

Nonparametric Inference of Hemodynamic Response for Multi-Subject fMRI Data under Multi-Stimulus Design

Tingting Zhang^{a,1}, Fan Li^b, Lane Beckes^a, James A. Coan^a

^aUniversity of Virginia

^bDuke University

Abstract

Many existing methods for estimating the hemodynamic response functions (HRF) break down when analyzing multi-subject fMRI data from multi-stimulus designs due to the large and inhomogeneous variances across subjects. Within the context of the General Linear Model (GLM), we propose two new nonparametric estimators of the HRF for fMRI data arising from such designs. We first introduce a kernel-smoothed nonparametric estimator, based on which we conduct hypothesis tests to identify task relevant neural activity. To cope with the inherent large data variance, we further impose a Tikhonov regularization to the kernel-smoothed HRF estimator, on top of which a bias-correction procedure is introduced by utilizing the multi-subject averaged HRF. A fast algorithm with linear computational time is also developed to select optimal regularization and smoothing parameters. We apply these methods to the fMRI data collected under a psychology design employing the Monetary Incentive Delay (MID) task. The proposed method successfully identified regions of neural activation related to affective responses to monetary incentives and penalties that were not identified using the traditional GLM approach. Moreover, the HRF estimates reveal that the magnitude of the penalty-related neural activity is significantly correlated with an individual state anxiety measure.

Keywords: general linear model, hemodynamic response function, kernel smoothing, regularization.

¹Corresponding author at: Halsey Hall 106, University of Virginia, Charlottesville, VA, 22904, USA. Fax: (434) 924-3076. E-mail address: tz3b@virginia.edu (Tingting Zhang)

1. Introduction

There is a vast literature on estimating the hemodynamic response function (HRF) from functional magnetic resonance imaging (fMRI) data within the framework of the General Linear Model (GLM) (Friston et al., 1995a; Friston et al., 1995b; Worsley and Friston, 1995). Most of the existing methods focus on analyzing the fMRI data in a specific brain region with a high signal-to-noise ratio under single-stimulus designs. Multi-stimulus designs are common in fMRI experiments, however, especially in psychological studies. For example, in the Monetary Incentive Delay (MID) task (Knutson et al., 2000) that motivates this analysis, the experimental fMRI paradigm consists of 6 stimuli corresponding to three different monetary outcomes (more details in Section 4). Analysis of multi-subject fMRI data from such multi-stimulus designs is challenging due to (1) low signal-to-noise ratio, (2) large and inhomogeneous variances across subjects, (3) subject-varying low-frequency drift error, and (4) possibly overwhelming computational cost due to the massive size of these kinds of data sets. To address these issues, we propose two new nonparametric HRF estimators in this article. The resulting estimates can be easily utilized to detect the distinction between brain responses to different stimuli and to model the relationship between subjects' brain functions with their individual biological or behavior traits.

A range of parametric and nonparametric approaches have been proposed in the literature to estimate HRFs. Standard parametric methods rely on assuming *a priori* a functional form of the HRF with a number of free parameters, such as the canonical form of mixtures of gamma functions (Worsley and Friston, 1995), poisson function (Friston et al. 1994), inverse logit function (Lindquist and Wager, 2007; Lindquist, 2008), and radial basis functions (Riera et al., 2004). Parametric models with just one or two parameters usually lack enough flexibility to accommodate heterogeneous brain responses across different brain regions, outer stimuli and subjects. On the other hand, estimation based on parametric models with an even moderate number of parameters can be computationally time-consuming and unstable. By contrast, nonparametric approaches are usually com-

putationally fast. One such example is the most flexible smooth finite impulse response (SFIR) method (Glover, 1999; Goutte et al., 2000; Ollinger et al., 2001) which treats the HRF at every time point as a free parameter and exploits a regularization term to obtain a smooth estimate. However, the SFIR has several complications when analyzing fMRI data from multi-stimulus designs: (1) selection for the optimal regularization parameter involves computationally intensive algorithms, (2) the resulting estimate has a large bias and can be sensitive to the parameter choice when data are noisy, and (3) the SFIR regularization incurs unequal biases towards estimating HRFs under distinct stimuli, so they may not be suitable for comparing brain responses to multiple stimuli. A third common class of approaches are based on representing the HRFs by a set of basis functions (e.g., Aguirre et al., 1998; Woolrich et al., 2004; Zarahn 2002). The main challenge of such methods lies in finding appropriate functional bases to represent the HRF parsimoniously, especially for data with weak signals. For example, we have analyzed the aforementioned MID data using several standard bases such as the sine and cosine transform bases (Christensen 1994,1999; Ashburner and Friston 1999), the spline bases, and the Gaussian model (Rajapakse et al., 1998); our analyses revealed that the chosen number of bases based on the common criteria such as mean square error (MSE) and the Bayesian information criterion (BIC) tends to be comparable to the number of time units that the HRF covers. As such, the HRF estimate is almost identical to the SFIR estimate with little or no regularization, which often has an unnaturally high frequency of noise. A popular approach based on functional bases is to represent the HRF as a linear combination of a canonical HRF and its derivatives with respect to time and dispersion (Friston et al., 1998; Henson et al., 2001; Henson et al., 2002), but it works well only when the underlying HRF is not far from the canonical form.

To overcome the difficulties of the existing methods in analyzing multi-subject fMRI data from multi-stimulus designs, we introduce a kernel-smoothed nonparametric estimator whose value at each time point is a weighted average of the ordinary least square (OLS) estimates of the HRF at neighboring time points. The new estimator has several

advantages: (1) the drift error term can be easily removed by conducting OLS regression first, (2) by “borrowing” information across neighboring data points, the kernel-smoothed estimator has a much smaller variance than the OLS estimator yet without conceding any model flexibility, (3) simple test statistics can be constructed for population-wide exploratory analysis, such as identifying responsive brain regions and comparing brain responses to different stimuli in any regions, and (4) a fast algorithm with linear computational time available. We then propose a Tikhonov regularization to the kernel-smoothed estimator to further reduce variance and stabilize the estimation. A fast algorithm is developed for selecting the optimal regularization and smoothing parameters in terms of minimizing MSE. Our method draws comparison to the nonparametric SFIR method that also executes smoothing and regularization. Distinct from the SFIR, we exploit the averaged HRF from multi-subject data to conduct a bias-correction, which is one of the main thrusts of this article. The common practice of inferring the HRFs of one subject solely from this subject’s data may be inefficient and unstable as the signal from single-subject data is usually weak. By gaining efficiency from the subjects-averaged HRF, the proposed bias-corrected estimator demonstrates significant improvement over existing approaches.

The rest of the article is organized as follows. In Section 2, we first briefly review the GLM framework and propose the nonparametric kernel-smoothed estimator for hypothesis testing on brain activity. Then we refine the estimator by adding the Tikhonov regularization and applying our bias-correction approach. In Section 3, we describe fast algorithms for conducting parameter selection. In Section 4, we then describe an application of our proposed analytic approach using data collected during the MID task. Finally, Section 5 concludes the paper with a discussion of our results.

