



## OPEN ACCESS

## EDITED BY

Klaus Lehnertz,  
University of Bonn, Germany

## REVIEWED BY

Markus Wilhelm Abel,  
University of Potsdam, Germany  
Onera Korhonen,  
Aalto University, Finland

## \*CORRESPONDENCE

Jakub Sawicki,  
zergon@gmx.net

## SPECIALTY SECTION

This article was submitted to Networks in the Brain System, a section of the journal Frontiers in Network Physiology

RECEIVED 01 April 2022

ACCEPTED 11 July 2022

PUBLISHED 29 August 2022

## CITATION

Sawicki J, Hartmann L, Bader R and Schöll E (2022), Modelling the perception of music in brain network dynamics. *Front. Netw. Physiol.* 2:910920. doi: 10.3389/fnetp.2022.910920

## COPYRIGHT

© 2022 Sawicki, Hartmann, Bader and Schöll. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). 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.

# Modelling the perception of music in brain network dynamics

Jakub Sawicki<sup>1,2,3,4\*</sup>, Lenz Hartmann<sup>5</sup>, Rolf Bader<sup>5</sup> and Eckehard Schöll<sup>1,4,6</sup>

<sup>1</sup>Potsdam Institute for Climate Impact Research, Potsdam, Germany, <sup>2</sup>Institut für Musikpädagogik, Universität der Künste Berlin, Berlin, Germany, <sup>3</sup>Fachhochschule Nordwestschweiz FHNW, Basel, Switzerland, <sup>4</sup>Institut für Theoretische Physik, Technische Universität Berlin, Berlin, Germany, <sup>5</sup>Institute of Systematic Musicology, University of Hamburg, Hamburg, Germany, <sup>6</sup>Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität, Berlin, Germany

We analyze the influence of music in a network of FitzHugh-Nagumo oscillators with empirical structural connectivity measured in healthy human subjects. We report an increase of coherence between the global dynamics in our network and the input signal induced by a specific music song. We show that the level of coherence depends crucially on the frequency band. We compare our results with experimental data, which also describe global neural synchronization between different brain regions in the gamma-band range in a time-dependent manner correlated with musical large-scale form, showing increased synchronization just before transitions between different parts in a musical piece (musical high-level events). The results also suggest a separation in musical form-related brain synchronization between high brain frequencies, associated with neocortical activity, and low frequencies in the range of dance movements, associated with interactivity between cortical and subcortical regions.

## KEYWORDS

synchronization, coupled oscillators, neuronal network dynamics, pattern formation: activity and anatomic, external driven, electroencephalography (EEG)

## 1 Introduction

Dealing with the dynamics of neural networks, one repeatedly encounters the phenomenon of synchronization. In the brain, a high degree of synchronization is related to (slow-wave) sleep (Steriade et al., 1993; Rattenborg et al., 2000) or transitions from wakefulness to sleep (Schwartz and Roth, 2008; Moroni et al., 2012). Often, only a part of the brain is synchronized. This phenomenon of so-called partial synchronization Schöll (2021) has recently become a reference point for the explanation of unihemispheric sleep (Rattenborg et al., 2000, 2016; Mascetti, 2016; Ramlow et al., 2019) and the first-night effect (Tamaki et al., 2016), which describes troubled sleep in a novel environment. Furthermore, synchronized dynamics plays an integral role in the dynamics of epileptic seizures (Gerster et al., 2020), where the synchronization of a part of the brain causes dangerous consequences for the persons concerned. By contrast, synchronization is also used to explain brain processes serving the development of syntax and its perception (Koelsch et al., 2013; Large et al., 2015; Bader, 2020). Generally,

synchronization theory is of great importance for the analysis and understanding of musical acoustics and music psychology (Bader, 2013; Sawicki et al., 2018a; Hou et al., 2020; Shainline, 2020).

Although the neurophysiological processes involved in listening to music are still being researched, it is believed that some degree of synchrony can be observed in listening to music and building expectations. Event-related potentials, measured by electroencephalography (EEG) of participants while listening to music, show synchronized dynamics between different brain regions (Hartmann and Bader, 2014, 2020). These studies indicate that the synchronization dynamics represents musical large-scale form perception. The coupling of oscillatory neural signals within the usual frequency bands has been thought to be a mechanism that is related to a broad range of perceptual, sensorimotor, and cognitive processes, such as Gestalt perception and binding (Gray and Singer, 1989; Tallon et al., 1995; Keil et al., 1999; Rodriguez et al., 1999; Tallon-Baudry and Bertrand, 1999; Engel et al., 2001; Engel and Singer, 2001), timing and expectation (Buhusi and Meck, 2005, 2009), attention (Womelsdorf and Fries, 2007; Fries, 2009; Nikolić et al., 2013), consciousness (Baars, 2006; Dehaene et al., 2011; Engel and Fries, 2016; Owen and Guta, 2019), or motor functions (Thaut et al., 2015) as well as in music perception (Bhattacharya et al., 2001; Zanto et al., 2005; Bonetti et al., 2021).

According to (Engel and Fries, 2016), oscillatory brain activity is usually clustered into several frequency bands: delta (0.5–3.5 Hz), theta (4–7 Hz), alpha (8–12 Hz), beta (13–30 Hz) and gamma (> 30 Hz). Since the gamma-band is the ‘youngest’ frequency band which has become of interest (from about the late 1990s), the ranges and definitions vary from source to source. Here, we refer to the classification of (Freeman and Quian Quiroga, 2013), who speak of a low gamma range for frequencies above 30 Hz up to 60 Hz, and high gamma for frequencies above 60 Hz up to about 120 Hz. For everything above 120 Hz, we use the term ‘fast oscillations’ as employed by Buzsáki (2006). The gamma-band frequency range is of particular interest in the context of large-scale synchronization since it is thought to be a mechanism that integrates information from different parts of the cortex. In more detail, for specific frequency bands the increase and decrease of synchronization are following the large-scale form of the listened music in a coherent way. Moreover, it has been observed that areas of the whole brain are involved in neural dynamics during perception (Bader, 2020).

The musical form as the hierarchically highest level of musical structure and its perception is related to some of the mentioned processes above (Lerdahl and Jackendoff, 1990; Hartmann and Bader, 2020). Perceptually, notes, bars, and phrases are grouped and integrated into a high-level part of the form by the Gestalt laws (Leman, 1997; Deutsch, 2013; Neuhaus, 2013; Deliège and Melen, 2014). The contrast of the form’s parts, such as the concatenation of verse and chorus in a song, the sonata form of classical music, or the continuous night-

long tension build-up and decay in Techno, House or Electronic Dance Music, characterize the musical form and the learned knowledge about the underlying structures leads to the build-up of expectation and their fulfillment as well as to modulated attention. On an emotional level, this can be expressed in the terms of tension and relaxation (Koelsch, 2014; Lehne and Koelsch, 2015). Also, the transition from “potential energy” (expectations) into “kinetic energy” (dancing) as proposed by (Kurth, 1931) can be related to the processing of musical form in the sense of entrainment of neurons in the motor cortex by neurons from the auditory cortex (Thaut et al., 2015).

The characteristic of contrasting parts can be revealed not only by music analysis using pen and paper but also by different computational methods by the music information retrieval discipline, like the amplitude of a piece of music that corresponds to the subjective perception of loudness. Also other properties of the stimulus, such as the spectral centroid that corresponds to the perceived brightness of a sound, or the fractal correlation dimension (Grassberger and Procaccia, 1983a,b) corresponding to the perceived density and thereby representing the complexity of a piece of music, are drivers of the musical form (Bader, 2013; Hartmann and Bader, 2020; Bader, 2021; Bader et al., 2021; Linke et al., 2021).

Recently, the general influence of sound on a dynamical system with complex network connectivities (derived from empirical Diffusion Tensor Imaging (DTI) measurements) has been investigated (Sawicki and Schöll, 2021). It has been shown that an external sound source, which is connected to the auditory cortex of the human brain, induces partial synchronization patterns. Nevertheless, this study has neglected the complexity of music and its distinct effects in different frequency bands within the brain oscillations. There are a variety of recognized modeling approaches with respect to neural systems in general (Kacprzyk and Pedrycz, 2015; Bassett and Sporns, 2017; Bassett et al., 2018; Petkoski et al., 2018; Petkoski and Jirsa, 2019) and related to music in particular (Friston and Friston, 2013). In this paper, we model the spiking dynamics of the neurons by the paradigmatic FitzHugh–Nagumo model, and investigate possible coherence between the dynamics of the brain network and an external music source, which is connected to the auditory cortex of the human brain. Moreover, we present experimental data which we successfully reproduce numerically with the help of our network model, which combines simple node dynamics with complex network connectivities derived from empirical measurements.

An intriguing synchronization phenomenon in multilayer networks is relay synchronization between layers which are not directly connected, and interact via an intermediate (relay) layer (Leyva et al., 2018). Multilayer networks can give a general framework to describe and model real life examples of various systems, e.g., the two hemispheres of the brain or two cortical regions connected by the hippocampus (Gollo et al., 2011). Relay synchronization, a regime where pairs of nodes synchronize

despite their large distances on the network graph, has been shown to depend on the network symmetries (Bergner et al., 2012; Nicosia et al., 2013; Gambuzza et al., 2013; Zhang et al., 2017a,b). Recently the notion of relay synchronization has been extended from completely synchronized states to partial synchronization patterns. It has been shown that the multilayer structure of a network allows for (partial) synchronization in the outer layers via the relay layer (Sawicki et al., 2018b,c; Sawicki, 2019; Winkler et al., 2019; Drauschke et al., 2020; Sawicki et al., 2021).

Going towards more realistic models, time-delay plays an important role in the modeling of the dynamics of complex networks. In brain networks, the communication speed is affected by the distance between regions and therefore a stimulus applied to one region needs time to reach a different region. In such delayed system, it is possible to predict if the effects of stimulation remain focal or spread globally (Muldoon et al., 2016). More generally, time delays due to propagation over the white-matter tracts have been shown to organize the brain network synchronization dynamics for different types of oscillatory nodes (Petkoski and Jirsa, 2019). Within the scope of this paper, we focus on the requirements for a simple model to exhibit partial synchronization patterns, which have been experimentally observed (Hartmann and Bader, 2014, 2020). Therefore, we defer the consideration of time delays for now.

This article is organized as follows. In Section 2, we discuss the transformation of music to a neural input signal using a detailed cochlea model. In Section 3, we introduce the neural network model based upon empirical connectivities with neural input to the auditory cortex generated by music. In Section 4, we introduce some methods to characterize the neural output. Section 5 presents the results of the computer simulations and discusses the dynamical scenarios. Section 6 presents a comparison with experiments on human subjects, and Section 7 finally concludes.

## 2 From sound to neural spikes

The transformation of sound into neural spikes is the subject of much current research (Tritsch et al., 2010; Mizrahi et al., 2014; Bader, 2015, 2017, 2018; Guo et al., 2021). Music, speech, or any sound enters through the outer and middle ear as sound pressure, then acting on the oval window of the cochlea. The movement of the oval window is then transferred to a pressure in the lymph liquid of the cochlea surrounding the basilar membrane, which again acts on the basilar membrane, causing traveling waves there. Due to spatial differences in stiffness and damping on the membrane, sinusoidal waves with a single frequency show an increase in amplitude up to a point with maximum amplitude, the position of the so-called best-frequency, with a fast decay afterwards. Therefore, different positions on the basilar membrane represent different

frequencies, making the cochlea a Fourier analyzer. The stereocilia on the basilar membrane at the position of respective best-frequency are then transferring the mechanical energy into neural spikes. The frequency distribution on the basilar membrane is logarithmic. Movements of neighboring frequencies lead to interactions, causing roughness perception up to a frequency band of a musical major third. These bands are called critical bands, and the basilar membrane consists of 24 such bands. The spikes leaving the respective bands are fed into the auditory pathway, consisting of several neural nuclei, where the nucleus cochlearis or the trapezoid body are the first two. The interaction between these neural nuclei is manifold with several feedback loops and binaural connections (Schofield, 2011) ending at the auditory cortex on both hemispheres. Still up to the A1 region of the auditory cortex, the critical bands are maintained, where neural connections of higher nuclei are connected to bands on the basilar membrane, which is called tonotopy.

Many auditory features are present, extracted, or perceived already in this pathway, like sound localization, pitch, or timbre (Lyon and Shamma, 1996), although research has not concluded on further processing in the cortex (Bader, 2021). Music perception of larger temporal content, like song or sonata form, are not part of processing in the auditory pathway up to the cortex, as far as we know. Still the feedback loops within the pathway are both directions, up and down, afferent and efferent, so e.g. there is one connection down from the cortex to the cochlea with only one nucleus in between, tuning the basilar membrane tension through efferent nerves, according to cortex activity (Schofield, 2011).

