States) using SYBR Green (TAKARA Bio Inc., Japan) The cycling program consists of initial incubation at 95°C for 10 min, followed by 40 cycles at 95°C for 15 s, 60°C for 45 s, and a final step at 95°C for 15 s and reactions were performed in a final volume of 25 μ l. The threshold cycle (C_T) was collected from each reaction, and the relative expression of normalized data was calculated by the comparative $2^{-\Delta\Delta C_T}$ method (Livak and Schmittgen, 2001; Zhang et al., 2017).

RESULTS

RNA-Seq and Sequence Assembly

The results of RNA sequencing from ACB-AhR and ACB-BtS were ranged from 41,703,706 to 62,099,678 (Table 1). The clean sequences per library were ranged from 40,607,798 to 59,909,406 reads. Moreover, GC contents were ranged from 48.01 to 51.02%. The number of the reads ranged from 40.09 to 44.83%, were mapped to the trinity spliced transcriptomes. A total of 73,229 unigenes assembled from cDNA libraries of both resistant and susceptible strains with an average length of 844 bp and N50 length of 1,018 bp (Table 2).

³<http://frodo.wi.mit.edu/primer5/>

TABLE 1 | Summary of reads in Cry1Ah-resistant strain (ACB-AhR) and susceptible strain (ACB-BtS) of *Ostrinia furnacalis* transcriptomes.

Samples	Raw reads	Clean reads	Total mapped	Q30 (%)	GC (%)
ACB-BtS_1	54,874,002	53,134,612	21,299,364 (40.09%)	90.90	50.55
ACB-BtS_2	41,970,856	40,607,798	16,906,344 (41.63%)	91.55	51.02
ACB-BtS_3	54,469,658	52,770,170	21,864,498 (41.43%)	91.50	50.40
ACB-AhR_1	41,703,706	40,498,994	17,492,734 (43.19%)	91.81	48.11
ACB-AhR_2	62,099,678	59,909,406	24,716,016 (41.26%)	91.19	50.71
ACB-AhR_3	47,695,548	46,435,464	20,817,278 (44.83%)	92.06	48.01

TABLE 2 | Summary of RNA-seq metrics from *Ostrinia furnacalis* midgut transcriptomes.

Metrics	Number
Number of genes	73,229
*N50	1,018
Maximum length	29,484
Minimum length	201
Average length	844
Total assembled bases	132,050,968

*N50 is the 50% length of all genes.

The total numbers of sequences detected by mass spectrometry of ACB proteome were 585,828, which represented 29,314 peptide spectra and 5,900 proteins were matched (Table 3). The total DEPs between ACB-AhR and ACB-BtS were 182.

Differentially Expressed Genes Between Cry1Ah-Resistant and Susceptible Strains of ACB

A total of 4,209 down-regulated and 2,798 up-regulated genes were differentially expressed ($P < 0.05$ and $|\log_2\text{-ratio}| \geq 1$) (Figure 1A) in both ACB-AhR and ACB-BtS strains. These comparison results revealed that most of the genes were significantly down-regulated compared to up-regulated including APN, ALP, and member of the ABC the transporter family (Supplementary Table S1). Furthermore, genes significantly down-regulated in the high severity in ACB-AhR strain with threshold group ($q\text{-value} < 1$ and $\log_2(\text{fold-change}) \leq -2$), several genes were annotated as previously known Bt resistance genes including members of the APN gene family, *apn3* paralogs and *apn8*, an ABC transporter in subgroup G, *abcg*, and serine protease genes. The up-regulated genes ($q\text{-value} < 1$ and $\log_2(\text{fold-change}) \geq 2$) were significantly smaller in number for ACB-AhR strain compared to down-regulated genes. The up-regulated genes in ACB-AhR strain included heat shock proteins and carboxylesterase genes (Supplementary Table S1).

Supplementary Table S2 shows the GO classification of genes that were differentially expressed between ACB-AhR and ACB-BtS midgut tissues (≥ 2 -fold change, $\text{FDR} \leq 0.001$). With Blast2Go, 7,007 DEGs were assigned to 51 GO classes (Figure 2A), which cover three domains: biological process, cellular components, and molecular functions. In terms of biological process mostly genes are assigned to oxidation–reduction process and DNA integration. In case of

oxidation reduction reaction, 277 DEGs were associated, where 162 were down-regulated and 115 DEGs were up-regulated in ACB-AhR (Supplementary Table S2). In case of cellular components terms, mostly fatty acid synthesis complex, and cytosolic part represented most of the genes. In the molecular function category, oxidoreductase activity, peptidase activity, and dehydrogenase activity were the most abundant (Supplementary Table S2).

In the KEGG database, 27 pathways were substantially enriched ($P \leq 0.05$), including “Valine, leucine and isoleucine degradation” and “Galactose metabolism” (Figure 3 and Supplementary Table S3). Specifically, 51 genes encoding enzymes involved in fatty acid elongation and metabolism of xenobiotics by cytochrome P450 pathways were highly enriched, including dehydrogenase, glutathione S-transferase (GSTs), and nicotinamide adenine dinucleotide phosphate (NADPHs) (Supplementary Table S4). The up-regulated genes included acetyltransferase, dehydrogenase, GST, and carbonyl reductase NADPH. Whereas, down-regulated genes enriched in galactose metabolism pathways, included steroid dehydrogenase and UDP-glucosyltransferase (Supplementary Table S4).

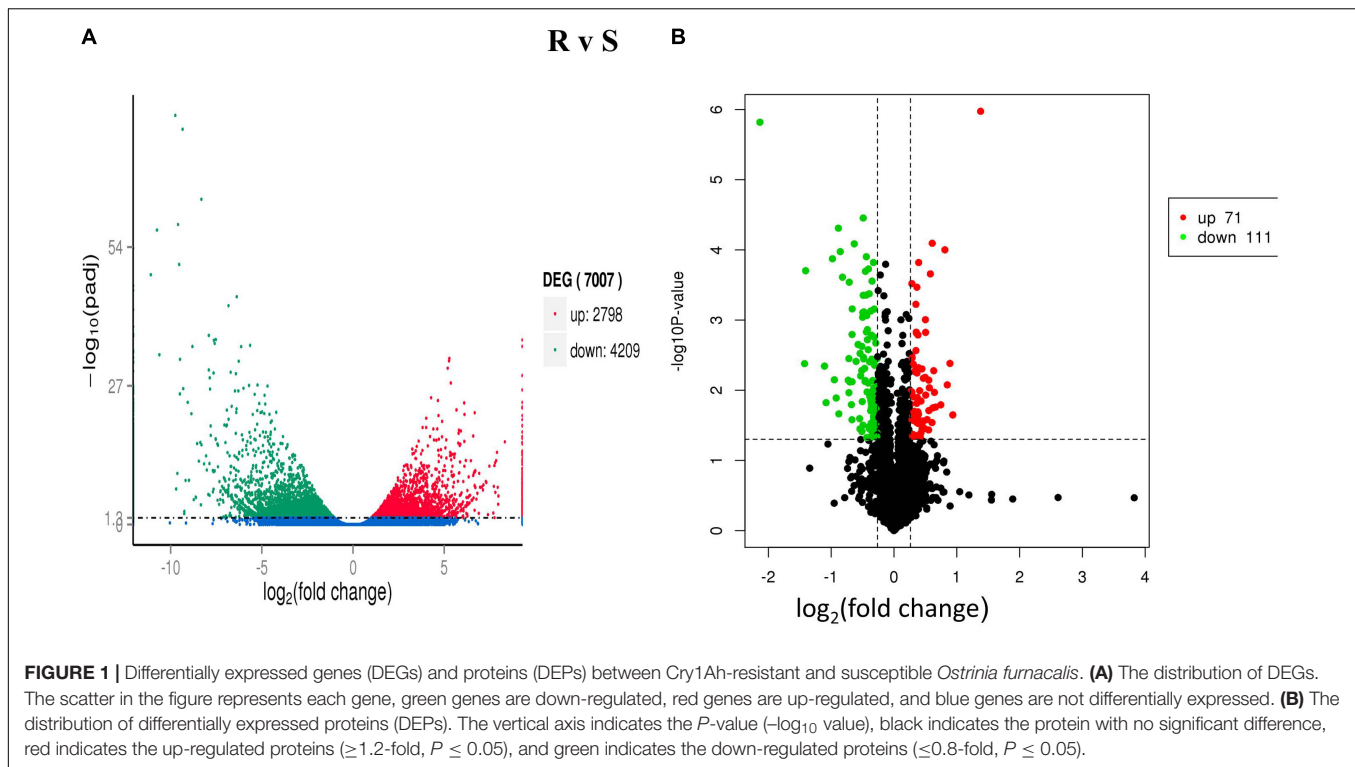
Cry1Ah-Induced Differentially Expressed Proteins Between ACB-AhR and ACB-BtS Strains

After Cry1Ah-treatment, 182 DEPs ($P \leq 0.05$) were identified between ACB-AhR and ACB-BtS strains of ACB (Figure 1B). Among them, 111 proteins were down-regulated (≤ 0.8 -fold, $P \leq 0.05$) and 71 proteins were up-regulated (≥ 1.2 -fold, $P \leq 0.05$) (Supplementary Table S5). Following in-gel digestion by trypsin, proteins were identified by liquid chromatography-electrospray ionization multistage mass spectrometry (LC-ESI-MS/MS). APN and ABCC proteins which are involved in Bt resistance were down-regulated by -0.45 - and -0.51 -fold, respectively, in ACB-AhR strain relative to the ACB-BtS strain. Others down-regulated proteins in resistance included trypsin (-1.41 -fold), which are considered the main proteases involved in Bt proteolysis activation and detoxification, GST (-0.67 -fold), and DIMBOA-induced cytochrome P450 (-0.46 -fold). The proteins that were up-regulated in Cry1Ah-resistant insects of ACB are fatty acid binding protein 1 (0.41-fold), aldose 1-epimerase (0.50-fold) involved in carbohydrate metabolic process, lipase (0.58-fold), plays an essential role in the digestion, transport and metabolism and UDP-glycosyltransferase (0.42-fold), involved in inactivation and excretion of endogenous and exogenous compounds. Additionally, proteins related to energy regulations, transportation of proteins, oxidation–reduction process, binding, and metabolism were also differentially expressed between ACB-AhR and ACB-BtS strains of ACB (Supplementary Table S5).

The relationship of correlation between the DEGs and DEPs showed that there were only 90 genes/proteins related to resistance that were either up-regulated or down-regulated identified in RNA-seq and iTRAQ techniques (Figure 4 and Supplementary Table S6). Among

TABLE 3 | Summary of iTRAQ metrics from the Cry1Ah-resistant strain (ACB-AhR) and susceptible strain (ACB-BtS) of *Ostrinia furnacalis* proteomes.

Metrics	Number
Total spectra	585,828
Unique spectra	29,314
Matched proteins	5,900
Differentially expressed proteins	182



them 63 genes/proteins were with same trend and 27 genes/proteins showed opposite trend either up-regulated or down-regulated in both analyses (Supplementary Table S6).

Gene Ontology and Pathway Enrichment

Among the 182 DEPs, 34 were subcategorized into 15 hierarchically structured GO classes, including 3 biological processes, 3 cellular components, and 9 molecular functions (Figure 2B). Specifically, “oxidation–reduction process” and single-organism metabolic process were highly represented in “Biological process”. While extracellular space was the most common categories in “Cellular components”. Likewise, iron ion binding, heme binding, and transition metal ion binding were the most top categories in “Molecular function”.

Fifty-nine DEPs were allocated to reference pathways in KEGG when exposed to Cry1Ah toxin. As a result, 10 pathways were enriched $P \leq 0.05$, Supplementary Table S7), including “glycine, serine, and threonine metabolism” and “galactose metabolism” which have the lowest *P*-value. The top 20 highly enriched pathways are shown in Figure 5.

Correlation of the enriched pathways for DEGs and DEPs showed that there were four mainly identical pathways related to metabolic process playing a role in resistance, including determining, galactose metabolism, glycerolipid metabolism, metabolism of xenobiotics by cytochrome P450 and glycine, serine, and threonine metabolism (Figures 3, 5). KEGG pathway analysis also revealed that the most enriched peptides, including phosphoglycerate dehydrogenase, *N*-acetylgalactosaminidase, NADPH, and UDP-glycosyltransferase were involved in glycine,

serine, and threonine metabolism, galactose metabolism, and metabolism of xenobiotics by cytochrome P450 (Supplementary Table S8).

Validation of Differentially Expressed Genes by qRT-PCR

According to fold-change calculations by qRT-PCR analyses, the results supported the differentially expressed on gene level. All the tested genes were in the same trend with the omics results except the chitin synthase which presented the down-regulation in Cry1Ah-resistant (ACB-AhR) strain compared to susceptible strain (ACB-BtS). However, the higher expression level was observed in ACB-AhR by qRT-PCR analysis (Figure 6). Most of the selected genes were down-regulated in ACB-AhR; only HSP 70 showed higher expression in ACB-AhR compared to ACB-BtS (Figure 6).

DISCUSSION

Insect resistance to *Bacillus thuringiensis* (Bt) is a significant threat to the enduring success of most extensively used genetically modified crops (Tabashnik et al., 2003, 2013). To counter the threat of resistance, it is important to understand the molecular mechanism of resistance of ACB to Bt toxins. In this study, ACB-AhR and ACB-BtS were sequenced for the transcriptomics and proteomics analyses, and we obtained a total of 73,229 genes with an average length of 844 bp from the transcriptome analysis. The average length of the genes was longer than those observed in ACB (Xu et al., 2015; Zhang et al., 2016; Cui et al., 2017),

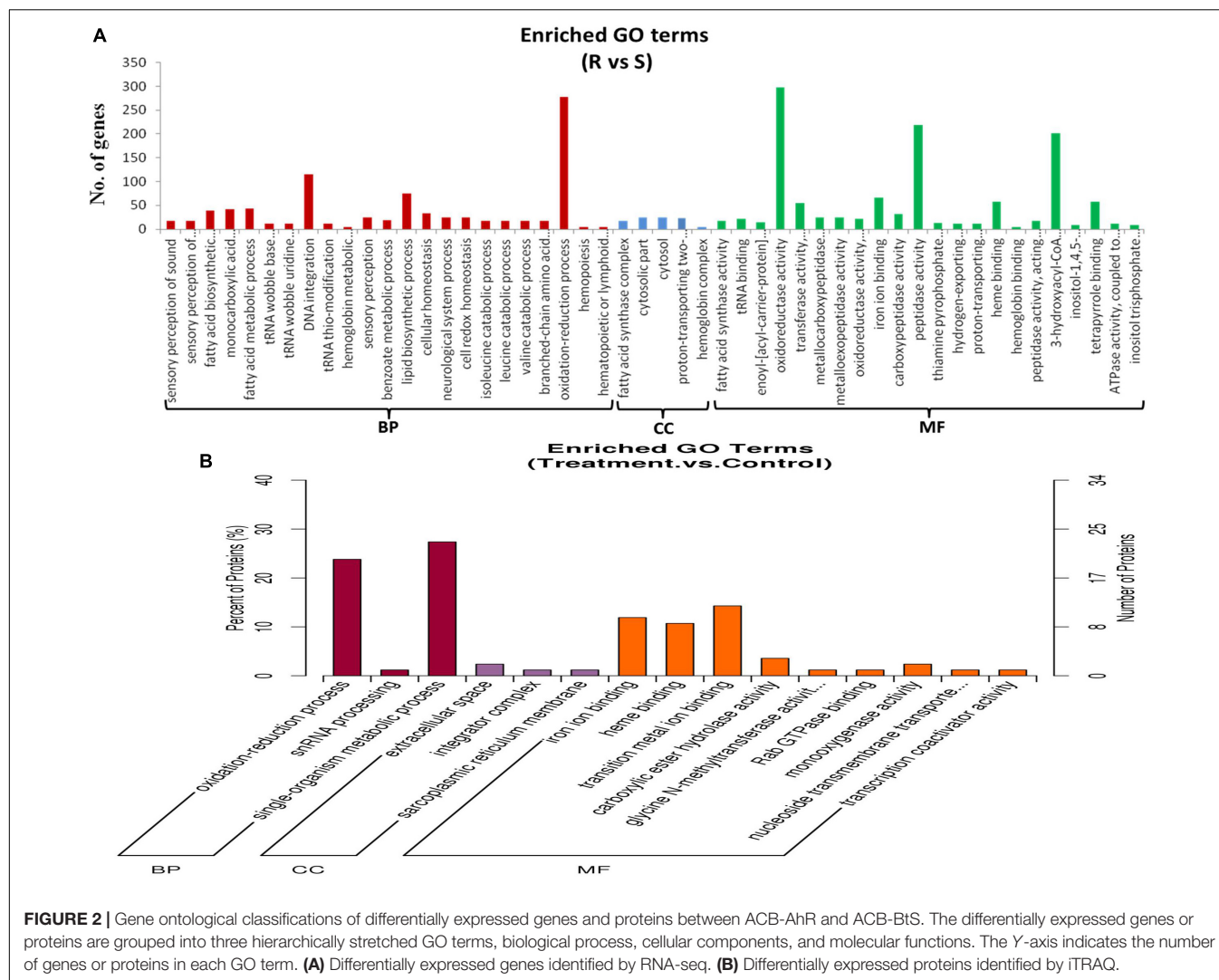


FIGURE 2 | Gene ontological classifications of differentially expressed genes and proteins between ACB-AhR and ACB-BtS. The differentially expressed genes or proteins are grouped into three hierarchically stretched GO terms, biological process, cellular components, and molecular functions. The Y-axis indicates the number of genes or proteins in each GO term. **(A)** Differentially expressed genes identified by RNA-seq. **(B)** Differentially expressed proteins identified by iTRAQ.

and *Plutella xylostella* (Lin et al., 2013). The genes length may be correlated to sequence techniques and the application of assembly tools. Mostly, assembled genes were not significantly matched with available databases due to their short sequences or because they characterized significantly novel genes. Comparatively, a low number of the genes had been annotated previously as compared to our findings. Therefore our Illumina sequencing and analysis described improvements over earlier studies (Xiang et al., 2010; Li et al., 2013).

Particularly, comparative analysis of midgut transcripts and proteins between ACB-AhR and ACB-BtS strains discovered a distinctive set of genes/proteins differentially expressed. Both transcriptomic and proteomic sequences showed more down-regulation of genes/proteins than up-regulations in ACB-AhR strain (Figure 1). Specifically, our results are in agreement with a previous transcriptomic analysis showing down-regulation of genes in resistant strains using a digital gene expression tag profiling (DGETP) approach (Paris et al., 2012; Tetreau et al., 2012). Similarly, significant alteration of the ACB transcriptome was observed in a Cry1Ab resistant strain (ACB-AhR) including,

3,157 genes being down-regulated and 636 were up-regulated after exposure to Cry1Ab toxin (Xu et al., 2015). Moreover, in a previous study, an analysis of DEGs directed that 1,026 DEGs were down-regulated and 189 were up-regulated, expressed between resistant and susceptible strains of *P. xylostella* (Lin et al., 2013). However, a study of transcriptome response to Cry1Ac toxin indicated more up-regulated genes as compared to down-regulated genes in a Cry1Ac-resistant strain of *P. xylostella* (Lei et al., 2014). The observation of different trends among experiments was possibly due to the technical differences and the variations in the materials examined, as a whole body of target insects at various developmental stages was used in susceptible and Cry1Ab-resistant strains of *P. xylostella* (Lin et al., 2013). However, midgut tissue was used from Cry1Ac-resistant and susceptible strains of *P. xylostella* (Lei et al., 2014). These results suggest that mechanisms of resistance to Cry toxins can be conferred by deficient activation of protoxins or reduced binding of toxins to the membrane (Griffitts and Aroian, 2005).

A correlation analysis of DEGs and DEPs from the larval midgut displayed the same trend of a subset of genes

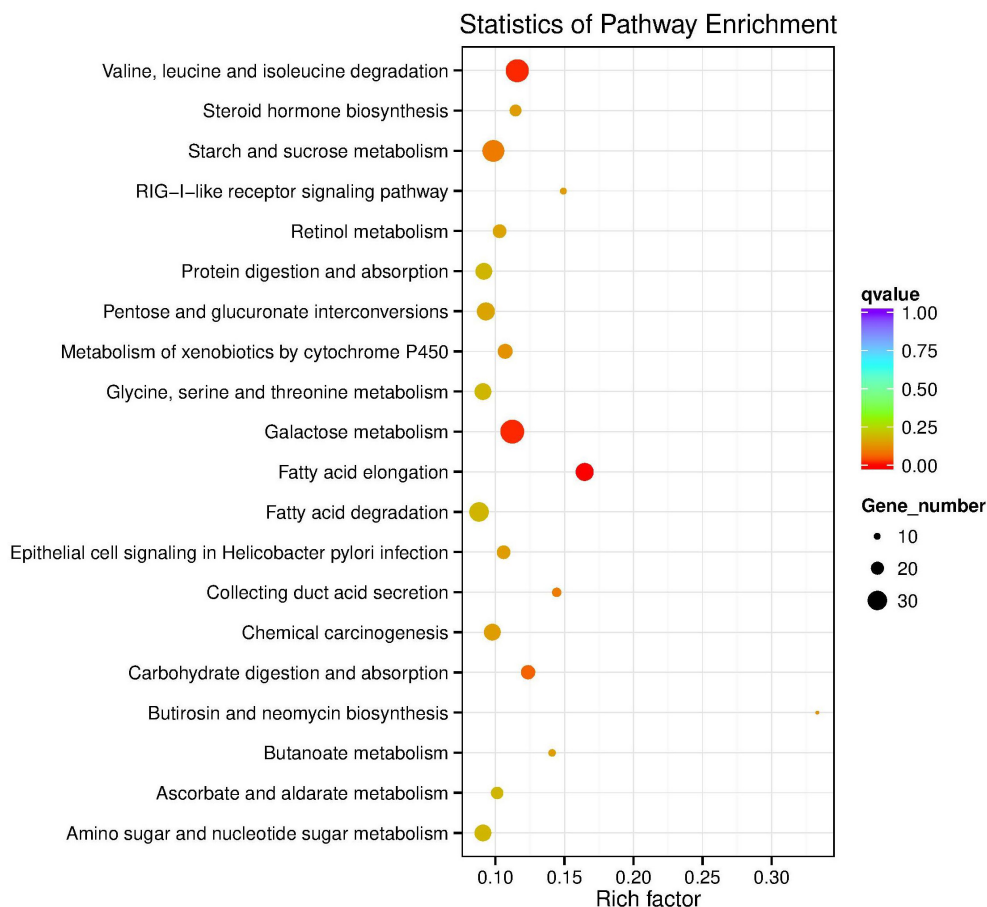


FIGURE 3 | KEGG pathway enrichment scatter plot. The vertical axis represents the path name, and the horizontal axis represents the path factor corresponding to the Rich factor. The size of the q -value is represented by the color of the point. The smaller the q -value, the closer the color is to the red color. The number of differential genes included in each pathway are expressed by the size of the point. (Top 20 enriched pathways are represented in scatter plot).

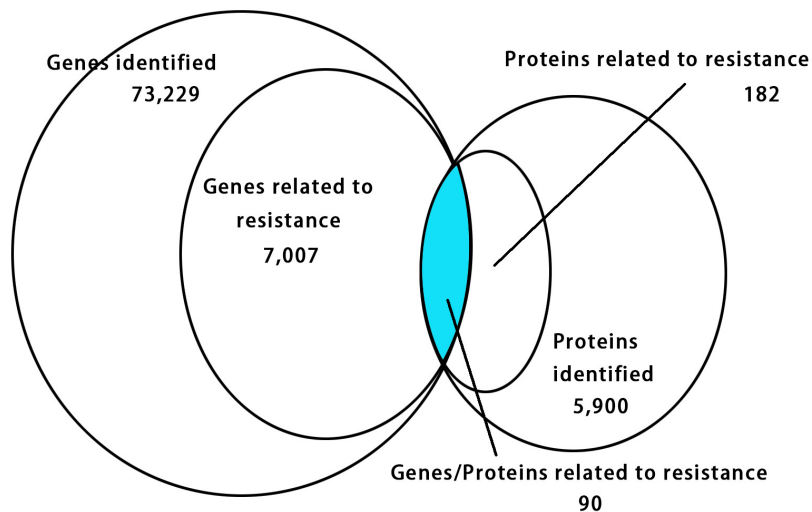
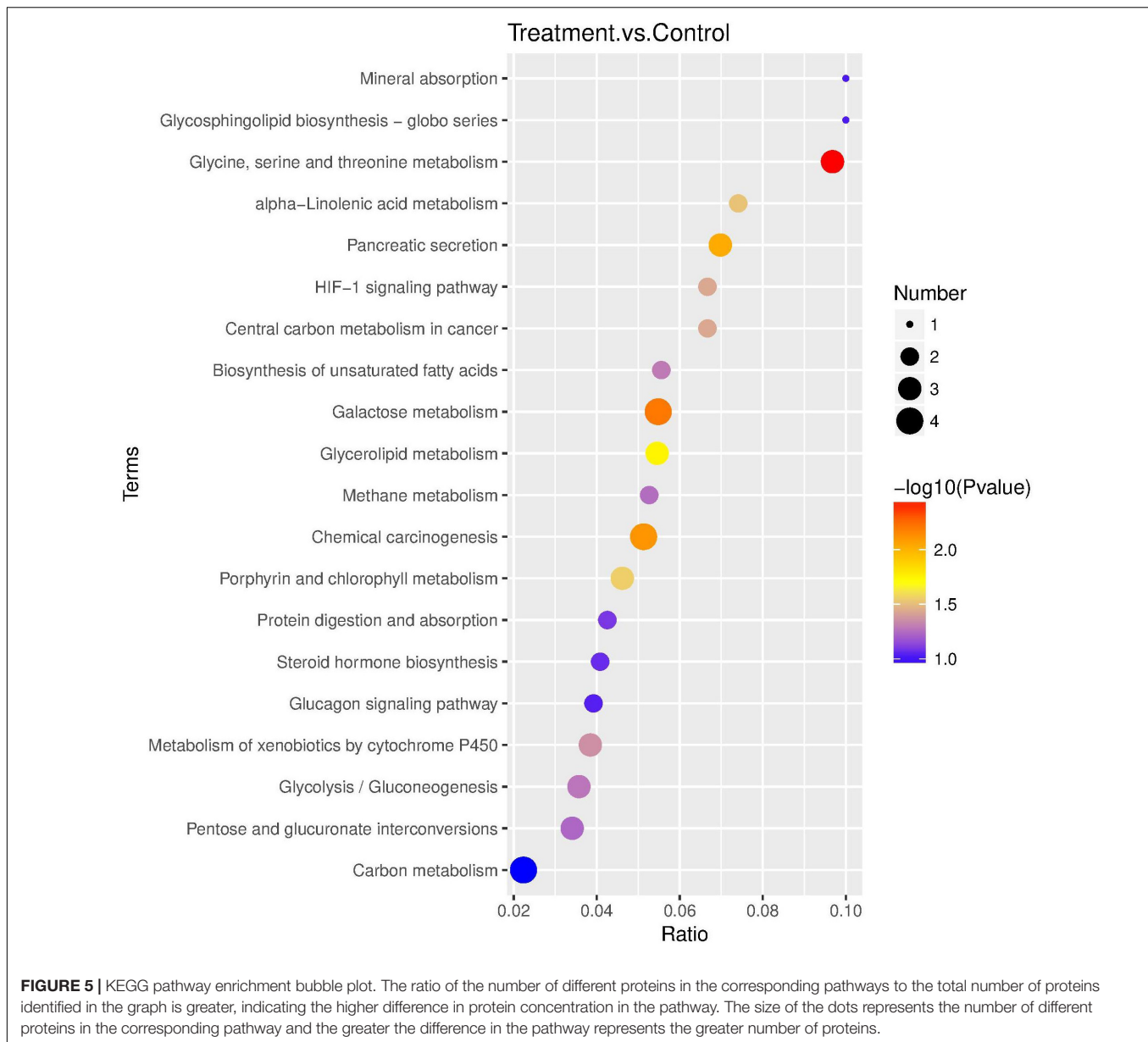


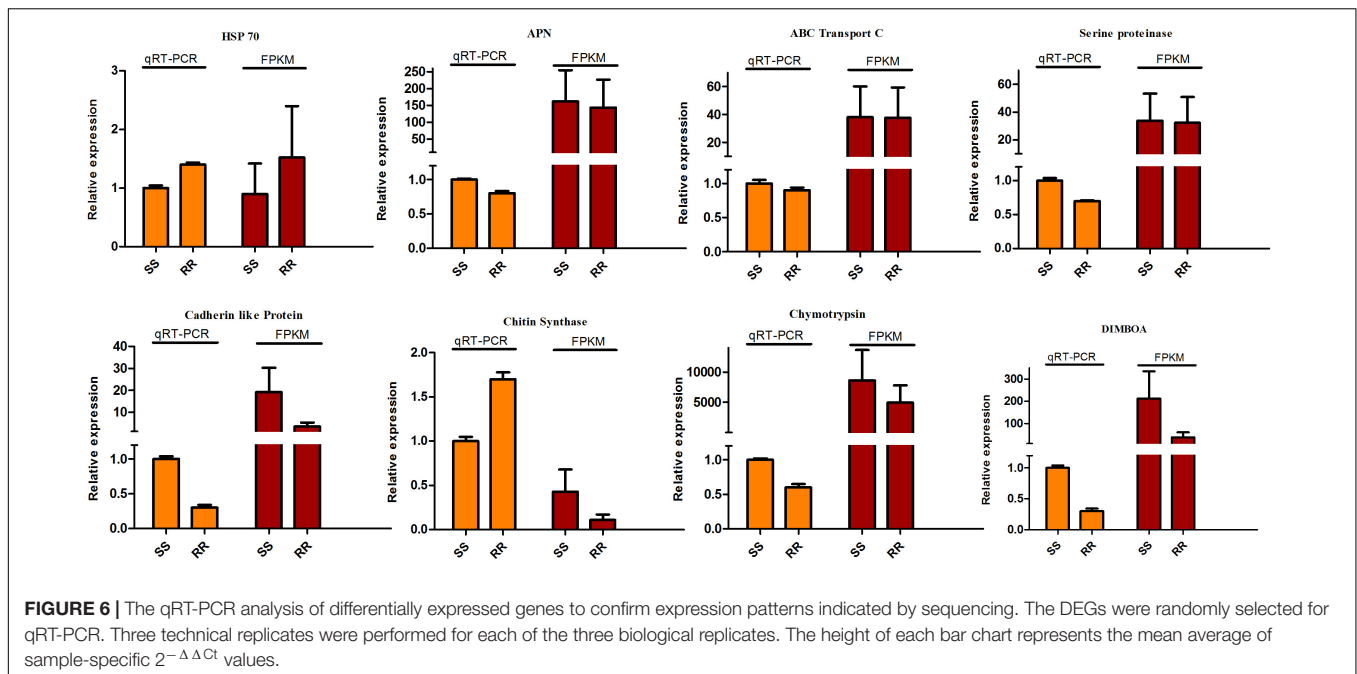
FIGURE 4 | Venn diagram showing relationship between the differentially expressed proteins and genes. The values in each circle represents the quantity of genes or proteins, including identified genes and proteins and genes and proteins related to resistance, respectively, genes/proteins related to resistance together. The cut-off value of \log_2 fold changes for up-regulated and down-regulated gene/protein was $+1/-1$.



and proteins (**Supplementary Table S6**). Genes including ABC transporter C2, DIMBOA-induced cytochrome P450, cadherin-like protein, and chymotrypsin-like serine protease were down-regulated, whereas aldehyde dehydrogenase and *N*-acetylgalactosaminidase were up-regulated at both transcriptional and translational levels (**Supplementary Table S6**). Likewise, physiologically similar responses were documented in *Sarcophaga crassipalpis*, *Drosophila melanogaster*, and *Caenorhabditis elegans* transcriptomes (Ragland et al., 2010). However, we found some genes with the opposite trends, like the trypsin-like serine protease and NADH dehydrogenase were up-regulated at the transcriptional level and down-regulated at the translational level. This effect could be attributed to the difference in expression time (Ragland et al., 2010). Moreover, expression profiles of mRNA and protein levels do not always

correlate (Nie et al., 2006), and differences in directional changes between proteomic and transcriptome are possibly due to the single sampling time-point and changes in protein versus genes *in vivo* are rarely studied (Popescu et al., 2010). Similarly, the difference between differentially expressed transcripts and proteins will most likely be the normal rather than exception, without a fully sequenced ACB genome.

In the present study, several transcripts which are down-regulated in the ACB-AhR strain were previously documented as important candidate Bt resistance genes/proteins or other genes involved in insecticide resistance in numerous insects including APN, ABCC3, V-ATPase, trypsin-like serine protease, DIMBOA-induced cytochrome P450, ALP, GST, chymotrypsin-like serine protease family members and chitin synthase (**Supplementary Table S1**). The significant



correlation between transcriptome/proteomic and qRT-PCR results further verified the gene expression data, providing assurance in the reliability of our data (Figure 6). Different isoforms of APNs and CAD together with ALP have been reported as Cry toxin receptors (Pigott and Ellar, 2007). The same phenomenon of down-regulation of cadherin as a Cry toxin receptor was previously described in ACB-AhR, and ACr strains in both microarrays and qRT-PCR results (Zhang et al., 2017), supporting the results of a prior study which indicated the down-regulation of *Ocad* gene in Cry1Ac-resistant strain (Jin et al., 2014). Down-regulation of APN transcripts in resistant strains has been shown to be involved in the Bt mode of action and mechanisms of the resistance are reported through proteomics and molecular analyses to different Cry toxins (Nanath et al., 2015; Zhang et al., 2017). Interestingly, we also found dozens of genes annotated to APN were over-expressed in ACB-AhR strain. GO, and up-regulation (2.47 to 5.65-times) of APN1 (ABQ51393.1), APN2 (ACF34999.1), APN3 (AEO12689.1), and APN4 (ACF34998.2) of APN in Cry1Ab resistance in ACB-AhR (Xu et al., 2015). It was also reported that APN encoded by the Unigenes59183-mk was significantly up-regulated in a Cry1Ac-resistant strain of *P. xylostella* (Lei et al., 2014), and AAEL012774 annotated to APN were over-expressed found by proteomic approaches in LiTOX strain (Tetreau et al., 2012). According to pore formation model, the expression of Bt receptors genes like cadherin should be down-regulated in the resistant insects (Peng et al., 2010; Vachon et al., 2012). However, the current findings were not always consistent with this approach. Based on our observations, along with previous studies, we speculated that APN and cadherin-like protein should have a significant role in Cry1Ah resistance of ACB, and resistance might be associated with the expression of multiple receptors between ACB-AhR and

ACB-BtS strains. In this study, the GPI-anchored metabolic pathway was detected in GO annotation, and KEGG pathway analysis and GPI-anchored proteins like ALP were identified as Cry-toxin receptors. ALP expression was under-expressed in *H. virescens* population in a laboratory experiment (Jurat-fuentes et al., 2003). The identification of ALP has been described as Cry-toxin receptors for Cry1Ac (Chen et al., 2015; Jin et al., 2015), Cry11Aa (Fernandez et al., 2006), and Cry4Ba toxins (Moonsom et al., 2007). Generally, the Bt resistance confers changes in the structure of Cry toxin receptors rather than in their expression (Griffitts and Aroian, 2005). These changes in the expression of Cry receptors are likely the result of different genetic mechanisms involving mutations in regulatory regions or genome rearrangements which cause rapid adaptations to new environmental pressure such as an insecticide treatment.

Moreover, GO function and KEGG pathway enrichment were analyzed for DEGs of ACB-AhR and ACB-BtS to find other Cry1Ah-resistance related genes in ACB, as these pathway analyses provide a valuable understanding of the biological process, cellular components and molecular functions of target sites (Ji et al., 2012). The results revealed that the majority of these DEGs were down-regulated in ACB-AhR both from RNA-seq and iTRAQ analyses. These results are in agreement with the Cry1Ab resistance study which showed down-regulation (85.8%) of DEGs in the ACB-AhR strain (Xu et al., 2015). However, the majority of DEGs was significantly up-regulated in a Cry1Ac-resistant strain of *P. xylostella* (Lei et al., 2014). These findings proposed that Cry1Ah-resistance mechanism in ACB can differ from *P. xylostella*, or expression level of up-regulation of genes could be compensated for the loss of other catalytic genes to reduce the fitness costs of Cry toxin resistance. In the present study, expression of mostly genes annotated to GSTs, ATPase, ABCC3, trypsin, and P450 was lower in ACB-AhR.

In previous studies, GSTs and P450 genes were reported to confer resistance and were involved in detoxifications of xenobiotic (Xu et al., 2015; Pavlidi et al., 2018), as well as trypsin, which is considered the main proteinase involved in Bt toxin activation and detoxification (Liu et al., 2014). The ABC proteins are membrane bound transporters associated with the movement of solutes across the lipid membranes and have been linked to Bt toxin resistance in the midgut of Cry1Ac and Cry1Ab resistant larvae (Dermauw and Van, 2014; Tabashnik, 2015). In this study, differentially expressed ABC transporters between ACB-BtS and ACB-AhR strains included ABCC1, ABCC2, ABCC3, ABCC4, and Abcc10 and the majority of them were down-regulated. Previously, ABCC2 has been reported to be involved in Cry1Ac resistance in three lepidopterans (Gahan et al., 2010; Baxter et al., 2011). Additionally, eight genes annotated to ABCC2 were detected in Cry1Ac resistance strain of *P. xylostella*, and the majority of them were down-regulated (Lei et al., 2014). Nevertheless, ABCC2 can function as Cry1A toxin receptors (Degen, 2004), and further investigations are required to elucidate the role of genes within Bt resistance mechanisms.

Generally, Cyt toxins identified in the case of previously documented Cry1Ah toxin as receptors of Cry toxins (Perez et al., 2005), possibly contribute to overcoming receptor alterations in ACB-AhR strain. As previously reported in several Bt resistant insects, Cry-toxin resistance might be linked with multiple receptors, and there is a possibility that Cry1Ah resistance is associated with differential expression of Bt toxin receptors between ACB-AhR and ACB-BtS strains. In conclusion, this is the first combination of a study using RNA-Seq

and iTRAQ approaches on Cry1Ah-toxin binding, which led to the identification of a longer length of genes in ACB. Besides, Cry1Ah-resistance in ACB is involved in metabolic and catalytic pathways. DEGs and DEPs would be used for further studies on the membrane receptors which are associated with Cry1Ah-resistance and could lead to the analysis of genetic differences between Bt resistant and susceptible strains of ACB.

AUTHOR CONTRIBUTIONS

KH and MS designed the experiments. MS and TZ performed the experiments and analyzed the data. ZW and KH provided the insect, reagents, and materials. MS drafted the manuscript. KH and TZ reviewed and edited the manuscript.

FUNDING

This research was funded by the National Science and Technology Major Project of China (2016ZX080011003).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fphys.2019.00027/full#supplementary-material>

REFERENCES

- Anders, S., and Huber, W. (2010). Differential expression analysis for sequence count data. *Genome Biol.* 11:R106. doi: 10.1186/gb-2010-11-10-r106
- Audic, S., and Claverie, J. M. (1997). The significance of digital gene expression profiles. *Genome Res.* 7, 986–995. doi: 10.1101/gr.7.10.986
- Baxter, S. W., Badenes-Perez, F. R., Morrison, A., Vogel, H., Crickmore, N., Kain, W., et al. (2011). Parallel evolution of *Bacillus thuringiensis* toxin resistance in Lepidoptera. *Genetics* 189, 675–679. doi: 10.1534/genetics.111.130971
- Bravo, A., and Soberon, M. (2008). How to cope with insect resistance to Bt toxins? *Trends Biotechnol.* 26, 573–579. doi: 10.1016/j.tibtech.2008.06.005
- Chen, C., Ai, H., Ren, J., Li, W., Li, P., Qiao, R., et al. (2011). A global view of the porcine transcriptome in three tissues from a full-sib pair with extreme phenotypes in growth and fat deposition by paired-end RNA sequencing. *BMC Genomics* 12:448. doi: 10.1186/1471-2164-12-448
- Chen, W., Liu, C., Xiao, Y., Zhang, D., Zhang, Y., Li, X., et al. (2015). A toxin-binding alkaline phosphatase fragment synergizes Bt toxin Cry1Ac against susceptible and resistant *Helicoverpa armigera*. *PLoS One* 10:e0126288. doi: 10.1371/journal.pone.0126288
- Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization, and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676. doi: 10.1093/bioinformatics/bti610
- Cui, L., Rui, C., Yang, D., Wang, Z., and Yuan, H. (2017). De novo transcriptome and expression profile analyses of the Asian corn borer (*Ostrinia furnacalis*) reveals relevant flubendiamide response genes. *BMC Genomics* 18:20. doi: 10.1186/s12864-016-3431-6
- Daniel, A., Sangadala, S., Dean, D. H., and Adang, M. J. (2002). Denaturation of either *Manduca sexta* aminopeptidase N or *Bacillus thuringiensis* Cry1A toxins exposes binding epitopes hidden under non-denaturing conditions. *Appl. Environ. Microbiol.* 68, 2106–2112. doi: 10.1128/AEM.68.5.2106-2112.2002
- Degen, T. (2004). High genetic variability of herbivore-induced volatile emission within a broad range of maize inbred lines. *Plant Physiol.* 135, 1928–1938. doi: 10.1104/pp.104.039891
- Dermauw, W., and Van, L. T. (2014). The ABC gene family in arthropods: comparative genomics and role in insecticide transport and resistance. *Insect Biochem. Mol. Biol.* 45, 89–110. doi: 10.1016/j.ibmb.2013.11.001
- Fernandez, L. E., Aimanova, K. G., Gill, S. S., Bravo, A., and Soberón, M. (2006). A GPI-anchored alkaline phosphatase is a functional midgut receptor of Cry11Aa toxin in *Aedes aegypti* larvae. *Biochem. J.* 394, 77–84. doi: 10.1042/BJ20051517
- Gahan, L. J., Pauchet, Y., Vogel, H., and Heckel, D. G. (2010). An ABC transporter mutation is correlated with insect resistance to *Bacillus thuringiensis* Cry1Ac toxin. *PLoS Genet.* 6:e1001248. doi: 10.1371/journal.pgen.1001248
- Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2013). Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883
- Griffitts, J. S., and Aroian, R. V. (2005). Many roads to resistance: how invertebrates adapt to Bt toxins. *Bioessays* 27, 614–624. doi: 10.1002/bies.20239
- He, K. L., Wang, Z. Y., Zhou, D., Wen, L., and Song, Y. (2003). Evaluation of transgenic Bt corn for resistance to the Asian corn borer (Lepidoptera: Pyralidae). *J. Econ. Entomol.* 96, 935–940. doi: 10.1603/0022-0493-96.3.935
- Ji, Z., Wang, G., Xie, Z., Wang, J., Zhang, C., Dong, F., et al. (2012). Identification of novel and differentially expressed microRNAs of dairy goat mammary gland tissues using solexa sequencing and bioinformatics. *PLoS One* 7:e49463. doi: 10.1371/journal.pone.0049463
- Jin, T., Chang, X., Gatehouse, A. M. R., Wang, Z., Edwards, M. G., and He, K. (2014). Downregulation and mutation of a Cadherin gene associated with

- Cry1Ac resistance in the Asian corn borer, *Ostrinia furnacalis* (Guenée). *Toxins* 6, 2676–2693. doi: 10.3390/toxins6092676
- Jin, T., Duan, X., Bravo, A., Soberón, M., Wang, Z., and He, K. (2015). Identification of an alkaline phosphatase as a putative Cry1Ac binding protein in *Ostrinia furnacalis* (Guenée). *Pestic. Biochem. Physiol.* 131, 80–86. doi: 10.1016/j.pestbp.2015.12.008
- Jurat-Fuentes, J. L., and Adang, M. J. (2007). A proteomic approach to study Cry1Ac binding proteins and their alterations in resistant *Heliothis virescens* larvae. *J. Invertebr. Pathol.* 95, 187–191. doi: 10.1016/j.jip.2007.01.008
- Jurat-fuentes, J. L., Gould, F. L., Michael, J., and Adang, M. J. (2003). Dual resistance to *Bacillus thuringiensis* Cry1Ac and Cry2Aa toxins in *Heliothis virescens* suggests multiple mechanisms of resistance. *Appl. Environ. Microbiol.* 69, 5898–5906. doi: 10.1128/AEM.69.10.5898-5906.2003
- Lei, Y., Zhu, X., Xie, W., Wu, Q., Wang, S., Guo, Z., et al. (2014). Midgut transcriptome response to a Cry toxin in the diamondback moth, *Plutella xylostella* (Lepidoptera: Plutellidae). *Gene* 533, 180–187. doi: 10.1016/j.gene.2013.09.091
- Li, L. T., Zhu, Y. B., Ma, J. F., Li, Z. Y., and Dong, Z. P. (2013). An analysis of the *Athetis lepigone* transcriptome from four developmental stages. *PLoS One* 8:e73911. doi: 10.1371/journal.pone.0073911
- Lin, Q., Jin, F., Hu, Z., Chen, H., Yin, F., Li, Z., et al. (2013). Transcriptome analysis of chlorantraniliprole resistance development in the diamondback moth *Plutella xylostella*. *PLoS One* 8:e72314. doi: 10.1371/journal.pone.0072314
- Liu, C., Xiao, Y., Li, X., Oppert, B., Tabashnik, B. E., and Wu, K. (2014). Cis-mediated down-regulation of a trypsin gene associated with Bt resistance in cotton bollworm. *Sci. Rep.* 4:7219. doi: 10.1038/srep07219
- Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262
- Moonsom, S., Chaisri, U., Kasinrer, W., and Angsuthanasombat, C. (2007). Binding characteristics to mosquito-larval midgut proteins of the cloned domain II-III fragment from the *Bacillus thuringiensis* Cry4Ba toxin. *J. Biochem. Mol. Biol.* 40, 783–790. doi: 10.5483/BMBRep.2007.40.5.783
- Nanoth, V. N., Wang, H., Eyun, S., Moriyama, E. N., Coates, B. S., Miller, N. J., et al. (2015). Transcriptional analysis of susceptible and resistant European corn borer strains and their response to Cry1F protoxin. *BMC Genomics* 16:558. doi: 10.1186/s12864-015-1751-6
- Nicolas, J. A., Tamayo, N. V., Caocili, B. L., Reddy, M. S., Ilao, R. I., and Faylon, P. S. (2013). “Improving the yield of glutinous white corn by distance of planting and use of biocontrol agents for management of Asian corn borer, *Ostrinia furnacalis* Guenee,” in *Proceedings of 3rd Asian Conference on Plant Growth-Promoting Rhizobacteria (PGPR) and other Microbials*, Manila, 50–74.
- Nie, L., Wu, G., and Zhang, W. (2006). Correlation between mRNA and protein abundance in *Desulfovibrio vulgaris*: a multiple regression to identify sources of variations. *Biochem. Biophys. Res. Commun.* 339, 603–610. doi: 10.1016/j.bbrc.2005.11.055
- Noirel, J., Evans, C., Salim, M., Mukherjee, J., Ow, S. Y., Pandhal, J., et al. (2011). Methods in quantitative proteomics: setting iTRAQ on the right track. *Curr. Proteomics* 8, 17–30. doi: 10.2174/157016411794697408
- Paris, M., Melodelima, C., Coissac, E., Tetreau, G., Reynaud, S., David, J. P., et al. (2012). Transcription profiling of resistance to Bti toxins in the mosquito *Aedes aegypti* using next-generation sequencing. *J. Invertebr. Pathol.* 109, 201–208. doi: 10.1016/j.jip.2011.11.004
- Pavlidis, N., Vontas, J., and Van, L. T. (2018). The role of glutathione S-transferases (GSTs) in insecticide resistance in crop pests and disease vectors. *Curr. Opin. Insect Sci.* 27, 97–102. doi: 10.1016/j.cois.2018.04.007
- Peng, D., Xu, X., Ye, W., Yu, Z., and Sun, M. (2010). *Helicoverpa armigera* cadherin fragment enhances Cry1Ac insecticidal activity by facilitating toxin-oligomer formation. *Appl. Microbiol. Biotechnol.* 85, 1033–1040. doi: 10.1007/s00253-009-2142-1
- Perez, C., Fernandez, L. E., Sun, J., Folch, J. L., Gill, S. S., Soberon, M., et al. (2005). *Bacillus thuringiensis* subsp. *israelensis* Cyt1Aa synergizes Cry11Aa toxin by functioning as a membrane-bound receptor. *Proc. Natl. Acad. Sci. U.S.A.* 102, 18303–18308. doi: 10.1073/pnas.0505494102
- Pigott, C. R., and Ellar, D. J. (2007). Role of receptors in *Bacillus thuringiensis* crystal toxin activity. *Microbiol. Mol. Biol. Rev.* 71, 255–281. doi: 10.1128/MMBR.00034-06
- Popescu, J. T., Martyniuk, C. J., Denslow, N. D., and Trudeau, V. L. (2010). Rapid dopaminergic modulation of the fish hypothalamic transcriptome and proteome. *PLoS One* 5:e12338. doi: 10.1371/journal.pone.0012338
- Ragland, G. J., Denlinger, D. L., and Hahn, D. A. (2010). Mechanisms of suspended animation are revealed by transcript profiling of diapause in the flesh fly. *Proc. Natl. Acad. Sci. U.S.A.* 107, 14909–14914. doi: 10.1073/pnas.1007075107
- Shabbir, M. Z., Quan, Y., Wang, Z. Y., Bravo, A., Soberón, M., and He, K. L. (2018). Characterization of the Cry1Ah resistance in Asian corn borer and its cross-resistance to other *Bacillus thuringiensis* toxins. *Sci. Rep.* 8:234. doi: 10.1038/s41598-017-18586-2
- Soberón, M., Gill, S. S., and Bravo, A. (2009). Signaling versus punching hole: how do *Bacillus thuringiensis* toxins kill insect midgut cells? *Cell. Mol. Life Sci.* 66, 1337–1349. doi: 10.1007/s00018-008-8330-9
- Song, J., Sun, R., Li, D., Tan, F., Li, X., Jiang, P., et al. (2012). An improvement of shotgun proteomics analysis by adding next-generation sequencing transcriptome data in orange. *PLoS One* 7:e39494. doi: 10.1371/journal.pone.0039494
- Tabashnik, B. E. (2015). ABCs of insect resistance to Bt. *PLoS Genet.* 11:e1005646. doi: 10.1371/journal.pgen.1005646
- Tabashnik, B. E., Brévault, T., and Carrière, Y. (2013). Insect resistance to Bt crops: lessons from the first billion acres. *Nat. Biotechnol.* 31, 510–521. doi: 10.1038/nbt.2597
- Tabashnik, B. E., Carrière, Y., Dennehy, T. J., Morin, S., Sisterson, M. S., Roush, R. T., et al. (2003). Insect resistance to transgenic Bt crops: lessons from the laboratory and field. *J. Econ. Entomol.* 96, 1031–1038. doi: 10.1603/0022-0493-96.4.1031
- Tetreau, G., Bayyareddy, K., Jones, C. M., Stalinski, R., Riaz, M. A., Paris, M., et al. (2012). Larval midgut modifications associated with Bti resistance in the yellow fever mosquito using proteomic and transcriptomic approaches. *BMC Genomics* 13:248. doi: 10.1186/1471-2164-13-248
- Tiewsi, K., and Wang, P. (2011). Differential alteration of two aminopeptidases N associated with resistance to *Bacillus thuringiensis* toxin Cry1Ac in cabbage looper. *Proc. Natl. Acad. Sci. U.S.A.* 108, 14037–14042. doi: 10.1073/pnas.1102555108
- Vachon, V., Laprade, R., and Schwartz, J. L. (2012). Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: a critical review. *J. Invertebr. Pathol.* 111, 1–12. doi: 10.1016/j.jip.2012.05.001
- Wang, X., Li, X., Shen, A., and Wu, Y. (2010). Baseline susceptibility of the diamondback moth (Lepidoptera: Plutellidae) to chlorantraniliprole in china. *J. Econ. Entomol.* 103, 843–848. doi: 10.1603/EC09367
- Wang, Y., Wang, Y., Wang, Z. Y., Bravo, A., Soberón, M., and He, K. L. (2016). Genetic basis of Cry1F-resistance in a laboratory selected Asian corn borer strain and its cross-resistance to other *Bacillus thuringiensis* toxins. *PLoS One* 11:0161189. doi: 10.1371/journal.pone.0161189
- Wang, Z. Y., He, K. L., Zhang, F., Lu, X., and Babendreier, D. (2014). Mass rearing and release of *Trichogramma* for biological control of insect pests of corn in China. *Biol. Control* 68, 136–144. doi: 10.1016/j.biocontrol.2013.06.015
- Xiang, L. X., He, D., Dong, W. R., Zhang, Y. Y., and Shao, J. Z. (2010). Deep sequencing-based transcriptome profiling analysis of bacteria-challenged *Lateolabrax japonicus* reveals insight into the immune-relevant genes in marine fish. *BMC Genomics* 11:472. doi: 10.1186/1471-2164-11-472
- Xu, L., Ferry, N., Wang, Z. Y., Zhang, J., Edwards, M. G., Gatehouse, A. M. R., et al. (2013). A proteomic approach to study the mechanism of tolerance to Bt toxins in *Ostrinia furnacalis* larvae selected for resistance to Cry1Ab. *Transgenic Res.* 22, 1155–1166. doi: 10.1007/s11248-013-9718-3
- Xu, L., Wang, Z. Y., Zhang, J., Ferry, N., Edwards, M. G., Gatehouse, A. M. R., et al. (2014). Characterization of four midgut aminopeptidase N isozymes from *Ostrinia furnacalis* strains with different susceptibilities to *Bacillus thuringiensis*. *J. Invertebr. Pathol.* 115, 95–98. doi: 10.1016/j.jip.2013.11.001
- Xu, L. N., Wang, Y. Q., Wang, Z. Y., Hu, B. J., Ling, Y. H., and He, K. L. (2015). Transcriptome differences between Cry1Ab resistant and susceptible strains of Asian corn borer. *BMC Genomics* 16:173. doi: 10.1186/s12864-015-1362-2
- Xu, X., Yu, L., and Wu, Y. (2005). Disruption of a cadherin gene associated with resistance to Cry1Ac δ -endotoxin of *Bacillus thuringiensis* in *Helicoverpa armigera*. *Appl. Environ. Microbiol.* 71, 948–954. doi: 10.1128/AEM.71.2.948
- Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11:R14. doi: 10.1186/gb-2010-11-2-r14

- Zhang, T., Coates, B. S., Wang, Y. Q., Wang, Y., Bai, S., Wang, Z. Y., et al. (2017). Down-regulation of aminopeptidase N and ABC transporter subfamily G transcripts in Cry1Ab and Cry1Ac resistant Asian corn borer, *Ostrinia furnacalis* (Lepidoptera: Crambidae). *Int. J. Biol. Sci.* 13, 835–851. doi: 10.7150/ijbs.18868
- Zhang, T., He, K., and Wang, Z. (2016). Transcriptome comparison analysis of *Ostrinia furnacalis* in four developmental stages. *Sci. Rep.* 6:35008. doi: 10.1038/srep35008
- Zhang, T., He, M., Gatehouse, A. M. R., Wang, Z. Y., Edwards, M. G., Li, Q., et al. (2014). Inheritance patterns, dominance, and cross-resistance of Cry1Ab- and Cry1Ac-selected *Ostrinia furnacalis* (Guenée). *Toxins* 6, 2694–2707. doi: 10.3390/toxins6092694
- Zhang, X., Candas, M., Griko, N. B., Taussig, R., and Bulla, L. A. (2006). A mechanism of cell death involving an adenylyl cyclase/PKA signaling pathway is induced by the Cry1Ab toxin of *Bacillus thuringiensis*. *Proc. Natl. Acad. Sci. U.S.A.* 103, 9897–9902. doi: 10.1073/pnas.0604017103
- Zhang, Y., Liu, Y., Ren, Y., Liu, Y., Liang, G., Song, F., et al. (2013). Overexpression of a novel *Cry1Ie* gene confers resistance to Cry1Ac-resistant cotton bollworm in transgenic lines of maize. *Plant Cell Tissue Organ Cult.* 115, 151–158. doi: 10.1007/s11240-013-0348-5
- Conflict of Interest Statement:** 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.