2. Statistical Theory

2.1. The GLM

We conduct voxel-wise or massive-univariate analysis of the fMRI data in the context of the GLM. Since the same approach applies to each voxel, the subscript for voxel is

omitted in the article. Let $y_i(t)$ for $t = 1, \dots, T$ and $i = 1, \dots, N$ be the fMRI time series for a pre-specified voxel of subject i , where T is the total observation time and N is the number of subjects. Suppose the design has K stimuli. Let $v_{i,k}(t)$ be the k th ($k = 1, \dots, K$) stimulus function for subject i with $v_{i,k}(t) = 1$ if the stimulus is evoked at time t and 0 otherwise. In real applications, the stimulus function $v_{i,k}(t)$ can vary across subjects. The GLM represents the observed fMRI time series as a convolution of the HRF and the stimuli: $y_i(t) = \sum_{k=1}^K \int_0^m h_{i,k}(u)v_{i,k}(t-u)du + \varepsilon_i(t)$, where $h_{i,k}$ is the voxel's HRF of subject i under stimulus k , m is a known positive constant beyond which the HRF equals zero, and $\varepsilon_i(t)$ is an identically distributed error term. Typically m takes a value between 20 and 30 seconds, and $\varepsilon_i(t)$ is assumed to follow the autoregressive model AR(1). In practice, the blood oxygen level dependent (BOLD) fMRI signal also contains a low-frequency drift due to physiological noise or subject motion (Smith et al., 1999; Brosch, et al., 2002; Luo and Puthusserypady, 2008); this can be modeled by adding a polynomial term of time t to the above GLM as

$$y_i(t) = d_{0,i} + d_{1,i} \cdot t + d_{2,i} \cdot t^2 + \sum_{k=1}^K \int_0^m h_{i,k}(u)v_{i,k}(t-u)du + \varepsilon_i(t), \quad (1)$$

where the drift parameters $d_{0,i}$, $d_{1,i}$, and $d_{2,i}$ are allowed to vary across subjects.

2.2. Kernel Smoothed Nonparametric Estimator

To allow flexible modeling of brain activities across different brain regions, stimuli and subjects, we adopt a nonparametric approach where the value of each HRF at every time unit is taken as a free parameter, similar to the SFIR method. For a fixed subject i , let $\mathbf{Y}_i = (y_i(1), \dots, y_i(T))'$ be the observed fMRI time series. Also denote the HRF values at each time unit under stimulus k by $\boldsymbol{\beta}_{i,k} = (\beta_{i,k}(1), \dots, \beta_{i,k}(m))'$, and let $\boldsymbol{\beta}_i = (\boldsymbol{\beta}'_{i,1}, \dots, \boldsymbol{\beta}'_{i,K})'$, where $\beta_{i,k}(t) = \int_{t-1}^t h_{i,k}(u)du$ in a block design or $\beta_{i,k}(t) = h_{i,k}(t)$ in an event design. Denoting all the coefficients in model (1) by $\boldsymbol{\eta}_i = (d_{0,i}, d_{1,i}, d_{2,i}, \boldsymbol{\beta}'_i)'$, then the GLM (1) can

be written in a matrix form as

$$\mathbf{Y}_i = \mathbf{X}_i \boldsymbol{\eta}_i + \boldsymbol{\varepsilon}_i, \quad (2)$$

where \mathbf{X}_i is the design matrix corresponding to the time covariates and the stimulus functions for subject i , and $\boldsymbol{\varepsilon}_i = (\varepsilon_1, \dots, \varepsilon_T)' \sim N(0, \sigma_i^2 \Gamma_i)$ with unknown variance σ_i^2 and correlation matrix Γ_i . We assume Γ_i to be an identity matrix for simplicity, and simulations showed that the proposed method is robust to this assumption. Since $h_{i,k}(t)$ is random across subjects, the coefficients β_i are also random. As a result, model (2) is in fact a linear random-effect model. For each subject, we can remove the drift term through OLS regression and obtain an unbiased OLS estimate of β_i , denoted by $\hat{\beta}_i = (\hat{\beta}_{i,1}(1), \dots, \hat{\beta}_{i,1}(m), \hat{\beta}_{i,2}(1), \dots, \hat{\beta}_{i,K}(m))'$. As noted in Goutte et al. (2000), $\hat{\beta}_i$ usually has an artificial high-frequency noise due to the large number of parameters under estimation and the experimental designs with interleaved stimuli and inter-stimulus intervals. This can be clearly seen from the simulation example in Figure 1(a). Therefore, smoothing techniques are often employed to reduce the unnatural ruggedness of the estimates. Previous approaches usually apply temporal smoothing directly to $y_i(t)$ to increase the statistical power for detection (e.g., Friston et al., 1994; Friston et al., 1995b; Worsley and Friston, 1995). Since our interest lies in estimating the HRF, and the degree of smoothness may vary across the HRFs under different stimuli, we propose to conduct kernel smoothing on the OLS estimates $\hat{\beta}_i$ instead. Specifically, we propose to use the Nadaraya-Watson kernel estimator given by

$$\tilde{\beta}_{i,k}(t) = \sum_{u=t-l}^{t+l} W_{t,u} \cdot \hat{\beta}_{i,k}(u), \quad \text{with } W_{t,u} = \frac{f(\frac{t-u}{h})/h}{\sum_{u=t-l}^{t+l} f(\frac{t-u}{h})/h}. \quad (3)$$

Here h is a pre-specified bandwidth controlling the degree of smoothing, $f(t)$ is a given symmetric density (kernel) function, and l is a pre-specified constant giving an upper bound on the number of data points used for the estimation. In this article, we let $f(t)$ be a standard Gaussian density and $l = m$. Existing results suggest that choice of these

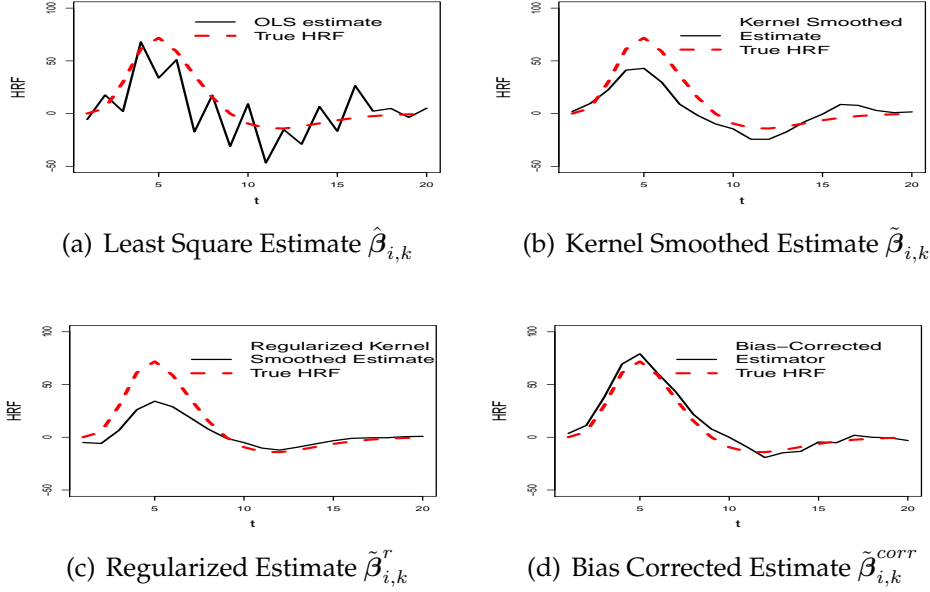


Figure 1: **Four Estimates of one random HRF** 1(a) The OLS estimate of the HRF with a randomly generate fMRI data. 1(b) The kernel-smoothed estimate using the selected h . 1(c) The Tikonov-regularized and smoothed estimate using the selected h and λ . 1(d) The bias-corrected estimate using the selected h and λ .

