

## C H A P T E R

## 14

## Convolution Models for fMRI

R. Henson and K. Friston

## INTRODUCTION

This chapter reviews issues specific to the analysis of functional magnetic resonance imaging (fMRI) data. It extends the general linear model (GLM) introduced in Chapter 8 to convolution models, in which the blood oxygenation-level-dependent (BOLD) signal is modelled by neuronal causes that are expressed via a haemodynamic response function (HRF). We begin by considering linear convolution models and introduce the concept of temporal basis functions. We then consider the related issues of temporal filtering and temporal autocorrelation. Finally, we extend the convolution model to include non-linear terms and conclude with some example analyses of fMRI data.

## THE HAEMODYNAMIC RESPONSE FUNCTION (HRF)

A typical BOLD response to a single, impulsive stimulation ('event') is shown in Figure 14.1. The response peaks approximately 5s after stimulation, and is followed by an undershoot that lasts as long as 30s (at high magnetic fields, an initial undershoot can sometimes be observed) (Malonek and Givald, 1996). Early event-related studies therefore used a long time between events (i.e. a long stimulus onset asynchrony (SOA)) to allow the response to return to baseline between stimulations. However, although the responses to successive events will overlap at shorter SOAs, this overlap can be modelled explicitly within the GLM via a convolution model and a haemodynamic response function (HRF), as described in below. Short SOAs of a few seconds are desirable because they are comparable to those typically used in behavioural and electrophysiological studies, and because they are

generally more efficient from a statistical perspective, as we will see in the next chapter.

The shape of the BOLD impulse response appears similar across early sensory regions, such as V1 (Boynton *et al.*, 1996), A1 (Josephs *et al.*, 1997) and S1 (Zarahn *et al.*, 1997)[\*\*14.1]. However, the precise shape has been shown to vary across the brain, particularly in higher cortical regions (Schacter *et al.*, 1997), presumably due mainly to variations in the vasculature of different regions (Lee *et al.*, 1995). Moreover, the BOLD response appears to vary considerably across people (Aguirre *et al.*, 1998).<sup>1</sup>

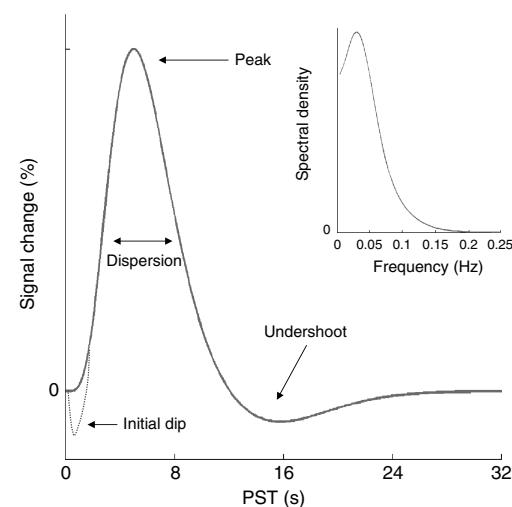


FIGURE 14.1 Typical (canonical) BOLD impulse response (power spectrum inset).

<sup>1</sup> This has prompted some to use subject-specific HRFs derived from a reference region known to respond to a specific task (e.g. from central sulcus during a simple manual task performed during a pilot scan on each subject (Aguirre *et al.*, 1998). However, while this allows for inter-subject variability, it does not allow for inter-regional variability within subjects (or potential error in estimation of the reference response).

These types of variability can be accommodated by expressing the HRF in terms of a set of temporal basis functions.

A linear convolution model assumes that successive responses summate (superpose). However, there is good evidence for non-linearity in the amplitude of the BOLD response, as a function of the stimulus duration or stimulus magnitude (e.g. Vasquez and Noll, 1998), and as a function of SOA (Pollman *et al.*, 1998; Friston *et al.*, 1998a; Miezin *et al.*, 2000). These non-linearities also appear to vary across different brain regions (Birn *et al.* 2001; Huettel and McCarthy, 2001). The non-linearity found as a function of SOA is typically a 'saturation' whereby the response to a run of events is smaller than would be predicted by the summation of responses to each event alone. This saturation is believed to arise in the mapping from blood flow to BOLD signal (Friston *et al.*, 2000a), though may also have a neuronal locus, particularly for very short SOAs or long stimulus durations (for biophysical models that incorporate such non-linearities, see Chapter 27). Saturation has been found for SOAs below approximately 8 s, and the degree of saturation increases as the SOA decreases. For typical SOAs of 2–4 s, however, its magnitude can be small (typically less than 20 per cent) (Miezin *et al.*, 2000). Later we will see how the linear convolution model is extended to handle such non-linearities via a Volterra expansion.

