

On Decomposing Stimulus and Response Waveforms in Event-Related Potentials Recordings

Gang Yin and Jun Zhang*

Abstract—Event-related potentials (ERPs) reflect the brain activities related to specific behavioral events, and are obtained by averaging across many trial repetitions with individual trials aligned to the onset of a specific event, e.g., the onset of stimulus (s-aligned) or the onset of the behavioral response (r-aligned). However, the s-aligned and r-aligned ERP waveforms do not purely reflect, respectively, underlying stimulus (S-) or response (R-) component waveform, due to their cross-contaminations in the recorded ERP waveforms. Zhang [*J. Neurosci. Methods*, 80, pp. 49–63, 1998] proposed an algorithm to recover the pure S-component waveform and the pure R-component waveform from the s-aligned and r-aligned ERP average waveforms—however, due to the nature of this inverse problem, a direct solution is sensitive to noise that disproportionately affects low-frequency components, hindering the practical implementation of this algorithm. Here, we apply the Wiener deconvolution technique to deal with noise in input data, and investigate a Tikhonov regularization approach to obtain a stable solution that is robust against variances in the sampling of reaction-time distribution (when number of trials is low). Our method is demonstrated using data from a Go/NoGo experiment about image classification and recognition.

Index Terms—Component waveform, event-related potentials (ERPs), null space, stimulus–response decomposition, Tikhonov regularization, Wiener deconvolution.

I. INTRODUCTION

IN event-related potentials (ERP) studies, ERP averages are typically calculated with individual trials (in an ensemble of many repetitions) aligned with respect to the onset of stimulus, as in the study of various ERP elements thought to be related to stimulus processing, such as P300 [1], [2] and N400 [3]. When behavioral response is required, ERP averages have also been calculated with individual trials aligned with respect to the onset of behavioral response, as in the study of error-related negativity (ERN) [4], [5], and lateralized readiness potential (LRP).

Manuscript received November 26, 2009; revised April 6, 2010 and July 19, 2010; accepted October 1, 2010. Date of publication October 28, 2010; date of current version May 18, 2011. This work was supported by the Air Force Office of Scientific Research (AFOSR) under Grant FA9550-06-0298 awarded to the University of Michigan. This work was completed while G. Yin was an exchange student at the University of Michigan supported in part by a fellowship from the China Scholarship Council and while J. Zhang was on assignment to AFOSR from the University of Michigan. Asterisk indicates corresponding author.

G. Yin was with the School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China. He is now with Sichuan Cancer Hospital, Chengdu 610041, China (e-mail: gangyin@umich.edu).

*J. Zhang is with the Department of Psychology, University of Michigan, Ann Arbor, MI 48109 USA (e-mail: Junz@umich.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TBME.2010.2090152

Therefore, as long as the reaction time (RT), defined as the time elapse between the stimulus onset and the behavioral response onset for each trial, is not uniform across the ensemble of trials and has a nonzero variance, the stimulus-aligned ERP average and the response-aligned ERP average will differ in their waveforms. The stimulus-aligned (response-aligned) ERP average does not contain purely those elements that are time-locked to stimulus or response onset, since both these component waveforms contribute towards the stimulus (response)-aligned ERP average waveforms, and are hence cross-contaminated therein.

To see this, let $f_s(t)$, $f_r(t)$ denote, respectively, the underlying stimulus-locked and response-locked component waveforms. Denote the RTs as t_i . Then the stimulus-aligned ERP average waveform $F_s(t)$ and its response-aligned waveform $F_r(t)$ are, respectively:

$$\begin{aligned} F_s(t) &= \frac{1}{N} \sum_i (f_s(t) + f_r(t - t_i) + \xi_i(t)) \\ &\approx f_s(t) + \frac{1}{N} \sum_i f_r(t) * \delta(t - t_i) \end{aligned} \quad (1)$$

$$\begin{aligned} F_r(t) &= \frac{1}{N} \sum_i (f_s(t + t_i) + f_r(t) + \eta_i(t)) \\ &\approx f_r(t) + \frac{1}{N} \sum_i f_s(t) * \delta(t + t_i) \end{aligned} \quad (2)$$

where ξ_i and η_i are white noise (measured with respect to stimulus onset and response onset) on trial i , N is total trial number, $\delta(t)$ is the Kronecker delta function, and $\xi_i(t) = \eta_i(t - t_i) \Leftrightarrow \xi_i(t + t_i) = \eta_i(t)$; they are averaged “out” during ensemble averaging, assuming the noise is “white.” But the summation terms in the expression of $F_s(t)$ and $F_r(t)$ do not vanish in general; they represent residual contaminations: $F_s(t) \neq f_s(t)$, $F_r(t) \neq f_r(t)$.

Zhang [6] proposed an algorithm that recovers the stimulus- and response-locked waveforms $f_s(t)$, $f_r(t)$ from stimulus-aligned and response-aligned ERP averages $F_s(t)$, $F_r(t)$, as long as the empirically obtained RT distribution is given. The algorithm (as follows) provided a unique solution that can be implemented, in theory, through either the Fourier transform method or an iterative procedure (for solving integral equations) that is guaranteed to converge. Zhang’s [6] method has recently been extended to separate N components in ERP recordings given N behavioral event times (especially the $N = 3$ case useful for many paradigms) [7]. The basic assumptions for these algorithms are that noise is stationary across trials, and that the stimulus- and response-locked waveforms have the same shapes

and amplitude across single trials (see Section V for a relaxation to allow latency jitter).

One of the challenges in applying Zhang's [6] method to real ERP datasets is the issue of input noise, especially when the trial number in an experiment is low. It will be shown below that because the power of the low-frequency Fourier components (in the limit, dc component) of the RT distribution is close to 1, making the recovery of low-frequency components an ill-posed problem, noise-reduction techniques are called for in implementing the method of Zhang [6]. Recently, Takeda *et al.* [8], using an essentially identical method, applied a high-pass filter to EEG recordings before the decomposition of the stimulus-locked and response-locked component waveforms. This, we believe, leads to an undue attenuation of low-frequency signal (slow waveform) while not resolving the instability problem in the recovered waveforms caused by a lack of consistent estimation of RT distribution (which is due to limited number of experimental trials).

In this paper, we apply two common techniques to deal with input noise in the decomposition algorithm of Zhang [6]. First, we consider Wiener deconvolution method for noise control. Wiener filter theory was developed to find an optimal filter that adapts to the SNR by estimating the power spectra of both the noise and the signal and then minimizing mean-square error (MSE) for signal recovery. Wiener theory for optimal noise filtering of biomedical signals was adopted by, for instance, Walter [9], modified by, for instance, Doyle [10], and widely applied to resolve deconvolution problems [11]. Here, in departure from the standard treatment, we have to consider a 2-D signal, arising from being locked to stimulus and to response, which are themselves coupled through the same RT distribution function. Therefore, specialized formulas need to be developed for the current application. The second approach for noise reduction we consider here is the Tikhonov regularization method for computing stabilized solutions to an ill-posed problem. The main challenge to be addressed is to determine the appropriate value of the regularization parameter. Below, we compare the Wiener deconvolution approach with Tikhonov regularization approach (including methods of determining the optimal regularization parameter) for noise control in our problem. We compare these algorithms on simulated single-trial data under different noise levels, different variances of RT distributions, and different amount of available trial numbers. Finally, we apply our noise control method to the decomposition of ERP components in a Go/NoGo experiment.

II. MATHEMATICAL ANALYSIS

A. S-R Decomposition: A Revisit to [6]

Suppose, we have the following experimental data: 1) the stimulus-aligned ERP average waveform denoted $F_s(t)$, 2) the response-aligned ERP average waveform $F_r(t)$, and 3) the distribution of RTs $g(t)$. The problem is how to recover the stimulus component waveform time-locked to stimulus onset ("S-component") $f_s(t)$ and the response component waveform time-locked to response onset ("R-component") $f_r(t)$. From (1) and (2), and casting them in the limiting case of continuous time

with infinite trials, we can derive the following two mathematical equations:

$$\begin{aligned} F_s(t) &= f_s(t) + \frac{1}{N} \sum_{i=1}^N f_r(t) * \delta(t - t_i) \\ &= f_s(t) + f_r(t) * g(t) \end{aligned} \quad (3)$$

$$\begin{aligned} F_r(t) &= f_r(t) + \frac{1}{N} \sum_{i=1}^N f_s(t) * \delta(t + t_i) \\ &= f_r(t) + f_s(t) * g(-t) \end{aligned} \quad (4)$$

where N is the total number of trials. We define $g(t) = \frac{1}{N} \sum_{i=1}^N \delta(t - t_i)$, with its Fourier transform

$$\tilde{g}(w) = \int g(t) e^{-iwt} dt$$

where "~~" denotes frequency-domain representation. Solving for the frequency components $\tilde{f}_s(w)$, $\tilde{f}_r(w)$, and then applying inverse-Fourier transform, Zhang [6] obtained a closed-form solution

$$f_s(t) = \frac{1}{2\pi} \int \frac{\tilde{F}_s(w) - \tilde{F}_r(w) \cdot \tilde{g}(w)}{1 - |\tilde{g}(w)|^2} e^{iwt} dw \quad (5)$$

$$f_r(t) = \frac{1}{2\pi} \int \frac{\tilde{F}_r(w) - \tilde{F}_s(w) \cdot \tilde{g}(-w)}{1 - |\tilde{g}(w)|^2} e^{iwt} dw. \quad (6)$$