two values only has a small effect on the estimation (Eubank, 1988; Müller, 1988; Härdle, 1990; Wand and Jones, 1994; Fan and Gijbels, 1996). Choice of the key bandwidth parameter h is elaborated in Section 3. The underlying idea of kernel smoothing is to borrow information from the neighboring data, as the estimate $\tilde{\beta}_{i,k}(t)$ is a weighted average of the neighboring OLS estimates and the weight $W_{t,u}$ is negatively correlated with the distance $|u - t|$. The boundary condition of $\beta_{i,k}(t) = 0$ for $t < 0$ and $t > m$ is imposed by setting $\hat{\beta}_{i,k}(u) = 0$ for $u < 1$ and $u > m$ in the estimator (3). Letting $\hat{\beta}_{i,k} = (\hat{\beta}_{i,k}(1), \dots, \hat{\beta}_{i,k}(m))'$ and $\tilde{\beta}_{i,k} = (\tilde{\beta}_{i,k}(1), \dots, \tilde{\beta}_{i,k}(m))'$, the kernel estimator can be easily rewritten as a matrix transformation of the OLS estimators as follows:

$$\tilde{\beta}_{i,k} = \mathbf{B}_h \hat{\beta}_{i,k} \quad \text{and} \quad \tilde{\beta}_i = \mathbf{A}_h \hat{\beta}_i, \quad (4)$$

where \mathbf{B}_h is an $m \times m$ matrix such that $\mathbf{B}_h(t, u_1) = W_{t,u_1}$ for $u_1, t = 1, \dots, m$, $\mathbf{A}_h = \mathbf{I}_K \otimes \mathbf{B}_h$ with \otimes denoting the Kronecker product, and \mathbf{I}_K is a $K \times K$ identity matrix. Figure 1(b)

demonstrates the improvement of the $\tilde{\beta}_{i,k}$ over the OLS estimate (Figure 1(a)).

2.3. Tikhonov Regularized Smoothed Estimator with Bias Correction

The kernel estimator $\tilde{\beta}_i$ is designed to increase the temporal continuity of the HRF estimate. However, large variation in the magnitude of $\tilde{\beta}_{i,k}(t)$ across time may still remain, especially when the signal-to-noise ratio is low, as is often the case in fMRI data. Regularization is a common technique to prevent overfitting and reduce variation of regression estimates. Here we propose to insert the widely-used Tikhonov regularization to $\tilde{\beta}_i$ to further reduce volatility of the kernel estimates. Let \mathbf{R}_λ^i be the lower $(K \times m)$ -by- $(K \times m)$ square sub-matrix of $(\mathbf{X}'_i \mathbf{X}_i + \lambda \cdot \mathbf{D}_1)^{-1} (\mathbf{X}'_i \mathbf{X}_i)$, where \mathbf{D}_1 is a $(3 + K \times m)$ -by- $(3 + K \times m)$ diagonal matrix whose first 3 diagonal entries equal zero and the rest equal 1, and λ is a given positive constant. The new estimator is given by

$$\hat{\beta}_i^r = \mathbf{A}_h \mathbf{R}_\lambda^i \hat{\beta}_i. \quad (5)$$

Calculation of $\hat{\beta}_i^r$ is in fact equivalent to a two-step procedure: first, perform ridge regression on η_i in model (2) with a regularization term $\lambda \cdot \mathbf{D}_1$; second, conduct smoothing on the ridge estimate of β_i . The role of the regularization matrix \mathbf{R}_λ^i is to shrink $\hat{\beta}_i$ to a smaller norm. We let the first 3 diagonal elements of \mathbf{D}_1 equal zero to avoid shrinkage on the estimate of drift parameters. The ridge estimate of β_i can still be rough, so an additional smoothing step is executed on top of it. Here, we let $\tilde{\beta}_{i,k}^r$ be the sub-vector of $\hat{\beta}_i^r$ corresponding to $\beta_{i,k}$. As demonstrated in Figure 1(c), $\tilde{\beta}_{i,k}^r$ further smooths out the remaining small "rugged" parts in $\tilde{\beta}_{i,k}$.

Tikhonov regularization and kernel smoothing greatly reduce the variances in estimating the HRFs, but can also lead to large biases without further adjustment. Therefore a bias correction step is necessary. It is easy to show from equation (5) that the bias of $\hat{\beta}_i^r$ equals $[\mathbf{A}_h \mathbf{R}_\lambda^i - \mathbf{I}_{K \cdot m}] \beta_i$ and depends on the underlying true β_i . A close approximation to β_i is usually unavailable in real applications, but when there are multiple subjects, one can use the sample average to approximate β_i and conduct bias correction. Specifically,

we first calculate the average of $\hat{\beta}_i$ across all subjects, $\bar{\beta}$, and then apply the smoothing matrix \mathbf{A}_{h_0} to $\bar{\beta}$ as in (4) to obtain a smoothed population-averaged HRF estimate, denoted by $\tilde{\beta}_{\cdot(h_0)}$. The initial smoothing bandwidth h_0 is chosen to be $1/\sqrt{7/\text{TR}}$ as suggested by Goutte et al. (2000), where TR is the repetition time unit of the experiment. Since h_0 is comparably small, $\tilde{\beta}_{\cdot(h_0)}$ has a small variance and bias. Taking $\tilde{\beta}_{\cdot(h_0)}$ as an approximation to the true β_i for all subjects, we propose the following bias-corrected Tikhonov-regularized estimator:

$$\tilde{\beta}_i^{cor} = \tilde{\beta}_i^r - [\mathbf{A}_h \mathbf{R}_\lambda^i - \mathbf{I}_{K \cdot m}] \tilde{\beta}_{\cdot(h_0)}. \quad (6)$$

Utilization of the multiple-subject information through sample mean is the main improvement of estimator (6) over most of the existing methods that estimate the HRFs solely based on single-subject data. The simulation example in Figure 1(d) shows that the bias-corrected estimate almost exactly matches the underlying truth.

2.4. Kernel Estimate based Hypothesis Testing

In this section, we describe the use of the HRF estimates to perform two population-wide explorations of fMRI data using multi-stimulus designs. First, we identified brain voxels responsive to a specific stimulus. Next, we identified brain voxels that functioned differently in response to different stimuli. We formulate these two inference goals as two corresponding hypothesis tests for each voxel: (1) $H_0 : E(\beta_{i,k}) = \mu_k$ for some given stimulus k , where the expectation is with respect to all the subjects, and μ_k is a pre-specified vector of constants which is usually a zero vector; (2) $H_0 : E(\beta_{i,k}) = E(\beta_{i,k'})$ for $k \neq k'$.

Distinct from the existing tests on single aspect of the HRF, e.g., latency or magnitude, hypotheses (1) and (2) aim to detect any possible deviation from the null hypothesized HRF on the whole time domain, as they involve the entire vector of the HRF. We choose to construct test statistics based on the kernel-smoothed estimator $\tilde{\beta}_{i,k}$ proposed in Section 2.2 instead of the Tikhonov estimator $\tilde{\beta}_i^{cor}$. This is because $E(\tilde{\beta}_{i,k})$ equals $E(\tilde{\beta}_{i,k'})$ as long as $E(\beta_{i,k}) = E(\beta_{i,k'})$ for any fixed bandwidth h . However, this is not true for $\tilde{\beta}_i^{cor}$, because

regularization renders unequal biases in estimating different HRFs, leading to a test with an erroneous type I error. The same problem applies to the SFIR estimator, which also utilizes regularization. Conceptually the test statistics can be also constructed directly based on the OLS estimates $\hat{\beta}_{i,k}$. However, since the kernel-smoothed estimator $\tilde{\beta}_{i,k}$ has a much smaller variance, test-statistics based on the $\tilde{\beta}_{i,k}$ have a higher statistical power.