Copyright © 2019 Shabbir, Zhang, Wang and He. 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.



Extracranial Estimation of Neural Mass Model Parameters Using the Unscented Kalman Filter

Lara Escuain-Poole^{1*}, Jordi Garcia-Ojalvo² and Antonio J. Pons¹

¹ Physics Department, Polytechnic University of Catalonia, Terrassa, Spain, ² Department of Experimental and Health Sciences, Pompeu Fabra University, Barcelona, Spain

OPEN ACCESS

Edited by:

Axel Hutt,
German Meteorological Service,
Germany

Reviewed by:

Nicola Politi,
Politecnico di Torino, Italy
Lili Lei,
Nanjing University, China
Sergiy Zhuk,
IBM Research (Ireland), Ireland

*Correspondence:

Lara Escuain-Poole
lara.escuain@upc.edu

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 22 May 2018

Accepted: 14 September 2018

Published: 15 October 2018

Citation:

Escuain-Poole L, Garcia-Ojalvo J and
Pons AJ (2018) Extracranial Estimation
of Neural Mass Model Parameters
Using the Unscented Kalman Filter.
Front. Appl. Math. Stat. 4:46.
doi: 10.3389/fams.2018.00046

Data assimilation, defined as the fusion of data with preexisting knowledge, is particularly suited to elucidating underlying phenomena from noisy/insufficient observations. Although this approach has been widely used in diverse fields, only recently have efforts been directed to problems in neuroscience, using mainly intracranial data and thus limiting its applicability to invasive measurements involving electrode implants. Here we intend to apply data assimilation to non-invasive electroencephalography (EEG) measurements to infer brain states and their characteristics. For this purpose, we use Kalman filtering to combine synthetic EEG data with a coupled neural-mass model together with Ary's model of the head, which projects intracranial signals onto the scalp. Our results show that using several extracranial electrodes allows to successfully estimate the state and a specific parameter of the model, whereas one single electrode provides only a very partial and insufficient view of the system. The superiority of using multiple extracranial electrodes over using only one, be it intra- or extra-cranial, is shown in different dynamical behaviours. Our results show potential toward future clinical applications of the method.

Keywords: Unscented Kalman filter, data assimilation, EEG, neural mass model, parameter estimation

1. INTRODUCTION

After several decades studying its morphology and dynamics [1], the basic mechanisms that describe the functioning of the brain are still far from being completely understood. There are different reasons that explain this arduous route toward understanding this organ. First, the neurons that form the brain are very diverse morphologically [2] and dynamically [3]. Second, these neurons are connected to each other in extremely large numbers and forming very complex networks [4], whose structural characteristics are still mostly unknown. And third, brain dynamics are very irregular and complex [5, 6]. The opposed views of an essentially noisy brain and a deterministic brain exhibiting chaotic activity have been often contrasted. On the one hand there is multiple evidence, both theoretical and experimental, that justifies a stochastic view of the brain [7, 8]. On the other hand, other studies reveal deterministic, or rather reproducible, dynamical behaviour [9, 10] both at the microscopic scale [11] and at the mesoscale recorded by electroencephalograms (EEG) or magnetoencephalograms (MEG) [12]. The reality is probably a combination of the two views. The fact that the brain receives continuous external inputs from the sensory system also makes its dynamical and experimental interpretation more complex because, even though experiments are designed to minimise uncontrolled inputs, they cannot completely rule them out. Another important limitation for studying the brain is that experimental recordings (such as EEG or fMRI) are almost always indirect reflections of the underlying neural activity [13].

A way of facing the complexities described above is by systematically comparing the experimental observations of brain activity with mathematical models based on specific hypotheses, which can thereby be validated or disproven. Modelling cerebral activity has been attempted both with top-down and bottom-up approaches [14–18]. Many of these theoretical models are simplifications that capture the basic ingredients of brain dynamics, while others are detailed accounts of the dynamics of neurons that necessarily forgo the description of the whole brain. In that context, a more feasible scale of study is the mesoscopic scale [19–25]. Many of the modern experimental techniques record information coming from populations of neurons working together. Neural mass models describe the activity of these populations mathematically using reasonably simple equations [26, 27]. These models can describe both the intrinsic oscillatory behaviour recorded at the mesoscale or event-related responses [28, 29] with morphologically plausible assumptions for their construction.

In all modelling strategies, however, identifying realistic values for the parameters of the model is a challenging task. One way to address this problem is by integrating experimental information into the models using Bayesian inference [30–35]. This strategy has started to be pursued by using Kalman filtering to integrate experimental data at both the microscopic scale of neuronal networks [36–38] and the mesoscopic scale of neural mass models [39–42]. This data assimilation approach aims to tackle the high level of noise in neuronal activity, and allows to estimate both the state and the parameters of the theoretical model using the experimental data available. The method has been used to estimate, for example, the effective connectivity that characterises epileptic seizures on a patient-specific basis (see [43] and references therein). Kalman filtering has also been used to analyse the suppression of epileptic seizures in coupled neural mass models [40, 44], and the induction of the anesthetized state by drugs [45]. But these studies use mainly invasive intracranial signals, and it would be desirable to extend them to non-invasive extracranial measurements such as EEG. Intracranial signals can be translated into EEG signals in a forward manner [46, 47], and, in the opposite direction, solving the inverse problem allows to infer intracranial signals from EEG recordings [48–50]. In this paper we advance the applications of Kalman filtering in neuroscience by extending the current procedures with a model of the head, exploring the possibilities of using non-invasive scalp measurements.

2. METHODS

To obtain a reliable estimation of the state and the dynamics of the brain, we require a biologically inspired mathematical model of its dynamics, experimental data (as non-invasive as possible), and the means of fusing both sources of information together. In this paper, for the purpose of providing a proof-of-concept of our proposed data assimilation approach, we use *in silico* data, instead of real experimental observations, generated by Jansen and Rit's model [26, 51], as a way to represent the dynamical evolution of the cortical structures. We then use the unscented

Kalman filter as our data assimilation algorithm to estimate the state and a specific parameter of the model jointly [52–54].

2.1. Mesoscopic Neural Mass Model

Jansen and Rit's model [26, 51] describes the mesoscopic activity of a population of neurons [55, 56], providing a good compromise between physiological realism and computational simplicity. This model reduces the neuronal diversity of a cortical column to three interacting populations: pyramidal neurons, excitatory interneurons, and inhibitory interneurons. The larger pyramidal population excites both groups of interneurons, which in turn feed back into the pyramidal cells. In our approximation, the pyramidal population is also driven by neighbouring columns and by excitatory noise representing the input from distant areas of the brain. The model is given by the following set of coupled second-order differential equations [26, 57]:

$$\ddot{x}_0(t) + 2a\dot{x}_0(t) + a^2x_0(t) = Aa \text{Sigm}[x_1^i(t) - x_2^i(t)], \quad (1)$$

$$\ddot{x}_1^i(t) + 2a\dot{x}_1^i(t) + a^2x_1^i(t) = Aa \left(p^i(t) + k \sum_{j=1}^{N_d} K^{ij} \text{Sigm}(x_1^j(t - \tau^{ij}) - x_2^j(t - \tau^{ij})) + C_2 \text{Sigm}[C_1x_0^i(t)] \right), \quad (2)$$

$$\ddot{x}_2^i(t) + 2b\dot{x}_2^i(t) + b^2x_2^i(t) = Bb (C_4 \text{Sigm}[C_3x_0^i(t)]), \quad (3)$$

where x_0 is the average excitatory postsynaptic potential (PSP) coming to the two interneuron populations, and x_1 (x_2) is the average excitatory (inhibitory) PSP which inputs to the pyramidal population. The superindex $i = 1 \dots N_d$ runs over all the coupled cortical columns (dipole sources) of the model. The quantity $x_1 - x_2$ is the net PSP of the pyramidal neurons, which produces the signal detected by extracranial electrodes, and is therefore our observable. The sigmoid function $\text{Sigm}(v)$ converts the net average PSP of a population, v , into an average firing rate:

$$\text{Sigm}(v) = \frac{2e_0}{1 + e^{\gamma(v_0 - v)}}, \quad (4)$$

where e_0 is the maximum firing rate of the population, γ controls the slope of the sigmoid, and v_0 is the post-synaptic potential for which a 50% firing rate is obtained. The resulting firing rate is then transformed back into an average PSP by the second-order differential Equations 1–3.

The parameters A and B in the right-hand side of Equations 1–3 are the amplitudes of the excitatory and inhibitory post-synaptic potentials, and a and b are the lumped representations of the sums of the reciprocal of the time constant of the passive membrane, and all other spatially distributed delays in the dendritic network. The parameters C_1 to C_4 are connectivity constants that govern the interactions between populations, $p^i(t)$ is a stochastic external input that adds dynamic noise to the system, and the summation term represents the input from other coupled cortical columns. The strength of the coupling is modulated by k , with K denoting the adjacency matrix. When generating the *in silico* data we consider that column i receives the signal of column j with a delay τ^{ij} [58]. This is because we want to generate data, in a controlled way, with a model as complex and rich in dynamics as possible to mimic real data. However, Kalman filtering, as defined

TABLE 1 | Description and default values of the parameters for the system of neural masses.

Param.	Description	Value
A	EPSP amplitude	3.25 mV
B	IPSP amplitude	22.00 mV
a	Rate constant for the excitatory population*	100 s ⁻¹
b	Rate constant for the inhibitory population*	50 s ⁻¹
C ₁	Strength of synaptic connections from PC to EI	135
C ₂	Strength of synaptic connections from II to PC	108
C ₃	Strength of synaptic connections from PC to II	33.75
C ₄	Strength of synaptic connections from EI to PC	33.75
e ₀	Maximum firing rate of the population	2.5 s ⁻¹
v ₀	Mean threshold of the population	6 mV
γ	Steepness of the sigmoidal transformation	0.56 mV ⁻¹
k	Coupling constant	5 or 10 according to the experiment
K	Adjacency matrix	$K_{ij} = 1, i \neq j$ $K_{ii} = 0, i = j$
τ	Delay	Between 15 and 21 ms, according to distance [58]
p ₀	External input	200 s ⁻¹
ε	Intensity of the noise	100 s ⁻¹

*Lumped representation of the sum of the reciprocal of the time constant of passive membrane and all other spatially distributed delays. See the Results section for details of the configuration of each numerical experiment. Here, PC refers to pyramidal cells, EI to excitatory interneurons, II to inhibitory interneurons, EPSP to excitatory post-synaptic potential, and IPSP to inhibitory post-synaptic potential.

in Equations 10–11 below, does not include temporal delays. Therefore, for simplicity, the model used during the filtering does not consider delays. **Table 1** provides the descriptions and values of these parameters. The electrical activity detected by the electrodes on the scalp is originated by the weighted sum of the averaged membrane potential of the pyramidal cells of all the cortical columns, $x^i(t) = x_1^i(t) - x_2^i(t)$ [59], using a head model as described below.

2.2. Head Model

The main contribution of this paper is the use of multichannel extracranial data to obtain information about the neuronal populations inside the brain using data assimilation. To accomplish this, we use synthetic EEG data generated *in silico* using Jansen and Rit's model and Ary's head model. To that end, we transform the output $\mathbf{x}(t)$ of the neural masses to EEG signals $\mathbf{z}(t)$ in the electrodes (see **Figure 1**). This transformation is mediated by a lead field matrix [47], which builds on the basic idea of calculating the electric potential caused by a dipole source [13] on a three-layer isotropic hemisphere of radius 1 [46, 60] that represents the three main tissues that impact brain activity readings (brain, skull, and scalp). The lead field matrix also contains information about the geometry of the problem (e.g., locations of cortical columns and electrodes) and about the

electrophysiology of the head (e.g., conductivities of the different tissues). The following equations show the potential $V^{e,i}$ on an electrode e , located at \mathbf{r}_e^e [61], caused by the dipole $\mathbf{q}^i(t) = x^i(t)\hat{\mathbf{q}}^i$ generated by the cortical column i , located at \mathbf{r}_q^i and oriented as $\hat{\mathbf{q}}^i$. In these equations, $e = 1, \dots, N_e$, where N_e is the total number of electrodes, and $i = 1, \dots, N_d$, where N_d is the total number of dipoles. Vectors are typeset in bold and modules are in regular type.

$$V^{e,i}(\mathbf{r}_e^e; \mathbf{r}_q^i, \mathbf{q}^i) \cong v^1(\mathbf{r}_e^e; \mu_1 \mathbf{r}_q^i, \rho_1 \mathbf{q}^i) + v^2(\mathbf{r}_e^e; \mu_2 \mathbf{r}_q^i, \rho_2 \mathbf{q}^i) + v^3(\mathbf{r}_e^e; \mu_3 \mathbf{r}_q^i, \rho_3 \mathbf{q}^i), \quad (5)$$

$$v^1(\mathbf{r}_e^e; \mathbf{r}_q^i, \mathbf{q}^i) = \left((c_1^{e,i,1} - c_2^{e,i,1}(\mathbf{r}_e^e \cdot \mathbf{r}_q^i)) \mathbf{r}_q^i + c_2^{e,i,1}(\mathbf{r}_q^i)^2 \mathbf{r}_e^e \right) \cdot \mathbf{q}^i, \quad (6)$$

$$v^2(\mathbf{r}_e^e; \mathbf{r}_q^i, \mathbf{q}^i) = \left((c_1^{e,i,2} - c_2^{e,i,2}(\mathbf{r}_e^e \cdot \mathbf{r}_q^i)) \mathbf{r}_q^i + c_2^{e,i,2}(\mathbf{r}_q^i)^2 \mathbf{r}_e^e \right) \cdot \mathbf{q}^i, \quad (7)$$

$$v^3(\mathbf{r}_e^e; \mathbf{r}_q^i, \mathbf{q}^i) = \left((c_1^{e,i,3} - c_2^{e,i,3}(\mathbf{r}_e^e \cdot \mathbf{r}_q^i)) \mathbf{r}_q^i + c_2^{e,i,3}(\mathbf{r}_q^i)^2 \mathbf{r}_e^e \right) \cdot \mathbf{q}^i. \quad (8)$$

In these expressions,

$$c_1^{e,i,s} = \frac{1}{4\pi\sigma^s(\mathbf{r}_q^i)^2} \left(2 \frac{\mathbf{d}^{e,i} \cdot \mathbf{r}_q^i}{(\mathbf{d}^{e,i})^3} + \frac{1}{\mathbf{d}^{e,i}} - \frac{1}{\mathbf{r}_e^e} \right),$$

$$c_2^{e,i,s} = \frac{1}{4\pi\sigma^s(\mathbf{r}_q^i)^2} \left(\frac{2}{(\mathbf{d}^{e,i})^3} + \frac{\mathbf{d}^{e,i} + \mathbf{r}_e^e}{\mathbf{r}_e^e \Gamma(\mathbf{r}_e^e, \mathbf{r}_q^i)} \right), \quad (9)$$

$$\Gamma(\mathbf{r}_e^e, \mathbf{r}_q^i) = \mathbf{d}^{e,i} (\mathbf{r}_e^e \mathbf{d}^{e,i} + (\mathbf{r}_e^e)^2 - (\mathbf{r}_q^i \cdot \mathbf{r}_e^e)).$$

The tangential conductivity of each layer is represented by σ^s [60] and ρ_s and μ_s are the Berg parameters relative to it [62] (see **Table 2**). The parameter $\mathbf{d}^{e,i} = \mathbf{r}_e^e - \mathbf{r}_q^i$ is the relative position of the electrode e under consideration with respect to the position of the dipole i .

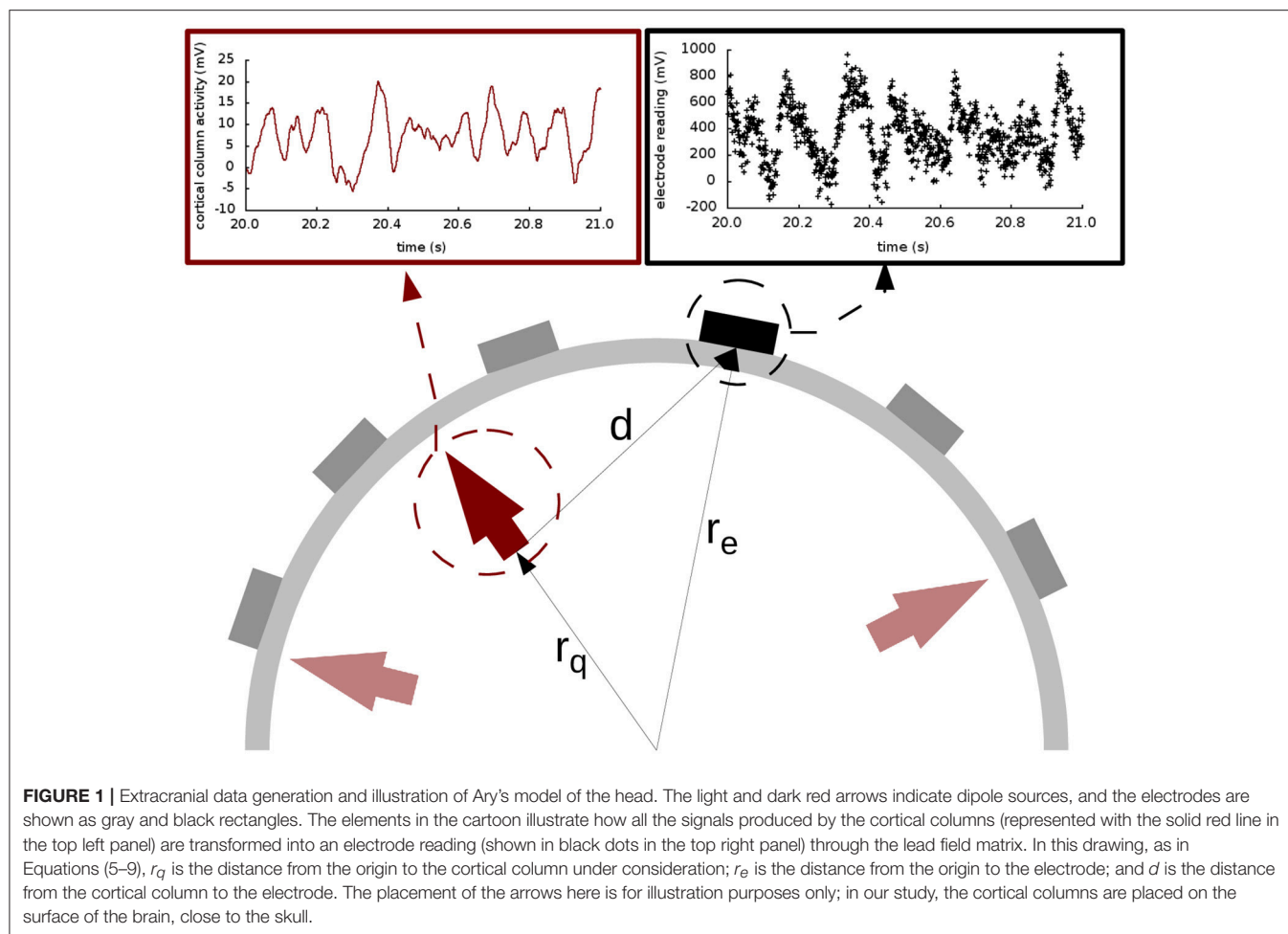
2.3. The Unscented Kalman Filter for Data Assimilation

The Unscented Kalman Filter (UKF) is our algorithm of choice to bring together the dynamical state of the model and the *in silico* data. It is a standard tool in the field of systems and control engineering, and has been shown to be both computationally efficient and robust even when dealing with stochastic nonlinear systems [63]. In our case, the computational burden— $O(n^3)$, where n is the size of the state—is acceptable for a biologically reasonable number of sources. In order to simultaneously estimate the state and parameters of the model described by Equations (1)–(3), we regard it as a discrete-time state-space dynamical system of the following form:

$$\mathbf{x}_{k+1} = \mathbf{F}(\mathbf{x}_k) + \mathbf{v}_k \quad (10)$$

$$\mathbf{z}_k = \mathbf{H}(\mathbf{x}_k) + \mathbf{w}_k \quad (11)$$

where $\mathbf{x} = (x_0^1, x_1^1, x_2^1, x_0^2, \dots, x_2^{N_d}, \theta^1, \dots, \theta^{N_p}) \in \mathbb{R}^{n_x}$ is the state vector (related to the variables and parameters of the model), with



θ^P being the parameters to estimate, which obey the equations $\dot{\theta}^P = 0$. (In our joint estimation of the parameters, these are included in the state vector together with the system variables). The vector $z \in \mathbb{R}^{n_z}$ is the measurement vector (our *in silico* EEG readings). The vectors v and w are uncertainty terms that account for process noise and measurement noise, respectively, with Gaussian distributions $p(v) \sim N(0, Q)$ and $p(w) \sim N(0, R)$, respectively. The process transition F is obtained with a numerical implementation of Equations (1)–(3), as described below. Finally, H relates the state to measurement space, which is either $V^{e,i}$ (in the case of simulated EEG), or x^i (in the case of simulated electrocorticography). Interestingly, where EEG is concerned, this basic part of the Kalman filter is in our case implemented by the skull, the effect of which is represented by the lead field matrix, based on Ary's head model and introduced above.

The UKF is a recursive predictor-corrector-type algorithm that aims to minimise the mean square error of the estimated states and parameters over time. For each time step it calculates a prediction of the state and parameters of the system, which is corrected when the information from a measurement is incorporated. The amount of confidence given to the model and measurement is quantified by the Kalman gain K , which

TABLE 2 | Values of the Berg parameters for the three layers [60, 62].

Parameter	Layer 1	Layer 2	Layer 3
Tangential conductivity σ^S	1.0	0.0125	1.0
Berg parameter ρ_s	0.9901	0.7687	0.4421
Berg parameter μ_s	0.0659	0.2389	0.3561

is calculated at each time step based on prediction covariances as well as model and measurement error covariances (Q and R , respectively). For more details on the implementation of the filter, the reader is referred to the **Appendix** and to Kalman [54], Merwe and Wan [52], Julier and Uhlmann [53], and Solonen et al. [64].

2.4. Generation of *in silico* Datasets

For this paper three different *in silico* datasets were generated. We consider both simulated electrocorticography (ECoG, intracortical) and electroencephalography (EEG, extracranial) readings (using Ary's model in the latter case). We chose to use three sources because this provides a considerable spatial

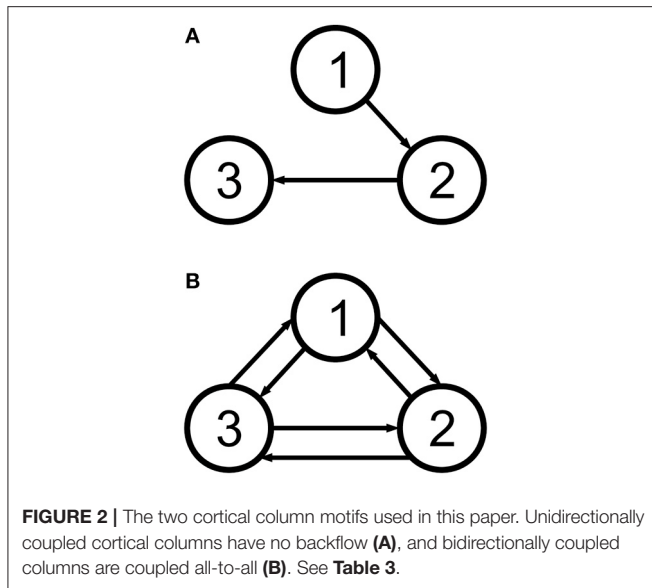


TABLE 3 | Cartesian coordinates of the dipoles used throughout the study.

	x	y	z
dipole 1	0.1688	0.2242	0.2597
dipole 2	0.3766	-0.8520	0.2597
dipole 3	0.6622	-0.2242	-0.1948

The origin of coordinates is the centre of the perimeter of the head.

and temporal richness in the resulting signals, while keeping the system reasonably simple and still biologically plausible [65, 66].

The presence of additional dipoles in the brain, and its influence on the sources of study, is accounted for in the stochastic external input to the sources ($p(t)$, see Equation 2):

$$p(t) = p_0 + \xi(t), \quad (12)$$

where $p_0 = 200 \text{ s}^{-1}$ and $\xi(t)$ is Gaussian white noise [67] of zero mean and correlation $\langle \xi(t)\xi(t') \rangle = 2\epsilon\delta(t - t')$ [68]. At the extracranial level, the other sources also affect the final EEG signal, as well as the different tissues (brain, skull, scalp, and even hair). This is modelled by adding Gaussian noise with zero mean and standard deviation 100 mV (unless otherwise stated) to the simulated EEG.

All datasets used the same locations for the cortical columns [66]. The electrodes were placed using a subset of the equidistant layout, a standard layout for EEG [69] (roughly illustrated in Figures 6–8). The strength of the coupling was set at a medium value so that the cortical columns have a visible effect on one another without fully synchronising behaviours and locking their dynamics (between $k = 5$ and $k = 10$), and the configurations of the couplings are as shown in Figure 2. Table 1 shows representative values for the parameters used in all analyses unless otherwise specified. In this paper we focus on estimating the amplitudes A of the EPSPs of the different cortical columns, and therefore we choose values for these amplitudes that produce signals that reflect various dynamic regimes that we

wish to explore. (The rest of the parameters were fixed to their standard values [26, 51], as described in Table 1).

The numerical solver used to generate the *in silico* time series was the Heun algorithm [70] with a time step of $\Delta t = 1 \text{ ms}$. The length of the data is 100 s in all cases. Using the Heun algorithm together with Equations 1–3 to update the state variables and the lead field matrix (in order to get the potential in the electrodes of the scalp in Equations 5–9), we generate the required map to apply Kalman filtering in Equations 10 and 11. The following equations implement the stochastic Heun algorithm used to update \mathbf{x}_k :

$$\mathbf{x}_{k+1} = \mathbf{x}_k + \frac{1}{2} (F(\mathbf{x}_k) + F(\tilde{\mathbf{x}}_k)) \Delta t + \frac{1}{2} \sum (g(\mathbf{x}_k) + g(\tilde{\mathbf{x}}_k)) X, \quad (13)$$

$$\tilde{\mathbf{x}}_k = \mathbf{x}_k + F(\mathbf{x}_k) \Delta t + g(\mathbf{x}_k) X. \quad (14)$$

Where $g(\dots)$, together with Equation 12, introduces the noise term in Equation 2 and is zero for Equations 1 and 3. In $X = \sqrt{2\epsilon\Delta t}\gamma$, γ are gaussianly distributed random numbers with zero mean and unit variance. At different instants of time, these random numbers are independent from one another.

2.4.1. Three Unidirectionally Coupled Cortical Columns

For the first study the cortical columns were coupled unidirectionally (Figure 2A), as described in Liu and Gao [71]. The parameters were set to standard values [26] for the three cortical columns (see Table 1), except for the first column, in which A_1 was set to 3.58 mV to make it hyperexcitable. Additionally, the three cortical columns had $p_0 = 90 \text{ s}^{-1}$ and $\epsilon = 2 \text{ s}^{-1}$. This first hyperexcitable column causes a spiking cascade in the other two columns. With this experiment, we aimed to compare how extra- and intra-cranial electrodes perform in the case of a behaviour being induced by an input from another column, and not by the column's own parameter configuration. The resulting data can be found in the Data Sheets 1, 2 in Supplementary Material. Please, see the README file (Data Sheet 6) for more information.

2.4.2. Three Bidirectionally Coupled Cortical Columns: Coarse Parameter Estimation

The three cortical columns are located as in the previous section, but coupled bidirectionally (Figure 2B). Additionally, the maximum amplitudes of the excitatory PSPs were set to $A_1 = 4.25 \text{ mV}$, $A_2 = 10.00 \text{ mV}$, and $A_3 = 3.25 \text{ mV}$. These values were chosen to cause the three cortical columns to be in very different dynamical regimes: cortical column 1 operates in a spiking regime; cortical column 2 oscillates with alpha frequency but with an amplitude similar to that of the spikes; and cortical column 3 oscillates in a more standard regime, as described in [26]. Also, the external input $p(t)$ for each of the three cortical columns was set using $p_0 = 200 \text{ s}^{-1}$ and $\epsilon = 100 \text{ s}^{-1}$. Our aim here was to study how the filter performs in an extreme situation, in which the dynamics of the columns are widely different from one another. We intended to explore the outcome of estimating

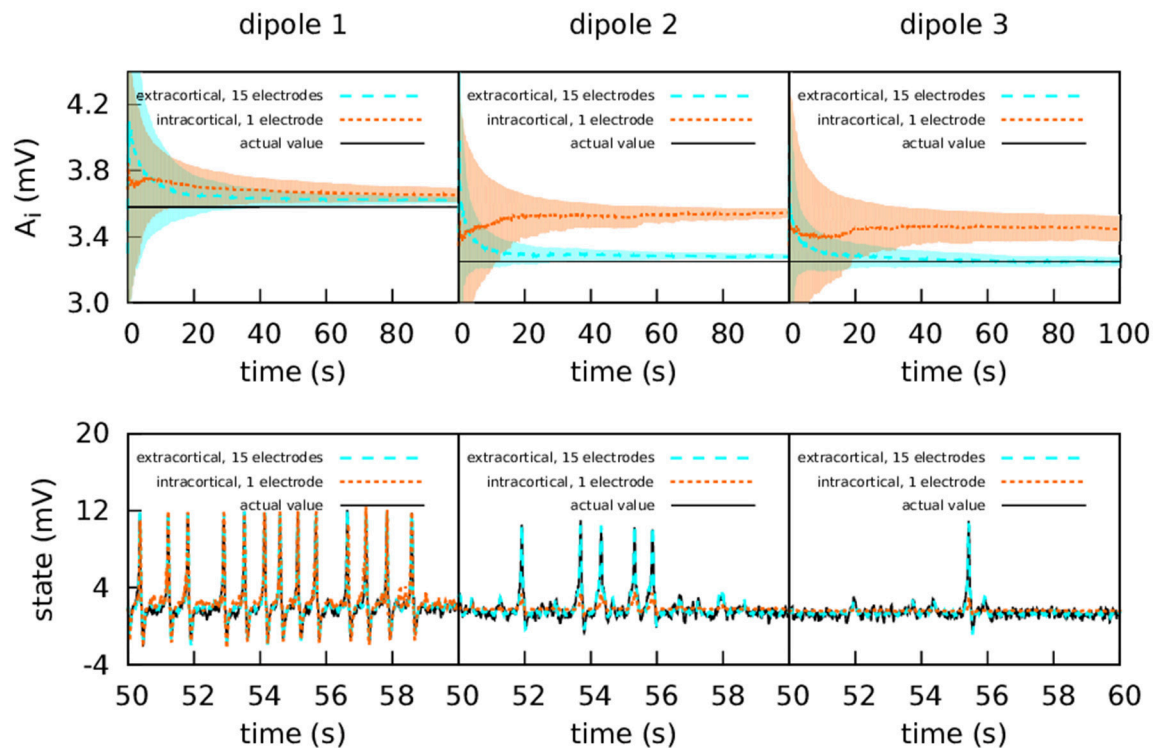


FIGURE 3 | Intracranial and extracranial fittings with propagated excitation along unidirectionally coupled cortical columns. The upper panels show the estimation of parameter A , and the lower panels show the estimations of the observed states. The lines show the averages of the 50 realisations of the estimation, and the shadowed areas indicate the standard deviation. The actual values of the A parameters are $A_1 = 3.58$ mV, $A_2 = 3.25$ mV, and $A_3 = 3.25$ mV, the other parameters being set to standard values (Table 1). All three cortical columns received an external input, $p(t)$, with $p_0 = 90$ s $^{-1}$ and $\epsilon = 2$ s $^{-1}$. The coupling constant was set to $k = 10$. The measurements were corrupted with Gaussian noise of mean 0 and standard deviation 100 mV for extracranial measurements and standard deviation 5 mV for intracortical measurements. Except for cortical column 1, with intracortical data the filter converges to a much higher value than the target, whereas with extracranial data the filter converges to a value which is accurate. In the lower panels it is shown that extracranial estimations of the state are also accurate, whereas intracortical estimations fail to reproduce the spikes correctly.

with single extracranial electrodes as well as the complete set, and to compare with intracranial estimation (Data Sheets 3, 4).

2.4.3. Three Bidirectionally Coupled Cortical Columns: Fine Parameter Estimation

In the previous section, the value of A of one of the cortical columns was much larger than the other two. We now consider the same coupling motif, but with values of the A parameter that are much closer together in value: $A_1 = 3.58$ mV, $A_2 = 3.25$ mV, and $A_3 = 3.10$ mV. (The values defining the external input $p(t)$ remain the same as in the previous experiment). Our goal was to check if the filter can discriminate between the values when they are closer together (Data Sheet 5).

2.5. Filtering

For each of the experiments we conducted 50 realisations of each estimation for the complete state vector, with different initial conditions; all the figures show averages of the 50 estimations, unless otherwise specified. The initial conditions for state and parameter estimations were randomly generated with a normal distribution of zero mean and unit variance; the parameters, however, were constrained to deviate no more than 90% of their actual value as an initial assumption.

The noise covariances Q and R were chosen according to the best knowledge of the system and of the noise corrupting the data. Therefore, Q was set to contemplate the incoming noise to each dipole, i.e., it was set to a null matrix except for the term corresponding to the equation that contains the input $p(t)$ (see Equation 2 and [68]). The matrix R was set to $1000I$ mV 2 . (In practice, in most applications of the Kalman filter, the matrix R is fairly easy to set with the knowledge of the measurement precision as a starting point, but Q is often set by trial and error).

2.6. Ethics Statement

All data used in this manuscript come from numerical simulations of a mathematical model. No human or animal data have therefore been used, and ethics approval was not necessary.

3. RESULTS

In order to compare the performance of the extra- and intracranial approaches to Kalman filtering, we have analysed three different cortical column configurations, each using one of the two motifs shown in Figure 2. Where relevant, two different types of estimations have been used: intracranial and extracranial. Intracranial estimation uses simulated data that would have

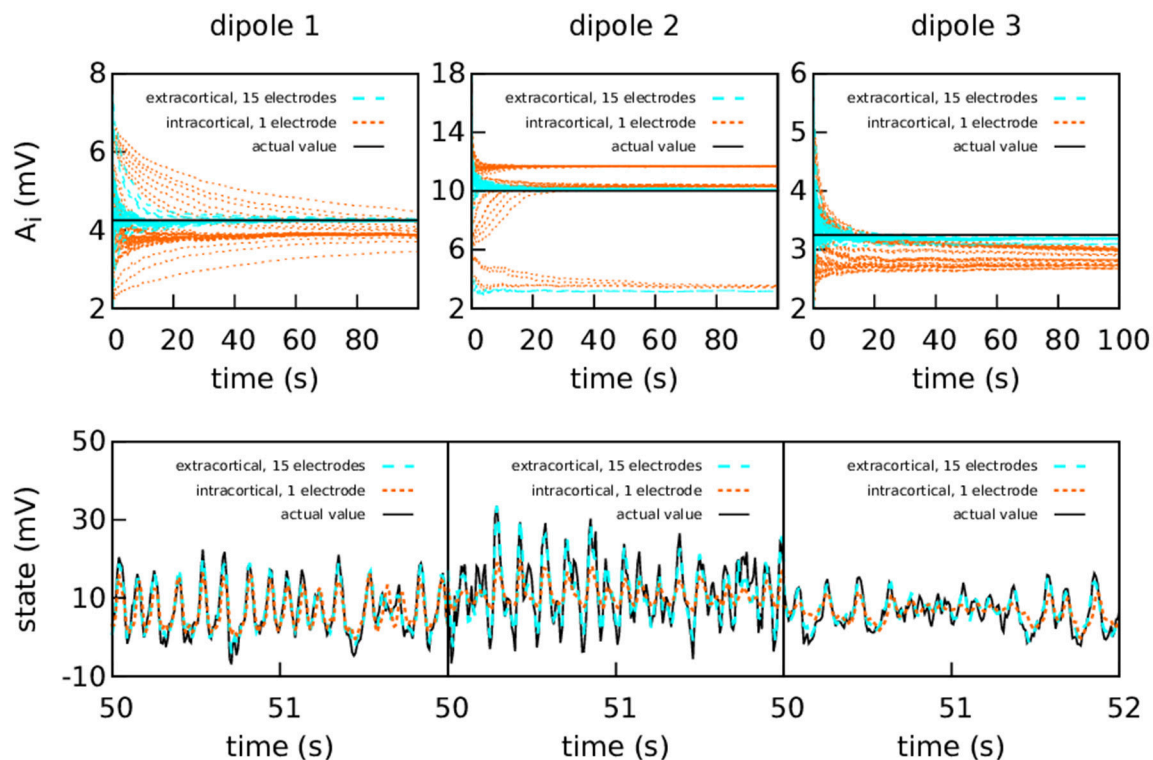


FIGURE 4 | Intracranial and extracranial fittings for coarse parameter estimation in the case of bidirectional coupling. As in the previous figure, the upper panels show the estimation of A for each cortical column and the lower show the estimations of the observed states. The results are shown here without averaging. The actual values of the amplitudes of the EPSPs are $A_1 = 4.25$ mV, $A_2 = 10.00$ mV, and $A_3 = 3.25$ mV; the rest of the parameters were set to standard values (Table 1). The external input $p(t)$ for the three cortical columns had $p_0 = 200$ s $^{-1}$ and $\epsilon = 100$ s $^{-1}$. The coupling constant was set to $k = 5$. The intracortical measurements were corrupted with Gaussian noise of mean 0 and standard deviation 5 mV, while the noise in the extracranial measurements has standard deviation 100 mV. Extracranial estimations of the parameters are both faster and more accurate than intracortical estimations; this applies also to the state, whose dynamics are more faithfully reproduced using multi-electrode extracranial estimation (as shown in the lower panels).

hypothetically been obtained from electrocorticography, that is, using a single intracortical electrode, and is estimated with the data provided by a single location—in other words, the direct output of Jansen and Rit's model. Extracranial estimation, on the other hand, employs simulated data originated from EEG recordings, using several electrodes placed on the skull, and is implemented here with the projection on the head of the model output. We now discuss the results for the three different situations that we have considered.

3.1. Three Unidirectionally Coupled Cortical Columns

In this case, information flows unidirectionally because of the way the cortical columns are coupled [71]. As can be seen in the lower panels of Figure 3, the first cortical column has a random spiking activity, due to the increased value of A and the presence of noise [20]. Due to the architecture of the coupling, cortical column 1 causes cortical columns 2 and 3 to spike also, when otherwise they would have simply fluctuated around their resting level.

The upper panels of Figure 3 show the intracortical and extracranial estimations of A for the three cortical columns. The

estimation for A_1 of the first column converges to its correct value, with both the intra- and extracortical approaches. This was to be expected, since the first cortical column receives no inputs from other elements of the system. In contrast, the intracortical estimations for cortical columns 2 and 3 converge to values significantly higher than their actual value of 3.25 mV. We conjecture that this is caused by the spiking of these two cortical columns, which as mentioned above is due to the influence of cortical column 1. Multi-channel extracranial information, in contrast, allows to see the complete picture of the coupled cortical columns and treat them as a single composed system, contrary to the partial picture obtained from the information provided by the single intracranial recordings. Therefore, estimation is better when using extracranial information with several electrodes, as shown in the upper panels of the figure. The lower panels of Figure 3 show the estimation of the state. The UKF shows great efficacy when the estimation is extracranial, but performs poorly in the case of intracortical estimation (with the exception of cortical column 1, because it has no input from other cortical columns). This highlights the value of extracranial estimation, in which it is possible to take the whole brain into account in a non-invasive manner.

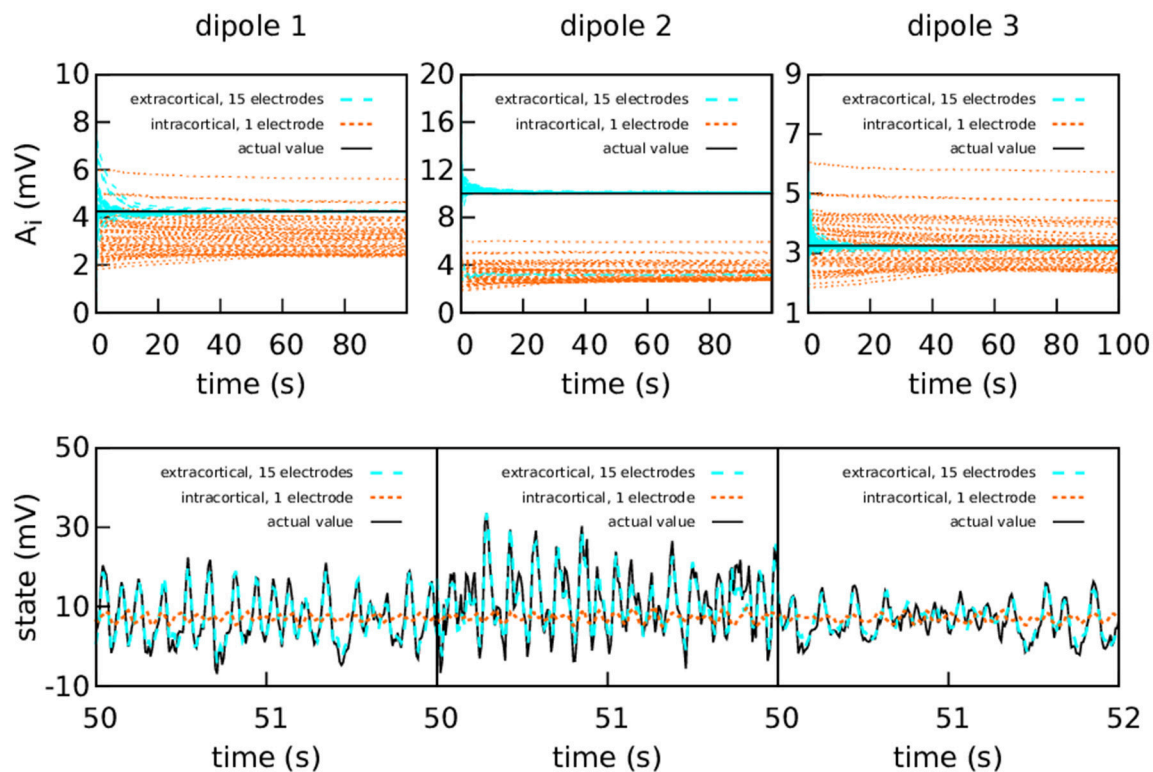


FIGURE 5 | Intracranial and extracranial fittings for coarse parameter estimation, with a higher amount of intracortical measurement noise. The upper panels show the estimation of the EPSPs for each cortical column and the lower panels show the estimations of the observed states. The results are shown here without averaging. The actual values of the amplitudes of the EPSPs are $A_1 = 4.25$ mV, $A_2 = 10.00$ mV, and $A_3 = 3.25$ mV; the rest of the parameters were set to standard values (**Table 1**). The external input $p(t)$ for the three cortical columns had $p_0 = 200$ s $^{-1}$ and $\epsilon = 100$ s $^{-1}$. The coupling constant was set to $k = 5$. The intracortical measurements were corrupted with Gaussian noise of mean 0 and standard deviation 100 mV—about an order of magnitude higher than the noise in the previous graph—, while the noise in the extracranial measurements has standard deviation 100 mV. Extracranial estimations of the parameters are also faster and more accurate than intracortical estimations, more markedly so in this case; as to the state, in this more extreme case, the intracortical estimation does not mimic the evolution of the system in any way.

3.2. Three Bidirectionally Coupled Cortical Columns: Coarse Parameter Estimation

The second experiment aims to explore the possibilities of the filter in more extreme situations, as the parameters were chosen to reflect more diverse dynamical regimes. The following sections describe the results of single- and multichannel estimations.

3.2.1. Moderate Intracortical Measurement Noise

Figure 4 shows again the performance obtained using the simulated data from a set of extracranial electrodes compared to using individual intracortical electrodes for each cortical column. In this case we show the 50 realisations of each filtering, without showing the average. The extracranial data for this experiment were corrupted with a measurement Gaussian noise of zero mean and standard deviation 100 mV; the intracortical data were corrupted with a measurement noise of standard deviation 5 mV in order to maintain similar levels of signal-to-noise ratio.

As shown in **Figure 4**, the intracortical parameter estimations do not approximate the target value very well. In particular, the estimations of A for cortical column 2 converge to three different values depending on the initial conditions. The state estimation

follows the actual state of the system closely only for cortical column 1. The situation is very different when with extracranial electrodes, where all 50 realisations of the estimations converge with much more precision to the correct values for both state and parameters (with the exception of A_2 , which still tends to lower values in a very small quantity of the realisations). Again, extracranial performance is better, in general, to intracortical.

3.2.2. High Intracortical Measurement Noise

The difference between intracranial and extracranial estimation is even larger for higher measurement noise (**Figure 5**). In this case, the amount of noise in the intracortical data was set to the same value as the noise in the extracranial data. The value of R was tuned to reflect the increase in measurement noise, but the intracortical estimations failed to obtain the correct values for the parameters and reproduce the state.

3.2.3. Using One Single Extracranial Electrode

Using the same dataset, we aimed to investigate the outcome of using each extracranial electrode individually [43], as opposed to using the complete subset as until now. Therefore, we used each

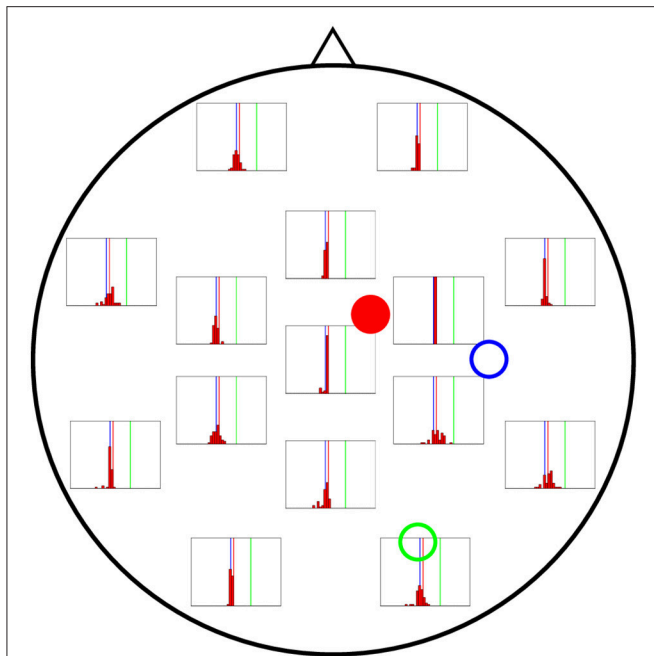


FIGURE 6 | Distribution of 50 realisations of A estimations from a single electrode for cortical column 1 (solid red circle). The histograms are placed at the location of the corresponding measuring electrode, and the location of the three cortical columns generating the activity are shown with coloured circles (with the full circle corresponding to the column whose value of A is being estimated in this figure). Vertical lines with the same colours as the circles mark the corresponding actual A values. The distributions tend to be narrowest in the vicinities of cortical column 1. Nevertheless, they do not group around the target value of $A_1 = 4.25$ mV (vertical red line), as they should, but around that of $A_3 = 3.25$ mV (vertical blue line).

electrode separately to estimate the state and parameters of the complete system, with 50 realisations of the estimation for each electrode. By doing so, we show that the quality of the estimations is strongly dependent on the relative positions of sources and electrodes.

In **Figures 6–8** we present the results for the estimation of parameter A of each of the three cortical columns separately. The histograms show the distribution of the 50 estimations of A using each electrode, placed in the respective position of the electrode in question. Vertical coloured lines in the histograms mark the value of the three A parameters being estimated (one in each figure). The histograms show a strong dependence on space of the quality of the estimations. As a general trait, the estimations are better when the electrodes are near the cortical column whose value of A is being estimated, whereas the more distant electrodes show a wider distribution of final values for the parameter.

In **Figure 6** the distribution of the estimations of A_1 are shown. The distributions tend to be narrowest in the vicinities of the cortical column whose A value is being estimated. However, it is noteworthy that the histograms obtained from the observations in distant electrodes tend to group not around the actual value of $A_1 = 4.25$ mV (red vertical line), but of $A_3 = 3.25$ mV (blue vertical line). This result suggests that the algorithm is unable

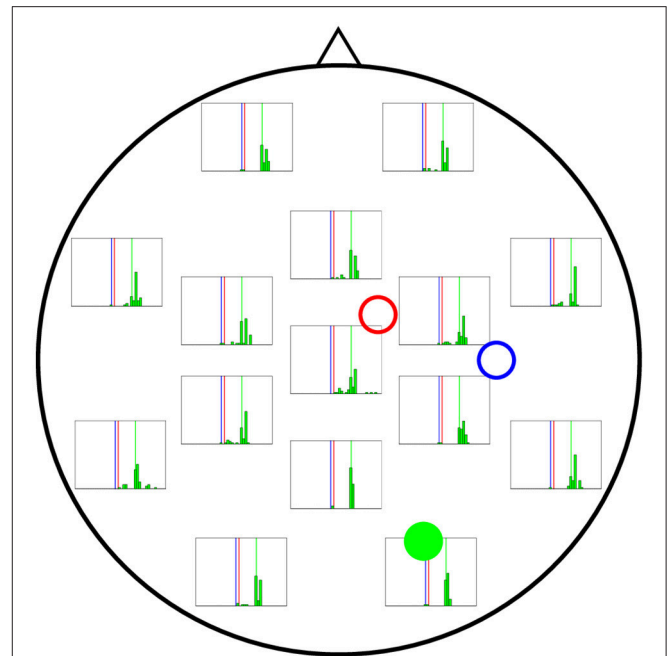


FIGURE 7 | Distribution of 50 realisations of A estimations from a single electrode for cortical column 2 (solid green circle). The distributions here are wider than for A_1 and A_3 , although they still tend to be more accurate near the cortical column (solid green circle) and group around the target value of $A_2 = 10.00$ mV (vertical green line).

to distinguish the origin of the EEG activity when sources and electrodes are distant from each other.

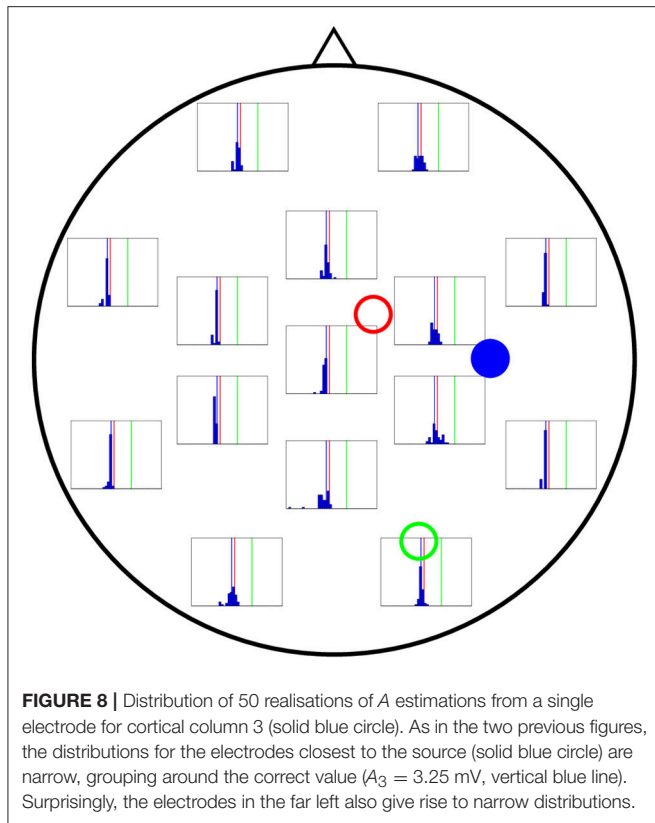
Figure 7 shows the results of the estimation of A_2 (actual value shown by vertical green lines), revealing wider distributions in general, which indicates a stronger dependence on initial conditions. Although it is true that the electrodes near cortical column 2 perform better in estimating A for that column, the difference with more distant electrodes is not as large as for the estimates of A for cortical columns 1 and 3.

Finally, **Figure 8** shows the performance of each electrode when A_3 is being estimated (actual value shown by vertical blue lines in the figure). Interestingly, even the electrodes located at the far left of the figure lead to a good estimate of A , comparable to that coming from the electrodes in the far right, which are closer to column 3 and could therefore be expected to provide a much more accurate estimation.

While the estimations arising from single electrodes are reasonably accurate in some cases, using the complete set of 15 electrodes invariably yields better results. This is because, in Kalman filtering, combining many sources of information always improves the final estimation, even if some of the sources are inaccurate or incomplete [72].