### Linear time-invariant (convolution) models

It is useful to treat a session of fMRI scans as a time-series. This is because the data tend to be correlated across successive scans, given that the typical measurement interval,  $T_R$ , of 1–3 s, is less than the duration of the BOLD response. The GLM can be expressed as a function of time (Friston *et al.*, 1994):

$$\begin{aligned} y(t) &= X(t)\beta + \varepsilon(t) \\ \varepsilon(t) &\sim N(0, \sigma^2 \Sigma) \end{aligned} \quad 14.1$$

where the data,  $y(t)$ , comprise the fMRI time series, the explanatory variables,  $X(t)$  are now functions of time,  $\beta$  are (time-invariant) parameters, and  $\Sigma$  is the noise auto-correlation. Though  $y(t)$  and  $X(t)$  are really discrete (sampled) time-series (normally represented by the vector  $y$  and design matrix  $X$  respectively), we will initially treat the data and model in terms of continuous time. For simplicity, we will consider the case of a single cause or regressor.

The explanatory variables  $X(t)$  represents the predicted BOLD time course arising from neuronal activity,  $u(t)$ , up to some scaling factor. This neuronal activity (e.g. the mean synaptic activity of an ensemble of neurons – see Chapter 32) is assumed to be caused by a sequence of

experimental manipulations and is usually referred to as the stimulus function. If we assume that the BOLD signal is the output of a linear time-invariant (LTI) system (Boynton *et al.*, 1996), i.e. that the BOLD response to a brief input has a finite duration and is independent of time, and that the responses to successive inputs superpose in a linear fashion, then we can express  $X(t)$  as the convolution of the stimulus function with an impulse response, or HRF,  $h(t)$ :

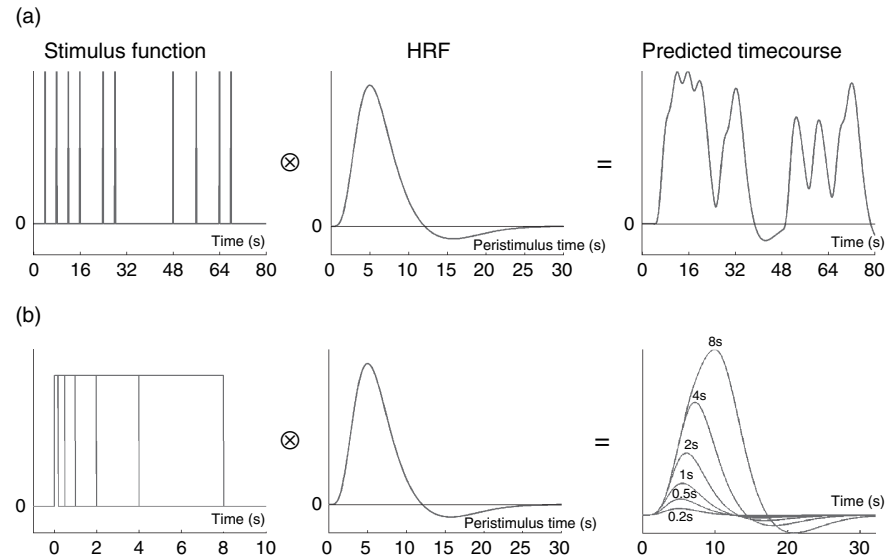
$$X(t) = u(t) \otimes h(t) = \int_0^T u(t - \tau)h(\tau)d\tau \quad 14.2$$

where  $\tau$  indexes the peristimulus time (PST), over which the BOLD impulse response is expressed. The HRF is equivalent to the first-order Volterra kernel described below. The stimulus function  $u(t)$  is usually a stick-function or boxcar function encoding the occurrence of an event or epoch. The result of convolving a random sequence of neuronal events with a 'canonical' HRF (see Figure 14.1) is shown in Figure 14.2(a). The smoothness of the resulting response is why the HRF is often viewed as a low-pass filter. The result of convolving more sustained periods of neuronal activity (called epochs in SPM) with the canonical HRF is shown in Figure 14.2(b). Note that the dominant effect of increasing the duration of neuronal activity, up to a few seconds, is to increase the peak amplitude of the BOLD response. In other words, the BOLD response integrates neuronal activity over a few seconds. The corollary is that a difference in the amplitude of the BOLD response (as tested conventionally) does not necessarily imply a difference in the mean level of neuronal activity: the difference could reflect different durations of neuronal activity at same mean level. One way to distinguish between these scenarios is to test for differences in the peak latency of the BOLD impulse response (Henson *et al.*, 2001a).

In practice, the convolution must be performed in discrete time. Given that significant information may exist in the predicted BOLD time course beyond that captured by typical  $T_R$ s of 1–3 s, SPM performs the convolution at a higher temporal resolution with  $N$  time points per scan (i.e. with resolution,  $\Delta t = T_R/N$  seconds). This means, for example, that stimulus onsets do not need to be synchronized with scans (they can be specified in fractions of scans).<sup>2</sup> To create the explanatory variables, the predicted BOLD time course is then down-sampled every  $T_R$  with reference to a specified time point  $T_0$  (Plate 8, see colour plate sections).

