



# Modal-Polar Representation of Evoked Response Potentials in Multiple Arousal States

Rawan K. El-Zghir<sup>1,2\*</sup>, Natasha C. Gabay<sup>1,2</sup> and Peter A. Robinson<sup>1,2</sup>

<sup>1</sup> School of Physics, University of Sydney, Sydney, NSW, Australia, <sup>2</sup> Center for Integrative Brain Function, University of Sydney, Sydney, NSW, Australia

An expansion of the corticothalamic transfer function into eigenmodes and resonant poles is used to derive a simple formula for evoked response potentials (ERPs) in various states of arousal. The transfer function corresponds to the cortical response to an external stimulus, which encodes all the information and properties of the linear system. This approach links experimental observations of resonances and characteristic timescales in brain activity with physically based neural field theory (NFT). The present work greatly simplifies the formula of the analytical ERP, and separates its spatial part (eigenmodes) from the temporal part (poles). Within this framework, calculations involve contour integrations that yield an explicit expression for ERPs. The dominant global mode is considered explicitly in more detail to study how the ERP varies with time in this mode and to illustrate the method. For each arousal state in sleep and wake, the resonances of the system are determined and it is found that five poles are sufficient to study the main dynamics of the system in waking eyes-open and eyes-closed states. Similarly, it is shown that six poles suffice to reproduce ERPs in rapid-eye movement sleep, sleep state 1, and sleep state 2 states, whereas just four poles suffice to reproduce the dynamics in slow wave sleep. Thus, six poles are sufficient to preserve the main global ERP dynamics of the system for all states of arousal. These six poles correspond to the dominant resonances of the system at slow-wave, alpha, and beta frequencies. These results provide the basis for simplified analytic treatment of brain dynamics and link observations more closely to theory.

## OPEN ACCESS

### Edited by:

Chun Siong Soon,  
National University of Singapore,  
Singapore

### Reviewed by:

Daniela Dentico,  
Independent Researcher, Puglia, Italy  
Johan Van Der Meer,  
The University of Queensland,  
Australia

### \*Correspondence:

Rawan K. El-Zghir  
relz2030@uni.sydney.edu.au

### Specialty section:

This article was submitted to  
Cognitive Neuroscience,  
a section of the journal  
Frontiers in Human Neuroscience

**Received:** 16 December 2020

**Accepted:** 10 May 2021

**Published:** 07 June 2021

### Citation:

El-Zghir RK, Gabay NC and  
Robinson PA (2021) Modal-Polar  
Representation of Evoked Response  
Potentials in Multiple Arousal States.  
*Front. Hum. Neurosci.* 15:642479.  
doi: 10.3389/fnhum.2021.642479

**Keywords:** neural field theory, brain dynamics, evoked response potentials, brain resonances, eigenmodes

## 1. INTRODUCTION

Evoked response potentials (ERPs) reflect the electrical activity of the brain triggered by sensory stimuli or events (Niedermeyer and Lopes Da Silva, 1999). ERPs have been widely used to provide windows on cognitive processes such as attention and perception (Luck and Kappenman, 2013). Most notably, attentional modulation enhances both early and late features of an ERP (Hillyard and Anllo-Vento, 1998; Herrmann and Knight, 2001; Vázquez Marrufo et al., 2001; Yamagishi et al., 2003). It has been shown that many characteristics of the stimulus and subject affect ERP waveforms (Picton et al., 2000; Womelsdorf et al., 2007; Woodman, 2010), notably including the state of arousal of the subject, with very different ERPs in sleep than wake, for example (Feng et al., 2012).

Traditional phenomenological analysis reduces ERPs to a small set of “components” defined by amplitudes and latencies (time delays after the stimulus) each of which correspond to a peak or trough in the waveform (Luck, 2014); each component is presumed to be generated by a group of excitatory or inhibitory neurons that have a certain cognitive role (Luck and Kappenman, 2013), but there is no explicit link to physiology and it is common to omit most data points and focus on only amplitudes and latencies of a few components. Moreover, it has been widely recognized that ERPs can be treated as impulse responses whose building blocks are damped sinusoids that reflect the dynamics of the underlying physical system that generates them (Kelly and Reilly, 1983).

Over the last 20 years, many quantitative brain studies have been performed based on a corticothalamic neural field theory (NFT) (Wright and Liley, 1996; Robinson et al., 1997, 2002, 2004, 2005). It has been shown that much brain activity is approximately linear and that ERPs can thus be described by a system transfer function that describes the response to a delta function input (Rennie et al., 1999; Robinson et al., 2005; Kerr et al., 2008). Such a function can also be used to calculate responses to arbitrary stimuli, and to derive other activity-dependent quantities such as correlations (Robinson et al., 2018; Robinson, 2019). Corticothalamic NFT averages over scales of a few tenths of a mm and has been an essential tool to accurately predict significant electroencephalographic (EEG) features such as ERPs that have been previously calculated and verified against experiment (Rennie et al., 1999; Robinson et al., 2005; Kerr et al., 2008). The corticothalamic model introduces physiologically based parameters that correspond to different physical quantities, and each real brain state is described by a particular set of parameters; consequently, transfer functions can be related to underlying physiology via NFT.

Within NFT, the transfer function can be expanded in terms of eigenfunctions which are the natural modes of the brain (Robinson et al., 1997, 2001) and the building blocks of normal brain dynamics (Pinotsis et al., 2013; Robinson et al., 2016). Eigenmodes of the corticothalamic system manifest on the cortical surface at length scales detectable with EEG, magnetoencephalography (MEG), and fMRI. Using NFT these modes can be described in terms of physiological parameters such as corticocortical, intrathalamic, and corticothalamic feedback loop strengths and inverse synaptodendritic decay and rise times (Gabay and Robinson, 2017). Such a modal perspective has yielded fruitful results regarding dynamic brain connectivity via spectral analysis (Gabay and Robinson, 2017; Gabay et al., 2018).

NFT expressions for the time dependence of activity are relatively complicated, as are the expressions for the transfer functions, which involve transcendental equations and are not analytically tractable. This poses a problem when comparing with experimental results that often involve expressions in terms of characteristic frequencies and time delays. Hence, it is desirable to recast the outcomes of NFT in terms of observable quantities in order to bridge more directly between theory and experiment. Recent work has begun to address this problem via rational function approximations to the temporal transfer function that express it compactly in terms of a small number of poles, as

is often done in control theory (Babaie-Janvier and Robinson, 2018). Interestingly, this approach has yielded interpretations of EEG resonances in terms of proportional-integral-derivative (PID) filters that are commonly used in engineering control systems for predicting the future course of inputs within their corresponding frequency ranges (Kwakernaak and Sivan, 1972; Ogata, 2010; Babaie-Janvier and Robinson, 2019).

In the present work we use NFT to analyze cortical activity by representing the arousal states of the brain in terms of its normal modes. This approach has also been used to approximate brain activity by decomposing the transfer function first in terms of eigenfunctions (spatial modes) then the temporal response of each eigenfunction in terms of poles (resonant frequencies), but not many results about resonances are available in the literature, except for waking states of arousal (Gabay et al., 2018). In order to better understand how many poles to include for accurate representations, we systematically investigate the modal-polar representation of the transfer function in various states of arousal and provide parameters of the poles in each state for the first time. Hence, the present work provides a modal-polar representation of corticothalamic NFT that bridges between the mathematical NFT and the applied work done by neuroscientists in measuring resonances and impulse responses.

In this paper, the temporal transfer function is decomposed in terms of poles or resonances to derive a general formula for ERPs which greatly simplifies their analytical form. This framework simplifies NFT results, and links NFT to observables by expanding previous work that has calculated the poles for the waking states only (Gabay et al., 2018). It also enables observed resonances to be interpreted in terms of the transfer function's poles, and hence data filters of a type previously shown to implement prediction and attention (Babaie-Janvier and Robinson, 2018, 2019). Section 2 provides an overview of the corticothalamic NFT. In section 3, we illustrate the derivation of ERP formula based on modal-polar expansion. Then, in section 4 we show and discuss our results. Finally, section 5 summarizes our work.

## 2. MATERIALS AND METHODS

In this section we first review the relevant corticothalamic NFT in section 2.1, based on previous studies (Robinson et al., 1997, 2002; Kerr et al., 2008; Babaie-Janvier and Robinson, 2018; Gabay et al., 2018). Further details can be found in the references cited. Then, we use the modal-polar representation of the corticothalamic NFT as a tool for finding an explicit expression of the transcendental transfer function in terms of the brain's natural modes and their corresponding poles. Finally, in section 2.3 we discuss the spectral features and parameters characterizing each arousal state.

### 2.1. Corticothalamic Neural Field Theory

Neural field theory averages neural quantities over scales of a few tenths of a millimeter. It has been widely used to interpret and reproduce key features of experimental findings in EEG and fMRI. The neural field model consists of cortical excitatory ( $e$ )

and inhibitory (*i*) populations, thalamic specific relay populations (*s*), thalamic reticular populations (*r*), and external sensory inputs (*n*). This model incorporates key anatomic connectivities between those populations, as shown in **Figure 1**, where  $\phi_{ab}$  is the mean activity field reaching population *a* due to signals from population *b*. The strength of connection to population *a* from population *b* is Robinson et al. (2005)

$$v_{ab} = s_{ab}N_{ab}, \tag{1}$$

where  $s_{ab}$  is the mean time-integrated strength of the response in neurons *a* per incoming signal from neurons *b*, and  $N_{ab}$  is the mean number of synapses to neurons *a* from *b*.

It is found that population *a*'s average firing rate  $Q_a$  is linked to the corresponding average membrane potential  $V_a$ , relative to resting, by a nonlinear sigmoid function

$$Q_a = S(V_a) = \frac{Q_{\max}}{1 + \exp[-(V_a - \theta)/\sigma']}, \tag{2}$$

where  $Q_{\max}$  is the maximum firing rate,  $\theta$  is the mean threshold voltage, and  $\sigma'\pi/\sqrt{3}$  is the standard deviation of the threshold distribution.

Due to synaptodendritic dynamics and soma capacitance, presynaptic inputs *b* from different types of neurons *a* are

summed after being filtered and smeared out in time, giving rise to the potential  $V_a$  such that

$$V_a(\mathbf{r}, t) = \sum_b V_{ab}(\mathbf{r}, t). \tag{3}$$

We can convert from the signal/stimulus to the population response using the following equation:

$$D_\alpha(t)V_{ab}(\mathbf{r}, t) = v_{ab}\phi_{ab}(\mathbf{r}, t - \tau_{ab}), \tag{4}$$

where  $D_\alpha$  is the wave propagation operator, given by

$$D_\alpha(t) = \frac{1}{\alpha\beta} \frac{d^2}{dt^2} + \left(\frac{1}{\alpha} + \frac{1}{\beta}\right) \frac{d}{dt} + 1, \tag{5}$$

where  $\mathbf{r}$  is the position vector on the brain,  $1/\beta$  and  $1/\alpha$  are the rise and decay times, respectively, of the potential at the cell body elicited by an impulse response at the dendritic synapse, and  $\tau_{ab}$  is the time delay due to anatomical separations between neural populations *a* and *b*, as specified in **Table 1**. The only nonzero time delays correspond to propagation time from cortex to thalamus and vice versa ( $\tau_{ab} \approx 0$  in the case of intrathalamic and intracortical connections).