Due to the inhomogeneous variances across subjects, for the first hypothesis, we use $\mathbf{z}_{i,k} = (\tilde{\beta}_{i,k} - \mathbf{B}_h \boldsymbol{\mu}_k) / \hat{\sigma}_i$, $i = 1, \dots, N$, where $\hat{\sigma}_i$ is the OLS estimate of σ_i , to construct the Hotelling's T-square test statistic:

$$U_k = N \bar{\mathbf{z}}_k' \Lambda_k^{-1} \bar{\mathbf{z}}_k, \quad (7)$$

where $\bar{\mathbf{z}}_k$ and Λ_k are respectively the sample average and the sample variance-covariance matrix of $\mathbf{z}_{i,k}$. Given the type I error α , the null is rejected if $(N - m)U_k / (m(N - 1))$ is larger than $100(1 - \alpha)\%$ percentile of the F distribution with degrees of freedom $(m, N - m)$. Similarly for the second hypothesis, letting $\mathbf{Z}_{k,k'}^i = (\tilde{\beta}_{i,k} - \tilde{\beta}_{i,k'}) / \hat{\sigma}_i$, $i = 1, \dots, N$, we use the following test statistic

$$Q_{k,k'} = N \bar{\mathbf{Z}}_{k,k'}' \Omega_{k,k'}^{-1} \bar{\mathbf{Z}}_{k,k'}, \quad (8)$$

where $\bar{\mathbf{Z}}_{k,k'}$ and $\Omega_{k,k'}$ are the sample average and variance-covariance matrix of $\mathbf{Z}_{k,k'}^i$. The null hypothesis is rejected at α level if $(N - m)Q_{k,k'} / (m(N - 1))$ is larger than $100(1 - \alpha)\%$ percentile of the F distribution with degrees of freedom $(m, N - m)$.

3. Algorithms for Parameter Selection

Choosing proper values of the parameters h and λ that control the extent of smoothing and regularization is crucial in balancing variance and bias of the HRF estimates. However, it is also a nontrivial task computationally. Here we develop a fast algorithm for choosing h and λ with linear computational time - $O(T \times N \times K \times m)$ operations, for each voxel.

Typically, MSE is used as the criterion to choose the optimal parameters. In real ap-

plications, the σ_i^2 s are highly heteroscedastic across subjects, so it is more appropriate to weight each subject's fMRI data inversely proportional to σ_i^2 in the population-wide inference. Following this idea, we propose to select the bandwidth h based on minimizing the weighted MSE. For the kernel-smoothed estimator $\tilde{\beta}_{i,k}$, only bandwidth h needs to be selected. Denote $\sum_{i=1}^N E\|\tilde{\beta}_{i,k} - \beta_{i,k}\|^2/N$ by $\text{MSE}_k(h)$, where $\|\cdot\|$ is the L^2 norm, and define the weighted MSE (WMSE) as

$$\text{WMSE}_k(h) = \frac{1}{N} \sum_{i=1}^N E\|\tilde{\beta}_{i,k} - \beta_{i,k}\|^2/\sigma_i^2.$$

Since the true $\beta_{i,k}$'s are unknown, we would approximate $\text{WMSE}_k(h)$ for each candidate h and select the one that minimizes the weighted MSE estimate. We defer the technical derivations of $\text{MSE}_k(h)$ and $\text{WMSE}_k(h)$ to the Appendix and present the selection algorithm below.

Algorithm 1 for selecting bandwidth h for the kernel-smoothed estimator.

1. Starting from an initial bandwidth $h_0 = 1/\sqrt{7/\text{TR}}$, for each stimulus k , obtain $\tilde{\beta}_{i,k}$, $i = 1, \dots, N$, and calculate their average, denoted by $\tilde{\beta}_{\cdot k(h_0)}$.
2. For every subject i , calculate the OLS estimate $\hat{\sigma}_i^2$ of the variance σ_i^2 of the regression error in model (2). For each candidate h , calculate matrix \mathbf{B}_h and \mathbf{A}_h . Let

$$(\tau_{1,1}^i, \dots, \tau_{1,m}^i, \dots, \tau_{K,1}^i, \dots, \tau_{K,m}^i)' = \text{diag}(\mathbf{A}_h \Psi_i \mathbf{A}_h'),$$

where Ψ_i is the lower $(K \times m)$ -by- $(K \times m)$ square sub-matrix of $(\mathbf{X}_i' \mathbf{X}_i)^{-1}$, and $\text{diag}(\cdot)$ represents the diagonal vector of a square matrix.

3. For each candidate h , get an estimate of $\text{WMSE}_k(h)$ as

$$\widehat{\text{WMSE}}_k(h) = \sum_{i=1}^N \frac{1}{N} \left(\sum_{t=1}^m \tau_{k,t}^i \right) + \left(\frac{1}{N} \sum_{i=1}^N \frac{1}{\hat{\sigma}_i^2} \right) \tilde{\beta}'_{\cdot k(h_0)} (\mathbf{B}_h - \mathbf{I}_m)' (\mathbf{B}_h - \mathbf{I}_m) \tilde{\beta}_{\cdot k(h_0)}.$$

4. Choose the h that leads to the smallest $\widehat{\text{WMSE}}_k(h)$ for each k , or choose a universal h that minimizes $\sum_{k=1}^K \widehat{\text{WMSE}}_k(h)$.

For the Tikhonov regularized estimator $\tilde{\beta}_{i,k}^r$, both the bandwidth h and the regularization parameter λ need to be selected. Similar as above, the selection is based on minimizing the weighted MSE $\text{WMSE}_k(h, \lambda) = \sum_{i=1}^n E \|\tilde{\beta}_{i,k}^r - \beta_{i,k}\|^2 / (N\sigma_i^2)$ and the algorithm is presented below.

Algorithm 2 for selecting h, λ in the Tikhonov estimator .

1. Starting from the initial bandwidth $h_0 = 1/\sqrt{7/\text{TR}}$, compute $\tilde{\beta}_{\cdot,k(h_0)}$ as in Algorithm 1.
2. Estimate σ_i^2 in model (2) by its OLS estimate $\hat{\sigma}_i^2$ for each subject i .
3. For each candidate h , calculate matrix \mathbf{B}_h and \mathbf{A}_h . Let

$$(b_1^i(1), \dots, b_1^i(m), b_2^i(1), \dots, b_K^i(m))' = (\mathbf{A}_h \mathbf{R}_\lambda^i - \mathbf{I}_{K \cdot m}) \tilde{\beta}_{\cdot,k(h_0)},$$

$$\text{and } (\tau_{1,1}^i, \dots, \tau_{1,m}^i, \dots, \tau_{K,1}^i, \dots, \tau_{K,m}^i)' = \text{diag}(\mathbf{A}_h \mathbf{R}_\lambda^i \Psi_i (\mathbf{R}_\lambda^i)' \mathbf{A}_h').$$

4. For each combination of candidate (h, λ) , get an estimate of $\text{WMSE}_k(h, \lambda)$ as

$$\widehat{\text{WMSE}}_k(h, \lambda) = \sum_{i=1}^N \frac{1}{N} \sum_{t=1}^m \tau_{k,t}^i + \sum_{i=1}^N \frac{1}{N\hat{\sigma}_i^2} \sum_{t=1}^m (b_k^i(t))^2.$$

5. Choose the h and λ that minimize $\widehat{\text{WMSE}}_k(h, \lambda)$ for each stimulus k or select a universal pair (h, λ) that gives rise to the smallest $\sum_{k=1}^K \widehat{\text{WMSE}}_k(h, \lambda)$.

Note in the above two algorithms, for simplicity, the OLS estimates $\hat{\sigma}_i^2$ are calculated based on the assumption that the error terms ε_i are i.i.d.. Simulations confirm that our procedure is robust to this assumption.

4. Empirical Demonstrations

4.1. Monetary Incentive Delay Experiment

Subjects Nineteen subjects (10 male, 9 female) participated in exchange for financial payment (\$40.00 minimum, plus whatever money they managed to win during the study task). Subjects were recruited from a larger representative longitudinal community sample (e.g., Allen et al., 2007). All participants were between 22 and 25 years of age at the time of participation, 37% identified as black, and 63% identified as white.