<sup>2</sup> In SPM, an 'event' is defined as having a duration of 0, which in practice corresponds to a single non-zero value for one time bin of duration  $\Delta t$ , where the value of the stimulus function is  $1/\Delta t$ . For epochs, the stimulus function is scaled so that it sums to one over a second.

**FIGURE 14.2** Linear convolution with a canonical HRF illustrated for (a) randomly presented events, and (b) epochs of neuronal activity with durations increasing from 200 ms to 16 s.



## TEMPORAL BASIS FUNCTIONS

Given that the precise shape of the HRF may vary over brain regions and over individuals, variability in its shape needs to be accommodated. The simplest way to achieve this within the GLM is via an expansion in terms of  $K$  temporal basis functions,  $f_k(\tau)$ :

$$h(\tau) = \sum_{k=1}^K \beta_k f_k(\tau) \quad 14.3$$

If the stimulation and resulting neuronal activity were a sequence of  $J$  impulses at times  $o_j$ , we can construct a stimulus stick-function:

$$u(t) = \sum_{j=1}^J \alpha_j \delta(t - o_j) \quad 14.4$$

where  $\delta(t)$  is the Dirac delta function. Note that variations in the stimulus – for example, its magnitude  $\alpha_j$  on each trial – can be accommodated by modulating the delta-functions, prior to convolution with the HRF. These are called ‘parametric modulations’ in SPM. We will see an example of this parametric modulation in the last section. It is quite common to use a series of modulated stick-functions to model a single event-type by using a polynomial or other expansions of  $\alpha_j$ . For simplicity, we will assume that  $\alpha_j = 1$  (i.e. a zeroth-order term).

Having specified the form of the stimulus and haemodynamic response functions in this way, the GLM equation in Eqn. 14.1 can be written:

$$y(t) = \sum_{j=1}^J \sum_{k=1}^K \beta_k f_k(t - o_j) + \varepsilon(t) \quad 14.5$$

where  $\beta_k$  are the parameters to be estimated. Several temporal basis sets are offered in SPM, though not all are true ‘basis’ sets in the (mathematical) sense that they span the space of all possible impulse response shapes (over the finite duration of that response), the term ‘basis’ is used to reflect the user’s assumption that the set of functions chosen capture BOLD impulse response shapes that occur in reality.

## FIR and Fourier sets

The most flexible basis sets are the finite impulse response (FIR) and Fourier basis sets, which make the least assumptions about the shape of the response. The FIR set consists of contiguous boxcar functions of PST, each lasting  $T/K_{\text{FIR}}$  seconds (see Plate 9(a)), where  $T$  is duration of the HRF. The Fourier set (see Plate 9(b)) consists of a constant and  $K_F$  sine and  $K_F$  cosine functions of harmonic periods  $T, T/2, \dots, T/K_F$  seconds (i.e.  $K = 2K_F + 1$  basis functions). Linear combinations of the (orthonormal) FIR or Fourier basis functions can capture any shape of response up to a specified timescale ( $T/K_{\text{FIR}}$  in the case of the FIR) or frequency ( $K_F/T$  in the case of the Fourier set).

### Relationship between FIR and ‘selective averaging’

In practice, there is little to choose between the FIR and Fourier sets. The Fourier set can be better suited when the sampling of peristimulus time (as determined by the relationship between the SOA and  $T_R$ ) is non-uniform, whereas the parameter estimates for the FIR functions have a more direct interpretation in terms of the ‘averaged’ peristimulus time histogram (PSTH). Indeed, in the

special case when  $T/K_{\text{FIR}} = T_R = \Delta t$ , the FIR functions can be considered as approximate delta-functions:

$$h(\tau) = \sum_{k=1}^K \beta_k \delta(\tau - k) \quad 14.6$$

where  $\tau$  ranges over post-stimulus scans. Then the model becomes (with some abuse of the delta-function):

$$\begin{aligned} y(t) &= \sum_{j=1}^J \sum_{k=1}^K \beta_k \delta(t - k - o_j) + \varepsilon(t) \Rightarrow \\ y &= X\beta + \varepsilon \\ X_{jk} &= \sum_{i=1}^I \delta(t - k - o_i) \end{aligned} \quad 14.7$$

For the special case of non-overlapping responses and independent and identically distributed (IID) error (i.e.,  $\Sigma = 1$ ), the maximum likelihood estimates of the FIR parameters are equivalent to the simple trial-averaged data (much like with ERPs):

$$\begin{aligned} \hat{\beta}_k &= \frac{1}{J} \sum_{j=1}^J y(o_j + k - 1) \Rightarrow \\ \beta &= \frac{1}{J} X^T y \end{aligned} \quad 14.8$$

Estimating the haemodynamic response like this has been called ‘selective averaging’ (Dale and Buckner, 1997). However, in practice, this estimator is biased and suboptimal because it requires the information matrix (called the ‘overlap correction matrix’ (Dale, 1999)) to be a leading diagonal matrix, i.e.  $X^T X = JI$ , in which case the ordinary least-squares estimates becomes the selective average:

$$\hat{\beta} = (X^T X)^{-1} X^T y = \frac{1}{J} X^T y \quad 14.9$$

With careful counterbalancing of different stimuli and the use of variable SOAs (e.g. via null events; see Chapter 15), this requirement can be met approximately. However, selective averaging as a procedure is redundant and represents a special case of the general deconvolution that obtains when simply inverting a linear convolution model. Selective averaging rests on undesirable and unnecessary constraints on the experimental design and is seldom used anymore.