3.3. Three Bidirectionally Coupled Cortical Columns: Fine Parameter Estimation

In the previous section, the aim was to generate widely different dynamics in each column. We now consider the results of

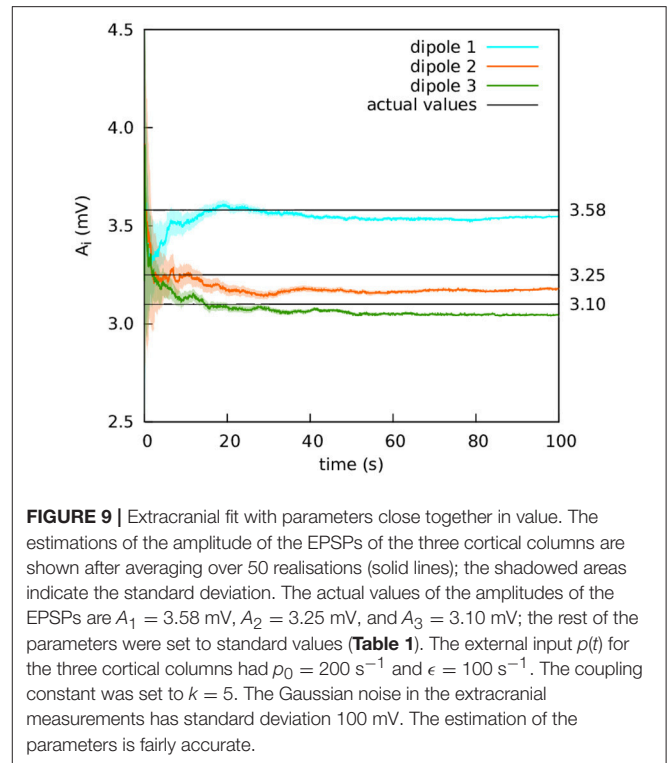


estimating parameters which are much closer to one another. The purpose of this test was to ascertain whether the filter could differentiate between parameters with smaller differences in value. This ability is very important if we expect to use the technique in clinical applications. **Figure 9** shows the extracranial estimation of the A parameters using the complete subset of 15 electrodes. The estimations converge to the actual values with enough accuracy as to give hopes of using the filter in a clinical setting.

4. DISCUSSION

The most important limitation of current data assimilation processes in neuroscience is that the appropriate experimental recordings are usually intracranial. Despite this fact, using Kalman filtering to fit these data to neural mass models shows promise in several contexts and applications. In this study we have modified this type of approach by extending it with a head model, with the aim of integrating non-invasive experimental recordings taken from the scalp (EEG). By increasing the range of recordings in this way, the application of the data assimilation protocols opens up to the large set of situations in which scalp recordings are used. We keep the exploration of the technique using real EEG experimental data in mind, but here we have explored the limitations and advantages of our model using *in silico* data in very well controlled conditions.

Our main goal in this paper has been to show that data assimilation employing multiple non-invasive EEG electrodes



(as coming from scalp EEG measurements) provides a better estimate of the brain's dynamical state than using a single invasive (intracranial) EEG electrode. In particular, we have aimed at contrasting our results with existing work using the latter approach, which has employed a filtering method, namely the unscented Kalman filter [39]. Filtering methods have been so far the method of choice in data assimilation problem in neuroscience [37, 38, 40–43], with variational methods having been used very sparsely [73]. We thus chose to work with a filtering algorithm, the UKF, that is already relatively well characterized in neural model, and which we could therefore use as a benchmark.

We have considered a system comprised of three cortical columns, modelled according to Jansen and Rit's equations and coupled following two different motifs. The cortical columns are all driven by a noisy input coming from the columns of the rest of the brain and sensory stimuli. The signal from the cortical columns is then transferred to the skull, after which it is corrupted with Gaussian noise to simulate electrode readings from EEG. These are then used to estimate the amplitude of the excitatory post-synaptic potentials.

Even though the quality of the experimental measurements at the scalp might be, in general, worse than the intracranial recordings, EEG can always be measured from several positions. This allows to obtain measurements for patients without intracranial implants and also to compensate the potentially low quality of the data by having many recordings at the same time. Besides, the spatial distribution of the electrodes on the scalp allows the information arriving from the whole cortex to be available during the assimilation process. In order to

address these strengths and weaknesses of the scalp recordings with respect to intracranial measurements, we have analysed situations where assimilation with only intracortical recordings may be wanting, where diverse dynamical regimes coexist due to large differences in control parameters in the cortical columns, or where fine changes of the parameters make the discrimination difficult.

The first study considered here involves three columns that are coupled unidirectionally with no backflow. The first cortical column is made hyperexcitable by increasing the excitatory post-synaptic potential to $A_1 = 3.58$ mV; this cortical column causes the second cortical column and, indirectly, the third, to modify their behaviour by inducing spiking. For the intracranial estimations, single intracortical electrodes measured the evolution of the three cortical columns independently; for the extracranial estimations, 15 extracranial electrodes were used simultaneously. Applying the Kalman filter to the extracranial data provided a good estimation of the A parameters and of the dynamical state of the model; the intracortical measurements, however, yielded mixed results. The estimation for cortical column 1 was accurate, whereas for cortical columns 2 and 3 the estimation of A was above the target value and very close to the estimation for cortical column 1 (see orange dashed lines in **Figure 3**). The estimation of the dynamical state of cortical columns 2 and 3 was also worse than the estimation for cortical column 1. We attribute this to the fact that columns 2 and 3 are excited by column 1, which spikes due to a higher value of A . As a consequence, when independently evaluated using the intracranial information, the estimation is higher than the actual value. Therefore we suggest that one intracranial electrode provides only a partial view of the system, and thus cannot capture the behaviours of all three cortical columns and the interactions between them; the use of many electrodes provides a more complete view of the system.

Next we considered a situation in which the dipoles were coupled bidirectionally in an all-to-all configuration. The A parameters were chosen such as to cause different dynamic behaviours in the cortical columns. Three types of fitting via Kalman filtering were performed, using (i) independent intracortical recordings of single cortical columns, (ii) the complete subset of 15 extracranial electrodes, and (iii) single extracranial electrodes. The intracortical data were corrupted with two different levels (medium and high) of measurement noise. For both cases, the multi-electrode extracranial estimation surpasses the intracortical results in both speed of convergence and quality; the difference, however, is more marked in the presence of higher measurement noise in the intracortical recordings. In all these cases, the representation of the dynamical state of the three cortical columns using the complete set of 15 extracranial electrodes nicely matched the actual dynamical state, contrary to the limited match obtained using single intracranial or extracranial recordings. The results for the single electrodes show a significant influence of space on the quality of the estimations, in the sense that estimations of electrodes close to the source are relatively accurate, and electrodes further away from the source might not allow to discriminate the source of the

information correctly, or might completely fail to represent the system.

Finally, we considered the situation of an identical cortical column configuration—in terms of situation and coupling—, except for the values of the EPSPs of the cortical columns. This dataset was filtered only extracranially, with the purpose of evaluating the filter's ability to discriminate parameter values within narrower ranges. The results in this case were also reasonably good, even though the real values of the parameter were much closer to one another, which makes data assimilation more challenging.

Even though the results shown here are better when considering extracranial electrodes, the method has, of course, limitations. For instance, the head model introduces new parameters which should be realistic. The use of Jansen's model, while being a very standard choice in the field, is not mandatory and could be substituted by others. There several alternatives to Ary's head model too. The succesful application of the method with different combinations of these models will, for sure, guide researchers to choose which models are more suitable for the theoretical description of the mesoscale in the brain. Even though the exploration of the dynamics for the different neural mass models or of the different head models might be worth exploring in future works, it lays outside of the scope of this work.

Applications of the method presented here will certainly appear in the field of brain-machine interface, long-term tracking for early diagnosis of degenerative diseases, or short-term tracking during rehabilitation of traumas and strokes. However, the succesful application of the method in each of these fields will require further research.

Taken as a whole, our results show that, independently of the need to explore more realistic situations, extracranial EEG recordings constitute a good candidate to be used together with neural mass models and Kalman filters, provided the method is extended with a head model. With its management of the noise in the system and of the inherent simplifications in neurological models, the Kalman filter is an appropriate tool for tackling the challenges of brain data processing. Using non-invasive techniques in these processes widens the applications of Kalman-based data assimilation methods in neuroscience.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

FUNDING

This work was partially supported by the Spanish Ministry of Economy and Competitiveness and FEDER (project FIS2015-66503) and by the Catalan Government (AGAUR grant FI-DGR 2014-2017). JG-O also acknowledges support from the Catalan Government (project 2014SGR0947), the ICREA Academia programme, and from the María de Maeztu programme for Units of Excellence in R&D (Spanish Ministry of Economy and Competitiveness, MDM-2014-0370).

ACKNOWLEDGMENTS

The authors wish to thank Drs. Maciej Jedynak and Daniel Malagarriaga for helpful discussions.

REFERENCES

- Yuste R. From the neuron doctrine to neural networks. *Nature Rev Neurosci.* (2015) **16**:487–97. doi: 10.1038/nrn3962
- Braitenberg V, Schütz A. Anatomy of the cortex. In: *Statistics and Geometry, No. 18 in Studies of Brain Function*. Berlin: Springer (1991).
- Izhikevich EM, Edelman GM. Large-scale model of mammalian thalamocortical systems. *Proc Natl Acad Sci USA.* (2008) **105**:3593–8. doi: 10.1073/pnas.0712231105
- Callaway EM. *Micro-, Meso- and Macro-Connectomics of the Brain*. Springer (2016). Available online at: <http://link.springer.com/10.1007/978-3-319-27777-6>
- Freeman WJ. Simulation of chaotic EEG patterns with a dynamic model of the olfactory system. *Biol Cybern.* (1987) **56**:139–50. doi: 10.1007/BF00317988
- Faure P, Korn H. Is there chaos in the brain? I. Concepts of nonlinear dynamics and methods of investigation. *C R Acad Sci III.* (2001) **324**:773–93. doi: 10.1016/S0764-4469(01)01377-4
- Shadlen MN, Newsome WT. Noise, neural codes and cortical organization. *Curr Opin Neurobiol.* (1994) **4**:569–79. doi: 10.1016/0959-4388(94)90059-0
- Faisal AA, Selen LPJ, Wolpert DM. Noise in the nervous system. *Nature Rev Neurosci.* (2008) **9**:292–303. doi: 10.1038/nrn2258
- Schiff SJ, Jerger K, Duong DH, Chang T, Spano ML, Ditto WL. Controlling chaos in the brain. *Nature* (1994) **370**:615–20. doi: 10.1038/370615a0
- van Vreeswijk C, Sompolinsky H. Chaos in neuronal networks with balanced excitatory and inhibitory activity. *Science (NY)* (1996) **274**:1724–6. doi: 10.1126/science.274.5293.1724
- Celletti A, Froeschle C, Tetko IV, Villa AEP. Deterministic behaviour of short time series. *Meccanica.* (1999) **34**:147–54. doi: 10.1023/A:1004668310653
- Stam CJ. Nonlinear dynamical analysis of EEG and MEG: review of an emerging field. *Clin Neurophysiol.* (2005) **116**:2266–301. doi: 10.1016/j.clinph.2005.06.011
- Buzsáki G, Anastassiou Ca, Koch C. The origin of extracellular fields and currents—EEG, ECoG, LFP and spikes. *Nat Rev Neurosci.* (2012) **13**:407–20. doi: 10.1038/nrn3241
- Wright JJ, Liley DTJ. Dynamics of the brain at global and microscopic scales: neural Networks and the EEG. *Behav Brain Sci.* (1996) **19**:285. doi: 10.1017/S0140525X00042679
- Rabinovich MI, Varona P, Selverston AI, Abarbanel HDI. Dynamical principles in neuroscience. *Rev Modern Phys.* (2006) **78**:1213. doi: 10.1103/RevModPhys.78.1213
- Eliasmith C, Stewart TC, Choo X, Bekolay T, DeWolf T, Tang C, et al. A large-scale model of the functioning brain. *Science* (2012) **338**:1202–5. doi: 10.1126/science.1225266
- Deco G, Ponce-Alvarez A, Hagmann P, Romani G, Mantini D, Corbetta M. How local excitation-inhibition ratio impacts the whole brain dynamics. *J Neurosci.* (2014) **34**:7886–98. doi: 10.1523/JNEUROSCI.5068-13.2014
- Nogaret A, Meliza CD, Margoliash D, Abarbanel HDI. Automatic construction of predictive neuron models through large scale assimilation of electrophysiological data. *Sci Rep.* (2016) **6**:32749. doi: 10.1038/srep32749
- David O, Friston KJ. A neural mass model for MEG/EEG: coupling and neuronal dynamics. *Neuroimage* (2003) **20**:1743–55. doi: 10.1016/j.neuroimage.2003.07.015
- Grimbert F, Faugeras O. Bifurcation analysis of Jansen's neural mass model. *Neural Comput.* (2006) **18**:3052–68. doi: 10.1162/neco.2006.18.12.3052
- Cona F, Zavaglia M, Massimini M, Rosanova M, Ursino M. A neural mass model of interconnected regions simulates rhythm propagation observed via TMS-EEG. *Neuroimage* (2011) **57**:1045–58. doi: 10.1016/j.neuroimage.2011.05.007
- Coombes S. Large-scale neural dynamics: Simple and complex. *Neuroimage* (2010) **52**:731–9. doi: 10.1016/j.neuroimage.2010.01.045
- Babajani A, Soltanian-Zadeh H. Integrated MEG/EEG and fMRI model based on neural masses. *IEEE Trans Biomed Eng.* (2006) **53**:1794–1801. doi: 10.1109/TBME.2006.873748
- Babiloni F, Cincotti F, Babiloni C, Carducci F, Mattia D, Astolfi L, et al. Estimation of the cortical functional connectivity with the multimodal integration of high-resolution EEG and fMRI data by directed transfer function. *Neuroimage* (2005) **24**:118–31. doi: 10.1016/j.neuroimage.2004.09.036
- Bojak I, Oostendorp TF, Reid AT, Kötter R. Connecting mean field models of neural activity to EEG and fMRI data. *Brain Topogr.* (2010) **23**:139–49. doi: 10.1007/s10548-010-0140-3
- Jansen BH, Rit VG. Electroencephalogram and visual evoked potential generation in a mathematical model of coupled cortical columns. *Biol Cybern.* (1995) **73**:357–66. doi: 10.1007/BF00199471
- Wendling F, Bartolomei F, Bellanger JJ, Chauvel P. Epileptic fast activity can be explained by a model of impaired GABAergic dendritic inhibition. *Euro J Neurosci.* (2002) **15**:1499–508. doi: 10.1046/j.1460-9568.2002.01985.x
- David O, Harrison L, Friston KJ. Modelling event-related responses in the brain. *Neuroimage* (2005) **25**:756–70. doi: 10.1016/j.neuroimage.2004.12.030
- Spiegler A, Knösche TR, Schwab K, Haueisen J, Atay FM. Modeling brain resonance phenomena using a neural mass model. *PLoS Comput Biol.* (2011) **7**:e1002298. doi: 10.1371/journal.pcbi.1002298
- Wright JJ, Robinson PA, Rennie CJ, Gordon E, Bourke PD, Chapman CL, et al. Toward an integrated continuum model of cerebral dynamics: the cerebral rhythms, synchronous oscillation and cortical stability. *BioSystems* (2001) **63**:71–88. doi: 10.1016/S0303-2647(01)00148-4
- Schiff SJ, Sauer T. Kalman filter control of a model of spatiotemporal cortical dynamics. *J Neural Eng.* (2008) **5**:1–8. doi: 10.1088/1741-2560/5/1/001
- Kiebel SJ, Garrido MI, Moran RJ, Friston KJ. Dynamic causal modelling for EEG and MEG. *Cogn Neurodyn.* (2008) **2**:121–36. doi: 10.1007/s11571-008-9038-0
- Shine JM, Koyejo O, Bell PT, Gorgolewski KJ, Gilat M, Poldrack RA. Estimation of dynamic functional connectivity using Multiplication of Temporal Derivatives. *Neuroimage* (2015) **122**:399–407. doi: 10.1016/j.neuroimage.2015.07.064
- Ma X, Chou CA, Sayama H, Chaovalitwongse WA. Brain response pattern identification of fMRI data using a particle swarm optimization-based approach. *Brain Informat.* (2016) **3**:181–92. doi: 10.1007/s40708-016-0049-z
- Alswailhi J, Potthast R, Bojak I, Saddy D, Hutt A. Kernel reconstruction for delayed neural field equations. *J Math Neurosci.* (2018) **8**:3. doi: 10.1186/s13408-018-0058-8
- Voss HU, Timmer J, Kurths J. Nonlinear dynamical system identification from uncertain and indirect measurements. *Int J Bifurcat Chaos* (2004) **14**:1905–33. doi: 10.1142/S0218127404010345
- Hamilton F, Berry T, Peixoto N, Sauer T. Real-time tracking of neuronal network structure using data assimilation. *Phys Rev E.* (2013) **88**:052715. doi: 10.1103/PhysRevE.88.052715
- Hamilton F, Cressman J, Peixoto N, Sauer T. Reconstructing neural dynamics using data assimilation with multiple models. *EPL (Europhys Lett).* (2014) **107**:68005. doi: 10.1209/0295-5075/107/68005
- Freestone DR, Kuhlmann L, Chong MS, Nesic D, Grayden DB, Aram P, et al. Patient-specific neural mass modeling - stochastic and deterministic, methods. In: *Recent Advances in Predicting and Preventing Epileptic Seizures*. (2013) p. 63–82. Available online at: https://hal.archives-ouvertes.fr/hal-00876475/%5Cnhttp://www.worldscientific.com/doi/abs/10.1142/9789814525350_0005

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fams.2018.00046/full#supplementary-material>

40. Shan B, Wang J, Deng B, Wei X, Yu H, Li H. UKF-based closed loop iterative learning control of epileptiform wave in a neural mass model. *Cogn Neurodyn*. (2015) 9:31–40. doi: 10.1007/s11571-014-9306-0
41. López-Cuevas A, Castillo-Toledo B, Medina-Ceja L, Ventura-Mejía C. State and parameter estimation of a neural mass model from electrophysiological signals during the status epilepticus. *Neuroimage*. (2015) 113:374–86. doi: 10.1016/j.neuroimage.2015.02.059
42. Li S, Li J, Li Z. An improved unscented kalman filter based decoder for cortical brain-machine interfaces. *Front Neurosci*. (2016) 10:587. doi: 10.3389/fnins.2016.00587
43. Freestone DR, Karoly PJ, Nešić D, Aram P, Cook MJ, Grayden DB. Estimation of effective connectivity via data-driven neural modeling. *Front Neurosci*. (2014) 8:383. doi: 10.3389/fnins.2014.00383
44. Cao Y, Ren K, Su F, Deng B, Wei X, Wang J. Suppression of seizures based on the multi-coupled neural mass model. *Chaos* (2015) 25:103120. doi: 10.1063/1.4931715
45. Kuhlmann L, Freestone DR, Manton JH, Heyse B, Vereecke HEM, Lipping T, et al. Neural mass model-based tracking of anesthetic brain states. *Neuroimage* (2016) 133:438–56. doi: 10.1016/j.neuroimage.2016.03.039
46. Zhang Z. A fast method to compute surface potentials generated by dipoles within multilayer anisotropic spheres. *Phys Med Biol*. (1995) 40:335–49. doi: 10.1088/0031-9155/40/3/001
47. Mosher JC, Leahy RM, Lewis PS. EEG and MEG: forward solutions for inverse methods. *IEEE Trans Biomed Eng*. (1999) 46:245–59. doi: 10.1109/10.748978
48. Haufe S, Tomioka R, Dickhaus T, Sannelli C, Blankertz B, Nolte G, et al. Large-scale EEG/MEG source localization with spatial flexibility. *Neuroimage* (2011) 54:851–9. doi: 10.1016/j.neuroimage.2010.09.003
49. Verhellen E, Boon P. EEG source localization of the epileptogenic focus in patients with refractory temporal lobe epilepsy, dipole modelling revisited. *Acta Neurol Belg*. (2007) 107:71–7. Available online at: <http://www.actaneurologica.be/pdf.php?annee=2007&trimestre=3>
50. Gotman J. Noninvasive methods for evaluating the localization and propagation of epileptic activity. *Epilepsia*. (2003) 44(Suppl 1):21–9. doi: 10.1111/j.0013-9580.2003.12003.x
51. Jansen BH, Zouridakis G, Brandt ME. A neurophysiologically-based mathematical model of flash visual evoked potentials. *Biol Cybern*. (1993) 68:275–83. doi: 10.1007/BF00224863
52. Merwe RVD, Wan EA. The square-root unscented Kalman filter for state and parameter-estimation. in: *Proceedings of (ICASSP '01) 2001 IEEE International Conference on Acoustics, Speech, and Signal Processing*, Vol. 6. Salt Lake City (2001). p. 3461–4.
53. Julier SJ, Uhlmann JK. Unscented filtering and nonlinear estimation. *Proc IEEE* (2004) 92:401–22. doi: 10.1109/JPROC.2003.823141
54. Kalman RE. A new approach to linear filtering and prediction problems. *Trans ASME J Basic Eng*. (1960) 82:35–45. doi: 10.1115/1.3662552
55. Silva FHL, Hoeks A, Smits H, Zetterberg LH. Model of brain rhythmic activity. *Kybernetik* (1974) 15:27–37. doi: 10.1007/BF00270757
56. Faugeras O, Touboul J, Cessac B. A constructive mean-field analysis of multi-population neural networks with random synaptic weights and stochastic inputs. *Front Comput Neurosci*. (2009) 3:1. doi: 10.3389/neuro.10.001.2009
57. Grimbert F, Faugeras O. *Analysis of Jansen's Model of a Single Cortical Column*. Research Report INRIA (2006).
58. Pons AJ, Cantero JL, Atienza M, Garcia-Ojalvo J. Relating structural and functional anomalous connectivity in the aging brain via neural mass modeling. *Neuroimage* (2010) 52:848–61. doi: 10.1016/j.neuroimage.2009.12.105
59. Silva FLD. 5. In: *EEG: Origin and Measurement*. World Scientific Publishing Co. (2011). p. 63–82. Available online at: http://www.worldscientific.com/doi/abs/10.1142/9789814525350_0005
60. Ary JP, Klein SA, Fender DH. Location of sources of evoked scalp potentials: corrections for skull and scalp thicknesses. *IEEE Trans Biomed Eng* (1981) 28:447–52. doi: 10.1109/TBME.1981.324817
61. Jurcak V, Tsuzuki D, Dan I. 10/20, 10/10, and 10/5 systems revisited: their validity as relative head-surface-based positioning systems. *Neuroimage* (2007) 34:1600–11. doi: 10.1016/j.neuroimage.2006.09.024
62. Berg P, Scherg M. A fast method for forward computation of multiple-shell spherical head models. *Electroencephalogr Clin Neurophysiol*. (1994) 90:58–64. doi: 10.1016/0013-4694(94)90113-9
63. Merwe RVD, Wan EA. The unscented kalman filter for nonlinear estimation. In: *Adaptive Systems for Signal Processing, Communications, and Control Symposium 2000, AS-SPCC*. Lake Louise, CA: IEEE (2000). p. 153–8.
64. Solonen A, Hakkarainen J, Ilin A, Abbas M, Bibov A. Estimating model error covariance matrix parameters in extended Kalman filtering. *Nonlinear Process Geophys*. (2014) 21:919–27. doi: 10.5194/npg-21-919-2014
65. Cantero JL, Atienza M, Gomez-Herrero G, Cruz-Vadell A, Gil-Neciga E, Rodriguez-Romero R, et al. Functional integrity of thalamocortical circuits differentiates normal aging from mild cognitive impairment. *Hum Brain Mapp*. (2009) 30:3944–57. doi: 10.1002/hbm.20819
66. Richards JE. Recovering dipole sources from scalp-recorded event-related-potentials using component analysis: principal component analysis and independent component analysis. *Int J Psychophysiol*. (2004) 54:201–20. doi: 10.1016/j.ijpsycho.2004.03.009
67. Crispin W Gardiner. *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. Berlin: Springer (2004).
68. Garcia-Ojalvo J, Sancho J. *Noise in Spatially Extended Systems*, 1999th Edn. New York, NY: Institute for Nonlinear Science, Springer (1999).
69. Easycap. *Easycap Equidistant Layouts*. (2018). Available online at: <https://www.easycap.de/wordpress/wp-content/uploads/2018/02/Easycap-Equidistant-Layouts.pdf>
70. Toral R, Colet P. *Stochastic Numerical Methods: An Introduction for Students and Scientists*. John Wiley & Sons (2014).
71. Liu X, Gao Q. Parameter estimation and control for a neural mass model based on the unscented Kalman filter. *Phys Rev E*. (2013) 88:042905. doi: 10.1103/PhysRevE.88.042905
72. Schiff SJ. Neural control engineering: the emerging intersection between control theory and neuroscience. In: *Computational Neuroscience*. MIT Press (2012). Available online at: <https://books.google.es/books?id=P9UvTQtnqKwC>
73. Moya MJ, Diekmann CO. Data assimilation methods for neuronal state and parameter estimation. *J Math Neurosci*. (2018) 8:11. doi: 10.1186/s13408-018-0066-8

Conflict of Interest Statement: 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.

Copyright © 2018 Escuain-Poole, Garcia-Ojalvo and Pons. 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.

APPENDIX: THE UNSCENTED KALMAN FILTER (UKF) ALGORITHM

UKF is a predictor-corrector algorithm that estimates the state and parameters at a given time step k in two phases. The first one predicts the state based solely on the dynamical information of the system, i.e., the model. The second incorporates a measurement with which to correct the first estimation. **Table A1** presents the symbols used in this paper for the variables of the Kalman filter.

The first step of the algorithm involves computing the expectation of the state and of the state covariance at time instant $k + 1$, known as the *a priori* estimation. For this we use a numerical implementation (using Heun's solver) of Jansen and Rit's model of a cortical column [26, 51], as described in the section 2.2.

The nature of the nonlinearities of this model prevents us from using a simple linearisation approach to propagating the statistics of the state variables across the transformation, as would be the case if we used the extended Kalman filter, for example. Therefore, we incorporate the unscented transform (UT) in our formulation of the Kalman filter, which, instead of attempting to propagate a distribution through the nonlinearity, first propagates a series of deterministically chosen points through the nonlinearity and then recovers the statistical information of the distribution from these.

TABLE A1 | Variables of the Unscented Kalman Filter.

Parameter	Description
\hat{x}	State estimate
\hat{x}^-	<i>a priori</i> state estimate
Σ	Sigma points
X^*	Transformed sigma points
X	Redrawn sigma points [52]
Υ	Sigma points projected to measurement space
y^-	Estimated measurement
z	Measurement
f	Model of the system
H	Observation model
n	State size
λ	Scaling parameter
α	Primary scaling factor
β	Secondary scaling factor
κ	Tertiary scaling factor
W^m	Weight vector for the mean
W^{cov}	Weight vector for the covariance
P	State covariance estimate
P^-	<i>a priori</i> state covariance estimate
P_{yy}	Predicted measurement covariance
P_{xy}	State-measurement cross-covariance
Q	State error covariance*
R	Measurement error covariance*
K	Kalman gain

Therefore, the *a priori* estimation of the state, \hat{x}_k^- , is obtained as follows, beginning with the calculation and projection of the $2n + 1$ (where n is the state size) sigma points,

$$\begin{aligned}\Sigma_{k-1,0} &= \hat{x}_{k-1} \\ \Sigma_{k-1,i} &= \hat{x}_{k-1} + \left(\sqrt{(n+\lambda)P_{k-1}} \right)_i, \quad i = 1, \dots, n \\ \Sigma_{k-1,i} &= \hat{x}_{k-1} - \left(\sqrt{(n+\lambda)P_{k-1}} \right)_{i-n}, \quad i = n+1, \dots, 2n\end{aligned}\quad (A1)$$

where P_{k-1} is the estimated state covariance matrix for the previous time step. The square root of this matrix is well-defined, and can be calculated efficiently via a Cholesky decomposition [52]. This continues with the condensation of the projected sigma points into the *a priori* state estimate:

$$X_{k|k-1}^* = f(\Sigma_{k-1}) \quad (A2)$$

$$\hat{x}_k^- = \sum_{i=0}^{2L} W_i^m X_{i,k|k-1}^* \quad (A3)$$

$$P_k^- = \sum_{i=0}^{2L} W_i^{cov} [X_{i,k|k-1}^* - \hat{x}_k^-] [X_{i,k|k-1}^* - \hat{x}_k^-]^T + Q \quad (A4)$$

where Q is the state error covariance and W^m and W^{cov} are the weight vectors, defined as

$$\begin{aligned}W_0^m &= \frac{\lambda}{n+\lambda} \\ W_0^{cov} &= \frac{\lambda}{n+\lambda} + 1 - \alpha^2 + \beta \\ W_i^m = W_i^{cov} &= \frac{1}{2(n+\lambda)}, i = 1, \dots, 2n\end{aligned}\quad (A5)$$

In Equations A1 and A5, α , β and κ are scaling factors, and λ , which is crucial to guarantee a positive semi-definite covariance matrix P , is calculated as $\lambda = \alpha^2(n + \kappa) - n$. The primary scaling factor α determines the spread of the sigma points around the mean and is set at 0.001, it being usually set between 0.001 and 1 [63] and chosen according to the quality of the resulting estimation. The secondary scaling factor β contains prior information about the distribution of x ; for Gaussian distributions, its optimal value is 2. Finally, κ , the tertiary scaling parameter, is set to 0, as is a usual practice [63].

We now use a measurement to correct the state estimation, which implies the mapping of the *a priori* estimate onto the measurement space for comparison. In our case, this transformation is a linear matrix H that relates the state of the cortical columns to an EEG reading (see section 2.2 for details). The sigma points $\Sigma_{k|k-1}$ are projected into the measurement space [52]

$$\Upsilon_{k|k-1} = H[\Sigma_{k|k-1}], \quad (A6)$$

from which the estimation of the measurement, \hat{y}_k^- , is calculated:

$$\hat{y}_k^- = \sum_{i=0}^{2L} W_i^m \Upsilon_{i,k|k-1} \quad (\text{A7})$$

The second step of the algorithm corrects the *a priori* estimation of state and covariance by using the information available from the most recent measurement (in our case, an EEG reading). The impact of the measurement is determined by the Kalman gain K_k , which essentially expresses the level of confidence on the accuracy of the model and the level of noise in the data.

$$P_{y_k y_k} = \sum_{i=0}^{2L} W_i^{cov} [\Upsilon_{i,k|k-1} - \hat{y}_k^-][\Upsilon_{i,k|k-1} - \hat{y}_k^-]^T + R \quad (\text{A8})$$

$$P_{x_k y_k} = \sum_{i=0}^{2L} W_i^{cov} [X_{i,k|k-1} - \hat{x}_k^-][\Upsilon_{i,k|k-1} - \hat{y}_k^-]^T \quad (\text{A9})$$

$$K_k = P_{x_k y_k} P_{y_k y_k}^{-1} \quad (\text{A10})$$

$$\hat{x}_k = \hat{x}_k^- + K_k (z_k - \hat{y}_k^-) \quad (\text{A11})$$

$$P_k = P_k^- - K_k P_{y_k y_k} K_k^T \quad (\text{A12})$$

where $P_{y_k y_k}$ is the predicted measurement covariance, $P_{x_k y_k}$ is the state-measurement cross-covariance, R is the measurement error covariance, and z_k is the measurement for the current time step.



Statistical Data Assimilation: Formulation and Examples From Neurobiology

Anna Miller^{1*}, Dawei Li¹, Jason Platt¹, Arij Daou², Daniel Margoliash³ and Henry D. I. Abarbanel^{1,4}

¹ Department of Physics, University of California, San Diego, La Jolla, CA, United States, ² Biomedical Engineering Program, American University of Beirut, Beirut, Lebanon, ³ Department of Anatomy and Organismal Biology, University of Chicago, Chicago, IL, United States, ⁴ Marine Physical Laboratory (Scripps Institution of Oceanography), Department of Physics, University of California, San Diego, La Jolla, CA, United States

OPEN ACCESS

Edited by:

Axel Hutt,
German Meteorological Service,
Germany

Reviewed by:

Meysam Hashemi,
INSERM U1106 Institut de
Neurosciences des Systèmes, France
Lili Lei,
Nanjing University, China

*Correspondence:

Anna Miller
a8miller@ucsd.edu

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 10 September 2018

Accepted: 02 November 2018

Published: 26 November 2018

Citation:

Miller A, Li D, Platt J, Daou A,
Margoliash D and Abarbanel HDI
(2018) Statistical Data Assimilation:
Formulation and Examples From
Neurobiology.
Front. Appl. Math. Stat. 4:53.
doi: 10.3389/fams.2018.00053

For the Research Topic *Data Assimilation and Control: Theory and Applications in Life Sciences* we first review the formulation of statistical data assimilation (SDA) and discuss algorithms for exploring variational approximations to the conditional expected values of biophysical aspects of functional neural circuits. Then we report on the application of SDA to (1) the exploration of properties of individual neurons in the HVC nucleus of the avian song system, and (2) characterizing individual neurons formulated as very large scale integration (VLSI) analog circuits with a goal of building functional, biophysically realistic, VLSI representations of functional nervous systems. Networks of neurons pose a substantially greater challenge, and we comment on formulating experiments to probe the properties, especially the functional connectivity, in song command circuits within HVC.

Keywords: data assimilation, neuronal dynamics, HVC, ion channel properties, variational annealing, neuromorphic, VLSI

1. INTRODUCTION

A broad class of “inverse” problems presents itself in many scientific and engineering inquiries. The overall question addressed by these is how to transfer information from laboratory and field observations to candidate models of the processes underlying those observations.

The existence of large, information rich, well curated data sets from increasingly sophisticated observations on complicated nonlinear systems has set new challenges to the information transfer task. Assisting with this challenge are new substantial computational capabilities.

Together they have provided an arena in which principled formulation of this information transfer along with algorithms to effect the transfer have come to play an essential role. This paper reports on some efforts to meet this class of challenge within neuroscience. Many of the ideas are applicable much more broadly than our focus, and we hope the reader will find this helpful in their own inquiries.

In this special issue entitled *Data Assimilation and Control: Theory and Applications in Life Sciences*, of the journal *Frontiers in Applied Mathematics and Statistics–Dynamical Systems*, we participate in the broader quantitative setting for this information transfer. The procedures are called “data assimilation” following its use in the effort to develop realistic numerical weather prediction models [1, 2] over many decades. We prefer the term “statistical data assimilation” (SDA) to emphasize that key ingredients in the procedures involved in the transfer rest on noisy

data and on recognizing errors in the models to which information in the noisy data is to be transferred.

This article begins with a formulation of SDA with some additional clarity beyond the discussion in Abarbanel [3]. We also discuss some algorithms helpful for implementing the information transfer, testing model compatibility with the available data, and quantitatively identifying how much information in the data can be represented in the model selected by the SDA user. Using SDA will also remind us that data assimilation efforts are well cast as problems in statistical physics [4].

After the discussion of SDA, we turn to some working ideas on how to perform the high dimensional integrals involved in SDA. In particular we address the “standard model” of SDA where data is contaminated by Gaussian noise and model errors are represented by Gaussian noise, though the integrals to be performed are, of course, not Gaussian. The topics include the approximation of Laplace [5] and Monte Carlo methods.

With these tools in hand, we turn to neurobiological questions that arise in the analysis of individual neurons and, in planning, for network components of the avian song production pathway. These questions are nicely formulated in the general framework, and we dwell on specifics of SDA in a realistic biological context. The penultimate topic we address is the use of SDA to calibrate VLSI analog chips designed and built as components of a developing instantiation of the full songbird song command network, called HVC. Lastly, we discuss the potential utilization of SDA for exploring biological networks.

At the outset of this article we may expect that our readers from Physics and Applied Mathematics along with our readers from Neurobiology may find the conjunction of the two “strange bedfellows” to be incongruous. For the opportunity to illuminate the natural melding of the facets of both kinds of questions, we thank the editors of this special issue.

2. MATERIALS AND METHODS

2.1. General Overview of Data Assimilation

We will provide a structure within which we will frame our discussion of transfer of information from data to a model of the underlying processes producing the data.

We start with an observation window in time $[t_0, t_F]$ within which we make a set of measurements at times $t = \{\tau_1, \tau_2, \dots, \tau_k, \dots, \tau_F\}$; $t_0 \leq \tau_k \leq t_F$. At each of these measurement times, we observe L quantities $\mathbf{y}(\tau_k) = \{y_1(\tau_k), y_2(\tau_k), \dots, y_L(\tau_k)\}$. The number L of observations at each measurement time τ_k is typically less, often much less, than the number of degrees of freedom D in the observed system; $D \gg L$.

These are views into the dynamical processes of a system we wish to characterize. The quantitative characterization is through a model we choose. It describes the interactions among the states of the observed system. If we are observing the time course of a neuron, for example, we might measure the membrane voltage $y_1(\tau_k) = V_m(\tau_k)$ and the intracellular Ca^{2+} concentration $y_2(\tau_k) = [\text{Ca}^{2+}](\tau_k)$. From these data we want to estimate the unmeasured states of the model as a function of time as well as estimate biophysical parameters in the model.

The processes characterizing the state of the system (neuron) we call $x_a(t)$; $a = 1, 2, \dots, D \geq L$, and they are selected by the user to describe the dynamical behavior of the observations through a set of equations in continuous time

$$\frac{dx_a(t)}{dt} = F_a(\mathbf{x}(t), \mathbf{q}), \quad (1)$$

or in discrete time $t_n = t_0 + n\Delta t$; $n = 0, 1, \dots, N$; $t_N = t_F$ via

$$x_a(t_{n+1}) = x_a(n+1) = f_a(\mathbf{x}(t_n), \mathbf{q}) = f_a(\mathbf{x}(n), \mathbf{q}), \quad (2)$$

where \mathbf{q} is a set of fixed parameters associated with the model. $\mathbf{f}(\mathbf{x}(n), \mathbf{q})$ is related to $\mathbf{F}(\mathbf{x}(t), \mathbf{q})$ through the choice the user makes for solving the continuous time flow for $\mathbf{x}(t)$ through a numerical solution method of choice [6].

Considering neuronal activity, Equation 1 could be coupled Hodgkin-Huxley (HH) Equations [7, 8] for voltage, ion channel gating variables, constituent concentrations, and other ingredients. If the neuron is isolated *in vitro*, such as by using drugs to block synaptic transmission, then there would be no synaptic input to the cell to describe. While if it is coupled to a network of neurons, their functional connectivity would be described in $\mathbf{F}(\mathbf{x}(t), \mathbf{q})$ or $\mathbf{f}(\mathbf{x}(n), \mathbf{q})$. Typical parameters might be maximal conductances of the ion channels, reversal potentials, and other time-independent numbers describing the kinetics of the gating variables. In many experiments L is only 1, namely, the voltage across the cell membrane, while D may be on the order of 100; Hence $D \gg L$.

As we proceed from the initiation of the observation window at t_0 we must move our model equation variables $\mathbf{x}(0)$, Equation 2, from t_0 to τ_1 where a measurement is made. Then using the model dynamics we move along to τ_2 and so forth until we reach the last measurement time τ_F and finally move the model from $\mathbf{x}(\tau_F)$ to $\mathbf{x}(t_F)$. In each stepping of the model equations (Equation 2) we may make many steps of Δt in time to achieve accuracy in the representation of the model dynamics. The full set of times t_n at which we evaluate the model $\mathbf{x}(t_n)$ we collect into the path of the state of the model through D -dimensional space: $\mathbf{X} = \{\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(n), \dots, \mathbf{x}(N) = \mathbf{x}(F)\}$. The dimension of the path is $(N+1)D + Nq$, where Nq is the number of parameters \mathbf{q} in our model.

It is worth a pause here to note that we have now collected two of the needed three ingredients to effect our transfer of the information in the collection of all measurements $\mathbf{Y} = \{\mathbf{y}(\tau_1), \mathbf{y}(\tau_2), \dots, \mathbf{y}(\tau_F)\}$ to the model $\mathbf{f}(\mathbf{x}(n), \mathbf{q})$ along the path \mathbf{X} through the observation window $[t_0, t_F]$: (1) data \mathbf{Y} and (2) a model of the processes in \mathbf{Y} , devised by our experience and knowledge of those processes. The notation and a visual presentation of this is found in **Figure 1**.

The **third** ingredient, comprised of methods to generate the transfer from \mathbf{Y} to properties of the model, will command our attention throughout most of this paper. If the transfer methods are successful and, according to some metric of success, we arrange matters so that at the measurement times τ_k , the L model variables $\mathbf{x}(t)$ associated with $\mathbf{y}(\tau_k)$ are such that $x_l(\tau_k) \approx y_l(\tau_k)$, we are *not* finished. We have then only demonstrated that the model is consistent with the known data \mathbf{Y} . We must use the

Data Assimilation in a time window $[t_0, t_F]$:

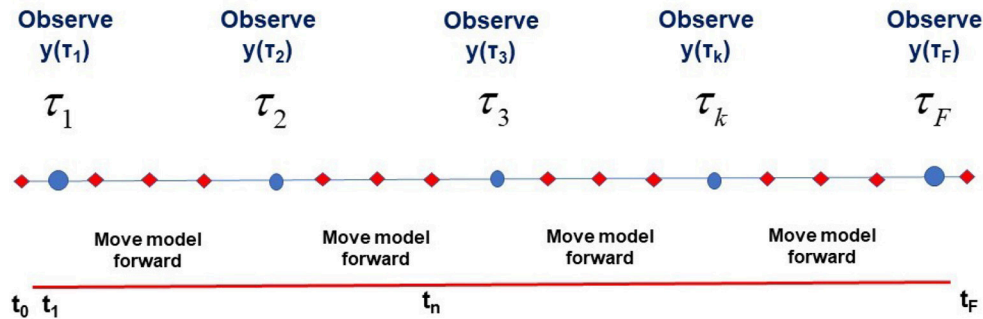


FIGURE 1 | A visual representation of the window $t_0 \leq t \leq t_F$ in time during which L -dimensional observations $y(\tau_k)$ are performed at observation times $t = \tau_k$; $k = 1, \dots, F$. This also shows times at which the D -dimensional model developed by the user $\mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{q})$ is used to move forward from time n to time $n+1$: $t_n = t_0 + n\Delta t$; $n = 0, 1, \dots, N$. $D \geq L$. The path of the model $\mathbf{X} = \{\mathbf{x}(0), \mathbf{x}(1), \dots, \mathbf{x}(n), \dots, \mathbf{x}(N) = \mathbf{x}(F)\}$ and the collection \mathbf{Y} of L -dimensional observations at each observation time τ_k , $\mathbf{Y} = \{y(\tau_1), y(\tau_2), \dots, y(\tau_k), \dots, y(\tau_F)\}$ ($\mathbf{y} = \{y_1, y_2, \dots, y_L\}$) is also indicated.

model, completed by the estimates of \mathbf{q} and the state of the model at t_F , $\mathbf{x}(t_F)$, to predict forward for $t > t_F$, and we should succeed in comparison with measurements for $y(\tau_r)$ for $\tau_r > t_F$. As the measure of success of predictions, we may use the same metric as utilized in the observation window.

As a small aside, the same overall setup applies to supervised machine learning networks [9] where the observation window is called the training set; the prediction window is called the test set, and prediction is called generalization.

2.1.1. The Data Are Noisy; The Model Has Errors

Inevitably, the data we collect is noisy, and equally the model we select to describe the production of those data has errors. This means we must, at the outset, address a conditional probability distribution $P(\mathbf{X}|\mathbf{Y})$ as our goal in the data assimilation transfer from \mathbf{Y} to the model. In Abarbanel [3] we describe how to use the Markov nature of the model $\mathbf{x}(n) \rightarrow \mathbf{x}(n+1) = \mathbf{f}(\mathbf{x}(n), \mathbf{q})$ and the definition of conditional probabilities to derive the recursion relation:

$$P(\mathbf{X}(n+1)|\mathbf{Y}(n+1)) = \frac{P(\mathbf{y}(n+1), \mathbf{x}(n+1), \mathbf{X}(n)|\mathbf{Y}(n))}{P(\mathbf{y}(n+1)|\mathbf{Y}(n))P(\mathbf{x}(n+1), \mathbf{X}(n)|\mathbf{Y}(n))} \cdot \frac{P(\mathbf{x}(n+1)|\mathbf{x}(n))P(\mathbf{X}(n)|\mathbf{Y}(n))}{\exp[CMI(\mathbf{y}(n+1), \mathbf{x}(n+1), \mathbf{X}(n)|\mathbf{Y}(n))]} \cdot \frac{P(\mathbf{x}(n+1)|\mathbf{x}(n))P(\mathbf{X}(n)|\mathbf{Y}(n))}{P(\mathbf{x}(n+1)|\mathbf{x}(n))P(\mathbf{X}(n)|\mathbf{Y}(n))}, \quad (3)$$

where we have identified $CMI(a, b|c) = \log \left[\frac{P(a, b|c)}{P(a|c)P(b|c)} \right]$. This is Shannon's conditional mutual information [10] telling us how many bits (for \log_2) we know about a when observing b conditioned on c . For us $a = \{\mathbf{y}(n+1)\}$, $b = \{\mathbf{x}(n+1), \mathbf{X}(n)\}$, $c = \{\mathbf{Y}(n)\}$. We can simplify this further with the assumption that an observation at any time depends only on the state of the system.

$$P(\mathbf{X}(n+1)|\mathbf{Y}(n+1)) = P(\mathbf{y}(n+1)|\mathbf{X}(n+1))P(\mathbf{x}(n+1)|\mathbf{x}(n))P(\mathbf{X}(n)|\mathbf{Y}(n)) \quad (4)$$

Using this recursion relation to move backwards through the observation window from $t_F = t_0 + N\Delta t$ through the

measurements at times τ_k to the start of the window at t_0 , we may write, up to factors independent of \mathbf{X}

$$P(\mathbf{X}|\mathbf{Y}) = \left\{ \prod_{k=1}^F P(\mathbf{y}(\tau_k)|\mathbf{X}(\tau_k)) \prod_{n=0}^{N-1} P(\mathbf{x}(n+1)|\mathbf{x}(n)) \right\} P(\mathbf{x}(0)). \quad (5)$$

If we now write $P(\mathbf{X}|\mathbf{Y}) \propto \exp[-A(\mathbf{X})]$ where $A(\mathbf{X})$, the negative of the log likelihood, we call the action, then conditional expected values for functions along the path \mathbf{X} are defined by

$$E[G(\mathbf{X})|\mathbf{Y}] = \langle G(\mathbf{X}) \rangle = \frac{\int d\mathbf{X} G(\mathbf{X}) e^{-A(\mathbf{X})}}{\int d\mathbf{X} e^{-A(\mathbf{X})}}, \quad (6)$$

$d\mathbf{X} = \prod_{n=0}^N d^D \mathbf{x}(n)$, and all factors in the action independent of \mathbf{X} cancel out here. The action takes the convenient expression

$$A(\mathbf{X}) = - \left\{ \sum_{k=1}^F \log[P(\mathbf{y}(\tau_k)|\mathbf{X}(\tau_k))] + \sum_{n=0}^{N-1} \log[P(\mathbf{x}(n+1)|\mathbf{x}(n))] \right\} - \log[P(\mathbf{x}(0))], \quad (7)$$

which is the sum of the terms which modify the conditional probability distribution when an observation is made at $t = \tau_k$ and the sum of the stochastic version of $\mathbf{x}(n) \rightarrow \mathbf{x}(n+1) - \mathbf{f}(\mathbf{x}(n), \mathbf{q})$ and finally the distribution when the observation window opens at t_0 .

What quantities $G(\mathbf{X})$ are of interest? One natural one is the path $G(\mathbf{X}) = \mathbf{X}_\mu$; $\mu = \{a, n\}$ itself; another is the covariance around that mean $\langle \mathbf{X}_\mu \rangle = \bar{\mathbf{X}}_\mu = \langle \mathbf{X}_\mu \rangle : \langle (\mathbf{X}_\mu - \bar{\mathbf{X}}_\mu)(\mathbf{X}_\nu - \bar{\mathbf{X}}_\nu) \rangle$. Other moments are of interest, of course. If one has an anticipated form for the distribution at large \mathbf{X} , then $G(\mathbf{X})$ may be chosen as a parametrized version of that form and those parameters determined near the maximum of $P(\mathbf{X}|\mathbf{Y})$.

The action simplifies to what we call the "standard model" of data assimilation when (1) observations \mathbf{y} are related to their

model counterparts via Gaussian noise with zero mean and diagonal precision matrix \mathbf{R}_m , and (2) model errors are associated with Gaussian errors of mean zero and diagonal precision matrix \mathbf{R}_f :

$$A(\mathbf{X}) = \sum_{k=1}^F \sum_{l=1}^L \frac{R_m(k)}{2} (x_l(\tau_k) - y_l(\tau_k))^2 + \sum_{n=0}^{N-1} \sum_{a=1}^D \frac{R_f(a)}{2} (x_a(n+1) - f_a(\mathbf{x}(n), \mathbf{q}))^2. \quad (8)$$

If we have knowledge of the distribution $P(\mathbf{x}(0))$ at t_0 we may add it to this action. If we have no knowledge of $P(\mathbf{x}(0))$, we may take its distribution to be uniform over the dynamic range of the model variables, then it, as here, is absent, canceling numerator and denominator in Equation (6).

Our challenge is to perform integrals such as Equation (6). One should anticipate that the dominant contribution to the expected value comes from the maxima of $P(\mathbf{X}|\mathbf{Y})$ or, equivalently the minima of $A(\mathbf{X})$. At such minima, the two contributions to the action, the measurement error and the model error, balance each other to accomplish the explicit transfer of information from the former to the latter.

We note, as before, that when $\mathbf{f}(\mathbf{x}(n), \mathbf{q})$ is nonlinear in \mathbf{X} , as it always is in interesting examples, the expected value integral Equation (6) is not Gaussian. So, some thinking is in order to approximate this high dimensional integral. We turn to that now. After consideration of methods to do the integral, we will return to a variety of examples taken from neuroscience.

The two generally useful methods available for evaluating this kind of high dimensional integral are Laplace's method [5] and Monte Carlo techniques [6, 11, 12]. We address them in order. We also add our own new and useful versions of the methods.

2.1.2. Laplace's Method

To locate the minima of the action $A(\mathbf{X}) = -\log[P(\mathbf{X}|\mathbf{Y})]$ we must seek paths $\mathbf{X}^{(j)}$; $j = 0, 1, \dots$ such that $\partial A(\mathbf{X})/\partial \mathbf{X}|_{\mathbf{X}^{(j)}} = 0$, and then check that the second derivative at $\mathbf{X}^{(j)}$, the Hessian, is a positive definite matrix in path coordinates. The vanishing of the derivative is a necessary condition.

Laplace's method [5] expands the action around the $\mathbf{X}^{(j)}$ seeking the path $\mathbf{X}^{(0)}$ with the smallest value of $A(\mathbf{X})$. The contribution of $\mathbf{X}^{(0)}$ to the integral Equation (6) is approximately $\exp[A(\mathbf{X}^{(1)}) - A(\mathbf{X}^{(0)})]$ bigger than that of the path with the next smallest action.

This sounds more or less straightforward; however, finding the global minimum of a nonlinear function such as $A(\mathbf{X})$ is an NP-complete problem [13]. In a practical sense one cannot expect to succeed with such problems. However there is an attractive feature of the form of $A(\mathbf{X})$ that permits us to accomplish more.

We now discuss two algorithmic approaches to implementing Laplace's method.

2.1.3. Precision Annealing for Laplace's Method

Looking at Equation (8) we see that if the precision of the model is zero, $R_f = 0$, the action is quadratic in the L measured variables $x_l(n)$ and independent of the remaining states. The global minimum of such an action comes with $x_l(\tau_k) = y_l(\tau_k)$

and any choice for the remaining states and parameters. Choose the path with these values of $\mathbf{x}(\tau_k)$ and values from a uniform distribution of the other state variables and parameters covering the expected dynamic range of those, and call it path \mathbf{X}_{init} . In practice, we recognize that the global minimum of $A(\mathbf{X})$ is degenerate at $R_f = 0$, so we select many initial paths. We choose N_I of them, and initialize whatever numerical optimization program we have selected, to run on each of them. We continue to call the collection of N_I paths \mathbf{X}_{init} .

- Now we increase R_f from $R_f = 0$ to a small value R_{f0} . Use each of the N_I paths in \mathbf{X}_{init} as initial conditions for our numerical optimization program chosen to find the minima of $A(\mathbf{X})$, and we arrive at N_I paths \mathbf{X}_0 . Evaluate $A(\mathbf{X}_0)$ on all N_I paths \mathbf{X}_0 .
- We increase $R_f = R_{f0} \rightarrow R_{f0}\alpha$; $\alpha > 1$, and now use the N_I paths \mathbf{X}_0 as the initial conditions for our numerical optimization program chosen to find the minima of $A(\mathbf{X})$, we arrive at N_I paths \mathbf{X}_1 . Evaluate $A(\mathbf{X}_1)$ on all N_I paths \mathbf{X}_1 .
- We increase $R_f = R_{f0}\alpha \rightarrow R_{f0}\alpha^2$. Now use the N_I paths \mathbf{X}_1 as the initial conditions for our numerical optimization program chosen to find the minima of $A(\mathbf{X})$, we arrive at N_I paths \mathbf{X}_2 . Evaluate $A(\mathbf{X}_2)$ on all N_I paths \mathbf{X}_2 .
- Continue in this manner increasing R_f to $R_f = R_{f0}\alpha^\beta$; $\beta = 0, 1, \dots$, then using the selected numerical optimization program to arrive at N_I paths \mathbf{X}_β . Evaluate $A(\mathbf{X}_\beta)$ on all N_I paths \mathbf{X}_β .
- As a function of $\log_\alpha \left[\frac{R_f}{R_{f0}} \right]$ display all N_I values of $A(\mathbf{X}_\beta)$ vs. β for all $\beta = 0, 1, 2, \dots, \beta_{\text{max}}$.

We call this method *precision annealing* (PA) [14–17]. It slowly turns up the precision of the model collecting paths at each R_f that emerged from the degenerate global minimum at $R_f = 0$. In practice it is able to track N_I possible minima of $A(\mathbf{X})$ at each R_f . When not enough information is presented to the model, that is L is too small, there are many local minima at all R_f . This is a manifestation of the NP-completeness of the minimization of $A(\mathbf{X})$ problem. None of the minima may dominate the expected value integral of interest.

As L increases, and enough information is transmitted to the model, for large R_f one minimum appears to stand out as the global minimum, and the paths associated with that smallest minimum yields good predictions. We note that there are always paths, not just a single path, as we have a distribution of paths, N_I of which are sampled in the PA procedure, within a variation of size $1/\sqrt{R_m}$. A clear example of this is seen in Shirman [18] in a small, illustrative model.

In the even that the chosen model is inconsistent with the data, or there is too much noise in the model error term, a single minimum of the action will not appear for large R_f . As in the case of too few measurements, there will be multiple local minima. An example of this can be seen in Ye et al. [14].

2.1.4. "Nudging" Within Laplace's Method

In meteorology one approach to data assimilation is to add a term to the deterministic dynamics which move the state of a model toward the observations [19]

$$x_a(n+1) = f_a(\mathbf{x}(n), \mathbf{q}) + u(n)(y_l(n) - x_l(n))\delta_{al}, \quad (9)$$

where $u(n) > 0$ and vanishes except where a measurement is available. This is referred to as “nudging.” It appears in an *ad hoc*, but quite useful, manner.