Up to now, no model of the whole auditory pathway exists on a detailed neural level. The model used in this paper therefore concentrates on main findings, i.e., the transition from sound to neural spikes, the tonotopy of neural connections up to the cortex, as well as partial synchronization of phases in the cochlea, which are also present as coincidence detection in the auditory pathway. A Finite-Difference Time Domain (FDTD) physical model of the cochlea is used (Bader, 2015). The basilar membrane is about 3.5 cm long and only between 0.1–0.12 cm wide, so it is more a rod than a membrane. Therefore, the present model assumes a differential equation of a membrane like

$$\frac{K(x)}{\mu(x)} \frac{\partial^2 u}{\partial x^2} - d(x) \frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial t^2} + f(t), \quad (1)$$

with basilar membrane displacement  $u$  along a one-dimensional axis  $x$ , basilar membrane stiffness  $K(x) = 2 \times 10^9 e^{-3.4x}$  dyn/cm<sup>3</sup> changing along  $x$ , and linear mass density  $\mu(x) = m/A(x)$  with mass  $m$  over cross section  $A$  again changing along the basilar membrane and  $A(x) = 0.1 \text{ cm} \times (0.1 \text{ cm} + 0.02 \text{ cm} \times x/l)$  with basilar membrane length  $l = 3.5 \text{ cm}$  taking into account the slight widening of the basilar membrane over its length. The boundary conditions of the basilar membrane are homogeneous Dirichlet boundary conditions which do not allow for displacements on

the boundaries, but any derivative is allowed in accordance with the physiological conditions. Comparison between a membrane and a rod model shows no considerable differences, therefore a rod model is used. Here  $d$  is damping, and  $f(t)$  is the driving force of the lymph fluid which drives the basilar membrane.

To calculate the spikes omitted by the cochlea, the recording of the musical piece used is fed into the cochlea model. Here the amplitudes of the digital musical sound file are taken as sound pressures acting on the oval window of the cochlea and therefore immediately on the peri- and endolymph around the basilar membrane. As the speed of sound in the lymph ( $\sim 1,500$  m/s) is much larger than the speed of waves on the basilar membrane which is between  $\sim 100$  m/s at the oval window and down to  $\sim 10$  m/s at the helicotrema, an instantaneous action of the pressure at the oval window on the basilar membrane is reasonable and known as long-wave approximation (de Boer, 1991). This holds for frequencies up to  $\sim 4$  kHz, where pitch perception stops and humans only hear a very high sound. This approximation is used in the model. It leads to the force  $f(t)$  in Eq. 1 which represents the amplitudes of the digital musical sound file acting instantaneously on all points of the basilar membrane at each time point respectively. It is interesting to see that the traveling wave on the basilar membrane is therefore not caused by an external input slowly traveling through the cochlea but by the intrinsic solution of the inhomogeneous differential equation of the basilar membrane driven by a periodic force over its whole length instantaneously.

Depending on the brain region, neurological measurements reveal different time scales (Spitmaun et al., 2020). In our work we choose 50 ms as a time integration step as this is consistent with a characteristic time scale in music as well as in visual perception. In music 50 ms correspond to the second integration time, below which two events cannot be distinguished one from another. This leads to a threshold of 20 Hz, above which musical pitches are perceived and below which adjacent events are heard as rhythms. In vision, 18–24 frames per second lead to a continuous visual perception, again corresponding to about 50 ms time intervals. Therefore, in terms of hearing and seeing, the brain seems to update perceptual input on a time-scale of 50 ms (Bader, 2013).

The transition between mechanical displacement and electrical spike is performed using two conditions according to literature (Hubbard and Mountain, 1996). A neural spike at one point  $X$  on the basilar membrane at time  $\tau$  is excited if two conditions hold.

$$u(X, \tau) > u(X - 1, \tau), u(X + 1, \tau) \quad (2a)$$

$$u(X, \tau) > u(X, \tau - 1), u(X, \tau + 1). \quad (2b)$$

Condition (2a) means a maximum shearing of two nervous fibers as a necessary condition to an opening of the ion channels at the fibers. This only happens with a positive slope, as only then the stereocilia are driven away from each other. With a negative slope the cilia are getting closer and therefore no stress appears at the

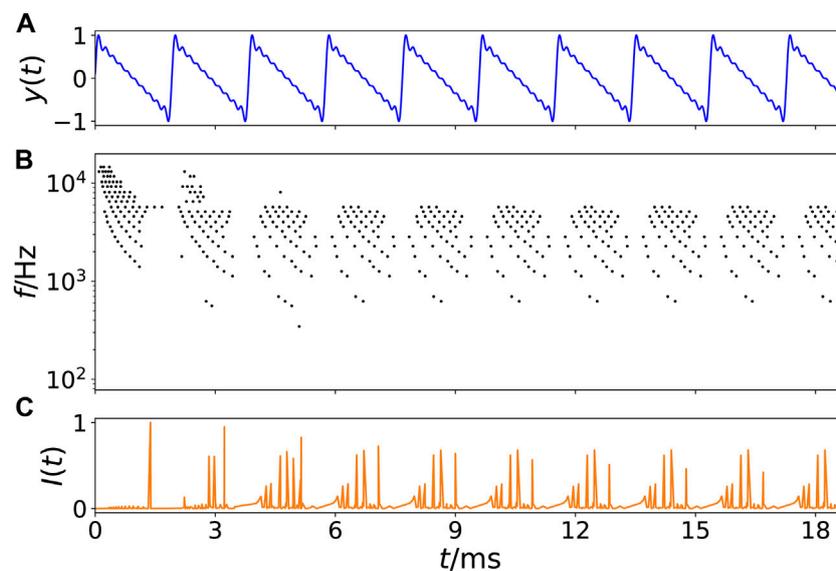
tip links between them. This corresponds to the rectification process in gammatone filter banks. Condition (2b) is a temporal maximum positive peak of the basilar membrane displacement. It is the temporal equivalent to the spatial condition of a maximum acceleration, where the tip link between the cell and its neighboring cells is most active.

To calculate the spikes omitted by the cochlea, the recording of the musical piece used is fed into the cochlea model. Therefore, the original piece, available as a digital recording of 44.1 kHz sample rate (CD-Quality) is upsampled to 192 kHz to meet Finite-Difference Time Domain (FDTD) stability criteria. The cochlea model is then run with a time-discretization step of  $\Delta t = 1/192,000$  s. Each time when a neural spike appears, the time point, strength, and critical band of the spike is stored. Therefore, after processing, a time series  $I(t)$  of all spikes leaving the cochlea is obtained.

Figure 1A displays an example of an artificially generated so-called *tone complex* with  $f_0 = 475$  Hz and ten partial tones (harmonics) with amplitudes  $1/m$  where  $m = 1, 2, 3, \dots, 10$ . The respective spike output of the basilar membrane model is shown in Figure 1B. Each time when the sound wave has a maximum amplitude, a pressure pulse is traveling over the basilar membrane, which emits electrical spikes at respective best-frequency positions on the membrane in accordance with the frequencies in the activating sound. As traveling waves on the membrane start at the basal end, next to the oval window, where high frequencies have their best-frequency location, and travel down the membrane towards the upper end, the helicotrema, where low frequencies are located, low frequencies show a time-delay with respect to higher frequencies. If the spikes of all critical bands are summed up for a certain point in time, a time series  $I(t)$  of all neural spikes leaving the cochlea can be generated, as exemplarily shown in Figure 1C. The simplification that the output of the cochlea model is summed up at one time point is motivated by the results of (Joris et al., 1994): In an experiment with cats, the authors could show that the scattered output of the cochlea is synchronized in the trapezoid body.

### 3 Neural network model

In this section, we introduce an empirical structural brain network as shown in Figure 2A where every region of interest is modeled by a single FitzHugh-Nagumo (FHN) oscillator. The weighted adjacency matrix  $\mathbf{A} = \{A_{kj}\}$  of size  $90 \times 90$ , with node indices  $k, j \in N = \{1, 2, \dots, 90\}$  was obtained from averaged diffusion-weighted magnetic resonance imaging data measured in 20 healthy human subjects. For details of the measurement procedure including acquisition parameters, see (Melicher et al., 2015), for previous utilization of the structural networks to analyze chimera states see (Chouzouris et al., 2018; Ramlow et al., 2019; Gerster et al., 2020; Schöll, 2021). The data were analyzed using probabilistic tractography as implemented in the



**FIGURE 1**

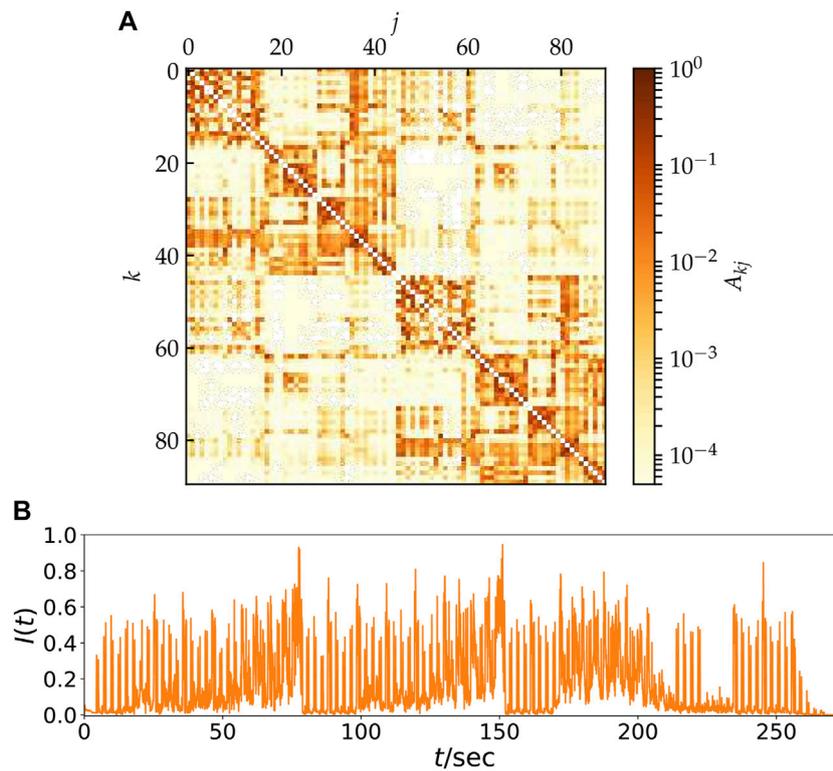
Example of transformation of a sound wave into a spike pattern of the cochlea model. **(A)** Time series of an artificially generated tone complex  $y(t)$  versus time  $t$  in ms with  $f_0 = 475$  Hz and ten partial tones (harmonics) with amplitudes  $1/m$  where  $m = 1, 2, 3, \dots, 10$ . **(B)** Spikes (black dots) leaving the cochlea as calculated from the model (Bader, 2015), where the vertical axis represents the cochlea position with best-frequency  $f$  in Hz indicated, i.e., categorized into 24 so-called critical bands. **(C)** Time series  $I(t)$  of the sum of all spike weights leaving the cochlea at a certain time  $t$ . Note that the first 5 ms are transients.

FMRIB Software Library, where FMRIB stands for Functional Magnetic Resonance Imaging of the Brain ([www.fmrib.ox.ac.uk/fsl/](http://www.fmrib.ox.ac.uk/fsl/)). The anatomic network of the cortex and subcortex is measured using Diffusion Tensor Imaging (DTI) and subsequently divided into 90 predefined regions according to the Automated Anatomical Labeling (AAL) Atlas (Tzourio-Mazoyer et al., 2002), see Table 1. Each node of the network corresponds to a brain region. Note that in contrast to the original AAL indexing, where sequential indices correspond to homologous brain regions, the indices in Figure 2A are rearranged such that  $k \in N_L = \{1, 2, \dots, 45\}$  corresponds to left and  $k \in N_R = \{46, \dots, 90\}$  to the right hemisphere. Thereby the hemispheric structure of the brain, i.e., stronger intra-hemispheric coupling compared to inter-hemispheric coupling, is highlighted (Figure 2A).

The structural connectivity matrices serve as a realistic input for modeling, rather than as exact information concerning the existence and strength of each connection in the human brain. The pipeline for constructing such connectivity information using diffusion tractography is known to face a range of challenges (Schilling et al., 2019). While some estimates of the strength and direction of structural connections from measurements of brain activity can in principle be attempted, the relation of these can vary dramatically with (experimentally unknown) parameters of the local dynamics and coupling function (Hlinka and Coombes, 2012).