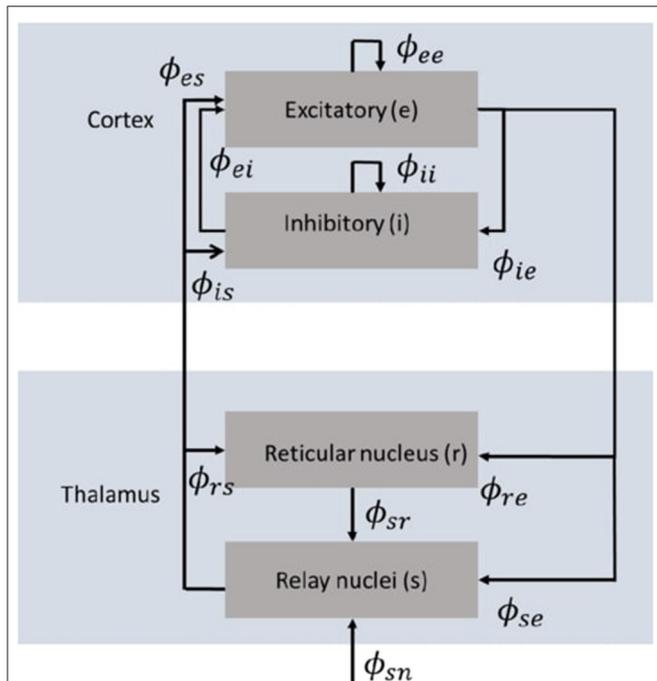
The mapping of the source firing rates  $Q_b(\mathbf{r}, t)$  into the axonal signal is achieved within the following equations (Robinson et al., 1997)

$$D_{ab}(\mathbf{r}, t)\phi_{ab}(\mathbf{r}, t) = Q_b(\mathbf{r}, t), \tag{6}$$

$$D_{ab}(\mathbf{r}, t) = \frac{1}{\gamma_{ab}^2} \frac{\partial^2}{\partial t^2} + \frac{2}{\gamma_{ab}} \frac{\partial}{\partial t} + 1 - r_{ab}^2 \nabla^2, \tag{7}$$

which describes the propagation of a mean activity field  $\phi_{ab}(\mathbf{r}, t)$  that obeys a damped wave equation where  $\gamma_{ab} = v_{ab}/r_{ab}$  is the temporal damping rate,  $r_{ab}$  is the mean characteristic range of axons to population *a* from population *b*, and  $v_{ab}$  is the propagation velocity in axons to population *a* from population *b*.

The mean axonal ranges for all populations except for excitatory cortical neurons are very short so we can write  $r_{ab} \approx 0$  for  $b = i, r, s$ , yielding  $D_{ab} = 1$  for these neural populations. Let us assume that the number of intracortical synapses is proportional to the number of neurons involved, then  $N_{ib} = N_{eb}$  for all *b* (Wright and Liley, 1996; Robinson et al., 1997;



**FIGURE 1** | Schematic diagram of the corticothalamic model (Gabay and Robinson, 2017) that incorporates key anatomic connectivities between neural populations, where  $\phi_{ab}$  is the mean activity field reaching population *a* due to signals from population *b*. There are two cortical populations of excitatory (*e*) and inhibitory (*i*) neurons, and two thalamic populations corresponding to the thalamic reticular nucleus (*r*) and the thalamic relay nuclei (*s*).

**TABLE 1** | Nominal parameters of corticothalamic neural field theory based on previous work (Babaie-Janvier and Robinson, 2019).

Symbol	Quantity	Value	Units
$\alpha$	Synaptodendritic decay rate	80	s <sup>-1</sup>
$\beta$	Synaptodendritic rise rate	320	s <sup>-1</sup>
$\tau_{es}$	Thalamocortical axonal delay	0.02	s
$\tau_{se}$	Corticothalamic relay axonal delay	0.06	s
$\gamma_{ee}$	Cortical damping rate	116	s <sup>-1</sup>
$r_{ee}$	Excitatory axon range	0.086	m

Braitenberg and Schüz, 1998). We also assume that the input stimulus  $\phi_{sn}$  is not too large such that the system has a fixed point with low firing rates (Robinson et al., 2002). Hence, we can apply linear approximation where we henceforth treat each dynamic quantity ( $\phi_{ab}, Q_a, V_a$ ) as a linear perturbation from its steady state value, which is denoted by the superscript (0). Then

$$Q_a(\mathbf{r}, t) = \rho_a V_a(\mathbf{r}, t), \tag{8}$$

where

$$\rho_a = S' \left( V_a^{(0)} \right), \tag{9}$$

is the derivative with respect to voltage of the sigmoid function evaluated at the steady state.

By taking the Fourier transform of the above equations, we get

$$V_{ab}(\mathbf{k}, \omega) = L(\omega) \nu_{ab} \phi_{ab}(\mathbf{k}, \omega) e^{i\omega\tau_{ab}}, \tag{10}$$

$$L(\omega) = \left( 1 - \frac{i\omega}{\alpha} \right)^{-1} \left( 1 - \frac{i\omega}{\beta} \right)^{-1}, \tag{11}$$

$$D_{ab}(\mathbf{k}, \omega) \phi_{ab}(\mathbf{k}, \omega) = Q_b(\mathbf{k}, \omega), \tag{12}$$

where  $\mathbf{k}$  is the wave vector and  $L(\omega)$  embodies a low-pass filter response function. Activity  $Q_b$  generates fields of activity  $\phi_{ab}$  that propagate through the brain to affect populations  $a$ . We express the firing rate  $\phi_{ee}$  in terms of the external stimulus  $\phi_{sn}$  via the transfer function

$$T(\mathbf{k}, \omega) = \frac{\phi_{ee}(\mathbf{k}, \omega)}{\phi_{sn}(\mathbf{k}, \omega)}, \tag{13}$$

$$= \frac{L^2 G_{es} G_{sn} e^{i\omega\tau_{es}}}{(1 - L^2 G_{sr} G_{rs})(1 - G_{ei} L)} \frac{1}{k^2 r_{ee}^2 + q^2(\omega) r_{ee}^2}, \tag{14}$$

where we assign

$$q^2(\omega) r_{ee}^2 = \left( 1 - \frac{i\omega}{\gamma_{ee}} \right)^2 - \frac{1}{1 - G_{ei} L} \left[ L G_{ee} + \frac{L^2 G_{es} (G_{se} + L G_{sr} G_{re}) \exp[i\omega(\tau_{es} + \tau_{se})]}{1 - L^2 G_{sr} G_{rs}} \right]. \tag{15}$$

and  $G_{ab}$  is the gain of responses in population  $a$  due to signals from population  $b$  such that  $G_{ab} = \rho_a \nu_{ab}$ . The above form of the transfer function is transcendental, which is not easy to work with and does not link easily to observable features such as resonances in the EEG spectrum, so we seek to simplify it via the approximations below. Note that, in a linear system, all the other  $\phi_{ab}$  are linearly related to  $\phi_{ee}$ .

Let  $D'(\omega) = q^2(\omega) r_{ee}^2$ , then in the absence of external stimulus Equation (14) yields

$$D'(\omega) \phi_{ee} = -r_{ee}^2 \nabla^2 \phi_{ee}, \tag{16}$$

where  $\nabla^2$  is the Laplace-Beltrami operator.

We analyze the spatiotemporal dynamics of the brain in terms of discrete modes labeled  $\eta$ . To solve Equation (16) for these eigenmodes, we introduce the ansatz

$$\phi_{ee} = u_\eta(\mathbf{r}) e^{-i\omega_\eta t}, \tag{17}$$

where  $u_\eta(\mathbf{r})$  represents the spatial eigenmode on the cortical surface oscillating at an eigenfrequency  $\omega_\eta$ . Substituting (Equation 17) into (Equation 16) and using separation of variables yields the Helmholtz equation for the spatial eigenmode (Gabay and Robinson, 2017)

$$\nabla^2 u_\eta(\mathbf{r}) = -k_\eta^2 u_\eta(\mathbf{r}), \tag{18}$$

where  $u_\eta(\mathbf{r})$  are the eigenmode solutions of the equation,  $k_\eta$  are the wavenumbers and  $k_\eta^2$  are the eigenvalues. By substituting the eigenvalues  $k_\eta^2$  from Equation (18) into the dispersion relation

$$k_\eta^2 + q_\eta^2(\omega) = 0, \tag{19}$$

we get

$$D'_\eta(\omega) = -k_\eta^2 r_{ee}^2. \tag{20}$$

There are an infinite number of solutions of the dispersion relation because it is transcendental, although, as has been previously shown, a small number dominate the dynamics. These solutions correspond to the eigenfrequencies of the eigenmodes  $u_\eta(\mathbf{r})$  and they correspond to the poles of the transfer function  $T(k_\eta^2, \omega)$  (Gabay et al., 2018).

## 2.2. Modal-Polar Expansion

The transfer function represents the cortical response to an external stimulus. It encodes all the properties of the linear system. By expanding the transfer function in terms of modes, we can decompose (Equations 13, 14) into two separate parts: the modal part which is spatial, and the temporal part, as in Equation (17). Separation of variables is more general, but if we assume that the transfer function is spatially symmetric, then the set of eigenmodes  $u_\eta$  is complete and orthonormal, which means that any connectivity or activity whatsoever can be expressed in terms of  $u_\eta$  (Robinson et al., 2018), as

$$T(\mathbf{r}, \mathbf{r}', \omega) = \sum_\eta u_\eta(\mathbf{r}) u_\eta^*(\mathbf{r}') \theta_\eta(\omega), \tag{21}$$

where  $u_\eta(\mathbf{r})$  are the eigenmodes, and  $\theta_\eta(\omega)$  is the temporal part of the transfer function defined in Equation (23). Also, we assume that the brain is a static network on the timescales of interest, then the transfer function depends on the time difference  $t - t_0$  (Robinson, 2019); i.e.,

$$T(\mathbf{r}, t; \mathbf{r}_0, t_0) = T(\mathbf{r}, \mathbf{r}_0, t - t_0). \tag{22}$$

Because the dispersion relation in Equation (20) is transcendental, it has an infinite number of solutions. This means that there are infinite number of poles (eigenfrequencies)

for the corresponding eigenmode  $u_\eta(\mathbf{r})$ . However, transfer functions are ratios of exponential polynomials of  $-i\omega$ . Then, applying rational approximation enables us to write the transcendental transfer function in Equation (15) in the form (Gabay et al., 2018)

$$\theta_\eta(\omega) \approx \frac{\sum_{p=0}^m B_p(k_\eta^2)(-i\omega)^p}{\sum_{q=0}^n A_q(k_\eta^2)(-i\omega)^q}, \tag{23}$$

the above rational approximation is Padé approximation of order  $(m/n)$  which is used to convert the transcendental transfer function into a rational polynomial (Equation 23), where  $m$  and  $n$  are the degrees of the numerator and denominator, respectively, and  $n > m$  (Golub and Van Loan, 1983). The coefficients  $B_p$  and  $A_q$  are real and depend on  $k_\eta^2$  (the eigenvalue of the mode  $\eta$ ), and  $n$  represents the number of poles where the denominator vanishes. Increasing accuracy is attained as we consider more poles in the expression. However, an important aim is to find the smallest number of poles that retains the main dynamics of the system while maintaining a high accuracy. Note that the form in Equation (23) has many advantages because of its analytic simplicity and pole structure (Gabay et al., 2018), as we see below.

Applying partial fraction decomposition of the rational approximate transfer function in Equation (23) yields (Varga, 1991; Babaie-Janvier and Robinson, 2018)

$$\theta_\eta(\omega) \approx \sum_{j=1}^n \frac{r_{\eta j}(k_\eta^2)}{\omega - \omega_{\eta j}(k_\eta^2)}, \tag{24}$$

where  $r_{\eta j}$  and  $\omega_{\eta j}$  are the residues and poles, respectively, and we truncate the sum at  $j = n$ . We assume that the modes are nondegenerate with wave numbers  $k_\eta$  and have complex eigenfrequencies  $\omega_{\eta j}$ . The resonant frequencies are of the form  $\omega_{\eta j} = \Omega_{\eta j} - i\gamma_{\eta j}$ , where  $\Omega_{\eta j}$  and  $\gamma_{\eta j}$  are real and represent the angular frequency and damping rate, respectively, of the  $j^{\text{th}}$  pole at the  $\eta^{\text{th}}$  mode. We omit the explicit dependence on  $k_\eta^2$ , for simplicity, and write  $T$  for mode  $\eta$  as  $\theta_\eta(\omega)$  in accordance with Equation (21). Then,

$$\theta_\eta(\omega) \approx \sum_{j=1}^n \frac{r_{\eta j}}{\omega - \omega_{\eta j}}. \tag{25}$$

The form (21) with the above expression of  $\theta_\eta(\omega)$  is termed modal-polar representation of corticothalamic transfer function.