Within the structure we have developed, one may see that the “nudging term” arises through the balance between the measurement error term and the model error term in the action. This is easy to see when we look at the continuous time version of the data assimilation standard model

$$A(\mathbf{x}(t), \dot{\mathbf{x}}(t)) = \int_{t_0}^{t_f} dt \left\{ \sum_{l=1}^L \frac{R_m(t, l)}{2} (x_l(t) - y_l(t))^2 + \sum_{a=1}^D \frac{R_f(a)}{2} (\dot{x}_a(t) - F_a(\mathbf{x}(t), \mathbf{q}))^2 \right\}. \quad (10)$$

The extremum of this action is given by the Euler-Lagrange equations for the variational problem [20]

$$\left[\delta_{ab} \frac{d}{dt} + \frac{\partial F_b(\mathbf{x}(t))}{\partial x_a(t)} \right] \left[\dot{x}_b(t) - F_b(\mathbf{x}(t)) \right] = \frac{R_m(a, t)}{R_f(a)} \delta_{al} (x_l(t) - y_l(t)), \quad (11)$$

in which the right hand side is the “nudging” term appearing in a natural manner. Approximating the operator $\delta_{ab} \frac{d}{dt} + \frac{\partial F_b(\mathbf{x}(t))}{\partial x_a(t)}$ we can rewrite this Euler-Lagrange equation in “nudging” form

$$\frac{dx_a(t)}{dt} = F_a(\mathbf{x}(t)) + u(t) \delta_{al} (x_l(t) - y_l(t)). \quad (12)$$

We will use both the full variation of the action, in discrete time, as well as its nudging form in our examples below.

2.1.5. Monte Carlo Methods

Monte Carlo methods [6, 11, 17, 21] are well covered in the literature. We have not used them in the examples in this paper. However, the development of a precision annealing version of Monte Carlo techniques promises to address the difficulties with large matrices for the Jacobian and Hessians required in variational principles (Wong et al., unpublished). When one comes to network problems, about which we comment later, this method may be essential.

3. RESULTS

3.1. Using SDA to Analyze the Avian Song System

We take our examples of the use of SDA in neurobiology from experiments on the avian song system. These have been performed in the University of Chicago laboratory of Daniel Margoliash, and we do not plan to describe in any detail the experiments nor the avian song production pathways in the avian brain. We give the essentials of the experiments and direct the reader to our references to develop the full biologically oriented idea why this system is enormously interesting.

Essentially, however, the manner in which songbirds learn and produce their functional vocalization—song—is an elegant

non-human example of a behavior that is cultural: the song is determined both by a genetic substrate and, interestingly, by refinement on top of that substrate by juveniles learning the song from their (male) elders. The analogs to learning speech in humans [22] are striking.

Our avian singer is a zebra finch. They, as do most other songbirds, learn vocal expression through auditory feedback [22–26]. This makes the study of the song system a good model for learning complex behavior [25, 27, 28]. Parts of the song system are analogous to the mammalian basal ganglia and regions of the frontal cortex [25, 29, 30]. Zebra finch in particular have the attractive property of singing only a single learned song, and with high precision, throughout their adult life.

Beyond the auditory pathways themselves, two neural pathways are principally responsible for song acquisition and production in zebra finch. The first is the Anterior Forebrain Pathway (AFP) which modulates learning. The second is a posterior pathway responsible for directing song production: the Song Motor Pathway (SMP) [24, 26, 31]. The HVC nucleus in the avian brain uniquely contributes to both of these [26].

There are two principal classes of projection neurons which extend from HVC: neurons which project to the robust nucleus of the arcopallium (HVC_{RA}), and neurons which project to Area X (HVC_X). HVC_{RA} neurons extend to the SMP pathway and HVC_X neurons extend to the AFP [26, 32]. These two classes of projection neurons combined with classes of HVC interneurons, make up the three broad classes of neurons within HVC. **Figure 2** [33] displays these structures in the avian brain.

In vitro observations of each HVC cell type have been obtained through patch-clamp techniques making intracellular voltage measurements in a reduced, brain slice preparation [23]. In this configuration, the electrode can simultaneously inject current into the neuron while measuring the whole cell voltage response [34]. From these data, one can establish the physical parameters of the system [23]. Traditionally this is done using neurochemicals to block selected ion channels and measuring the response properties of others [35]. Single current behavior is recorded and parameters are found through mathematical fits of the data. This procedure has its drawbacks, of course. There are various technical problems with the choice of channel blockers. Many of even the modern channel blockers are not subtype specific [36] and may only partially block channels [37]. A deeper conceptual problem is that is difficult to know what channels one may have missed altogether. Perhaps there are channels which express themselves only outside the bounds of the experimental conditions.

Our solution to such problems is the utilization of statistical data assimilation (SDA). This a method developed by meteorologists and others as computational models of increasingly large dynamical systems have been desired. Data assimilation has been described in our earlier sections.

In this paper, we focus on the song learning pathway, reporting on experiments involving the HVC_X neuron. The methods are generally applicable to the other neurons in HVC, and actually, to neurons seen as dynamical systems in general.

We start with a discussion about the neuron model. First we demonstrate the utility of our precision annealing methods

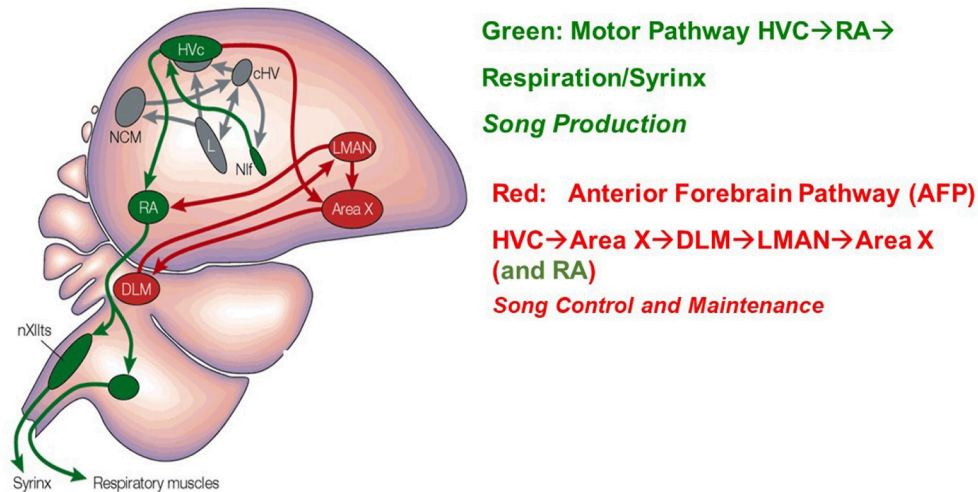


FIGURE 2 | A drawing of the Song Production Pathway and the Anterior Forebrain Pathway of avian songbirds. Parts of the auditory pathways are shown in gray. Pathways from the brainstem that ultimately return to HVC are not shown. The HVC network image is taken from Brainard and Doupe [33]. Reprinted by permission from Copyright Clearance Center: Springer Nature.

through the use of *twin experiments*. These are numerical experiments in which “data” is generated through a known model (of HVC_X), then analyzed via precision annealing. In a twin experiment, we know everything, so we can verify the SDA method by looking at predictions after a observation window in which the model is trained, and we may also compare the estimations of *unobserved state variables and parameters* to the ingredients and output of the model.

Twin experiments are meant to mirror the circumstances of the real experiment. We start by taking the model that we think describes our physical system. Initial points for the state variables and parameters are chosen at random from a uniform distribution within the state/parameter bounds, which are used along with the model to numerically integrate forward in time. This leaves us with complete information about the system. Noise is added to a subset of the state variables to emulate the data to be collected in a lab experiment. We then perform PA on these simulated data, as if they were real data. The results of these numerical experiments can be used to inform laboratory experiments, and indeed help design them, by identifying the necessary measurements and stimulus needed to accurately electrophysiologically characterize a neuron.

The second set of SDA analyses we report on using “nudging,” as described above, to estimate some key biophysical properties of HVC_X neurons from laboratory data. This SDA procedure is applied to HVC_X neurons in two different birds. The results show that though each bird is capable of normal vocalization, their complement of ion channel strengths is apparently different. We report on a suggestive example of this, leaving a full discussion to Daou and Margoliash (in review).

In order to obtain good estimation results, we must choose a forcing or stimulus with the model in mind: the dynamical range of the neuron must be thoroughly explored. This suggests a few key properties of the stimulus:

- The current waveform of $I_{injected}(t)$ must have sufficient amplitudes (\pm) and must be applied sufficiently long in time that it explores the full range of the neuron variation.
- The frequency content of the stimulus current must be a low enough that it does not exceed the low-pass cutoff frequency associated with the RC time constant of the neuron. This cutoff is typically in the neighborhood of 50–100 Hz.
- The current must explore all time scales expressed in the neuron’s behavior.

3.2. Analysis of HVC_X Data

The model for an HVC_X neuron is substantially taken from Daou et al. [23] and described in **Supplementary Data Sheet 1**. We now use this model in a “twin experiment” in which PA is utilized, and then using “nudging” we present the analysis of experimental data on two Zebra Finch.

3.2.1. Twin Experiment on HVC_X Neuron Model

A twin experiment is a synthetic numerical experiment meant to mirror the conditions of a laboratory experiment. We use our mathematical model with some informed parameter choices in order to generate numerical data. Noise is added to observable variables in the model, here $V(t)$. These data are then put through our SDA procedure to estimate parameters and unobserved states of the model. The neuron model is now calibrated or completed.

Using the parameters and the full state of the model at the end t_F of an observation window $[t_0, t_F]$, we take a current waveform $I_{injected}(t \geq t_F)$ to drive the model neuron and predict the time course of all dynamical variables in the prediction window $[t_F, \dots]$. This validation of the model is the critical test of our SDA procedure, here PA. In a laboratory experiment we have no specific knowledge of the parameters in the model and, by definition, cannot observe the unobserved state variables; here we can do that. So, “fitting” the observed data within the observation

window $[t_0, t_F]$ is not enough, we must reproduce all states for $t \geq t_F$ to test our SDA methods.

We use the model laid out in the **Supplementary Data Sheet 1**. We assume that the neuron has a resting potential of -70 mV and set the initial values for the voltage and each gating variable accordingly. We assume that the internal calcium concentration of the cell is $C_{in} = 0.1 \mu M$. We use an integration time step of 0.02 ms and integrate forward in time using an adaptive Runge-Kutta method [6]. Noise is added to the voltage time course by sampling from a Gaussian distribution $\mathcal{N}(0, 2)$ in units of mV.

The waveform of the injected current was chosen to have three key attributes: (1) It is strong enough to cause spiking in the neuron, (2) it dwells a long time in a hyperpolarizing region, and (3) its overall frequency content is low enough to not be filtered out by the neuron's RC time constant. On this last point, a neuron behaves like an RC circuit, it has a cut off frequency limited by the time constant of the system. Any input current which has a frequency higher than that of the cut off frequency won't be "seen" by the neuron. The time constant is taken to be the time it takes to spike and return back to 37% above its resting voltage. We chose a current where the majority of the power spectral density exists below 50 Hz.

The first two seconds of our chosen current waveform is a varying hyperpolarizing current. In order to characterize $I_h(t)$ and $I_{CaT}(t)$, it is necessary to thoroughly explore the region where the current is active. $I_h(t)$ is only activated when the neuron is hyperpolarized. The activation of $I_h(t)$ deinactivates $I_{CaT}(t)$, thereby allowing its dynamics to be explored. In order to characterize $I_{Na}(t)$ and $I_K(t)$, it is necessary to cause spiking in the neuron. The depolarizing current must be strong enough to hit the threshold potential for spike activation.

The parameters used to generate the data used in the twin experiment are in **Table 1**, and the injection current data and the membrane voltage response may be seen in **Figures 3A,B**.

The numbers chosen for the data assimilation procedure in this paper are $\alpha = 1.4$ and β ranging from 1 to 100. $R_{f,0,V} = 10^{-4}$ for voltage and $R_{f,0,j} = 1$ for all gating variables. These numbers are chosen because the voltage range is 100 times large than the gating variable range. Choosing a single $R_{f,0}$ value would result in the gating variable equations being less enforced than the voltage equation by a factor of 10^4 . The α and β numbers are chosen because we seek to make $\frac{R_f}{R_{f0}}$ sufficiently large. The α and β values chosen allow $\frac{R_f}{R_{f0}}$ to reach 10^{15} .

During estimation we instructed our methods to estimate the inverse capacitance and estimate the ratio $g' = \frac{g}{C_m}$ instead of g and C_m independently. That separation can present a challenge to numerical procedures. We also estimated the reversal potentials as a check on the SDA method.

Within our computational capability we can reasonably perform estimates on 50,000 data points. This captures a second of data when $\Delta t = 0.02$ ms. However, there are time constants in the model neuron which are on order 1 second. In order for us to estimate the behavior of these parameters accurately, we need to see multiple instances of the full response. We need a window on the order of 2–3 s. We can obtain this by downsampling the

TABLE 1 | Parameter values used to numerically generate the HVC_X data. The source of these values comes from Daou et al. [23]. Data was generated using an adaptive Runge-Kutta method, and can be seen in **Figures 3A,B**.

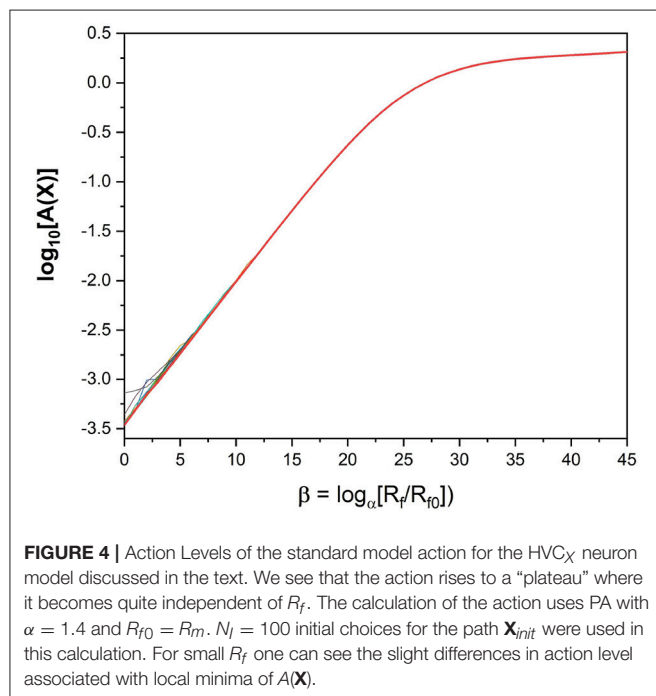
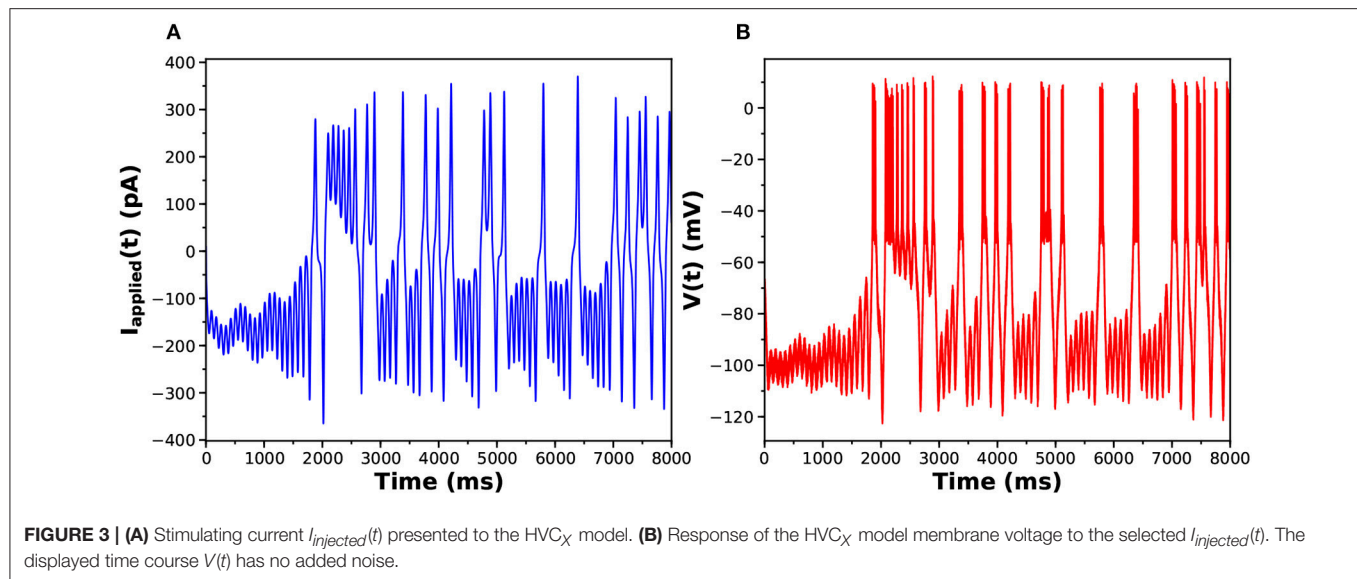
Parameter	Value	Parameter	Value	Parameter	Value
g_{Na}	450 nS	g_L	2 nS	k_f	0.3
E_{Na}	50 mV	E_L	-70 mV	θ_{mp}	-40 mV
g_K	50 nS	g_{Nap}	1 nS	σ_{mp}	-6 mV
E_K	-90 mV	g_{CaL}	19 nS	θ_s	-20 mV
g_{CaT}	2.65 nS	θ_m	-35 mV	σ_s	-0.05 mV
g_{SK}	6 nS	σ_m	-5 mV	θ_{hp}	-48 mV
g_H	4 nS	θ_n	-30 mV	σ_{hp}	6 mV
E_H	-30 mV	σ_n	-5 mV	$\bar{\tau}_{hp}$	1000 ms
C_m	100 pF	$\bar{\tau}_n$	10 ms	θ_e	-60 mV
θ_a	-20 mV	θ_{ff}	-105 mV	σ_e	5 mV
σ_a	-10 mV	σ_{ff}	5 mV	τ_e	20 ms
θ_{rs}	-105 mV	θ_{aR}	-65 mV	θ_b	0.4 mV
σ_{rs}	25 mV	σ_{aT}	-7.8 mV	σ_b	-0.1 mV
θ_{rT}	-67 mV	θ_{rrT}	68 mV	f	0.1
σ_{rT}	2 mV	σ_{rrT}	2.2 mV	ϵ	$0.0015 \frac{\mu M}{pA \cdot ms}$
τ_{r0}	200 ms	τ_{r1}	87.5 ms	p_{rf}	100
k_{Ca}	0.3 ms^{-1}	b_{Ca}	$0.1 \mu M$	k_s	$0.5 \mu M$

data. We know from previous results that downsampling can lead to better estimations [38]. We take every i^{th} data point, depending on the level of downsampling we want to do. In this data assimilation run, we downsampled by a factor of 4 to incorporate 4 s of data in the estimation window.

We look at a plot of the action as a function of β ; that is, $\log[R_f/R_{f0}]$. We expect to see a leveling off of the action [16] as a function of R_f . If the action becomes independent of R_f , we can then explore how well our parameter estimations perform when integrating them forward as predictions of the calibrated model. Looking at the action plot in **Figure 4**, we can see there is a region in which the action appears to level off, around $\beta = 40$. It is in this region where we look for our parameter estimates.

We examine all solutions around this region of β and utilize their parameter estimates in our predictions. We compare our numerical prediction to the "real" data from our synthetic experiment. We evaluate good predictions by finding the correlation coefficient between these two curves. This metric is chosen instead of a simple root mean square error because slight variations in spike timings yield a high amount of error even if the general spiking pattern is correct. The prediction plot and parameters for the best prediction can be seen in **Figure 5** and **Table 2**. The voltage trace in red is the estimated voltage after data assimilation is completed. It is overlaid on the synthetic input data in black. The blue time course is a prediction, starting at the last time point of the red estimated $V(t)$ trace and using the parameter estimates for $t \leq 4,000$ ms.

The red curve matches the computed voltage trace quite well. There is no significant deviation in the frequency of spikes, spike amplitudes, or the hyperpolarized region of the cell. Looking at the prediction window, we can see that there is some deviation



in the hyperpolarized voltage trace after 7,000 ms. Our predicted voltage does not become nearly as hyperpolarized as the real data. This is an indication that our parameter estimates for currents activated in this region are not entirely correct. Comparing parameters in **Table 2**, we can see that E_h is estimated as lower than its actual value. Despite that, we still are able to reproduce neuron behavior fairly well.

3.2.2. Analysis of Biophysical Parameters From HVC_X Neurons in Two Zebra Finch

Our next use of SDA employs the “nudging” method described in Eq. (9). In this section we used some of the data [Daou and

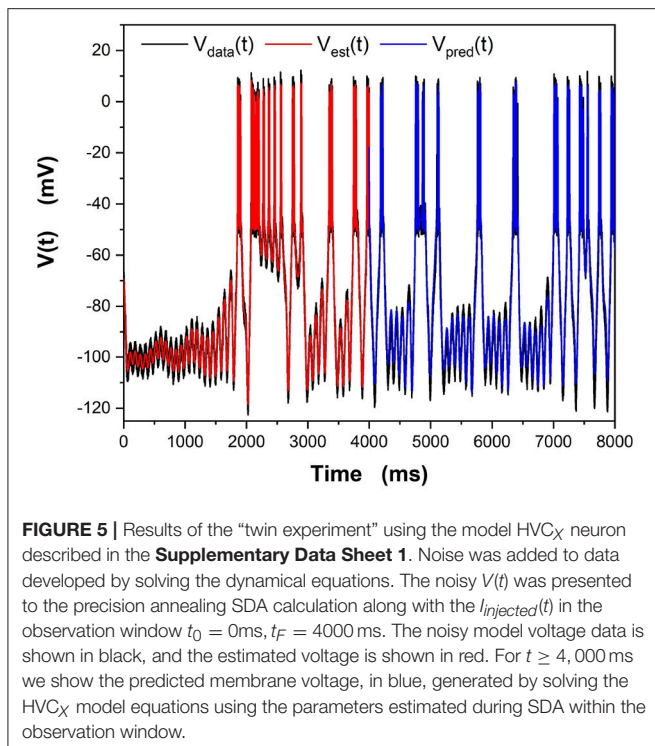
TABLE 2 | Parameter Estimates from the Best Predictions.

Parameter	Bounds	Best estimate	Actual value	Units
g'_{Na}	0.1, 10	4.98	4.5	nS/pF
E_{Na}	1, 100	43.2	50	mV
g_K	0.01, 5	0.907	0.5	nS/pF
E_K	-140, -10	-127.4	-90	mV
g'_{CaT}	0.001, 1	0.0326	0.0265	nS/pF
g_{SK}	0.001, 1	0.0373	0.06	nS/pF
g_h	0.001, 1	0.0432	0.04	nS/pF
E_h	-100, -1	-44.1	-30	mV
C_{inv}	0.001, 0.5	0.011	0.01	pF ⁻¹

The best prediction is chosen by finding the highest correlation coefficient between the predicted voltage and “real” voltage. This comparison can be made on experimental data. It represents an attractive alternative to the familiar least squares metric commonly used. That metric is very sensitive to spike times in data with action potentials: small errors in spike times may result in large errors in a least squares metric.

Margoliash (in review)] taken in experiments on multiple HVC_X neurons from different zebra finches. The questions we asked was whether we could, using SDA, identify differences in biophysical characteristics of the birds. This question is motivated by prior biological observations [Daou and Margoliash (in review)].

Using the same HVC_X model as before, we estimated the maximal conductances $\{g_{Na}, g_K, g_{CaT}, g_{SK}, g_h\}$ holding fixed other kinetic parameters and the Nernst/Reversal potentials. The baseline characteristics of an ion channel are set by the properties of the cell membrane and the complex proteins penetrating the membrane forming the physical channel. Differences among birds would then come from the density or numbers of various channels as characterized by the maximal conductances. If such differences were identified, it would promote further investigation of the biologically exciting proposition that these differences arise in relation to some aspect of the song learning experience of the birds [Daou and Margoliash (in review)].



In **Figures 6A,B** we display the stimulating current and membrane voltage response from one of 9 neurons in our large sample. The analysis using SDA was of four neurons from one bird and seven neurons from another. The results for $\{g_{CaT}, g_{Na}, g_{SK}\}$ is displayed in **Figure 7**. The maximal conductances from one bird are shown in blue and from the other bird, in red. There is a striking difference between the distributions of maximal conductances.

We do not propose here to delve into the biological implications of these results. Nevertheless, we note that the neurons from each bird occupy a small but distinct region of the parameter space (**Figure 7**). This result and its implications for birdsong learning, and more broadly for neuroscience, are described in Daou and Margoliash (in review). Here, however, we display this result as an example of the power of SDA to address a biologically important question in a systematic, principled manner beyond what is normally achieved in analyses of such data.

To fully embrace the utility of SDA for these experiments, however, further work is needed. A limitation of the present result is that the SDA estimates for g_{SK} for a subset of the neurons/observations for Bird One reach the bounds of the observation window (**Figure 7**). Addressing such issues would be prelude to the exciting possibility of estimating more parameters than just the principle ion currents in the Hodgkin-Huxley equations. This could use SDA numerical techniques to calculate, over hours or days, estimates of parameters that could require months or years of work to measure with traditional biological and biophysical approaches, in some cases requiring specialized equipment beyond that available for most *in vitro* recording set

ups. In contrast, applying SDA to such data sets requires only a computer.

3.3. Analysis of Neuromorphic VLSI Instantiations of Neurons

An ambitious effort in neuroscience is the creation of low power consumption analog neural-emulating VLSI circuitry. The goals for this effort range from the challenge itself to the development of fast, reconfigurable circuitry on which to incorporate information revealed in biological experiments for use in

- creating model neural circuits of “healthy” performance to be compared to subsequent observations on the same circuitry informed by “unhealthy” performances. If the comparison can be made rapidly and accurately, the actual instantiations in the VLSI circuitry could be used to diagnose the changes in neuron properties and circuit connectivity perhaps leading to directions for cures, and
- in creating VLSI realizations of neural circuits with desired functions—say, learning syntax in interesting sequences—might allow those functions to be performed at many times increased speed than seen in the biological manifestation. If those functionalities are of engineering utility, the speed up could be critical in applications.

One of the curious roadblocks in achieving critical steps of these goals is that after the circuitry is designed and manufactured into VLSI chips, what comes back from a fabrication plant is not precisely what we designed. This is due to the realities of the manufacturing processes and not inadequacies of the designers.

To overcome this barrier in using the VLSI devices in networks, we need an algorithmic tool to determine just what did return from the factory, so we know how the nodes of a silicon network are constituted. As each chip is an electronic device built on a model design, and the flaws in manufacturing are imperfections in the realization of design parameters, we can use data from the actual chip and SDA to estimate the actual parameters on the chip.

SDA has an advantageous position here. If we present to the chip input signals with much the same design as we prepared for the neurobiological experiments discussed in the previous section, we can measure everything about each output from the chip and use SDA to estimate the actual parameters produced in manufacturing. Of course, we do not know those parameters *a priori* so after estimating the parameters, thus “calibrating” the chip, we must use those estimated parameters to predict the response of the chip to a new stimulating currents. That will validate (or not) the completion of the model of the actual circuitry on the chip and permit confidence in using it in building interesting networks.

We have done this on chips produced in the laboratory of Gert Cauwenberghs at UCSD using PA [38, 39] and using “nudging” as we now report.

The chip we worked with was designed to produce the simplest spiking neuron, namely one having just Na, K, and leak channels [7, 8] as in the original HH experiments. This neuron

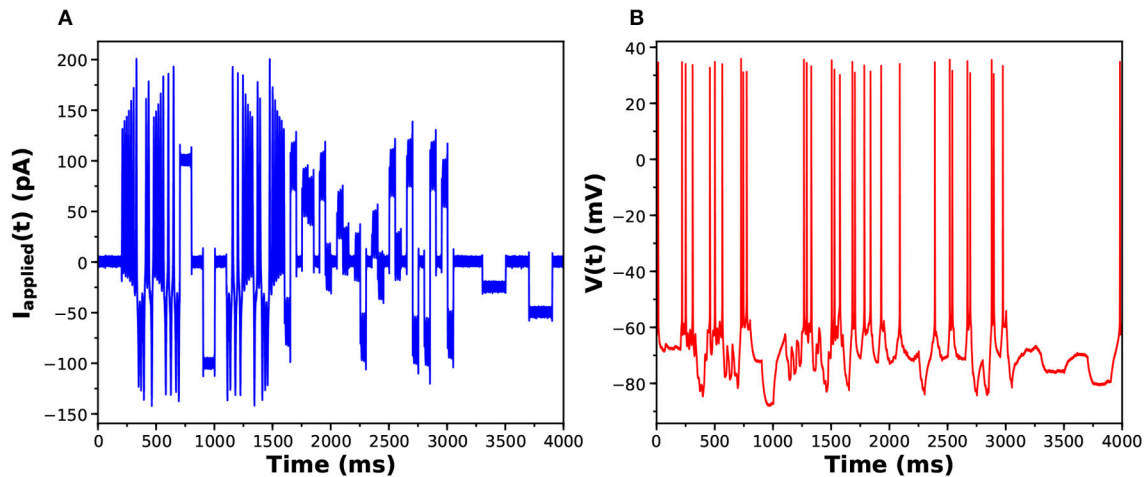


FIGURE 6 | (A) One of the library of stimuli used in exciting voltage response activity in an HVC_X neuron. The cell was prepared *in vitro*, and a single patch clamp electrode injected $I_{injected}(t)$ (this waveform) and recorded the membrane potential. **(B)** The voltage response. One of the library of stimuli used in exciting voltage response activity in an HVC_X neuron. The cell was prepared *in vitro*, and a single patch clamp electrode injected $I_{injected}(t)$ (this waveform) and recorded the membrane potential.

has four state variables $\{V(t), m(t), h(t), n(t)\}$:

$$C \frac{dV(t)}{dt} = g_{Na} m^3(t) h(t) (E_{Na} - V(t)) + g_K n^4(t) (E_K - V(t)) + g_L (E_L - V(t)) + I_{injected}(t)$$

in which the gating variables $w(t) = \{m(t), h(t), n(t)\}$ satisfy

$$\frac{dw(t)}{dt} = \frac{(w_{\infty}(V(t)) - w(t))}{\tau_w(V(t))}, \quad (13)$$

and the functions $w_{\infty}(V)$ are discussed in depth in Johnston and Wu [7] and Sterratt et al. [8].

In our experiments on a “NaKL” chip we used the stimulating current displayed in **Figure 8**,

and measured **all** the neural responses $\{V(t), m(t), h(t), n(t)\}$. These observations were presented to the designed model within SDA to estimate the parameters in the model.

We then tested/validated the estimations by using the calibrated model to predict how the VLSI chip would respond to a different injected current. In **Figure 9** we show the observed $V_{data}(t)$ in black, the estimation of the voltage through SDA in red, and the prediction of $V(t)$ in blue for times after the end of the observation window.

While one can be pleased with the outcome of these procedures, for our purposes we see that the use of our SDA algorithms gives the user substantial confidence in the functioning characteristics of the VLSI chips one wishes to use at the nodes of a large, perhaps even very large, realization of a desired neural circuit in VLSI. We are not unaware of the software developments to allow efficient calibration of very large numbers of manufactured silicon neurons. A possible worry about also determining the connectivity, both the links and their strength and time constants, may be alleviated by realizing these links through a high bandwidth bi-directional connection of the massive array of chips and the designation of connectivity characteristics on an off-chip computer.

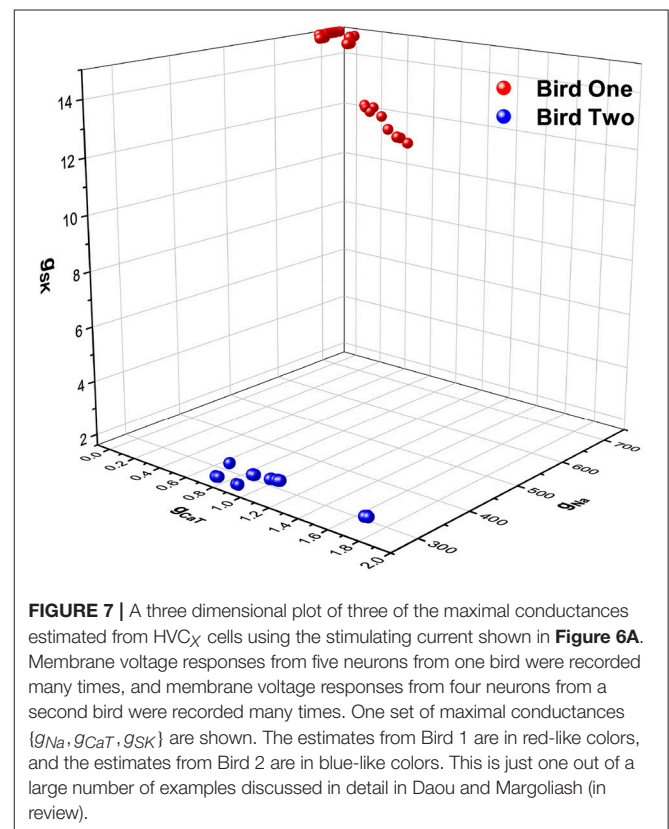
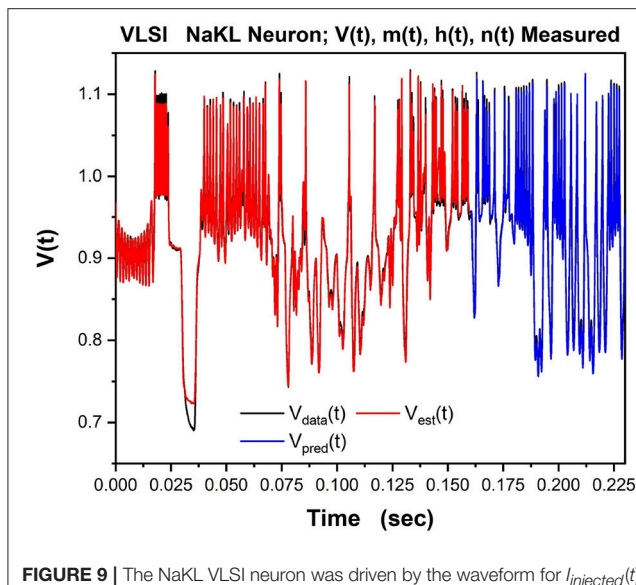
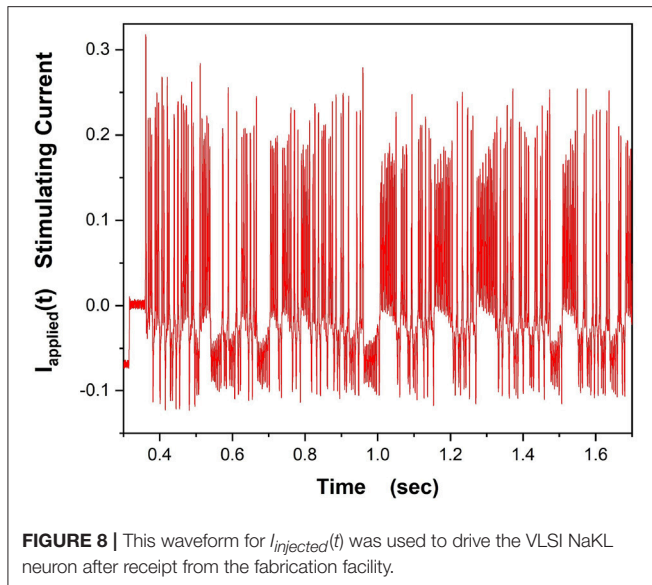


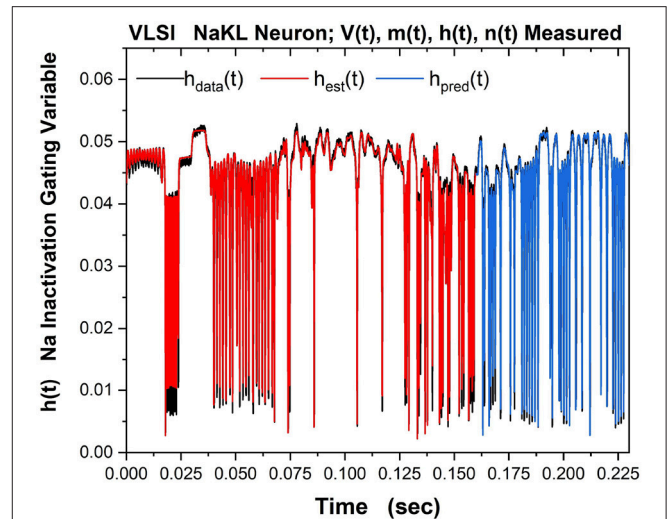
FIGURE 7 | A three dimensional plot of three of the maximal conductances estimated from HVC_X cells using the stimulating current shown in **Figure 6A**. Membrane voltage responses from five neurons from one bird were recorded many times, and membrane voltage responses from four neurons from a second bird were recorded many times. One set of maximal conductances $\{g_{Na}, g_{CaT}, g_{SK}\}$ are shown. The estimates from Bird 1 are in red-like colors, and the estimates from Bird 2 are in blue-like colors. This is just one out of a large number of examples discussed in detail in Daou and Margoliash (in review).

Part of the same analysis is the ability to observe, estimate and predict the experimentally unobservable gating variables. This serves, in this context, as a check on the SDA calculations. The Na inactivation variable $h(t)$ is shown in **Figure 10** as its measured time course $h_{data}(t)$ in black, its estimated time course $h_{est}(t)$ in red, and its predicted time course $h_{pred}(t)$ in blue.



4. DISCUSSION

Our review of the general formulation of statistical data assimilation (SDA) started our remarks. Many details can be found in Abarbanel [3] and subsequent papers by the authors. Recognizing that the core problem is to perform, approximately of course, the integral in Equation (6) is the essential take away message. This task requires well “curated” data and a model of the processes producing the data. In the context of experiments in life sciences or, say, acquisition of data from earth system sensors,



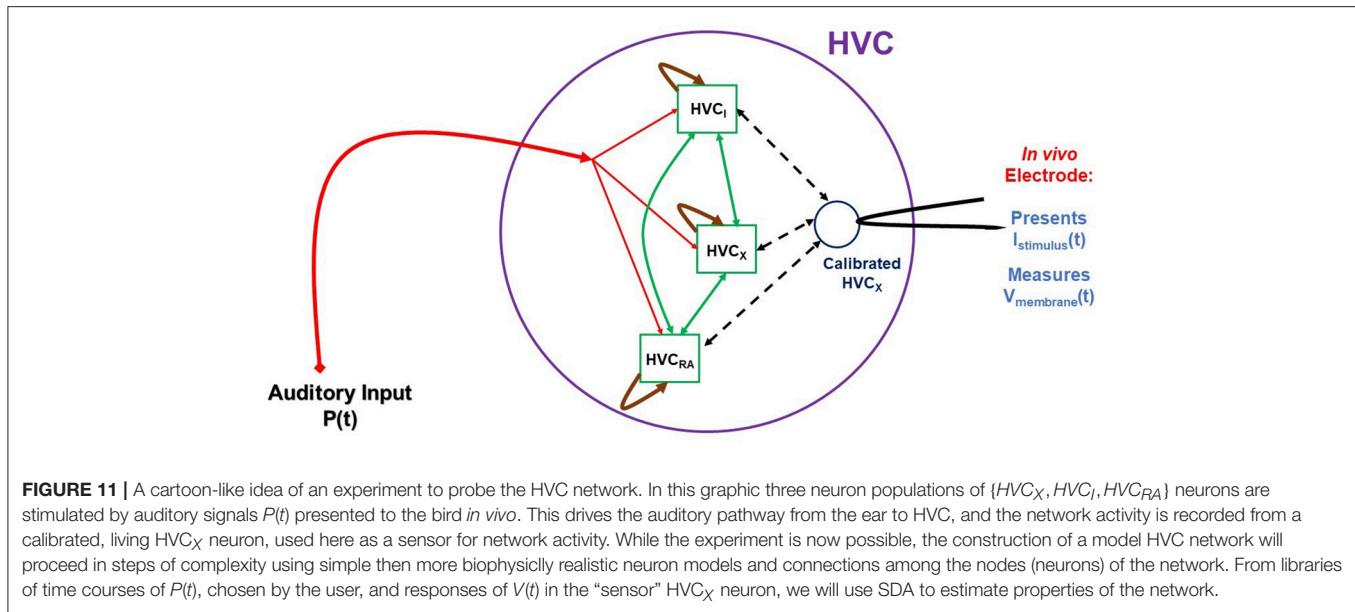
curation includes an assessment of errors and the properties of the instruments as well.

One we have the data and a model, we still need a set of procedures to transfer the information from the data to the model, then test/validate the model on data not used to train the model. The techniques we covered are general. Their application to examples from neuroscience comprise the second part of this paper.

In the second part we first address properties of the avian songbird song production pathway and a neural control pathway modulating the learning and production of functional vocalization—song. We focus our attention on one class of neurons, HVC_X, but have also demonstrated the utility of SDA to describe the response properties of other classes of neurons in HVC, such as HVC_{RA} [40] and HVC_I [41]. Indeed, the SDA approach is generally applicable wherever there is insight to relate biophysical properties of neurons to their dynamics through Hodgkin Huxley equations.

Our SDA methods considered variational algorithms that seek the highest conditional probability distributions of the model states and parameters conditioned on the collection of observations over a measurement window in time. Other approaches, especially Monte Carlo algorithms were not discussed here, but are equally attractive.

We discussed methods of testing models of HVC_X neurons using “twin experiments” in which a model of the individual neuron produces synthetic data to which we add noise with a selected signal-to-noise-ratio. Some state variable time courses from the library of these model produced data, for us the voltage



across the cell membrane, is then part of the action Equation (8), specifically in the measurement error term. Errors in the model are represented in the model error term of the action.

Using a precision annealing protocol to identify and track the global minimum of the action, the successful twin experiment gives us confidence in this SDA method from information trans from data to the model.

We then introduced a “nudging” method as an approximation to the Euler-Lagrange equations derived from the numerical optimization of the action Equation (8)—this is Laplace’s method in our SDA context. The nudging method, introduced in meteorology some time ago, was used to distinguish between two different members of the Zebra Finch collection. We showed, in a quite preliminary manner, that the two, unrelated birds of the same species, express different HVC network properties as seen in a critical set of maximal conductances for the ion channels in their dynamics.

Finally we turned to a consideration of the challenge of implementing in VLSI technology the neurons in HVC toward the goal of building a silicon-HVC network. The challenge at the design and fabrication stages of this effort where illuminated by our use of SDA to determine what was actually returned from the manufacturing process for our analog neurons.

4.1. Moving Forward to Network Analysis

Finally, we have a few comments associated with the next stage of analysis of HVC. In this, and previous papers, we analyzed individual neurons in HVC. These analyses were assisted by our using SDA, through twin experiments, to design laboratory experiments though the selection of effective stimulating injected currents.

Having characterized the electrophysiology of an individual neuron within the framework of Hodgkin-Huxley (HH) models, we may now proceed beyond the study of individual neurons [42] *in vitro*. Once we have characterized an HVC neuron through a biophysical HH model, we may then use it *in vivo* as a sensor of the activity of the HVC network where it is connected to HVC_{RA} , HVC_I , and other HVC_X neurons. The schema for this

kind of experiment is displayed in **Figure 11**. These experiments require the capability to perform measurements on HVC neurons in the living bird. That capability is available, and experiments as suggested in our graphic are feasible, if challenging.

The schematic indicates that the stimulating input to the experiments is auditory signals, chosen by the user, presented to the bird’s ear and reaching HVC through the auditory pathway. The stimuli from this signal is then distributed in a manner to be deduced from experiment and then produces activity in the HVC network that we must model. The goal is, at least initially, to establish, again within the models we develop, the connectivity of HVC neuron classes as it manifests itself in the function of the network. We have some information about this [43, 44], and these results will guide the development of the HVC model used in these whole-network experiments. An important point to address is what changes to the *in vitro* model might be necessary to render it a model for *in vivo* activity.

DATA AVAILABILITY

The raw data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

AM generated and analyzed numerical data for the twin experiment on HVCX model equations. DL analyzed the biological HVCX data taken by AD. JP analyzed data from the VLSI chip. AM and HA wrote the manuscript with comments from DM, AD, JP, and DL.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fams.2018.00053/full#supplementary-material>

REFERENCES

- Lorenz EN. Predictability: a problem partly solved. In: Palmer T, Hagedorn R, editors. *Predictability of Weather and Climate*. Cambridge University Press, Cambridge (2006), p. 342–61.
- Evensen G. *Data Assimilation: The Ensemble Kalman Filter*. Berlin: Springer (2009).
- Abarbanel HDI. *Predicting the Future: Completing Models of Observed Complex Systems*. New York, NY: Springer (2013).
- Tong D. *Statistical Physics*. (2011). Available online at: www.damtp.cam.ac.uk/user/tong/statphys/sp.pdf
- Laplace PS. Memoir of the probability of causes of events. *Stat. Sci.* (1986) 1:354–78. doi: 10.1214/ss/1177013621
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP. *Numerical Recipes: The Art of Scientific Computing*, 3rd ed. New York, NY: Cambridge University Press (2007).
- Johnston D, Wu SMS. *Foundations of Cellular Neurophysiology*. Cambridge, MA: Bradford Books, MIT Press (1995).
- Sterratt D, Graham B, Gillies A, Willshaw D. *Principles of Computational Modelling in Neuroscience*. New York, NY: Cambridge University Press (2011).
- Abarbanel HDI, Rozdeba PJ, Shirman S. Machine learning: deepest learning as statistical data assimilation problems. *Neural Comput.* (2018). 30:2025–55. doi: 10.1162/neco_a_01094
- Fano RM. *Transmission of Information; A Statistical Theory of Communication*. Cambridge, MA: MIT Press (1961).
- Kostuk M, Toth BA, Meliza CD, Margoliash D, Abarbanel HDI. Dynamical estimation of neuron and network properties II: path integral Monte Carlo methods. *Biol Cybern.* (2012) 106:155–67. doi: 10.1007/s00422-012-0487-5
- Neal R. MCMC using hamiltonian dynamics. In: Gelman A, Jones G, Meng XL, editors. *Handbook of Markov Chain Monte Carlo*. Sound Parkway, NJ: Chapman and Hall, CRC Press (2011), p. 113–63.
- Murty KG, Kabadi SN. Some NP-complete problems in quadratic and nonlinear programming. *Math Program.* (1987) 39:117–29. doi: 10.1007/BF02592948
- Ye J, Kadakia N, Rozdeba PJ, Abarbanel HDI, Quinn JC. Improved variational methods in statistical data assimilation. *Nonlin Process Geophys.* (2015) 22:205–13. doi: 10.5194/npg-22-205-2015
- Ye J, Rey D, Kadakia N, Eldridge M, Morone UI, Rozdeba P, et al. Systematic variational method for statistical nonlinear state and parameter estimation. *Phys Rev E* (2015) 92:052901. doi: 10.1103/PhysRevE.92.052901
- Ye J. *Systematic Annealing Approach for Statistical Data Assimilation*. La Jolla, CA: University of California San Diego (2016).
- Quinn JC. *A path integral approach to data assimilation in stochastic nonlinear systems*. La Jolla, CA: University of California San Diego (2010).
- Shirman S. *Strategic Monte Carlo and Variational Methods in Statistical Data Assimilation for Nonlinear Dynamical Systems*. University of California San Diego (2018).
- Anthes R. Data assimilation and initialization of hurricane prediction models. *J Atmos Sci.* (1974) 31. doi: 10.1175/1520-0469(1974)031<0702:DAAIOH>2.0.CO;2
- Gelfand IM, Fomin SV. *Calculus of Variations*. Dover Publications, Inc. (1963).
- Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *J Chem Phys.* (1953) 21:1087–92. doi: 10.1063/1.1699114
- Doupe AJ, Kuhl PK. Birdsong and human speech: common themes and mechanisms. *Annu. Rev. Neurosci.* (1999) 22:567–631. doi: 10.1146/annurev.neuro.22.1.567
- Daou A, Ross M, Johnson F, Hyson R, Bertram R. Electrophysiological characterization and computational models of HVC neurons in the zebra finch. *J Neurophysiol.* (2013) 110:1227–45. doi: 10.1152/jn.00162.2013
- Bolhuis JJ, Okanoya K, Scarff C. Twitter evolution: converging mechanisms in birdsong and human speech. *Nat Rev Neurosci.* (2010) 11:747–59. doi: 10.1038/nrn2931
- Fee MS, Scharff C. The songbird as a model for the generation and learning of complex sequential behaviors. *ILAR J.* (2010) 51:362–77. doi: 10.1093/ilar.51.4.362
- Mooney R. Neuronal mechanisms for learned birdsong. *Learn Mem.* (2009) 16:655–69. doi: 10.1101/lm.1065209
- Simonyan K, Horwitz B, Jarvis E. Dopamine regulation of human speech and bird song: a critical review. *Brain Lang.* (2012) 122:142–50. doi: 10.1016/j.bandl.2011.12.009
- Teramitsu I, Kudo LC, London SE, Geschwind DH, White SA. Parallel FoxP1 and FoxP2 expression in songbird and human brain predicts functional interaction. *J Neurosci.* (2004) 24:3152–63. doi: 10.1523/JNEUROSCI.5589-03.2004
- Jarvis ED, Güntürkün O, Bruce L, Csillag A, Karten H, Kuenzel W, et al. Avian brains and a new understanding of vertebrate brain evolution. *Nat Rev Neurosci.* (2005) 6:151–9. doi: 10.1038/nrn1606
- Doupe AJ, Perkel DJ, Reiner A, Stern EA. Birdbrains could teach basal ganglia research a new song. *Trends Neurosci.* (2005) 28:353–63. doi: 10.1016/j.tins.2005.05.005
- Nottebohm F, Stokes TM, Leonard CM. Central control of song in the canary, *Serinus canarius*. *J Compar Neurol.* (1976) 165:457–86. doi: 10.1002/cne.901650405
- Fortune ES, Margoliash D. Parallel pathways and convergence onto HVC and adjacent neostriatum of adult zebra finches (*Taeniopygia guttata*). *J Compar Neurol.* (1995) 360:413–41. doi: 10.1002/cne.903600305
- Brainard MS, Doupe AJ. Auditory feedback in learning and maintenance of vocal behaviour. *Nat Rev Neurosci.* (2000) 1:31–40. doi: 10.1038/35036205
- Hamill OP, Marty A, Neher E, Sakmann B, Sigworth FJ. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflügers Archiv.* (1981) 391:85–100. doi: 10.1007/BF00656997
- Hernández-Ochoa EO, Schneider MF. Voltage clamp methods for the study of membrane currents and SR Ca^{2+} release in adult skeletal muscle fibres. *Prog Biophys Mol Biol.* (2012) 108:98–118. doi: 10.1016/j.pbiomolbio.2012.01.001
- Bagal SK, Marron BE, Owen RM, Storer RI, Swain NA. Voltage gated sodium channels as drug discovery targets. *Channels (Austin).* (2015) 9:360–6. doi: 10.1080/19336950.2015.1079674
- Frolov RV, Weckström M. Harnessing the flow of excitation: TRP, voltage-gated Na^{+} , and voltage-gated Ca^{2+} channels in contemporary medicine. In *Ion Channels as Therapeutic Targets, Part A*. Cambridge, MA: Academic Press (2016), p. 25–95.
- Breen D. *Characterizing Real World Neural Systems Using Variational Methods of Data Assimilation*. La Jolla, CA: University of California San Diego (2017).
- Wang J, Breen D, Akinin A, Broccard F, Abarbanel HDI, Cauwenberghs G. Assimilation of biophysical neuronal dynamics in neuromorphic VLSI. *Biomed Circuits Syst.* (2017) 11:1258–70. doi: 10.1109/TBCAS.2017.2776198
- Kadakia N, Armstrong E, Breen D, Daou A, Margoliash D, Abarbanel H. Nonlinear statistical data assimilation for HVCRA neurons in the avian song system. *Biol Cybern.* (2016) 110:417–34. doi: 10.1007/s00422-016-0697-3
- Breen D, Shirman S, Armstrong E, Kadakia N, Abarbanel H. HVCRA Neuron Properties from Statistical Data Assimilation (2016).
- Nogaret A, Meliza CD, Margoliash D, Abarbanel. Automatic construction of predictive neuron models through large scale assimilation of electrophysiological data. *Sci Rep.* (2016) 6:32749. doi: 10.1038/srep32749
- Mooney R, Prather JF. The HVC microcircuit: the synaptic basis for interactions between song motor and vocal plasticity pathways. *J Neurosci.* (2005) 25:1952–64. doi: 10.1523/JNEUROSCI.3726-04.2005
- Kosche G, Vallentin D, Long MA. Interplay of inhibition and excitation shapes a premotor neural sequence. *J Neurosci.* (2015) 35:1217–27. doi: 10.1523/JNEUROSCI.4346-14.2015

Conflict of Interest Statement: 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.

Copyright © 2018 Miller, Li, Platt, Daou, Margoliash and Abarbanel. 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.



Data-Driven Modeling and Prediction of Complex Spatio-Temporal Dynamics in Excitable Media

Sebastian Herzog^{1,2}, Florentin Wörgötter² and Ulrich Parlitz^{1,3,4*}

¹ Max Planck Institute for Dynamics and Self-Organization, Göttingen, Germany, ² Third Institute of Physics and Bernstein Center for Computational Neuroscience, University of Göttingen, Göttingen, Germany, ³ Institute for Nonlinear Dynamics, University of Göttingen, Göttingen, Germany, ⁴ DZHK (German Centre for Cardiovascular Research), Partner Site Göttingen, Göttingen, Germany

OPEN ACCESS

Edited by:

Axel Hutt,
German Weather Service, Germany

Reviewed by:

Christian Andreas Welzbacher,
German Weather Service, Germany
Xin Tong,
National University of Singapore,
Singapore

*Correspondence:

Ulrich Parlitz
ulrich.parlitz@ds.mpg.de

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 04 September 2018

Accepted: 26 November 2018

Published: 11 December 2018

Citation:

Herzog S, Wörgötter F and Parlitz U
(2018) Data-Driven Modeling and
Prediction of Complex
Spatio-Temporal Dynamics in
Excitable Media.
Front. Appl. Math. Stat. 4:60.
doi: 10.3389/fams.2018.00060

Spatio-temporal chaotic dynamics in a two-dimensional excitable medium is (cross-) estimated using a machine learning method based on a convolutional neural network combined with a conditional random field. The performance of this approach is demonstrated using the four variables of the Bueno-Orovio-Fenton-Cherry model describing electrical excitation waves in cardiac tissue. Using temporal sequences of two-dimensional fields representing the values of one or more of the model variables as input the network successfully cross-estimates all variables and provides excellent forecasts when applied iteratively.