## Gamma functions

More parsimonious basis sets can be chosen that are informed about the shape of the HRF. For example, since

the HRF is assumed to be bounded at zero for  $\tau \leq 0$  and  $\tau \geq T$ , the Fourier basis functions can also be windowed (e.g. by a Hanning window) within this range. An alternative is based on the gamma function:

$$f(t) = \left( \frac{t - o}{d} \right)^{p-1} \left( \frac{\exp(-(t - o)/d)}{d(p-1)!} \right)^{p-1} \quad 14.10$$

where  $o$  is the onset delay,  $d$  is the time-scaling, and  $p$  is an integer phase-delay (the peak delay is given by  $pd$ , and the dispersion by  $pd^2$ ). This function is bounded and positively skewed (unlike a Gaussian for example). A single gamma function has been shown to provide a reasonably good fit to the BOLD impulse response (Boynton *et al.*, 1996), though it lacks an undershoot (Fransson *et al.*, 1999; Glover, 1999). A set of gamma functions of increasing dispersions can be obtained by increasing  $p$  (see Plate 9(c)). In SPM, these functions (as with all basis functions) are orthogonalized with respect to one another. This set is more parsimonious, in that fewer functions are required to capture the typical range of BOLD impulse responses than required by Fourier or FIR sets. This precludes overfitting and reduces the model’s degrees of freedom, to provide more powerful tests.

## The ‘informed’ basis set and the canonical HRF

Another more parsimonious basis set, suggested by Friston *et al.* (1998b), is based on a ‘canonical HRF’ and its partial derivatives (see Plate 9(d)). The canonical HRF is a ‘typical’ BOLD impulse response characterized by two gamma functions, one modelling the peak and one modelling the undershoot. The canonical HRF is parameterized by a peak delay of 6 s and an undershoot delay of 16 s, with a peak-undershoot amplitude ratio of six; these values were derived from a principal component analysis of the data reported in Friston *et al.* (1998a). To allow for variations about the canonical form, the partial derivatives of the canonical HRF with respect to its delay and dispersion can be added as further basis functions. For example, if the real BOLD impulse response is shifted by a small amount in time  $\tau$ , then by the first-order Taylor expansion:

$$h(t + \tau) \approx h(t) + \tau h'(t) \quad 14.11$$

This is the same as adding a small amount of the temporal derivative of  $h'(t)$ . Thus, if  $h(t)$  and  $h'(t)$  are used as two basis functions in the GLM to estimate the parameters  $\beta_1$  and  $\beta_2$  respectively, then small changes in the latency of the response can be captured by the parameter estimate for the temporal derivative (more precisely,

$\tau \approx \hat{\beta}_2/\hat{\beta}_1$ ; see Henson *et al.*, 2002a; Liao *et al.*, 2002, for a fuller treatment). In SPM, the temporal derivative is created from the finite difference between a canonical HRF and a canonical HRF shifted by one second. Thus, using the temporal derivative as a further response can capture differences in the latency of the BOLD response up to plus or minus a second (beyond this range, the first-order Taylor approximation breaks down). A similar logic applies to the use of dispersion derivative to capture [small] differences in the duration of the peak response. Together, these three functions comprise SPM's 'informed' basis set, in that they are informed by the range of typical BOLD impulse response shapes observed. Subsequent work, using more biophysically informed models of the haemodynamic response, revealed that the informed set is almost identical to the principal components of variation, with respect to the parameters of the Balloon model described in Chapter 27.

The temporal derivatives of an assumed HRF can also be used to allow for differences in the acquisition times of different slices with echo-planar imaging (EPI) sequences, in order to address the so-called 'slice-timing problem' (see Chapter 15). The ability of the temporal derivative to capture these latency differences is appropriate for a  $T_R$  of up to 2s (after synchronizing the model with the slice acquired half way through each scan), assuming that the true BOLD impulse responses match the canonical HRF in all brain regions (i.e. all slices; Henson *et al.*, 1999).