### 2.3. Arousal State Characterization

In this section we apply our methods to arousal states that correspond to the levels defined by Rechtschaffen and Kales (1968): wake (W) which includes both the eyes-open (EO) and eye-closed (EC) states; sleep state 1 (S1) which refers to light sleep; sleep state 2 (S2) which is a deeper level of sleep; slow-wave sleep (SWS) which is the deepest level; rapid-eye movement (REM); and sleep spindles that Rechtschaffen and Kales (1968) defined them as “a burst of oscillatory brain activity visible on an EEG that occurs during S2.” In order to distinguish

between each arousal state, we examine its corresponding EEG power spectrum. For normal adult humans, the frequency is conventionally divided into the following bands (Niedermeyer and Lopes Da Silva, 1999; Buzsáki and Draguhn, 2004): infraslow oscillations (0.01–0.1 Hz); delta (1.5–4 Hz); theta (4–7.5 Hz); alpha (7.5–13 Hz); spindle (11–16 Hz); and beta (13–30 Hz). The frequencies between 0.1 and 1.5 Hz are referred to as slow-waves, and they are characteristic waves in the delta band. Buzsáki and Draguhn (2004) divided these slow-wave frequencies into four classes (slow 1, slow 2, slow 3, and slow 4).

The waking EC state is distinguished by a distinctive strong alpha peak, whereas the waking EO state has a weaker alpha peak but similar alpha-band power (Rechtschaffen and Kales, 1968; Niedermeyer and Lopes Da Silva, 1999; Chiang et al., 2008). Eventually, as we move from the waking EC to S1, the alpha peak strength decreases significantly with an increase in the delta power, which continues to rise in the deep sleep states (Iber et al., 2007; Van Albada and Robinson, 2013).

Over the last 25 years, it has been shown that the changes in peak and trough timings and amplitudes are a natural consequence of the dominant EEG frequencies in the various states of arousal. In this paper, we are modeling what the impulse response would be in the various background EEG states, which is relevant if an external impulsive stimulus is applied, as in ERP experiments. This approach has been successfully used to model ERPs in multiple previous studies, however, the new technique in this framework is the use and verification that a few-pole approximation suffices to produce the same results, thereby leading to a more compact representation.

The responses of the system transfer functions for all populations are based on the nominal parameters shown in **Tables 1, 2**. The gains in **Table 2** were derived by Abeysuriya et al. (2015) by fitting NFT predictions to experimental EEG spectra. For wake states, they used a data set of 2100 subjects with EEG recordings from the Brain Resource International Database, an archive of electrophysiological and psychophysiological measures, psychometric tests, and demographics (Gordon et al., 2005). For sleep states, they used manually scored polysomnograms from healthy controls (Wang et al., 2005; D’Rozario et al., 2013). These are spontaneous parameters and fits to observed ERPs could also be done using the same approach. Note that in **Table 2**, spindles are shown in one column with different gain parameters. Although spindles are considered to characterize Stage 2 sleep, previous findings suggest that spindles can be considered as transient substates and can span a different region of parameter space than simply S2 does (Abeysuriya et al., 2014). A state between SWS and Spindle would mix characteristics and produce a k-complex, a gallery of examples of these intermediate waveforms was shown in Figure 1 in Zobaer et al. (2017). Spindle dominates when that pole is only weakly damped; likewise, for slow waves. There is a region of parameter space where both poles are weakly damped and both oscillations can be seen. Figure 7 in Robinson et al. (2005) represents the stability zone and shows that the SWS and spindle zones meet. Gain parameters can in general be modulated by local feedback

between populations in response to incoming stimuli, which has been useful for analyzing how evoked responses are shaped by attention (Babaie-Janvier and Robinson, 2020). Hence, we could tweak the parameters, but the core aim in this work is to determine how many poles are required to get a reasonable approximation to analytically predicted ERPs for the same parameters over the whole range of arousal states. Parameters could be changed to account for the fact that ERPs may be generated by a subsystem with somewhat different properties, for example (Robinson, 2005). However, the gains remain fixed in the present work which focuses on the theoretical ERP forms and their simplifications. However, the gains remain fixed in the present work which focuses on the theoretical ERP forms and their simplifications. It should be noted that these model parameters correspond to representative brain states from each of the sleep stages, where each sleep stage is associated with a range of parameters, reflecting individual differences in EEG between subjects at the same arousal stage (Abey Suriya et al., 2015).

### 3. RESULTS

This section contains two sets of results: new theoretical developments and application to states of arousal. In the first part, we derive a simple expression for the ERP based on the modal-polar transfer function. In the second part, we apply the results to the different arousal states by plotting the locations of poles for different cases of number of poles and finding the root mean square error to study the convergence of  $T(f)$  (where the frequency in Hz is  $f = \omega/2\pi$ ) and  $ERP(t)$  to their exact results. This enables us to estimate the number of poles needed to attain high accuracy in studying each arousal state.

#### 3.1. ERP Derivation

For simplicity, throughout the present work we restrict attention to the spatially uniform global mode ( $k_\eta^2 = 0$ ), because this mode has the lowest damping rate and is the least stable, so it is the most easily excited and dominates the response (Robinson, 2003; Gabay and Robinson, 2017).

**TABLE 2 |** Gain parameters used in the present work based on previous studies from Abey Suriya et al. (2015) for the EO, EC, REM, S1, S2, SWS, and Spindles arousal states.

Gain	EO	EC	REM	S1	S2	SWS	Spindles
$G_{ee}$	10.50	2.07	5.87	7.45	16.86	19.52	18.52
$-G_{ei}$	13.22	4.11	6.61	8.30	17.93	19.74	18.96
$G_{es}$	1.21	0.77	0.21	0.31	3.89	5.30	2.55
$G_{se}$	5.78	7.77	0.66	1.67	0.07	0.22	0.73
$-G_{sr}$	2.83	3.30	0.28	0.40	0.14	0.22	0.26
$G_{sn}$	14.23	8.10	0.68	3.90	2.38	1.70	2.78
$G_{re}$	0.85	0.66	2.08	7.47	4.96	1.90	4.67
$G_{rs}$	0.25	0.20	4.59	4.44	8.33	1.35	16.92

The ERP for a stimulus at  $(\mathbf{r}_0, t_0)$  is identical to the transfer function; i.e.,

$$ERP(\mathbf{r}, \mathbf{r}_0, t) = T(\mathbf{r}, \mathbf{r}_0, t - t_0). \tag{26}$$

If we let  $t_0 = 0$  without loss of generality and Fourier transform, then

$$ERP(\mathbf{r}, \mathbf{r}_0, \omega) = T(\mathbf{r}, \mathbf{r}_0, \omega), \tag{27}$$

so

$$ERP(\mathbf{r}, \mathbf{r}_0, \omega) \approx \sum_{\eta} u_{\eta}(\mathbf{r}) u_{\eta}^*(\mathbf{r}_0) \sum_{j=1}^n \frac{r_{\eta j}}{\omega - \omega_{\eta j}}. \tag{28}$$

Applying an inverse Fourier transform we get

$$ERP(\mathbf{r}, \mathbf{r}_0, t) = \sum_{\eta} u_{\eta}(\mathbf{r}) u_{\eta}^*(\mathbf{r}_0) \theta_{\eta}(t), \tag{29}$$

$$ERP(\mathbf{r}, \mathbf{r}_0, t) \approx \sum_{\eta} u_{\eta}(\mathbf{r}) u_{\eta}^*(\mathbf{r}_0) \sum_{j=1}^n \int \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}}, \tag{30}$$

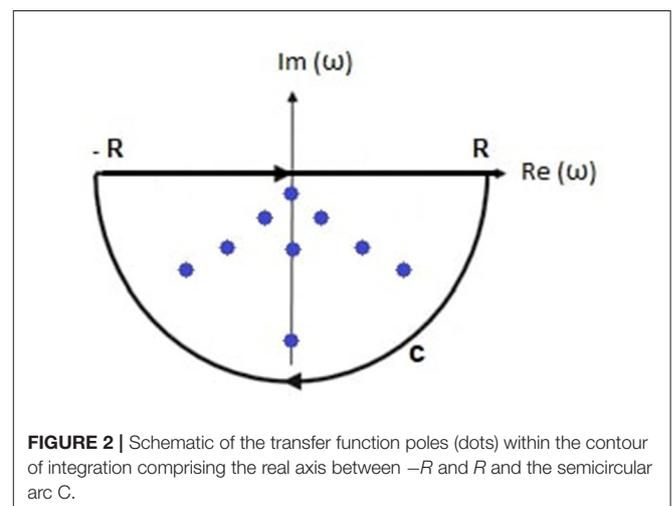
where

$$\theta_{\eta}(t) \approx \sum_{j=1}^n \theta_{\eta j}(t) = \sum_{j=1}^n \int \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}}, \tag{31}$$

is the temporal part of the ERP due to the mode  $\eta$ .

##### 3.1.1. Integration Contour

In this section, we evaluate Equation (31) by contour integration. The damping rate must be real and positive in a stable system, so  $\text{Im}(\omega_{\eta j}) < 0$ , as shown in **Figure 2**, and the appropriate integration contour lies in the negative imaginary half-plane



**FIGURE 2 |** Schematic of the transfer function poles (dots) within the contour of integration comprising the real axis between  $-R$  and  $R$  and the semicircular arc  $C$ .

of complex angular frequency and encloses all  $n$  poles being considered. Then

$$\oint \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}} = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{dx}{2\pi} e^{-ixt} \frac{Re(r_{\eta j})}{x - Re(\omega_{\eta j})} + \lim_{R \rightarrow \infty} \int_C \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}}. \quad (32)$$

By Jordan's Lemma (Mitrinović, 1984),

$$\lim_{R \rightarrow \infty} \int_C \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}} = 0. \quad (33)$$

So

$$\oint \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}} = \lim_{R \rightarrow \infty} \int_{-R}^R \frac{dx}{2\pi} e^{-ixt} \frac{Re(r_{\eta j})}{x - Re(\omega_{\eta j})}. \quad (34)$$

### 3.1.2. Cauchy Residue Theorem

The integrand of  $\theta_{\eta j}(t)$  is analytic except at the frequencies  $\omega_{\eta j}$ , and the integral is over the closed contour in **Figure 2**, Cauchy's residue theorem then implies

$$\oint \frac{d\omega}{2\pi} e^{-i\omega t} \frac{r_{\eta j}}{\omega - \omega_{\eta j}} = -i \sum_{\omega_{\eta j}} \text{Res} \left[ \frac{r_{\eta j} e^{-i\omega t}}{\omega - \omega_{\eta j}} \right], \quad (35)$$

where Res denotes the residue of the function in brackets in Equation (35) and the sum is over all  $n$  poles.

For nondegenerate modes, there are only simple poles at  $\omega = \omega_{\eta j}$ , and the residue is

$$\lim_{\omega \rightarrow \omega_{\eta j}} (\omega - \omega_{\eta j}) \frac{r_{\eta j} e^{-i\omega t}}{\omega - \omega_{\eta j}} = r_{\eta j} e^{-i\omega_{\eta j} t}. \quad (36)$$

Remember that we have been setting  $t_0 = 0$ , and writing  $t$  instead of  $t - t_0$  when evaluating the temporal part of the ERP. Thus, the general form of the temporal part of the ERP for arbitrary  $t_0$  is

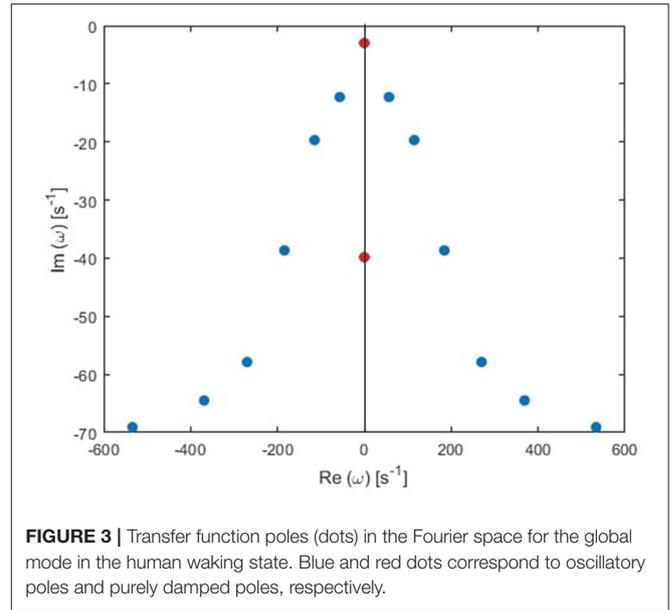
$$\theta_{\eta j}(t) \approx -i \sum_{j=1}^n r_{\eta j} e^{-i\omega_{\eta j}(t-t_0)} = -i \sum_{j=1}^n r_{\eta j} e^{-i\Omega_{\eta j}(t-t_0)} e^{-i\gamma_{\eta j} t}, \quad (37)$$

for  $t \geq t_0$ , and  $\theta_{\eta j}(t) = 0$  for  $t < t_0$  to ensure causality, where  $\omega_{\eta j}$ ,  $\Omega_{\eta j}$ , and  $\gamma_{\eta j}$  are defined in section 2.2.