B. Methodological Limitation

In subsequent discussions, RT distribution $g(t)$ is estimated using the empirical distribution (frequency count of proportion of trials associated with a particular RT value t as measured from a fixed stimulus onset moment, denoted as $t = 0$, to the moment of behavioral response on individual trials). It is convenient to work in the frequency (as opposed to time) domain. Denote the matrix $\mathbf{H}(w)$ as follows:

$$\mathbf{H}(w) = \begin{bmatrix} 1 & \tilde{g}(w) \\ \tilde{g}(-w) & 1 \end{bmatrix} \quad (7)$$

with which (3) and (4) can now be converted to matrix form

$$\mathbf{y} = \mathbf{H}\mathbf{x}; \quad \mathbf{y} = \begin{bmatrix} \tilde{F}_s(w) \\ \tilde{F}_r(w) \end{bmatrix} \quad \text{and} \quad \mathbf{x} = \begin{bmatrix} \tilde{f}_s(w) \\ \tilde{f}_r(w) \end{bmatrix}. \quad (8)$$

Note that the complex conjugate (denoted by C) of $g(w)$ is

$$\tilde{g}^C(w) = \int g(t) e^{iwt} dt = \int g(t) e^{-i(-w)t} dt = \tilde{g}(-w).$$

This means $|\tilde{g}(w)| = |\tilde{g}(-w)|$, which implies that the \mathbf{H} matrix in (7) is Hermitian, that is, its conjugated transpose (denoted by $\#$) equals itself

$$\mathbf{H}^\#(w) = \mathbf{H}(w).$$

Therefore, \mathbf{H} has real eigenvalues, with associated eigenvectors orthogonal to each other. Solving the eigenequation, $\mathbf{H}\mathbf{t} = \lambda\mathbf{t}$, we obtain that the two eigenvalues of \mathbf{H} are

$$\lambda_1 = 1 + |\tilde{g}(w)|, \quad \lambda_2 = 1 - |\tilde{g}(w)|$$

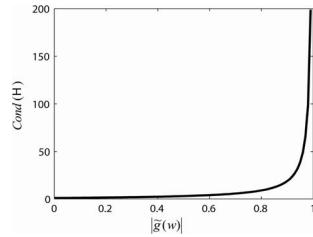


Fig. 1. Relationship between $|\tilde{g}(w)|$ and $\text{Cond}(\mathbf{H})$. Note that when $w \rightarrow 0$ (at lower frequency approaching dc) then $|\tilde{g}(w)| \rightarrow 1$, so $\text{Cond}(\mathbf{H}) \rightarrow +\infty$.

with corresponding orthonormal eigenvectors

$$\mathbf{t}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ e^{i\phi} \end{bmatrix}, \quad \mathbf{t}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -e^{i\phi} \end{bmatrix}.$$

Denote the condition number of \mathbf{H} as the ratio of its two eigenvalues λ_1, λ_2

$$\text{Cond}(\mathbf{H}) = \frac{\lambda_1}{\lambda_2} = \frac{1 + |\tilde{g}(w)|}{1 - |\tilde{g}(w)|}.$$

The condition number determines how input (at the corresponding frequency component) will be amplified—the larger the value of $\text{Cond}(\mathbf{H})$, the greater the amplification of small fluctuating noise. Since $0 \leq |\tilde{g}(w)| \leq 1$, $\text{Cond}(\mathbf{H})$ is monotonically related to $|\tilde{g}(w)|$ and bounded between 1 and ∞ . Fig. 1 shows the relationship between $|\tilde{g}(w)|$ and $\text{Cond}(\mathbf{H})$ —as frequency goes to 0, $|\tilde{g}(w)|$ approaches 1, resulting in the explosion of the condition number of \mathbf{H} , and hence disproportionate amplification of input noise. Fig. 2 demonstrates how artifacts might be introduced in the method of Zhang [6] when different amounts of noise are injected. For example, we can create, by (21), the ground-truth waveforms $f_s(t), f_r(t)$ with a known RT distribution. We then generate an ensemble of trial, and from which, the s-aligned and r-aligned ERP average waveforms. Plugging these s- and r-aligned waveforms as well as the RT distribution as input to (5) and (6) allows us to calculate $f_s(t), f_r(t)$. With different low-frequency noise added [SNR = 20, 30, and 40 dB, define by (24)], (5) and (6) give different outputs of $f_s(t), f_r(t)$; the recovered waveforms have more distortions with more low-frequency noise added. Put in another way, a pair of (carefully chosen) S- and R-component waveforms may cancel themselves out during the trial-by-trial stimulus- and response-aligned averaging, thus resulting in only noise in the ERP average waveforms. In short, the S–R decomposition algorithm of Zhang [6] suffers from disproportionate noise amplification at low-frequency—low-frequency distortions in the recovered waveforms may occur if noise is not handled properly.

C. Wiener Deconvolution

Wiener deconvolution refers to the application of Wiener filter to deal with noise inherent in a linear input–output system. Working in the frequency domain, the goal of the deconvolution technique is to minimize the impact of noise at all frequencies. Given a linear input–output system

$$y(t) = h(t) * x(t) + n(t)$$

where $x(t)$ is an unknown input signal, $h(t)$ is the known response, $n(t)$ is an unknown additive noise independent of $x(t)$, and $y(t)$ is an observed signal. The goal of Wiener deconvolution is to find some $d(t)$ so that we can estimate $x(t)$ as follows:

$$\hat{x}(t) = d(t) * y(t) \quad (9)$$

where $\hat{x}(t)$ is an estimate of $x(t)$ based on minimizing the MSE defined as

$$\text{MSE} = E[\|\hat{x}(t) - x(t)\|^2]. \quad (10)$$

The Wiener deconvolution filter $\tilde{d}(w)$ provides such a solution, which is conveniently given in the frequency domain

$$\tilde{d}(w) = \frac{\tilde{h}^C(w)}{|\tilde{h}(w)|^2 + n(w)/s(w)}.$$

Here, $s(w)$ and $n(w)$ are, respectively, the mean power spectrum of the signal $x(t)$ and of the noise $n(t)$, and $\tilde{h}^C(w)$ is the complex conjugate of the Fourier spectrum $\tilde{h}(w)$ of the transfer function $h(t)$. In actual application, one can: 1) find an initial estimate of signal $x_0(t)$ (by assuming zero noise, for instance) and calculate its signal power spectrum $s_0(w)$; 2) subtract the signal $x_0(t)$ from trial-by-trial data to obtain an estimate of noise power spectrum $n_0(w)$; 3) substitute the signal and noise power spectra into the Wiener filter (9) to get new estimate of signal $x_1(t)$; 4) repeat steps 1) to 3) until convergence. The final solution $x(t)$ has the property that $E[\|n(w)\|^2]$ is minimized.

Our signal model assumes that for every single trial, noise is independent of signal. With noise considered, (3) and (4) become, in matrix form,

$$\begin{bmatrix} \tilde{F}_s(w) \\ \tilde{F}_r(w) \end{bmatrix} = \begin{bmatrix} 1 & \tilde{g}(w) \\ \tilde{g}(-w) & 1 \end{bmatrix} \begin{bmatrix} \tilde{f}_s(w) \\ \tilde{f}_r(w) \end{bmatrix} + \begin{bmatrix} \tilde{\xi}_s(w) \\ \tilde{\xi}_r(w) \end{bmatrix} \quad (11)$$

in which $\tilde{\xi}_s(w)$ and $\tilde{\xi}_r(w)$ are the Fourier transform of noise waveforms time-locked to stimulus onset and to response onset, respectively. In vector notation, it is

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \boldsymbol{\xi}. \quad (12)$$

In order to apply the Wiener deconvolution method to the vectorial equations (11) and (12), we need to invert \mathbf{H} . Denote

$$\hat{\mathbf{H}} = \begin{bmatrix} 1 & -\tilde{g}(w) \\ -\tilde{g}(-w) & 1 \end{bmatrix}$$

such that

$$\hat{\mathbf{H}}\mathbf{H} = \mathbf{H}\hat{\mathbf{H}} = \lambda_1 \lambda_2 \mathbf{I}$$

where \mathbf{I} is the identity matrix. We can then convert (12) into (by multiplying both sides with $\hat{\mathbf{H}}$)

$$\hat{\mathbf{H}}\mathbf{y} = \lambda_1 \lambda_2 \mathbf{x} + \hat{\boldsymbol{\xi}}$$

with $\hat{\boldsymbol{\xi}} = \hat{\mathbf{H}}\boldsymbol{\xi}$, and then apply Wiener deconvolution filter to each of the vector component

$$\mathbf{x} = \frac{\lambda_1 \lambda_2}{(\lambda_1 \lambda_2)^2 + \text{SNR}^{-1}} \hat{\mathbf{H}}\mathbf{y}$$