The auditory cortex is the part of the temporal lobe that processes auditory information in humans. It is a part of the auditory system, performing basic and higher functions in hearing and is located bilaterally, roughly at the upper sides of the temporal lobes, i.e., corresponding to the AAL indexing  $k = 41, 86$  (temporal sup L/R). The auditory cortex takes part in the spectrotemporal analysis of the input passed on from the ear. Figure 2B displays the time-series of impulses which are supplied to the brain by means of the auditory cortex. These neural impulses were obtained by the method of Bader described in Section 2 (Bader, 2015, 2017, 2018). Here, in contrast to Figure 1, a real piece of music was used, namely the hip hop music song *One Mic*, composed by the American rapper *Nas* and released in 2002. During the transition from acoustic mechanical to electrical excitation within the cochlea, synchronization appears to improve perception of pitch, speech, or localization. The sampling rate of these impulses obtained by Bader's method is  $f_s = 192$  kHz.

Each node corresponding to a brain region is modeled by the FitzHugh-Nagumo (FHN) model with external stimulus, a paradigmatic model for neural spiking (FitzHugh, 1961; Nagumo et al., 1962; Bassett et al., 2018). Note that while the FitzHugh-Nagumo model is a simplified model of a single neuron, it is also often used as a generic model for excitable media on a coarse-grained level (Chernihovskiy et al., 2005; Chernihovskiy and Lehnertz, 2007). Thus the dynamics of the network reads:



**FIGURE 2**

(A) Model for the hemispheric brain structure: Weighted adjacency matrix  $A_{kj}$  of the averaged empirical structural brain network derived from twenty healthy human subjects by averaging over the coupling between two brain regions  $k$  and  $j$ . The brain regions  $k, j$  are taken from the Automated Anatomic Labeling Atlas (Tzourio-Mazoyer et al., 2002), but re-labeled such that  $k = 1, \dots, 45$  and  $k = 46, \dots, 90$  correspond to the left and right hemisphere, respectively. After (Gerster et al., 2020). (B) Time-series of the neural input signal  $I(t)$  obtained from the music song *One Mic* transformed by a method developed by Bader (Bader, 2020). The song has a length of about 270 s and was released in 2002 by American rapper Nas.

$$\begin{aligned} \epsilon \dot{u}_k &= u_k - \frac{u_k^3}{3} - v_k \\ &+ \sigma \sum_{j \in N_H} A_{kj} [B_{uu}(u_j - u_k) + B_{uv}(v_j - v_k)] \\ &+ \zeta \sum_{j \notin N_H} A_{kj} [B_{uu}(u_j - u_k) + B_{uv}(v_j - v_k)], \\ &+ C_k I(t) \end{aligned} \tag{3a}$$

$$\begin{aligned} \dot{v}_k &= u_k + a \\ &+ \sigma \sum_{j \in N_H} A_{kj} [B_{vu}(u_j - u_k) + B_{vv}(v_j - v_k)] \\ &+ \zeta \sum_{j \notin N_H} A_{kj} [B_{vu}(u_j - u_k) + B_{vv}(v_j - v_k)], \end{aligned} \tag{3b}$$

With  $k \in N_H$  where  $N_H$  denotes either the set of nodes  $k$  belonging to the left ( $N_L$ ) or the right ( $N_R$ ) hemisphere. Parameter  $\epsilon = 0.05$  describes the timescale separation between the fast activator variable (neuron membrane potential)  $u$  and the slow inhibitor (recovery variable)  $v$  (FitzHugh, 1961). Depending on the threshold parameter  $a$ , the FHN model may exhibit excitable behavior ( $|a| > 1$ ) or self-sustained oscillations ( $|a| < 1$ ). We use the FHN model in the oscillatory regime

and thus fix the threshold parameter at  $a = 0.5$  sufficiently far from the Hopf bifurcation point. The coupling within the hemispheres is given by the coupling strength  $\sigma$  while the coupling between the hemispheres is given by the inter-hemispheric coupling strength  $\zeta$ . As we are looking for partial synchronization patterns we fix  $\sigma = 0.7$  and  $\zeta = 0.15$  similar to numerical studies of synchronization phenomena during unihemispheric sleep (Ramlow et al., 2019) where partial synchronization patterns have been observed. The interaction scheme between nodes is characterized by a rotational coupling matrix:

$$\mathbf{B} = \begin{pmatrix} B_{uu} & B_{uv} \\ B_{vu} & B_{vv} \end{pmatrix} = \begin{pmatrix} \cos \phi & \sin \phi \\ -\sin \phi & \cos \phi \end{pmatrix}, \tag{4}$$

with coupling phase  $\phi = \frac{\pi}{2} - 0.1$ , causing primarily an activator-inhibitor cross-coupling. This particular scheme was shown to be crucial for the occurrence of partial synchronization patterns in ring topologies (Omelchenko et al., 2013) as it reduces the stability of the completely synchronized state. Also in the

**TABLE 1** Cortical and subcortical regions, according to the Automated Anatomical Labeling Atlas (AAL). Note that the numbering of the brain regions is different from the original numbering (Tzourio-Mazoyer et al., 2002).

Label L/R	Region	Lobe
1/46	Precentral	Central region
2/47	Frontal Sup	Frontal lobe
3/48	Frontal Sup Orb	Frontal lobe
4/49	Frontal Mid	Frontal lobe
5/50	Frontal Mid Orb	Frontal lobe
6/51	Frontal Inf Oper	Frontal lobe
7/52	Frontal Inf Tri	Frontal lobe
8/53	Frontal Inf Orb	Frontal lobe
9/54	Rolandic Oper	Central Region
10/55	Supp Motor Area	Frontal lobe
11/56	Olfactory	Frontal lobe
12/57	Frontal Sup Medial	Frontal lobe
13/58	Frontal Med Orb	Frontal lobe
14/59	Rectus	Frontal lobe
15/60	Insula	Insula
16/61	Cingulum Ant	Limbic lobe
17/62	Cingulum Mid	Limbic lobe
18/63	Cingulum Post	Limbic lobe
19/64	Hippocampus	Limbic lobe
20/65	ParaHippocampal	Limbic lobe
21/66	Amygdala	Sub cort. gray nuc
22/67	Calcarine	Occipital lobe
23/68	Cuneus	Occipital lobe
24/69	Lingual	Occipital lobe
25/70	Occipital Sup	Occipital lobe
26/71	Occipital Mid	Occipital lobe
27/72	Occipital Inf	Occipital lobe
28/73	Fusiform	Occipital lobe
29/74	Postcentral	Central region
30/75	Parietal Sup	Parietal lobe
31/76	Parietal Inf	Parietal lobe
32/77	Supramarginal	Parietal lobe
33/78	Angular	Parietal lobe
34/79	Precuneus	Parietal lobe
35/80	Paracentral Lobule	Frontal lobe
36/81	Caudate	Sub cort. gray nuc
37/82	Putamen	Sub cort. gray nuc
38/83	Pallidum	Sub cort. gray nuc
39/84	Thalamus	Sub cort. gray nuc
40/85	Heschl	Temporal lobe
41/86	Temporal Sup	Temporal lobe
42/87	Temporal Pole Sup	Limbic lobe
43/88	Temporal Mid	Temporal lobe
44/89	Temporal Pole Mid	Limbic lobe
45/90	Temporal Inf	Temporal lobe

modeling of epileptic-seizure-related synchronization phenomena (Gerster et al., 2020), where a part of the brain synchronizes, it turned out that such a cross-coupling is important. The subtle interplay of excitatory and inhibitory interaction is typical of the critical state at the edge of different dynamical regimes in which the brain operates (Massobrio et al., 2015; Shi et al., 2022), and gives rise to partial synchronization patterns which are not found otherwise.

The external stimulus  $I(t)$  describes the impulses evoked by the music piece *One Mic* by *Nas* and is applied to the brain areas  $k = 41, 86$  associated with the auditory cortex, i.e.,  $C_k = 1$  if  $k = 41$  or  $86$  and zero otherwise. Since  $I(t)$  is a time series which is calculated from a real piece of music, see Section 2, it has a physical dimension in seconds. On the other hand, the FitzHugh-Nagumo model has no explicit time scale. Its intrinsic angular frequency is dimensionless and given by  $\omega_k = \omega_{\text{FHN}} = 2\pi f_{\text{FHN}} \approx 2.51$  (corresponding to dimensionless frequency  $f_{\text{FHN}} \approx 0.4$ ). In order to compare our simulations with real data and include the time signal  $I(t)$  correctly in our dimensionless model, we must transform the dimensionless time units of the FHN oscillator model to real time units by comparing the FHN oscillation period of a single FHN oscillator  $T \approx 2.5$  to the characteristic frequencies  $n_b$  in Hz of an empirical time series. Depending upon the frequency band  $n_b$  (in Hz) chosen, the simulation time is converted to real time by  $1 \text{ s} = 2.5n_b$  simulation time units, or the simulated frequency (in Hz) is

$$f_b = n_b / f_{\text{FHN}}. \quad (5)$$

In this way, the parameter  $n_b$  effectively removes the time scale from the input, but on the other hand, it can also be seen as creating a link between our dimensionless model and the input signal  $I(t)$ .

## 4 Synchrony measures

We explore the dynamical behavior by calculating the mean phase velocity  $\omega_k = 2\pi M_k / \Delta T$  for each node  $k$ , where  $\Delta T$  denotes the time interval during which  $M_k$  complete rotations are realized. Throughout the paper, we denote the length of the input signal  $I(t)$  as  $\Delta T$ . For the numerical integration an adaptive Runge–Kutta integration method has been applied (python scipy: solve\_ivp, RK45). For all simulations we use initial conditions randomly distributed on the circle  $u_k^2 + v_k^2 = 4$  and a transient time of  $t_{\text{trans}} = 10,000$  before the input signal  $I(t)$  is supplied to the system. In case of an uncoupled system ( $\sigma = \zeta = 0$ ), the mean phase velocity (or natural frequency) of each node is  $\omega_k = \omega_{\text{FHN}} = 2\pi f_{\text{FHN}} \approx 2.51$ .

First, we introduce the spatially averaged mean phase velocity:

$$\bar{\omega} = \frac{1}{90} \sum_{k=1}^N \omega_k. \quad (6)$$

Thus  $\bar{\omega}$  corresponds to the mean phase velocity averaged over the left and right hemisphere.

Second, we take advantage of an abstract dynamical phase  $\theta_k$  that can be obtained from the standard geometric phase  $\tilde{\phi}_k(t) = \arctan(v_k/u_k)$  by a transformation which yields constant phase velocity  $\dot{\theta}_k$ . For an uncoupled FHN oscillator the function  $t(\tilde{\phi}_k)$  is calculated numerically, assigning a value of time  $0 < t(\tilde{\phi}_k) < T$  for every value of the geometric phase, where  $T$  is the oscillation period. The dynamical phase is then defined as  $\theta_k = 2\pi t(\tilde{\phi}_k)/T$ , which yields  $\dot{\theta}_k = \text{const}$ . Thereby identical, uncoupled oscillators have a constant phase relation with respect to the dynamical phase. By means of the dynamical phase  $\theta_k$  we can calculate the Kuramoto order parameter

$$R(t) = \frac{1}{90} \left| \sum_{k=1}^N \exp[i\theta_k(t)] \right|, \quad (7)$$

where the fluctuations of the order parameter  $R$  caused by the FHN model's slow-fast time scales are suppressed and a change in  $R$  indeed reflects a change in the degree of synchronization. The Kuramoto order parameter may vary between 0 and 1, where  $R = 1$  corresponds to complete phase synchronization, and small values characterize spatially desynchronized states.

Third, we introduce a new measure which specifies the coherence between the Kuramoto order parameter and the input signal by using the time average of the Kuramoto order parameter weighted with the input signal

$$\gamma = \frac{1}{\Delta T} \int_0^{\Delta T} R(t)I(t) dt \quad (8)$$

to quantify the overlap of coherent episodes ( $R$  large) with large input signals, averaged over time. The coherence  $\gamma$  is maximum if the synchronization is large whenever the signal is large. It is small if the overall synchronization is low, or if the modulation of the synchronization in time is not in phase with the modulation of the input signal amplitude. For  $\gamma = 0$  the Kuramoto order parameter and the input signal do not overlap at any time point. An increased value of  $\gamma \in [0, 1]$  means increased overlap between the Kuramoto order parameter and the input signal. The motivation for introducing the measure  $\gamma$  lies in the fact that in the human brain the increase and decrease of synchronization follows the large-scale form of the listened music in a coherent way (Hartmann and Bader, 2014, 2020).