## Other methods

Other methods for estimating the shape of the BOLD impulse response use non-linear (iterative) fitting techniques, beyond the GLM. These approaches are more powerful, but computationally more expensive. Various parameterizations have been used, such as a Gaussian function parameterized by amplitude, onset latency and dispersion (Rajapakse *et al.*, 1998), a gamma function parameterized by amplitude, onset latency and peak latency (Miezin *et al.*, 2000), or even SPM's canonical HRF, with the amplitude, onset latency and peak latency parameters free to vary (Henson and Rugg, 2001). A problem with unconstrained iterative fitting techniques is that the parameter estimates may not be optimal, because of local minima in the search space. Parameters that have correlated effects compound this problem (often requiring a re-parameterization into orthogonal components). One solution is to put priors on the parameters in a Bayesian estimation scheme (Chapter 34) in order to 'regularize' the solutions (see Gossel *et al.*, 2001, and Woolrich *et al.*, 2004, for other examples). Indeed, more

recent Bayesian methods not only provide posterior densities for HRF parameters, but also provide metrics of the 'goodness' of different HRF models, using Bayesian model evidence (Friston, 2002; Penny *et al.*, 2005).

## Which temporal basis set?

Inferences using multiple basis functions are made with  $F$ -contrasts (see Chapter 9). An example  $F$ -contrast that tests for any difference in the event-related response to two trial-types modelled by SPM's informed basis set is shown in Plate13(c). If the real response matches an assumed HRF, models using just that HRF are statistically more powerful (Ollinger *et al.*, 2001). In such cases,  $t$ -tests on the parameter estimate for the HRF can be interpreted directly in terms of the 'amplitude' of the response. However, when the real response differs appreciably from the assumed form, tests on the HRF parameter estimates are biased (and unmodelled structure will exist in the residuals). In such cases, the parameter estimate for the canonical HRF, for example, can no longer necessarily be interpreted in terms of amplitude. The addition of partial derivatives of the HRF can ameliorate this problem: the inclusion of a temporal derivative, for example, can reduce the residual error by capturing systematic delays relative to the assumed HRF. Nonetheless, for responses that differ by more than a second in their latency (i.e. when the first-order Taylor approximation fails), different canonical HRF parameters will be estimated even when the responses have identical peak amplitudes (Henson *et al.*, 2002a).<sup>3</sup>

An important empirical question then arises: how much variability exists in the BOLD impulse response? Henson *et al.* (2001) addressed this question for a dataset involving rapid motor responses to the brief presentations of faces across twelve subjects. By modelling the event-related response with a canonical HRF, its partial derivatives *and* an FIR basis set, the authors assessed the contribution of the different basis functions using a series of  $F$ -contrasts (that collapsed over subjects within a single first-level design matrix). Significant additional variability was captured by both the temporal derivative and dispersion derivative, confirming that different regions exhibit variations around the canonical form (see

<sup>3</sup>Note that the inclusion of the partial derivatives of SPM's canonical HRF does not necessarily affect the parameter estimate for the HRF itself, since the basis functions are orthogonalized (unless correlations between the regressors arise due to under-sampling by the  $T_R$ , or by temporal correlations between the onsets of events of different types). Thus, their inclusion does not necessarily affect second-level  $t$ -tests on the canonical HRF parameter estimate alone.

Plate 10(a)). Little additional variability was captured by the FIR basis set. This suggests that the canonical HRF and its two partial derivatives are sufficient to capture the majority of experimental variability (at least in regions that were activated in this task). The same conclusion was reached using a second-level model and the twelve parameter estimates of a (pure) FIR model; by testing *F*-contrasts that specify the 'null-space' of either the canonical HRF or the canonical HRF plus its partial derivatives. Significant variability was not captured by the canonical HRF alone but there was little significant variability that could not be captured once the two partial derivatives were added (see Plate 10(b)). The latter data and analyses can be downloaded from the SPM website (<http://www.fil.ion.ucl.ac.uk/spm/data>).

This sufficiency of the informed basis set may be specific to this dataset and reflect the fact that neuronal activity was reasonably well approximated by a delta function. It is unlikely to hold for more complex experimental trials, such as working memory trials where information must be maintained for several seconds (e.g. Ollinger *et al.*, 2001). Nonetheless, such trials may be better accommodated by more complex neuronal models. This usually entails using multiple stimulus functions for different components of each trial (e.g. onset, delay-period, offset, etc.) while still using an informed model for the HRF. This allows more direct inferences about stimulus, response and delay components of a trial for example (Zarahn, 2000). More generally, the question of which basis set and how many components to use becomes a problem of model selection that can be addressed simply using *F*-contrasts or Bayesian techniques (Penny *et al.*, 2005).