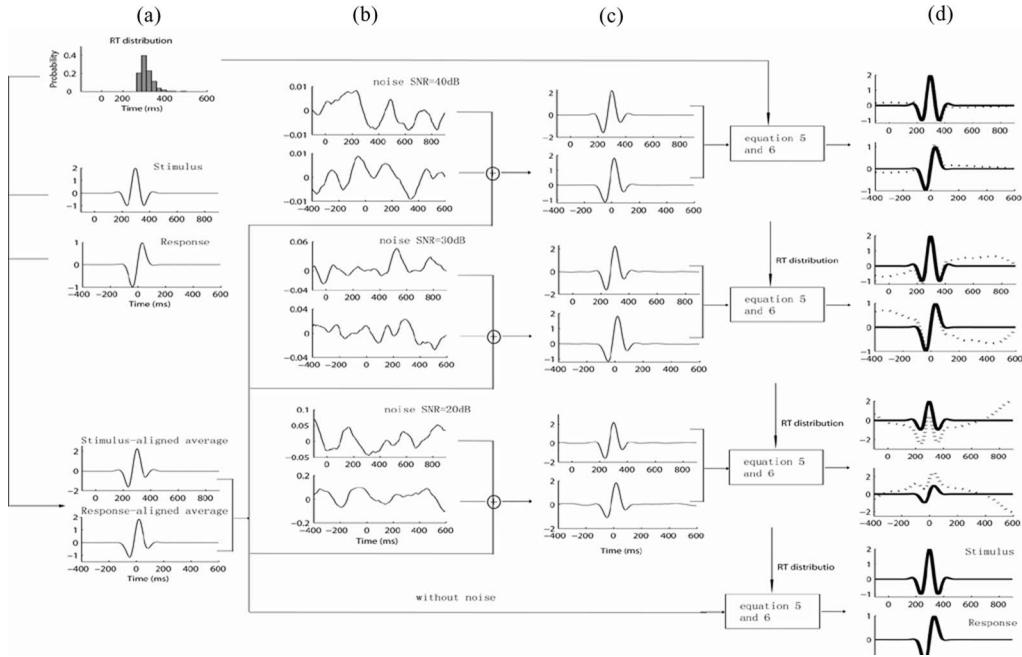


Fig. 2. Illustration of how artifacts in stimulus–response decomposition may arise. (a) Simulated S- and R-component waveforms created by (21), RT distribution, and resulting stimulus- and response-aligned waveforms. (b) Different low-frequency noise (SNR = 20, 30, and 40 dB) produced by (22) and (23). Note the difference in ordinate scales. (c) With different amount of low-frequency noise added into the ensemble of trials, resulting stimulus- and response-aligned waveforms change only slightly. (d) Recovered S- and R-component waveforms (dotted line) from (5) and (6) under different SNR condition and original S- and R-component waveform (thick lines). Clearly, input noises greatly affect recovery even though their effect on ERP averages is small.

or more explicitly,

$$\begin{bmatrix} \tilde{f}_s(w) \\ \tilde{f}_r(w) \end{bmatrix} = \frac{\lambda_1 \lambda_2}{(\lambda_1 \lambda_2)^2 + \text{SNR}^{-1}} \begin{bmatrix} \tilde{F}_s(w) - \tilde{g}(w)\tilde{F}_r(w) \\ \tilde{F}_r(w) - \tilde{g}(-w)\tilde{F}_s(w) \end{bmatrix}. \quad (13)$$

Here, SNR is the estimated signal-to-noise ratio which we take as $E[\|\tilde{f}_s\|^2]/E[\|\tilde{\xi}_s\|^2]$ in the calculation of the first vector (stimulus) component and $E[\|\tilde{f}_r\|^2]/E[\|\tilde{\xi}_r\|^2]$ in the calculation of the second vector (response) component. To estimate signal and noise power spectra, we use an iterative procedure by subtracting the currently-estimated waveforms from the raw, single trial data (see Section II-D). Since in the above procedure, Wiener filters are applied to the S-component waveform f_s and R-component waveform f_r directly, we call this method Wiener deconvolution with coupling of filters.

We have also investigated Wiener deconvolution with uncoupled filters, which turns out to work much better (as follows). First, we observe that the orthonormal eigenvectors of \mathbf{H} satisfy

$$\mathbf{t}_1^\# \cdot \mathbf{t}_1 = \mathbf{t}_2^\# \cdot \mathbf{t}_2 = 1, \quad \mathbf{t}_1^\# \cdot \mathbf{t}_2 = \mathbf{t}_2^\# \cdot \mathbf{t}_1 = 0.$$

We may expand the experimentally obtained waveform \mathbf{y} by

$$\mathbf{y} = (\mathbf{t}_1^\# \cdot \mathbf{y})\mathbf{t}_1 + (\mathbf{t}_2^\# \cdot \mathbf{y})\mathbf{t}_2 \equiv \mathbf{y}_1 + \mathbf{y}_2$$

where \mathbf{y}_1 and \mathbf{y}_2 represent the projections of \mathbf{y} onto \mathbf{t}_1 and \mathbf{t}_2 directions (the two terms in summation), respectively. Similarly, we can also expand the unknown waveform \mathbf{x} as

$$\mathbf{x} = (\mathbf{t}_1^\# \cdot \mathbf{x})\mathbf{t}_1 + (\mathbf{t}_2^\# \cdot \mathbf{x})\mathbf{t}_2 \equiv \mathbf{x}_1 + \mathbf{x}_2$$

so that

$$\begin{aligned} \mathbf{H}\mathbf{x} &= (\mathbf{t}_1^\# \cdot \mathbf{x})\mathbf{H}\mathbf{t}_1 + (\mathbf{t}_2^\# \cdot \mathbf{x})\mathbf{H}\mathbf{t}_2 \\ &= \lambda_1(\mathbf{t}_1^\# \cdot \mathbf{x})\mathbf{t}_1 + \lambda_2(\mathbf{t}_2^\# \cdot \mathbf{x})\mathbf{t}_2 = \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2. \end{aligned}$$

Of course, we do not know \mathbf{x} yet, so its projection onto the two eigendirections would have to be calculated.

According to our signal model equation (12), the noise component $\boldsymbol{\xi}$ may also project in the two eigendirections

$$\boldsymbol{\xi} = (\mathbf{t}_1^\# \cdot \boldsymbol{\xi})\mathbf{t}_1 + (\mathbf{t}_2^\# \cdot \boldsymbol{\xi})\mathbf{t}_2 \equiv \boldsymbol{\xi}_1 + \boldsymbol{\xi}_2$$

with their respective power spectra to be estimated. Note, however, that it is the noise power $\boldsymbol{\xi}_2$ in the direction (subspace) of \mathbf{t}_2 sustained by the second eigenvalue λ_2 that needs to be estimated accurately, as this will be the null space of the operator $\mathbf{H}(w)$ when the frequency goes to zero.

With the above notations, the signal model equation (12) breaks into two equations

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{H}\mathbf{x}_1 + \boldsymbol{\xi}_1 \quad \text{or} \quad (\mathbf{t}_1^\# \cdot \mathbf{y})\mathbf{t}_1 = \lambda_1\mathbf{x}_1 + (\mathbf{t}_1^\# \cdot \boldsymbol{\xi})\mathbf{t}_1 \\ \mathbf{y}_2 &= \mathbf{H}\mathbf{x}_2 + \boldsymbol{\xi}_2 \quad (\mathbf{t}_2^\# \cdot \mathbf{y})\mathbf{t}_2 = \lambda_2\mathbf{x}_2 + (\mathbf{t}_2^\# \cdot \boldsymbol{\xi})\mathbf{t}_2. \end{aligned} \quad (14)$$

The left-hand sides are what we can calculate from the data, and \mathbf{x}_1 and \mathbf{x}_2 are unknown 2-D vectors. Applying Wiener deconvolution method, we obtain

$$\mathbf{x}_i = \frac{\lambda_i}{\lambda_i^2 + (\text{SNR}_i)^{-1}} (\mathbf{t}_i^\# \cdot \mathbf{y})\mathbf{t}_i, \quad i = 1, 2 \quad (15)$$

where $\text{SNR}_i = E[\|\mathbf{x}_i\|^2]/E[\|\boldsymbol{\xi}_i\|^2]$, estimated separately for $i = 1, 2$ components. This is what we call Wiener deconvolution with “decoupled filters.” Note that since λ_1 is always greater than or equal to 1, the denominator $(\text{SNR}_1)^{-1}$ can (in most cases of noise range) be dropped in the estimate of \mathbf{x}_1 . The more crucial application of Wiener filter is to obtain the \mathbf{x}_2 component robustly against input noise, as noted earlier. In the next section, we provide a detailed algorithm for estimating SNR_2 .

D. Estimating Signal and Noise Power Spectra

Below, we describe in detail the estimation of signal and noise power spectra in the current context where “signals” are 2-D and coupled. We need to derive expressions that link signal power spectra (for both underlying S-component waveform and R-component waveform), noise power spectrum, and the difference of trial-by-trial waveform from the ensemble averaged waveforms.