Procedure We analyzed data collected from the 19 subjects using the MID task, which measures subjects' brain activity related to reward and penalty processing (Knutson et al., 2000). In the MID task, each participant completed a protocol comprised of 2 blocks of 72 6-s trials involving either no monetary outcome (control task), a potential reward (reward task), or a potential penalty (penalty task). During each trial, participants are shown a shape for 500 ms (anticipation condition), a variable interval delay of between 4000 and 4500 ms, and a white target square lasting between 160 and 260 ms (response condition). Participants are then instructed to respond with a button press. The cue shape (circle, square or triangle) shown at the start of each trial signals the type of the trial to be implemented, i.e., reward, penalty or no incentive respectively. Additionally, each reward and penalty shape included lines across the shape, which indicated the amount of money the participant could win or lose during the trial (i.e., 3 lines = \$5.00, 2 lines = \$1.00, and 1 line = \$0.20). Participants were also told that their reaction times to the white target would be recorded, and that receiving or preventing the monetary reward or punishment, respectively, depended on whether they responded within a given window of time. The order of trials in the protocol for each participant was randomized, with 25% of them control trials, 37.5% reward trials, and 37.5% punishment trials. In addition to the fMRI data, measures of each subject's state anxiety were collected using the state-trait anxiety inventory (Spielberger and Vagg, 1984).

fMRI Data Collection and Preprocessing Functional images were acquired using a Siemens 3.0 Tesla MAGNETOM Trio high-speed magnetic imaging device at UVA's Fontaine Re-

search Park, with a CP transmit/receive head coil with integrated mirror. Two hundred twenty four functional T2*-weighted Echo Planar images (EPIs) sensitive to BOLD contrast were collected per block, in volumes of twenty-eight 3.5-mm transversal echo-planar slices (1-mm slice gap) covering the whole brain (1-mm slice gap, TR=2000ms, TE=40ms, flip angle=90°, FOV= 192 mm, matrix= 64×64, voxel size= 3×3×3.5mm). Prior to collection of functional images, one hundred seventy-six high-resolution T1-magnetization-prepared rapid-acquisition gradient echo images were acquired to determine the localization of function (1-mm slices, TR=1900 ms, TE=2.53ms, flip angle= 9°, FOV=250mm, voxel size= 1×1×1mm). Data were preprocessed and analyzed using FMRIB's Software Library (FSL) software (Version 5.98; www.fmrib.ox.ac.uk/fsl, Worsley, 1994). Motion correction involved FMRIB's Linear Image Registration Tool, an intra-modal correction algorithm tool (MCFLIRT; Jenkinson et al., 2002), with slice scan-time correction and a high-pass filtering cutoff point of 100 seconds, removing irrelevant signals. We used BET (Smith, 2002) brain extraction, eliminating non-brain material voxels in the fMRI data, and a 5-mm full width at half minimum Gaussian kernel for smoothing. Images were registered to the Montreal Neurological Institute (MNI) space by FLIRT (Jenkinson et al., 2002). Regions of interest (ROI's) were determined structurally using the Harvard sub-cortical brain atlas, and were chosen for their likely involvement in affective processing based on a variety of previous studies unto affective neural processes (e.g., Knutson, et al., 2000). The chosen regions to be analyzed were the right putamen, right caudate, right nucleus accumbens, right pallidum, and right amygdala.

Data Analysis We included 6 stimuli in the GLM for MID data: the 3 signal stimuli for three types of monetary outcomes and the corresponding 3 reaction stimuli that the participants are required to respond to. The 6 stimuli are respectively referred to below as neutral anticipation, reward anticipation, penalty anticipation, neutral response, reward response, and penalty response. We focus on three goals in MID data analysis: first, identifying the brain voxels responsive to each stimulus, especially those involving monetary outcomes; second, identifying the voxels that react differentially to monetary reward and

punishment stimuli; and third, modeling the relationship between subjects' brain functions related to reward and punishment processing measured by the fMRI data with self-reported state anxiety.

We conducted the kernel-estimate-based hypothesis testing in Section 2.4 to pursue the first two goals. Specifically, the kernel-smoothed test statistic (7) was calculated for each voxel. Comparisons were drawn with two existing methods with similar order of computation: the basis set method using a canonical HRF and its temporal derivative to represent HRFs (referred to as the canonical method hereafter) and the SFIR. For these two methods, we first estimate the HRF for each voxel and extract the magnitude values from the estimated HRF, and then perform t-tests on these magnitude estimates to evaluate the significance of the HRF. The significance levels are set at 5% in the following tests. Other parametric approaches were not compared due to the intensive computation involved. Among the voxel-wise kernel-estimate-based tests on responsiveness to 6 stimuli, the tests corresponding to reward response and penalty anticipation stimuli identified the most responsive voxels in right caudate and right putamen, while the other 4 did not detect significantly many. The analysis results of the other three regions-right nucleus accumbens, right pallidum, and right amygdala-showed a similar pattern, and are not elaborated here. To determine whether there were significant clusters of activation within each ROI, we employed a small volume correction using a Monte Carlo procedure via AFNI's AlphaSim. At correction thresholds of individual P-values $< .05$ and cluster-wise P-values $< .05$, the kernel-smoothed method produced significant results after small volume correction, but the canonical HRF did not (see Table 1). Figures 2(a) and 2(d) respectively show voxels in right caudate responsive to reward response and penalty anticipation cues. Similarly, responsive voxels in right putamen are displayed in Figures 2(b) and 2(e). By contrast, few voxels in the right caudate were detected using the canonical method, as shown in Figures 2(c) and 2(f). In fact, using the canonical method, the neutral response stimulus was associated with the largest numbers of significant voxels both in the right caudate and right putamen. The testing results based on SFIR are omitted here,

because it incorrectly identified all the voxels to be responsive to all the stimuli.

To identify the brain voxels that react differentially to the monetary and neutral stimuli, we conducted tests on 6 pairs of HRFs, i.e., neutral vs. reward, neutral vs. penalty, and reward vs. penalty respectively for anticipatory and response stimuli. Using the kernel-smoothed test statistic (8), two pairs of HRF comparisons have identified significantly many voxels that possibly function differently under different stimuli based on expected false discovery rates. These pairwise tests include: (1) neutral anticipation and penalty anticipation, and (2) reward anticipation and penalty anticipation. The corresponding significant voxels in right caudate are respectively shown in Figures 3(a), and 3(b). As a contrast, significant voxels identified for the same two comparisons through canonical method are displayed in Figures 3(c), and 3(d). Only the kernel-smoothed-HRF test on the contrast reward anticipation versus penalty anticipation in the right caudate was significant after clusterwise small volume correction (see Table 1), and the canonical method could not identify significantly many voxels that respond differently to different anticipatory stimuli. However, the kernel-smoothed HRF was more sensitive in the contrast tests on anticipatory stimuli in the selected regions. The comparison results for the other three regions were similar, and are omitted here. Comparing the two methods, the kernel-based test (8) could potentially have more power than the canonical approach, possibly because the underlying HRF is different from the canonical form and the sample size under study is small.