Fourth, we make use of the Pearson correlation coefficient  $r$ , a linear cross-correlation, for simplicity taken without time delays. This is widely used as a non-directed measure of the strength of the correlation between two variables or sequences  $\{x_1, x_2, \dots, x_n\}$  and  $\{y_1, y_2, \dots, y_n\}$  (Glantz, 2002; Bastos and Schoffelen, 2015; Guevara Erra et al., 2017):

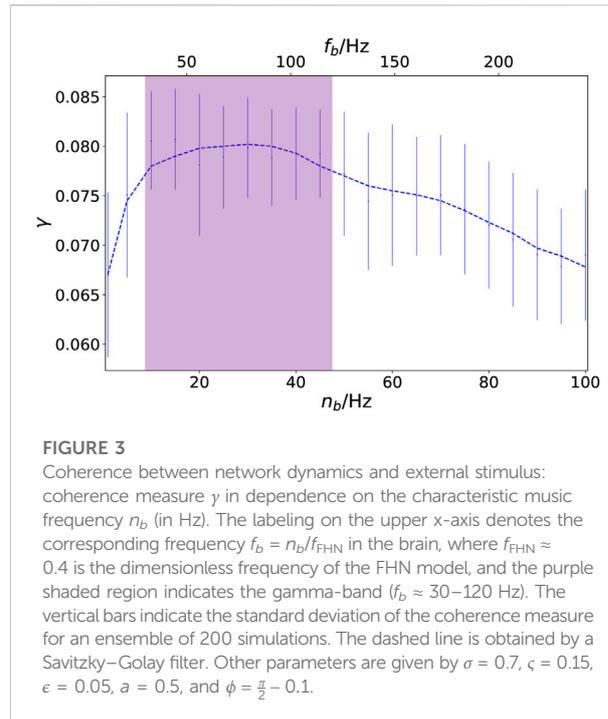


FIGURE 3

Coherence between network dynamics and external stimulus: coherence measure  $\gamma$  in dependence on the characteristic music frequency  $n_b$  (in Hz). The labeling on the upper x-axis denotes the corresponding frequency  $f_b = n_b/f_{\text{FHN}}$  in the brain, where  $f_{\text{FHN}} \approx 0.4$  is the dimensionless frequency of the FHN model, and the purple shaded region indicates the gamma-band ( $f_b \approx 30$ – $120$  Hz). The vertical bars indicate the standard deviation of the coherence measure for an ensemble of 200 simulations. The dashed line is obtained by a Savitzky–Golay filter. Other parameters are given by  $\sigma = 0.7$ ,  $\zeta = 0.15$ ,  $\epsilon = 0.05$ ,  $a = 0.5$ , and  $\phi = \frac{\pi}{2} - 0.1$ .

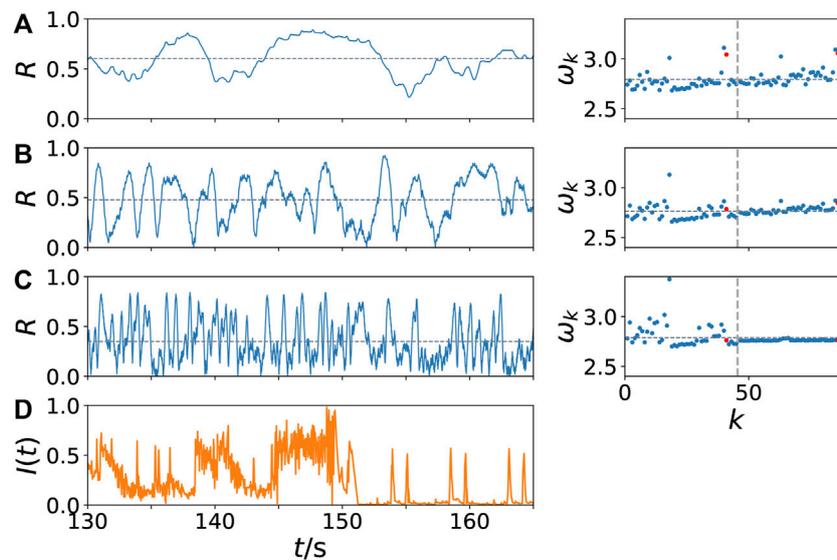
$$r = r_{x,y} = \frac{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (9)$$

where  $\bar{x}, \bar{y}$  denotes the mean of  $x, y$ , respectively. In recent decades, various methods for measuring synchronization have been introduced (Blinowska, 2011; Bastos and Schoffelen, 2015). The advantage of the Pearson correlation coefficient  $r$  is that it allows for easy and efficient calculation of the linear correlations between two variables or time series, and the results are very similar to those obtained by other common methods such as the phase-locking value (Lachaux et al., 1999). For a comparison of the different synchronization measures see (Jalili et al., 2014).

The input signal  $I(t)$  is obtained from the original music song *One Mic* by the cochlea model described in Section 2 (see Figure 1). The song has a length of about 4.5 min and the sampling rate of the obtained input signal is given by  $f_s = 192$  kHz. Sampling is the reduction of a continuous-time signal to a discrete-time signal, e.g., the conversion of a sound wave (a continuous signal) to a sequence of samples (a discrete-time signal). The sampling rate  $f_s$  is then the average number of samples obtained in one second. According to the Nyquist criterion, the frequency information of  $I(t)$  is then band-limited to  $f_b < \frac{1}{2}f_s$ .

## 5 Frequency bands and coherence

Next, we investigate dynamical scenarios emerging from an external stimulus in the auditory cortices of both hemispheres ( $k =$



**FIGURE 4**

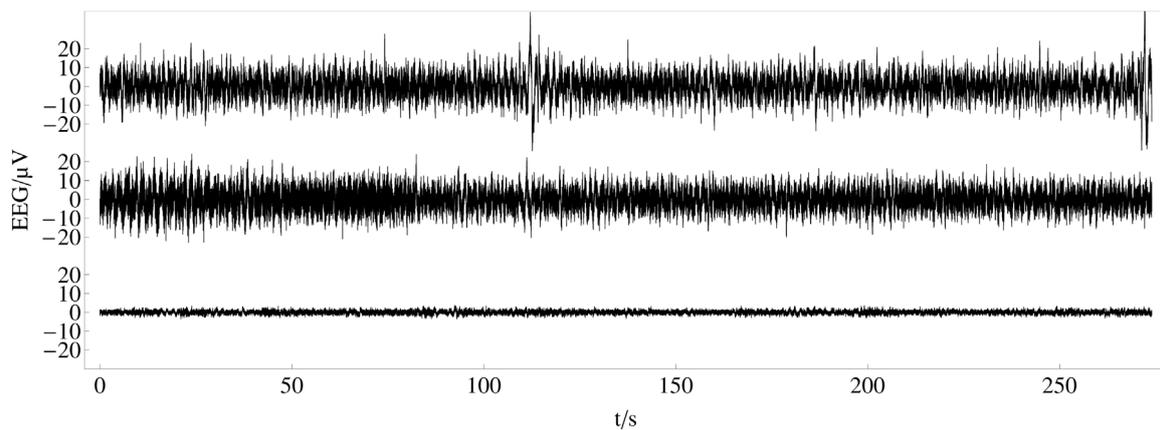
Dynamical scenarios: network dynamics for low and high values of coherence  $\gamma$ . Kuramoto order parameter  $R$  versus time in s (left column) and dimensionless mean phase velocity profile  $\omega_k = 2\pi f_k$  versus  $k$  (right column) for increasing values of the frequency  $n_b$  of the external stimulus  $I(t)$  (A)  $n_b = 5$  Hz (B)  $n_b = 30$  Hz and (C)  $n_b = 90$  Hz. In panel (D) the corresponding external stimulus  $I(t)$  is plotted, which is a blowup of a part of Figure 2B. The vertical dashed line in the right column separates the left and right brain hemisphere; the red dots mark the nodes of the auditory regions ( $k = 41, 86$ ). The horizontal grey dotted line indicates the temporal average of the Kuramoto order parameter  $R$  in the left column, and the spatial average of the mean-field frequency  $\bar{\omega}$  in the right column. Other parameters are as in Figure 3.

41, 86). In order to compare our simulations with the empirical analysis of the influence of music upon the brain (Hartmann and Bader, 2014, Hartmann and Bader, 2020, see also Section 6), we may choose different frequency bands  $n_b$ , and hence a different scaling of the time in the external stimulus. This can be visualized by plotting the coherence measure  $\gamma$  in dependence on the characteristic frequency  $n_b$  (in Hz), see Figure 3. We find a strong non-monotonic behavior of  $\gamma(n_b)$  and it turns out that by taking the frequency band  $n_b$  of the external stimulus as a control parameter, one can change the level of coherence between the system dynamics and the external stimulus. Although the standard deviation of the coherence measure is relatively large for an ensemble size of 200 simulations (indicated by the vertical bars), we find a pronounced maximum of the coherence  $\gamma$  for  $n_b = 12\text{--}48$  Hz corresponding to the gamma-band of brain waves ( $f_b \approx 30\text{--}120$  Hz) shown in Figure 3 by purple shading. This means that for that frequency  $n_b$  the level of synchronization follows the external signal most closely. It is in agreement with what has been observed in empirical brain analysis of the perception of music (Hartmann and Bader, 2014, 2020).

Figures 4A–C depicts the details of the change of the time series of the Kuramoto order parameter  $R(t)$  with increasing values of the frequency band  $n_b$  of the external stimulus  $I(t)$ , which is shown in Figure 4D. It represents a part of the neural input signal  $I(t)$  constructed from the music song *One Mic* and shown in Figure 2B. We take a closer look at the temporal evolution of  $R$  and the mean phase velocities  $\omega_k$  in the system for different values of

$n_b$  chosen from three different regimes in Figure 3: With increasing value of  $n_b$  in panels (A)–(C), the time scale of the simulated neural output in Hz changes from lower to higher frequencies  $f_b$  which is also seen in the temporal fluctuations of  $R(t)$ . Furthermore we observe on the one hand an increasing amplitude of the temporal fluctuations of  $R$ . On the other hand, the temporal average of the Kuramoto order parameter  $R$  decreases with increasing  $n_b$ , marked by a horizontal grey dotted line in the left column: While for a small value of  $n_b = 5$  Hz in Figure 4A the Kuramoto order parameter  $R$  assumes rather large values, and small values  $R < 0.2$  are not reached, for high values of  $n_b = 90$  Hz in Figure 4C rather small values of  $R$  are measured. This trend can be seen by means of the temporal average of the Kuramoto order parameter  $R$ . For  $n_b = 30$  Hz in Figure 4B, the temporal average of  $R$  takes a value  $\approx 0.5$  and the time evolution shows regular oscillations between low ( $R < 0.2$ ) and high values ( $R > 0.8$ ). This aspect will be further discussed in the next section, since it can also be observed in experiments.

As shown in Figure 3, in the case of  $n_b = 30$  Hz the coherence  $\gamma$  is maximum. Even though a higher value of the temporal average of  $R(t)$ , as observed in Figure 4A for  $n_b = 5$ , might imply a higher value of  $\gamma$  according to Eq. (8), Figure 4B shows that it is more important that  $R(t)$  and  $I(t)$  show a similar temporal modulation, as in Figure 4B for  $n_b = 30$ . Despite the averaging over 250 simulations over the whole simulation time in Figure 3, the time segment in Figure 4B shows such a similarity in the modulation: We can see simultaneous drops of  $R(t) < 0.1$  and  $I(t) < 0.1$  for example at  $t \approx 138, 140, 150$ , whereas the values in between are higher, even if they fluctuate.



**FIGURE 5**

Recorded and averaged electroencephalogram (EEG) data: top and middle plot show recorded EEG time series after pre-processing for one electrode (Fp1) from two different participants. The bottom plot shows the time series of the same electrode averaged over 25 subjects and three trials.

In the right column of [Figure 4](#) the dimensionless mean phase velocities  $\omega_k$  of all nodes are plotted, the horizontal grey dotted line indicates the spatial average, i.e., the collective mean-field frequency  $\bar{\omega}$ , which does not change for different  $n_b$  since it is determined by the intrinsic collective dynamics. In contrast, the node dynamics of the auditory regions ( $k = 41, 86$ ), indicated by red dots, depends on  $n_b$  since it receives the external input signal which has a higher frequency in dimensionless units if the time is scaled in larger units  $1/n_b$ . For  $n_b = 5$  Hz in [Figure 4A](#), the mean phase velocity of the auditory cortex is higher compared to the spatial average of the collective mean-field frequency  $\bar{\omega}$ . For  $n_b = 30$  Hz in [Figure 4B](#), the mean phase velocity of the auditory cortex approaches  $\bar{\omega}$  having a bigger impact on the dynamics of the whole system than in [Figure 4A](#) for  $n_b = 5$  Hz.