Our signal model (for any single trial i) is

$$\begin{aligned}s_i(t) &= f_s(t) + f_r(t - t_i) + n_i(t); \\ r_i(t) &= f_s(t + t_i) + f_r(t) + n_i(t + t_i)\end{aligned}$$

or, explicitly,

$$\begin{aligned}s_1(t) &= f_s(t) + f_r(t - t_1) + n_1(t); \\ r_1(t) &= f_s(t + t_1) + f_r(t) + n_1(t + t_1) \\ s_2(t) &= f_s(t) + f_r(t - t_2) + n_2(t); \\ r_2(t) &= f_s(t + t_2) + f_r(t) + n_2(t + t_2) \\ &\vdots \\ s_k(t) &= f_s(t) + f_r(t - t_k) + n_k(t); \\ r_k(t) &= f_s(t + t_k) + f_r(t) + n_k(t + t_k)\end{aligned}$$

where n_1, n_2, \dots, n_k are noise waveforms on each trial, assumed to be independent of the signal and independent of each other. Averaging across trials, we obtain

$$\begin{aligned}F_s(t) &= \frac{1}{k} \sum_{i=1}^k s_i(t) = f_s(t) + f_r(t) * g(t) + \frac{1}{k} \sum_{i=1}^k n_i(t) \\ F_r(t) &= \frac{1}{k} \sum_{i=1}^k r_i(t) = f_r(t) + f_s(t) * g(-t) + \frac{1}{k} \sum_{i=1}^k n_i(t + t_i)\end{aligned}$$

where

$$\xi_s(t) = \frac{1}{k} \sum_{i=1}^k n_i(t), \quad \xi_r(t) = \frac{1}{k} \sum_{i=1}^k n_i(t + t_i).$$

In order to estimate the power spectrum of the signal and of the noise, we subtract s-aligned average signal from single trial waveform

$$\begin{aligned}s_i(t) - F_s(t) &= (f_s(t) + f_r(t - t_i) + n_i(t)) - F_s(t) \\ &= (f_r(t - t_i) - f_r(t) * g(t)) + \left(n_i(t) - \frac{1}{k} \sum_{i=1}^k n_i(t) \right) \quad (16)\end{aligned}$$

where

$$f_r(t) * g(t) = \frac{1}{k} \sum_{i=1}^k f_r(t - t_i).$$

In frequency domain, (16) becomes

$$\begin{aligned}\tilde{s}_i(w) - \tilde{F}_s(w) &= (\tilde{f}_r(w) \cdot e^{-jw t_i} - \tilde{f}_r(w) \cdot \tilde{g}(w)) \\ &+ \left(\tilde{n}_i(w) - \frac{1}{k} \sum_{i=1}^k \tilde{n}_i(w) \right).\end{aligned}$$

As noise is independent of the signal,

$$\begin{aligned}E[|\tilde{s}_i(w) - \tilde{F}_s(w)|^2] &= E[|(\tilde{f}_r(w) \cdot e^{-jw t_i} - \tilde{f}_r(w) \cdot \tilde{g}(w))|^2] \\ &+ E\left[\left(\tilde{n}_i(w) - \frac{1}{k} \sum_{i=1}^k \tilde{n}_i(w)\right)^2\right].\end{aligned} \quad (17)$$

Evaluating the first term on the right-hand side

$$\begin{aligned}E[|(\tilde{f}_r(w) \cdot e^{-jw t_i} - \tilde{f}_r(w) \cdot \tilde{g}(w))|^2] &= |\tilde{f}_r(w)|^2 \\ &\cdot E[(1 + |\tilde{g}(w)|^2 - (\tilde{g}^C(w) \cdot e^{-jw t_i} + \tilde{g}(w) \cdot e^{jw t_i}))]\end{aligned}$$

and summing over i yields

$$\begin{aligned}\frac{1}{k} \sum_{i=1}^k E[|(\tilde{f}_r(w) \cdot e^{-jw t_i} - \tilde{f}_r(w) \cdot \tilde{g}(w))|^2] &= |\tilde{f}_r(w)|^2 (1 - |\tilde{g}(w)|^2).\end{aligned}$$

The second term on the right-hand side of (17) is

$$\begin{aligned}E\left[\left(\tilde{n}_i(w) - \frac{1}{k} \sum_{i=1}^k \tilde{n}_i(w)\right)^2\right] &= E[|\tilde{n}_i(w)|^2] - \frac{2}{k} E[|\tilde{n}_i(w)|^2] + \frac{1}{k^2} E\left[\sum_{i=1}^k |\tilde{n}_i(w)|^2\right] \\ &= \left(1 - \frac{1}{k}\right) E[|\tilde{n}_i(w)|^2]\end{aligned}$$

where we have used the property that noise is independent of each other on individual trials but with the same expected power spectrum,

$$E[|\tilde{n}_1(w)|^2] = E[|\tilde{n}_2(w)|^2] = \dots = E[|\tilde{n}_k(w)|^2] \equiv k |\tilde{\xi}_s(w)|^2.$$

Therefore, we obtain the relation

$$\begin{aligned}\frac{1}{k} \sum_{i=1}^k |\tilde{s}_i(w) - \tilde{F}_s(w)|^2 &= |\tilde{f}_r(w)|^2 \cdot (1 - |\tilde{g}(w)|^2) + (k - 1) |\tilde{\xi}_s(w)|^2.\end{aligned} \quad (18)$$

A similar expression can be derived for the r-aligned analysis

$$\begin{aligned}\frac{1}{k} \sum_{i=1}^k |\tilde{s}_i(w) - \tilde{F}_r(w)|^2 &= |\tilde{f}_s(w)|^2 \cdot (1 - |\tilde{g}(w)|^2) + (k - 1) |\tilde{\xi}_r(w)|^2.\end{aligned} \quad (19)$$

The left-hand sides of (18) and (19) can be calculated from the experimental data directly. Once $|\tilde{f}_r(w)|^2$ or $|\tilde{f}_s(w)|^2$ is known, then $|\tilde{\xi}_s(w)|^2$ or $|\tilde{\xi}_r(w)|^2$ can be estimated, which, in turn, can be used in the Wiener deconvolution filter to estimate S- and R-component waveforms. This creates an iterative procedure, just as in the vanilla Wiener deconvolution case described in Section II-C.

In Section II-C, we discussed two approaches, one based on coupling of filters (see (13)), and the other on uncoupling of filters (see (15)). The difference is whether we first project all

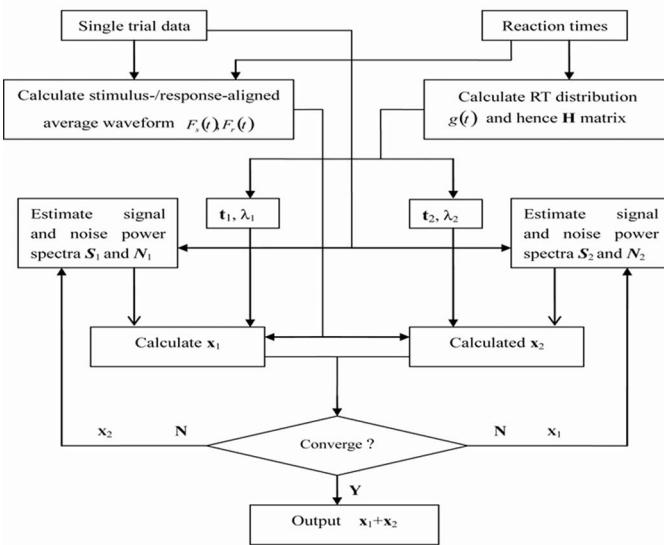


Fig. 3. Flow diagram of S-R waveform decomposition with noise control. During the first pass, x_1 , x_2 is calculated without estimating signal and noise spectra. During the second and subsequent passes, Wiener deconvolution filter is used for calculating x_1 , x_2 (as indicated by the open arrow, distinct from all other arrows).

waveforms in the eigendirections of the \mathbf{H} operator or not. In the “coupled filter” approach, no distinctions are made whether noise lie inside or outside the null space of the \mathbf{H} operator. In the “decoupled filter” approach, which is mathematically rigorous, we treat the projections of the noise onto \mathbf{t}_1 and \mathbf{t}_2 directions distinctly and apply Wiener filters separately; this is based on the noise decomposition given by (14). Our algorithm is shown in Fig. 3.

E. Tikhonov Regularization

Tikhonov regularization is the most commonly used method for regularizing ill-posed problems. In statistics, the method is also known as ridge regression. Given a system of linear equations $\mathbf{A}x = b$, where the matrix \mathbf{A} may be singular (and hence the problem is ill-conditioned), one may seek a solution that minimizes the residual $\|\mathbf{A}x - b\|^2$, where $\|\cdot\|$ is the Euclidean norm. In order to give preference to solutions with certain desirable properties (such as smoothest, or with minimal power), a regularization term is included in this minimization:

$$\|\mathbf{A}x - b\|^2 + \beta^2 \|x\|^2. \quad (20)$$

An explicit solution, denoted by \hat{x} , is given by

$$\hat{x} = (\mathbf{A}^T \mathbf{A} + \beta^2 \mathbf{I})^{-1} \mathbf{A}^T b$$

where \mathbf{I} is the identity matrix, and the regularization parameter β controls the quality of the final solution. An appropriate choice of the value of the regularization parameter should balance the goodness-of-fit, the first term in expression (20), with the amount of signal power in the null space of the \mathbf{A} operator, related to the second term in expression (20). The criteria for choosing a suitable regularization parameter is still a hotly researched problem, for which several methods have been pro-

posed [12], [13]. If the norm of the noise is known a priori, as in simulation studies, discrepancy principle (DP) can be used to choose regularization parameter, an approach investigated in some depth [14], [15]. Other methods that do not rely on *a priori* knowledge of noise levels include generalized cross validation (GCV) [16] and L-curve [17], [18], etc. In our current context, where noise estimate is difficult and only indirectly obtained, we shall compare GCV and L-curve methods in this paper.

F. GCV and L-Curve

The basic principle of the GCV method is to omit one of the data points and then determine the regularization parameter, with the goal of predicting the missing data point with greatest accuracy. This is done for all data points in the set, and a GCV score can be evaluated as a function of the regularization parameter. Adapting GCV method to our situation, as our decomposition algorithm is based on s-aligned ERP average and r-aligned ERP average waveforms as input, we omit one trial in our input data-set and determine the regularization parameter that predicts the missing trial waveform with greatest accuracy; this procedure is done for all trials in the input data-set to obtain a GCV score.

The L-curve method is an alternative method for determining the regularization parameter β . An L-curve is formed by plotting the norm squared of the penalty function $\|x\|^2$ against the residual from data-fitting $\|\mathbf{A}x - b\|^2$, for a broad range of β parameters considered. As the L-curve clearly depicts the compromise between minimization of these two quantities, which is at the heart of any regularization method, the point of maximum curvature (i.e., the L-shaped corner) is taken to be the point of optimal regularization parameter.