To explore the connection between reward-processing brain activities and subjective experience, measured by the state anxiety score, we conducted a multiple linear regression with the state anxiety as the dependent variable, and the 4 magnitude differences between HRFs of monetary and neutral stimuli as predictors. In this analysis, the bias-corrected estimator $\tilde{\beta}_{i,k}^{cor}$ is used to calculate the HRFs, and the magnitudes of HRF estimates are extracted for constructing predictors. For both right caudate and right putamen, we found that across the 4 predictors, the penalty anticipation stimulus corresponded with the largest number of significant voxels, while for the other three regions-right nu-

Table 1: Comparison of Hypothesis Testing through Kernel Method and Canonical Basis Method

Statistical Tests on HRF's	# of Significant Voxels, $p < .05$	# of Voxels in Largest Cluster	Centroid Coordinates			Approx. P-value for cluster
			x	y	z	
Nonparametric HRF's						
<i>Right Caudate</i>						
Penalty Anticipation	302	254	10	10	2	< .001
Reward Response	215	119	14	20	-4	.009
Reward vs. Penalty Anti. ¹	238	191	12	6	16	.006
Regression of State Anx. ²	141	64	18	18	-4	.11
<i>Right Putamen</i>						
Penalty Anticipation	325	249	22	4	-6	< .001
Reward Response	259	149	18	10	-4	.005
Reward vs. Penalty Anti. ¹	124	52	22	4	-6	.26
Regression of State Anx. ²	259	147	22	18	-2	.005
Canonical HRF's						
<i>Right Caudate</i>						
Penalty Anticipation	96	29	20	8	14	.58
Reward Response	37	11	8	46	12	.97
Reward vs. Penalty Anti. ¹	111	62	8	4	14	.13
Regression of State Anx. ²	43	22	12	20	8	.75
<i>Right Putamen</i>						
Penalty Anticipation	137	54	14	10	-8	.24
Reward Response	140	68	12	6	-12	.13
Reward vs. Penalty Anti. ¹	29	5	30	-24	10	.99
Regression of State Anx. ²	3	1	30	-4	6	1.0

¹ Voxels react differentially to reward anticipation stimulus versus penalty anticipation stimulus;

² voxels with difference in HRF magnitudes between penalty anticipation and neutral anticipation stimuli that are significantly predictive of the individual state anxiety score.

The top half of this table shows the location and size of significant clusters determined using the nonparametric HRF's and a small volume correction procedure within the right caudate and putamen. The bottom half shows the results of the same tests using the canonical HRF's.

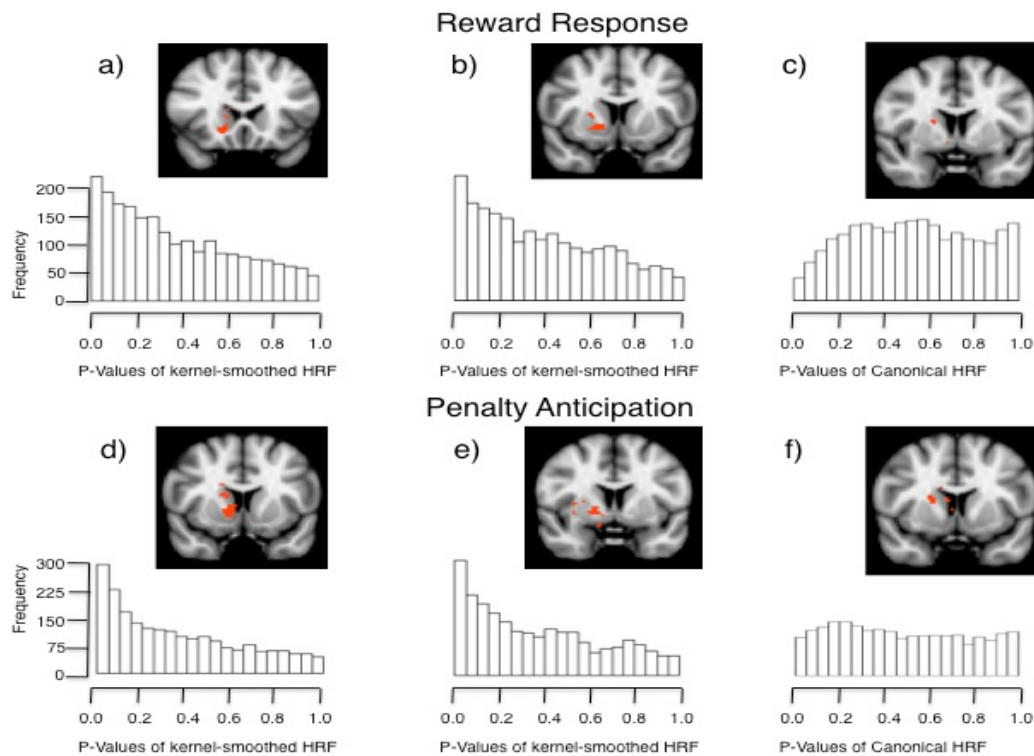


Figure 2: Tests on reward response and penalty anticipation stimuli 2(a),2(d) Tests based on kernel-smoothed HRF estimates for voxels in right caudate. 2(a),2(c) Tests based on kernel-smoothed HRF estimates for voxels in right putamen. 2(d),2(e) T-tests on magnitude estimates of HRFs through canonical method in right caudate.

cleus accumbens, right pallidum, and right amygdala—there was no significant result for any of the 4 predictors either by the canonical method or by the new proposed method. Within the right putamen, the regression of state anxiety on penalty minus neutral anticipation was significant even after clusterwise small volume correction (see Table 1). Figures 4(a) and 4(b) respectively show the significant voxels in the right caudate and right putamen related to the penalty anticipation stimulus. In addition, we also found that state anxiety is positively correlated with penalty stimuli, and negatively correlated with reward stimuli. By contrast, regression analyses with predictors constructed through the canonical method identified no significant clusters associated with any of the predictors. **Summary** These analyses suggest that the nonparametric HRF's employed in this analysis outperformed other ways of constructing the HRF, particularly the canonical method.

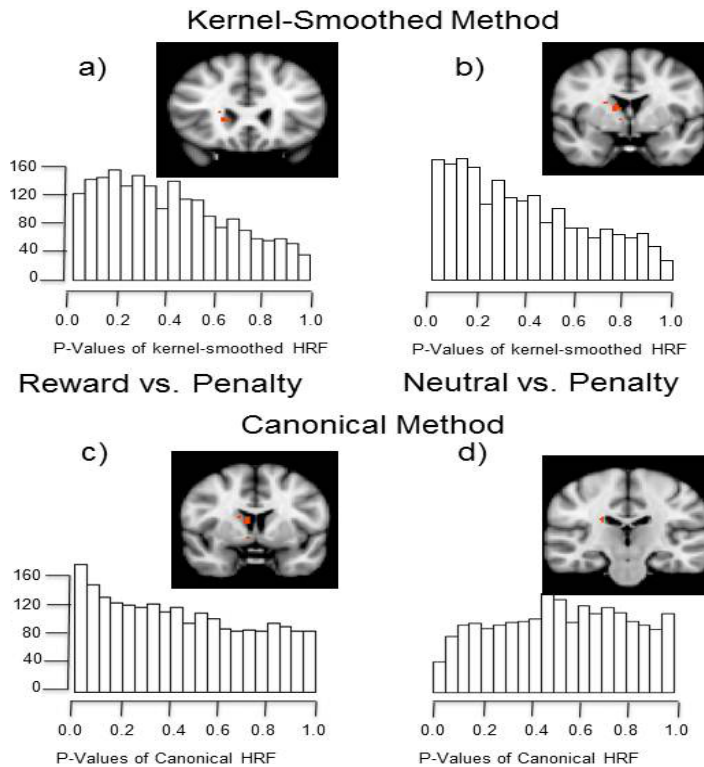


Figure 3: **Pairwise comparison of HRFs for voxels in right caudate** Figure 3(a) and 3(c) compare HRFs of neutral anticipation and penalty anticipation through kernel-smoothed and canonical methods respectively. 3(b) and 3(d) compare HRFs of reward anticipation and penalty anticipation through kernel-smoothed and canonical methods respectively.