Remarkable is the fact of a dynamical asymmetry shown by the mean phase velocities in [Figure 4C](#): While the nodes of the right hemisphere exhibit equal mean phase velocity, i.e., they are frequency synchronized, the left hemisphere remains desynchronized and exhibits on average faster dynamics. This may indicate that regardless of the input  $I(t)$  the system can exhibit partial synchronization. Such behavior is similar to the dynamics of unihemispheric sleep studied in ([Ramlow et al., 2019](#)), where no external input has been applied to the dynamical system. In such states one hemisphere is synchronized, whereas the other hemisphere is partially desynchronized.

## 6 Comparison with experiments

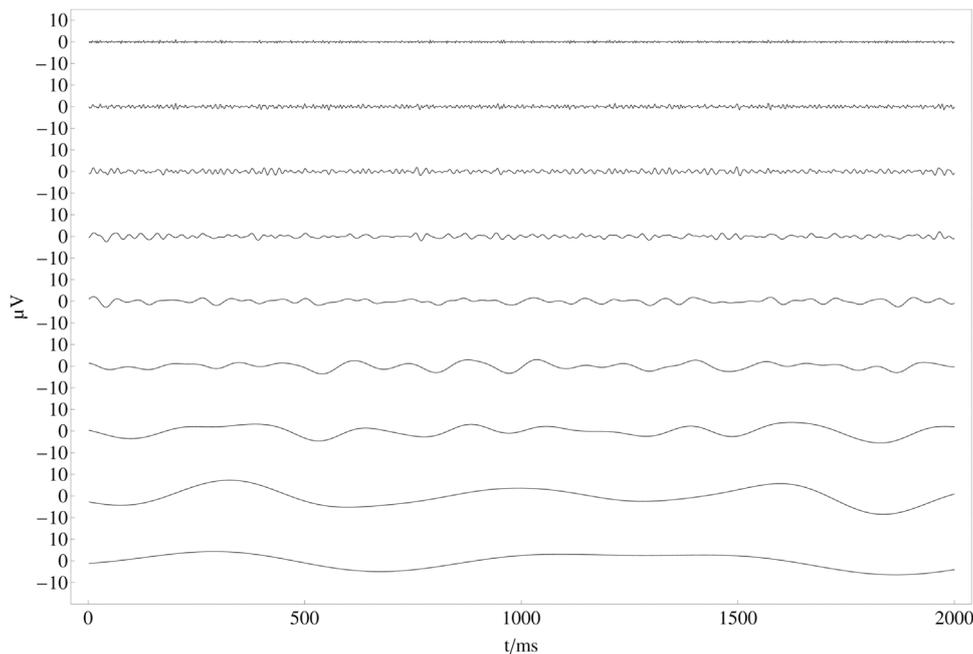
Based on the correlations between the processes associated with the perception of musical form and neural synchronization, we expect the dynamics of neural synchronization to correspond to the amplitude dynamics of the stimulus. Again, the musical

amplitude corresponds to perceived loudness, and is calculated as integration of energy over time intervals. Then synchronization between different brain regions is high when the amplitude of the musical piece is high, and synchronization is low when the amplitude of the piece is low. We expect such brain synchronization to be strong due to the prominence of the gamma-band in perception of musical parameters.

In an experiment, we have recorded the electroencephalogram (EEG) from human scalps to examine the perception of music large-scale form (see [Figure 5](#))<sup>1</sup>. 25 musically skilled subjects listened to the song *One Mic* from the artist *Nas* three times each. The song was released in 2001 on his Album *Stillmatic* on *Columbia Records*. The electroencephalogram (EEG) signals were recorded with a sample rate of 500 Hz from 32 electrodes, positioned following the 10–20 method of placement ([Jasper, 1958](#)). In this experiment, we are focused on the temporal dynamics of synchronization related to the time span of the musical form and therefore do not take advantage of methods for the inverse modeling of EEG data ([Schiffelen and Gross, 2009](#); [Palva and Palva, 2012](#)).

After artifact correction, recorded data for each channel has been averaged over subjects and trials to obtain a grand average of 75 trials for each channel to increase the signal to noise ratio and enhance event-related potentials. This type

<sup>1</sup> We have taken into account the usual guidelines regarding ethical procedure (informed consent). The subjects were mainly found and recruited through the Institute of Systematic Musicology Hamburg and had instrumental lessons on at least one instrument (mean duration 10.0 years, standard deviation 4.6 years) or corresponding experience as DJ. They participated in accordance with local ethics committee guidelines.



**FIGURE 6**

Nine frequency bands (FB) after wavelet transformation: Result of the continuous wavelet transform for the first 2 seconds of the averaged time series in Figure 5. From top to bottom frequency bands correspond to FB 1: 125 – 250 Hz, FB 2: 62.5 – 125 Hz, FB 3: 31.25 – 62.5 Hz, FB 4: 15.63 – 31.25 Hz, FB 5: 7.81 – 15.63 Hz, FB 6: 3.91 – 7.81 Hz, FB 7: 1.95 – 3.91 Hz, FB 8: 0.98 – 1.95 Hz, FB 9: 0.49–0.98 Hz.

of averaging reveals *evoked potentials* (in contrast to *induced potentials*) and is related to the presented stimulus in a classical event-related potential manner (Tallon-Baudry et al., 1996; Tallon-Baudry and Bertrand, 1999; Zanto et al., 2005). We are aware that our choice for evoked potentials pushes subjective, individual brain activity that is not stimulus-locked into the background. Indeed, it was found that this subjective, individual brain activity, often referred to as ‘noise’, contains valuable information that is lost when averaging over many subjects (Tallon-Baudry and Bertrand, 1999). On the other side, recent studies on this issue have shown strong overlap between subjects’ brain activity (Hasson et al., 2004; Dmochowski et al., 2012; Abrams et al., 2013; Kaneshiro et al., 2021). Therefore, we choose to take advantage of the improvement in the signal-to-noise ratio over the disadvantage of the individual portion of the perception. Individual perception might be subject to future studies. Also note that the choice of using a correlation analysis between single electrodes is not including redundant synchrony due to overlap of electrical fields between electrodes, since the positions of the electrodes do not differ over measurement time. Therefore, the differences in correlation strength between different electrodes cannot be explained by spurious synchrony (Holsheimer and Feenstra, 1977; Kayser and Tenke, 2006; Bhavsar et al., 2018). For a more detailed description of the experimental procedure, technical details and pre-processing, see (Hartmann and Bader, 2020).

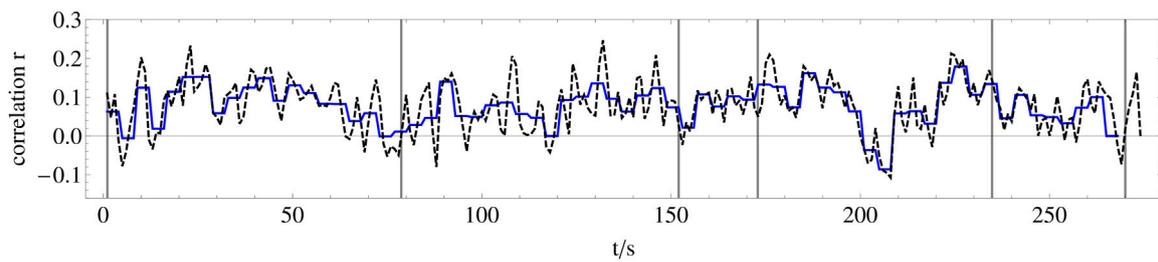
In Figure 6, all channels have been decomposed into nine independent frequency bands that correspond approximately to the frequency bands mentioned above by using a continuous wavelet transform with a Mexican Hat wavelet (Freeman and Quian Quiroga, 2013). In contrast to a bandpass filter with a subsequent Hilbert transform, using a Mexican hat wavelet for filtering is fast and efficient since one can decompose the recorded EEG data into the desired frequency bands in one step by defining the number of octaves. The continuous wavelet transform of a uniformly sampled sequence  $\{x_1, x_2, \dots, x_n\} = \{x(t_0), x(t_0 + \Delta t), \dots, x(t_0 + (n-1)\Delta t)\}$  is given by

$$w(u, s) = \frac{1}{\sqrt{s}} \sum_{k=1}^n x_k \psi\left(\frac{(k-u)\Delta t}{s}\right), \quad (10)$$

where  $s \in \mathbb{R}$  corresponds to the frequency of the EEG band and  $u = 1, \dots, n$  labels the wavelet coefficients with the number  $n$  of analyzed sample points defining the time window of observation. As wavelet function  $\psi$  a Mexican Hat wavelet is used, given by

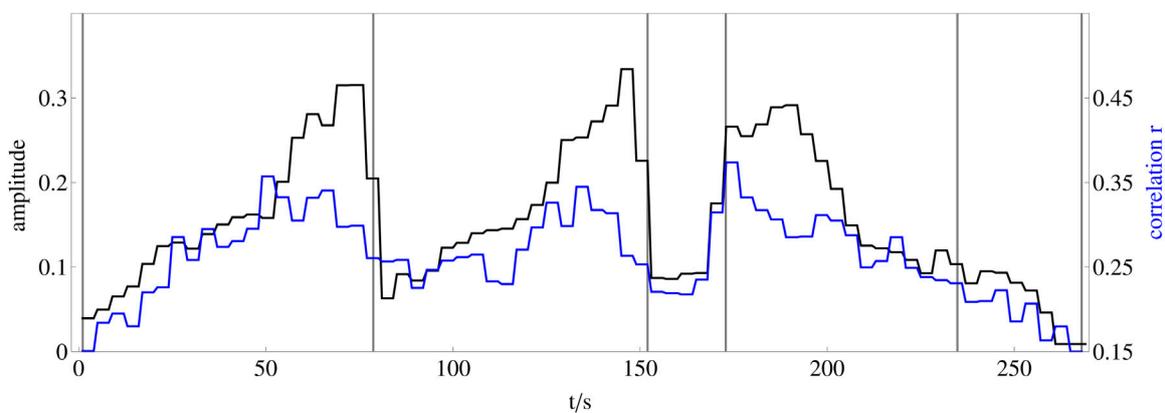
$$\psi(x) = \frac{-2}{\sqrt{\pi}\sqrt{3}\sigma} \left(\frac{x^2}{\sigma^2} - 1\right) \exp\left(-\frac{x^2}{2\sigma^2}\right), \quad (11)$$

where  $\sigma$  is the width of the wavelet. The EEG bands used align very well with a musical scale, where each higher band doubles the frequency of its respective lower band, corresponding to a musical octave. Please note that this relation might only be at chance, still it



**FIGURE 7**

Example of the synchronization dynamics between two electrodes. Dashed black line: Time series of the Pearson correlation coefficient  $r$  calculated for successive 1-s time windows ( $n = 500$  in Eq. 9 between averaged EEG recordings of electrode Fp1 (lower plot in Figure 5) and electrode T7. Blue line: Pearson correlation coefficient averaged over four consecutive 1-s time windows of the dashed black line.



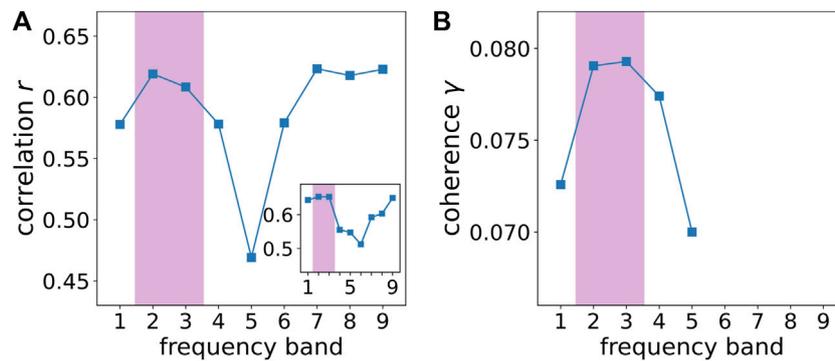
**FIGURE 8**

Comparison of whole brain synchronization dynamics and representation of the musical form of the stimulus. The black line shows the amplitude dynamics of the stimulus as a representation of the musical form, averaged over each of four consecutive seconds. The blue line shows the average of the 25 correlation time series between two electrodes from each frequency band that correlates most strongly with the amplitude dynamics of the stimulus.

may also relate to the fact that all human senses relate physics to perception in a logarithmic way (Schneider, 2018). It is therefore convenient to scale  $s$  in the wavelet transform in the same mathematical way as an equal-tempered musical scale like  $s_{\text{oct}} = \alpha 2^{\text{oct}-1}$ , where  $\text{oct} \in \{1, 2, \dots, 9\}$  is the octave number related to the nine frequency bands shown in Figure 6 and  $\alpha$  is the smallest wavelet scale.