Keywords: deep learning, conditional random fields, artificial neural network, cross-estimation, spatio-temporal chaos, excitable media, cardiac arrhythmias, non-linear observer

1. INTRODUCTION

In life sciences mathematical models based on first principles are scarce and often a variety of approximate models of different complexity exists for describing the given (experimental) dynamical process. For example, electrical excitation waves in cardiac tissue can be described using partial differential equations (PDEs) with 2 to more than 60 variables, covering the range from simple qualitative models [1, 2] to detailed ionic cell models including not only cell membrane voltage but also different ionic currents and gating variables [3, 4]. While there are several modalities for measuring membrane voltage (electrical sensors, fluorescent dyes [5]) it is in general much more difficult and expensive (if not impossible) to directly measure the other variables of the mathematical model, such as ionic currents or gating variables. In such cases it is desirable to (cross) estimate variables, which are difficult to assess from those that can be easily measured. In control theory this task is addressed by constructing an observer based on a given mathematical model describing the process of interest. Once all state variables of the model have been estimated, the model (e.g., a PDE) can be used to simulate and forecast the future evolution of the dynamical process. This combination of cross estimation and prediction of dynamical variables is the core of all data assimilation methods [6–10] where again the model equations are involved and have to be known. In this contribution, we present a machine learning method for estimating all state variables and forecasting their evolution from limited observations. This “black-box model” consists of a convolutional neural network (CNN) combined with a conditional random field (CRF) and will be introduced in section 2. For training and evaluating the network two dimensional spatio-temporal time series are used, which were generated by the Bueno-Orovio-Fenton-Cherry (BOCF) model [11] describing complex electrical excitation waves in cardiac tissue. This model is introduced in

section 3. As modeling tasks we consider cross estimation of variables, forecasting dynamics using an iterative feedback scheme, and a combination of forecasting and cross estimation providing future values of not measured variables. These results are presented in section 4. A summary and a brief discussion of potential future developments are given in section 5.

2. DATA DRIVEN MODELING

In data driven modeling mathematical models are not based on first principles (e.g., Newton's laws, Maxwell's equations, ...) but are directly derived from experimental measurement data or other physical observations. The model should describe the experiment as precisely as possible but it also should possess a high level of generalizability, i.e., the ability to provide a suitable and good description for data from a very similar experiment. Therefore, overfitting has to be avoided and all irrelevant aspects that are not necessary to describe the desired effect should be discarded when generating the model (without employing human expert knowledge). Many approaches for generating (dynamical) models from (training) data have been devised including autoregressive models [12], evolutionary algorithms in particular genetic algorithms [13], local modeling [14], reservoir computing [15–19], symbolic regression [20], or adaptive fuzzy rule-based models [21]. Furthermore, Monte Carlo techniques may be used for assessing uncertainty in model parameters [22]. In this work we present a modeling ansatz which combines deep convolutional neural networks [23] for feature extraction and dimension reduction with conditional random fields (CRFs) [24] for modeling the properties of temporal sequences.

2.1. Artificial Neural Network

Artificial neural networks (ANNs) [25–27] are parameterizable models for approximating a (unknown) function F implicitly given by the data. The actual function provided by the ANN:

$$f: \mathbb{R}^O \mapsto \mathbb{R}^P, \quad (1)$$

should be a good approximation of F , i.e., $f \simeq F$. Here $O \in \mathbb{N}$ and $P \in \mathbb{N}$ denote the dimension of the input and the output of f , respectively. A widely used type of ANN are feed-forward neural networks (FNN) where, in general, f is given by

$$f(X) = \psi(WX + b), \quad (2)$$

with a non-linear function ψ applied component-wise, an input vector $X \in \mathbb{R}^O$, a weight matrix $W \in \mathbb{R}^{P \times O}$, and a bias $b \in \mathbb{R}^P$. Equation (2) is called a one-layer FNN. By recursively applying the output of one layer as input to the next layer, a multi-layer FNN can be constructed:

$$f(X) = f^L(\dots f^2(f^1(X; W^1, b^1); W^2, b^2) \dots; W^L, b^L). \quad (3)$$

Equation (3) describes a multi-layer FNN with $L \in \mathbb{N}$ layers. In the following an input with several variables is considered and the input is given by $X \in \mathbb{R}^{h \times w \times d}$, with $h \in \mathbb{N}$ rows and $w \in \mathbb{N}$ columns of the input field, and the number of variables

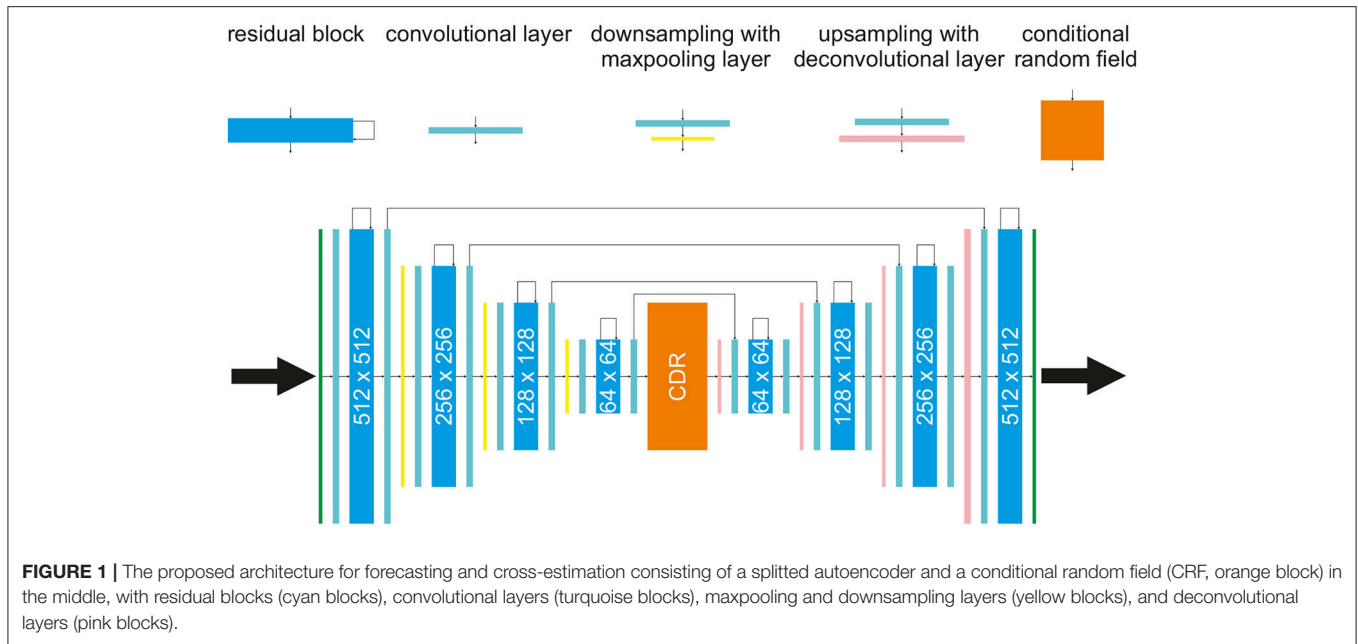
d . To improve the approximation properties of the network Equation (3), FNNs may contain additional convolutional layers leading to state-of-the-art models for data classification, so-called convolutional neural networks (CNNs) [23].

2.2. Network Architecture

The network used in the following for prediction of multivariate time series is built based on the architecture of a convolutional autoencoder [28], with residual connections [29] consisting of an encoding path (left half of the network, from 512×512 to 64×64) to retrieve the features of interest and a symmetric decoding path (right half of the network, from 64×64 back to 512×512). As illustrated in **Figure 1** each encoding/decoding path consists of multiple levels, i.e., resolutions, for feature extraction on different scales and noise reduction. The conditional random field block has a special role: Based on the selected feature, the CRF should map a sequence of features of a previous time step t to the next time step $t + \Delta t$. The other four components of the network are basic building blocks, like regular convolutional layers followed by rectified linear unit activation and batch normalization (these blocks are omitted in **Figure 1** for simplicity). Each residual layer consists of three convolutional blocks and a residual skip connection. A maxpooling layer is located between levels in the encoding path to perform downsampling for feature compression. The deconvolutional layer [30] is located between levels in the decoding path to up-sample the input data using learnable interpolations. The input for the network are all four system variables of the BOCF model which will be introduced in section 3.1 or a sequence of the four system variables as introduced in section 4.1. The output of the network always consists of four system variables.

2.3. Convolution Layer

Convolutional neural networks [23, 26, 27] receive a training data set $\mathcal{X} = \{X_1, X_2, \dots, X_m\}$, where $X_\alpha \in \mathbb{R}^{h \times w \times d}$. The data processing through the network is described layer-wise i.e., in the l -th convolutional layer the input $X^{(l)}$ will be transformed to the *raw output* $o^{(l)}$, which is in turn the input to the next layer $l + 1$, where the dimension changes depending on the number and size of convolutions, padding and stride of the layers as illustrated in **Figure 1**. The padding parameter $P^{(l)} \in \mathbb{N}$, for layer l , describes the number of zeros at the edges of a field by which the field is extended. This is necessary since every convolution being larger than 1×1 will decrease the output size. The stride parameter $S^{(l)} \in \mathbb{N}$ is the parameter determining how much the kernel is shifted in each step to compute the next spatial position (x, y) . This specifies the overlap between individual output pixels, and it is here set to 1. Each layer l is specified by its number of kernels $K^{(l)} = \{K^{(l,1)}, K^{(l,2)}, \dots, K^{(l,d^{(l)})}\}$, where $d^{(l)} \in \mathbb{N}$ is the number of kernels in layer l , and its additive bias terms $b^{(l)} = \{b^{(l,1)}, b^{(l,2)}, \dots, b^{(l,d^{(l)})}\}$ with $b^{(l,d)} \in \mathbb{R}$. Note that the input $X^{(l,d)} \in \mathbb{R}^{h^{(l)} \times w^{(l)}}$ in the l -th layer with size $h^{(l)} \times w^{(l)}$, kernel k , and depth $d^{(l)}$ is processed by a set of kernels $\{K^{(l,d)}\}$. For each kernel $K^{(l,d)} \in \mathbb{R}^{h_K^{(l)} \times w_K^{(l)}}$ with size $h_K^{(l)} \times w_K^{(l)}$ and $d \in \{1, \dots, d^{(l)}\}$, the *raw output* $o^{(l)} \in \mathbb{R}^{\frac{h^{(l)} - h_K^{(l)} - 1 + P^{(l)}}{S^{(l)}} \times \frac{w^{(l)} - w_K^{(l)} - 1 + P^{(l)}}{S^{(l)}}}$ is computed



element by element as:

$$\begin{aligned} o_{x,y}^{(l,d)} &= b^{(l,d)} + \left(K^{(l,d)} * X^{(l,d)} \right)_{x,y} \\ &= b^{(l,d)} + \sum_{k=1}^{d^{(l)}} \sum_{i=1}^{h_k^{(l)}} \sum_{j=1}^{w_k^{(l)}} K_{i,j}^{(l,d)} \cdot X_{x+i-1,y+j-1}^{(l,k)}. \end{aligned} \quad (4)$$

The result is clipped by an *activation function* ψ to obtain the *activation* $\psi(o_{x,y}^{(l,d)})$ of each unit in layer l :

$$\psi(o_{x,y}^{(l,d)}) = \max \{ 0, o_{x,y}^{(l,d)} \}. \quad (5)$$

To obtain $o^{(l)} = \{o^{(l,1)}, \dots, o^{(l,d^{(l)})}\}$, Equation (5) needs to be calculated $\forall d = 1, \dots, d^{(l)}$ and $\forall (x, y)$. Each spatial calculation of $o_{x,y}^{(l,d)}$ is considered as a *unit* and $\psi(o_{x,y}^{(l,d)})$ as the feedforward *activation* of the unit. The value of an element of a kernel ($K_{i,j}^{(l,d)}$) between two units is the weight of the feedforward connection. Such systems are well-suited for feature extraction [28], but their linear structure does not allow a direct modeling of temporal changes or the possibility to process a sequence of data. To enable temporal modeling, we employ linear-chain conditional random fields [31] that will be introduced in the next section.

2.4. Linear-Chain Conditional Random Fields

To implement a probabilistic forecasting block we consider the output of the convolutional layer \mathbf{o} and the corresponding forecast q as random variables \mathbf{O} and \mathbf{Q} , respectively. Both random variables \mathbf{O} and \mathbf{Q} are jointly distributed and in a predictive framework we aim at constructing a conditional model $P(\mathbf{Q}|\mathbf{O})$ from paired observation and forecast sequences. Let $G =$

(V, E) be a undirected graph such that $\mathbf{Q} = (\mathbf{Q}_v)_{v \in V}$, where \mathbf{Q} is indexed by the vertices of G . Each vertex in G represents a state, a history or a forecast. Then (\mathbf{O}, \mathbf{Q}) is a conditional random field (CRF), if conditioned on \mathbf{O} the random variables \mathbf{Q}_v obey the Markov property [24]. A linear-chain conditional random field, where \mathbf{o} is a sequence of historical extracted features and q a corresponding forecasted feature in the future, is given by:

$$\begin{aligned} P(q | \mathbf{o}, \theta) &= \sum_{\mathbf{h} \in \mathcal{H}} P(q, \mathbf{h} | \mathbf{o}, \theta) \\ &= \frac{\sum_{\mathbf{h} \in \mathcal{H}} \exp(\Psi(q, \mathbf{h}, \mathbf{o}; \theta))}{\sum_{q' \in \mathcal{Q}} \sum_{\mathbf{h} \in \mathcal{H}} \exp(\Psi(q', \mathbf{h}, \mathbf{o}; \theta))}, \end{aligned} \quad (6)$$

where $q \in \mathcal{Q}$, \mathcal{Q} is a set of future events, $\mathbf{h} \in \mathcal{H}$, \mathcal{H} is the set of layers of the CRF where each element h_i of \mathbf{h} represents a historical state of an event at time t . θ is the set of parameters. $\Psi(q, \mathbf{h}, \mathbf{o}; \theta)$ is a so called *potential function* (also called *local* or *compatibility function*) which measures the compatibility (i.e., high probability) between a forecast, a set of observed features, and a configuration of historical states, such that:

$$\begin{aligned} \Psi(q, \mathbf{h}, \mathbf{o}; \theta) &= \sum_{j=1}^n \phi_j(\mathbf{o}, \omega) \cdot \theta_h[h_j] \\ &+ \sum_{j=1}^n \theta_y[y, h_j] + \sum_{(i,j) \in \epsilon} \theta_e[q, h_j, h_k] + \frac{\phi(\mathbf{o}, \omega) \cdot \theta_p[q]}{k}, \end{aligned} \quad (7)$$

Here n is the number of historical states and $\phi_j(\mathbf{o}, \omega)$ is a vector that can include any feature of the observation specific for a specific time window ω , and $\theta = [\theta_h, \theta_q, \theta_e, \theta_p]$ are model parameters. To restrict the search space for possible parametrizations only sine, cosine, and a linear interpolation

function are allowed to be used as feature functions. $\theta_h[h_j]$ is the parameter that corresponds to the state h_j . The function $\theta_q[q, h_j]$ indicates the parameters that corresponds to the forecast q and the state h_j . $\theta_e[q, h_i, h_k]$ refers to parameters that describe the dependency relation between the nodes h_i and h_k . $\theta_p[q]$ defines the parameters for q given the features over the past, while the dot product $\phi_j(\mathbf{o}, \omega) \cdot \theta_n[h_j]$ measures the compatibility between the observed features and the state at time j . In contrast to this $\phi(\mathbf{o}, \omega) \cdot \theta_p[q]$ measures the compatibility between observation and the forecast. \mathbf{h} consists of $k = 1,024$ elements and the last term in Equation (7) captures the influence of the past features on the forecast. For training the following likelihood function is defined:

$$L(\theta) = \sum_{i=1}^n P(q_i | \mathbf{o}_i, \theta) - \frac{1}{2\sigma^2} \|\theta\|^2, \quad (8)$$

where n is the number of training examples. By maximizing the likelihood for the forecasted training data the optimal parameter set θ^* is determined. To find θ^* Equation (8) can be evaluated by the same gradient descent method which is used for optimizing/training the autoencoder. To forecast the input sequence with a linear-chain CRF it is necessary to compute the q sequence that maximizes the following equation:

$$\hat{q} = \arg \max_q P(q | \mathbf{o}; \theta^*) \quad (9)$$

The sequence maximizing this is then used by the deconvolutional part of the network to map the features back to the desired system variables at $t + \Delta t$.

3. MODELING EXCITABLE MEDIA

Excitable systems are non-linear dynamical systems with a stable fixed point. Small perturbations of the stable equilibrium decay, but stronger perturbations above some characteristic threshold lead to a high amplitude excursion in state space until the trajectory returns to the stable fixed point. In neural or cardiac cells this response leads to a so-called action potential. After such a strong response a so-called *refractory period* has to pass until the next excitation can be initialized by perturbing the system again. An excitable medium consists of excitable systems (e.g., cells), which are spatially coupled. Electric coupling of neighboring cardiac cells, for example, can be modeled by means of a diffusion term for local currents. The resulting partial differential equations (PDEs) describe the propagation of undamped solitary excitation waves. Due to the refractory time of local excitations spiral or scroll waves are very common and typical hallmarks of excitable media, which can lead to stable periodically rotating wave patterns or may break-up forming complex chaotic wave dynamics. From the large selection of different PDE models describing excitable media we have chosen the Bueno-Orovio-Cherry-Fenton (BOCF) model which was devised as an efficient model for cardiac tissue [11].

3.1. Bueno-Orovio-Cherry-Fenton Model

The Bueno-Orovio-Cherry-Fenton (BOCF) model [11] provides a compact description of excitable cardiac dynamics. We use this model as a benchmark to validate our approach for forecasting and cross-estimation of complex wave patterns in excitable media. The evolution of the four system variables of the BOCF model is given by four PDEs

$$\begin{aligned} \frac{\partial u}{\partial t} &= D \cdot \nabla^2 u - (J_{si} + J_{fi} + J_{so}) \\ \frac{\partial v}{\partial t} &= \frac{1}{\tau_v^-} (1 - H(u - \theta_v)) (v_\infty - v) - \frac{1}{\tau_v^+} H(u - \theta_v) v \\ \frac{\partial w}{\partial t} &= \frac{1}{\tau_w^-} (1 - H(u - \theta_w)) (w_\infty - w) - \frac{1}{\tau_w^+} H(u - \theta_w) w \\ \frac{\partial s}{\partial t} &= \frac{1}{2\tau_s} ((1 + \tanh(k_s(u - u_s))) - 2s), \end{aligned} \quad (10)$$

where u represents the membrane voltage and $H(\cdot)$ denotes the Heaviside function. The three currents J_{si} , J_{fi} and J_{so} are given by the equations

$$\begin{aligned} J_{si} &= -\frac{1}{\tau_{si}} H(u - \theta_w) w s \\ J_{fi} &= -\frac{1}{\tau_{fi}} v H(u - \theta_v) (u - \theta_v) (u_u - u) \\ J_{so} &= \frac{1}{\tau_o} (u - u_o) (1 - H(u - \theta_w)) + \frac{1}{\tau_{so}} H(u - \theta_w). \end{aligned} \quad (11)$$

Furthermore, seven voltage dependent variables

$$\begin{aligned} \tau_v^- &= (1 - H(u - \theta_v^-)) \tau_{v1}^- + H(u - \theta_v^-) \tau_{v2}^- \\ \tau_w^- &= \tau_{w1}^- + \frac{1}{2} (\tau_{w2}^- - \tau_{w1}^-) (1 + \tanh(k_w^- (u - u_w^-))) \\ \tau_{so}^- &= \tau_{so1} + \frac{1}{2} (\tau_{so2} - \tau_{so1}) (1 + \tanh(k_{so} (u - u_{so}))) \\ \tau_s &= (1 - H(u - \theta_w)) \tau_{s1} + H(u - \theta_w) \tau_{s2} \\ \tau_o &= (1 - H(u - \theta_o)) \tau_{o1} + H(u - \theta_o) \tau_{o2} \end{aligned} \quad (12)$$

$$v_\infty = \begin{cases} 1, & \text{if } u \leq \theta_v^- \\ 0, & \text{if } u \geq \theta_v^- \end{cases}$$

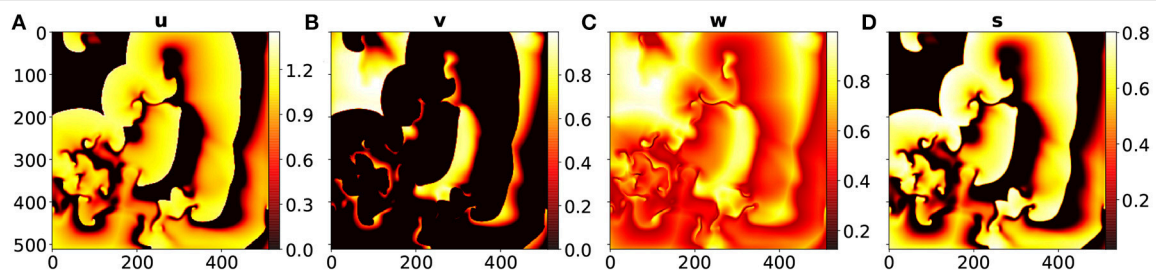
$$w_\infty = (1 - H(u - \theta_o)) (1 - \frac{u}{\tau_{w\infty}}) + H(u - \theta_o) w_\infty^*$$

are required. The characteristic model dynamics is determined through 28 parameters. In our simulations we used a set of parameters [11] given in **Table 1** for which the BOCF model exhibits chaotic excitation wave dynamics similar to the *Ten Tusscher-Noble-Noble-Panfilov* (TNNP) model [32].

The spatio-temporal chaotic dynamics of this system is actually transient chaos whose lifetime grows exponentially with system size [33, 34]. To obtain chaotic dynamics with a sufficiently long lifetime the system has been simulated on a domain of 512×512 grid points with a grid constant of $\Delta x = 1.0$ space units and a diffusion

TABLE 1 | TNNP model parameter values for the BOCF model [11].

u_o	0	τ_{v2}^-	1150	τ_{fi}	0.11	τ_{s1}	2.7342
u_u	1.58	τ_{v1}^+	1.4506	τ_{o1}	6	τ_{s2}	3
θ_v	0.3	τ_{w1}^-	70	τ_{o2}	6	k_s	2.0994
θ_w	0.015	τ_{w2}^-	20	τ_{so1}	43	u_s	0.9087
θ_v^-	0.015	k_w^-	65	τ_{so2}	0.2	τ_{si}	2.8723
θ_o	0.006	u_w^-	0.03	k_{so}	2	$\tau_{w\infty}$	0.07
τ_{v1}^-	60	τ_w^+	280	u_{so}	0.65	w_{∞}^*	0.94

**FIGURE 2** | Snapshots from the BOCF model at $t = 100$ of (A) the u variable, (B) the v variable, (C) the w variable, and (D) the s variable.

constant $D = 0.2$. Furthermore, an explicit Euler stepping in time with $\Delta t = 0.1$ time units¹, a 5 point approximation of the Laplace operator, and no-flux boundary conditions were used. **Figure 2** shows typical snapshots of the dynamics.

4. RESULTS

The proposed network model was trained with simulated data generated by the BOCF model with parameter values given in **Table 1**. Ten trajectories with different initial conditions for the variables u, v, w , and s were generated by simulating the BOCF model for a time series of 50,000 samples spanning a period of time of 5 s. Five of these data sets randomly chosen, were used to train the network, while the other solutions were used for validation. For training the Adam optimizer [35] was used, with a learning rate $lr = 0.0001$ and $\beta_1 = 0.9$, $\beta_2 = 0.999$.

In order to quantify the performance of the estimation and prediction methods the similarity of target fields and output fields of the network has to be quantified. For this purpose we use the Jensen-Shannon divergence (JSD) [36] applied to normalized fields of the variables of the BOCF model. The JSD of two discrete probability distributions A and B is defined as

$$JSD(A||B) = \frac{1}{2}D_{KL}(A||M) + \frac{1}{2}D_{KL}(B||M), \quad (13)$$

¹We consider all variables and parameters of the BOCF model as dimensionless. The parameter values given in **Table 1** are, however, consistent with the choice of a time unit equalling 1ms. In this case all t -values given in this article would correspond to milli seconds.

where $M = \frac{1}{2}(A + B)$ and $D_{KL}(A||M)$ is the Kullback-Leibler divergence [37]:

$$D_{KL}(A||M) = - \sum_i P(i) \log \left(\frac{A(i)}{M(i)} \right). \quad (14)$$

During training the JSD was used as objective function to be minimized (for a GPU implementation of the JSD see [38]). The JSD is bounded by 0 and 1 and a value below 0.02 was considered to indicate no discernible differences between the two distributions (fields). An alternative for quantifying the deviation would be the Fractions Brier Score [39]. For training the network, for each trajectory at each time step, sequences of lengths up to $m = 10$ were used as input.

The input of the network consisted of fields of variables that were assumed to be measured and random fields representing variables that were considered to be not available.

4.1. Forecast

For forecasting the input of the network consisted of sequences of length $m = 10$ of u, v, w , and s given by $\{u_{t-m+1}, \dots, u_t\}$, $\{v_{t-m+1}, \dots, v_t\}$, $\{w_{t-m+1}, \dots, w_t\}$, and $\{s_{t-m+1}, \dots, s_t\}$. The desired output of the network is then $u_{t+\Delta t}$, $v_{t+\Delta t}$, $w_{t+\Delta t}$ and $s_{t+\Delta t}$. By using the output of the network as a new input the system can be run iteratively in a closed loop for long term prediction. The development of the JSDs of u, v, w, s through time are shown in **Figure 3A**. Since the u and the s fields look quite similar (see **Figures 2A,D**) their JSD-values are almost the same. The w -field (**Figure 2C**) exhibits relatively high values at all spatial locations and therefore the JSD of two such fields is rather low. On the other hand, the v field (**Figure 2B**) possesses only very localized structures with high values and this leads to rather high values of the JSD for (slightly) different fields. **Figure 3B**

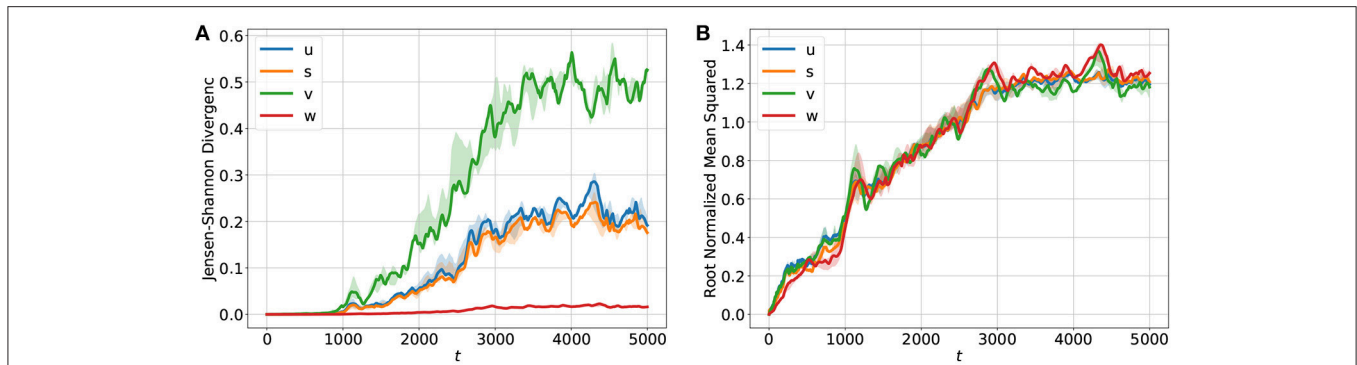


FIGURE 3 | Temporal development of (A) the Jensen-Shannon divergence (JSD) and (B) the root normalized mean squared error (RNMSE) for all variables u, v, w, s showing the deviation of the iterative network prediction (in a feedback mode) from the reference orbit obtained with the BOCF model. During the period $[0 - 1000]$ the predicted and the true fields agree very well as indicated by very small values of the JSD. In the time interval $(1000 - 3000)$ the JSD values increase until they saturate and the forecasts become very poor and useless. The RNMSE values show a similar increase in time but turn out to be more sensitive to minor deviations during the initial phase $[0 - 1000]$ of the forecast. The solid curves show median values of JSD and RNMSE obtained from ten different initial values of u, v, w, s . The transparent areas visualize the 0.25/0.75 percentile.

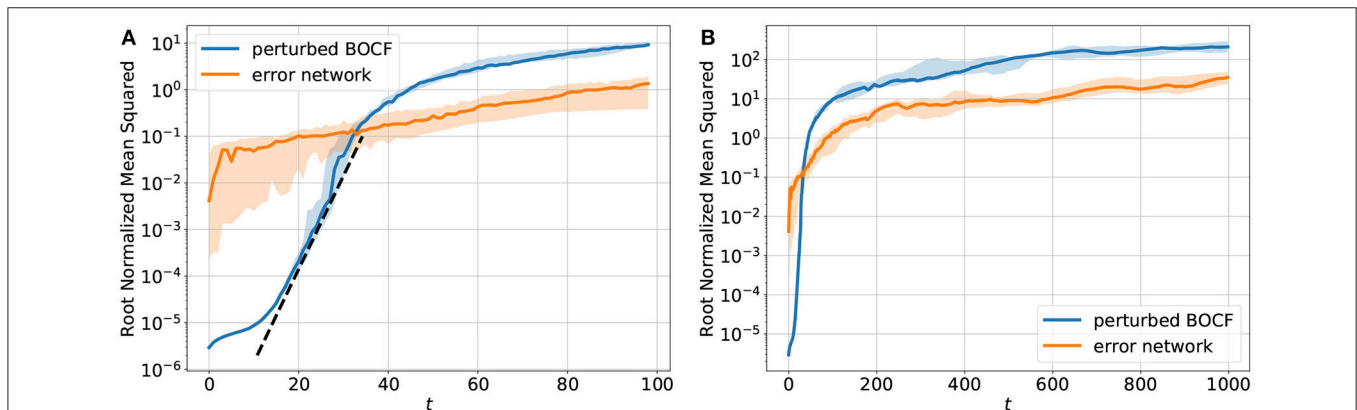


FIGURE 4 | Temporal development of the sum of the root normalized mean squared errors (RNMSE) of all variables u, v, w, s . (A) shows the NMSE for $t \in [1, 100]$ and (B) shows the NMSE for $t \in [1, 1000]$. The orange curve describes the deviation of the trajectory generated by the network from the reference orbit simulated with the BOCF model. For comparison the blue curve shows the distance between the reference orbit and a second solution of the BOCF model obtained by perturbing the initial conditions where each variable was perturbed at every spatial location using Gaussian random noise ($\mu = 0, \sigma^2 = 10^{-11}$). The error dynamics of ten perturbed trajectories was analyzed. These orbits were obtained by perturbing the reference orbit at different times $[0, 1000], [1000, 2000], \dots, [9000, 10000]$. The blue curve shows the median and the 0.25/0.75 percentile is visualized by the transparent areas. The dotted black line (A) denotes the slope the linear part of the $\log(\text{NMSE})$ vs. t curve which provides an estimate of the largest Lyapunov exponent [40] $\lambda_1 \approx 0.25$ (with respect to the natural logarithm).

shows for comparison the root normalized mean squared errors (RNMSE) of all variables u, v, w, s which is given by

$$\text{RNMSE}(v) = \sqrt{\frac{\text{MSE}(v)}{\text{MSE}(\bar{v})}} \quad (15)$$

where

$$\text{MSE}(v) = \frac{1}{M^2} \sum_{i=1}^M \sum_{j=1}^M \left(v_{ij}^{\text{BOCF}}(t) - v_{ij}(t) \right)^2. \quad (16)$$

Here \bar{v} denotes the temporal and spatial mean values of the BOCF sequence of length T_F , $M^2 = 512 \cdot 512$ is the number of grid points of the domain and v_{ij}^{BOCF} denotes the value of variable v at

grid point (i, j) for the reference solution generated by the BOCF model. As can be seen in Figure 3A all four curves possess very similar values and indicate an increase of the error already during the initial period for $t \in [0, 1000]$.

Figure 4 shows a comparison of the error dynamics of the forecast obtained with the iterated network with feedback (orange curve) and the dynamics of a BOCF model starting from slightly perturbed initial conditions (blue curve). Both curves give the root normalized mean squared error (RNMSE) with respect to the same reference orbit generated by the BOCF model. The perturbation of the initial condition of the second BOCF solution with respect to the initial condition of the reference orbits was chosen to be very small. Therefore, during the initial phase the deviation still remains so small that (with semi-logarithmic axes) a linear segment of the error curve occurs that

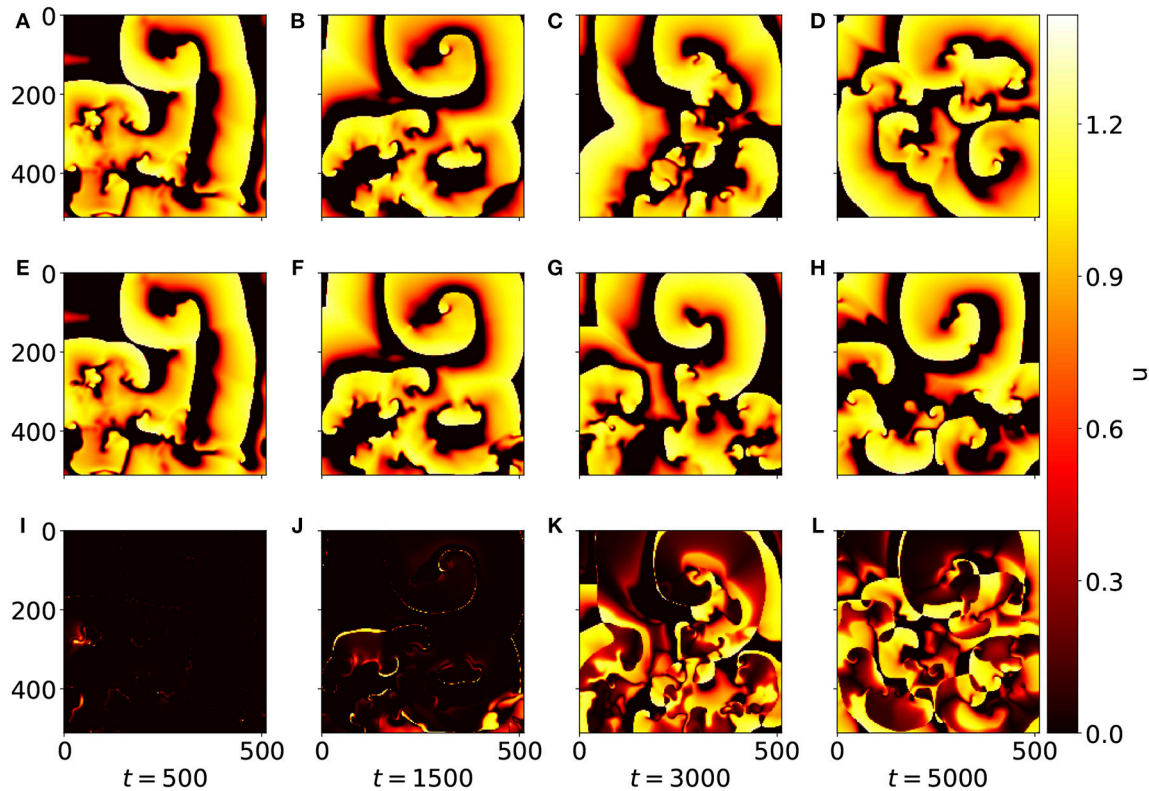


FIGURE 5 | Snapshots of u at different time steps. (A–D) Show the (reference) values from the BOCF simulation, while (E–H) display the values forecasted by the network. The diagrams (I–L) show the absolute deviation of the forecasted values from the reference values. At $t = 500$ the patterns (A,E) are still (almost) indistinguishable, and for $t = 1,500$ still only minor differences between (B,F) are noticeable.

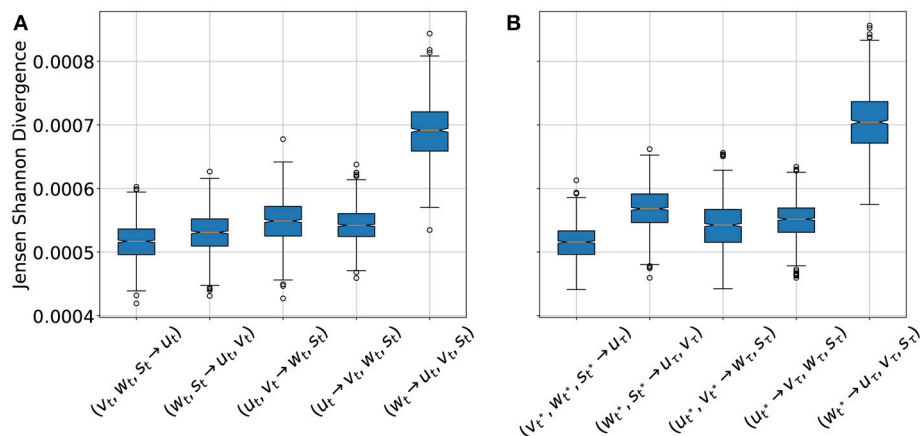
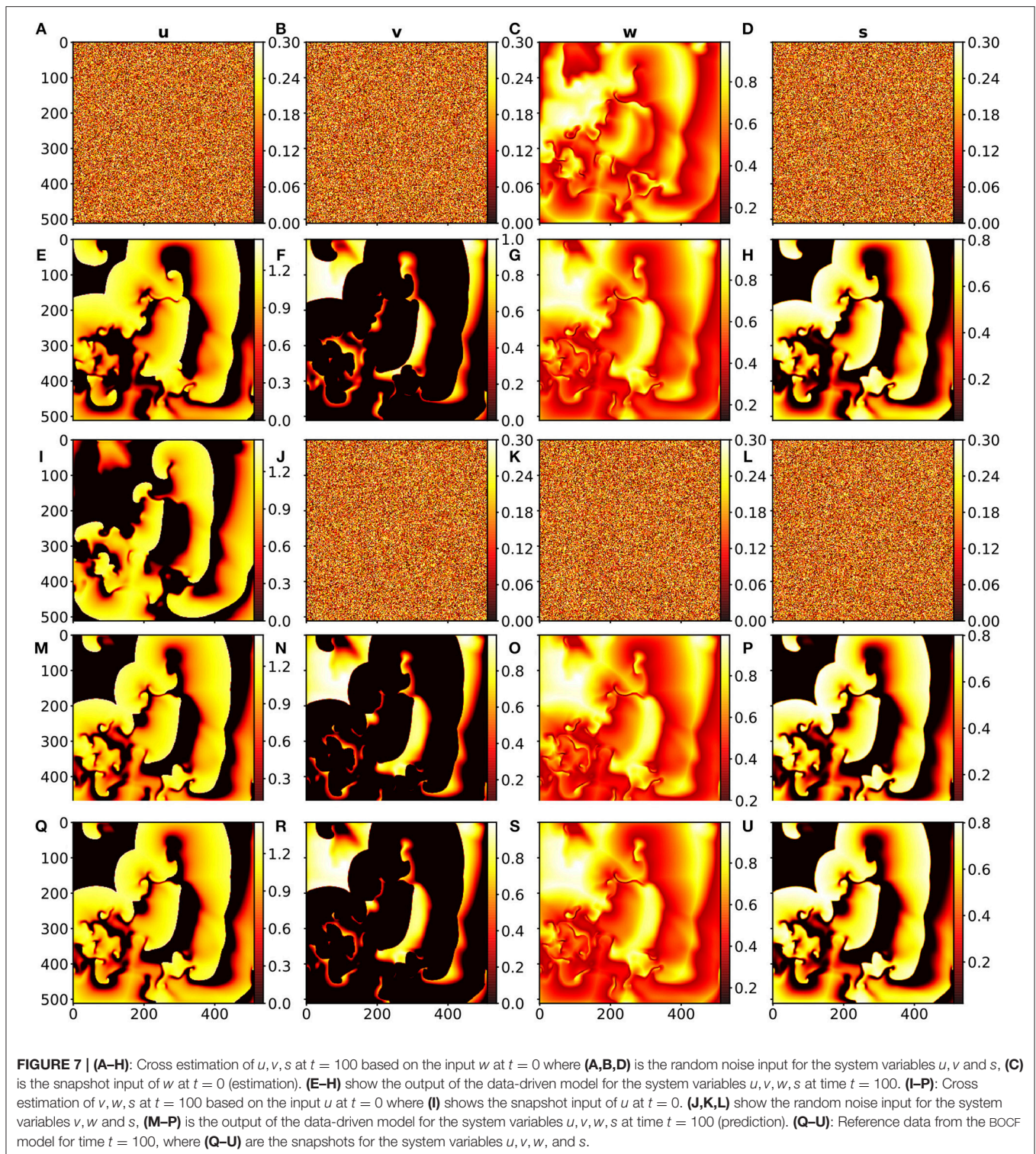


FIGURE 6 | Jensen-Shannon-Divergence (JSD) of true and estimated fields for different cross estimation tasks. In cases where more than one variable is estimated the mean value of the JSDs of the estimated variables is given. (A) Cross estimation for the cases $(V_t, W_t, S_t \rightarrow U_t)$, $(W_t, S_t \rightarrow U_t, V_t)$, $(U_t, V_t \rightarrow W_t, S_t)$, $(U_t \rightarrow V_t, W_t, S_t)$, and $(W_t \rightarrow U_t, V_t, S_t)$, based on the input from the BOCF simulation. (B) Cross estimation of future values of not measured variables for the cases $(V_{t^*}, W_{t^*}, S_{t^*} \rightarrow U_t)$, $(W_{t^*}, S_{t^*} \rightarrow U_t, V_t)$, $(U_{t^*}, V_{t^*} \rightarrow W_t, S_t)$, $(U_{t^*} \rightarrow V_t, W_t, S_t)$, and $(W_{t^*} \rightarrow U_t, V_t, S_t)$ based on the forecast of the data driven model for a period of $\tau = 1,000$, where t^* denotes 10 successive snapshots at times $0, 0.1, \dots, 0.9$ constituting the input. In both diagrams the orange line is the median value for each case, the box extends from the lower to upper quartile values. The whiskers extend from the box to show the range of the data. Flier points are those past the end of the whiskers.



can be used to estimate the largest Lyapunov exponent [40]. Once the error of the perturbed BOCF orbit (blue curve) reaches the level of the network prediction error (orange curve) both error curves continue to increase in the same way indicating that the network almost perfectly learned the true dynamics of the BOCF model.

To illustrate the deviation between the u field forecasted by the network and the (true) u field provided by the simulation of the BOCF PDE **Figure 5** shows snapshots at times $t = 500$, $t = 1,500$, $t = 3,000$, and $t = 5,000$. While at $t = 500$ original (A) and forecast (E) are almost indistinguishable the snapshots at $t = 1,500$ exhibit minor differences (**Figures 5B,F**). At time

$t = 3,000$ only rough structures agree (**Figures 5C,G**) until at $t = 5,000$ forecast and simulation appear completely decorrelated (**Figures 5D,H**). The full evolution of the forecast compared to the original dynamics generated with the BOCF model is also available as a movie (**Supplemental Data**). Compared to a typical spiral rotation period of approximately $T_{sp} = 350$ good forecasting results can be obtained for about five spiral rotations corresponding to $5 \cdot 350 / 4 = 437$ Lyapunov times $T_L = 1/\lambda_1 \approx 4$ given by the largest Lyapunov exponent $\lambda_1 \approx 0.25$ (see **Figure 4**).

4.2. Cross-Estimation

For cross-estimation only a part of the system variables are considered as being directly observable or measurable. Based on these available variables the other not measurable variables have to be estimated (a task also called cross prediction). In the context of the BOCF model we shall, for example, estimate v_t, w_t, s_t from observations of u_t , only. Since the network expects all system variables as input the not observed variables were replaced by uniform noise in the range of $0 - 0.3$. For this purpose for every $t \in [0, 1000]$ the data of the BOCF model were used as single time step input for the network and the cases $(v_t, w_t, s_t \rightarrow u_t)$, $(w_t, s_t \rightarrow u_t, v_t)$, $(u_t, v_t \rightarrow w_t, s_t)$, $(u_t \rightarrow v_t, w_t, s_t)$, and $(w_t \rightarrow u_t, v_t, s_t)$ were considered as estimation tasks. **Figure 6** shows the JSD statistics for all these cases. The low JSD values for $(v_t, w_t, s_t \rightarrow u_t)$ indicated that the variable u can be very well estimated by the variables v, w, s , which could be expected because the variable u is part of the PDEs of the other variables. Similarly good estimation results are obtained for $(u_t \rightarrow v_t, w_t, s_t)$ which is remarkable, because the membrane potential u is the variable, which can be measured most easily in experiments and the result shows that this information is sufficient to recover the other variables v, w , and s of the BOCF model. The worst performance is achieved if only w is used to cross estimate all other system variables. These cross estimation results are in very good agreement with the performance of an Echo State Network applied to similar data [19].

4.3. Forecast and Cross-Estimation

This investigation represents a combination of the two previous ones. In this case, however, not for every time step the data from the BOCF model were used, but only ten time steps from the BOCF model were used to initialize the forecast of the network. Depending on the case which variable should be estimated the BOCF variables for initialization were replaced by uniform noise, as before. **Figure 6B** shows the JSD statistics for the four estimation cases considered and in **Figure 7** snapshots of the input and the true and estimated fields are presented illustrating the very good performance at time $t = 100$.

5. DISCUSSION

Spatio-temporal non-linear dynamical systems like extended systems (described by PDEs) or networks of interacting

oscillators may exhibit very high dimensional chaotic dynamics. A typical example are complex wave pattern occurring in some excitable media. As a representative of this class of systems we used the BOCF model describing electrical excitation waves in cardiac tissue where chaotic dynamics is associated with cardiac arrhythmias. For future applications like monitoring and predicting the dynamical state of the heart or the impact of interventions, mathematical models are required describing the temporal evolution or the relation between different (physical) variables. As an alternative to the large number of simple qualitative or detailed (ionic) models (incorporating many biophysical details and corresponding variables) we presented a machine learning approach for data driven modeling of the spatio-temporal dynamics. A convolutional neural network combined with a linear-chain of conditional random fields was trained and validated with data generated by a simulation of the BOCF model. To mimic experimental limitations when measuring cardiac dynamics we considered different cases where only some of the variables of the BOCF model were assumed to be available as input of the generated model and the not measurable variables were replaced by random numbers. Running the trained network in a closed loop (feedback) configuration iterated prediction provided forecasts of the complex dynamics that turned out to follow the true (chaotic!) evolution of the BOCF simulation for about five periods of the intrinsic spiral rotations. These results clearly show that machine learning methods like those employed here provide faithful models of the underlying complex dynamics of excitable media that, when suitably trained can provide powerful tools for predicting the spatio-temporal evolution and for cross-estimating not directly observed variables.

AUTHOR CONTRIBUTIONS

SH performed numerical simulations. UP and SH identified the scientific topic. UP, FW, and SH devised the strategy for solving it, and wrote the manuscript.

FUNDING

SH acknowledges funding by the International Max Planck Research Schools of Physics of Biological and Complex Systems.

ACKNOWLEDGMENTS

We thank T. Lilienkamp and S. Luther for support with the BOCF model and for inspiring discussions about spatio-temporal dynamics in excitable media.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fams.2018.00060/full#supplementary-material>

Supplemental Data | Movie showing the temporal evolution of the u field from the simulation, the forecast and the absolute difference of both.

REFERENCES

- Barkley D. A model for fast computer simulation of waves in excitable media. *Physica D Nonlinear Phenomena* (1991) **49**:61–70.
- Bär M, Eiswirth M. Turbulence due to spiral breakup in a continuous excitable medium. *Phys Rev E* (1993) **48**:R1635–7.
- Cherry EM, Fenton F, Krogh-Madsen T, Luther S, Parlitz U. Introduction to focus issue: complex cardiac dynamics. *Chaos* (2017) **27**:093701. doi: 10.1063/1.5003940
- Clayton RH, Bernus O, Cherry EM, Dierckx H, Fenton FH, Mirabella L, et al. Models of cardiac tissue electrophysiology: progress, challenges and open questions. *Prog Biophys Mol Biol.* (2011) **104**:22–48. doi: 10.1016/j.pbiomolbio.2010.05.008
- Uzelac I, Ji YC, Hornung D, Schröder-Schelling J, Luther S, Gray RA, et al. Simultaneous quantification of spatially discordant alternans in voltage and intracellular calcium in langendorff-perfused rabbit hearts and inconsistencies with models of cardiac action potentials and Ca transients. *Front Physiol.* (2017) **8**:819. doi: 10.3389/fphys.2017.00819
- Evensen G, van Leeuwen P. An ensemble Kalman smoother for nonlinear dynamics. *Mon Weather Rev.* (2000) **128**:1852–67. doi: 10.1175/1520-0493(2000)128<1852:AEKSFN>2.0.CO;2
- Abarbanel HDI. *Predicting the Future - Completing Models of Observed Complex Systems.* New York, NY: Springer-Verlag (2013).
- Law K, Stuart A, Zygalakis K. *Data Assimilation - A Mathematical Introduction.* Cham: Springer International Publishers (2015).
- Hoffman MJ, LaVigne NS, Scorse ST, Fenton FH, Cherry EM. Reconstructing three-dimensional reentrant cardiac electrical wave dynamics using data assimilation. *Chaos* (2016) **26**:013107. doi: 10.1063/1.4940238
- Berg S, Luther S, Parlitz U. Synchronization based system identification of an extended excitable system. *Chaos* (2011) **21**:033104. doi: 10.1063/1.3613921
- Bueno-Orovio A, Cherry EM, Fenton FH. Minimal model for human ventricular action potentials in tissue. *J Theor Biol.* (2008) **253**:544–60. doi: 10.1016/j.jtbi.2008.03.029
- Box GE, Jenkins GM, Reinsel GC, Ljung GM. *Time Series Analysis: Forecasting and Control.* San Francisco, CA: John Wiley & Sons (2015).
- Solomatine DP, Ostfeld A. Data-driven modelling: some past experiences and new approaches. *J Hydroinformatics* (2008) **10**:3–22. doi: 10.2166/hydro.2008.015
- Parlitz U, Merkwirth C. Prediction of spatiotemporal time series based on reconstructed local states. *Phys Rev Lett.* (2000) **84**:1890. doi: 10.1103/PhysRevLett.84.1890
- Jaeger H. *The 'Echo State' Approach to Analysing and Training Recurrent Neural Networks - With an Erratum Note.* GMD Report (2001).
- Jäger H, Haas H. Harnessing nonlinearity: predicting chaotic systems and saving energy in wireless communication. *Science* (2004) **304**:78–80. doi: 10.1126/science.1091277
- Verstraeten D, Schrauwen B, D'Haene M, Stroobandt D. An experimental unification of reservoir computing methods. *Neural Netw.* (2007) **20**:391–403. doi: 10.1016/j.neunet.2007.04.003
- Pathak J, Hunt B, Girvan M, Lu Z, Ott E. Model-free prediction of large spatiotemporally chaotic systems from data: a reservoir computing approach. *Phys Rev Lett.* (2018) **120**:024102. doi: 10.1103/PhysRevLett.120.024102
- Zimmermann RS, Parlitz U. Observing spatio-temporal dynamics of excitable media using reservoir computing. *Chaos* (2018) **28**:043118. doi: 10.1063/1.5022276
- Quade M, Abel M, Nathan Kutz J, Brunton SL. Sparse identification of nonlinear dynamics for rapid model recovery. *Chaos* (2018) **28**:063116. doi: 10.1063/1.5027470
- Abebe A, Solomatine D, Venneker R. Application of adaptive fuzzy rule-based models for reconstruction of missing precipitation events. *Hydrol Sci J.* (2000) **45**:425–36. doi: 10.1080/02626660009492339
- Abebe AJ, Guinot V, Solomatine DP. Fuzzy alpha-cut vs. Monte Carlo techniques in assessing uncertainty in model parameters. In: *Proc. 4-th International Conference on Hydroinformatics.* Iowa City, IA (2000). p. 1–8.
- Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Advances in Neural Information Processing Systems.* Curran Associates Inc. (2012). p. 1097–105.
- Lafferty JD, McCallum A, Pereira FCN. Conditional Random Fields: Probabilistic Models for Segmenting and Labeling Sequence Data. In: *Proceedings of the Eighteenth International Conference on Machine Learning. ICML '01.* San Francisco, CA: Morgan Kaufmann Publishers Inc. (2001). p. 282–9.
- Vemuri V. *Artificial Neural Networks.* Rockville, MD: Computer Science Press Inc. (1988).
- LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* (2015) **521**:436–44. doi: 10.1038/nature14539
- Schmidhuber J. Deep learning in neural networks: an overview. *Neural Netw.* (2015) **61**:85–117. doi: 10.1016/j.neunet.2014.09.003
- Hinton GE, Salakhutdinov RR. Reducing the dimensionality of data with neural networks. *Science* (2006) **313**:504–7. doi: 10.1126/science.1129198
- He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. *arXiv:1512.03385* (2015).
- Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. *arXiv:1505.04366* (2015).
- Chatzis SP, Demiris Y. The infinite-order conditional random field model for sequential data modeling. *IEEE Trans Pattern Anal Mach Intell.* (2013) **35**:1523–34. doi: 10.1109/TPAMI.2012.208
- Ten Tusscher K, Noble D, Noble P, Panfilov AV. A model for human ventricular tissue. *Am J Physiol Heart Circ Physiol.* (2004) **286**:H1573–89. doi: 10.1152/ajpheart.00794.2003
- Strain MC, Greenside HS. Size-dependent transition to high-dimensional chaotic dynamics in a two-dimensional excitable medium. *Phys Rev Lett.* (1998) **80**:2306–9. doi: 10.1103/PhysRevLett.80.2306
- Lilienkamp T, Christoph J, Parlitz U. Features of chaotic transients in excitable media governed by spiral and scroll waves. *Phys Rev Lett.* (2017) **119**:054101. doi: 10.1103/PhysRevLett.119.054101
- Kingma DP, Ba J. Adam: a method for stochastic optimization. *arXiv preprint arXiv:1412.6980* (2014).
- Lin J. Divergence measures based on the Shannon entropy. *IEEE Trans Inform Theor.* (1991) **37**:145–51.
- Kullback S, Leibler RA. On information and sufficiency. *Ann Math Stat.* (1951) **22**:79–86.
- Gültas M, Düzgün G, Herzog S, Jäger SJ, Meckbach C, Wingender E, et al. Quantum coupled mutation finder: predicting functionally or structurally important sites in proteins using quantum Jensen-Shannon divergence and CUDA programming. *BMC Bioinformatics* (2014) **15**:96. doi: 10.1186/1471-2105-15-96
- Roberts N. Assessing the spatial and temporal variation in the skill of precipitation forecasts from an NWP model. *Meteorol Appl.* (2008) **15**:163–9. doi: 10.1002/met.57
- Parlitz U. Estimating Lyapunov Exponents from Time Series. In: Skokos C, Gottwald G, Laskar J, editors. *Chaos Detection and Predictability. Lecture Notes in Physics.* Vol 915. Berlin; Heidelberg: Springer (2016). p. 1–34.