Not only did the new HRF methods detect larger numbers of significant voxels in meaningful conditions and ROIs, it detected theoretically meaningful clusters of activation in the MID task whereas the canonical method did not. For example, the findings that the kernel-smoothed HRF produced small volume corrected clusters of significant activation for the reward response and penalty anticipation conditions in both the right caudate and putamen is consonant with previous work finding that both these regions are sensitive to the motivational value of stimuli, and is in line with previous MID findings (Knutson et al., 2000; Bjork et al., 2004; Knutson et al., 2001). Further, the finding that penalty anticipation in the right putamen is significantly positively correlated (after small volume correction) with state anxiety provides further evidence that the HRF's utilized are picking up on meaningful variance in the data. Given that penalty anticipation should

Regression of State Anxiety on Penalty Minus Neutral Contrast in the Anticipation Condition

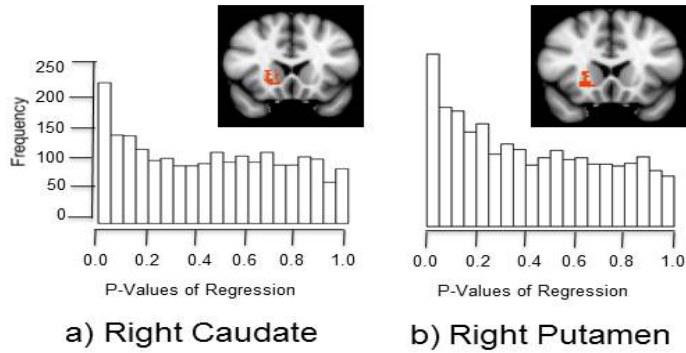


Figure 4: Voxels respectively in right caudate and right putamen whose HRF-magnitude difference under penalty anticipation and neutral anticipation stimuli are significantly correlated with subjects' state anxiety measures.

covary with state anxiety, we argue that this finding is particularly compelling in terms of demonstrating the utility of the novel HRF methods presented here.

5. Conclusions

In this article, we propose two new nonparametric HRF estimators within the GLM framework that provide faster computation, more flexible modeling of brain activities across different brain regions, stimuli and subjects, and most importantly, smaller estimation errors than the existing methods. The first kernel-smoothed estimator is developed to perform population-wide hypothesis tests on brain responses to stimuli. The second estimator is constructed by adding a Tikhonov regularization to the kernel-smoothed estimator, which further stabilizes the HRF estimate. On top of this estimator, a bias-correction step based on the multi-subject averaged HRF is proposed to rectify the possibly large bias induced by regularization and smoothing. There are two justifications for using the averaged HRF estimate: first, the HRFs of any specific voxel across different subjects are supposed to share a similar form, so the averaged HRF estimate would be informative for approximating such a functional form; second, each single subject's fMRI data has a low signal-to-noise ratio, and estimates solely based on such data may be unstable. The

bias-correction step enables borrowing information from the averaged data, and significantly increases estimation efficiency. Following this idea, we evaluate the estimation error of the two estimators by utilizing the averaged data, and propose fast algorithms with linear computational time for selecting the optimal regularization and smoothing parameters. Finally, the application of our methods to a real fMRI study demonstrates the improvement in power and efficiency in estimating the HRF.

Though powerful in estimating the magnitude of the HRF, the regularized smoothed estimator, similar to most nonparametric approaches, is not delicate enough to evaluate brain regions' reaction time or any measure of brain activities on the natural time scale instead of the experimental TR scale. This is because nonparametric approaches treat the HRF at each time unit as independent parameters, intrinsically ignoring their temporal correlation on the natural time scale. Kernel-smoothing can improve temporal continuity of the estimates but only in the scale of experimental repetition time unit, which is much larger than brain reaction time. Improvement over existing nonparametric methods on estimating brain reaction and duration time will be a major focus for our future investigation.

6. Appendix

We show the derivation of the MSE of our nonparametric estimator here. It is easy to see that

$$E(\tilde{\beta}_{i,k}(t) - \beta_{i,k}(t))^2 = E\left(\text{Var}(\tilde{\beta}_{i,k}(t)|i)\right) + E\left(E(\tilde{\beta}_{i,k}(t)|i) - \beta_{i,k}(t)\right)^2.$$

As the OLS estimate $\hat{\boldsymbol{\eta}}_i = \boldsymbol{\eta}_i + (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i \boldsymbol{\epsilon}_i$, we have

$$E(\hat{\boldsymbol{\beta}}_i|i) = \boldsymbol{\beta}_i \quad \text{and} \quad \text{Var}(\hat{\boldsymbol{\eta}}_i|i) = \sigma_i^2 (\mathbf{X}'_i \mathbf{X}_i)^{-1} \mathbf{X}'_i \boldsymbol{\Gamma}_i \mathbf{X}_i (\mathbf{X}'_i \mathbf{X}_i)^{-1}.$$

With $\Gamma_i = \mathbf{I}_{K \times (m+3)}$, $Var(\hat{\boldsymbol{\eta}}_i|i) = \sigma_i^2(\mathbf{X}'_i\mathbf{X}_i)^{-1}$. Denote $Var(\hat{\boldsymbol{\beta}}_i|i)$ by $\sigma_i^2\Psi_i$, which is the sub-matrix of $Var(\hat{\boldsymbol{\eta}}_i|i)$ corresponding to $\hat{\boldsymbol{\beta}}_i$. Then we have

$$\tilde{\boldsymbol{\beta}}_i = \mathbf{A}_h\hat{\boldsymbol{\beta}}_i, \quad E(\tilde{\boldsymbol{\beta}}_i|i) = \mathbf{A}_h\boldsymbol{\beta}_i \quad \text{and} \quad Var(\tilde{\boldsymbol{\beta}}_i|i) = \sigma_i^2\mathbf{A}_h\Psi_i\mathbf{A}'_h.$$

For estimating the bias $E(E(\tilde{\boldsymbol{\beta}}_{i,k}(t)|i) - \beta_{i,k}(t))^2$, since $E(\tilde{\boldsymbol{\beta}}_{i,k}|i) - \boldsymbol{\beta}_{i,k} = (\mathbf{B}_h - \mathbf{I}_m)\boldsymbol{\beta}_{i,k}$, then

$$\|E(\tilde{\boldsymbol{\beta}}_{i,k}|i) - \boldsymbol{\beta}_{i,k}\|^2 = \boldsymbol{\beta}'_{i,k}(\mathbf{B}_h - \mathbf{I}_m)'(\mathbf{B}_h - \mathbf{I}_m)\boldsymbol{\beta}_{i,k}.$$

Then $MSE_k(h) = \sum_{i=1}^N \sigma_i^2 \text{diag}(\mathbf{A}_h\Psi_i\mathbf{A}'_h)/N + \sum_{i=1}^N \boldsymbol{\beta}'_{i,k}(\mathbf{B}_h - \mathbf{I}_m)'(\mathbf{B}_h - \mathbf{I}_m)\boldsymbol{\beta}_{i,k}/N$. In practical approximation, σ_i^2 above is estimated by its OLS $\hat{\sigma}_i^2$, and $\boldsymbol{\beta}_{i,k}$ are all replaced with their sample average $\tilde{\boldsymbol{\beta}}_{\cdot k(h_0)}$.

We first look at the Tikhonov regularized/ridge regression estimate of $\boldsymbol{\eta}_i$ as

$$\hat{\boldsymbol{\eta}}_i^T = (\mathbf{X}'_i\mathbf{X}_i + \lambda \cdot \mathbf{D}_1)^{-1}\mathbf{X}_i\mathbf{Y}_i = (\mathbf{X}'_i\mathbf{X}_i + \lambda \cdot \mathbf{D}_1)^{-1}(\mathbf{X}'_i\mathbf{X}_i)\hat{\boldsymbol{\eta}}_i.$$

The regularized smoothed estimator $\tilde{\boldsymbol{\beta}}_i^r$ is essentially the kernel-smoothed $\hat{\boldsymbol{\beta}}_i^T$, which is the sub-vector of $\hat{\boldsymbol{\eta}}_i^T$ corresponding to $\boldsymbol{\beta}_i$. The reason for $\tilde{\boldsymbol{\beta}}_i^r$ equalling $\mathbf{A}_h\mathbf{R}_\lambda^i\hat{\boldsymbol{\beta}}_i$ is because the lower left $(K \times m)$ -by-3 sub-matrix of $(\mathbf{X}'_i\mathbf{X}_i + \lambda \cdot \mathbf{D}_1)^{-1}(\mathbf{X}'_i\mathbf{X}_i)$ equals zero. Then it is easy to see $\tilde{\boldsymbol{\beta}}_i^r$ is still linear of the OLS estimate, we can easily get its bias and variance as

$$E(\tilde{\boldsymbol{\beta}}_i^r|i) = \mathbf{A}_h\mathbf{R}_\lambda^i\boldsymbol{\beta}_i \quad \text{and} \quad Var(\tilde{\boldsymbol{\beta}}_i^r|i) = \sigma_i^2\mathbf{A}_h\mathbf{R}_\lambda^i\Psi_i(\mathbf{R}_\lambda^i)'\mathbf{A}'_h.$$

Based on above, the MSE of $\tilde{\boldsymbol{\beta}}_i^r$ can be derived similarly as above.