One issue arises when one wishes to use multiple basis functions to make inferences in second-level analyses (e.g. in 'random effects' analyses over subjects; see Chapter 12). Subject-specific contrast images created after fitting an FIR model in a first-level analysis could, for example, enter into a second-level model as a peristimulus time factor (differential *F*-contrasts which would correspond to a condition-by-time interaction in a conventional repeated-measures analysis of variance (ANOVA); Chapter 13). However, the parameter estimates are unlikely to be independent or identically distributed over subjects, violating the 'sphericity' assumption of univariate, parametric statistical tests (Chapter 10). This is one reason why researchers have tended to stick with *t*-tests on (contrasts of) the parameter estimate for a single canonical HRF at the second-level. This is at the expense of potentially missing response differences with a non-canonical form. One solution is to use multivariate tests (Henson *et al.*, 2000), though these are generally less sensitive (by virtue of making minimal assumptions about the data covariance) (Kiebel

*et al.*, 2004). Alternatively, restricted maximum likelihood (ReML) can be used to estimate the covariance components subtending any non-sphericity (Friston *et al.*, 2002; Chapter 22). In this case, one generally needs to model both unequal variances (given that different basis functions can have different scaling) and unequal covariances (given that parameter estimates for different basis functions are likely to be correlated across subjects). This allows one to make statistical inferences over multiple basis functions at the second-level, provided one is prepared to assume that the basic correlational structure of the error is invariant across 'activated' voxels (the 'pooling device'; see Chapter 10).

## TEMPORAL FILTERING AND AUTOCORRELATION

We can also view our time-series in terms of frequency components via the Fourier transform. A schematic of the power spectrum, typical of a subject at rest in the scanner, is shown in Plate 11(a). This 'noise' spectrum is dominated by low frequencies and has been characterized by a  $1/f$  form when expressed in amplitude (Zarahn *et al.*, 1997)[\*\*14.1]. The noise arises from physical sources, sometimes referred to as 'scanner drift' (e.g. slowly-varying changes in ambient temperature); from physiological sources (e.g. biorhythms, such as  $\sim 1$  Hz respiratory or  $\sim 0.25$  Hz cardiac cycles, which are aliased by the slower sampling rate); and from residual movement effects and their interaction with the static magnetic field (Turner *et al.*, 1998). When the subject is performing a task, signal components are added to this noise. For example, Plate 11(b) shows the approximate signal spectrum induced by a square-wave stimulation, with a duty cycle of 64 s. When averaging over all frequencies, this signal might be difficult to detect against background noise. However, by filtering the data with an appropriate high-pass filter (see Plate 11(c)), we can remove most of the noise. Ideally, the remaining noise spectrum would be flat (i.e. 'white' noise, with equal power at all frequencies).

### Highpass filtering

The choice of the highpass cut-off would ideally maximize the signal-to-noise ratio. However, we cannot distinguish signal from noise on the basis of the power spectrum alone. Usually, a cut-off period of approximately 128 s is used, based on observations that the noise becomes appreciable at frequencies below approximately 0.008 Hz (though this may vary considerably

across scanners and subjects). In other words, some loss of signal may be necessary to minimize noise. Experimental designs therefore try to avoid significant power at low frequencies (i.e. conditions to be contrasted should be presented too far apart in time; see Chapter 15).

In the time domain, a highpass filter can be implemented by a discrete cosine transform (DCT) with harmonic periods up to the cut-off. These basis functions can be made explicit as confounds in the design matrix  $X_0$  or they can be viewed as part of a filter matrix,  $S$  (as in current implementations of SPM).<sup>4</sup> This matrix is applied to both data and model:

$$\begin{aligned} y &= X\beta + X_0\beta_0 + \varepsilon & 14.12 \\ \Leftrightarrow \\ Sy &= SX\beta + S\varepsilon \\ S &= I - X_0X_0^+ \end{aligned}$$

The effect of applying a highpass filter to real data (taken from a 42s-epoch experiment; data available from the SPM website) is illustrated in Plate 11(d). Plate 11(e) shows the fitted responses after the filter  $S$  is applied to two boxcar models, one with and one without convolution with the HRF. The importance of convolving the neuronal model with an HRF is evident in the residuals (see Plate 11(f)); had the explanatory variables been directly equated with the stimulus function (or neuronal activity), significant temporal structure would remain in the residuals (e.g. as negative deviations at the start of each block, i.e. at higher frequency harmonics of the boxcar function).

## Temporal autocorrelations

There are various reasons why the noise component may not be white even after highpass filtering. These include unmodelled neuronal sources that have their own haemodynamic correlates. Because these components live in the same frequency range as the effects of interest, they cannot be removed by the highpass filter. These noise sources induce temporal correlation between the residual errors. Such autocorrelation is a special case of non-sphericity, which is treated more generally in Chapter 10. Here, we review briefly the various (historical) solutions to the specific problem of temporal autocorrelation in fMRI time-series (see Friston *et al.*, 2000b, for a fuller treatment).

<sup>4</sup> Though the matrix form expedites mathematical analysis, in practice highpass filtering is implemented by the computationally efficient subtraction  $Sy = y - X_0X_0^+y$ , where  $X_0$  is the matrix containing the DCT.