For each electrode pair of these nine data sets filtered in this way, the synchronization is calculated by means of the Pearson correlation coefficient  $r$  (see Eq. 9) in the next step. Thus, we can analyze the synchronization dynamics as a function of the frequency bands. Since we aim to reveal synchronization dynamics on the level of musical form, we calculate the correlation within successive 1-s time windows for each possible pair of electrodes of each wavelet-filtered dataset, which results in  $32 \cdot 31 / 2 \cdot 9 = 4,464$  time series of correlation coefficients representing the synchronization dynamics between electrode-pairs with a resolution of 1 s, and each of these time series has a length of 270 s corresponding to the stimulus length (see Figure 7).

In order to relate this huge number of time series of correlation coefficients to the amplitude dynamics of the stimulus, we first average the amplitude of the stimulus and the correlation coefficients calculated for the 496 electrode pairs and nine frequency bands within successive 4-s windows to avoid minor amplitude fluctuations and obtain a scaling corresponding to about two musical bars that fits to changes related to the musical form (Figure 7). In the second step, we correlate all 4,464 time series of correlation coefficients with the amplitude dynamics of the stimulus. In the third step, we select the 25 time series of correlation coefficients per frequency band that correlate most strongly with the amplitude dynamics of the stimulus, shown in Figure 8. Now, we average these 25 time series of correlation coefficients per frequency band, which results in a single time series of 270 s length for each frequency band, respectively. These averaged time series of correlation coefficients, representing the synchronization dynamics for each frequency band, are correlated over the whole recorded time with the amplitude dynamics of the



**FIGURE 9**

Comparison between experimental and numerical results **(A)** Experimentally recorded correlation  $r$  of the individual averages of the amplitude dynamics for each frequency band most strongly correlated with the stimulus as a function of frequency band (FB 1: 125 – 250 Hz, FB 2: 62.5 – 125 Hz, FB 3: 31.25 – 62.5 Hz, FB 4: 15.63 – 31.25 Hz, FB 5: 7.81 – 15.63 Hz, FB 6: 3.91 – 7.81 Hz, FB 7: 1.95 – 3.91 Hz, FB 8: 0.98 – 1.95 Hz, FB 9: 0.49 – 0.98 Hz). The inset depicts the Pearson correlation coefficient  $r$  as a function of frequency band where instead of the amplitude the fractal dimension (Grassberger and Procaccia, 1983a,b) has been used for the calculation of  $r$ . **(B)** Numerically simulated coherence  $\gamma$  between network dynamics and external stimulus, where the corresponding frequency bands are averaged from Figure 3. As in Figure 3, the purple shaded regions in both panels indicate the gamma-band ( $f_b \approx 30$ –120 Hz), respectively.

stimulus (see Figure 9A). It can be shown that the low and the high gamma-band (frequency bands 2–3) correlate strongly with the stimulus as expected, but also the slow oscillations (frequency bands 7–9) correlate very well (see discussion below). By this, we can reveal how good the synchronization dynamics in each frequency band corresponds to the amplitude dynamics of the stimulus on the level of musical form. In the next step, we average these time series representing the synchronization dynamics for each frequency band and correlate the resulting time series, representing the synchronization dynamics of the whole brain, with the amplitude dynamics of the stimulus as well. These two time series correlate with a Pearson coefficient of  $r = 0.76$ . Therefore, we can conclude that the higher the amplitude of the stimulus, the higher the synchronization between the most correlated time series of the different frequency bands. According to (Cohen, 1992), this is a strong effect.

As shown in Figure 8, the increased synchrony is not constant during music listening, but rather synchronization dynamics follows the sound amplitude. Note that the correlation between sound amplitude (perceived loudness) or other parameters like brightness or fractal correlation dimension (see inset of Figure 9A) and brain synchronization is not trivial. First, brain synchronization appears at frequencies much lower than most musical frequencies. Secondly, synchronization appears with multiple perceptual parameters. Thirdly, increasing, e.g., the sound amplitude might lead to an increase of the network amplitude, but here it leads to an enhanced synchronization, pointing to a highly nonlinear process in the network, caused by the activity of the brain when perceiving sound.

It is interesting to note that the correlation with the stimulus is highest when the time series from all frequency bands are averaged. The correlation coefficient of the averages of the 25 most correlated

time-series as a function of the individual frequency bands is shown in Figure 9A. It shows two regimes of high correlation, separated by a frequency band (FB 5) with low correlation. Here, the central nervous system in the spinal cord and its relation to the locomotor system are expected to be responsible for the dynamics in the frequency bands 6–9 due to their frequency range close to walking and dancing (van Noorden and Moelants, 1999). Note that the electroencephalogram (EEG) recordings are performed on the skull, and therefore represent the brain dynamics of the neocortex which is interacting with the brain stem. Therefore, the high correlations between synchronization and musical form in frequency bands 6–9 can be interpreted as caused by the interaction of the neocortex with subcortical brain regions. Likewise, the high correlations in frequency bands 2–3 are interpreted as activity of the neocortex solely, as expected. The results therefore also suggest a separation of musical form-related synchronization between cortical (frequency bands 2–3) and subcortical (frequency bands 6–9) regions.

The high correlations observed in frequency bands 2–3 for the sound amplitude (see Figure 9A) as well as for the fractal correlation dimension (see inset of Figure 9A) correspond to a frequency range of 31.25–125 Hz (gamma-band). On the other hand in Figure 3, the strongest coherence between the Kuramoto order parameter (measure for global neural synchronization) and the external input can be found for  $n_b = 10$ –40 Hz. Taking into account that the natural frequency of each node is  $f_{FHN} \approx 0.4$ , we can calculate the corresponding frequency band  $f_b = n_b/f_{FHN}$ . As shown by the upper x-axis in Figure 3, the strongest coherence in our model can be observed for a frequency band of  $f_b = 40$ –100 Hz, which agrees with the gamma-band in the brain. For comparison with the experiment, we show the corresponding numerically simulated results in Figure 9B, where the respective frequency bands are averaged from Figure 3. Both experimental and numerical results show a

pronounced maximum of correlation between stimulus and brain dynamics for the gamma-band (frequency bands 2–3) in Figure 9. Note that the second maximum in the experimental data (panel A), which is due to the interaction of the neocortex with subcortical brain regions as discussed above, is absent in the simulated data (panel B) since the computer simulation is only performed for the neocortex, using a cochlea input, but neglecting brain stem activity.

## 7 Conclusion

We have investigated the influence of music in a simulated network of FitzHugh-Nagumo oscillators with empirical structural connectivity obtained from healthy human subjects, and have compared it to measured electroencephalogram (EEG) data. We report an increase of coherence between the global dynamics and the input signal induced by a specific music song. We have shown that the level of coherence depends on the frequency band. We have compared our results with experimental data, which describe global neural synchronization between different brain regions in the gamma-band range and its increase just before transitions between different parts of the musical form (musical high-level events). Such synchronization increases before musical large-scale form boundaries, and decreases afterwards, therefore represents musical large-scale form perception.

The transformation of sound into neural spikes takes place in the cochlea, a part of the human ear which is directly connected to the auditory cortex. By means of the basilar membrane, the brain is able to perceive different frequencies organized in so-called critical bands. We have applied a cochlea model to transform a specific music song into an input signal representing neural spikes evoked by the music song. This input signal has then been supplied to a simulated network of neural oscillators with empirical structural connectivity. By the transformation of the dimensionless time units of the oscillator model to real time units, we have investigated dynamical scenarios in dependence on the introduced frequency band parameter. To quantify moreover the overlap between input signal and network dynamics, we have introduced a coherence measure. It has turned out that this coherence measure depends sensitively on the frequency band and has its maximum in the gamma-band. Therefore, depending on the frequency band, coherence can be induced between the dynamics of the system and its input signal.

These results are in accordance with our own and previous experiments (Hartmann and Bader, 2014, 2020) where music has also been found to induce a certain degree of synchrony in the human brain. We have shown that listening to music can have a remarkable influence on the brain dynamics, in particular, a periodic alternation between synchronization and desynchronization which is strongly related to the music perceived. We have experimentally analyzed in detail the influence of real music on the neural activity with respect to the common frequency bands in the brain. By means of the Pearson correlation coefficient of the sound amplitude as well

as the fractal correlation dimension, we have found the gamma-band to be important for musical form perception. Just as in the computer simulation, we have found a pronounced maximum for this frequency range. Moreover as in simulation, the increased gamma-band synchrony is not constant during music listening in our experiment, but rather synchronization dynamics follows the musical large-scale form represented by a perceptual related characteristic of the stimulus, i.e., the amplitude and fractal correlation dimension. Even though we chose a specific piece of music in this study, we expect future work to show that these results can be generalized.

Furthermore, the results suggest a separation in musical form-related brain synchronization between high brain frequencies, associated with neocortical activity, and low frequencies in the range of dance movements, associated with interactivity between cortical and subcortical regions. Besides, an alternation between synchronization and desynchronization reflects the variability of the system; this can be seen as a critical state between a fully synchronized and a desynchronized state. It is known that the brain is operating in a critical state at the edge of different dynamical regimes (Massobrio et al., 2015; Shi et al., 2022), exhibiting hysteresis and avalanche phenomena as seen in critical phenomena and phase transitions (Ribeiro et al., 2010; Steyn-Ross and Steyn-Ross, 2010; Kim et al., 2018).

By choosing appropriate parameters and measures, we have reported an intriguing dynamical behavior in dependence on the frequency bands, and have observed the induced increase of coherence both in numerical and experimental setups. To sum up, music supplied to the brain allows for a high coherence and correlation between musical input and brain dynamics especially in the gamma-band. This insight may be used to fathom the general modalities of the influence of music on the human brain.

## Data availability statement

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

## Ethics statement

Ethical review and approval was not required for the study on human participants in accordance with the local legislation and institutional requirements. The patients/participants provided their written informed consent to participate in this study.

## Author contributions

JS did the numerical simulations and the theoretical analysis, LH has performed the experiments. RB and ES

supervised the study. All authors designed the study and contributed to the preparation of the manuscript. All the authors have read and approved the final manuscript.

## Funding

This work was supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation, project No. 429685422) and the Open Access Publication Fund of TU Berlin.

## Acknowledgments

We are grateful to Antonín Škoch and Jaroslav Hlinka for preparing the example structural connectivity matrices.

## References

- Abrams, D., Ryali, S., Chen, T., Chordia, P., Khouzam, A., Levitin, D., et al. (2013). Inter-subject synchronization of brain responses during natural music listening. *Eur. J. Neurosci.* 34, 1458–1469. doi:10.1111/ejn.12173
- Baars, B. J. (2006). Global workspace theory of consciousness: Toward a cognitive neuroscience of human experience. *Prog. Brain Res.* 150, 45–53. doi:10.1016/s0079-6123(05)50004-9
- Bader, R. (2018). Cochlear spike synchronization and neuron coincidence detection model. *Chaos* 28, 023105. doi:10.1063/1.5011450
- Bader, R. (2021). *How music works*. Cham: Springer. doi:10.1007/978-3-030-67155-6
- Bader, R. (2020). Neural coincidence detection strategies during perception of multi-pitch musical tones. <http://arXiv.org/abs/2001.06212v1>.
- Bader, R. (2013). *Nonlinearities and synchronization in musical acoustics and music psychology*. Berlin: Springer.
- Bader, R. (2015). Phase synchronization in the cochlea at transition from mechanical waves to electrical spikes. *Chaos* 25, 103124. doi:10.1063/1.4932513
- Bader, R. (2017). Pitch and timbre discrimination at wave-to-spike transition in the cochlea. <http://ArXiv.org/abs/1711.05596>.
- Bader, R., Zielke, A., and Franke, J. (2021). Timbre-based machine learning of clustering Chinese and Western Hip Hop music. doi:10.31235/osf.io/8ef7g
- Bassett, D. S., and Sporns, O. (2017). Network neuroscience. *Nat. Neurosci.* 20, 353. EP. Review Article. doi:10.1038/nn.4502
- Bassett, D. S., Zurn, P., and Gold, J. I. (2018). On the nature and use of models in network neuroscience. *Nat. Rev. Neurosci.* 19, 566–578. doi:10.1038/s41583-018-0038-8
- Bastos, A. M., and Schoffelen, J. M. (2015). A tutorial review of functional connectivity analysis methods and their interpretational pitfalls. *Front. Syst. Neurosci.* 9, 175. doi:10.3389/fnsys.2015.00175
- Bergner, A., Frasca, M., Sciuto, G., Buscarino, A., Ngamga, E. J., Fortuna, L., et al. (2012). Remote synchronization in star networks. *Phys. Rev. E Stat. Nonlin. Soft Matter Phys.* 85, 026208. doi:10.1103/physreve.85.026208
- Bhattacharya, J., Petsche, H., and Pereda, E. (2001). Long-range synchrony in the gamma band: Role in music perception. *J. Neurosci.* 21, 6329–6337. doi:10.1523/jneurosci.21-16-06329.2001
- Bhavsar, R., Sun, Y., Helian, N., Davey, N., Mayor, D., and Steffert, T. (2018). The correlation between eeg signals as measured in different positions on scalp varying with distance. *Procedia Comput. Sci.* 123, 92–97. doi:10.1016/j.procs.2018.01.015
- Blinowska, K. J. (2011). Review of the methods of determination of directed connectivity from multichannel data. *Med. Biol. Eng. Comput.* 49, 521–529. doi:10.1007/s11517-011-0739-x
- Bonetti, L., Brattico, E., Carlomagno, F., Donati, G., Cabral, J., Haumann, N. T., et al. (2021). Rapid encoding of musical tones discovered in whole-brain

## Conflict of interest

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

The handling editor KL declared a past collaboration with the authors JS and ES.