### 3.1.3. Pairing Up Poles

In general, the poles of a transfer function can be grouped in pairs, such that each pair generates a real response in the time domain (Babaie-Janvier and Robinson, 2018). There are two types of pairs: the first type has two poles with the same damping rate ( $\gamma_{\eta j}$ ) and equal but opposite natural frequency ( $\Omega_{\eta j}$ ), and the second type consists of two pure imaginary poles; paired for mathematical convenience. For simplicity, we call the first type oscillatory poles, and the second type purely damped, as illustrated in **Figure 3**. Accordingly, the temporal part of the ERP is

$$\theta_{\eta}(t) = \sum_O I_{\eta O}(t) + \sum_D Q_{\eta D}(t), \quad (38)$$



**FIGURE 3** | Transfer function poles (dots) in the Fourier space for the global mode in the human waking state. Blue and red dots correspond to oscillatory poles and purely damped poles, respectively.

where  $I_{\eta O}(t)$  is the sum of  $\theta_{\eta}(t)$  over a single pair of oscillatory poles labeled  $o$  and  $o'$  such that  $O = [o, o']$ , and  $Q_{\eta D}(t)$  to a pair of purely damped poles labeled  $d$  and  $d'$  such that  $D = [d, d']$ . Note that the total number of poles  $n$  should be even in order to group them in pairs in our derivation. In the case when we retain an odd number of poles all the poles are paired except one purely damped pole.

To calculate  $I_{\eta O}(t)$ , we apply the reality condition (Howell, 2001; Babaie-Janvier and Robinson, 2018)

$$\omega_{\eta o'} = -\omega_{\eta o}^*, \quad (39)$$

and get

$$\omega_{\eta o} = \Omega_{\eta o} - i\gamma_{\eta o}, \quad (40)$$

$$\omega_{\eta o'} = -\Omega_{\eta o} - i\gamma_{\eta o}. \quad (41)$$

Then

$$I_{\eta O}(t) = -ie^{-\gamma_{\eta o} t} [r_{\eta o} e^{-i\Omega_{\eta o} t} + r_{\eta o'} e^{i\Omega_{\eta o} t}]. \quad (42)$$

Also, we have

$$r_{\eta o'} = -r_{\eta o}^*, \quad (43)$$

and can write

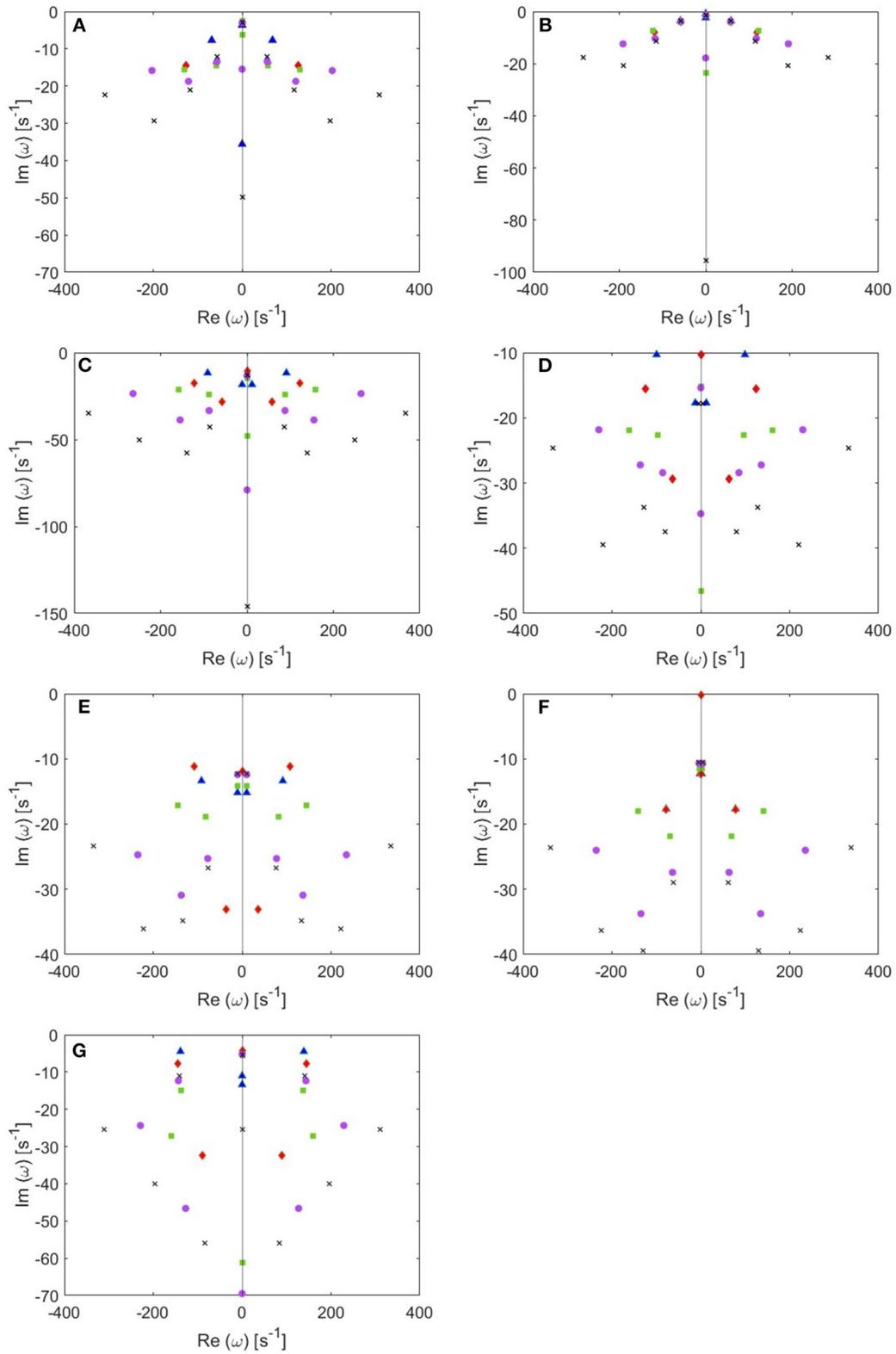
$$r_{\eta o} = |r_{\eta o}| e^{-i\psi_{\eta o}}, \quad (44)$$

$$\tan(\psi_{\eta o}) = \frac{\text{Im}(r_{\eta o})}{\text{Re}(r_{\eta o})}, \quad (45)$$

where  $\psi_{\eta o}$  is the complex argument of  $r_{\eta o}$ , so

$$I_{\eta O}(t) = ie^{-\gamma_{\eta o} t} |r_{\eta o}| \left[ e^{i(\psi_{\eta o} + \Omega_{\eta o} t)} - e^{-i(\psi_{\eta o} + \Omega_{\eta o} t)} \right], \quad (46)$$

$$= -2|r_{\eta o}| e^{-\gamma_{\eta o} t} \sin(\Omega_{\eta o} t + \psi_{\eta o}), \quad (47)$$



**FIGURE 4 |** Location of poles for 4-pole (blue triangles), 5-pole (red rhombuses), 6-pole (green squares), 8-pole (purple circles), and 10-pole (black crosses) approximations to the transfer function. **(A)** EO state, **(B)** EC state, **(C)** REM state, **(D)** S1 state, **(E)** S2 state, **(F)** SWS state, and **(G)** Spindles state.

which is a real response, as required.

Now consider the case of purely damped modes with

$$\Omega_{\eta j} = 0.$$

Then

$$\omega_{\eta d} = -i\gamma_{\eta d}, \tag{48}$$

$$\omega_{\eta d'} = -i\gamma_{\eta d'}. \tag{49}$$

So

$$Q_{\eta D}(t) = -i(r_{\eta d}e^{-\gamma_{\eta d}t} + r_{\eta d'}e^{-\gamma_{\eta d'}t}). \tag{50}$$

Note that  $r_{\eta d}$  and  $r_{\eta d'}$  are pure imaginary here, so

$$Q_{\eta D}(t) = \text{Im}(r_{\eta d})e^{-\gamma_{\eta d}t} + \text{Im}(r_{\eta d'})e^{-\gamma_{\eta d'}t}, \tag{51}$$

which is also a real response. Finally, by combining the above results, we find

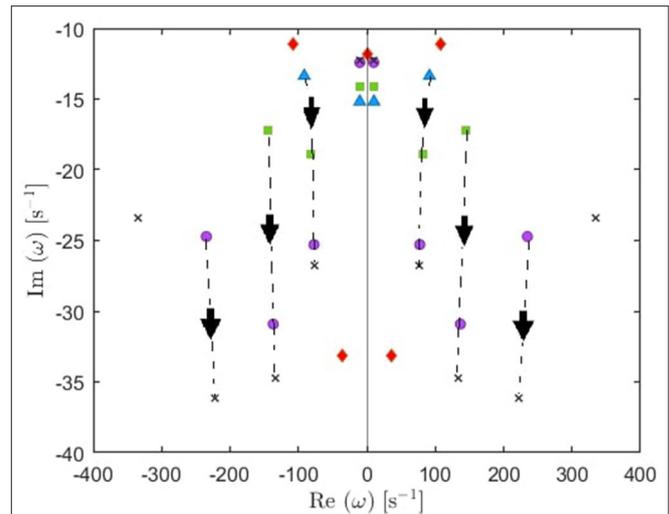
$$\text{ERP}(\mathbf{r}, \mathbf{r}_0, t) = \sum_{\eta} u_{\eta}(\mathbf{r})u_{\eta}^*(\mathbf{r}_0) \left[ \sum_O I_{\eta O}(t) + \sum_D Q_{\eta D}(t) \right]. \tag{52}$$

This equation expresses the ERP via a modal-polar representation of the transfer function. Each term has a spatial part  $u_{\eta}(\mathbf{r})u_{\eta}^*(\mathbf{r}_0)$  and a temporal factor in the square brackets. Remember that we have been considering an even number of oscillatory and purely damped poles in the temporal part.

### 3.2. Pole Locations

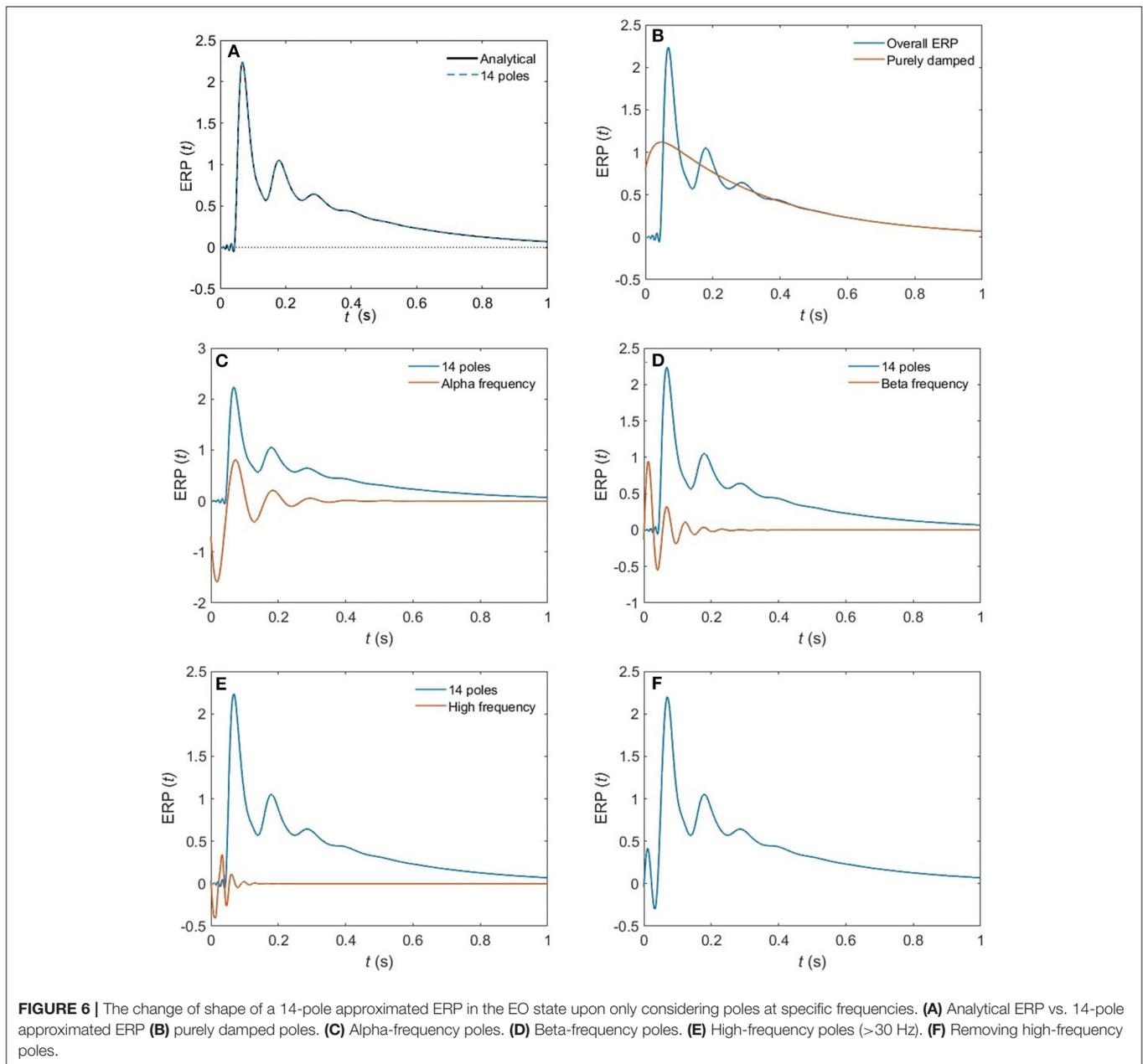
The locations of poles for several cases of the total number of poles are shown in **Figure 4**. Remember that the eigenfrequencies are of the form  $\omega = \Omega - i\gamma$ , where  $\Omega$  corresponds to the angular frequency and  $\gamma$  to the damping rate. Our main results are:

- (i) From these plots we figure out that as more poles are added, a given root usually moves downwards and approaches a limiting location (arrows in **Figure 5**), adding at most a few weakly damped poles; adding more poles in the analysis adds only strongly damped poles, including some at increasingly high frequencies. However, these poles contribute little to the overall dynamics in which the least damped roots are the dominant resonances. Because just a few weakly damped poles dominate the dynamics, addition of further poles only modifies the behavior slightly. Note that a very few poles move slightly upwards when we add more poles, as what happened to the beta frequency poles when changing from 4 to 5-pole approximation, but the general trend is moving downwards. Also, in **Figure 5**, we notice that some of the purely damped poles split into two oscillatory poles (red low frequency purely damped pole in a 5-pole approximation splits into two green oscillatory poles after changing to 6-pole approximation) and vice versa (two blue low frequency oscillatory poles in a 4-pole approximation add together to become one red purely damped pole after changing to 5-pole approximation).



**FIGURE 5** | Enlarged view of **Figure 4E** showing how poles move as we add more poles to the approximation. Starting with 4 poles (blue triangles), we track (by arrows) how the oscillatory poles in the right and left of the figure symmetrically move when changing to 5-pole approximation (red rhombuses), 6-pole approximation (green squares), 8-pole approximation (purple circles), and 10-pole approximation (black crosses).

- (ii) In the EO (**Figure 4A**) and EC (**Figure 4B**) waking states, the roots corresponding to alpha frequency shift slightly downwards as we add more poles, unlike the remaining sleep states (**Figures 4C–F**) that show a larger shift of the alpha roots as we add more poles. This is due to the need to reproduce the prominent alpha peak in the wake states, especially the EC state. We verify that by plotting the locations of poles of the Spindle condition in **Figure 4G** which shows that the poles corresponding to the spindle peak (at about 14 Hz) also stabilize at a limiting location like what happened in the waking states near their alpha and beta peak frequencies. This result is expected because it is well known in spectral analysis that whenever a peak exists, the damping rate is low (inversely proportional). Therefore, the distribution of poles (**Figure 4**) is a characteristic of the arousal state.
- (iii) Except for the least damped poles, the roots move rapidly toward large damping for the other sleep states as we increase the number of poles in the approximation, especially in the REM and S1 states. This outcome is predictable because REM and S1 states are classified as being the lightest sleep states, and their smooth spectra are fairly featureless, lack sharp peaks, and flatten at low frequencies. Therefore, these states only need the lowest  $\omega = 0$  poles to show their main dynamics.
- (iv) These poles enable us to plot ERPs for the different arousal states. **Figure 6** shows the shape of a 14-pole approximated ERP function upon only considering poles at specific frequencies in the EO state. In **Figure 6A**, we show the close match between analytical ERP and the 14-pole approximated ERP. The 14-pole approximated ERP



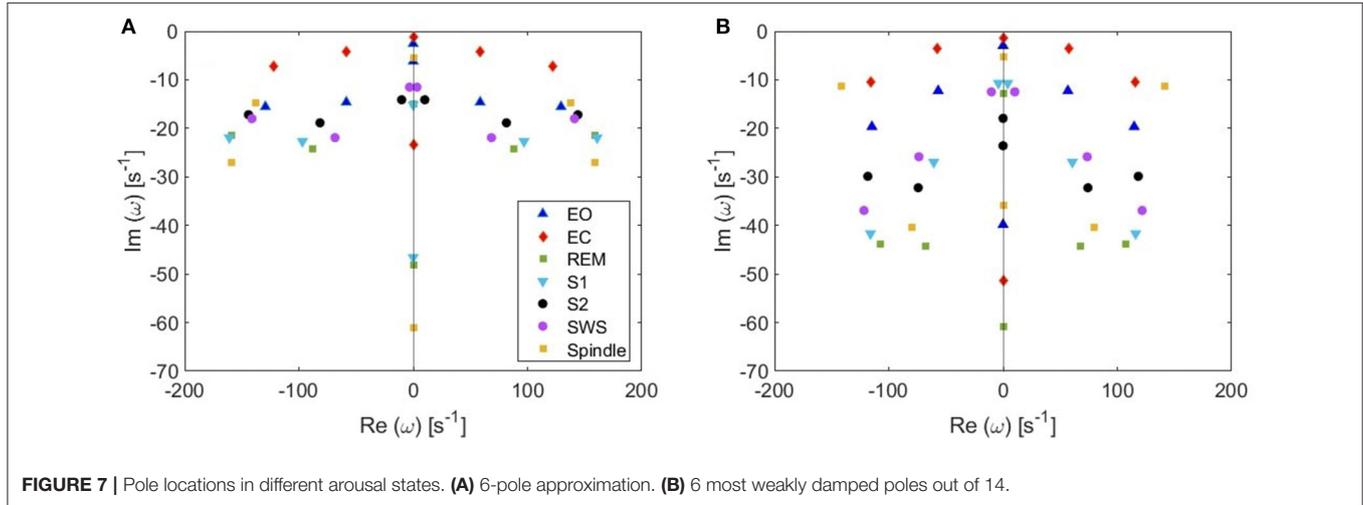
starts from zero with the appearance of very short and fast ripples before the stimulus hits the cortex (at  $\tau_{es} = 0.02$  ms), which then disappear and a sudden peak of the ERP takes place followed by consecutive smaller peaks decreasing in magnitude ending with an asymptotic decay toward zero after around 0.5 s. The ERP obtained by considering only purely damped poles ( $\Omega_{nj} = 0$ ), shown in **Figure 6B**, captures the asymptotic tail of the overall ERP and starts from a nonzero positive component, whereas the contribution of the alpha pole is negative (**Figure 6C**) at first. The ERPs formed by both the beta and high frequency poles (> 30 Hz) show small peaks before the stimulus

reaches the cortex (**Figures 6D,E**), and these small peaks disappear from the ERP if we remove the high frequency poles (**Figure 6F**). Therefore, when we add all poles together we get the ERP with very small residual ripples. This is analogous to ripples near a step change using finite number of Fourier coefficients. In real ERP analysis this is insignificant because there is always noise in the data.

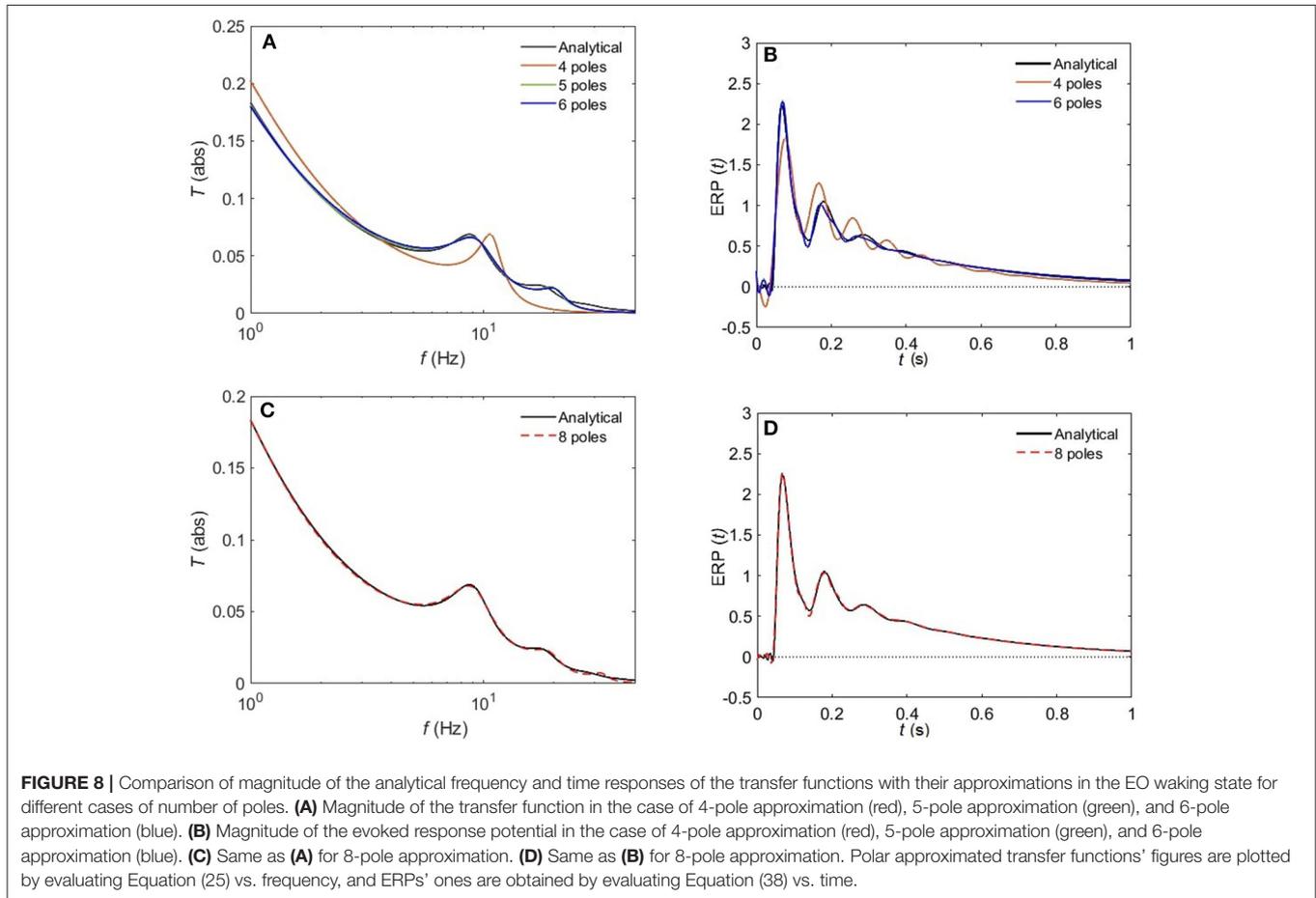
- (v) Because from **Figure 5** we deduce that the most weakly damped poles dominate, we plot in **Figure 7** the location of poles in a 6-pole approximation (**Figure 7A**), and the remaining 6 poles out of 14-pole approximation (**Figure 7B**) to remove highly damped poles across all

the states of arousal. Both figures look similar for the EO and EC waking states except for small shifts in the frequencies of the poles. However, in the other sleep states, poles have stronger damping in **Figure 7B** than **Figure 7A** especially in the REM and S1 states. In general, the two figures are consistent with each other which