## Pre-colouring

One solution proposed by Worsley and Friston (1995) is to apply temporal smoothing. This is equivalent to adding a lowpass filter component to  $S$  (such that  $S$ , together with the highpass filter, becomes a 'bandpass' filter). If the time-constants of the smoothing kernel are sufficiently large, the temporal autocorrelation induced by the smoothing can be assumed to swamp any intrinsic autocorrelation,  $\Sigma$ , such that:

$$V = S\Sigma S^T \approx SS^T \quad 14.13$$

The effective degrees of freedom can then be calculated using the classical Satterthwaite correction (see Appendix 8.2):

$$\begin{aligned} v &= \frac{\text{tr}(RV)^2}{\text{tr}(RVRV)} & 14.14 \\ R &= I - SX(SX)^+ \end{aligned}$$

solely via knowledge of the filter matrix. Low-pass filters derived from a Gaussian smoothing kernel with full-width at half maximum (FWHM) of 4–6s, or derived from the canonical HRF (see Figure 14.1, inset), have been suggested (Friston *et al.*, 2000b).

## Pre-whitening

An alternative solution is to estimate the intrinsic autocorrelation directly, which can be used to create a filter to 'pre-whiten' the data before fitting the GLM. In other words, the smoothing matrix is set to  $S = K^{-1}$ , where  $KK^T = \Sigma$  is the estimated autocorrelation matrix. If the estimation is exact, then:

$$V = S\Sigma S^T = I \quad 14.15$$

All methods for estimating the autocorrelation rest on a model of its components. These include autoregressive (AR) models (Bullmore *et al.*, 1996) and 1/f models (Zarahn *et al.*, 1997)[\*\*14.1]. An AR(p) is a pth-order autoregressive model, having the time domain form:

$$\begin{aligned} z_t &= a_1z_{t-1} + a_2z_{t-2} + \dots + a_pz_{t-p} + w_t \Rightarrow \\ z &= Az + w \\ w_t &\sim N(0, \lambda_w) & 14.16 \\ A &= \begin{bmatrix} 0 & 0 & 0 & \dots \\ a_1 & 0 & 0 & \\ a_2 & a_1 & 0 & \\ \vdots & & & \ddots \end{bmatrix} \end{aligned}$$

where  $w_t$  is an IID innovation or Gaussian process and  $A$  is a lower-triangular matrix containing the coefficients

in its lower leading diagonals. The regression coefficients  $a_i$  can be estimated by ordinary least-squares. Several authors (e.g. Bullmore *et al.*, 1996; Kruggel and von Cramon, 1999) use an AR(1) model, in which the autoregression parameters are estimated from the residuals after fitting the GLM. These estimates are then used to create the filter  $S = (I - A)^{-1}$  that is applied to the data before re-fitting the GLM (a procedure that can be iterated until the residuals are white).

The 1/f model is a linear model with the frequency  $\omega$  domain form:

$$\begin{aligned} s(\omega) &= b_1/\omega + b_2 \\ g(\omega) &= |s(\omega)|^2 \end{aligned} \quad 14.17$$

where  $g\omega$  is the power spectrum, whose parameters,  $b_1$  and  $b_2$ , can be estimated from the Fourier-transformed data. The advantage of these pre-whitening methods is that they produce the most efficient parameter estimates, under Gaussian assumptions (corresponding to Gauss-Markov or minimum variance estimators). Temporal smoothing is generally less efficient because it removes high-frequency components, which may contain signal. The disadvantage of the temporal autocorrelation models is that they can produce biased parameter estimates if the autocorrelation is not estimated accurately (i.e. they do not necessarily produce 'minimum bias estimators').

Friston *et al.* (2000b) argued that the AR(1) and 1/f models are not sufficient to estimate the typical temporal autocorrelation in fMRI data. This is illustrated in Plate 12(a), which shows the power spectra and 'autocorrelation functions'<sup>5</sup> for the residuals of an event-related dataset (used below). It can be seen that the AR(1) model underpredicts the intermediate-range correlations, whereas the 1/f model overpredicts the long-range correlations. Such a mismatch between the assumed and intrinsic autocorrelation will bias the statistics produced by pre-whitening the data.<sup>6</sup> This mismatch can be ameliorated by combining band-pass filtering (see Plate 12(b)) and modelling the autocorrelation, in which case both models provide a reasonable fit (see Plate 12(c)). Indeed, highpass filtering alone (with an appropriate cut-off) is normally sufficient to allow either model to fit the remaining autocorrelation (Friston *et al.*, 2000b).

<sup>5</sup> An autocorrelation function plots the correlation,  $\rho(t)$ , as a function of 'lag',  $t=0\dots n-1$ , and is the Fourier transform of the power spectrum,  $g\omega$ .

<sup>6</sup> More complex models of the temporal autocorrelation have since been shown to minimize bias, such as Tukey tapers (Woolrich *et al.*, 2001) and autoregressive moving average (ARMA) models, a special case of the latter being an AR(1)+white noise model (Burock and Dale, 2000).