Acknowledgments

This project was funded by a grant issued by the National Institute of Mental Health. The project described was supported by Award Number R01MH080725 to J.A. Coan. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute of Mental Health or the National Institutes of Health.

We acknowledge the support of Karen Hasselmo, Alexander Tatum, Zoe Englander

and Casey Brown.

References

- [1] Aguirre, G.K., Zarahn, E. and D'Esposito, M. (1998). The variability of human, BOLD hemodynamic responses. *NeuroImage* **8**, 360-369.
- [2] Allen, J. P., Porter, M., McFarland, F. C., McElhaney, K. B., Marsh, P. (2007). The relation of attachment security to adolescents' paternal and peer relationships, depression, and externalizing behavior. *Child Development* **78**, 1222-1239.
- [3] Ashburner J. and Friston K. J. (1999). Nonlinear spatial normalization using basis functions. *Human Brain Mapping* **7**, 254-66.
- [4] Bjork, J. M., Knutson, B., Fong, G. W., Caggiano, D. M., Bennett, S. M., and Hommer, D. (2004). Incentive-elicited brain activation in adolescents: Similarities and differences from young adults. *Journal of Neuroscience* **24**, 1793-1802.
- [5] Brosch, J., Talavage, T., Ulmer, J., and Nyenhuis, J. (2002). Simulation of human respiration in fMRI with a mechanical model. *IEEE Transactions on Biomedical Engineering* **49**, 700-707.
- [6] Christensen, G. E. (1994) Deformable shape models for anatomy. Doctoral thesis, Washington University, Sever Institute of Technology.
- [7] Christensen, G. E. (1999) Consistent linear elastic transformations for image matching. In *Proc Information Processing in Medical Imaging (IPMI)*, Kuba A, Sámal M, Todd-Pokropek A (eds), vol. 1613 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin and Heidelberg, 224-37.
- [8] Eubank, R.L.(1988). *Nonparametric Regression and Spline Smoothing*. Marcel Dekker, New York.

- [9] Fan, J. and Gijbels, I. (1996). *Local Polynomial Modelling and Its Applications*. Chapman and Hall, London.
- [10] Friston, K. J., Fletcher, P., Josephs, O., Holmes, A., Rugg, M. D., and Turner, R. (1998). Event-related fMRI: characterizing Differential Responses. *NeuroImage* **7**, 30-40.
- [11] Friston, K.J., Jezzard, P.J. and Turner R. (1994). Analysis of functional MRI time-series. *Human Brain Mapping* **1**, 153-171.
- [12] Friston, K.J., Holmes, A.P., Poline, J.B., Grasby, P.J., Williams, S.C., Frackowiak, R.S. and Turner, R. (1995a). Analysis of fMRI time-series revisited. *NeuroImage* **2**, 45-53.
- [13] Friston, K.J., Holmes, A.P., Worsley, K., Poline, P.J., Frith, C. and Frackowiak, R. (1995b). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping* **2**, 189-210.
- [14] Glover, G.H. (1999). Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* **9**, 416-429.
- [15] Goutte, C., Nielsen, F. A., and Hansen, L. K. (2000). Modeling the haemodynamic response in fMRI using smooth FIR filters. *IEEE Transactions on Medical Imaging* **19**, 1188-1201.
- [16] Härdle, W. (1990). *Applied Nonparametric Regression*. Cambridge University Press.
- [17] Henson, R., Rugg, M., and Friston, K. J. (2001). The choice of basis functions in event-related fMRI. *NeuroImage* **13**, 149.
- [18] Henson, R., Price, C. J., Rugg, M. D., Turner, R., and Friston, K. J. (2002). Detecting latency differences in event-related BOLD responses: application to words versus nonwords and initial versus repeated face presentations. *NeuroImage* **15**, 83-97.

- [19] Jenkinson, M., Bannister, P., Brady, M., and Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *NeuroImage* **17**, 825-41.
- [20] Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., and Hommer, D. (2001). Dissociation of reward anticipation and outcome with event-related fMRI. *Neuroreport* **12**(17), 3683-3687.
- [21] Knutson, B., Westdorp, A., Kaiser, E., and Hommer, D. (2000). FMRI visualization of brain activity during a monetary incentive delay task. *NeuroImage* **12**, 20-27.
- [22] Luo, H. and Puthusserypady, S. (2008). Analysis of fMRI Data With Drift: Modified General Linear Model and Bayesian Estimator. *IEEE Transactions on Biomedical Engineering* **55**, 1504-1511.
- [23] Lindquist, M.A. and Wager, T.D. (2007). Validity and power in hemodynamic response modeling: a comparison study and a new approach. *Human Brain Mapping* **28**, 764-784.
- [24] Lindquist, M.A. (2008). The statistical analysis of fMRI data. *Statistical Science*, **23**, 439-464.
- [25] Müller, H.G. (1988). *Nonparametric Regression Analysis of Longitudinal Data*. Springer, New York.
- [26] Ollinger, J.M., Corbetta, M., and Shulman, G.L. (2001). Separating processes within a trial in event-related functional MRI. *NeuroImage* **13** 218-229.
- [27] Rajapakse J. C., Kruggel F., Maisog J. M., von Cramon D. Y. (1998). Modeling hemodynamic response for analysis of functional MRI time-series. *Human Brain Mapping* **6** 283-300.

- [28] Riera, J.J., Watanabe, J., Kazuki, I., Naoki, M., Aubert, E., Ozaki, T., and Kawashima, R. (2004). A state-space model of the hemodynamic approach: Nonlinear filtering of bold signals. *NeuroImage* **21**, 547-567.
- [29] Smith, A., Lewis, B., Ruttinmann, U., et al. (1999). Investigation of low frequency drift in fMRI signal, *Neuroimage* **9**, 526-533.
- [30] Spielberger, C.D., and Vagg, P.R. (1984). Psychometric properties of the STAI: A reply to Ramanaiah, Franzen, and Schill. *Journal of Personality Assessment* **48**, 95-97.
- [31] Wand, M.P. and Jones, M.C. (1994). *Kernel Smoothing*. Chapman and Hall, London.
- [32] Woolrich, M.W., Behrens, T.E., and Smith, S.M. (2004). Constrained linear basis sets for HRF modelling using Variational Bayes. *NeuroImage* **21**, 1748-1761.
- [33] Worsley, K.J. (1994). Local maxima and the expected Euler characteristic of excursion sets of χ^2 , F and t fields. *Advanced Applied Probability* **26**, 13-42.
- [34] Worsley, K.J. and Friston, K.J. (1995). Analysis of fMRI time-series revisited again. *NeuroImage* **2**, 173-181.
- [35] Zarahn, E. (2002). Using larger dimensional signal subspaces to increase sensitivity in fMRI time series analyses. *Human Brain Mapping* **17**, 13-16.