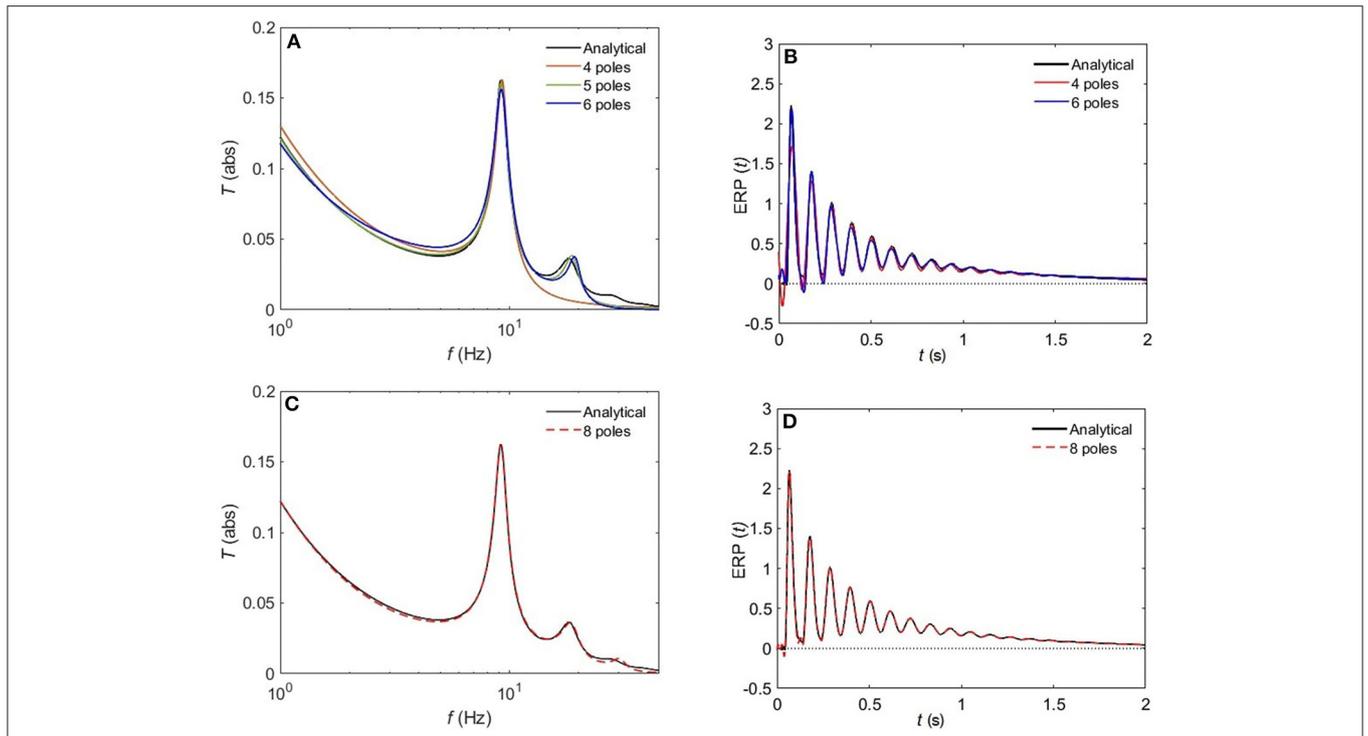
verifies that the three most weakly damped pairs of poles dominate, and correspond to the slow-wave, alpha, and beta resonances. Note that the 6-pole approximation figure is better (**Figure 7A**), and it can also be used as a reference to examine the shift between sleep states according to their poles' locations.



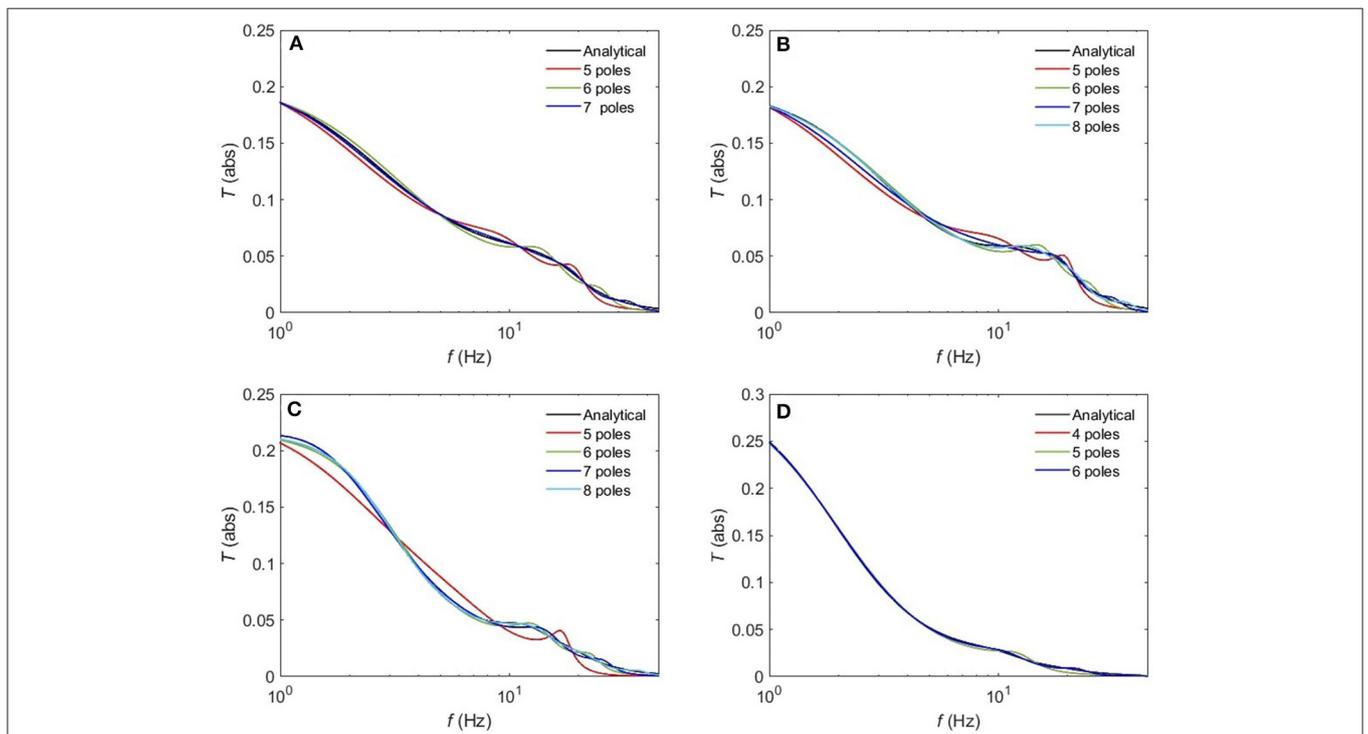
**FIGURE 7** | Pole locations in different arousal states. **(A)** 6-pole approximation. **(B)** 6 most weakly damped poles out of 14.



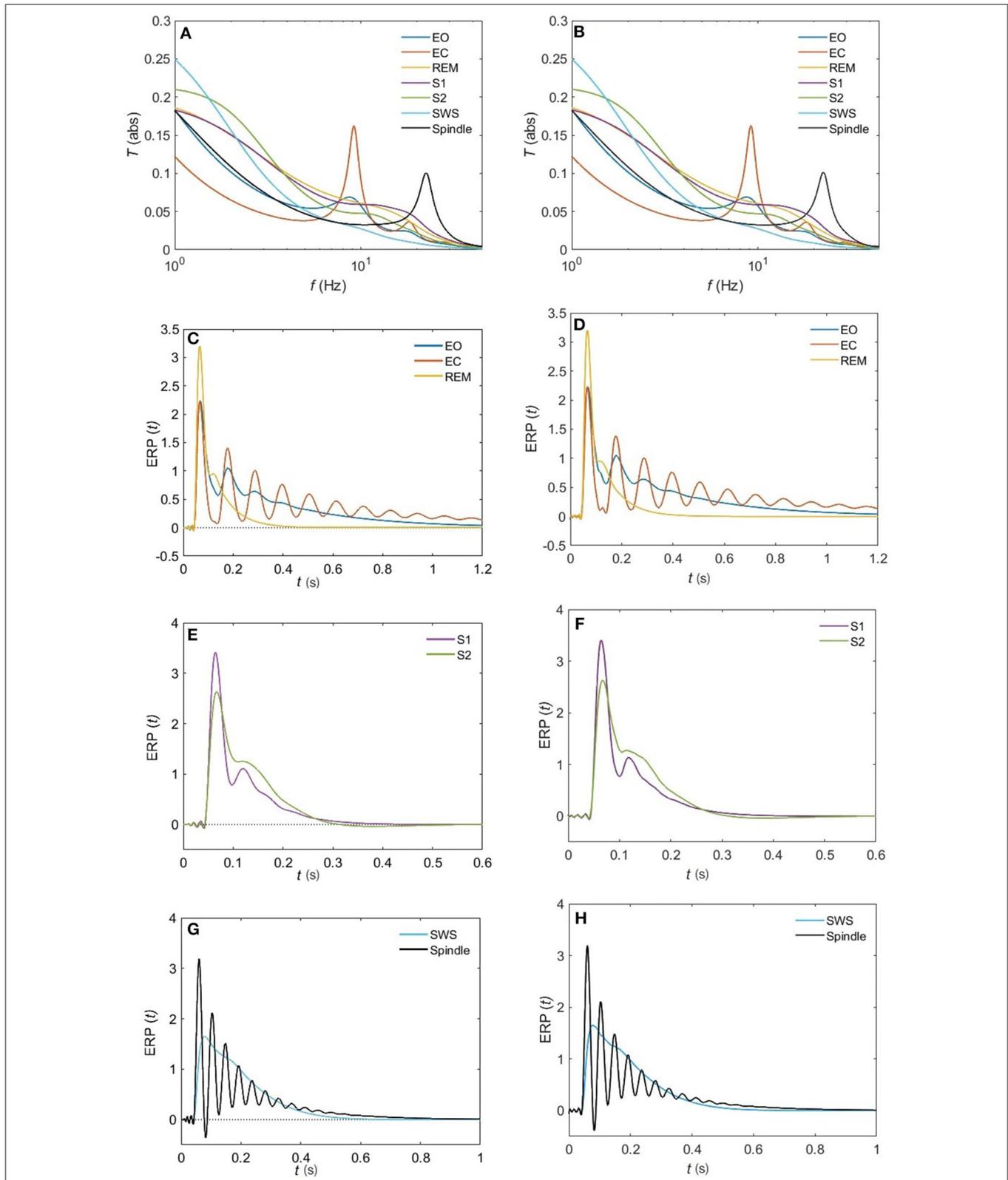
**FIGURE 8** | Comparison of magnitude of the analytical frequency and time responses of the transfer functions with their approximations in the EO waking state for different cases of number of poles. **(A)** Magnitude of the transfer function in the case of 4-pole approximation (red), 5-pole approximation (green), and 6-pole approximation (blue). **(B)** Magnitude of the evoked response potential in the case of 4-pole approximation (red), 5-pole approximation (green), and 6-pole approximation (blue). **(C)** Same as **(A)** for 8-pole approximation. **(D)** Same as **(B)** for 8-pole approximation. Polar approximated transfer functions' figures are plotted by evaluating Equation (25) vs. frequency, and ERPs' ones are obtained by evaluating Equation (38) vs. time.



**FIGURE 9 |** Comparison of magnitude of the analytical frequency and time responses of the transfer functions with their approximations in the EC waking state for different cases of number of poles. **(A)** Magnitude of the transfer function in the case of 4-pole approximation (red) and 5-pole approximation (green), and 6-pole approximation (blue). **(B)** Magnitude of the evoked response potential in the case of 4-pole approximation (red), 5-pole approximation (green), and 6-pole approximation (blue). **(C)** same as **(A)** for 8-pole approximation. **(D)** same as **(B)** for 8-pole approximation.



**FIGURE 10 |** Comparison of the magnitude of the analytical transfer function with its polar approximation for sleep states. **(A)** REM, **(B)** S1, **(C)** S2, and **(D)** SWS.



**FIGURE 11 |** Comparison of the magnitude of the analytical transfer functions and ERPs with their 10-pole approximations for all the states of arousal. **(A)** Analytical transfer function. **(B)** 10-pole approximated transfer function. **(C)** Analytical ERP for the EO, EC, and REM states. **(D)** 10-pole approximated ERP for the EO, EC, and REM states. **(E)** Analytical ERP for the S1 and S2 states. **(F)** 10-pole approximated ERP for the S1 and S2 states. **(G)** Analytical ERP for the SWS and Spindle states. **(H)** 10-pole approximated ERP for the SWS and Spindle states.