Conflict of Interest Statement: 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.

Copyright © 2018 Herzog, Wörgötter and Parlitz. 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.



Ultra Rapid Data Assimilation Based on Ensemble Filters

Roland Potthast^{1,2*} and Christian A. Welzbacher¹

¹ Deutscher Wetterdienst, Data Assimilation Unit, Offenbach, Germany, ² Department of Mathematics, University of Reading, Reading, United Kingdom

OPEN ACCESS

Edited by:

Wilhelm Stannat,
Technische Universität Berlin,
Germany

Reviewed by:

Xin Tong,
National University of Singapore,
Singapore
Meysam Hashemi,
INSERM U1106 Institut de
Neurosciences des Systèmes, France

*Correspondence:

Roland Potthast
roland.potthast@dwd.de

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 16 July 2018

Accepted: 14 September 2018

Published: 23 October 2018

Citation:

Potthast R and Welzbacher CA (2018)
Ultra Rapid Data Assimilation Based
on Ensemble Filters.
Front. Appl. Math. Stat. 4:45.
doi: 10.3389/fams.2018.00045

The goal of this work is to analyse and study an *ultra-rapid data assimilation* (URDA) method for adapting a given ensemble forecast for some particular variable of a dynamical system to given observation data which become available after the standard data assimilation and forecasting steps. Initial ideas have been suggested and tested by Etherthon 2006 and Madaus and Hakim 2015 in the framework of numerical weather prediction. The methods are, however, much more universally applicable to general non-linear dynamical systems as they arise in neuroscience, biology and medicine as well as numerical weather prediction. Here we provide a full analysis in the linear case, we formulate and analyse an ultra-rapid ensemble smoother and test the ideas on the Lorentz 63 dynamical system. In particular, we study the *assimilation and preemptive forecasting step* of an ultra-rapid data assimilation in comparison to a full ensemble data assimilation step as calculated by an ensemble Kalman square root filter. We show that for linear systems and observation operators, the ultra-rapid assimilation and forecasting is equivalent to a full ensemble Kalman filter step. For non-linear systems this is no longer the case. However, we show that we obtain good results even when rather strong nonlinearities are part of the time interval $[t_0, t_n]$ under consideration. Then, an *ultra-rapid ensemble Kalman smoother* is formulated and numerically tested. We show that when the numerical model under consideration is different from the true model, used to generate the nature run and observations, errors in the correlations will also lead to errors in the smoother analysis. The numerical study is based on the popular *Lorenz 1963 model system* used in geophysics and life sciences. We investigate both the situation where the full system forecast is calculated and the situation important to practical applications where we study *reduced data*, when only one or two variables are known to the URDA scheme.

Keywords: data assimilation (DA), ensemble filter, preemptive forecast, Lorenz 1963 system, rapid update

1. INTRODUCTION

Data assimilation is concerned with the use of observation data to control or determine the state of some dynamical system [1–3]. Data assimilation methods are indispensable ingredients to calculate forecasts of some system, with universal applicability ranging from neuroscience [4, 5] to weather forecasting [6–8], from systems engineering like traffic flow [9–11] to geophysical applications [12, 13].

Over time several generations of data assimilation methods have been developed, for example optimal interpolation in the 70th, variational methods in the 80 and 90th and ensemble data

assimilation since about 1995, with very intense research activities since about 2000 (e.g., [1–3, 14]). Today, ensemble data assimilation methods for example for numerical weather prediction are run daily on modern supercomputers by operational centers such as *Deutscher Wetterdienst* (DWD, Germany), *European Center for Medium Range Weather Forecast* (ECMWF, Reading, UK), or the *MetOffice* in the UK.

Usually, data assimilation takes *observations* y and combines them with *first guess* model states $x^{(b)}$ (also called *the background*) to estimate a best possible *analysis* $x^{(a)}$. Usually, the estimation of the analysis state is performed in turns with short-range forecasting, i.e., within a given temporal framework assimilations are carried out at times t_k for $k = 1, 2, 3, \dots$. Short range forecasts calculate the state $x_k^{(b)} = M(x_{k-1}^{(a)})$ by applying the model dynamics M to the initial state given by the analysis $x_{k-1}^{(a)}$ at time t_{k-1} . Then, the core analysis step is carried out at time t_k , based on observations which are available either at t_k or in the interval $[t_{k-1}, t_k]$. Alternating short-range forecasts and core analysis steps leads to the classical *data assimilation cycle*. Often, forecasts are then calculated based on selected analysis states and analysis times.

Today, many forecasting systems have moved away from pure deterministic forecasting and employ *ensemble prediction systems* (EPS), where several forecasts with different initial conditions (and sometimes different physical or stochastic parameters) are calculated. Based on an ensemble of states, the uncertainty of the forecast can be estimated. Further, the ensemble allows to determine dynamical spatial and temporal correlations, which help to improve the analysis itself and can serve as input for probabilistic diagnostics.

Often, for large-scale realistic systems, the model forecast as well as the analysis step needs huge computational resources. They limit the temporal resolution of the data assimilation cycle. Further restrictions are given by the availability of observations, which need to be measured and distributed to reach operational centers. For example, to run an assimilation cycle of 1 h for convection permitting high-resolution numerical weather models, top-500 supercomputers are needed to achieve a sufficient resolution and spatial extension of the model fields under consideration [8].

The core task addressed in this work is the problem of *ultra-rapid data assimilation* (URDA), in the case where standard data assimilation cycles have clear limits with respect to speed and flexibility. We assume that a classical data assimilation cycle is available, such that we can calculate an ensemble of forecasts for some time interval $[t_0, t_N]$. The next classical analysis is calculated for time t_N , such that a similar ensemble forecast will be available for a subsequent interval $[t_N, t_{N+1}]$. Here, we limit our interest in the ultra-rapid data assimilation for observations y_k which are available at points in time t_k with $t_0 < t_1 < t_2 < \dots \leq t_N$. The task is to provide an update of the ensemble forecast with high speed without using the full numerical model or a full-grown data assimilation system. In particular when we are interested only in the forecast of some layer or part of the state space, this is of high practical interest.

Usually, the classical forecast cycle in operational centers is based on a data assimilation cycle with frequency t_N of several hours. The term *rapid update cycle* (RUC) is used when cycling and forecasting is carried out hourly or subhourly. The term *ultra-rapid update cycle* is used when we go to a cycling interval which is much smaller, e.g., 5 min. Further, to achieve this speed we cannot initialize the full model in each step. The approach of ultra-rapid data assimilation—though embedded into a RUC or classical cycle—does not use the classical setup of cycling model and assimilation step any more for its updates. Further, it works with a subset of model variables only. The speed-up is achieved by the conceptional changes within the full cycle, not alone within the data assimilation step itself.

We will base ultra-rapid data assimilation on the ensemble transformation matrix given by the ensemble Kalman filter (EnKF) or ensemble Kalman square root filter (SRF), compare [15–19]. The basic idea of ultra-rapid data assimilation is to employ a reduced version of the state variables which are made available to the system. Measurements of some of these variables can be employed to calculate an ensemble Kalman transformation matrix¹. This Matrix is used to update both the analysis ensemble as well as the forecast ensemble. For linear model systems and linear observation operators, we will see that the forecasts based on the analysis ensemble and the transformed forecast ensemble are identical. This is true both for the full analysis and forecasting as well as for the case where we base our analysis and forecasting transformation on a reduced set of model variables or diagnostic ensemble output.

To study the quality of ultra-rapid data assimilation we apply the basic ideas to the Lorenz 1963 model [[20], see also for example [21–24] and [3] Chapter 6]. Here, we generate some truth by running the model with a particular setup. Observations are simulated and drawn with random perturbations. Then, a model with a different setup is used to assimilate the observations either with the ultra-rapid assimilation scheme and for comparison by running a full ensemble Kalman filter for each of the time-steps t_k , $k = 1, \dots, N$. We study the case of reduced variables and provide diagnostic results for the ensemble Kalman smoother over the full time interval $[t_0, t_N]$.

The approach discussed here was first suggested in the work of Etherton [25], where the term *preemptive forecast* was coined and the method was tested for a barotropic model. The ideas have been picked up by Madaus and Hakim [26], where the authors applied the approach to ensemble forecasts of numerical weather models, obtaining a so-called *ensemble forecast adjustment*. In the latter work the advantage of the method, that rapid updates of (subspaces of) model predictions without rerunning a full dynamical model again, are highlighted. They focused on global scales and corresponding time scales and observables for a particular application. Here, we provide additional contributions to the mathematical analysis for linear systems. Further, we formulate and investigate the ultra-rapid ensemble smoother and extensively study the non-linear Lorentz 63 system, which

¹Note that the calculation of this transform matrix takes place in ensemble space and is only a very small part of the total cost of the assimilation cycle and forecasting.

serves as a very popular reference system for geophysics and life sciences.

In section 2 we introduce our notation and basic results from ensemble data assimilation. In particular, we introduce the ensemble Kalman filter in the notation of Hunt et al. [16] and Nakamura and Potthast [3]. We also discuss the role of reduced variables for the ensemble Kalman square root filter. Section 3 serves to introduce and investigate details of the ultra-rapid data assimilation, with the data assimilation analysis and forecasting in section 3.1 and the ultra-rapid ensemble smoother in section 3.2. Numerical examples are shown in section 4, with generic results on the assimilation and forecasting in section 4.1 and the study of reduced variables in section 4.2. Conclusions are given in section 5.

2. ENSEMBLE DATA ASSIMILATION

This section serves to collect notation and basic results on the ensemble Kalman square root filter (SRF) following the notation of Hunt et al. [16] and Nakamura and Potthast [3]. The SRF is our reference for full-scale forecasting and it provides the core ingredients of our ultra-rapid data assimilation algorithms as described in section 3.

We consider a state space \mathbb{R}^n , an observation space \mathbb{R}^m with $n, m \in \mathbb{N}$, states $x \in \mathbb{R}^n$ and observations $y \in \mathbb{R}^m$. The basic idea of the ensemble Kalman filter type methods such as the SRF is to approximate the covariance matrix $B \in \mathbb{R}^{n \times n}$ of the system based on some ensemble $x^{b,(\ell)}$, $\ell = 1, \dots, L$ of $L \in \mathbb{N}$ states in the form

$$B^b = Q^b (Q^b)^T, \quad (1)$$

where

$$Q^b = \frac{1}{\sqrt{L-1}} (x^{b,(1)} - \bar{x}^b, \dots, x^{b,(L)} - \bar{x}^b), \quad (2)$$

is the matrix $Q^b \in \mathbb{R}^{n \times L}$ of centered differences (sometimes its columns are called the *centered ensemble*) with the *ensemble mean*

$$\bar{x}^b = \frac{1}{L} \sum_{\ell=1}^L x^{b,(\ell)}. \quad (3)$$

Note, by construction the space spanned by the member of the centered ensemble has dimension $L-1$ and one can define the full ensemble matrix

$$\begin{aligned} Q_{full}^b &= (x^{b,(1)}, \dots, x^{b,(L)}), \\ &= \bar{x}^b + \sqrt{L-1} Q^b. \end{aligned} \quad (4)$$

In order to assimilate observation data the *model equivalents* $y^{b,(\ell)}$ of the ensemble member are required, which are obtained by applying the observation operator $H: \mathbb{R}^n \rightarrow \mathbb{R}^m$ to the corresponding ensemble member

$$y^{b,(\ell)} = H(x^{b,(\ell)}), \quad \bar{y}^b = \frac{1}{L} \sum_{\ell=1}^L y^{b,(\ell)}. \quad (5)$$

With these quantities the matrix $T^b \in \mathbb{R}^{m \times L}$ can be defined analogously to Q^b

$$T^b := \frac{1}{\sqrt{L-1}} (y^{b,(1)} - \bar{y}^b, \dots, y^{b,(L)} - \bar{y}^b), \quad (6)$$

which one also denotes as $T^b = HQ^b$ assuming a linear operator H .

In the following, $+$ between a vector and a matrix indicates a column-wise summation $a + A = (a + a_1, \dots, a + a_L)$ with the columns a_ℓ , $\ell = 1, \dots, L$ of the matrix A , such that we can add column vectors and matrices in one joint notation. The generic update equation for an ensemble type data assimilation can be written in different forms, in particular

$$Q_{full}^a = \bar{x}^a + \sqrt{L-1} Q^a, \quad (7)$$

$$= \bar{x}^a + \sqrt{L-1} Q^b S, \quad (8)$$

$$= \bar{x}^b + Q^b \bar{s} + \sqrt{L-1} Q^b S, \quad (9)$$

$$= \bar{x}^b + Q^b (\bar{s} + \sqrt{L-1} S), \quad (10)$$

$$= \bar{x}^b + Q^b W, \quad (11)$$

$$= Q_{full}^b W_{full}, \quad (12)$$

with the transformation matrices $S, W_{full}, W \in \mathbb{R}^{L \times L}$ computed in ensemble space and $\bar{s} \in \mathbb{R}^L$, depending on the vectors and matrices $\bar{x}^b, Q^b, \bar{y}^b, T^b$ and the observation error correlation matrix $R \in \mathbb{R}^{m \times m}$. We quickly review the different versions as follows. An analysis update of the centered ensemble (see Equation 4) given by

$$Q^a = Q^b S, \quad (13)$$

leading to Equation(8). In Equation (9) we have used an update of the ensemble mean

$$\bar{x}^a - \bar{x}^b = Q^b \bar{s}, \quad (14)$$

with $\bar{s} \in \mathbb{R}^L$, which is naturally defined by the Ensemble Kalman Filter – details will be given below. Equation (10) just collects the increment in terms of Q^b . The definition of the transformation matrix

$$W = \bar{s} + \sqrt{L-1} S, \quad (15)$$

leads to the update Equation (11). For the full transform matrix W_{full} we obtain

$$W_{full} = \frac{\bar{s}}{\sqrt{L-1}} + S, \quad (16)$$

based on

$$\underbrace{(1, \dots, 1)}_{L \text{ times}} \left(\frac{\bar{s}}{\sqrt{L-1}} + S \right) = 1, \quad (17)$$

and by

$$\begin{aligned} Q_{full}^b W_{full} &= (x^b + \sqrt{L-1} Q^b) \left(\frac{\bar{s}}{\sqrt{L-1}} + S \right), \\ &= \underbrace{(x^b, \dots, x^b)}_{L \text{ times}} \left(\frac{\bar{s}}{\sqrt{L-1}} + S \right) + Q^b \bar{s} + Q^b S, \\ &= x^b + Q^b \bar{s} + Q^b S. \end{aligned} \quad (18)$$

Different notations have been used over time, depending on whether you want to keep your equations close to the classical Kalman filter equations or for a more practical focus. The quantities defined in Equations (2, 6) differ from the definitions of X^b and Y^b of Hunt et al., c.f. Equations (12, 18), by the normalization factors. The relations are

$$\begin{aligned} X^b &= \sqrt{L-1} Q^b, \\ Y^b &= \sqrt{L-1} T^b. \end{aligned} \quad (19)$$

The full ensemble matrix has different letters $Q_{full}^b = X_{full}^b$, we also note the identity $W^a = S$ between Hunt et al. [16] and Nakamura and Potthast [3]. Some equations are modified. For example Equation (4) changes to

$$\begin{aligned} X_{full}^b &= (x^{b,(1)}, \dots, x^{b,(L)}) , \\ &= \bar{x}^b + X^b, \end{aligned} \quad (20)$$

the update of the mean Equation (14) rewrites as

$$\bar{x}^a - \bar{x}^b = X^b w_m, \quad (21)$$

with $w_m = \bar{s}/\sqrt{L-1}$ and Equations (8, 10) are written as

$$X^a = X^b W^a, \quad (22)$$

$$\Leftrightarrow X_{full}^a = \bar{x}^a + X^b W^a, \quad (23)$$

$$= \bar{x}^b + X^b (w_m + W^a), \quad (24)$$

$$= \bar{x}^b + X^b W_X. \quad (25)$$

The transformation matrix in the sense of Equation (25), giving us the increment in ensemble space, is now given by

$$W_X = w_m + W^a. \quad (26)$$

2.1. Ensemble Kalman Square Root Filter

The ensemble Kalman filter combines the above introduced notion of an ensemble of model states to describe spatial and temporal correlations with the well-known Kalman Filter [27]. The pending task of generating an analysis ensemble obeying an obtained analysis correlation matrix can be completed by a square root filter (SRF), originating from the Kalman filter update for the correlation matrix applied to the ensemble representation

$$\begin{aligned} B^a &= (I - KH) B^b, \\ Q^a (Q^a)^T &= (I - KH) Q^b (Q^b)^T, \\ &= Q^b U (Q^b)^T, \\ &= Q^b S (Q^b)^T, \end{aligned} \quad (27)$$

with the Kalman gain matrix

$$K = Q^b (HQ^b)^T \left(R + HQ^b (HQ^b)^T \right)^{-1}, \quad (28)$$

and the transformation matrix S given by

$$S(S)^T = U. \quad (29)$$

Taking the square root of the symmetric matrix U results in

$$S = \sqrt{I - (HQ^b)^T \left(R + HQ^b (HQ^b)^T \right)^{-1} HQ^b}, \quad (30)$$

which is the transformation matrix of the update for the centered ensemble in Equation (8). Note, the notation in Equations (2, 6) is the one used by Nakamura and Potthast [3] and differs from the one introduced by Hunt et al. [16] (see Equation 19). However, by multiplying S in Equation (30) with the inverse of W^a defined by Hunt et al, the identity $W^a = S$ can be easily shown.

The update of the mean is obtained along the lines of the classical Kalman filter by

$$\bar{x}^a - \bar{x}^b = K(y^0 - \bar{y}^b), \quad (31)$$

with K given in Equation (28). Comparing Equation (14, 31) leads to

$$\bar{s} = (HQ^b)^T \left(R + HQ^b (HQ^b)^T \right)^{-1} (y^0 - \bar{y}^b), \quad (32)$$

in case of the SRF. The update for the full ensemble using the ensemble Kalman square root filter is therefore given by applying Equations (30, 31) to Equation (11, 15).

Here, we can now confirm the validity of Equation (17). From the definition of Q^b we know that the sum of the rows of Q is zero, such that the sum of the column of $\bar{s} = (Q^b)^T A$ with any matrix $A \in \mathbb{R}^{n \times L}$ is zero, and the sum of the columns of $I - (Q^b)^T A$ is one. If $I - (Q^b)^T A$ is symmetric, this means that the vector equal to 1 in each component is an eigenvector of $I - (Q^b)^T A$ with eigenvalue 1. But then it will also be an eigenvector with eigenvalue 1 for each power of $I - (Q^b)^T A$, such that (17) is satisfied.

2.2. Ensemble Data Assimilation With Reduced Data

Before we investigate ultra-rapid data assimilation based on reduced data, we need to recall how a standard ensemble Kalman square root filter will react when we base our analysis on a reduced set of model variables. Let us study the calculation of the ensemble analysis for the ensemble Kalman filter with reduced data. The basic formula for the ensemble Kalman filter can be expressed as $W = S + \bar{s}$ with S and \bar{s} given in Equations (30, 32)

Now, assume we observe $y \in \mathbb{R}^m$ which depends on some subset $x_1, \dots, x_{\tilde{n}}$ of the full set of variables x_1, \dots, x_n only. Given these reduced spaces the operator H will be of the form

$$H = \begin{pmatrix} H_{1,1} & \dots & H_{1,\tilde{n}} & 0 & \dots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ H_{m,1} & \dots & H_{m,\tilde{n}} & 0 & \dots & 0 \end{pmatrix}. \quad (33)$$

In this case, the terms $HQ^b \in \mathbb{R}^{m \times L}$ and $(HQ^b)^T \in \mathbb{R}^{L \times m}$ will be a linear combination of the variables $1, \dots, \tilde{n}$ of the ensemble members. If we are given the variables $x_1, \dots, x_{\tilde{n}}$ of the ensemble members only, the matrix W will not change. Also, $y - Hx^b$ will depend only on the variables $x_1^b, \dots, x_{\tilde{n}}^b$ of x^b . The solution $z \in \mathbb{R}^m$ of

$$(R + HQ^b (HQ^b)^T) z = y - Hx^b, \quad (34)$$

is calculated based on the variables $x_1, \dots, x_{\tilde{n}}$ of Q^b and $x_1^b, \dots, x_{\tilde{n}}^b$ of x^b only. We summarize the result of these arguments in the following lemma.

LEMMA 2.1. *If we have observations y dependent only on the variables $x_1, \dots, x_{\tilde{n}}$ for $\tilde{n} \in \{1, \dots, n\}$ of the full state $x \in \mathbb{R}^n$ of the state space of our dynamical system, the transformation matrix W of the ensemble Kalman square root filter update $x^a - x^b$ depends on these variables of the centered ensemble Q and the mean first guess \bar{x}^b only.*

A consequence of the above Lemma 2.1 is that, if we have reduced observations, the ensemble Kalman square root filter will give us an update matrix W which depends only on the variables under consideration.

But we need to pay attention to the update and propagation step. The update Equation (11) clearly updates all variables, since $W \in \mathbb{R}^{L \times L}$, and thus all variables of x are updated by the ensemble Kalman filter. If the model M is based on all variables, in general we expect model propagation to be dependent on all variables as well. In general, an update based on the transformation matrix W will change all variables of the initial state. This means that the first guess of the next assimilation step highly depends on the application of the matrix W to all variables, not only to the variables $x_1, \dots, x_{\tilde{n}}$.

Clearly, in general we cannot run the full ensemble Kalman filter on a reduced set of variables, just because you need all prognostic variables to run the numerical model. We will see later, that this limitation does no longer apply when we are in the framework of ultra-rapid data assimilation.

3. THE ULTRA-RAPID DATA ASSIMILATION AND FORECASTING STEP

This section serves to develop the main ideas of ultra-rapid analysis, forecasting and smoothing. We will first describe the idea of ultra-rapid analysis when observations y_k are given at point of time t_k , $k = 1, \dots, N$ throughout a time interval $[t_0, t_N]$ for which we are not able to employ a full data assimilation functionality. We assume that we have been able to perform some ensemble data assimilation scheme prior to the time t_1 at time t_0 and that a forecast ensemble has been calculated, such that

$$x_{0,\xi}^{f,(\ell)}, \quad \xi = 0, \dots, N, \quad \ell = 1, \dots, L, \quad (35)$$

is available at the points in time t_ξ , $\xi = 0, \dots, N$ and for the ensemble index $\ell \in \{1, \dots, L\}$. Note that $x_{0,0}^{f,(\ell)}$ corresponds to

the analysis of the full ensemble data assimilation. We are now successively at times t_1, t_2, \dots receive observations y_1, y_2, \dots . The goal is to provide ultra-rapid updates for estimation of our state at times t_1, t_2, \dots . When we are at time t_k , we would like to update the forecasts at the times t_ξ for $\xi = k, \dots, N$ and obtain the best possible estimate in an ultra-rapid forecasting step.

Note, the assimilation of observations at some point in time exhibits information about the past as well. This is called *smoothing*. We will describe an ultra-rapid ensemble smoother in a second step. We focus on the analysis and forecasting in section 3.1 and discuss smoothing in section 3.2.

3.1. Ultra-Rapid Analysis and Forecasting

Assume that we are given some ensemble $x_k^{a,(1)}, \dots, x_k^{a,(L)}$ of L states of our dynamical system at time $t_k \in \mathbb{R}$, which could be an analysis or a first guess from somewhere. Further, we assume that we have applied our model M to calculate forecasts based on $x_k^{a,(\ell)}$ at times t_{k+1}, \dots, t_N for $N > k$. The corresponding forecasts are denoted by $x_{k,\xi}^{f,(\ell)}$ for $\xi = k+1, \dots, N$, analogous to (35).

We employ the following matrix notation. The matrix \mathbf{F} is the matrix of the full forecast ensemble members in its columns, i.e.,

$$\mathbf{F}_{k,\xi} = (x_{k,\xi}^{f,(1)}, \dots, x_{k,\xi}^{f,(L)})^T, \quad (36)$$

of forecasts $x_{k,\xi}^{f,(\ell)}$ from t_k to t_ξ . The matrix $\mathbf{W}^{(k)}$ is the matrix of linear ensemble transform coefficients calculated based on the observations y_k at time t_k and the first guess ensemble at time t_k , i.e.,

$$\mathbf{W}^{(k)} = (W_{j,\ell}^{(k)})_{j,\ell=1,\dots,L}. \quad (37)$$

When the analysis ensemble at time t_k is given by a generic ensemble data assimilation approach, we know that

$$x_k^{a,(\ell)} = \sum_{j=1}^L x_k^{b,(j)} W_{j,\ell}^{(k)}, \quad \ell = 1, \dots, L, \quad (38)$$

with the matrix $W_{j,\ell}^{(k)}$, $j, \ell = 1, \dots, L$ given by Equation (16) with the two quantities \bar{s} and S being dictated by the specific ensemble data assimilation system (e.g., Equations 30, 32), where the time index k refers to the analysis time t_k for which $W_{j,\ell}$ is calculated.

Also, we note that the background $x_k^{b,(\ell)}$ is given by

$$x_k^{b,(\ell)} = M_{k-1,k} (x_{k-1}^{a,(\ell)}), \quad \ell = 1, \dots, L. \quad (39)$$

LEMMA 3.1. *Here, we assume that the model M is a linear model \mathbf{M} . In this case, the forecast ensemble $x_{k,\xi}^{f,(\ell)}$ at time t_ξ when observations at time k are assimilated by a linear data assimilation method as in Equation (12), the forecast ensemble can be calculated by*

$$x_{k,\xi}^{f,(\ell)} = \sum_{j=1}^L x_{k-1,\xi}^{f,(j)} W_{j,\ell}^{(k)} \quad (40)$$

Proof. In this case we have

$$\begin{aligned}
 x_{k,\xi}^{f,(\ell)} &= \mathbf{M}_{k,\xi} x_k^{a,(\ell)}, \\
 &= \mathbf{M}_{k,\xi} \sum_{j=1}^L x_k^{b,(j)} W_{j,\ell}^{(k)}, \\
 &= \sum_{j=1}^L \left(\mathbf{M}_{k,\xi} x_k^{b,(j)} \right) W_{j,\ell}^{(k)}, \\
 &= \sum_{j=1}^L \left(\mathbf{M}_{k,\xi} \mathbf{M}_{k-1,k} x_{k-1}^{a,(j)} \right) W_{j,\ell}^{(k)}, \\
 &= \sum_{j=1}^L x_{k-1,\xi}^{f,(j)} W_{j,\ell}^{(k)}, \quad (41)
 \end{aligned}$$

for $\ell = 1, \dots, L$ and $\xi \in \{1, \dots, N\}$, where we used $\mathbf{M}_{k-1,\xi} = \mathbf{M}_{k,\xi} \mathbf{M}_{k-1,k}$. \square

Before we continue with our introduction of ultra-rapid data assimilation, we would like to study the reduced variable case in the above Lemma 3.1. Clearly, to apply $\mathbf{M}_{k,\xi}$ to a state $x^{(a)}$ or $x^{(b)}$, we need to know the full state. If only a part of the state x is available, starting the model is no longer possible. However, the Equation (40) is still valid for each of its components, i.e., if W is known, the variable $x_{k,\xi}^{f,(\ell)}$ of $x_{k,\xi}^{f,(\ell)}$ can be calculated from the knowledge of $x_{k-1,\xi,i}^{f,(\ell)}$ for all $\ell = 1, \dots, L$, representing the i -th variable of the state vector of the l -th ensemble member obtained by a forecast from time t_{k-1} to time t_ξ .

COROLLARY 3.2 (REDUCED SET OF MODEL VARIABLES). *If the observation operator H depends on the variables $x_1, \dots, x_{\tilde{n}}$ of the state x only, then the transformation matrix $W^{(k)}$ for the assimilation of y_k can be calculated from a) the first guess ensemble data $x_1^{(b)}, \dots, x_{\tilde{n}}^{(b)}$ and b) the observation y_k . For a linear model M , for the variables with index i we have*

$$x_{k,\xi,i}^{f,(\ell)} = \sum_{j=1}^L x_{k-1,\xi,i}^{f,(j)} W_{j,\ell}^{(k)}, \quad (42)$$

for $i = 1, \dots, \tilde{n}$, i.e., the formula (40) is valid and the ensemble forecast based on the analysis with observation y_k can be calculated from the knowledge of the reduced set of variables only.

The consequence of Equation (41) is that for linear models we can calculate the forecast based on the analysis at time t_k by a superposition of the forecast from time t_{k-1} . The weight matrix $W_{\ell,j}^{(k)}$ is calculated from the ensemble analysis at time t_k given by the linear ensemble data assimilation scheme. We can also use Equation (41) recursively, which is formulated in the following Theorem.

THEOREM 3.3. *We assume we are given observations y_j , $j = 1, \dots, k$ at times t_1, \dots, t_k . The goal is to calculate the forecasts $x_{k,\xi}^{f,(\ell)}$ at time t_ξ based on the observations from t_1 to t_k and the*

initial ensemble $x_0^{a,(\ell)}$ at time t_0 with an ensemble data assimilation method as in Equation (11). If the model M is linear, we obtain

$$\mathbf{F}_{k,\xi} = \mathbf{F}_{0,\xi} \mathbf{W}^{(1)} \dots \mathbf{W}^{(k)}, \quad (43)$$

for $\xi = k+1, \dots, N$.

Proof. For a linear model, the generic step is given by Equation (41). Then, the same equation is applied to $x_{k-1,\xi}^{f,(j)}$, which leads to

$$\begin{aligned}
 x_{k,\xi}^{f,(\ell)} &= \sum_{j_1=1}^L x_{k-1,\xi}^{f,(j_1)} W_{j_1,\ell}^{(k)}, \\
 &= \sum_{j_1=1}^L \left(\sum_{j_2=1}^L x_{k-2,\xi}^{f,(j_2)} W_{j_2,j_1}^{(k-1)} \right) W_{j_1,\ell}^{(k)}, \quad (44)
 \end{aligned}$$

and by the same step η times to

$$x_{k,\xi}^{f,(\ell)} = \sum_{j_1, \dots, j_\eta=1}^L x_{k-\eta,\xi}^{f,(j_\eta)} W_{j_\eta,j_{\eta-1}}^{(k-(\eta-1))} \dots W_{j_1,\ell}^{(k)}, \quad (45)$$

for $\eta \leq k$ assimilation steps. In matrix notation and for $\eta = k$ this is Equation (43). \square

Note that the recursive application of Equation (41) implies that any transformation matrix $\mathbf{W}^{(i)}$ is obtained using the observation y_i and the full ensemble $\mathbf{F}_{i-1,\xi}$.

The results for reduced data are also valid for the core formula (43). We collect the relevant statements into the following corollary. The matrix $\mathbf{F}_{k,\xi}$ contains the different state variables in its rows and the columns represent the ensemble under consideration. We employ the notation $(\mathbf{F}_{k,\xi})_{i=1, \dots, \tilde{n}}$ for the rows with the variable indices $i = 1, \dots, \tilde{n}$.

COROLLARY 3.4 (REDUCED SET OF MODEL VARIABLES). *If the observation operator H depends on the variables $x_1, \dots, x_{\tilde{n}}$ of the state x only, then the transformation matrix $W^{(k)}$ for the assimilation of y_k can be calculated from a) the first guess ensemble data $(\mathbf{F}_{0,k})_{i=1, \dots, \tilde{n}}$, b) the observation y_k and c) the previous transformation matrices $\mathbf{W}^{(1)} \dots \mathbf{W}^{(k-1)}$ which depend on the corresponding observations $y_1 \dots y_{k-1}$. For a linear model M , for the variables with index i we have*

$$(\mathbf{F}_{k,\xi})_{i=1, \dots, \tilde{n}} = (\mathbf{F}_{0,\xi})_{i=1, \dots, \tilde{n}} \mathbf{W}^{(1)} \dots \mathbf{W}^{(k)}, \quad \xi = k+1, \dots, N, \quad (46)$$

i.e., the formula (43) is valid and the ensemble forecast based on the analysis with observation y_k can be calculated from the knowledge of the reduced set of variables only.

3.2. Ultra-Rapid Smoother Functionality

Smoothers are schemes which employ information from the future to improve the estimate about some present state. Alternatively, you could say that they use information now to update past states.

When we consider the scenario of ultra-rapid data assimilation, for the interval $[t_0, t_N]$ we are given an ensemble of

original states (35) over the full interval. When an observation is arriving at time t_k (ignoring delay usually needed for observation processing and transfer), we can employ the same techniques which are used for updating the analysis and forecast to the past interval $[t_0, t_k]$.

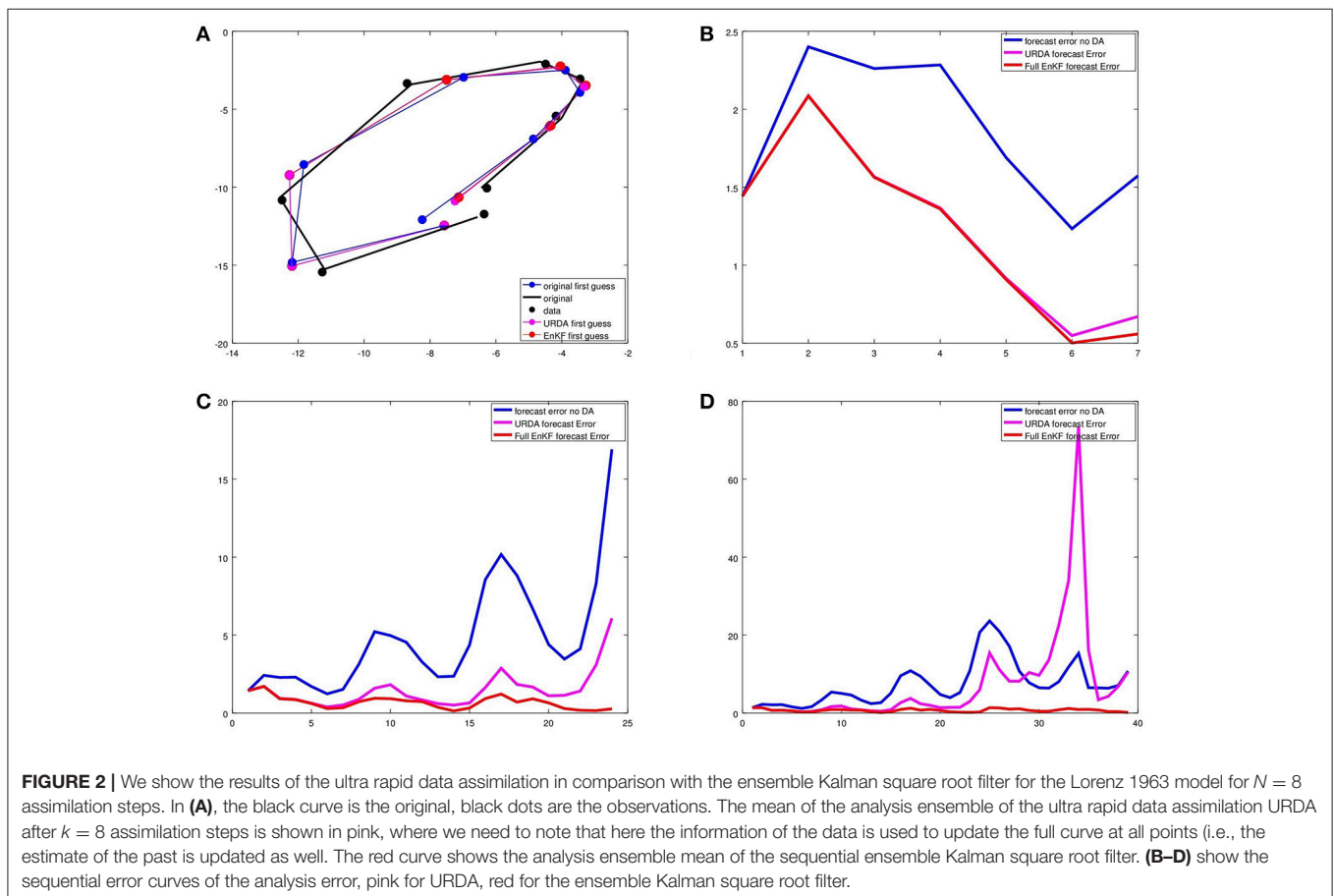
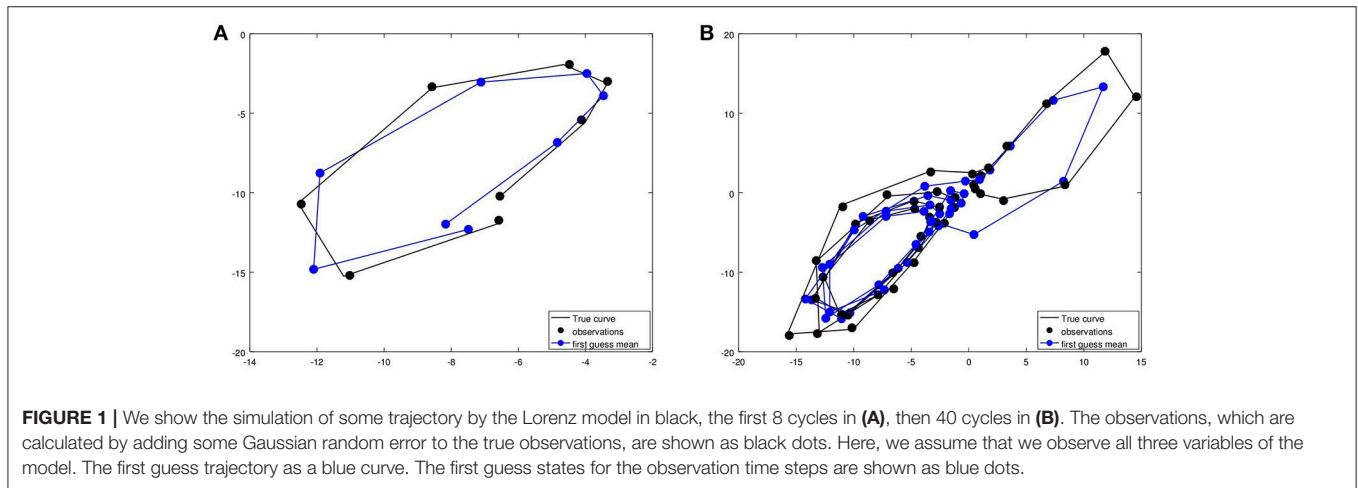
DEFINITION 3.5 (ULTRA-RAPID ENSEMBLE SMOOTHER).

Given the original first guess ensemble $F_{0,\xi}$ for $\xi = 0, \dots, N$ on

the time interval $[t_0, t_N]$ we define the ensemble analysis given the data y_1, \dots, y_k by

$$F_{k,\xi}^{(a)} := F_{0,\xi} W^{(1)} \dots W^{(k)}, \quad \xi = 0, \dots, N. \quad (47)$$

This analysis ensemble is defined for the full time interval.



In general, a convergence analysis of an ensemble Kalman smoother and its comparison to a four-dimensional variational data assimilation (4D-VAR) scheme over the time window $[t_0, t_n]$ can be found in Theorem 5.4.7 of Nakamura and Potthast [3]. For linear models and observation operators, the full Kalman smoother and 4D-VAR are equivalent.

Clearly, if we replace the full model M by the ensemble, this equivalence is no longer true. Also, if the numerical model M used to calculate the ensemble is different from the true model M_{true} , the temporal correlations, which are implicitly used when we employ the analysis matrix $W^{(k)}$ to update the ensemble in the past or in the future, may not be correct with respect to the true ensemble correlations. In this case, the information y_k in the future of t_0 may not improve the state estimate at time t_0 , but lead to additional errors in this state estimate. We will demonstrate this phenomenon in our numerical examples in section 4.

4. NUMERICAL EXAMPLES

The goal of this section is to study the ultra rapid data assimilation for simple generic examples. We want to show that the assimilation step can be carried out in a stable way and that the ultra-rapid forecasts indeed show an advantage over the ensemble forecasts without this step. Also, we would like to understand the range of skill which we can achieve when

we compare it with the full standard data assimilation and forecasting approaches.

4.1. Studying URDA for the Lorenz 63 Model System

Here, we start our study with the Lorenz 63 model Lorenz [20]. It is a very well-known chaotic ODE system with three unknowns, compare for example Nakamura and Potthast [3].

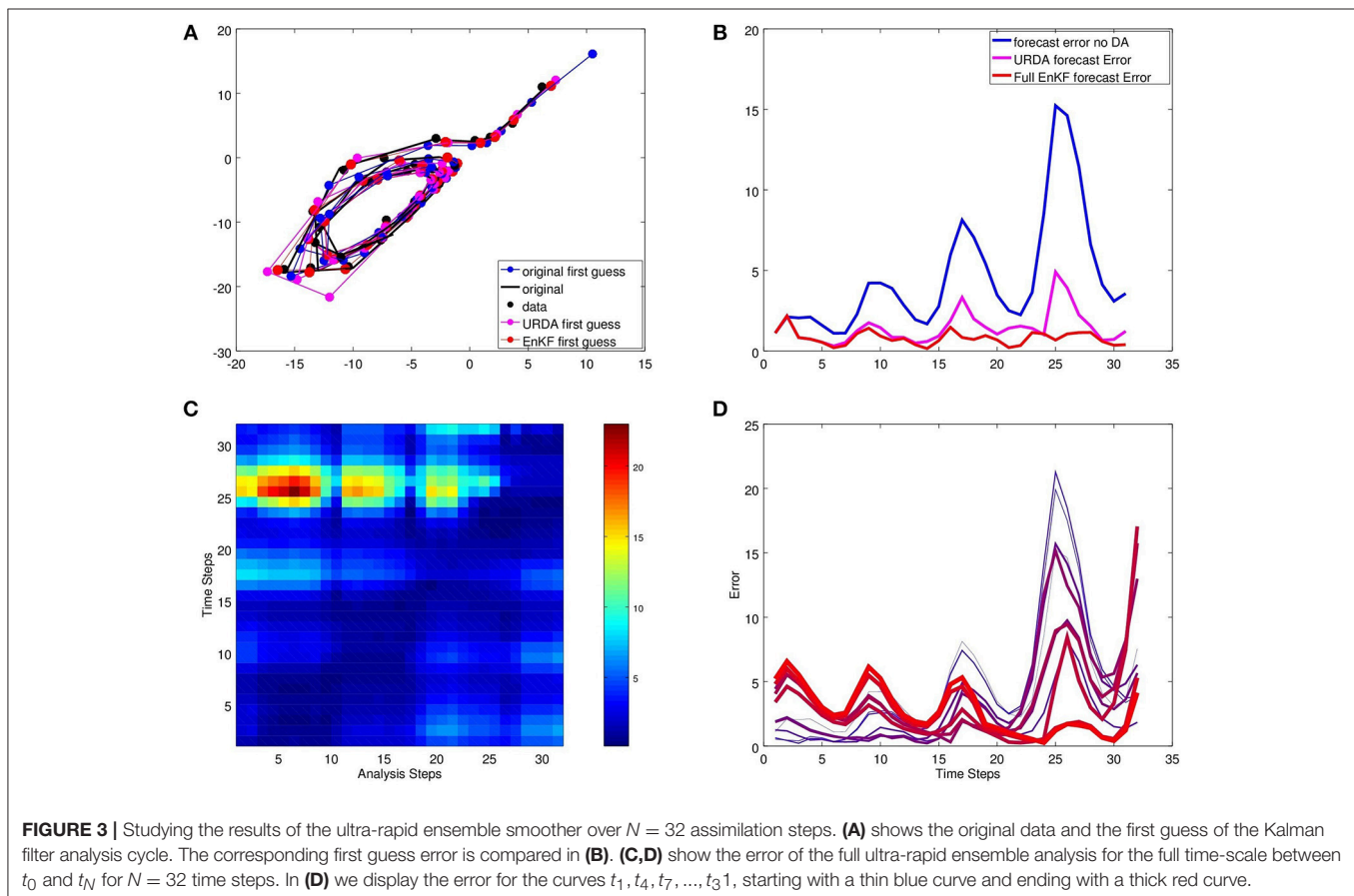
The Lorenz 1963 model is a system of three non-linear ordinary differential equations

$$\dot{x} = \sigma(y - x), \quad (48)$$

$$\dot{y} = x(\rho - z) - y, \quad (49)$$

$$\dot{z} = xy - \beta z, \quad (50)$$

with constants σ, ρ, β known as *Prandtl number*, the *Rayleigh number* and a non-dimensional *wave number*. Here, for the constants we take the classical values $\sigma = 10$, $\beta = 8/3$ and $\rho = 28$. The implementation of the system is usually carried out by a higher-order integration scheme such as 4th-order Runge-Kutta, which we have employed for our numerical testing. The setup for our case study is shown in **Figure 1A** with 8 cycles for better visibility and **Figure 1B** with 40 cycles for studying the error evolution.



Here, we want to test the feasibility of ultra rapid data assimilation. The original curve is shown in black in **Figure 1**. The measurements are calculated by adding a random Gaussian error to this curve at the measurement times t_1, t_2, \dots, t_k with $\Delta_t = t_{i+1} - t_i = 0.1$ (without units). For the original, we have used the above ODE system with $\sigma = 10$ to generate the *truth*. To test data assimilation we have employed a modified system where $\sigma = 12$ was chosen. The mean of the original first guess ensemble for the full time period under consideration is shown in **Figure 1** as a blue curve, with blue dots as the original first guess.

We have now followed two tracks. First, we have implemented an ensemble Kalman square root filter. We start with a first guess ensemble, which is generated at time t_0 by adding random Gaussian errors to the starting point of the original curve. Then we assimilate the observations (the black dots) using the Ensemble Kalman square root filter.

Second, the ultra-rapid data assimilation and forecasting cycle has been implemented. The ultra rapid data assimilation has been set up by first calculating the full first guess ensemble for the whole time interval under consideration. Then, a modified ensemble is calculated step by step following (43). We study N time steps (showing results for $N = 8$ and $N = 40$). In more detail, we have calculated the transformation matrix $W^{(k)}$ based on the observations y_k at time $t_k, k = 1, \dots, N$ and the transformed first guess ensemble $x_{k-1, \xi}^{b, (\ell)}$. Here, ξ is the time index of the

ensemble, i.e., $\xi = 1, \dots, N$. We carry out the assimilation for all time steps, changing the ensemble in the past as well as in the full future over the time interval under consideration.

The result of N time steps is shown in **Figure 2**. First, the example with $N = 8$ time steps is shown in **Figure 2A,B**, the first guess errors for $N = 25$ and $N = 40$ time steps in **Figure 2C,D**. Here, initially the ultra-rapid update is quite good, approximating well the full ensemble Kalman square root filter over 10 or 15 assimilation steps. Then, when the first guess ensemble and the true trajectory diverge further, the assimilation loses track and we obtain very large errors over time, as can be seen by the peak of the pink curve in **Figure 2D** at about t_k with $k = 34$.

Here, we also investigate the ultra rapid data assimilation tool as a smoother. We calculate the analysis ensemble $F_k^{(a)}$ defined in (47) based on the original first guess ensemble F_0 .

In **Figure 3** we study the filter and smoother results for a case with $N = 32$. For the latter we update both the future and the past. Errors of this with respect to the true curve are displayed in **Figure 3C,D**. Here, we need to note that we simulate a realistic setup in the sense that the true model M_{true} is different from the model M used to calculate the first guess ensemble. That has severe consequences for the convergence of the smoother. With the errors in the model, we obtain errors in the first guess ensemble and with this errors in the correlations and covariances

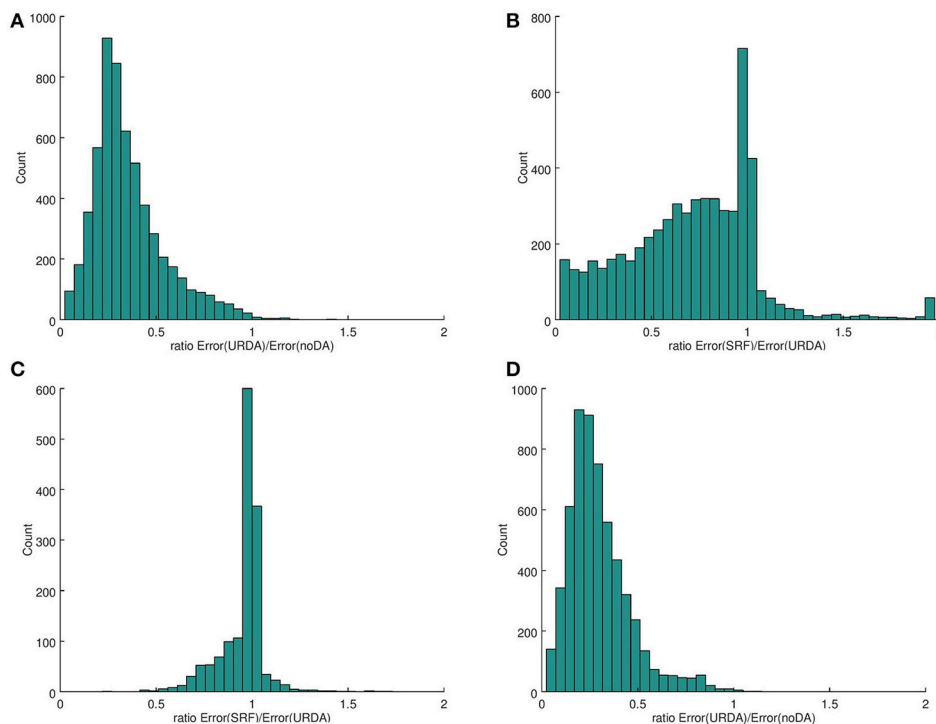


FIGURE 4 | We show the results of the mean-error of the ultra rapid data assimilation in comparison with the original first guess (no data assimilated) and the ensemble Kalman square root filter for the Lorenz 1963 model. We used $N_{stat} = 250$ different initializations of the random number generator to obtain different distributions for the observations and the initial ensemble. After assimilation of all data the mean error at each time step on the trajectory from the truth is counted. In **(A,B)** we used $L = 5$ and $N = 25$ to obtain histograms showing in **(A)** the ratio of the mean-error of URDA divided by the mean-error of the model forecast without data assimilation. For $L = 5$ in **(B)** the mean-error of the SRF divided by the one of URDA is displayed for $N = 25$ while the same is shown for $N = 8$ in **(C)**. The ratio of the mean-error of URDA divided by the initial forecast for $N = 25$ and $L = 25$ in **(D)**.

which are exploited by the ultra-rapid Kalman filter forecast and the update of the past in the ensemble Kalman smoother.

Studying **Figure 3C** we see that the error is smallest on the diagonal, i.e., for the analysis and short-term forecast based on the ultra-rapid ensemble Kalman analysis. The errors for the analysis or forecast increase with distance to the current point in time. We expect that the errors are large for larger lead times. But in general we do not expect that the errors in the past, i.e., at the beginning of the interval $[t_0, t_N]$ increase when we assimilate more and more data. When the ensemble reflects the correct correlations between the future and the past, the error should decrease. However, with a numerical model which is different from the true model, we also inherit errors into the temporal correlations. As a consequence we observe that the error at t_1 increases when we assimilate further data y_k for k in the second part of the interval $[t_0, t_N]$.

In **Figure 4** we evaluate the performance of URDA in a statistical manner by using different initialisations for the applied random number generator, which affects the observations drawn from a Gaussian distribution as well as the construction of the ensemble, and use different values of the starting point $x_0 = x(t_0)$, which is used to obtain the truth as well as the ensemble. We evaluate differences of the corresponding mean from the truth and take appropriate ratios. In **Figure 4A** the mean-error of URDA is divided by the mean error of the free forecast (no data assimilation, also abbreviated by no-DA) for $L = 5$ ensemble member and $N = 25$ time steps on the trajectory. A clear positive impact is visible with only very few cases where the free forecast

is better than URDA. **Figure 4B** shows the mean-error of the SRF divided by the one of URDA. As expected, the evaluation shows that in many cases the full SRF performs better compared to URDA. However, comparing with the results shown in **Figure 3C** this is what we expect due to the deviations after about $N = 8$ time steps. To test this, we show the result for the first 8 time steps of the run with $N = 8$ in **Figure 4C** and observe, that for a smaller N these compete indeed much better with the SRF. **Figure 4D** shows the result of the mean-error of URDA divided by the initial forecast with no data assimilation for $L = 25$ ensemble member and $N = 25$ time steps. We observe, that the improvement of URDA with $L = 25$ compared to $L = 5$ is not significant. This is no surprise since we deal with three prognostic variables where an ensemble of $L = 5$ is already sufficient to describe the relevant spread.

At the end of this section we highlight the impact of the time step in the model, which translates to the time the forecast from one point on the trajectory is performed. Note, this does not affect the performance of the Runge-Kutta-Scheme where the time step of the integration is kept fixed. We evaluate the ratio of the deviations from the mean error from the SRF to URDA. In **Figure 5** we show results for different sizes of the time step dt . Again we used $N_{stat} = 250$ and the total number of time steps $N = 25$ with the number of ensemble members $L = 5$. We observe, that for $1/4$ of the standard time step size $dt = 0.100$ the SRF and URDA perform almost equally. With increasing time step size we find more cases where the SRF outperforms URDA, which is still moderate for the standard time step size. For three

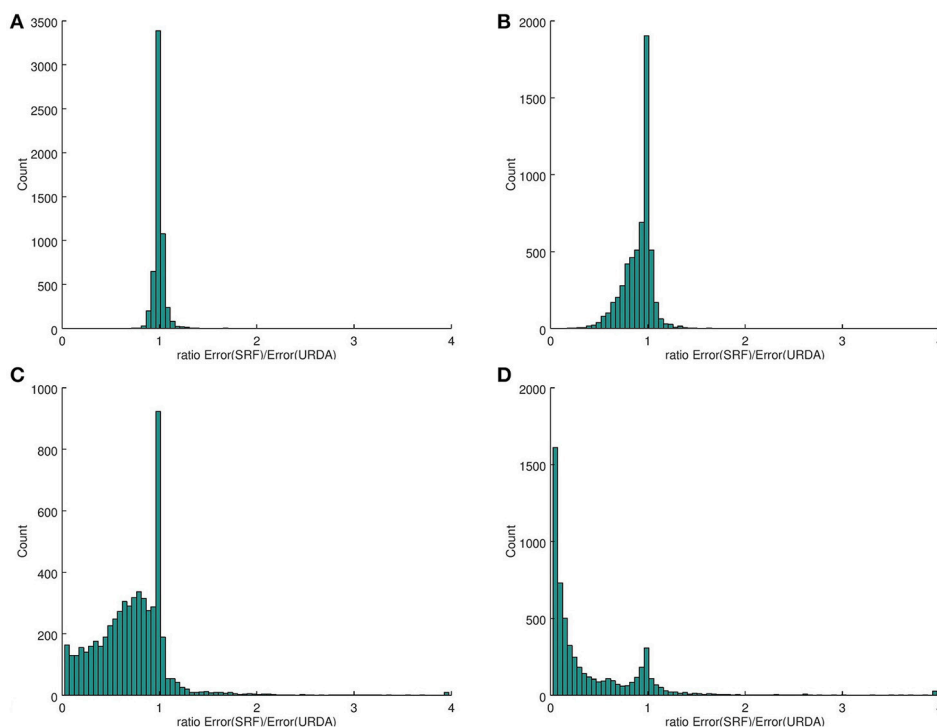


FIGURE 5 | We show the ratio of the mean-error of the SRF divided by the one of URDA for different step sizes in time. Specifically, in **(A)** we used $dt = 0.025$, in **(B)** $dt = 0.050$, in **(C)** the standard value $dt = 0.100$ and in **(D)** $dt = 0.300$.

times this step size we see a clear benefit for the SRF. Note, since we keep $N = 25$ fixed, the total length of the trajectory differs for the different time step sizes dt .