III. SIMULATION

A. Simulation Parameters and Performance Measures

1) *Simulated Waveform:* We consider simulated ERP waveforms as produced by the following equation:

$$f_k(t_i) = A_k \exp \left(- \left(2\pi\lambda_k \frac{t_i - \tau_k}{\gamma_k} \right)^2 \right) \times \cos(2\pi\lambda_k(t_i - \tau_k) + \alpha_k), \quad i = 1, \dots, n, \quad k = s, r \quad (21)$$

where $t_i = i * 0.004 \text{ s} = i * 4 \text{ ms}$ and $n = 500$. Here, A_k is the amplitude of the stimulus ($k = s$) or response ($k = r$) component waveform and τ_k is the random variable representing trial-by-trial onset time for stimulus event or response event. Without loss of generality, we take τ_s to be fixed at 100, and τ_r to be sampled from a Gamma distribution, with mean and variance given below. The value of α_k ($k = s, r$) is taken to be fixed between 0 and 2π . The parameter λ_k denotes the principal frequency of the simulated waveform, and γ_k is its time scale factor—the bigger γ_k , the longer duration of the waveform. ERP waveform on any individual trial is taken to be the sum of the above two components (one stimulus-locked and one response-locked), plus background noise described as follows.

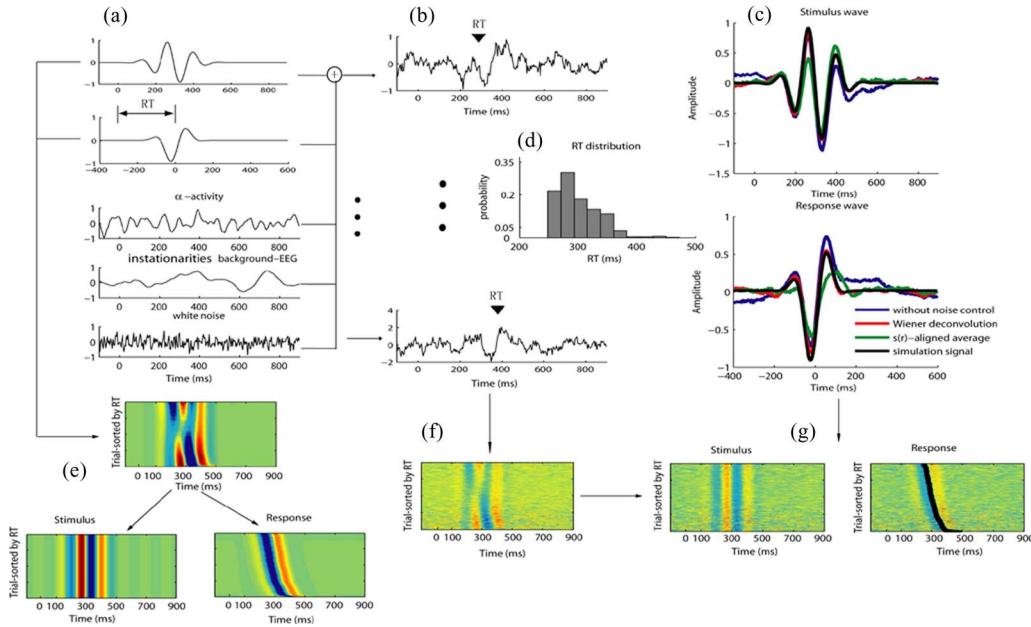


Fig. 4. Illustration of the S-R decomposition algorithm with or without noise control by the method of Wiener deconvolution. (a) S- and R-component waveforms and noise waveforms (α -activity, nonstationary background-EEG, and white noise) used for simulation. (b) Simulated single trial waveforms created through summing together S- and R-component waveforms, according to different RT, plus noise. (c) Comparison of original S- and R-component waveforms, s- and r-aligned ensemble average waveforms, and recovered S- and R-components either without noise control [calculated by equations (5) and (6)] or with Wiener deconvolution method. (d) RT distribution. (e) ERP image of the ensemble of trials for original stimulus-aligned data (Top) and for either S-component (Lower Left) or R-component (Lower Right) after trials sorted according to RT, with no noise injected. (f) ERP image of the ensemble of trials with noise added. This is the input to the S-R decomposition algorithm. (g) ERP image of ‘stimulus’ part and ‘response’ part, in which on each single trial, recovered R-component waveform (or S-component waveform) is subtracted, respectively. Here, different colors in an ERP image encode different amplitude values. The black line in ‘response’ part is RT curve.

To simulate the background EEG, two autoregressive (AR) processes with white noise are added into each trial. The coefficient of the AR2-processes is the same as that used in the paper of Krieger *et al.* [20]. The first AR2 process mimics alpha-band activity, and is specified as

$$v_1(n) = 1.721 \cdot v_1(n-1) - 0.819 \cdot v_1(n-2) + u(n); \quad (22)$$

the second AR2 process captures the nonstationarity of the background EEG, and is specified as

$$v_2(n) = 1.979 \cdot v_2(n-1) - 0.980 \cdot v_2(n-2) + u(n). \quad (23)$$

Here, $u(n)$ represents zero-mean Gaussian white noise.

2) *Performance Measures:* Two measures, i.e., relative error (RE) and correlation coefficient (COR), are used to evaluate the performance of the various recovery algorithms. Suppose \hat{x} is the estimated value for a true signal x , they are defined as

$$\text{RE} = \sqrt{\|x - \hat{x}\|^2 / \|x\|^2}, \quad \text{COR} = \langle x, \hat{x} \rangle / \|x\| \cdot \|\hat{x}\|$$

where $\langle \cdot \rangle$ denotes vector inner product. Here, waveforms (amplitude as a function of time) are considered as vectors in a (theoretically infinite-dimensional but practically finite-dimensional with dimensionality N) vector space, with vector addition defined by pointwise addition of evaluated function values.

In the simulations, the amount of noise is controlled by the SNR, defined as

$$\text{SNR} = 10 \cdot \log_{10} (\|\text{signal}\|^2 / \|\text{noise}\|^2). \quad (24)$$

B. Results

With stimulus onset time chosen as the reference zero, the response time (τ_r) relative to the stimulus time (τ_s) is taken from a Gamma distribution with a mean of 300 ms and four different values of standard deviations (see below). Four different SNR values were used (see below). For each condition of RT distribution and SNR value, the simulation is repeated 50 times to yield statistical averages. In every repeat, α_k is randomly selected within $0-2\pi$, and A_k are random values between 1.0–2.0, λ_s and λ_r are also random values between 5.0–7.0 and 4.0–6.0 respectively. Since, from (7), the smallest eigenvalue (we denote it λ_0) of matrix \mathbf{H} is below 1 and close to 0, while the largest is close to 2, we only select regularization parameter β that satisfies $\lambda_0 \leq \beta^2 \leq 2$ for GCV and L-corner methods.

Fig. 4 shows an example of decomposition results using the Wiener deconvolution method, under the following simulation parameters: $\gamma_s = 1.2$ and $\gamma_r = 0.8$, $\lambda_s = 5.9$ and $\lambda_r = 4.7$, $\alpha_s = 0.36$ and $\alpha_r = -0.42$, SNR = -10 dB, the standard deviation of the RT distribution is 20 ms, and the number of trials 100.

Fig. 5 shows the RE and COR performance measures for various noise control methods (Wiener deconvolution, LCV, GCV), along with no noise control (waveforms directly calculated by the original algorithm of Zhang [6] as given by (5) and (6) above). We conclude that the results obtained from the GCV method have larger errors for low SNR and smaller errors for high-SNR conditions. On the other hand, using the L-corner method, we can recover true signal in low-SNR condition, but

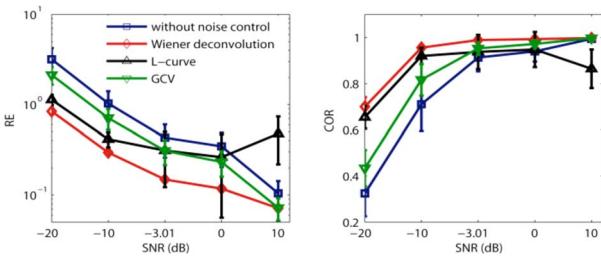


Fig. 5. Comparison of the results of various noise control techniques (GCV, L-corner, Wiener deconvolution) under different SNR conditions (standard deviation of RT is 20 ms, and total trial number is 200). Error bars represent standard deviation over 50 repeated simulations.

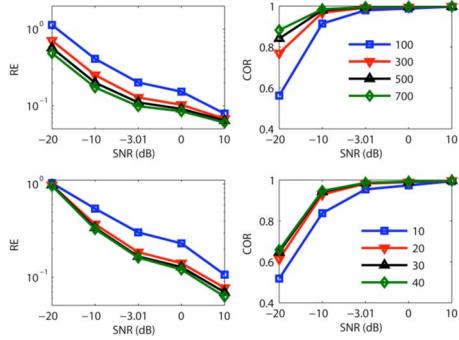


Fig. 6. Performance of Wiener deconvolution method of noise control when the number of trials are 100, 300, 500, and 700 (standard deviation of RT for each condition is fixed at 20 ms) (Top), and when the standard deviation of RT distribution is 10, 20, 30, and 40 ms (the number of trials for each condition is fixed at 100) (Bottom).

not for high-SNR condition, since usually the ‘‘corner’’ does not exist or cannot be determined with accuracy when SNR is high. Our Wiener deconvolution method performed well for both low- and high-SNR conditions—in fact, it is superior to both the GCV and the L-corner method across all SNR values tested.