## Publisher's note

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

connectivity. *Neuroimage* 245, 118735. doi:10.1016/j.neuroimage.2021.118735

Buhusi, C., and Meck, W. (2009). Relativity theory and time perception: Single or multiple clocks? *PLoS one* 4, e6268. doi:10.1371/journal.pone.0006268

Buhusi, C., and Meck, W. (2005). What makes us tick? Functional and neural mechanisms of interval timing. *Nat. Rev. Neurosci.* 6, 755–765. doi:10.1038/nrn1764

Buzsáki, G. (2006). *Rhythms of the brain*. Oxford, United Kingdom: Oxford University Press.

Chernihovskiy, A., and Lehnertz, K. (2007). Measuring synchronization with nonlinear excitable media. *Int. J. Bifurc. Chaos* 17, 3425–3429. doi:10.1142/s0218127407019159

Chernihovskiy, A., Mormann, F., Müller, M., Elger, C. E., Baier, G., Lehnertz, K., et al. (2005). EEG analysis with nonlinear excitable media. *J. Clin. Neurophysiol.* 22, 314–329. doi:10.1097/01.wnp.0000179968.14838.e7

Chouzouris, T., Omelchenko, I., Zakharova, A., Hlinka, J., Jiruska, P., Schöll, E., et al. (2018). Chimera states in brain networks: Empirical neural vs. modular fractal connectivity. *Chaos* 28, 045112. doi:10.1063/1.5009812

Cohen, J. (1992). A power primer. *Psychol. Bull.* 112, 155–159. doi:10.1037//0033-2909.112.1.155

Curran Associates, Inc (2018). “8th annual international conference on biologically inspired cognitive architectures, BICA 2017,” in *Procedia Computer Science* Volume 123, Moscow, Russia, August 1–6, 2017.

de Boer, E. (1991). Auditory physics. physical principles in hearing theory. iii. *Phys. Rep.* 203, 125–231. doi:10.1016/0370-1573(91)90068-w

Dehaene, S., Changeux, J. P., and Naccache, L. (2011). “The global neuronal workspace model of conscious access: From neuronal architectures to clinical applications,” in *Characterizing consciousness: From cognition to the clinic?* Editors S. Dehaene and Y. Christen (Berlin, Heidelberg: Springer-Verlag Berlin Heidelberg), 55–84. Research and Perspectives in Neurosciences. doi:10.1007/978-3-642-18015-6\_4

Deliège, I., and Melen, M. (2014). “Cue abstraction in the representation musical form,” in *Perception and cognition of music*. Editors I. Deliège and J. Sloboda (Hove: Psychology Press), 387–412.

Deutsch, D. (2013). *The psychology of music. Academic Press series in cognition and perception*. 3rd ed. edn. Oxford: Academic.

Dmochowski, J. P., Sajda, P., Dias, J., and Parra, L. C. (2012). Correlated components of ongoing eeg point to emotionally laden attention - a possible marker of engagement? *Front. Hum. Neurosci.* 6, 112. doi:10.3389/fnhum.2012.00112

Drauschke, F., Sawicki, J., Berner, R., Omelchenko, I., and Schöll, E. (2020). Effect of topology upon relay synchronization in triplex neuronal networks. *Chaos* 30, 051104. doi:10.1063/5.0008341

- Engel, A. K., and Fries, P. (2016). Chap. 3-Neuronal oscillations, coherence, and consciousness *The neurology of consciousness* (Second Edition) (Cambridge, Massachusetts: Academic Press), 49–60. doi:10.1016/b978-0-12-800948-2.00003-0
- Engel, A. K., Fries, P., and Singer, W. (2001). Dynamic predictions: Oscillations and synchrony in top-down processing. *Nat. Rev. Neurosci.* 2, 704–716. doi:10.1038/35094565
- Engel, A. K., and Singer, W. (2001). Temporal binding and the neural correlates of sensory awareness. *Trends Cogn. Sci.* 5, 16–25. doi:10.1016/s1364-6613(00)01568-0
- FitzHugh, R. (1961). Impulses and physiological states in theoretical models of nerve membrane. *Biophys. J.* 1, 445–466. doi:10.1016/s0006-3495(61)86902-6
- Freeman, W. J., and Quiroga, R. (2013). *Imaging brain function with EEG: Advanced temporal and spatial analysis of electroencephalographic signals*. New York: Springer.
- Fries, P. (2009). Neuronal gamma-band synchronization as a fundamental process in cortical computation. *Annu. Rev. Neurosci.* 32, 209–224. doi:10.1146/annurev.neuro.051508.135603
- Friston, K. J., and Friston, D. A. (2013). *A free energy formulation of music generation and perception: Helmholtz revisited*. Heidelberg: Springer International Publishing, 43–69. doi:10.1007/978-3-319-00107-4\_2
- Gambuzza, L. V., Cardillo, A., Fiasconaro, A., Fortuna, L., Gómez-Gardeñes, J., Frasca, M., et al. (2013). Analysis of remote synchronization in complex networks. *Chaos* 23, 043103. doi:10.1063/1.4824312
- Gerster, M., Berner, R., Sawicki, J., Zakharova, A., Skoch, A., Hlinka, J., et al. (2020). FitzHugh-Nagumo oscillators on complex networks mimic epileptic-seizure-related synchronization phenomena. *Chaos* 30, 123130. doi:10.1063/5.0021420
- Glantz, S. A. (2002). *Primer of biostatistics*. 5 edn. New York: McGraw-Hill.
- Gollo, L. L., Mirasso, C. R., Atienza, M., Crespo-García, M., and Cantero, J. L. (2011). Theta band zero-lag long-range cortical synchronization via hippocampal dynamical relaying. *PLoS ONE* 6, e17756. doi:10.1371/journal.pone.0017756
- Grassberger, P., and Procaccia, I. (1983a). Characterization of strange attractors. *Phys. Rev. Lett.* 50, 346–349. doi:10.1103/physrevlett.50.346
- Grassberger, P., and Procaccia, I. (1983b). Measuring the strangeness of strange attractors. *Phys. D. Nonlinear Phenom.* 9, 189–208. doi:10.1016/0167-2789(83)90298-1
- Gray, C. M., and Singer, W. (1989). Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. *Proc. Natl. Acad. Sci. U. S. A.* 86, 1698–1702. doi:10.1073/pnas.86.5.1698
- Guevara Erra, R., Perez Velazquez, J. L., and Rosenblum, M. (2017). Neural synchronization from the perspective of non-linear dynamics. *Front. Comput. Neurosci.* 11, 98. doi:10.3389/fncom.2017.00098
- Guo, X. X., Xiang, S. Y., Qu, Y., Han, Y. N., Wen, A. J., Hao, Y., et al. (2021). Enhanced prediction performance of a neuromorphic reservoir computing system using a semiconductor nanolaser with double phase conjugate feedbacks. *J. Light. Technol.* 39, 129–135. doi:10.1109/jlt.2020.3023451
- Hartmann, L., and Bader, R. (2020). Neural synchronization of music large-scale form. <http://arXiv.org/abs/2005.06938v1>.
- Hartmann, L., and Bader, R. (2014). Neuronal synchronization of musical large scale form: An eeg-study. *Proc. Meet. Acoust. 168th Meet. Acoust. Soc. Am.* 22, 1.
- Hasson, U., Nir, Y., Fuhrmann, G., and Malach, R. (2004). Intersubject synchronization of cortical activity during natural vision. *Science* 303, 1634–1640. doi:10.1126/science.1089506
- Hlinka, J., and Coombes, S. (2012). Using computational models to relate structural and functional brain connectivity. *Eur. J. Neurosci.* 36, 2137–2145. doi:10.1111/j.1460-9568.2012.08081.x
- Holsheimer, J., and Feenstra, B. W. A. (1977). Volume conduction and eeg measurements within the brain: A quantitative approach to the influence of electrical spread on the linear relationship of activity measured at different locations. *Electroencephalogr. Clin. Neurophysiol.* 43, 52–58. doi:10.1016/0013-4694(77)90194-8
- Hou, Y. S., Xia, G. Q., Jayaprasath, E., Yue, D. Z., and Wu, Z. M. (2020). Parallel information processing using a reservoir computing system based on mutually coupled semiconductor lasers. *Appl. Phys. B* 126, 40. doi:10.1007/s00340-019-7351-4
- Hubbard, A. E., and Mountain, D. C. (1996). “Analysis and synthesis of cochlear mechanical function using models,” in *Auditory computation*. Editors H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (New York: Springer), 62–120. chap. 3. doi:10.1007/978-1-4612-4070-9\_3
- Jalili, M., Barzegaran, E., and Knyazeva, M. G. (2014). Synchronization of eeg: Bivariate and multivariate measures. *IEEE Trans. Neural Syst. Rehabil. Eng.* 22, 212. doi:10.1109/tnrsre.2013.2289899
- Jasper, H. H. (1958). The ten-twenty electrode system of the international federation. *Electroencephalogr. Clin. Neurophysiol.* 10, 371–375.
- Joris, P. X., Carney, L. H., Smith, P. H., and Yin, T. C. T. (1994). Enhancement of neural synchronization in the anteroventral cochlear nucleus. I. Responses to tones at the characteristic frequency. *J. Neurophysiol.* 71, 1022–1036. doi:10.1152/jn.1994.71.3.1022
- Kacprzyk, J., and Pedrycz, W. (2015). “Springer handbook of computational intelligence,” in *Springer handbooks* (Heidelberg, Germany: Springer Berlin Heidelberg).
- Kaneshiro, B., Nguyen, D. T., Norcia, A. M., Dmochowski, J. P., and Berger, J. (2021). Inter-subject eeg correlation reflects time-varying engagement with natural music. *Preprint*. doi:10.1101/2021.04.14.439913
- Kayser, J., and Tenke, C. E. (2006). Principal components analysis of laplacian waveforms as a generic method for identifying erp generator patterns: I. Evaluation with auditory oddball tasks. *Clin. Neurophysiol.* 117, 348–368. doi:10.1016/j.clinph.2005.08.034
- Keil, A., Müller, M. M., Ray, W. J., Gruber, T., and Elbert, T. (1999). Human gamma band activity and perception of a gestalt. *J. Neurosci.* 19, 7152–7161. doi:10.1523/jneurosci.19-16-07152.1999
- Kim, H., Moon, J.-Y., Mashour, G. A., and Lee, U. (2018). Mechanisms of hysteresis in human brain networks during transitions of consciousness and unconsciousness: Theoretical principles and empirical evidence. *PLoS Comput. Biol.* 14, e1006424. doi:10.1371/journal.pcbi.1006424
- Koelsch, S. (2014). Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180. doi:10.1038/nrn3666
- Koelsch, S., Rohrmeier, M., Torrecuso, R., and Jentschke, S. (2013). Processing of hierarchical syntactic structure in music. *Proc. Natl. Acad. Sci. U. S. A.* 110, 15443–15448. doi:10.1073/pnas.1300272110
- Kurth, E. (1931). *Musikpsychologie*. Berlin: Hesse.
- Lachaux, J. P., Rodriguez, E., Martinerie, J., and Varela, F. J. (1999). Measuring phase synchrony in brain signals. *Hum. Brain Mapp.* 8, 194–208. doi:10.1002/(sici)1097-0193(1999)8:4<194::aid-hbm4>3.0.co;2-c
- Large, E. W., Herrera, J. A., and Velasco, M. J. (2015). Neural networks for beat perception in musical rhythm. *Front. Syst. Neurosci.* 9, 159. doi:10.3389/fnsys.2015.00159
- Lehne, M., and Koelsch, S. (2015). Toward a general psychological model of tension and suspense. *Front. Psychol.* 6, 79. doi:10.3389/fpsyg.2015.00079
- Leman, M. (1997). “Music, gestalt, and computing,” in *Studies in cognitive and systematic musicology* (Berlin, Heidelberg: Springer Berlin Heidelberg). vol. 1317 of SpringerLink Bücher. doi:10.1007/bfb0034102
- Lerdahl, F., and Jackendoff, R. (1990). *A generative theory of tonal music*. print edn. Cambridge, Mass: MIT Press, 4.
- Leyva, I., Sendiña-Nadal, I., Sevilla-Escoboza, R., Vera-Avila, V. P., Chholak, P., Boccaletti, S., et al. (2018). Relay synchronization in multiplex networks. *Sci. Rep.* 8, 8629. doi:10.1038/s41598-018-26945-w
- Linke, S., Bader, R., and Mores, R. (2021). Modeling synchronization in human musical rhythms using impulse pattern formulation (ipf).
- Lyon, R., and Shamma, S. (1996). “Auditory representations of timbre and pitch,” in *Auditory computation*. Editors H. L. Hawkins, T. A. McMullen, A. N. Popper, and R. R. Fay (New York: Springer), 221–270. chap. 6. doi:10.1007/978-1-4612-4070-9\_6
- Mascetti, G. G. (2016). Unihemispheric sleep and asymmetrical sleep: Behavioral, neurophysiological, and functional perspectives. *Nat. Sci. Sleep.* 8, 221–238. doi:10.2147/NSS.S71970
- Massobrio, P., de Arcangelis, L., Pasquale, V., Jensen, H. J., and Plenz, D. (2015). Criticality as a signature of healthy neural systems. *Front. Syst. Neurosci.* 9, 22. doi:10.3389/fnsys.2015.00022
- Melicher, T., Horacek, J., Hlinka, J., Spaniel, F., Tintera, J., Ibrahim, I., et al. (2015). White matter changes in first episode psychosis and their relation to the size of sample studied: A DTI study. *Schizophr. Res.* 162, 22–28. doi:10.1016/j.schres.2015.01.029
- Mizrahi, A., Shalev, A., and Nelken, I. (2014). Single neuron and population coding of natural sounds in auditory cortex. *Curr. Opin. Neurobiol.* 24, 103–110. doi:10.1016/j.conb.2013.09.007
- Moroni, F., Nobili, L., De Carli, F., Massimini, M., Francione, S., Marzano, C., et al. (2012). Slow eeg rhythms and inter-hemispheric synchronization across sleep and wakefulness in the human