4.2. URDA for Reduced Model Dynamics

In the second part of our numerical study we would like to understand how ultra-rapid ensemble data assimilation can be applied to the case where only a reduced set of variables is passed down from the standard ensemble data assimilation framework.

In the framework of the Lorenz model, we have carried out a study the use of the observation operators

$$H_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad H_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad H_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix}, \quad (51)$$

and study assimilation of observations of either the full state x , the first two variables of the state or the first variable of the state only.

We note that HQ will employ the corresponding selection of variables depending on the cases H_1 , H_2 or H_3 , i.e., H_1Q can be calculated from the knowledge of the first variable x_1 of $x = (x_1, x_2, x_3)^T$ only. Similarly, H_2Q can be calculated based on the knowledge of (x_1, x_2) of $x = (x_1, x_2, x_3)^T$.

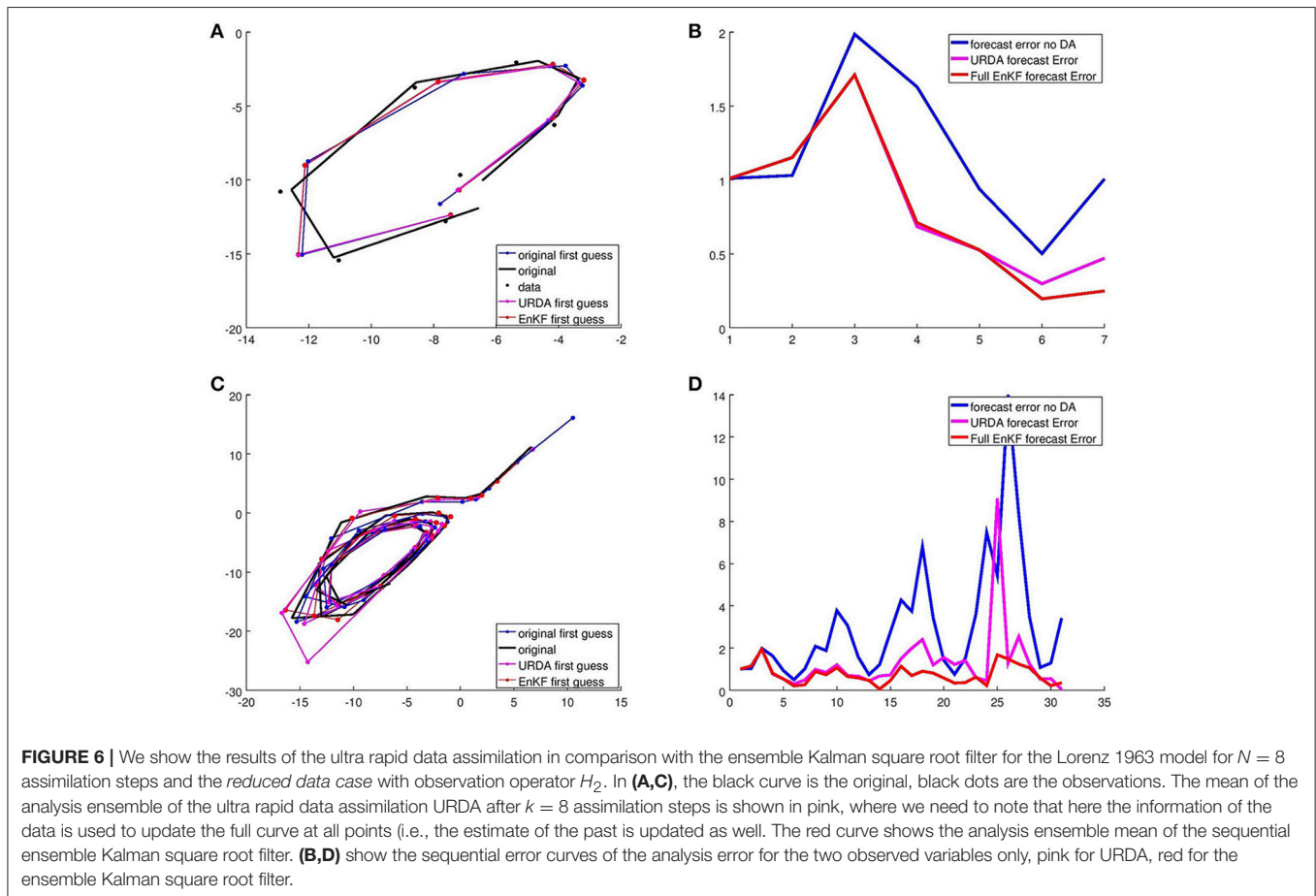
Here, we focus on the results for the use of H_2 in **Figure 6**. The effects are similar to the three-dimensional

version. **Figure 6A** displays the first 8 steps, and we see that the SRF analysis and the URDA analysis are very close to each other. The error is shown in **Figure 6B**, here only for the two variables under consideration. **Figure 6C,D** display 30 assimilation setups. After 20 and 25 steps we observe first cases where URDA is worse than no-DA. In all other cases it is a big increase from the no-DA case and its quality becomes close to the quality of the full square-root filter with subsequent forecast.

5. CONCLUSIONS

We analyse and investigate a ultra-rapid data assimilation scheme based on an ensemble square-root Kalman filter. Here, we have studied the analysis cycle, a preemptive forecasting step and also an ultra-rapid ensemble smoother.

For linear systems we have shown that the ultra-rapid data assimilation is equivalent to the full ensemble square-root filter. For non-linear systems, the Lorentz 63 system serves as a standard test case which is widely used within geophysics or the life sciences. We have carried out numerical tests of the URDA scheme, which shows highly encouraging results. For a significant number of assimilation and forecasting steps the URDA scheme shows a similar forecasting skill as the square-root filter with full model forecasts.



In particular, we have analyzed and tested the assimilation of observations which are influenced by a selection of state variables only, where the URDA scheme provides the possibility to touch only the variables of interest for the assimilation and preemptive forecasting or smoothing steps. This has very-high potential for many applications, where high-frequency analysis and/or forecasts need to be calculated, e.g., in the area of brain surgery in neuroscience or in nowcasting in geophysical applications.

This work aims to provide the basic theoretical inside and study a standard non-linear system of wide interest, the Lorenz 63 system. Initial tests on a real-world system in geophysics have been carried out in Etherton [25] and Madaus and Hakim[26]. Further work on error estimates for non-linear systems and the application of the method in neuroscience, biological systems or weather forecasting is still pending and will be our goal for the near future.

REFERENCES

1. Park SK, Park L, eds. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. II)*. Berlin; Heidelberg: Springer-Verlag (2013).
2. Kalnay E. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge: Cambridge University Press (2003).
3. Nakamura G, Potthast R. *Inverse Modeling*. 2053–2563. Bristol: IOP Publishing (2015).
4. Hamilton F, Berry T, Peixoto N, Sauer T. Real-time tracking of neuronal network structure using data assimilation. *Phys Rev E* (2013) **88**:052715. doi: 10.1103/PhysRevE.88.052715
5. Nogaret A, Meliza CD, Margoliash D, Abarbanel HDI. Automatic construction of predictive neuron models through large scale assimilation of electrophysiological data. *Sci Rep.* (2016) **6**:32749. doi: 10.1038/srep32749
6. Palmer T, Hagedorn R. *Predictability of Weather and Climate*. Cambridge: Cambridge University Press (2006).
7. Massimo B, Elias H, Lars I, Mike F. The evolution of the ECMWF hybrid data assimilation system. *Q J R Meteorol Soc.* (2015) **142**:287–303. doi: 10.1002/qj.2652
8. Schraff C, Reich H, Rhodin A, Schomburg A, Stephan K, Periañez A, et al. Kilometre-scale ensemble data assimilation for the COSMO model (KENDA). *Q J R Meteorol Soc.* (2016) **142**:1453–72. doi: 10.1002/qj.2748
9. Jaisson P, DeVuyt F. Data assimilation and inverse problem for fluid traffic flow models and algorithms. *Int J Numer Methods Eng.* (2008) **76**:837–61. doi: 10.1002/nme.2349
10. Fukuda D, Hong Z, Okamoto N, Ishida H. Data assimilation approach for traffic-state estimation and sensor location/spacing problems in an urban expressway. *J Jpn Soc Civil Eng Ser D3.* (2014) **70**:I1041–50. doi: 10.2208/jscejipm.70.I_1041
11. Nantes A, Ngoduy D, Bhaskar A, Miska M, Chung E. Real-time traffic state estimation in urban corridors from heterogeneous data. *Transport Res C Emerg Technol.* (2016) **66**:99–118. doi: 10.1016/j.trc.2015.07.005
12. Bocquet M, Pires CA, Wu L. Beyond gaussian statistical modeling in geophysical data assimilation. *Monthly Weather Rev.* (2010) **138**:2997–3023. doi: 10.1175/2010MWR3164.1
13. Alberto C, Marc B, Laurent B, Geir E. Data assimilation in the geosciences: an overview of methods, issues, and perspectives. *Wiley Interdiscipl Rev Clim Change* (2018) **9**:e535. doi: 10.1002/wcc.535
14. Daley R. *Atmospheric Data Analysis*. Cambridge Atmospheric and Space Science Series. Cambridge: Cambridge University Press (1993).
15. Evensen G. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. *J Geophys Res Oceans* (1994) **99**:10143–62. doi: 10.1029/94JC00572

AUTHOR CONTRIBUTIONS

Ideas for the investigation by RP. Execution of the numerical calculations were performed by RP and CW. Writing the publication was done by RP and CW.

ACKNOWLEDGMENTS

We thank Jeffrey Anderson from NCAR for the interesting discussions and pointing us to the paper of L. Madaus and G. Hakim. Furthermore we thank Z. Paschalidi and J. W. Acevedo Valencia for fruitful discussions on the topic, applications and further development of these ideas in the context of the projects Sinfony and Flottenwetterkarte at Deutscher Wetterdienst. This work was supported by the Deutscher Wetterdienst research program Innovation Programme for applied Researches and Developments (IAFE) in course of the SINFONY project.

16. Hunt BR, Kostelich EJ, Szunyogh I. Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter. *Physica D Nonlinear Phenomena* (2007) **230**:112–26. doi: 10.1016/j.physd.2006.11.008
17. Evensen G. *Data Assimilation: The Ensemble Kalman Filter*. Earth and Environmental Science. Heidelberg: Springer (2009).
18. Kalnay E, Hunt B, Ott E, Szunyogh I. *Ensemble Forecasting and Data Assimilation: Two Problems With the Same Solution*. Cambridge: Cambridge University Press (2006).
19. Freitag M, Potthast R. Synergy of inverse problems and data assimilation techniques. In: Cullen M, Freitag MA, Kindermann S, Scheichl R, editors. *Large Scale Inverse Problems. vol. 13 of Radon Series on Computational and Applied Mathematics*. Berlin: Walter de Gruyter (2013).
20. Lorenz EN. Deterministic nonperiodic flow. *J Atmos Sci.* (1963) **20**:130–41.
21. Gauthier P. Chaos and quadri-dimensional data assimilation: a study based on the Lorenz model. *Tellus A* (1992) **44**:2–17. doi: 10.3402/tellusa.v44i1.14938
22. Evensen G. Advanced data assimilation for strongly nonlinear dynamics. *Monthly Weather Rev.* (1997) **125**:1342–54.
23. Yang SC, Baker D, Li H, Cordes K, Huff M, Nagpal G, et al. Data assimilation as synchronization of truth and model: experiments with the three-variable lorenz system. *J Atmos Sci.* (2006) **63**:2340–54. doi: 10.1175/JAS3739.1
24. Tandeo P, Ailliot P, Ruiz J, Hannart A, Chapron B, Cuzol A, et al. Combining Analog Method and Ensemble Data Assimilation: Application to the Lorenz-63 Chaotic System. In: Lakshmanan V, Gilleland E, McGovern A, Tingley M. Editors. *Machine Learning and Data Mining Approaches to Climate Science*. Cham: Springer International Publishing (2015). p. 3–12.
25. Etherton BJ. Preemptive forecasts using an ensemble kalman filter. *Monthly Weather Rev.* (2007) **135**:3484–95. doi: 10.1175/MWR3480.1
26. Madaus LE, Hakim GJ. Rapid, short-term ensemble forecast adjustment through offline data assimilation. *Q J R Meteorol Soc.* (2015) **141**:2630–42. doi: 10.1002/qj.2549
27. Kalman RE. A New Approach to Linear Filtering and Prediction Problems. *J Basic Eng.* (1960) **82**:35–45.

Conflict of Interest Statement: 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.

Copyright © 2018 Potthast and Welzbacher. 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.



Forecast of Spectral Features by Ensemble Data Assimilation

Axel Hutt^{1,2*} and Roland Potthast^{1,2}

¹ Department FE12 - Data Assimilation, Deutscher Wetterdienst, Offenbach, Germany, ² Department of Applied Mathematics and Statistics, University of Reading, Reading, United Kingdom

OPEN ACCESS

Edited by:

Ulrich Parlitz,
Max-Planck-Institut für Dynamik und
Selbstorganisation, Germany

Reviewed by:

Hiromichi Suetani,
Oita University, Japan
Xin Tong,
National University of Singapore,
Singapore

*Correspondence:

Axel Hutt
axel.hutt@dwd.de

Specialty section:

This article was submitted to
Dynamical Systems,
a section of the journal
Frontiers in Applied Mathematics and
Statistics

Received: 10 April 2018

Accepted: 15 October 2018

Published: 01 November 2018

Citation:

Hutt A and Potthast R (2018) Forecast
of Spectral Features by Ensemble
Data Assimilation.
Front. Appl. Math. Stat. 4:52.
doi: 10.3389/fams.2018.00052

Data assimilation permits to compute optimal forecasts in high-dimensional systems as, e.g., in weather forecasting. Typically such forecasts are spatially distributed time series of system variables. We hypothesize that such forecasts are not optimal if the major interest does not lie in the temporal evolution of system variables but in time series composites or features. For instance, in neuroscience spectral features of neural activity are the primary functional elements. The present work proposes a data assimilation framework for forecasts of time-frequency distributions. The framework comprises the ensemble Kalman filter and a detailed statistical ensemble verification. The performance of the framework is evaluated for a simulated FitzHugh-Nagumo model, various measurement noise levels and for *in situ*-, nonlocal and speed observations. We discover a resonance effect in forecast errors between forecast time and frequencies in observations.

Keywords: Kalman filter, neural activity, prediction, dynamical system, verification

1. INTRODUCTION

Understanding the dynamics of natural complex systems is one of the great challenges in science. Various research domains have developed optimized analytical methods, computational techniques or conceptual frameworks to gain deeper insight into the underlying mechanisms of complex systems. In the last decades, more and more interdisciplinary research attracted attention building bridges between research domains by applying methodologies outside of domains. These cross-disciplinary techniques fertilize research domains and shed new light on their underlying properties. A prominent example is the mathematical domain of dynamical systems theory that traditionally is applied in physics and engineering and that has been applied very successfully in biology and neuroscience. For instance, taking a closer look at the spatio-temporal nonlinear dynamics of neural populations has allowed to identify epilepsy as a so-called dynamical disease [1]. This approach explains epileptic seizures as spatio-temporal instabilities hypothesizing that epileptic seizures emerge by phase transitions well-studied in physics. Another example is control theory that is well-established in electric engineering, e.g., in the cruise-control in automobiles or the flight control of airplanes. Similar control engineering techniques have been applied in neuroscience for some years now, e.g., to optimize electric deep brain stimulation in Parkinson disease [2, 3].

Weather forecasts are an everyday service provided by national and regional weather services that allows to plan business processes as well as private activity and serves as a warning system for extreme weather situations, such as floods or thunderstorms. Weather forecast is also an important research domain in meteorology that has been developed successfully in the last decades improving the forecasts for both global phenomena and local weather situations. In detail, today's weather services employ highly tuned and optimized meteorological models and data processing

techniques to compute reliable forecasts. Specifically the combination of an efficient model and measured meteorological data enables researchers to provide various types of predictions, such as the probability of rain or the expected temperature in certain local regions. This optimal combination of model and data is achieved by data assimilation [4] and yields corresponding optimal forecasts.

In other research domains, prediction methods are rare but highly requested. For instance, the prediction of epileptic seizures [5] would dramatically improve the life of epilepsy patients and spare some of them health-critical drug treatments. The typical approach of seizure prediction classifies measured neural activity [6, 7] into seizure-no seizure data, what however does not provide forecasts of neural activity. Although such forecasts are made possible by data assimilation techniques, until today research in neuroscience does apply data assimilation rarely. In recent years, data assimilation methods have been applied in neuroscience for model parameter identification primarily [3, 8–12]. The present work extends these studies by a framework to both compute and validate forecasts in neural problems. Although large parts of the methodology presented is well-established in meteorological forecasts [4], we extend the techniques by a focus on spectral features in measurement data. Such spectral data features play an important role in neuroscience since there is almost-proofed evidence that neural information processing is encoded in rhythmic activity. For instance, mammalian visual perception is achieved by synchronization in the frequency range [30 Hz; 60 Hz] [13] and unconsciousness and sleep is reflected in increased activity in the frequency range [0.5 Hz; 4 Hz] [14]. Moreover, epileptic seizures exhibit strong rhythmic patterns [1]. Consequently, we aim to forecast spectral distributions over time. To our best knowledge the present work is one of the first to predict spectral distributions optimally by data assimilation.

Most recent data assimilation studies apply the unscented Kalman filter [3, 10] that performs well for low-dimensional models. The present work considers the ensemble Kalman filter [15, 16] that has been shown to outperform the unscented Kalman filter and still performs well for high-dimensional models [17]. One of the major differences to previous studies is that the data assimilation cycle applied here does not estimate system parameters but providing reasonable forecasts. We provide a detailed description of the data assimilation elements and its extension to spectral feature forecasts. The additional verification of the ensemble forecasts gives insights into the power and weakness of spectral feature forecasts. For instance, we find a resonance effect between forecast time and the oscillation frequency of observations that yields improved verification metrics although forecasts are not improved.

The work is structured as follows. The Methods section introduces the model, simulated observations, the ensemble Kalman filter and the verification metrics applied. The subsequent section shows obtained results for *in situ*-, nonlocal, and speed observations and various measurement noise levels. A final discussion closes the work.

2. MATERIALS AND METHODS

2.1. The Model

Single biological neurons may exhibit various types of activity, such as no spike discharge, discharge of single spikes, regular spike discharge, or spike burst discharges. These activity modes can be described by high-dimensional dynamical models. A more simple model is the FitzHugh-Nagumo model [18, 19] describing spike discharges by two coupled nonlinear ordinary differential equations

$$\frac{dV}{dt} = V - \frac{1}{3}V^3 - w + I \quad (1a)$$

$$\tau \frac{dw}{dt} = (V + a - bw) \quad (1b)$$

with membrane potential V , recovery variable w and corresponding time scale τ , external input I , and physiological constants $a = 0.1, b = -0.15$. In our study, we consider two models. The nature model is non-homogeneous and time scales and input vary according to

$$\tau_n(t) = 10 + \frac{10t}{T}, \quad 0 \leq t \leq T \quad (2a)$$

$$I_n(t) = 0.35 + \frac{0.95t}{T}, \quad 0 \leq t \leq T \quad (2b)$$

with maximum time T . This model is supposed to describe the *true* dynamics in the system under study and typically that one does not know. The change of τ_n and I_n over time results in a shift of oscillation frequency of the system, i.e., from larger to smaller frequencies. Such a non-homogeneous temporal rhythm is well-known in neuroscience, e.g., in the presence of anesthetic drugs [20, 21]. The false model is not complete and represents just an estimate of the system under study. This is the model with which one describes systems and, typically, it is not correct. We assume that we do not know the non-homogeneous nature of the true model and assume temporally constant time scale and input

$$\tau_f = 20, \quad I_f = 1.3 \quad (3)$$

leading to a single oscillation frequency. We point out that $\tau_n(T) = \tau_f$ and $I_n(T) = I_f$ and both models converge to each other for $t \rightarrow T$.

The model integration over time uses a time step of 0.01 and every 50 steps a sample is written out running the integration over $5 \cdot 10^4$ steps in total. Initial conditions are $\mathbf{x}(t = 0) = (1.0, 0.2)^T$. After numerical integration, we re-scaled the unit-less time by $\alpha t \rightarrow t$ with $\alpha = 0.002s$ rendering the sample time to $\Delta t = 1$ ms and the maximum time to $t_{max} = 1s$. This sets the number of data points to $N = 1,000$.

To reveal non-stationary cyclic dynamics, we analyze the time-frequency distribution of data with spectral density $S(t_k, \nu_m)$, $k = 1, \dots, K$, $m = 1, \dots, M$ for number of time points K and number of frequencies M . The Morlet wavelet transform

$$\mathcal{W}[y](t, \nu) = \int_{-\infty}^{\infty} y(t') \Psi^* \left(\frac{t' - t}{a(\nu)} \right) dt'$$

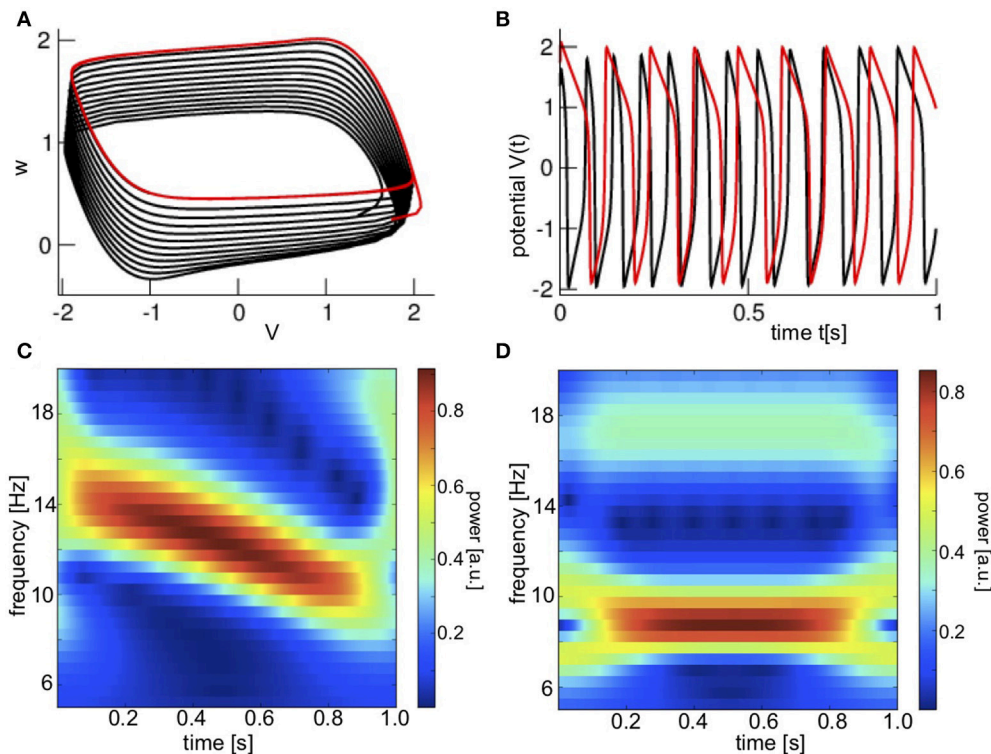


FIGURE 1 | Noise-free dynamics of the Fitzhugh-Nagumo model and corresponding observations. **(A)** The model trajectory in phase space for non-stationary (black) and stationary (red) dynamics. **(B)** The time series of corresponding observations $H[x]$ for non-stationary and stationary dynamics. **(C)** The time-frequency evolution of observations for non-stationary dynamics in the true model with Equation (2). **(D)** The time-frequency evolution of observations for the stationary dynamics with Equation (3).

applied uses a mother wavelet Ψ with central frequency $f_c = 8$ and the time-frequency distribution has a frequency resolution of $\Delta\nu = 0.5$ Hz in the range $\nu \in [5 \text{ Hz}; 20 \text{ Hz}]$. The parameter $a = f_c/\nu$ is the scale that depends on the pseudo-frequency ν . By the choice of the central frequency f_c , the mother wavelet has a width of 4 periods of the respective frequency. This aspect is important to re-call while interpreting temporal borders of time-frequency distributions. For instance, at a frequency of 15 Hz border disturbances occur in a window of 0.26 s from the initial and final time instant.

Figure 1A presents the phase space dynamics of the true model (black) and the false model (red) and one observes nonlinear cyclic dynamics. For illustration, **Figure 1B** shows the potential V . Oscillations of the true model (black) decelerate with time while the false model dynamics (red) is a stationary limit cycle. This can be seen even better in the time-frequency distribution shown in **Figures 1C,D** of the corresponding observations.

2.2. Observations

To relate model variables to observations, data assimilation introduces the notion of a measurement operator $\mathbf{H}: X \in \mathcal{M} \rightarrow Y \in \mathcal{O}$. This operator maps system variables $\mathbf{x} \in \mathcal{M}$ in model space \mathcal{M} to observable variables $\mathbf{y} \in \mathcal{O}$ in observation space \mathcal{O} .

The system dynamics can be observed in various ways and the observation operator is chosen correspondingly. Measurements directly in the system are called *in-situ* observations and, typically, the measured observable is proportional to a model variable. In this case, the operator is proportional to the identity. Examples for such observables are temperature or humidity in meteorology and intra-cellular potentials or Local Field Potentials in neurophysiology. Conversely, measurements outside the system are called *nonlocal* observations capturing the integral of activity from the system. Examples for such observations are satellite radiances or radar reflectivities in meteorology and encephalographic data and the BOLD response in functional Magnetic Resonance Imaging in neurophysiology.

The present study considers scalar *in-situ* observations, nonlocal observations and temporal derivatives and *in-situ* observations. We begin with *in-situ* observations $y(t)$ disturbed by measurement noise

$$y(t) = V(t) + \kappa \xi(t), \quad (4)$$

where $\xi(t)$ are Gaussian distributed uncorrelated random numbers with $\langle \xi(t) \rangle = 0$, $\langle \xi(t) \xi(t') \rangle = \delta(t - t')$, $\langle \cdot \rangle$ denotes the ensemble average and $V(t)$ is the membrane potential from model (1). The noise level κ is chosen to $\kappa = 0$ (no noise), $\kappa = 0.5$ (medium noise), and $\kappa = 0.8$ (large noise). **Figure 2** shows

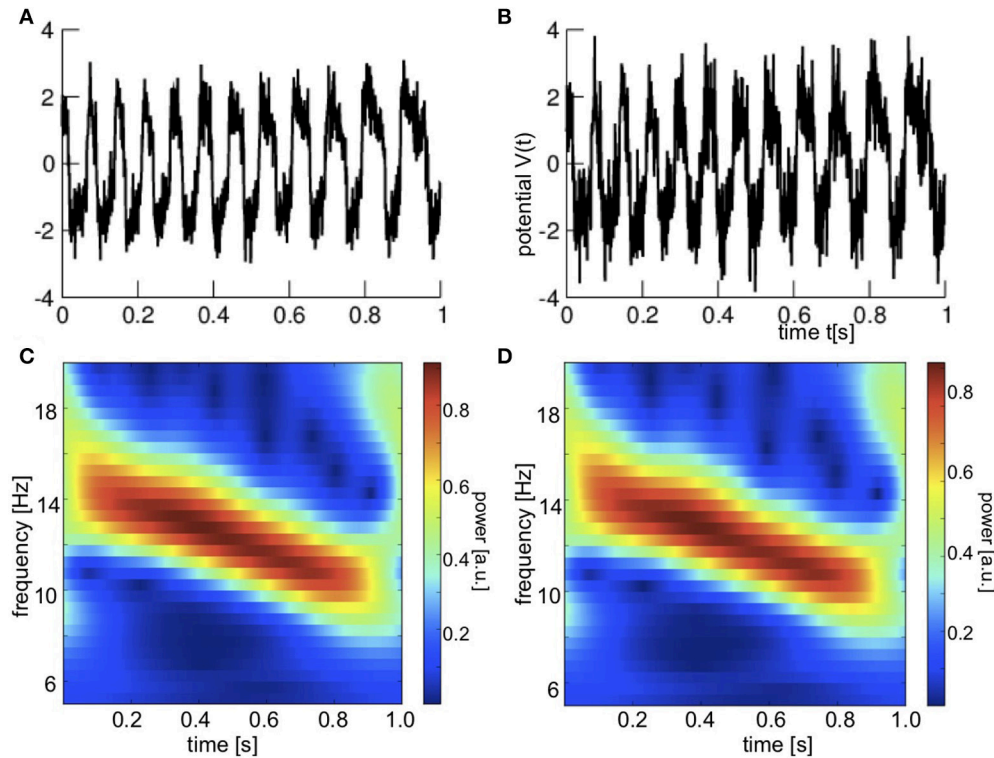


FIGURE 2 | Noisy *in-situ* observations. **(A)** Time series for medium noise level $\kappa = 0.5$. **(B)** Time series for large noise level $\kappa = 0.8$. **(C)** Time-frequency distribution for medium noise level $\kappa = 0.5$. **(D)** Time-frequency distribution for large noise level $\kappa = 0.8$.

the noisy observations under study. The oscillation frequency decreases corresponding to the non-homogeneous dynamics (2).

From Equation (4), one reads off the observation operator

$$H = (1 \ 0) \in \mathbb{R}^{1 \times 2}$$

with $y = \mathbf{H}\mathbf{x}$, $\mathbf{x} = (V, w)^t \in \mathbb{R}^2$.

For comparison, we also consider nonlocal observations with the observation operator

$$H = (1 \ 1) \in \mathbb{R}^{1 \times 2}$$

yielding

$$y(t) = V(t) + w(t) + \kappa \xi(t), \quad (5)$$

for the same noise levels κ as above. **Figure 3A** shows time series and corresponding time-frequency distributions. The frequency of the oscillation decreases over time similar to the *in-situ* observations.

As already stated, the aim of the present work is to introduce the idea to forecast temporal features. As a further step in this direction, let us consider temporal changes of the signal evolution, i.e., the speed of the system. To this end the definition of the observation operator \mathcal{H} is extended to

$$y(t) = \mathcal{H}\mathbf{x}(t)$$

with

$$\mathcal{H} = \begin{pmatrix} \frac{d}{dt} & 0 \end{pmatrix} \in \mathbb{R}^{1 \times 2} \quad (6)$$

yielding

$$y(t) = \frac{dV(t)}{dt} + \kappa \xi(t),$$

for two noise levels $\kappa = 0.0$ and $\kappa = 0.02$. Numerically, the derivative $dV(t)/dt$ is implemented as $V(t_n) - V(t_{n-1})$ at time instance t_n . **Figure 3B** shows the corresponding time series. We recognize the short time scale of the spike activity in *in-situ* observations as couples of sharp positive and negative spikes.

2.3. Ensemble Transform Kalman Filter

One of the major aims of data assimilation techniques is the optimal fit of model dynamics to observed data. Here, we introduce the major idea with a focus on the 2-dimensional model (1) and the scalar observation. Observations $y(t)$ evolve in the 1-dimensional observation space, while the model solutions are embedded in the 2-dimensional model phase space.

2.3.1. Analysis Ensemble

To merge observation $y(t)$ and model background state $\mathbf{x}_b(t)$ at time t optimally, the best new model state \mathbf{x}_a minimizes the cost

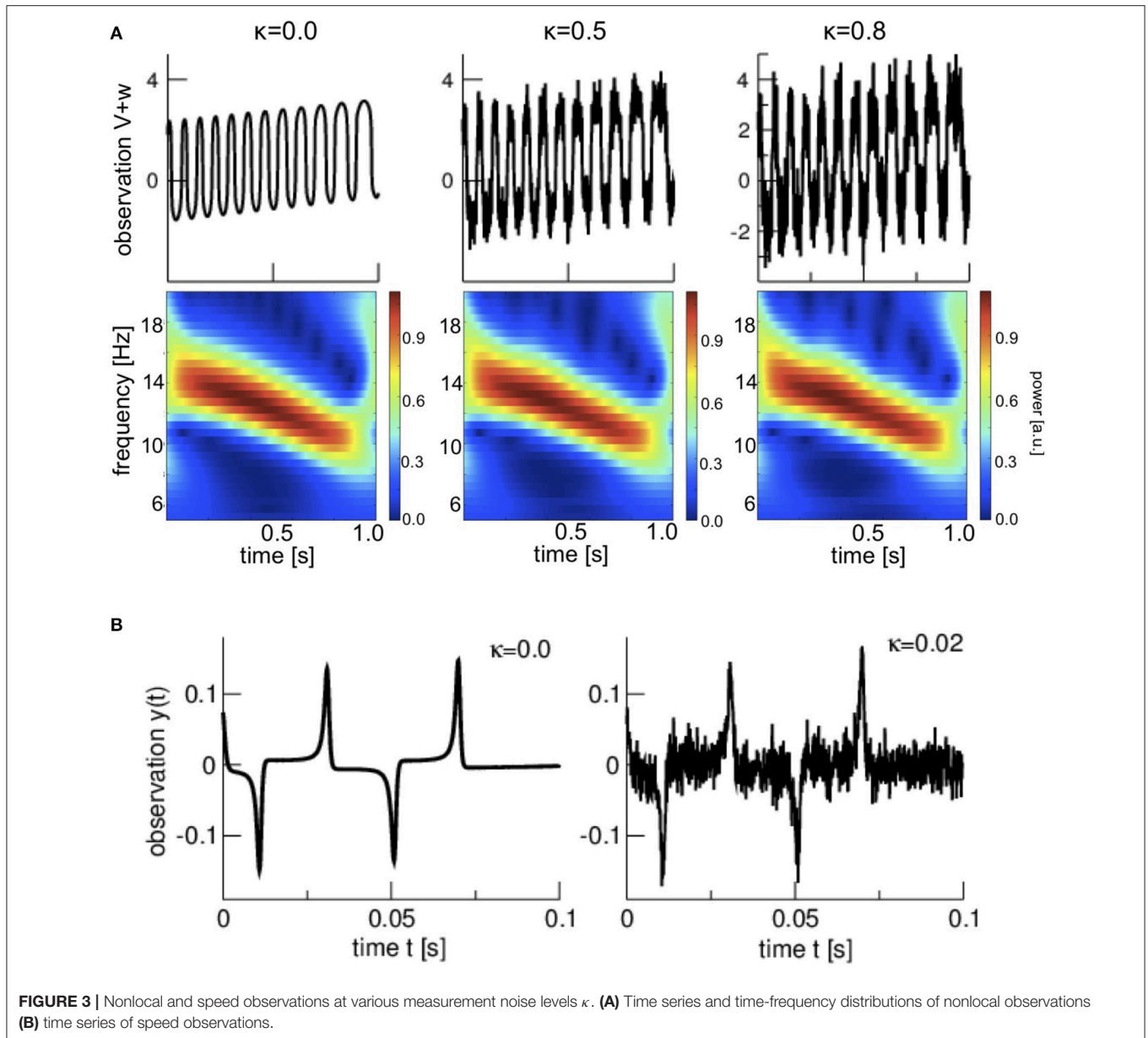


FIGURE 3 | Nonlocal and speed observations at various measurement noise levels κ . **(A)** Time series and time-frequency distributions of nonlocal observations **(B)** time series of speed observations.

function

$$C(\mathbf{x}_a) = (\mathbf{x}_a - \mathbf{x}_b)^t \mathbf{B}^{-1} (\mathbf{x}_a - \mathbf{x}_b) + (y - \mathbf{H}\mathbf{x}_a)^t (y - \mathbf{H}\mathbf{x}_a) / R$$

$$= \min!, \quad (7)$$

i.e., the solution is the minimum of the cost function C . Here, \mathbf{H} is the observation operator, \mathbf{x}_a is called the analysis, \mathbf{B} is the model error covariance matrix and R the observation error. If the assumed dynamical model and the assumed observation operator used in the data assimilation procedure are the true model and operator, respectively, then the assumed observation error is identical to the true error, i.e., $R = \kappa^2$. However, typically, one does not know the true observation error κ and R can just be estimated. This is the case we consider in the present work. In the present implementation $R = 1.5$. For given matrix \mathbf{B} and the

scalar R , the optimal new model state is

$$\mathbf{x}_a = \mathbf{x}_b + \frac{1}{R + \mathbf{H}\mathbf{B}\mathbf{H}^t} \mathbf{B}\mathbf{H}^t (y - \mathbf{H}\mathbf{x}_b). \quad (8)$$

This is the major result of the 3DVar technique for scalar observations [10].

Conversely, if the covariance error matrix \mathbf{B} is not known, it can be estimated from the model. To this end, one considers an ensemble of model states $\{\mathbf{x}_b^l\}$, $l = 1, \dots, L$ of L ensemble members and estimates \mathbf{B} by

$$\mathbf{B} \approx \frac{1}{L-1} \sum_{l=1}^L (\mathbf{x}_b^l - \bar{\mathbf{x}}_b) (\mathbf{x}_b^l - \bar{\mathbf{x}}_b)^t \quad (9)$$

$$= \frac{1}{L-1} \mathbf{X}\mathbf{X}^t \quad (10)$$

with the ensemble mean $\bar{\mathbf{x}}_b = \sum_{l=1}^L \mathbf{x}_b^l / L$ and $X_{kl} = (\mathbf{x}_b^l)_k$. In applications, we choose $L = 10$ if not stated differently. Introducing the equivalent of \mathbf{X} in observation space $\mathbf{Y} = \mathbf{H}\mathbf{X}$, Equation (8) reads in model space

$$\mathbf{x}_a - \mathbf{x}_b = \frac{1}{(L-1)R + \mathbf{Y}\mathbf{Y}^t} \mathbf{X}\mathbf{Y}^t (\mathbf{y} - \mathbf{H}\mathbf{x}_b) \quad (11)$$

and in observation space

$$y_a - y_b = \frac{\mathbf{Y}\mathbf{Y}^t}{(L-1)R + \mathbf{Y}\mathbf{Y}^t} (\mathbf{y} - y_b) \quad (12)$$

with $y_{a,b} = \mathbf{H}\mathbf{x}_{a,b}$. Since $\mathbf{Y}\mathbf{Y}^t$ and R are positive-definite scalars,

$$0 < \frac{y_a - y_b}{y - y_b} < 1 \quad (13)$$

stating that the analysis equivalent in observation space y_a is always closer to the observation as the background equivalent in observation space y_b .

The ensemble transform Kalman filter (ETKF) [22] optimizes observation and background ensemble members $\{\mathbf{x}_b^l\}$ to gain an analysis ensemble $\{\mathbf{x}_a^l\}$ in the ensemble space. This space is L -dimensional and is spanned by the ensemble members

$$\mathbf{x}_b = \bar{\mathbf{x}}_b + \mathbf{X}\mathbf{w}$$

with the ensemble space coordinates $\mathbf{w} \in \mathbb{R}^L$. Re-considering the optimization scheme (7) in this space

$$\bar{\mathbf{w}} = \mathbf{P}\mathbf{Y}^t (\mathbf{y} - \bar{y}_b) / R, \quad \mathbf{P} = ((L-1)\mathbf{I} + \mathbf{Y}^t \mathbf{Y} / R)^{-1} \in \mathbb{R}^{L \times L}$$

with $\bar{y}_b = \mathbf{H}\bar{\mathbf{x}}_b$ and the identity matrix $\mathbf{I} \in \mathbb{R}^{L \times L}$. Then the analysis ensemble mean $\bar{\mathbf{x}}_a$ and its covariance \mathbf{P}_a reads

$$\bar{\mathbf{x}}_a = \bar{\mathbf{x}}_b + \mathbf{X}\bar{\mathbf{w}} \quad (14)$$

$$\mathbf{P}_a = \mathbf{X}\mathbf{P}\mathbf{X}^t. \quad (15)$$

The analysis ensemble members can be calculated by

$$\mathbf{x}_a^l = \bar{\mathbf{x}}_a + \mathbf{X}\mathbf{w}_a^l, \quad (16)$$

with $\mathbf{w}_a^l \in \mathbb{R}^L, l = 1, \dots, L$. Let us define the deviations from the analysis mean

$$\mathbf{W}^l = \mathbf{w}_a^l - \bar{\mathbf{w}}, \quad l = 1, \dots, L \quad (17)$$

$$\mathbf{P}_a = \frac{1}{L-1} \mathbf{X}_a \mathbf{X}_a^t \quad (18)$$

corresponding to (10) and with $\mathbf{W}^l \in \mathbb{R}^L$. Defining the matrix $\mathbf{W} \in \mathbb{R}^{L \times L}$ with columns \mathbf{W}^l , the ansatz $\mathbf{P} = \mathbf{W}\mathbf{W}^t$, and Equation (15) yields $\mathbf{X}_a = \sqrt{L-1} \mathbf{X}\mathbf{W}$. With the singular value decomposition $\mathbf{P} = \mathbf{U}\mathbf{D}\mathbf{U}^t$, the orthogonal matrix \mathbf{U} and the diagonal matrix \mathbf{D} , essentially we gain

$$\mathbf{W} = \mathbf{U}\mathbf{D}^{1/2}\mathbf{U}^t,$$

where $D_{kk}^{1/2} = \sqrt{D_{kk}}$. This is the square-root filter implementation of the ETKF [23].

Equation (7) implies that all states, observations, covariances and operators are instantaneous. Extensions of this formulation are known, e.g., as the 4D-ENKF or the 4DVar [24–26]. Most of these previous extensions imply an instantaneous observation operator \mathcal{H} . In the previous section, we considered the speed of observations as the observations under study implying the temporal derivative of observed signals. This derivative is nonlocal in time and hence non-instantaneous. Here, we argue that the system evolves on a time scale that is much larger than the sampling time or, in other words, the sampling rate is high enough that the temporal derivative can be considered as being local in time. Consequently, Equation (7) may still hold in a good approximation.

2.3.2. Inflation

In each analysis step, the analysis equivalent in observation space y_a moves away from the model background state y_b closer to the observation y , cf. discussion of Equation (13). This assumes that observations reflect the true state. Of course, observations usually are erroneous due to measurement errors or errors in the observation operator. This is taken care of by the model error covariance matrix R . The uncertainty of the model state in observation space is described by the covariance estimator $\mathbf{Y}\mathbf{Y}^t$. However, the model has errors which are not completely reflected by the state estimate error covariance matrix $\mathbf{Y}\mathbf{Y}^t$, since this is calculated based on an ensemble of model forecasts with the same simulated model equations. To take care of the model error and draw the analysis closer to the background state, typically one enhances the ensemble spread by inflation.

For *in situ*- and nonlocal observations we have implemented constant multiplicative inflation by scaling \mathbf{w}_a^l in Equation (16) by a factor $\mathbf{w}_a^l \rightarrow 1.4 \cdot \mathbf{w}_a^l$. In addition, we employed additive covariance inflation by $\mathbf{B} \rightarrow \mathbf{B} + 0.15\mathbf{I}$ in Equation (10) with the 2×2 unity matrix \mathbf{I} . For speed observations, we have reduced the multiplicative inflation factor to 1.05 and the additive covariance inflation factor to 0.05.

2.4. Data Assimilation Cycling

Putting together models and data assimilation, the model evolution is controlled by observed data optimizing the initial state of the model iteration. Our data assimilation cycle starts with initial conditions from which the model evolves during the sampling interval. The model state after one sampling interval Δt is the background state or first guess \mathbf{x}_b . The subsequent data assimilation step estimates the analysis state \mathbf{x}_a that represents the initial state for the next model evolution step. In other words, data assimilation tunes the initial state for the model evolution after each sampling interval. Using the ETKF, this cycling is applied for all ensemble members which obey the model evolution and whose analysis state is computed in each data assimilation step. Initial ensemble member model states were $\mathbf{x}^l(0) = (\eta_1, \eta_2)^t, l = 1, \dots, L$ with random uniformly distributed numbers η_1, η_2 in the range $\eta_1, \eta_2 \in [0; 1]$.

2.5. Ensemble Prediction and Verification

The aim of the present work is to show how optimal forecasting can be done. Free ensemble forecasts are model evolutions over a time typically longer than the sampling time. This forecast time is called lead time. The initial state of the free forecasts are the analysis model states determined by data assimilation.

In the present work, we are interested in forecasts at every sample time instant. To this end we compute the model activity at a certain lead time. This forecast is computed for all ensemble members what renders it an ensemble prediction. The forecasts are solutions of the model $\mathbf{x}^f(t; t_a)$ at time $t \geq t_a$ with initial analysis state \mathbf{x}_a at time $t = t_a$ and lead time $T = t - t_a$. To compare them to observations, forecasts are mapped to observation space yielding model equivalents

$$y^f(t; t_a) = \mathbf{H}\mathbf{x}^f(t; t_a).$$

Later sections show free forecasts $y^f(t; t - T)$ with fixed lead time. In the following, model forecasts with the sampling time as lead time $T = \Delta t$ are called *first guess*.

Naturally, one expects that the forecasts diverge from observations with longer lead times but the question is which forecasts can still be trusted, i.e., are realistic. Essentially we ask the question how one can verify the forecasts. To this end, various metrics and scores have been developed [27]. Since most forecasts are validated against observations, metrics are based on model forecast equivalents in the observation space.

2.5.1. First Guess Departure Statistics

To estimate the deviation of forecast ensemble members $y_n^{f(l)}$ with forecast ensemble means $\bar{y}_n^f = \sum_{l=1}^L y_n^{f(l)}$, $n = 1, \dots, N$ from observations y_n , $n = 1, \dots, N$ of number N , we compute the mean error (bias)

$$\text{bias} = \frac{1}{N} \sum_{n=1}^N y_n - \bar{y}_n^f,$$

the root-mean square error

$$\text{rmse} = \sqrt{\frac{1}{N} \sum_{n=1}^N (y_n - \bar{y}_n^f)^2}$$

and the ensemble spread

$$\text{spread} = \frac{1}{N} \sum_{n=1}^N \sqrt{\frac{1}{L-1} \sum_{l=1}^L (y_n^{f(l)} - \bar{y}_n^f)^2}.$$

For scalar observations and corresponding forecasts, i.e., temporal time series, $y_n^f = y^f(t_n; t_n - T)$, $n = 1, \dots, N$ and N is the number of time points. Conversely, for time-frequency distributions $S(t, \nu)$ computed from the observation time series by a wavelet transform (cf. section 2.1) with K time points and M frequencies, $y_n^f = S(t_k, \nu_m)$, $k = 1, \dots, K$, $m = 1, \dots, M$, $n = (m-1)K + k$ and $N = KM$ is the number of all time-frequency elements.

The time-frequency distribution represents the spectral power distribution S at various time instances. Since spectral power is a positive-definite measure, the distance of two time-frequency distributions could be computed differently as a root mean square error. We can interpret the rmse as the Euclidean distance in high-dimensional signal space. However, the spectral power lies on a manifold in signal space and hence the distance between spectral power values is a Riemannian distance [28, 29]. Alternatively, the distance between time-frequency distributions may represent the temporal average of distances between two instantaneous power spectra $S_1(t_k, \nu)$, $S_2(t_k, \nu)$ at time instance t_k . A corresponding well-known distance measures is the time-averaged Itakura-Saito distance (ISD) [29, 30]

$$\text{ISD}_k = \frac{1}{M} \sum_{m=1}^M \frac{S_{obs}(t_k, \nu_m)}{S_{fc}(t_k, \nu_m)} - \ln \frac{S_{obs}(t_k, \nu_m)}{S_{fc}(t_k, \nu_m)} - 1,$$

$$\text{ISD} = \frac{1}{K} \sum_{k=1}^K \text{ISD}_k.$$

This distance measure is not symmetric in the spectral distributions and hence not a metric. As an alternative, one may also consider the log-spectral distance (LSD) [29, 31]

$$\text{LSD}_k = \sqrt{\frac{1}{M} \sum_{m=1}^M \left[10 \log_{10} \frac{S_{obs}(t_k, \nu_m)}{S_{fc}(t_k, \nu_m)} \right]^2}, \quad \text{LSD} = \frac{1}{K} \sum_{k=1}^K \text{LSD}_k$$

which has the advantage that it is symmetric in the distributions. In both latter measures S_{obs} and S_{fc} are the power spectra of observations and forecasts, respectively.

As pointed out above, we hypothesize that spectral features extracted from forecasts can be predicted in a better or more precise way than forecasts themselves. Since measurement noise plays an important role in experimental data, we evaluate predictions for medium and large noise levels κ compared to $\kappa = 0$. The *skill score* [32]

$$\text{SS}(\kappa) = 1 - \frac{\text{rmse}(\kappa)}{\text{rmse}(\kappa = 0)}, \quad \kappa = 0.5, 0.8$$

reflects the deviation of forecast errors at medium and large noise levels from noiseless forecasts. For $\text{SS} = 0$, forecasts have identical rmse and $\text{SS} < 0$ ($\text{SS} > 0$) reflects larger (smaller) rmse, i.e., worse (better) forecasts. The skill score SS is less sensitive to the bias as the rmse, and that also plays an important role in the evaluation of forecasts (similarly to the standard deviation). However, for small bias $\text{SS} > 0$ is a strong indication of improved forecasts.

According to Equation (10), the ensemble is supposed to describe well the model error. The ensemble spread represents the variability of the model and an optimal ensemble stipulates $\text{spread} = \text{rmse}$ [33]. The *spread-skill relation* [34]

$$\text{SSR} = \frac{\text{spread}}{\text{rmse}}$$

quantifies this relation. If $SSR > 1$, the ensemble spread is too large yielding bad estimates of the analysis ensemble and free forecasts, whereas $SSR < 1$ reflects a too small spread giving observations too much weight and yielding bad estimates of analysis ensemble and forecasts as well.

2.5.2. Ensemble Distribution Statistics

A representative forecast ensemble has the same distribution as the observations. This can be quantified by computing the rank of an observation in a forecast ensemble [35, 36]. If this rank is uniformly distributed, then the ensemble describes well the variability of the observations. Conversely, if the rank distribution has an U-shape (inverse U-shape) then most observations lie outside (inside) the range of the ensemble and the forecast ensemble is not representative. To estimate the shape of the rank distribution, we parameterize it by a beta-function

$$f(x) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1}, \quad x \in [0, 1]$$

with the gamma-function $\Gamma(x)$ and two parameters $\alpha, \beta > 0$. For a uniform distribution $\alpha = \beta = 1$, and U-shape (inverse U-shape) distributions have $\alpha, \beta < 1$ ($\alpha, \beta > 1$). Computing the sample of ranks $r \in [0, L]$ from the set of forecast ensembles and observations, their mean μ and variance σ^2 permits to estimate the function parameters by

$$\hat{\alpha} = \frac{\mu}{L} \left(\frac{\mu(L - \mu)}{\sigma^2} - 1 \right)$$

$$\hat{\beta} = \left(1 - \frac{\mu}{L} \right) \left(\frac{\mu(L - \mu)}{\sigma^2} - 1 \right).$$

The derived β -score [35]

$$\beta_c = 1 - 1/\sqrt{\hat{\alpha}\hat{\beta}}$$

equals 0 for a uniform distribution and $\beta_c > 0$ ($\beta_c < 0$) reflects the ensemble overestimation (underestimation) of the model uncertainty for an inverse U-shaped (U-shaped) distribution. In addition, the β -bias [35]

$$\beta_b = \hat{\beta} - \hat{\alpha}$$

quantifies the skewness of the rank distribution and $\beta_b = 0$ reflects symmetric distributions. β -bias values $\beta_b > 0$ ($\beta_b < 0$) reflect a weight to lower (higher) ranks and the majority of ensemble members is larger (smaller) than observations.

3. RESULTS

At first, we consider *in-situ* observations and evaluate the data assimilation cycle to illustrate some properties of the ETKF. Subsequently, we present forecasts for *in-situ* observations as time series and time-frequency distributions and evaluate the corresponding ensemble forecasts by statistical metrics well-known from verification in meteorology. To understand the specific nature of *in-situ* observations, subsequently we

also consider nonlocal observations and speed observations and present corresponding verification results. Eventually, we compute advanced statistical estimates specific for spectral power distributions and verify corresponding forecasts.

3.1. Data Assimilation Cycle—*in-situ* Observations

To start, we consider *in-situ* observations. **Figure 4** shows observations, the ensemble mean of first guess and analysis equivalents in observations space. We observe that the analysis (red) is always closer to the observation (black) than the first guess (blue). This validates Equation (13). Moreover, visual inspection tells that higher noise levels yields worse fits of the first guess and the analysis to observation. This will be quantified in more detail in later section 3.3.

To illustrate the ensemble evolution, **Figure 5** shows observations and the ensemble mean (blue solid line) and the single ensemble members (blue dots) of the first guess in an initial and final time interval. We observe that the ensemble starts with a narrow distribution while it diverges rapidly after several time steps. The ensemble spread about the ensemble mean reached after the initial transient phase remains rather constant over time.

3.2. Forecast—*in-situ* Observations

Now let us turn to the forecasts. In the data assimilation cycle, after one model step and hence one sampling time interval, the analysis is computed and initializes the phase space trajectory of the model evolution for the subsequent model step. In free forecasts $y^f(t; t_a)$, the model is integrated over a certain lead time $T = t - t_a$ initialized by the analysis at each time instant t_a . **Figures 6A–C** shows time series of observations and forecast ensemble mean equivalents for two lead times. For the short lead time $T = 10$ ms the first guess equivalent follows rather closely the observation, whereas it is phase-shifted to the observation for large lead time 40 ms. This holds true for all noise levels.