To further explore the Wiener deconvolution approach for noise control, we investigate how its performance degradation (as SNR is increased) depends on the standard deviation of the RT distribution and on the number of trials available for each simulation condition. The results are depicted in Fig. 6. Generally, performance of our algorithm degraded as the trial number decreased or as the standard deviation of the RT distribution increased. However, the changes are gradual, so degradation is graceful.

To confirm the stability of our algorithm, we choose the same parameter values, as shown in Fig. 4, using an ensemble of 1000 trials, and then randomly subsample 100, 300, 500, 700, and 900 trials to apply our recovery algorithm to. The results are shown in Fig. 7. Clearly, our method is quite robust with respect to significant reduction of number of trials available as input to the decomposition algorithm. This is important, since in many real experiments, trial number per experimental condition cannot be arbitrary large due to constraints on experimental design.

Lastly, we compare the results of two variants of Wiener deconvolution, one where the filters were ‘‘coupled’’ and the other ‘‘decoupled’’ (see Section II-D for details). Simulation parameters: $A_s = A_r = 1$; $\gamma_s = 2$ and $\gamma_r = 1$, $\lambda_s = 3.0778$ and

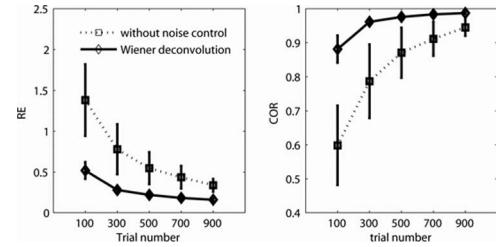


Fig. 7. Comparison of the performance measures between direct calculation (i.e., without noise control) and Wiener deconvolution method, using random subsamples of 100, 300, 500, 700, and 900 trials from a fixed ensemble of 1000 trials.

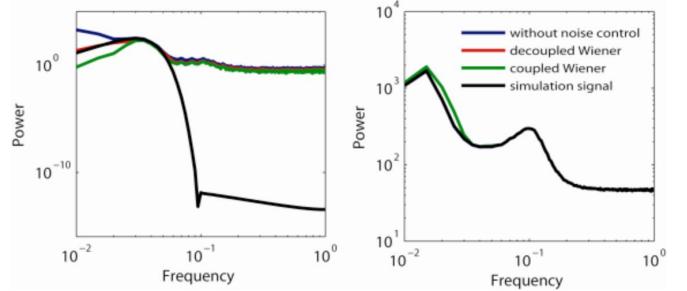


Fig. 8. Comparison of signal-to-noise power spectra across different recovery methods (power spectrum is normalized across frequencies). (Left) Power spectra of the recovered signal by different recovery methods. (Right) Power spectra of noise calculated by different recovery methods. Noise spectra are calculated by subtracting, from each trial, the signal waveforms recovered from the various methods, computing the power of the residue waveform (noise) of that trial, and averaging across all trials. The blue curve is completely overlapped with the black curve.

$\lambda_r = 3.17$, $\alpha_s = 1.36$ and $\alpha_r = -1.42$, the standard deviation of the RT distribution is 20 ms, with SNR = -10 dB and the number of trials 100, and the average results are calculated over 50 repeats. Power spectra of recovered signal by different methods and mean power spectra of noise of all trials are shown in Fig. 8. We can see that while Wiener deconvolution with uncoupled filters almost perfectly recovered true power spectrum of the simulated waveform and mean power spectra of noise of all trials, power spectra at lower frequencies are *underestimated* in the Wiener deconvolution with coupled filters and *overestimated* in direct calculation method without noise control. This means that the uncoupled Wiener filters are successful in controlling for noise, especially at low-frequency range, which has been problematic for applying Zhang [6].

IV. ILLUSTRATIVE EXAMPLE

A. Experimental Data

In this section, we apply our S–R decomposition algorithm to experiment data on a Go/NoGo task. Data used were obtained from an internet open source (http://www.sccn.ucsd.edu/~arno/fam2data/publicly_available_EEG_data.html) [21].

EEG was recorded by the experimenters with the 32-channel system at a sampling rate of 1000 Hz, and Cz was taken as the reference. Epochs contaminated with excessive eye movements, blinks, muscle artifact, or amplifier blocking were manually removed prior to averaging. The ERPs were re-referenced to the grand average reference during offline data processing.

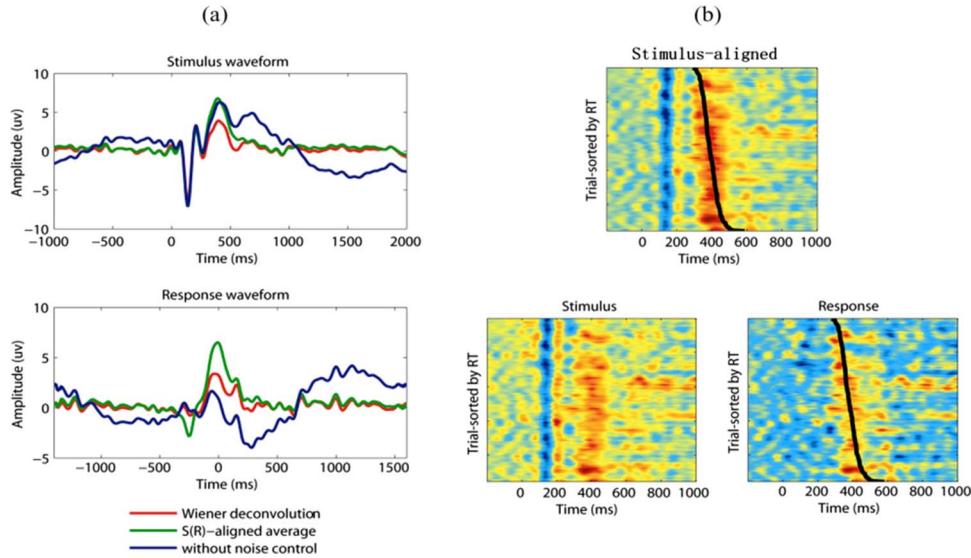


Fig. 9. Result of S-R decomposition of Pz channel for one subject in a Go/NoGo task. (a) Recovered S- and R-component waveform, obtained with Wiener deconvolution for noise control (red line) or without noise control (blue line), is compared with s-aligned ERP average, where 0 marks stimulus (or response) onset. (b) ERP images of the empirical data-set and after S-R decomposition. (Top) ERP image of all single trials sorted according to ascending RT; black line is the RT curve. (Bottom) ERP image of “stimulus” part and “response” part, in which recovered R-component or S-component has been, respectively, subtracted from every single trial.

B. Results

1) *Result From the Recognition Task:* For demonstration purpose, we first apply our Wiener deconvolution technique to the recognition task (recognition of “hard” animal images) of the aforementioned experiment. ERP data were taken from one subject (fsa) at electrode Pz with a total 248 trials under analysis. Results of Fig. 9 show that our Wiener deconvolution algorithm (using uncoupled filters) offers a remarkable improvement over the direct calculation method since, without noise control, the latter method disproportionately magnified low-frequency noise, and therefore, introduced significant distortions, causing an artifact even in the baseline regions of the waveforms. Visualizing the results using the ERP-image method (introduced by Jung *et al.* [24]) show that our algorithm can faithfully recover the underlying S- and R-component waveforms.

2) *Result from the Classification Task:* Here, we analyze ERPs recorded from the categorization task. Fig. 10(a; left) depicts the pure S-component waveform extracted from Go trials, and compares it to s-aligned ERP averages on both Go and NoGo trials across subjects on typical channels Fz, Pz, and Oz. It can be seen that, for Fz channel, the N2 component is smaller than that in the uncorrected s-aligned average; the P3 component in Go trials resembles that in NoGo trials. Our results thus reveal a significant amount of contamination, by response-locked ERP component, in stimulus-aligned ERP average on Go trials.

To validate our results, we split the real EEG data on Go trials into two categories – those with short RTs and those with long RTs (see Fig. 10(a), right). It is reasonable to assume that this cross contamination would be less or weaker in long RT trials (due to their larger separation) than in short RT trials so the original s- and r-aligned averages would be closer to “pure” S- and R-components in long RT trials than in short

RT trials. Our results (see Fig. 10(b)) show that for Fz and Oz channels, RE values (s- and r-aligned averages versus recovered S- and R-components) of long RT trials are significantly smaller than RE values from short RT trials (Fz: $t = 2.3$; $p = 0.04$, Oz: $t = 4.45$; $p = 0.001$). For COR measure, however, no significant difference is observed across all three channels, presumably because COR measures how similar two waveforms are across the entire period and not only the period between the stimulus onset and response onset. We then calculate the L_2 -distance (see Fig. 10(c) light-colored bars) of the original s- and r-aligned average waveforms (together forming the y-vector) constructed from the short-RT trials versus those constructed from long-RT trials, and the L_2 -distance (see Fig. 10(c) dark-colored bars) of the recovered S- and R-component waveforms (together forming the x-vector) constructed from the short RT trials versus those constructed from the long-RT trials. Results show that the L_2 -distance (between the short- and long-RT trials) is significantly larger in the uncorrected original waveforms than in the recovered waveforms across all three channels (Fz: $t = 7.1$; $p = 0$, Pz: $t = 2.66$; $p = 0.02$, Oz: $t = 4.94$; $p = 0.0003$).