- hippocampus. *Neuroimage* 60, 497–504. doi:10.1016/j.neuroimage.2011.11.093
- Muldoon, S. F., Pasqualetti, F., Gu, S., Cieslak, M., Grafton, S. T., Vettel, J. M., et al. (2016). Stimulation-based control of dynamic brain networks. *PLoS Comput. Biol.* 12, e1005076. doi:10.1371/journal.pcbi.1005076
- Nagumo, J., Arimoto, S., and Yoshizawa, S. (1962). An active pulse transmission line simulating nerve axon. *Proc. IRE* 50, 2061–2070. doi:10.1109/jrproc.1962.288235
- Neuhauss, C. (2013). Processing musical form: Behavioural and neurocognitive approaches. *Musicae Sci.* 17, 109–127. doi:10.1177/1029864912468998
- Nicosia, V., Valencia, M., Chavez, M., Díaz-Guilera, A., and Latora, V. (2013). Remote synchronization reveals network symmetries and functional modules. *Phys. Rev. Lett.* 110, 174102. doi:10.1103/physrevlett.110.174102
- Nikolić, D., Fries, P., and Singer, W. (2013). Gamma oscillations: Precise temporal coordination without a metronome. *Trends Cogn. Sci.* 17, 54–55. doi:10.1016/j.tics.2012.12.003
- Omelchenko, I., Omel'chenko, O. E., Hövel, P., and Schöll, E. (2013). When nonlocal coupling between oscillators becomes stronger: Patched synchrony or multichimera states. *Phys. Rev. Lett.* 110, 224101. doi:10.1103/physrevlett.110.224101
- Owen, M., and Guta, M. P. (2019). Physically sufficient neural mechanisms of consciousness. *Front. Syst. Neurosci.* 13, 24. doi:10.3389/fnsys.2019.00024
- Palva, J. M., and Palva, S. (2012). Infra-slow fluctuations in electrophysiological recordings, blood-oxygenation-level-dependent signals, and psychophysical time series. *Neuroimage* 62, 2201–2211. doi:10.1016/j.neuroimage.2012.02.060
- Petkoski, S., and Jirsa, V. K. (2019). Transmission time delays organize the brain network synchronization. *Philos. Trans. A Math. Phys. Eng. Sci.* 377, 20180132. doi:10.1098/rsta.2018.0132
- Petkoski, S., Palva, J. M., and Jirsa, V. K. (2018). Phase-lags in large scale brain synchronization: Methodological considerations and *in-silico* analysis. *PLoS Comput. Biol.* 14, e1006160. doi:10.1371/journal.pcbi.1006160
- RamLOW, L., Sawicki, J., Zakharova, A., Hlinka, J., Clausen, J. C., Schöll, E., et al. (2019). Partial synchronization in empirical brain networks as a model for unihemispheric sleep. *EPL* 126, 50007. doi:10.1209/0295-5075/126/50007
- Rattenborg, N. C., Amlaner, C. J., and Lima, S. L. (2000). Behavioral, neurophysiological and evolutionary perspectives on unihemispheric sleep. *Neurosci. Biobehav. Rev.* 24, 817–842. doi:10.1016/s0149-7634(00)00039-7
- Rattenborg, N. C., Voirin, B., Cruz, S. M., Tisdale, R., Dell'omo, G., Lipp, H. P., et al. (2016). Evidence that birds sleep in mid-flight. *Nat. Commun.* 7, 12468. doi:10.1038/ncomms12468
- Ribeiro, T. L., Copelli, M., Caixeta, F., Belchior, H., Chialvo, D. R., Nicolelis, M. A., et al. (2010). Spike avalanches exhibit universal dynamics across the sleep-wake cycle. *PLoS One* 5, e14129. doi:10.1371/journal.pone.0014129
- Rodriguez, E., George, N., Lachaux, J. P., Martinerie, J., Renault, B., Varela, F. J., et al. (1999). Perception's shadow: Long-distance synchronization of human brain activity. *Nature* 397, 430–433. doi:10.1038/17120
- Sawicki, J., Abel, M., and Schöll, E. (2018a). Synchronization of organ pipes. *Eur. Phys. J. B* 91, 24. doi:10.1140/epjb/e2017-80485-8
- Sawicki, J. (2019). *Delay controlled partial synchronization in complex networks*. Heidelberg: Springer. Springer Theses. doi:10.1007/978-3-030-34076-6\_5
- Sawicki, J., Koulen, J. M., and Schöll, E. (2021). Synchronization scenarios in three-layer networks with a hub. *Chaos* 31, 073131. doi:10.1063/5.0055835
- Sawicki, J., Omelchenko, I., Zakharova, A., and Schöll, E. (2018b). Delay controls chimera relay synchronization in multiplex networks. *Phys. Rev. E* 98, 062224. doi:10.1103/physrev.98.062224
- Sawicki, J., Omelchenko, I., Zakharova, A., and Schöll, E. (2018c). Synchronization scenarios of chimeras in multiplex networks. *Eur. Phys. J. Spec. Top.* 227, 1161–1171. doi:10.1140/epjst/e2018-800039-y
- Sawicki, J., and Schöll, E. (2021). Influence of sound on empirical brain networks. *Front. Appl. Math. Stat.* 7, 662221. doi:10.3389/fams.2021.662221
- Schilling, K. G., Daducci, A., Maier-Hein, K., Poupon, C., Houde, J.-C., Nath, V., et al. (2019). Challenges in diffusion mri tractography—lessons learned from international benchmark competitions. *Magn. Reson. Imaging* 57, 194–209. doi:10.1016/j.mri.2018.11.014
- Schneider, A. (2018). "Fundamentals," in *Springer handbook of systematic musicology*. Editor R. Bader (Berlin and Heidelberg: Springer), 559–603. doi:10.1007/978-3-662-55004-5\_30
- Schoffelen, J. M., and Gross, J. (2009). Source connectivity analysis with meg and eeg. *Hum. Brain Mapp.* 30, 1857–1865. doi:10.1002/hbm.20745
- Schofield, B. R. (2011). *Auditory and vestibular efferents*. New York: Springer. doi:10.1007/978-1-4419-7070-1
- Schöll, E. (2021). Partial synchronization patterns in brain networks. *Europhys. Lett.* 136, 18001. doi:10.1209/0295-5075/ac3b97
- Schwartz, J. R. L., and Roth, T. (2008). Neurophysiology of sleep and wakefulness: Basic science and clinical implications. *Curr. Neuropharmacol.* 6, 367–378. doi:10.2174/157015908787386050
- Shainline, J. M. (2020). Fluxonic processing of photonic synapse events. *IEEE J. Sel. Top. Quantum Electron.* 26, 1–15. doi:10.1109/jstqe.2019.2927473
- Shi, J., Kirihara, K., Tada, M., Fujioka, M., Usui, K., Koshiyama, D., et al. (2022). Criticality in the healthy brain. *Front. Netw. Physiol.* 1, 755685. doi:10.3389/fnetp.2021.755685
- Spitmaan, H., and Seo, M., Lee, D., and Soltani, A. (2020). Multiple timescales of neural dynamics and integration of task-relevant signals across cortex. *Proc. Natl. Acad. Sci. U. S. A.* 117, 22522–22531. doi:10.1073/pnas.2005993117
- Steriade, M., McCormick, D. A., and Sejnowski, T. J. (1993). Thalamocortical oscillations in the sleeping and aroused brain. *Science* 262, 679–685. doi:10.1126/science.8235588
- Steyn-Ross, A., and Steyn-Ross, M. (2010). *Modeling phase transitions in the brain*, 509. Berlin: Springer. doi:10.1007/978-1-4419-0796-7
- Tallon, C., Bertrand, O., Bouchet, P., and Pernier, J. (1995). Gamma-range activity evoked by coherent visual stimuli in humans. *Eur. J. Neurosci.* 7, 1285–1291. doi:10.1111/j.1460-9568.1995.tb01118.x
- Tallon-Baudry, C., Bertrand, O., Delpuech, C., and Pernier, J. (1996). Stimulus specificity of phase-locked and non-phase-locked 40 hz visual responses in human. *J. Neurosci.* 16, 4240–4249. doi:10.1523/JNEUROSCI.16-13-04240.1996
- Tallon-Baudry, C., and Bertrand, O. (1999). Oscillatory gamma activity in humans and its role in object representation. *Trends Cogn. Sci.* 3, 151–162. doi:10.1016/s1364-6613(99)01299-1
- Tamaki, M., Bang, J. W., Watanabe, T., and Sasaki, Y. (2016). Night watch in one brain hemisphere during sleep associated with the first-night effect in humans. *Curr. Biol.* 26, 1190–1194. doi:10.1016/j.cub.2016.02.063
- Thaut, M. H., McIntosh, G. C., and Hoemberg, V. (2015). Neurobiological foundations of neurologic music therapy: Rhythmic entrainment and the motor system. *Front. Psychol.* 5, 1185. doi:10.3389/fpsyg.2014.01185
- Tritsch, N. X., Rodríguez-Contreras, A., Crins, T. T., Wang, H. C., Borst, J. G., Bergles, D. E., et al. (2010). Calcium action potentials in hair cells pattern auditory neuron activity before hearing onset. *Nat. Neurosci.* 13, 1050–1052. doi:10.1038/nn.2604
- Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., et al. (2002). Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* 15, 273–289. doi:10.1006/nimg.2001.0978
- van Noorden, L., and Moelants, D. (1999). Resonance in the perception of musical pulse. *J. New Music Res.* 28, 43–66. doi:10.1076/jnmr.28.1.43.3122
- Winkler, M., Sawicki, J., Omelchenko, I., Zakharova, A., Anishchenko, V., Schöll, E., et al. (2019). Relay synchronization in multiplex networks of discrete maps. *EPL* 126, 50004. doi:10.1209/0295-5075/126/50004
- Womelsdorf, T., and Fries, P. (2007). The role of neuronal synchronization in selective attention. *Curr. Opin. Neurobiol.* 17, 154–160. doi:10.1016/j.conb.2007.02.002
- Zanto, T., Large, E. W., Fuchs, A., and Kelso, J. A. S. (2005). Gamma-band responses to perturbed auditory sequences: Evidence for synchronization of perceptual processes. *Music Percept.* 22, 531–547. doi:10.1525/mp.2005.22.3.531
- Zhang, L., Motter, A. E., and Nishikawa, T. (2017a). Incoherence-mediated remote synchronization. *Phys. Rev. Lett.* 118, 174102. doi:10.1103/physrevlett.118.174102
- Zhang, Y., Nishikawa, T., and Motter, A. E. (2017b). Asymmetry-induced synchronization in oscillator networks. *Phys. Rev. E* 95, 062215. doi:10.1103/physrev.95.062215