The time-frequency distribution of the observations and forecast equivalents is shown in **Figures 6D–F**. The time-frequency distribution of forecasts at short lead time resembles well the time-frequency distribution of observations, whereas prominent differences between large lead time-forecasts and observations occur, especially at the temporal borders.

3.3. Verification—*in-situ* Observations

To quantify the differences between forecasts and observations detected by visual inspection in section 3.2, we compute the forecast departure statistics subject to the lead time. **Figure 7A** shows that rmse of time-frequency data increases monotonically with lead time and it increases and finally decreases when based on time series data. The periodicity of rmse results from the increasing forecast-observation delay that increases with the lead time. Hence at a phase lag of π when the lead time is half the mean oscillation period the rmse is maximum. This explains why the two rmse minima have a temporal distance of ~ 70 ms what corresponds to one period of the mean system frequency of 14 Hz. Moreover, the bias decreases monotonically with the lead

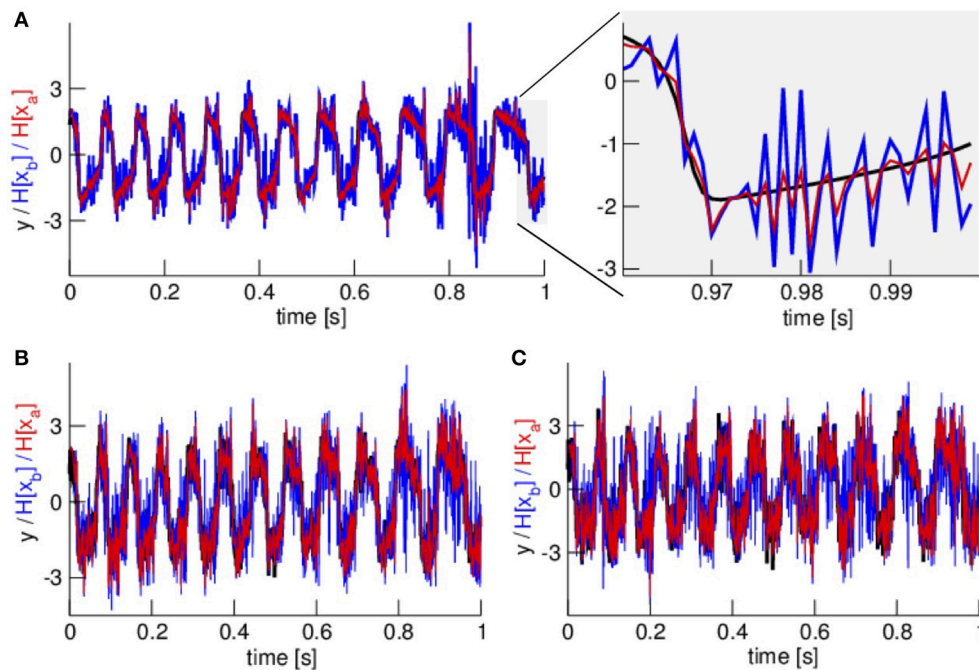


FIGURE 4 | *In-situ* observations y , ensemble mean of first guess $H[\bar{x}_b]$ and ensemble mean of analysis $H[\bar{x}_a]$ in observation space for no noise $\kappa = 0$ (A), medium noise level $\kappa = 0.5$ (B), and large noise level $\kappa = 0.8$ (C). In (A) the right panel zooms in a signal part illustrating that the analysis mean (in observation space, red color) is closer to the observations (black color) than the first guess mean (in observation space, blue color) in accordance to theory, see section 2. In all panels, observations are color-coded in black, first guess equivalents in observation space in blue and analysis equivalents in observation space in red.

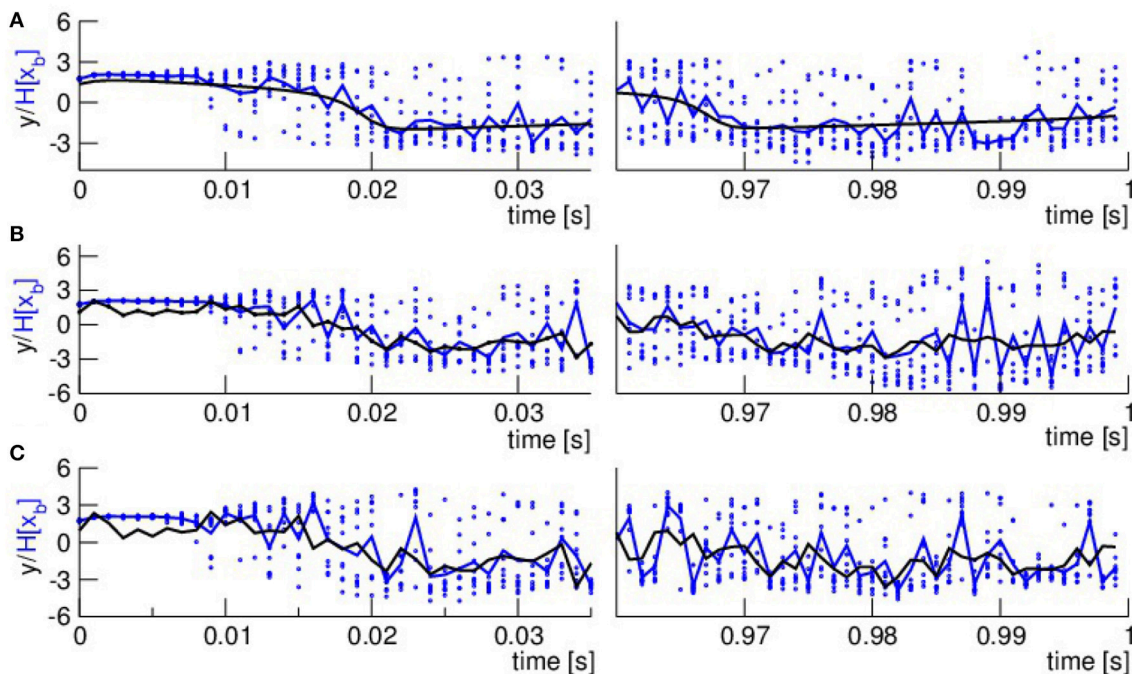


FIGURE 5 | Illustration of the temporal evolution of ensemble spread in observation space. (A) $\kappa = 0$, (B) $\kappa = 0.5$, (C) $\kappa = 0.8$. Observations are color-coded in black, the ensemble mean of the first guess is color-coded in blue and solid line and the single ensemble members are color-coded in blue and single dots.

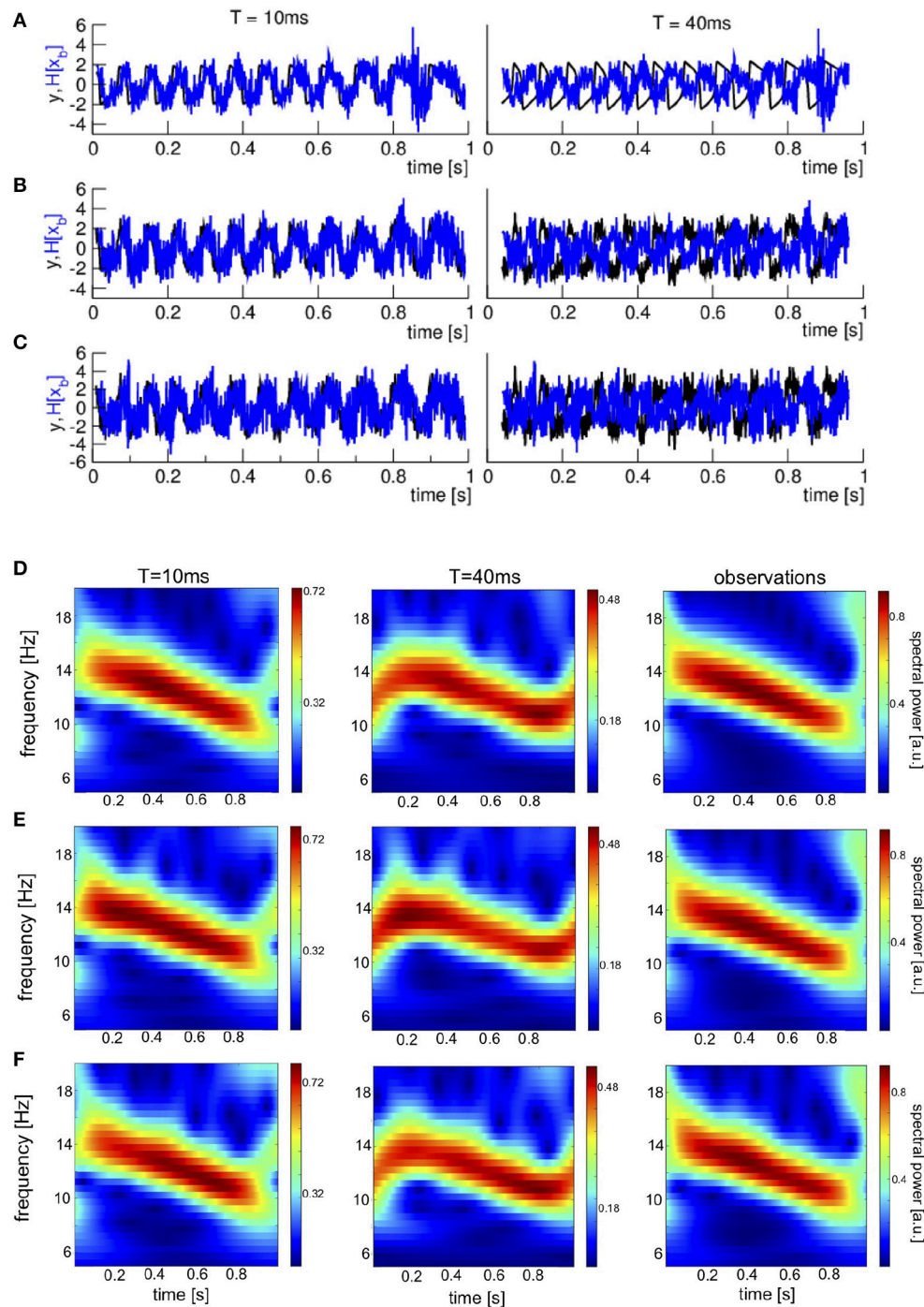


FIGURE 6 | *In-situ* time series and time-frequency distributions of free ensemble forecasts for two lead times T and different noise levels. **(A–C)** Time series, observations are color-coded in black, the ensemble mean of first guess equivalents in observation space $H[\tilde{x}_f]$ is color-coded in blue. **(D–F)** Time-frequency distributions. The forecasts shown are ensemble means of first guess equivalents in observation space $H[\tilde{x}_f]$. It is important to re-call that disturbances occur in a window of $4/f$ from the left and right temporal borders where f is the corresponding frequency, cf. section 2. **(A,D)** $\kappa = 0$, **(B,E)** $\kappa = 0.5$, and **(C,F)** $\kappa = 0.8$.

time for time series and increases for time-frequency data. To summarize these findings, we compute the skill score SS . Since the rmse for different noise levels approach each other for large

lead times the skill score approaches $SS = 0$ (Figure 7B). We observe that the skill score of time-frequency data exceeds SS of time series data.

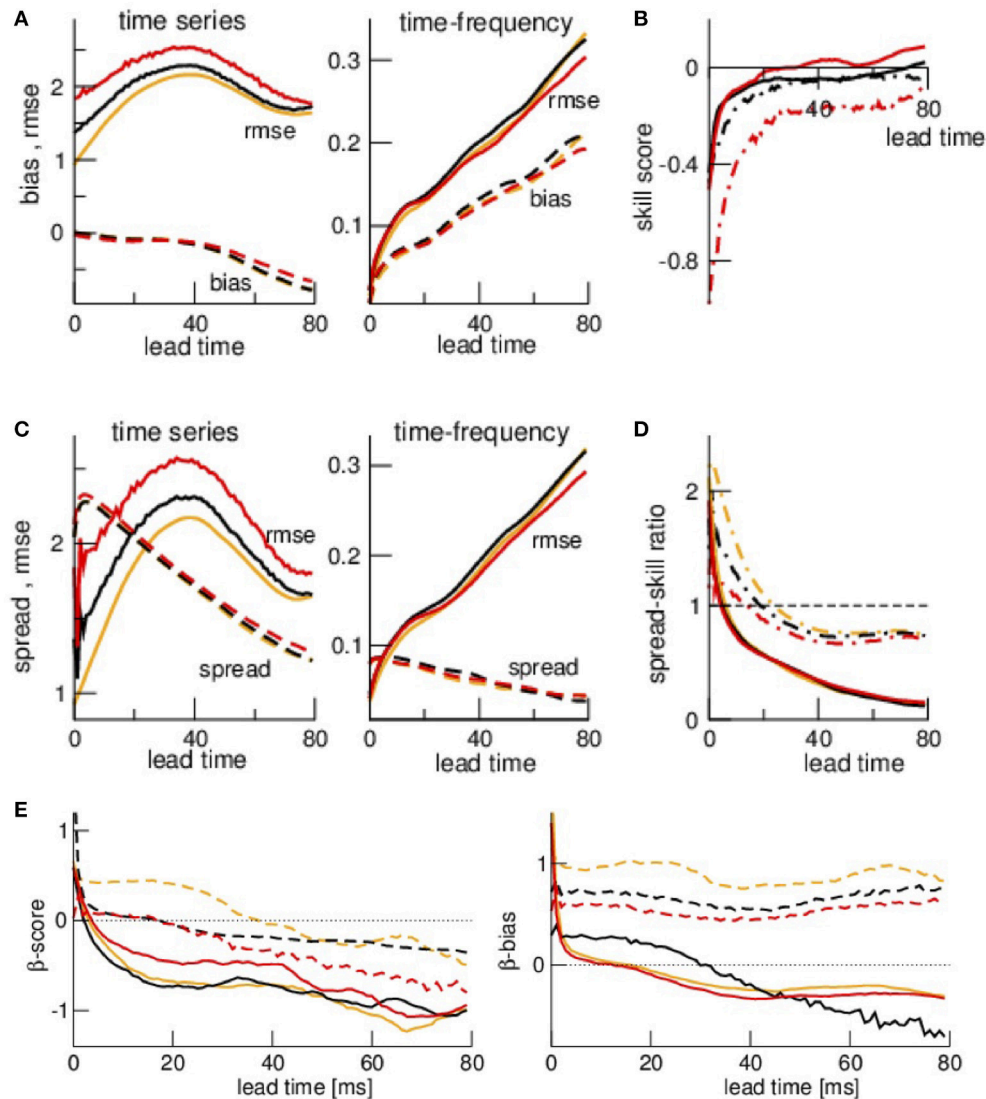


FIGURE 7 | Ensemble verification metrics and scores with respect to lead times. **(A)** rmse (solid line) and bias (dashed line) based on time series and time-frequency distributions. **(B)** Skill score $SS(\kappa) = 1 - \text{rmse}(\kappa)/\text{rmse}(0)$. **(C)** rmse (solid line) and spread (dashed line) based on time series and time-frequency distributions. **(D)** Spread-skill ratio $SSR = \text{spread}/\text{rmse}$. **(E)** Features of ensemble rank histogram β -score and β -bias with respect to lead times. Colors in **(A,C,D,E)** encode $\kappa = 0$ (orange), $\kappa = 0.5$ (black), and $\kappa = 0.8$ (red), line types in **(B,D,E)** encode time series data (dashed-dotted) and time-frequency distributions (solid). The estimates bias, rmse, and spread are averages over $N = 1,000$ observations for each lead time.

The ensemble spread decreases with lead time in both time series data and time-frequency data to values smaller than the rmse. This yields a decreasing spread-skill relation where SSR is well below $SSR = 1$ for both time series data and time-frequency distribution data. We note that SSR falls faster to lower values for time-frequency distribution data. Since one expects of good filters that the ensemble variations (spread) explain well the error (rmse), here the forecast ensemble of time series explains better the observations than time-frequency data since their SSR is closer to $SSR = 1$.

The reliability of the ensemble forecasts can be evaluated by rank histograms, i.e., the β -score β_c and β -bias β_b . **Figure 7E**

shows that β_c decreases from positive to negative values both for time series and time-frequency distribution data. This reveals an underestimation of the model uncertainty. The β -bias remains positive-definite for time series data whereas β_b of time-frequency distribution data decreases from positive to negative values. This result reveals that the majority of ensemble members are larger than the time series observations and smaller than the time-frequency spectral power observations.

To understand better why the ensemble spread shrinks at large lead time, **Figure 8** compares the ensemble mean of the model forecasts in phase space with the true phase space data. The forecasts exceed the true data at lead time $T = 1$ ms.

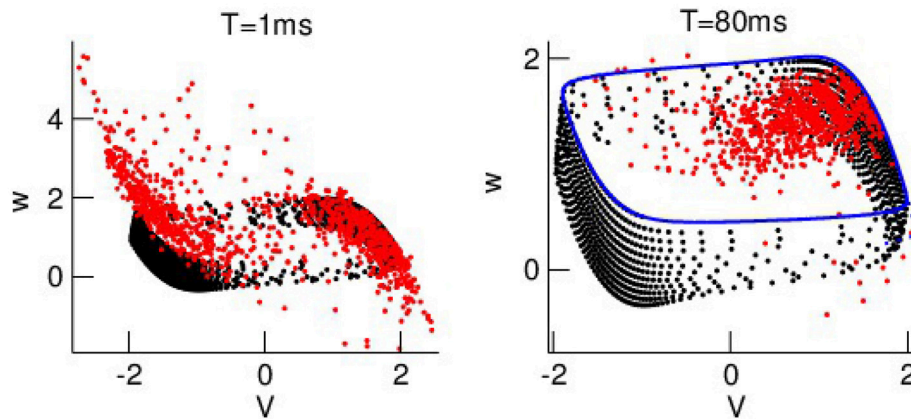


FIGURE 8 | Phase space dynamics for short and long lead times. The ensemble mean of forecasts and the true data are color-coded in red and black, respectively. The blue-coded points represent the false model data.

Conversely the forecast spread is much smaller than the true data at $T = 80$ ms since the forecasts obey the false model dynamics that evolves on a smaller phase space regime. Consequently the spread shrinkage with the lead time results from the smaller phase space regime of the false model.

3.4. Nonlocal Observations

To understand how specific the gained results from *in-situ* observations are, we compare them to statistics of other data type. Now let us consider nonlocal observations subjected to various noise levels. **Figures 9B–D** shows the time-frequency distributions for three noise levels and three lead times T . Forecasts at medium lead time T differ clearly to observations and forecasts at short and long lead time.

To understand this, we take a closer look at the forecast time series at $T = 40$ and compare it to the observations, cf. **Figure 9A**. Re-call that the analysis sets the initial condition for forecasts. For a lead time $T = 40$ ms the forecasts are in a fixed phase relation to an observation oscillations with $\nu_0 = 12.5$ Hz since then $T = 1/\nu_0$ is exactly one period of this oscillation. This fixed phase relation is observed in **Figure 9A** at ~ 0.5 s. Before and after that time, the observation frequency is larger and smaller, respectively, see also **Figure 6**, and the forecasts are out of phase. In addition, in the beginning and end the forecasts do not evolve rhythmically yielding missing spectral power, cf. **Figures 9B–D**. Summarizing, forecasts may resonate with oscillatory observations at frequency $\nu_0 = 1/T$.

The departure statistics between forecasts and observations resembles the findings for *in-situ* observations, cf. **Figure 10A**. Time-frequency distributions have almost optimal skill score SS for medium and large lead times, however with too small ensemble spread (SSR is very small). Conversely, time series data yield worse skill score but larger ensemble spread. Moreover, the rmse and bias have a maximum at about $T = 25$ ms and a minimum at about $T = 45$ ms. The minimum is explained

above as a resonance between forecast time and observation frequency.

These results are in good accordance to the rank histogram features β_c and β_b seen in **Figure 10B**. Very short lead times yield $\beta_c > 0$ reflecting an overestimation of the spread, otherwise $\beta_c < 0$ reflecting a too small ensemble spread. This holds true for all data types and all noise levels. The β -bias is similar to **Figure 7** and shows that the majority of the ensemble members is larger than the time series observations and smaller than the spectral power values.

Summarizing, the ensemble varies much with the lead time what indicates a fundamental problem in the ensemble forecast.

3.5. Speed Observations

Spectral power takes into account data at several time instances. Since to our knowledge Kalman filters have not been developed yet for observation operators nonlocal in time, we take a first step and consider speed observations subjected to two noise levels. **Figure 11A** compares observations, first guess and analysis in data assimilation cycling for the same number of ensemble members as in the previous assimilation examples. We observe that the first guess and analysis do not fit at all to the observations and hence the assimilation performs badly.

To improve the assimilation cycle, we diminish the observation error to $R = 0.01$ drawing the analysis closer to the observations. In addition, a larger ensemble improves the estimation of the model covariance inflation and we increase the number of ensemble members to $L = 50$ while decreasing the inflation factors to 1.05 (multiplicative inflation) and 0.05 (additive inflation). **Figure 11B** demonstrates that these modifications well improve the assimilation cycle. Now the first guess and analysis fit much better to the observations. An increased noise level renders the first guess and analysis less accurate.

The forecasts in **Figure 12** show that the assimilation cycle captures the upper observation spikes for $T = 40$ ms whereas forecasts at larger lead times are worse.

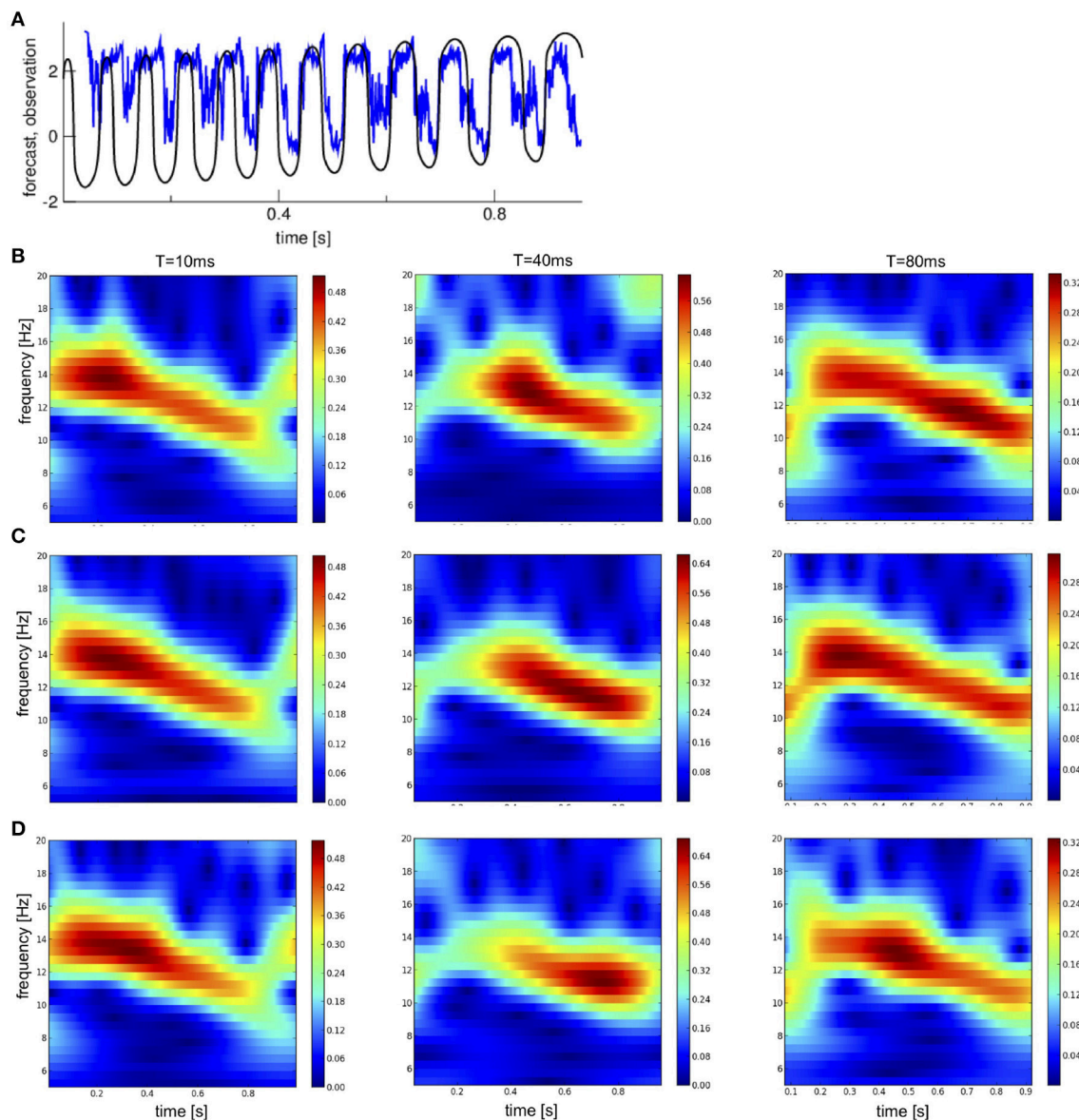


FIGURE 9 | Time series and time-frequency distributions of free ensemble mean forecasts compared to nonlocal observations. **(A)** Forecast for $T = 40$ ms (blue) and noise-free observations (black). **(B–D)** Forecasts for two lead times T and observations at different noise levels. **(B)** $\kappa = 0$, **(C)** $\kappa = 0.5$, and **(D)** $\kappa = 0.8$.

This can be quantified by departure statistics metrics as shown in **Figure 13**. The rmse increases slightly with lead time, i.e., the forecast error is larger for larger forecast times, while the bias is rather lead-time independent. Moreover, we observe that the spread is much smaller than the rmse. Since reliable ensemble forecasts should have a unity spread-skill ratio, this too small spread reflects a too small analysis inflation factor.

These results are in good accordance to the rank histogram features β_c and β_b seen in **Figure 14**. The negative values of β_c for all lead times reflects the underestimation of the spread and the β -bias $\beta_b \approx 0$ indicates that this underestimation is present for all forecast values. This holds true for both noise levels.

3.6. Advanced Statistical Measures

Since the rmse is not an optimal measure to quantify the difference between time-frequency distributions, we compute more advanced measures specific for power spectra. The Itakura-Saito distance (ISD) and the log-spectral distance (LSD) increase with the lead time for *in-situ* observations with a light local maximum at about $T = 40$ ms, cf. **Figure 15A**. A closer look at **Figure 6** reveals that the forecast spectral power at $T = 40$ is much smaller than the observation spectral power explaining this local increase of distance. The time-frequency distribution distances are rather similar in all noise levels. Moreover, spectral distances between nonlocal observations and forecasts

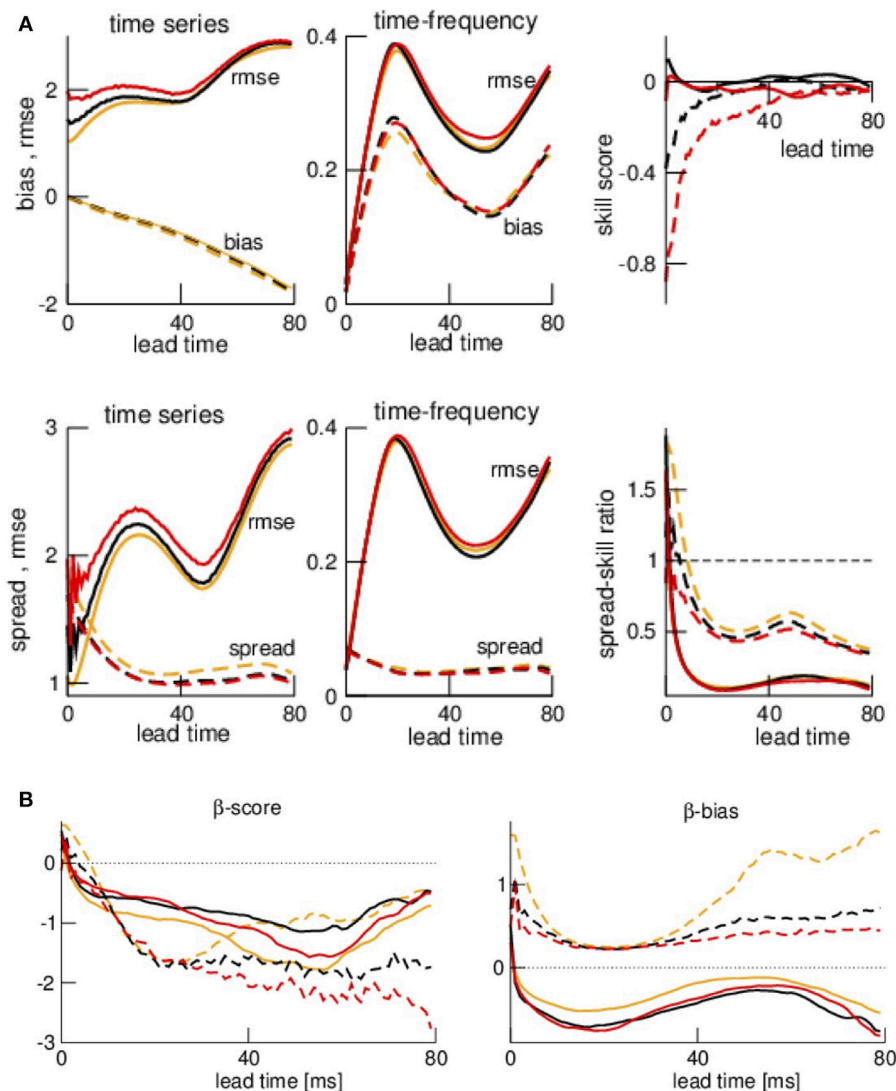


FIGURE 10 | The departure metrics Bias, rmse and spread of forecasts of nonlocal observations, the corresponding skill score SS and spread-skill ratio SSR (A) and the rank statistics β_c and β_b (B). The color- and line-coding is identical to the Figure 7. The estimates Bias, rmse, and spread are averages over $N = 1,000$ observations for each lead time.

exhibit a strongly non-monotonic dependence of the lead time. This is in good accordance to the results with the rmse in Figure 10.

Time frequency distributions appear to represent instantaneous spectral power. However, the spectral power distributions at subsequent time instances are strongly correlated dependent on the frequency. The correlation length is $\tau = 4/f$ leading to distortions at the temporal borders. Since the major spectral power occurs in the frequency interval [11 Hz; 15 Hz], i.e., for correlation times $0.27 \leq \tau \leq 0.36$, we define distorted time intervals with width 0.3 s and estimate improved time-frequency distribution distances neglecting the distorted initial and final time interval. Figure 15B shows the corresponding results. We observe that rmse, ISD and LSD depend similarly on the lead

time for both data types. Moreover, ISD and LSD are slightly smaller than their equivalents for the full time interval shown in Figure 15A.

4. DISCUSSION

The present work applies well-established techniques known in meteorology to find out whether they can be useful to forecast spectral features in other science domains where spectral dynamics plays an important role, such as in neuroscience. For *in situ*- and nonlocal observations, the assimilation of spectral features is indirect since the features are computed after the computation of conventional forecasts, i.e., in time series. We

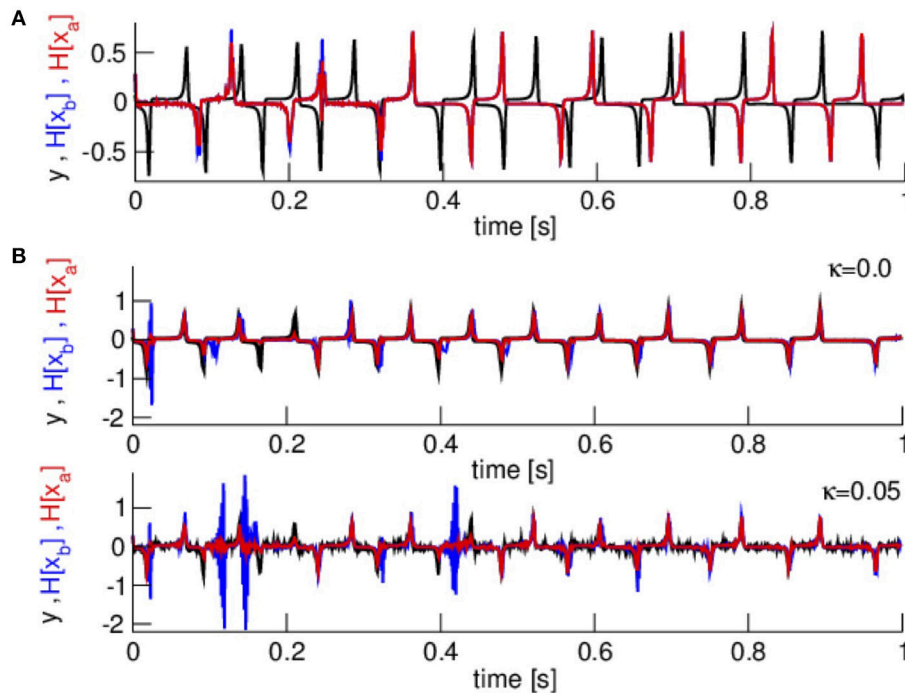


FIGURE 11 | Observations, first guess and analysis for speed data. **(A)** Noise-free observations and low number of ensemble members $L = 10$. **(B)** Two measurement noise levels and the modified assimilation parameters $R = 0.01$, $L = 50$, and the modified inflation factors.

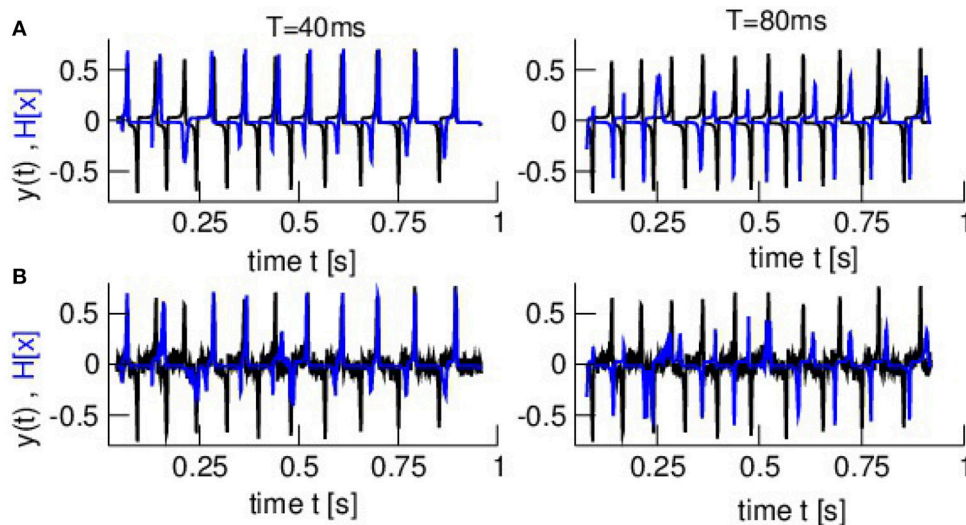


FIGURE 12 | Time series of observations and forecasts (model equivalent in observation space) at two lead times and two noise levels. **(A)** $\kappa = 0.0$ **(B)** $\kappa = 0.05$. Observations are color-coded in black and forecasts in blue. Parameters are identical to parameters in **Figure 11**.

show that they strongly improve skill scores (**Figures 7, 10**) for large lead times, whereas their spread is worse than for conventional forecasts for large lead times. This holds true for all measurement noise levels under study. In general, the ensemble forecast verification points to problems with the ensemble spread in all data types. This may result from a poor estimation of the

model error covariance **B** by too few ensemble members and a non-optimal choice of the inflation factor.

Since time-frequency distributions show time-variant spectral power, it is necessary to verify forecasts by spectral power-specific measures and take care of spectral power-specific artifacts, cf. **Figure 15**. The conventional estimate rmse and the

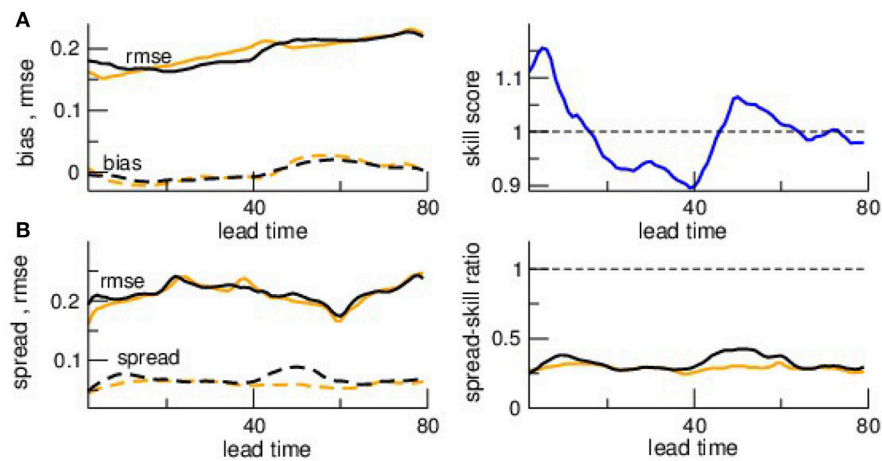


FIGURE 13 | The departure metrics bias, rmse, and spread of forecasts to speed observations and the corresponding skill score SS and spread-skill ratio SSR. Comparison of **(A)** bias and rmse and **(B)** spread and rmse. The skill score relates the rmse at both noise levels. The colors encode the noise level $\kappa = 0$ (orange) and $\kappa = 0.05$ (black). Parameters are identical to parameters in **Figure 11**.

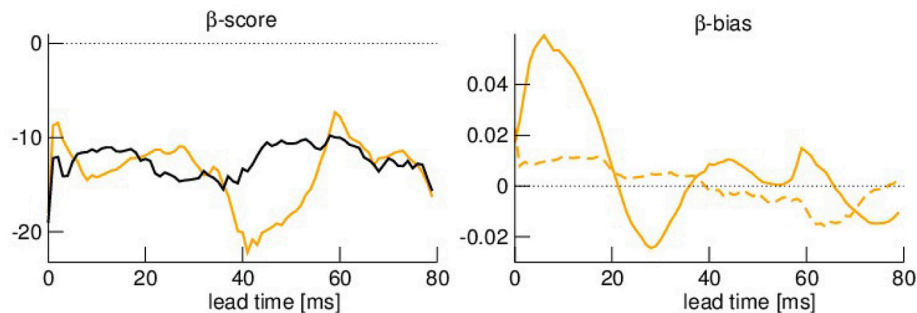


FIGURE 14 | Rank statistics β_C and β_B for the ensemble in the presence of speed observations. The color-coding is identical to **Figure 13**. Parameters are identical to parameters in **Figure 11**.

spectral-power specific estimates ISD and LSD behave similarly with respect to the lead time. Small differences between rmse and both ISD and LSD originates from the fact that ISD and LSD are time-averages over instantaneous spectral distance measures, whereas rmse averages over all frequencies and time instances and hence smoothes differences. Consequently ISD and LSD appear to be better verification measures of time-frequency distributions. Since LSD is a metric but ISD is not, future work will derive score measures based on LSD equivalent to the skill score SS. Moreover, we find that the border artifacts introduced by the wavelet transform do not affect our results qualitatively. Nevertheless, we recommend to exclude these artifacts in future work.

Conversely, speed observations consider the dynamical evolution of the system and are a very first approximation to a direct spectral feature. This is true since speed observations do not take into account the system state and observation at a single time instance only. Future work will extend this approach to a larger time window what allows to compute the power spectrum that can be mapped to a single time instance. Since generalizations or differential operators are integral

operators [37], future work will consider integral observation operators.

Since spectral feature forecasts are sensitive to certain frequencies, they are sensitive to lead time-observation frequency resonances. Such resonances seem to improve the forecast although these resonances are artifacts. To our best knowledge, the current work is the first to uncover these resonances that may play an important role in the interpretation of forecasts.

The ensemble data assimilation cycle involves several modern techniques, such as multiplicative and additive covariance inflation that well improves the forecasts. As a disadvantage, the spread for short lead times is too large. Future work will improve the ensemble statistics by adaptive inflation factors [38] and quality control methods, e.g., first guess checks [39] to remove outliers in every data assimilation step. This will surely contribute to improve ensemble forecasts.

The ensemble Kalman filter applied is one possible technique to gain forecasts. Other modern powerful techniques are the variational methods 3D- and 4D-Var [40], hybrids of ensemble and variational techniques like the EnVar [41] and particle filters [42, 43]. These techniques have been applied

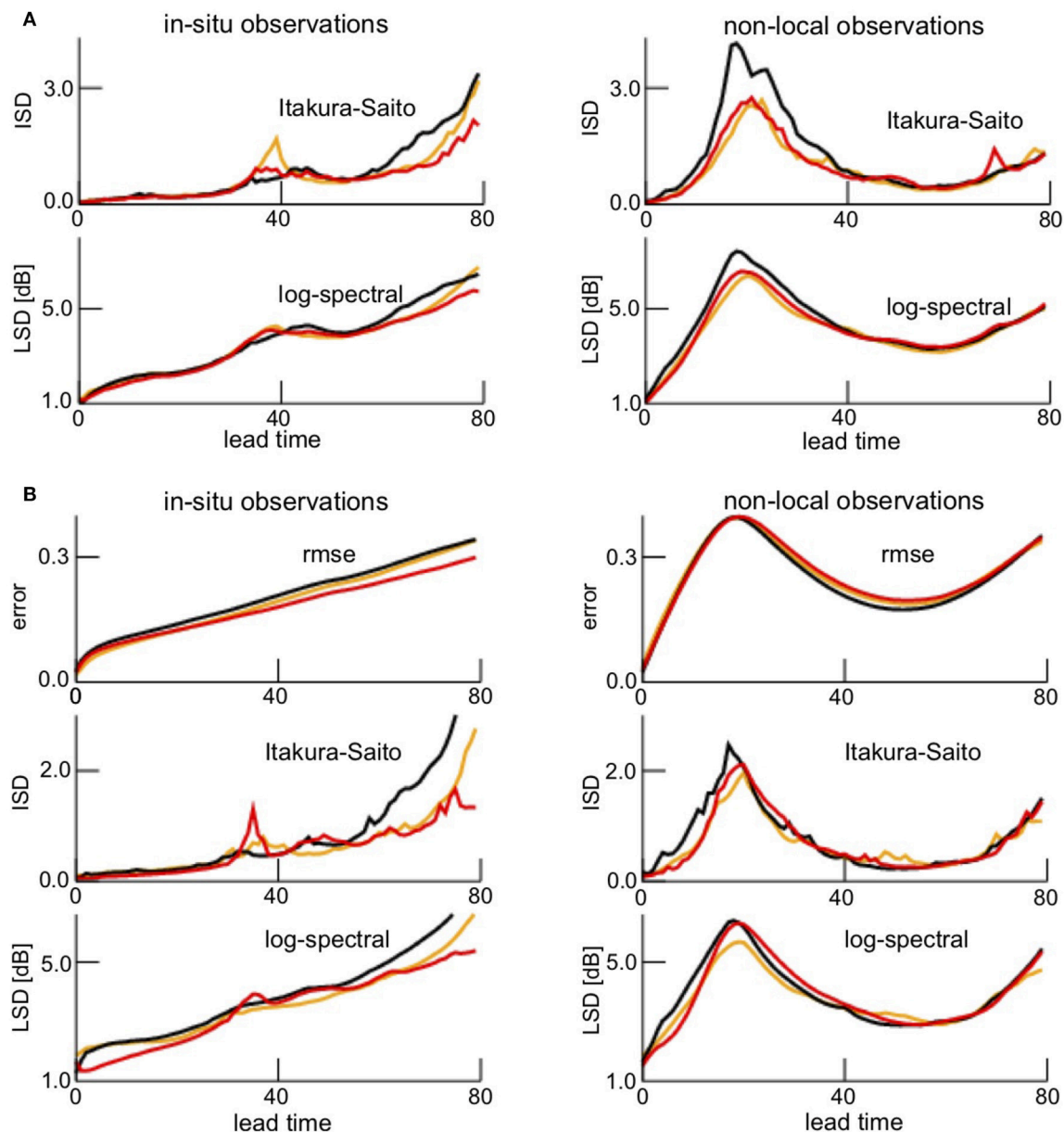


FIGURE 15 | Improved verification statistics of *in-situ* and non-local observations. **(A)** Itakura-Saito distance (ISD) and log-spectral distance (LSD) between forecasts and *in-situ* observations (left panel) and between forecasts and non-local observations (right panel) at various lead times. The distance measures are averages over the full time interval. **(B)** rmse, Itakura-Saito distance (ISD), and log-spectral distance (LSD). Here, the distance measures are averages over the constraint time interval [0.3 s; 0.7 s] to take care of the border effects generated the wavelet transform.

successfully in meteorological services world-wide and future work will investigate their performance in forecasting of power spectra.

Eventually, the present study considers a specific model system that exhibits a single time scale due to a single oscillation frequency. However, natural complex systems exhibit multiple time scales what may render the Kalman filter less effective and the superiority of the time-frequency data less obvious. In the future, it will be an important task to extend the present work to multi-scale Kalman filters [44, 45].

AUTHOR CONTRIBUTIONS

AH conceived the study and performed all simulations. AH and RP planned the manuscript structure and have written the manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Felix Fundel, Michael Denhardt, and Andreas Rhodin for valuable discussions.

REFERENCES

- Milton J, Jung P, (eds.). *Epilepsy as a Dynamic Disease*. New York, NY: Springer-Verlag (2002).
- Schiff S, Sauer T. Kalman filter control of a model of spatiotemporal cortical dynamics. *J Neural Eng.* (2008) 5:1–8. doi: 10.1088/1741-2560/5/1/001
- Schiff S. *Neural Control Engineering*. London, UK: MIT University Press (2011).
- Kalnay E. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge, UK: Academic Press (2002).
- Assi E, Nguyen D, Rihana S, Sawan M. Towards accurate prediction of epileptic seizures: a review. *Biomed Sign Proc Control* (2017) 34:144–57. doi: 10.1016/j.bspc.2017.02.001
- Park Y, Luo L, Parhi K, Netoff T. Seizure prediction with spectral power of EEG using cost-sensitive support vector machines. *Epilepsia* (2011) 52:1761–70. doi: 10.1111/j.1528-1167.2011.03138.x
- Chisci L, Mavino A, Perferi G, Sciandrone M, Anile C, Colicchio G, et al. Real-time epileptic seizure prediction using AR models and support vector machines. *IEEE Trans Biomed Eng.* (2010) 57:1124–32. doi: 10.1109/TBME.2009.2038990
- Ullah G, Schiff S. Assimilating seizure dynamics. *PLoS Comput Biol.* (2010) 6:e1000776. doi: 10.1371/journal.pcbi.1000776
- Sedigh-Sarvestani M, Schiff S, Gluckman B. Reconstructing mammalian sleep dynamics with data assimilation. *PLoS Comput Biol.* (2012) 8:e1002788. doi: 10.1371/journal.pcbi.1002788
- Nakamura G, Potthast R. *Inverse Modeling - An Introduction to the Theory and Methods of Inverse Problems and Data Assimilation*. Bristol, UK: IOP Publishing (2015).
- Alswailhi J, Potthast R, Bojak I, Saddy D, Hutt A. Kernel reconstruction for delayed neural field equations. *J Math Neurosci.* (2018) 8:3. doi: 10.1186/s13408-018-0058-8
- Hashemi M, Hutt A, Buhry L, Sleight J. Optimal model parameter estimation from EEG power spectrum features observed during general anesthesia. *Neuroinformatics* (2018) 16:231–51. doi: 10.1007/s12021-018-9369-x
- Fries P. Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu Rev Neurosci.* (2009) 32:209–24. doi: 10.1146/annurev.neuro.051508.135603
- Hutt A, (ed.). *Sleep and Anesthesia: Neural Correlates in Theory and Experiment*. No. 15 in Springer Series in Computational Neuroscience. New York, NY: Springer (2011).
- Roth M, Hendeby G, Fritsche C, Gustafsson F. The Ensemble Kalman Filter: a signal processing perspective. *EURASIP J Adv Sign Process.* (2017) 2017:56. doi: 10.1186/s13634-017-0492-x
- Evensen G. *Data Assimilation: The Ensemble Kalman Filter*. Heidelberg: Springer (2009).
- Roth M, Fritsche C, Hendeby G, Gustafsson F. The Ensemble Kalman Filter and its relations to other nonlinear filters. In: *Proceedings of the 2015 European Signal Processing Conference (EUSIPCO 2015)*, Institute of Electrical and Electronics Engineers (IEEE) (New York, NY) (2015). p. 1236–40.
- FitzHugh R. Mathematical models of threshold phenomena in the nerve membrane. *Bull Math Biophys.* (1955) 17:257–78.
- Nagumo J, Arimoto S, Yoshizawa S. An active pulse transmission line simulating nerve axon. *Proc IRE* (1962) 50:2061–270.
- Hutt A, Lefebvre J, Hight D, Sleight J. Suppression of underlying neuronal fluctuations mediates EEG slowing during general anaesthesia. *Neuroimage* (2018) 179:414–28. doi: 10.1016/j.neuroimage.2018.06.043
- Hashemi M, Hutt A, Sleight J. How the cortico-thalamic feedback affects the EEG power spectrum over frontal and occipital regions during propofol-induced anesthetic sedation. *J Comput Neurosci.* (2015) 39:155–79. doi: 10.1007/s10827-015-0569-1
- Hunt B, Kostelich E, Szunyogh I. Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter. *Phys D* (2007) 230:112–26. doi: 10.1016/j.physd.2006.11.008
- Tippett M, Anderson J, Bishop C, Hamill T, Whitaker J. Ensemble square root filters. *Mon Weather Rev.* (2003) 131:1485–90. doi: 10.1175/1520-0493(2003)131<1485:ESRF>2.0.CO;2
- Fertig E, Harlim J, Hunt B. A comparative study of 4D-Var and a 4D ensemble kalman filter: perfect model simulations with Lorenz-96. *Tellus* (2007) 59A:96–100. doi: 10.1111/j.1600-0870.2006.00205.x
- Le Dimet F, Talagrand O. Variational algorithm for analysis and assimilation of meteorological observations: theoretical aspects. *Tellus* (1986) 38A:97–110.
- Hunt B, Kalnay E, Ott E, Patil D, Sauer T, Szunyogh I, et al. Four-dimensional ensemble kalman filtering. *Tellus* (2004) 56A:273–7. doi: 10.3402/tellusa.v56i4.14424
- Mason S. Understanding forecast verification statistics. *Meteorol Appl.* (2008) 15:13040. doi: 10.1002/met.51
- Li Y, Wong KM. Riemannian distances for signal classification by power spectral density. *IEEE J Select T Sign Proc.* (2013) 7:655–69. doi: 10.1109/JSTSP.2013.2260320
- Jiang X, Ning L, Georgiou TT. Distances and Riemannian metrics for multivariate spectral densities. *IEEE Trans Auto Contr.* (2012) 57:1723–35. doi: 10.1109/TAC.2012.2183171
- Iser B, Schmidt G, Minker W. *Bandwidth Extension of Speech Signals*. Heidelberg: Springer (2008).
- Rabiner L, Juang B. *Fundamentals of Speech Recognition*. Upper Saddle River, NJ: PTR Prentice Hall (1993).
- Murphy AH. Skill scores based on the mean square error and their relationships to the correlation coefficient. *Monthly Weather Rev.* (1988) 116:2417–24.
- Weigel A. Ensemble verification. In: Joliffe I, Stephenson D, editors. *Forecast Verification: A Practitioner's Guide in Atmospheric Science*. 2nd Edn. Chichester: John Wiley and Sons (2011). p. 141–66.
- Hopson T. Assessing the ensemble spread-error relationship. *Mon Weather Rev.* (2014) 142:1125–42. doi: 10.1175/MWR-D-12-00111.1
- Keller J, Hense A. A new non-Gaussian evaluation method for ensemble forecasts based on analysis rank histograms. *Meteorol Zeitsch.* (2011) 20:107–17. doi: 10.1127/0941-2948/2011/0217
- Hamill T. Interpretation of rank histograms for verifying ensemble forecasts. *Mon Weather Rev.* (2000) 129:550–60. doi: 10.1175/1520-0493(2001)129<0550:IORHFV>2.0.CO;2
- Hutt A. Generalization of the reaction-diffusion, Swift-Hohenberg, and Kuramoto-Sivashinsky equations and effects of finite propagation speeds. *Phys Rev E* (2007) 75:026214. doi: 10.1103/PhysRevE.75.026214
- Anderson J. An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus* (2007) 59A:210–24. doi: 10.1111/j.1600-0870.2006.00216.x
- Geer A, Bauer P. Observation errors in all-sky data assimilation. *Q J R Meteorol Soc.* (2011) 137:2024–37. doi: 10.1002/qj.830
- Gauthier P, Tanguay M, Laroche S, Pellerin S, Morneau J. Extension of 3DVar to 4DVar: implementation of 4DVar at the Meteorological Service of Canada. *Monthly Weather Rev.* (2007) 135:2339–54. doi: 10.1175/MWR3394.1
- Bannister RN. A review of operational methods of variational and ensemble-variational data assimilation. *Q J R Meteorol Soc.* (2017) 143:607–33. doi: 10.1002/qj.2982
- van Leeuwen P.J. Particle filtering in geophysical systems. *Monthly Weather Rev.* (2009) 137:4098–114. doi: 10.1175/2009MWR2835.1
- Potthast R, Walter A, Rhodin A. A localised adaptive particle filter within an operational NWP Framework. (2018). *Monthly Weather Rev.* doi: 10.1175/MWR-D-18-0028.1. [Epub ahead of print].
- Hickmann KS, Godinez HC. A multiresolution ensemble Kalman filter using the wavelet decomposition. *arXiv:1511.01935* (2017).
- Nadeem A, Potthast R, Rhodin A. On sequential multiscale inversion and data assimilation. *J Comput Appl Math.* (2018) 336:338–52. doi: 10.1016/j.cam.2017.08.013

Conflict of Interest Statement: 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.

Copyright © 2018 Hutt and Potthast. 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.

Advantages of publishing in Frontiers



OPEN ACCESS

Articles are free to read
for greatest visibility
and readership



FAST PUBLICATION

Around 90 days
from submission
to decision



HIGH QUALITY PEER-REVIEW

Rigorous, collaborative,
and constructive
peer-review



TRANSPARENT PEER-REVIEW

Editors and reviewers
acknowledged by name
on published articles

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org

Contact us: info@frontiersin.org | +41 21 510 17 00



REPRODUCIBILITY OF RESEARCH

Support open data
and methods to enhance
research reproducibility



DIGITAL PUBLISHING

Articles designed
for optimal readership
across devices



FOLLOW US

@frontiersin



IMPACT METRICS

Advanced article metrics
track visibility across
digital media



EXTENSIVE PROMOTION

Marketing
and promotion
of impactful research



LOOP RESEARCH NETWORK

Our network
increases your
article's readership