In order to illustrate spatial effect of our decomposition algorithm, we choose three typical ERP components (P1—about 100 ms after stimulus onset; P3—about 350 ms after stimulus onset; response—about 25 ms after response onset), and compare the ERP topography before and after S-R decomposition, see Fig. 10(d). Results show that for the P1 component, original waveforms and the recovered waveforms have very similar spatial distribution. For the P3 and response components, significant differences are revealed in the prefrontal area and posterior parietal areas, with much larger amplitudes, due to S-R cross contamination (statistical test results will be shown in a separate report).

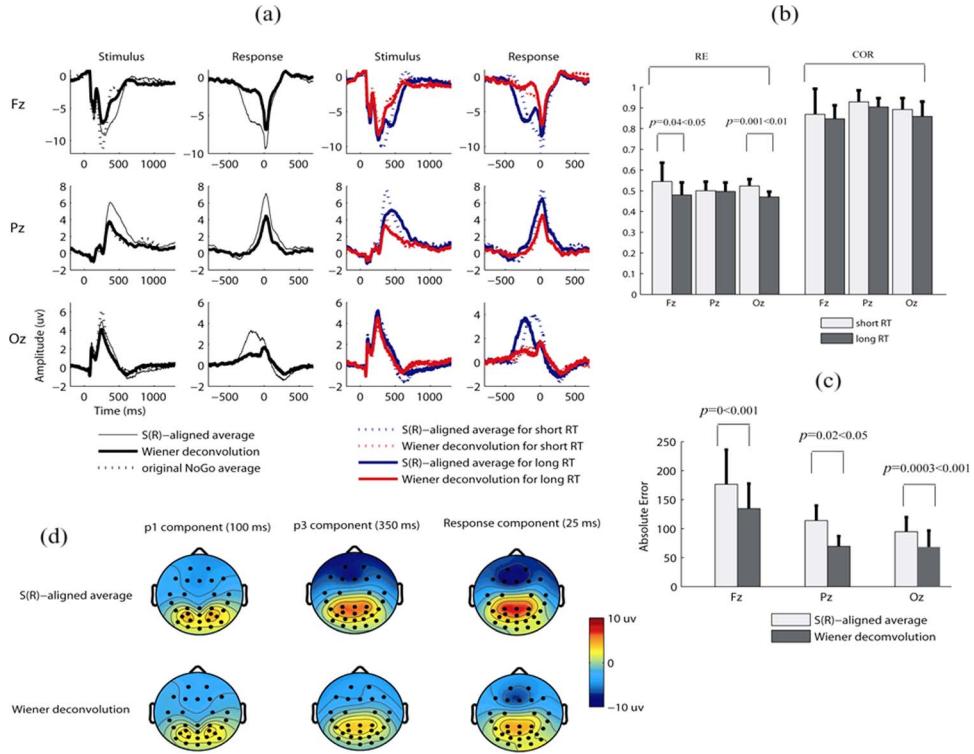


Fig. 10. Result of S-R decomposition for the classification task. (a) (Left) Extracted pure S- and R-component waveforms for Go trials, and the stimulus-aligned ERP averages for both Go and NoGo trials on channels Fz, Pz, and Oz across all subjects. (Right) Extracted pure S- and R-component waveform for short-RT and long-RT trials, respectively. (b) Comparison of original s- and r-aligned ERP average with recovered S- and R-component for short-RT trials and for long-RT trials, respectively. (c) Comparison of L_2 distances between s- and r-aligned ERP average on short-RT trials with s- and r-aligned ERP average on long-RT trials, and between recovered S- and R-components from short-RT trials and recovered S- and R-component from long-RT trials. (d) Topography of the P1 component (100 ms), the P3 component (350 ms) and the response component (25 ms) of response onset of original average ERP and recovered waveforms, respectively, for Go trials.

V. DISCUSSION

Our main aim in this report is to investigate the robustness of the algorithm of Zhang [6] to decompose stimulus and response components in ERP averages, when trial-by-trial variation in RT raises issues in interpreting stimulus-aligned and response-aligned average waveforms. The main challenge is that, due to the ill-conditioned matrix \mathbf{H} in (12), the solution is very susceptible to noise at low frequency. Hence, recovery of component waveforms is susceptible to distortion, where a large artifact may be introduced in both the recovered S-component waveform and R-component waveform such that they approximately cancel each other out. Put another way, there is a null space of waveforms, with nonvanishing energy at low-frequency components corresponding to the (approximately) zero eigenvalue of the \mathbf{H} matrix, that can be added to any recovered waveform, still leading to approximate satisfaction of (3) and (4), as demonstrated by Fig. 2. This makes the recovery of S- and R-component waveforms an ill-posed problem, and calls for suitable regularization and noise control. To this end we considered Tikhonov regularization approach and the Wiener deconvolution approach.

A. Tikhonov Regularization

It is common knowledge that the main challenge for Tikhonov regularization is the choice of the regularization parameter β .

If β is too small, the problem is not sufficiently regularized for noise to be properly controlled. If β is too large, then goodness of fit of the solution may be sacrificed. Obviously, optimal choice of β represents a compromise, which depends heavily on the SNR. In our simulation work, we considered the GCV method and the L-corner method, and found (see Fig. 5) that the GCV method works well in high-SNR condition but not in low-SNR condition, while the L-corner method works well in low-SNR condition but not in high-SNR condition. In practice, as single-trial EEG data contain low SNR, so L-corner is preferred as the method to determine the regularization parameter. However, our paradigm of recovering stimulus- and response-locked components are based on ERP averages which contain high-SNR. In such a case, the ‘corner’ of the L-corner curve may be hard to locate and determine, leading to practical difficulty. On the other hand, the GCV method would rely on having a large number of trials with identical statistical properties. In the current problem, there is a large variation of RTs across trials, and the number of trials with identical RT is very limited. Therefore, the basic premise for the GCV method is at question.

B. Wiener Deconvolution

Wiener deconvolution is a very common method for noise control in solving the inverse problem for linear input-output systems. The main advantage is that there is no need to estimate

the regularization parameter; instead, one only needs to estimate power spectra for both signal and noise. In our current application, two challenges emerge: 1) adapting the Wiener filter to only the low-frequency components, and 2) estimating signal and noise power spectra.

Our signal model treats noise that lie in the null space of the \mathbf{H} matrix (i.e., the directions corresponding to λ_2) separately from noise that lie in the other directions. Therefore, we apply separate Wiener filters after projecting the data into those directions that make up such null space and those that are orthogonal to the null space, respectively. Signal and noise spectra are estimated after this projection is done. This (what we call) “decoupled Wiener filters” approach effectively controls noise and avoids low-frequency artifact in a robust fashion. As a comparison, Wiener filters with coupling did not perform as well (see Fig. 8).

The resulting algorithm, using uncoupled Wiener deconvolution filters, is shown to outperform GCV and L-corner methods for all signal-to-noise levels (see Fig. 5), and is robust against a wide range of variances in RT distribution (see Fig. 6, bottom) and total trial numbers (see Fig. 6, top, and Fig. 7). The recovered waveform matches the signal almost perfectly for our simulated data (see Fig. 4), and sensibly in data from a real ERP experiment (see Fig. 9 and Fig. 10). Taken together, the current algorithm successfully accomplishes our goal of noise control in S–R waveform decomposition.

C. Experiment Data in Go/NoGo Task

As an application of our method, we extracted the stimulus- and response-locked ERP components for a Go/NoGo task. In Go/NoGo tasks, two major ERP components have been of interest to investigators, a negative potential within a latency range of 200–300 ms (N2 component), which is reported to have larger amplitude in NoGo trials as compared with Go trials, and a positive wave peaking within a latency range of 300–500 ms (P3 component), which is reported to have larger amplitude for frontal channels in the NoGo trials than in the Go trials [22], [23]. Our analysis showed that there is a big difference between the original stimulus- and response-aligned EEG averages and recovered stimulus- and response-locked component waveforms for channels Fz and Pz—P3 amplitude in the original stimulus-/response-aligned ERP averages are larger than the true underlying S-component/R-component waveforms. On the other hand, there appears to be no difference in the P3 time patch between recovered S-component waveform on Go signal and the stimulus-aligned ERP average on NoGo trial (see Fig. 10, left column). This implies that the difference in P3 between Go and NoGo trials may result from the presence of a response-locked component in Go trials, rather than any difference in stimulus processing between Go and NoGo trials. That the P3 component (peaking around 350 ms post stimulus onset) and the response component (peaking around 50 ms post-response onset) suffer from stronger cross-contamination is understandable since the mean RT is only 429 ms. In contrast, the earlier component P1 suffers little cross-contamination, as evidenced by the close resemblance of the spatial distributions

of the original stimulus-aligned ERP average and the recovered stimulus-locked waveform. A more detailed report, including tests of statistical significance for these effects, is forthcoming (Yin and Zhang, in preparation).

D. Basic Assumption and Limitation

Our mathematical model [6] makes the fundamental assumption that the true, underlying S- and R-components are the same in each and every single trial, in terms of timing and overall amplitude of the waveform. The issue of amplitude and latency variation of an ERP component across single trials has been addressed [25], and methods based on maximal likelihood estimates [26] and Bayesian maximal *a posteriori* (MAP) estimates [27] have been proposed. An examination of our analysis reveals that we can easily relax the assumption on exact timing and allow random latency jittering—in this case, the recovered waveforms are simply the jittered version of the underlying S- and R-component waveforms. The assumption on equal amplitude of the underlying waveform across individual trials is more critical—it amounts to a multiplication of the RT distribution by the RT-dependent amplitude modulation function and hence will change depending on the RT distribution. Future research will combine the template matching technique [27] with our S–R decomposition technique to achieve better results in extracting underlying waveforms that give rise to trial-by-trial evoked potential.