### 3.3. Accuracy vs. Number of Poles

The model parameters identifying brain states representing each of the Rechtschaffen and Kales sleep states are shown in **Table 2**, where we use them to evaluate the model transfer function. Here we apply our modal-polar approximation in frequency space to the transfer function corresponding to each arousal state separately in order to test our model in multiple arousal states. To check the accuracy of our approximation, we calculate the normalized root mean square error (rms) according to

$$\epsilon = \int \frac{\sqrt{(|T| - |T_{app}|)^2}}{\sqrt{|T|^2}} d\omega. \quad (53)$$

In the above rms error formula, we are considering the magnitude difference between the analytical and approximated transfer functions. The phase shift will not lead to much of a difference in ERPs because (at least in part) the magnitude of the high frequency components is very small. This is because the damping in the system attenuates high frequencies so the system only responds weakly to rapidly changing inputs. We could also use the Akaike information criterion (AIC) to choose the best fit on the basis of balancing accuracy against model complexity (i.e., number of poles in the present context) (Bozdogan, 1987). However, in the present work our aim is to achieve a given percentage accuracy, whereas AIC might trade off a poorer accuracy in favor of having fewer parameters.

**Figure 8** compares the approximated and analytical transfer functions (left) and their corresponding ERPs (right). In the alert EO state, as shown in **Figure 8A**, the analytical transfer function is characterized by possessing two consecutive peaks. The first peak corresponds to the alpha frequency (8.7 Hz), and the second one to the beta frequency (17.9 Hz). A 4-pole approximation of the EO transfer function yields to a large shift between the exact and approximated alpha peak in addition to missing the beta contribution. As we increase the number of poles to five, we observe a refined result, preserving the features of the transfer function with  $\epsilon \approx 4\%$ . As we keep on adding poles to the approximation the results become more accurate, with about 1% error achieved by using eight poles (**Figure 8C**). In **Figure 9**, we repeat the same strategy for the EC waking state which is characterized by a prominent alpha peak followed by a beta peak. Similarly, four poles are insufficient to represent the corresponding transfer function accurately. However, five poles yield about 5% error, and  $\epsilon < 2\%$  for an 8-pole approximation.

The normalized rms error  $\epsilon$  is worse for the sleep states (REM, S1, S2) compared to the waking states upon using a 5-pole approximation, with only about 10% error. However, **Figure 10** reveals that a 6-pole approximation reproduces acceptable results for the REM, S1 and S2 states ( $\epsilon < 5\%$ ); but, we notice that there is a weak additional peak near the beta frequency which is absent in the corresponding analytical transfer functions. Adding one more pole cancels that extra peak in the REM state and leads to further enhancement of the transfer function with about 2% error, whereas, for the S1 and S2 states, two more poles are needed to attain a better resolution, with 2 and 0.8% error, respectively. For the deepest sleep state SWS in **Figure 10D**, the 4-pole approximation shows a very close match with the analytical SWS

transfer function, with about 4% error. By increasing the number of poles to six,  $< 2\%$  error is achieved. **Figure 11** shows a close match ( $\epsilon < 1\%$ ) between the analytical and polar representation of both the transfer function and ERPs for all arousal states in a 10-pole approximation.

The normalized rms error values corresponding to each sleep state shown in **Table 3** are illustrated in **Figure 12**, where we observe how these values vary with the number of poles. We find that in the wake states, a 5-pole approximation is better than 6-pole approximation; this is due to the fact that the location of poles in a 5-pole approximation are approximately the same as for a 6-pole approximation (**Figures 4A,B**) except for the low frequency root which splits into two roots in the 6-pole approximation, reducing accuracy at low frequencies and increasing it at high frequencies (**Figures 4A,B**). In **Figure 12**, the rms error values  $\epsilon$  of the magnitude of the transfer function vs. number of poles is shown for the different arousal states. In order to obtain a general number of poles to be used when studying any arousal state, we plot the worst cases of the maximum values of  $\epsilon$  vs. the number of poles in **Figure 12H** across all the arousal states. We find that six poles are sufficient to attain better than 8% error, thereby preserving the main dynamics of the system.

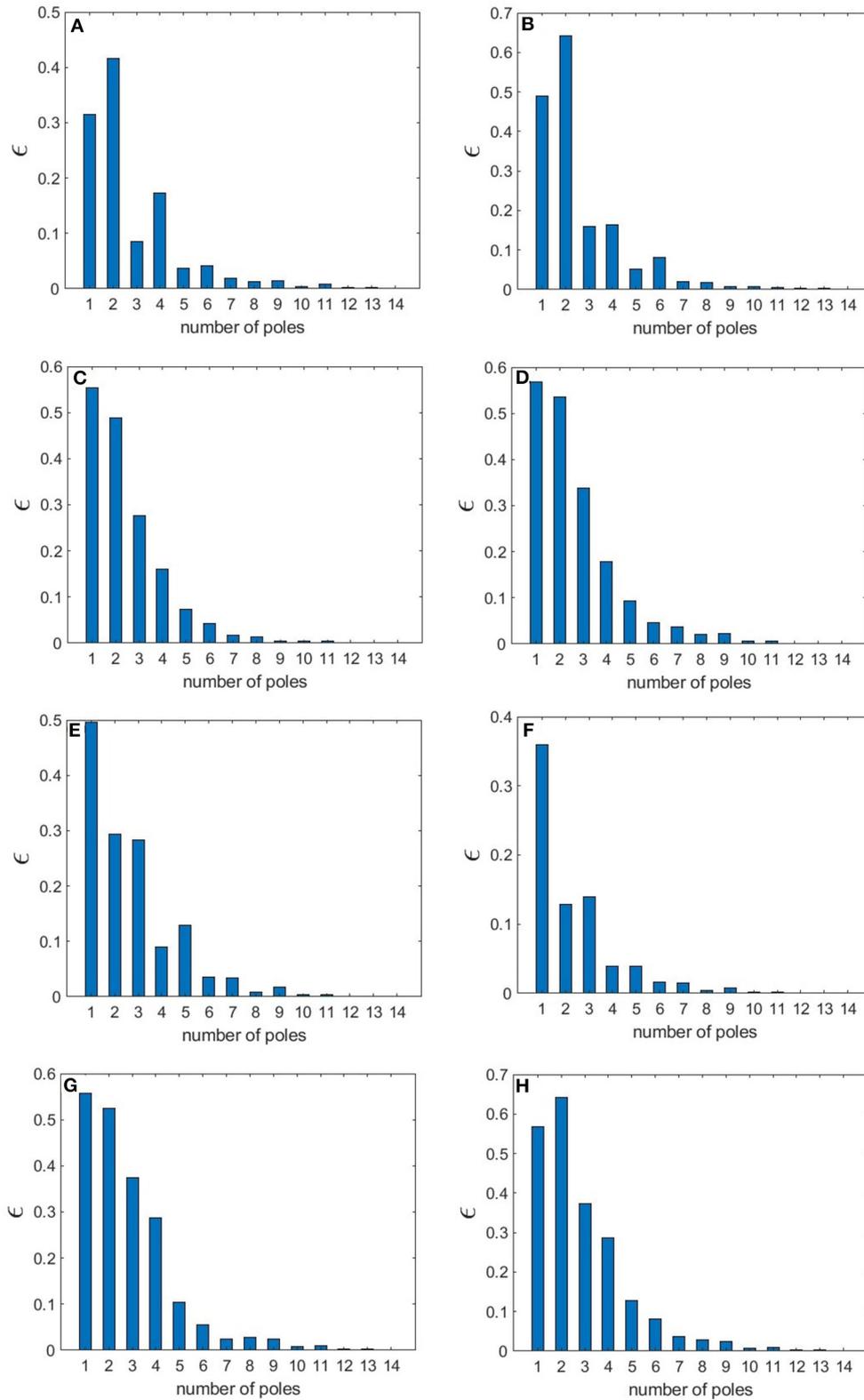
## 4. SUMMARY AND DISCUSSION

We have bridged between physiologically-based NFT and experimental observations of brain resonances by decomposing the corticothalamic transfer function into spatial (eigenmodes) and temporal (poles) components. Modal-polar representation of NFT provides a unified platform to investigate a formula for ERPs and explore the number of poles needed to study the main dynamics of the system in each arousal state. Our main results are as follows:

**TABLE 3 |** Root mean square error  $\epsilon$  (in%) of the magnitude of the transfer function for each arousal state, where  $n$  is the number of poles.

$n$	EO	EC	REM	S1	S2	SWS	Spindle
1	32	49	55	57	50	36	57
2	42	64	49	54	29	13	52
3	9	16	28	34	28	14	0.37
4	17	16	16	18	8.9	3.9	2.9
5	3.7	5.1	7.3	9.4	13	3.9	10
6	4.2	8.2	4.2	4.6	3.6	1.7	5.5
7	2.0	1.9	1.8	3.7	3.4	1.5	2.4
8	1.3	1.9	1.4	2.0	0.85	0.42	2.8
9	1.5	0.67	0.41	2.2	1.7	0.74	2.4
10	0.38	0.70	0.40	0.62	0.42	0.20	0.80
11	0.78	0.59	0.39	0.58	0.40	0.19	0.93
12	0.18	0.22	0.10	0.09	0.06	0.04	0.24
13	0.21	0.22	0.08	0.09	0.06	0.04	0.23
14	0.09	0.17	0.06	0.02	0.01	0.02	0.11

*This table corresponds to the data illustrated in Figure 12.*



**FIGURE 12 |** Root mean square error of the magnitude of the transfer function for different number of poles for various states. **(A)** EO state **(B)** EC state **(C)** REM state **(D)** S1 state **(E)** S2 state **(F)** SWS state **(G)** S2-Spindle state. In **(H)**, the highest  $\epsilon$  value across all the states is shown vs.  $n$ .

- (i) An approximation of the transfer function that describes the linear response of cortical brain activity to any input was achieved that allows it to be written as a ratio of polynomials. For each eigenmode, a pole of the transfer function corresponds to a given eigenfrequency or resonance. We restricted detailed analysis to the global mode ( $k_{\eta}^2 = 0$ ), because it has uniform spatial properties across the two brain hemispheres. As a first approach, we applied the modal-polar expansion of the transfer function to derive a simpler formula for the ERP which is otherwise not analytically tractable. The calculations involved complex integrals that were solved via contour integration and the Cauchy-residue theorem.
- (ii) We validated the new expression of the ERP by comparing it with the exact ERP (calculated from NFT) for various states of arousal.
- (iii) We investigated how the poles for each arousal state shift as the approximation order increases and becomes more accurate. We found that as the approximation order increases, most poles move downwards in the complex plane, corresponding to these eigenfrequencies becoming more damped. For waking states, the least damped poles are mostly robust to these changes and adding more poles to the approximation involves shifts in the strongly damped poles, which contribute little to the overall dynamics. In these waking states, the roots corresponding to the alpha resonance shift slightly as we add more poles to the approximation, unlike the other sleep states. This is due to the significant alpha peak in the wake states, especially the EC which is characterized by its prominent alpha peak. For sleep states, especially REM and S1, we observed larger downward shifts in the poles toward more strongly damped eigenfrequencies, reflecting the smooth almost featureless spectra of these states.
- (iv) We found that six roots or three pairs of eigenfrequencies (poles) suffice to preserve the main dynamics of all arousal states to within  $\approx 4\%$  error. These roots correspond to the three resonances (low frequencies, alpha, and beta) and have been found to resemble a response expressed in terms of PID filters via control theory which is analogous to using a group of controllers in order to enhance the control system performance (Babaie-Janvier and Robinson, 2018).

Overall, this framework lays foundation for simplifying NFT results, and connecting NFT with observations of brain resonances, which in turn leads to an effective simplification of analysis of ERPs by expressing exact ERPs (based on NFT) in terms of those resonances. The present work has expanded previous work that approximated the transfer function and calculated poles for the waking state only (Gabay et al., 2018). Hence, this links observable electrophysiological resonances more directly to underlying dynamics and function, and yields interpretations of EEG resonances in terms of PID filters via control theory that provides insights on cognitive processes. The present analysis was restricted to the global mode, but future work could investigate the multimodal case as well as calculate other quantities from the transfer function such as coherence, correlation functions, and plasticity.