Our signal model (see (3) and (4)) assumed that single-trial waveforms result entirely from a stimulus-locked component waveform and a response-locked component waveform, plus noise. Based on this assumption, unique recovery of these component waveforms is mathematically guaranteed. However, by scrutinizing the results (ERP images in Fig. 9), one finds that the recovered S-component waveform and R-component waveform share considerable common activities (on a window from about 400 ms after stimulus onset). Such common activity seems to reflect a component waveform that may neither be time-locked to stimulus nor to response (perhaps to stimulus-to-response “decision”), but yet has been, unwarantedly, attributed to an S- and R-component waveform. In [6], the possibility of a “decision-related” component, which is time-locked to the putative transition of a stimulus processing and a response processing stage, has been investigated. Future research will implement the solution strategy outlined in that paper to isolate such decision-related component if it exists.

VI. CONCLUSION

In this paper, we provided a robust method for decomposition of stimulus-locked and response-locked components in ERP recordings through careful noise control. Compared against different regularization-based methods, such as GCV and L-corner methods, we found that denoising based on the Wiener deconvolution technique, applied to those (low) frequency components that fall under the null space of an operator (controlled by the RT distribution), gives the most robust results. Our conclusion is confirmed by both simulated data and experimental data from a Go/NoGo task. The techniques described by this report will

make the method of Zhang [6] practically applicable to datasets where, despite considerable variation in trial-by-trial RTs, the total number of trials may be limited.

ACKNOWLEDGMENT

The authors thank an anonymous reviewer for pointing out apparent common activities in the recovered S-component waveform and R-component waveform (see Section V).

REFERENCES

- [1] N. K. Squires, K. C. Squires, and S. A. Hillyard, "Two varieties of long-latency positive waves evoked by unpredictable auditory stimuli in man," *Electroencephalogr. Clin. Neurophysiol.*, vol. 38, pp. 387–401, 1975.
- [2] S. A. Hillyard and T. W. Picton, "Electrophysiology of cognition," in *Handbook of Physiology: The Nervous System*, vol. V, E. Plum Ed., Bethesda, MD: Amer. Physiol. Soc., 1987, pp. 519–584.
- [3] M. Kutias and S. A. Hillyard, "Event-related brain potentials to semantically inappropriate and surprisingly large words," *Biol. Psychol.*, vol. 11, pp. 99–116, 1980.
- [4] W. J. Gehring, B. Goss, M. G. Coles, D. E. Meyer, and E. Docherty, "A neural system for error detection and compensation," *Psychol. Sci.*, vol. 4, pp. 385–390, 1993.
- [5] M. Falkenstein, N. A. Koshlykova, V. N. Kiroi, J. Hoermann, and J. Hohnsbein, "Late ERP components in visual and auditory go/nogo tasks," *Electroencephalogr. Clin. Neurophysiol.*, vol. 96, pp. 36–43, 1995.
- [6] J. Zhang, "Decomposing stimulus and response component waveforms in ERP," *J. Neurosci. Methods*, vol. 80, pp. 49–63, 1998.
- [7] G. Yin, J. Zhang, T. Yin, and D. Yao, "A multi-component decomposition algorithm for event-related potentials," *J. Neurosci. Methods*, vol. 178, pp. 219–227, 2009.
- [8] Y. Takeda, K. Yamanaka, and Y. Yamamoto, "Temporal decomposition of EEG during a simple reaction time task into stimulus- and response-locked components," *NeuroImage*, vol. 39, pp. 742–754, 2008.
- [9] D. O. Walter, "A posteriori "Wiener filtering" of average evoked responses," *Electroencephalogr. Clin. Neurophysiol.*, vol. 27, pp. 61–70, 1969.
- [10] D. J. Doyle, "Some comments on the use of Wiener filtering for the estimation of evoked potentials," *Electroencephalogr. Clin. Neurophysiol.*, vol. 38, pp. 533–534, 1975.
- [11] T. Wang, Ö. Özdamar, J. Bohórquez, Q. Shen, and M. Cheour, "Wiener filter deconvolution of overlapping evoked potentials," *J. Neurosci. Methods*, vol. 158, pp. 260–270, 2006.
- [12] P. C. Hansen, "Regularization tools version 4.0 for Matlab 7.3," *Numer. Algorithms*, vol. 46, pp. 189–194, 2007.
- [13] T. Lunttila, J. Nenonen, and E. Somersalo, "Regularization in cardiac source imaging," *FIMH, LNCS*, vol. 2674, pp. 101–111, 2003.
- [14] C. R. Johnson, "The generalized inverse problem in electrocardiography," in *Proc. 12th Ann. Intl. Conf. IEEE Eng. Med. Biol. Soc.*, Nov. 1990, pp. 593–594.
- [15] P. R. Johnston and R. M. Gulrajani, "A new method for regularization parameter determination in the inverse problem of electrocardiography," *IEEE Trans. Biomed. Eng.*, vol. 44, pp. 19–39, Jan. 1997.
- [16] G. H. Golub, M. Heath, and G. Wahba, "Generalized cross-validation as a method for choosing a good ridge parameter," *Technometric*, vol. 21, pp. 215–223, 1979.
- [17] P. R. Johnston and R. M. Gulrajani, "Selecting the corner in the L-curve approach to Tikhonov regularization," *IEEE Trans. Biomed. Eng.*, vol. 47, pp. 1293–1296, Sep. 2000.
- [18] B. Messnarz, B. Tilg, R. Modre, G. Fischer, and F. Hanser, "A new spatiotemporal regularization approach for reconstruction of cardiac transmembrane potential patterns," *IEEE Trans. Biomed. Eng.*, vol. 51, pp. 273–281, Feb. 2004.
- [19] D. Yao, "A method to standardize a reference of scalp EEG recording to a point at infinity," *Physiol. Meas.*, vol. 22, pp. 693–711, 2001.
- [20] S. Krieger, J. Timmer, S. Lis, and H. M. Olbrich, "Some considerations on estimating event-related brain signals," *J. Neural Transm. GenSect.*, vol. 99, pp. 103–129, 1995.
- [21] A. Delorme, G. Rousset, M. Mace, and M. Fabre-Thorpe, "Interaction of bottom-up and top-down processing in the fast visual analysis of natural scenes," *Cogn. Brain Res.*, vol. 19, pp. 103–113, 2004.
- [22] A. J. Fallgatter and W. K. Strik, "The NoGo-anteriorization as a neurophysiological standard-index for cognitive response control," *Int. J. Psychophysiol.*, vol. 32, pp. 233–238, 1999.
- [23] E. Jodo and Y. Kayama, "Relation of a negative ERP component to response inhibition in a Go/No-go task," *Electroencephalogr. Clin. Neurophysiol.*, vol. 82, pp. 477–482, 1999.
- [24] T. P. Jung, S. Makeig, M. Westerfield, J. Townsend, E. Courchesne, and T. J. Sejnowski, "Analysis and visualization of single-trial event-related potentials," *Human Brain Mapping*, vol. 14, pp. 166–185, 2001.
- [25] C. D. McGillen, J. I. Aunon, and K. Yu, "Signal and noise in evoked brain potentials," *IEEE Trans. Biomed. Eng.*, vol. 32, pp. 1012–1016, 1985.
- [26] P. Jaskowski and R. Verleger, "Amplitudes and latencies of single-trial ERP's estimated by a maximum-likelihood method," *IEEE Trans. Biomed. Eng.*, vol. 46, no. 8, pp. 987–993, 1999.
- [27] W. Truccolo, K. H. Knuth, A. Shah, C. Schroeder, S. L. Bressler, and M. Ding, "Estimation of single-trial multi-component ERPs: Differentially variable component analysis (dVCA)," *Biol. Cybern.*, vol. 89, pp. 426–438, 2003.
- [28] C. D. Woody, "Characterisation of an adaptive filter for the analysis of variable latency neuroelectric signals," *Med. Biol. Eng.*, vol. 5, pp. 539–553, 1967.



Gang Yin was born in Sichuan Province, China, in 1973. He received the B.S. and M.S. degrees in 1997 and 2004, respectively, from the University of Electronic Science and Technology of China (UESTC), Chengdu, China, where he obtained the Ph.D. degree with the School of Life Science and Technology in 2010.

He is currently with Sichuan Cancer Hospital, Chengdu, China. His current research interests include event-related potential (ERP) signal.

Jun Zhang received the B.Sc. degree in theoretical physics from Fudan University in 1985, and Ph.D. degree in neurobiology from the University of California, Berkeley, in 1992.

Supported by a fellowship from McDonnell-Pew Foundation, he worked briefly at the Computation Neuroscience Laboratory headed by Terrance Sejnowski at the Salk Institute. He then joined the faculty of the Department of Psychology, University of Michigan, as an Assistant Professor, Associate Professor, and Full Professor (1992–now). He has also held visiting positions at the University of Melbourne in Australia, the University of Waterloo in Canada, RIKEN Brain Science Institute in Japan, and the Institute of Biophysics, Chinese Academy of Sciences in China. He has published over 48 peer-reviewed journal papers in diverse disciplines such as vision, mathematical psychology, cognitive psychology, game theory, voting and social choice, information geometry, machine learning, etc.

Currently, Dr. Zhang is serving as an Associate Editor for the *Journal of Mathematical Psychology*, an executive Board Member of the Society for Mathematical Psychology, and was in the past the President of the Society for Mathematical Psychology. His research has been funded by the National Science Foundation (NSF) and Air Force Office of Scientific Research (AFOSR). During 2007–2010, he served as a Program Manager in the Directorate of Mathematics, Information and Life Sciences at AFOSR.