## DATA AVAILABILITY STATEMENT

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

## AUTHOR CONTRIBUTIONS

PR conceived the research. PR and RE-Z developed the theory, verified the analytical methods, and along with NG discussed all the results. NG wrote the computer code that was used and RE-Z added to this code, wrote the majority of the manuscript, and performed all the mathematical derivations. All authors contributed to the article and approved the submitted version.

## FUNDING

This work was supported by the Australian Research Council under Center of Excellence grant no. CE140100007 and Laureate Fellowship grant no. FL140100025.

## ACKNOWLEDGMENTS

We thank Dr. James Henderson and Dr. Ben Fulcher for helpful discussions.

## REFERENCES

- Abeyuriya, R., Rennie, C., and Robinson, P. (2014). Prediction and verification of nonlinear sleep spindle harmonic oscillations. *J. Theor. Biol.* 344, 70–77. doi: 10.1016/j.jtbi.2013.11.013
- Abeyuriya, R. G., Rennie, C. J., and Robinson, P. A. (2015). Physiologically based arousal state estimation and dynamics. *J. Neurosci. Methods* 253, 55–69. doi: 10.1016/j.jneumeth.2015.06.002
- Babaie-Janvier, T., and Robinson, P. A. (2018). Neural field theory of corticothalamic prediction with control systems analysis. (report). *Front. Hum. Neurosci.* 12:334. doi: 10.3389/fnhum.2018.00334
- Babaie-Janvier, T., and Robinson, P. A. (2019). Neural field theory of corticothalamic attention with control system analysis. *Front. Neurosci.* 13:1240. doi: 10.3389/fnins.2019.01240
- Babaie-Janvier, T., and Robinson, P. A. (2020). Neural field theory of evoked response potentials with attentional gain dynamics. *Front. Hum. Neurosci.* 14:293. doi: 10.3389/fnhum.2020.00293
- Bozdogan, H. (1987). Model selection and akaike's information criterion (aic): the general theory and its analytical extensions. *Psychometrika* 52, 345–370. doi: 10.1007/BF02294361
- Braitenberg, V., and Schüz, A. (1998). *Cortex: Statistics and Geometry of Neuronal Connectivity*. Berlin: Springer.
- Buzsáki, G., and Draguhn, A. (2004). Neuronal oscillations in cortical networks. *Science* 304, 1926–1929. doi: 10.1126/science.1099745
- Chiang, A. K. I., Rennie, C. J., Robinson, P. A., Roberts, J. A., Rigozzi, M. K., Whitehouse, R. W., et al. (2008). Automated characterization of multiple alpha peaks in multi-site electroencephalograms. *J. Neurosci. Methods* 168, 396–411. doi: 10.1016/j.jneumeth.2007.11.001

- D’Rozario, A., Kim, J., Wong, K., Bartlett, D., Marshall, N., Dijk, D.-J., et al. (2013). A new EEG biomarker of neurobehavioral impairment and sleepiness in sleep apnea patients and controls during extended wakefulness. *Clin. Neurophysiol.* 124, 1605–1614. doi: 10.1016/j.clinph.2013.02.022
- Feng, C., Wang, L., Liu, C., Zhu, X., Dai, R., Mai, X., et al. (2012). The time course of the influence of valence and arousal on the implicit processing of affective pictures. *PLoS ONE* 7:e29668. doi: 10.1371/journal.pone.0029668
- Gabay, N. C., Babiak-Janvier, T., and Robinson, P. A. (2018). Dynamics of cortical activity eigenmodes including standing, traveling, and rotating waves. *Phys. Rev. E* 98:042413. doi: 10.1103/PhysRevE.98.042413
- Gabay, N. C., and Robinson, P. A. (2017). Cortical geometry as a determinant of brain activity eigenmodes: Neural field analysis. *Phys. Rev. E* 96:032413. doi: 10.1103/PhysRevE.96.032413
- Golub, G. H., and Van Loan, C. F. (1983). *Matrix Computations / Gene H. Golub, Charles F. Van Loan*. Baltimore, MD: Johns Hopkins University Press Baltimore.
- Gordon, E., Cooper, N., Rennie, C. J., Hermens, D., and Williams, L. (2005). Integrative neuroscience: the role of a standardized database. *Clin. EEG Neurosci.* 36, 64–75. doi: 10.1177/155005940503600205
- Herrmann, C. S., and Knight, R. T. (2001). Mechanisms of human attention: event-related potentials and oscillations. *Neurosci. Biobeh. Rev.* 25, 465–476. doi: 10.1016/S0149-7634(01)00027-6
- Hillyard, S. A., and Anllo-Vento, L. (1998). Event-related brain potentials in the study of visual selective attention. *Proc. Natl. Acad. Sci. U.S.A.* 95, 781–787. doi: 10.1073/pnas.95.3.781
- Howell, K. (2001). *Principles of Fourier Analysis. Textbooks in Mathematics*. Florida: CRC Press.
- Iber, C., Ancoli-Israel, S., Chesson, A. L., and Quan, S. (2007). *The AASM Manual for the Scoring of Sleep and Associated Events: Rules, Terminology and Technical Specifications*. Westchester, IL: American Academy of Sleep Medicine.
- Kelly, J. T., and Reilly, E. L. (1983). *EEG and Evoked Potentials in Psychiatry and Behavioral Neurology*. Boston, MA: Butterworths.
- Kerr, C., Rennie, C., and Robinson, P. A. (2008). Physiology-based modeling of cortical auditory evoked potentials. *Biol. Cybern.* 98, 171–184. doi: 10.1007/s00422-007-0201-1
- Kwakernaak, H., and Sivan, R. (1972). *Linear Optimal Control Systems*, Vol. 1. New York, NY: Wiley-Interscience.
- Luck, S. J. (2014). *An Introduction to the Event-Related Potential Technique*, 2nd edn. Cambridge, MA: MIT Press.
- Luck, S. J., and Kappenman, E. S. (2013). *The Oxford Handbook of Event-Related Potential Components*. New York, NY: Oxford library of psychology. Oxford University Press.
- Mitrinović, D. S. (1984). *The Cauchy Method of Residues: Theory and Applications*. Dordrecht: D. Reidel.
- Niedermeyer, E., and Lopes Da Silva, F. (1999). *Electroencephalography: Basic Principles, Clinical Applications, and Related Fields*, 4th edn. Philadelphia, PA: Lippincott Williams & Wilkins.
- Ogata, K. (2010). *Modern Control Engineering*. New Jersey, NJ: Prentice hall.
- Picton, T. W., Bentin, S., Berg, P., Donchin, E., Hillyard, S. A., Johnson, R., et al. (2000). Guidelines for using human event-related potentials to study cognition: Recording standards and publication criteria. *Psychophysiology* 37, 127–152. doi: 10.1111/1469-8986.3720127
- Pinotsis, D. A., Hansen, E., Friston, K. J., and Jirsa, V. K. (2013). Anatomical connectivity and the resting state activity of large cortical networks. *NeuroImage* 65, 127–138. doi: 10.1016/j.neuroimage.2012.10.016
- Rechtschaffen, A., and Kales, A. (1968). *A Manual of Standardized Terminology, Techniques and Scoring System for Sleep Stages of Human Subjects*. Los Angeles, CA: Brain Information Service/Brain Research Institute.
- Rennie, C. J., Robinson, P. A., and Wright, J. J. (1999). Effects of local feedback on dispersion of electrical waves in the cerebral cortex. *Phys. Rev. E* 59, 3320–3329. doi: 10.1103/PhysRevE.59.3320
- Robinson, P. A. (2003). Neurophysical theory of coherence and correlations of electroencephalographic and electrocorticographic signals. *J. Theor. Biol.* 222, 163–175. doi: 10.1016/S0022-5193(03)00023-7
- Robinson, P. A. (2005). Propagator theory of brain dynamics. *Phys. Rev. E* 72:011904. doi: 10.1103/PhysRevE.72.011904
- Robinson, P. A. (2019). Physical brain connectomics. *Phys. Rev. E* 99:012421. doi: 10.1103/PhysRevE.99.012421
- Robinson, P. A., Loxley, P. N., O’Connor, S. C., and Rennie, C. J. (2001). Modal analysis of corticothalamic dynamics, electroencephalographic spectra, and evoked potentials. *Phys. Rev. E* 63:041909. doi: 10.1103/PhysRevE.63.041909
- Robinson, P. A., Pagès, J. C., Gabay, N. C., Babiak-Janvier, T., and Mukta, K. N. (2018). Neural field theory of perceptual echo and implications for estimating brain connectivity. *Phys. Rev. E* 97:042418. doi: 10.1103/PhysRevE.97.042418
- Robinson, P. A., Rennie, C. J., and Rowe, D. L. (2002). Dynamics of large-scale brain activity in normal arousal states and epileptic seizures. *Phys. Rev. E* 65(4 Pt 1):041924. doi: 10.1103/PhysRevE.65.041924
- Robinson, P. A., Rennie, C. J., Rowe, D. L., and O’Connor, S. C. (2004). Estimation of multiscale neurophysiologic parameters by electroencephalographic means. *Hum. Brain Mapp.* 23, 53–72. doi: 10.1002/hbm.20032
- Robinson, P. A., Rennie, C. J., Rowe, D. L., O’Connor, S. C., and Gordon, E. (2005). Multiscale brain modelling. *Philos. Trans. R. Soc. Ser. B Biol. Sci.* 360, 1043–1050. doi: 10.1098/rstb.2005.1638
- Robinson, P. A., Rennie, C. J., and Wright, J. J. (1997). Propagation and stability of waves of electrical activity in the cerebral cortex. *Phys. Rev. E* 56, 826–840. doi: 10.1103/PhysRevE.56.826
- Robinson, P. A., Zhao, X., Aquino, K. M., Griffiths, J. D., Sarkar, S., and Mehta-Pandey, G. (2016). Eigenmodes of brain activity: neural field theory predictions and comparison with experiment. *NeuroImage* 142, 79–98. doi: 10.1016/j.neuroimage.2016.04.050
- Van Albada, S. V., and Robinson, P. A. (2013). Relationships between electroencephalographic spectral peaks across frequency bands. *Front. Hum. Neurosci.* 7:56. doi: 10.3389/fnhum.2013.00056
- Varga, A. (1991). “Balancing free square-root algorithm for computing singular perturbation approximations,” in [1991] *Proceedings of the 30th IEEE Conference on Decision and Control* (Brighton), Vol. 2, 1062–1065.
- Vázquez Marrufo, M., Vaquero, E., Cardoso, M. J., and Gómez, C. M. (2001). Temporal evolution of  $\alpha$  and  $\beta$  bands during visual spatial attention. *Cogn. Brain Res.* 12, 315–320. doi: 10.1016/S0926-6410(01)00025-8
- Wang, D., Teichtahl, H., Drummer, O., Goodman, C., Cherry, G., Cunningham, D., et al. (2005). Central sleep apnea in stable methadone maintenance treatment patients. *Chest* 128, 1348–1356. doi: 10.1378/chest.128.3.1348
- Womelsdorf, T., Schoffelen, J. M., Oostenveld, R., Singer, W., Desimone, R., Engel, A. K., et al. (2007). Modulation of neuronal interactions through neuronal synchronization. *Science* 316, 1609–1612. doi: 10.1126/science.1139597
- Woodman, G. F. (2010). A brief introduction to the use of event-related potentials in studies of perception and attention. *Atten. Percept. Psychophys.* 72, 2031–2046. doi: 10.3758/BF03196680
- Wright, J. J., and Liley, D. T. J. (1996). Dynamics of the brain at global and microscopic scales: Neural networks and the EEG. *Behav. Brain Sci.* 19, 285–295. doi: 10.1017/S0140525X00042679
- Yamagishi, N., Callan, D. E., Goda, N., Anderson, S. J., Yoshida, Y., and Kawato, M. (2003). Attentional modulation of oscillatory activity in human visual cortex. *NeuroImage* 20, 98–113. doi: 10.1016/S1053-8119(03)00341-0
- Zobaer, M. S., Anderson, R. M., Kerr, C. C., Robinson, P. A., Wong, K. K. H., and D’Rozario, A. (2017). K-complexes, spindles, and erps as impulse responses: unification via neural field theory. *Biol. Cybern.* 111, 149–164. doi: 10.1007/s00422-017-0713-2

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

Copyright © 2021 El-Zghir, Gabay and Robinson. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.