## Estimating non-sphericity hyperparameters

The estimation of the autocovariance parameters or *hyperparameters* described so far is based on the residuals of the time-series and represents rather *ad hoc* procedures. They are *ad hoc* and biased because they do not allow for uncertainty about the fitted components that are removed from the data to produce the residuals. In other words, they fail to account for the loss of degrees of freedom due to parameter estimation *per se*. Current implementations of SPM avoid this shortcoming by partitioning the data covariance (rather than the residuals) using restricted maximum likelihood. This removes the bias resulting from correlations among the residuals induced by removing modelled effects (Friston *et al.*, 2002; though there are ways of reducing this bias, Worsley *et al.*, 2002).

### Restricted maximum likelihood

Restricted maximum likelihood (ReML), allows simultaneous estimation of model parameters and hyperparameters, with proper partitioning of the effective degrees of freedom (see Chapter 22 for more details). ReML can be used with any temporal autocorrelation model. Friston *et al.* (2002) use an 'AR(1)+white noise' model (Purdon and Weisskoff, 1998) with an autoregressive error term,  $z_t$  and a white noise term  $e_t$ :

$$\begin{aligned} y_t &= X_t\beta + z_t + e_t \\ z_t &= a_1 z_{t-1} + w_t \\ e_t &\sim N(0, \lambda_e) \\ w_t &\sim N(0, \lambda_w) \end{aligned} \quad 14.18$$

The autocorrelation coefficient  $a_1 = \exp(-1)$  was fixed, leaving two unknown hyperparameters;  $\lambda_e$  and  $\lambda_w$ . The white-noise component contributes to the zero-lag autocorrelation, which allows the AR(1) model to capture better the shape of the autocorrelation at longer lags. Note that this approach still requires a highpass filter to provide accurate fits (see Plate 12(d)), though a subtle difference from the residual-based approaches is that the highpass filter is also treated as part of the complete model to be estimated, rather than a pre-whitening filter.

### Pooled hyperparameter estimates

Iterative schemes like ReML are computationally expensive when performed at every voxel. Furthermore, the hyperparameter estimates from a single voxel can be quite imprecise. An expedient solution to both these issues is to assume that the relative values of the hyperparameters  $\lambda$  are stationary over voxels. This allows the data to be pooled over voxels in order to estimate the



hyperparameters and implicitly  $\Sigma(\lambda)$  for all voxels considered, in a single iterative procedure (see Chapter 22 for details). The ensuing autocorrelation matrix  $\Sigma(\lambda)$  is extremely precise because thousands of voxel time-series have been used to estimate it. This means it can now be used to estimate the parameters in the usual way, assuming known non-sphericity. This ReML approach to modelling serial correlations or temporal non-sphericity retains the efficiency of pre-whitening approaches, properly accounts for the loss of degrees of freedom when estimating the parameters, and allows for spatial variability in the error variance. This obviates the need for temporal smoothing, a consequence particularly important for event-related designs, in which appreciable signal can exist at high frequencies.

It should be noted that if the temporal autocorrelation varies over voxels (Zarahn *et al.*, 1997)[\*\*14.1] pooling may not be appropriate. For example, serial correlations are thought to be higher in grey than white matter (Woolrich *et al.*, 2001). This can be accommodated by estimating voxel-specific hyperparameters with some spatial regularization (Worsley *et al.*, 2002). However, this means that different voxels can have different effective degrees of freedom, which complicates the application of random field theory (Chapter 17). The solution we prefer is to pool over a homogeneous subset of voxels that are likely to show the same serial correlations (e.g. all those that respond to the paradigm).

## NON-LINEAR CONVOLUTION MODELS

The convolution model assumed thus far has been based on a linear approximation, for which there is counter-evidence, e.g. for events close together in time (see above). To allow non-linearities, a generalized convolution model can be used. This is based on the Volterra expansion (Friston *et al.*, 1998a; Josephs and Henson, 1999), which can be thought of as a generalization of the Taylor series approximation to dynamic systems and has the form:

$$y(t) = h_0 + \int_{-\infty}^{\infty} h_1(\tau_1) \cdot u(t - \tau_1) \cdot d\tau_1 + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} h_2(\tau_1, \tau_2) \cdot u(t - \tau_1) \cdot u(t - \tau_2) \cdot d\tau_1 d\tau_2 + \dots$$

14.19

(with only terms up to second order shown here), where  $h_n$  is the  $n$ -th order Volterra kernel. This expansion can model any analytic time-invariant system, and is often

used where the state equations (e.g. biophysical model) determining that system are unknown. In the present context, we assume we have a 'causal' system with finite memory (i.e. the integrals run from 0 to  $T$ ) and that a second-order approximation is sufficient.

## Basis functions and generalized HRFs

Again, temporal basis functions can be used to model the Volterra kernels:

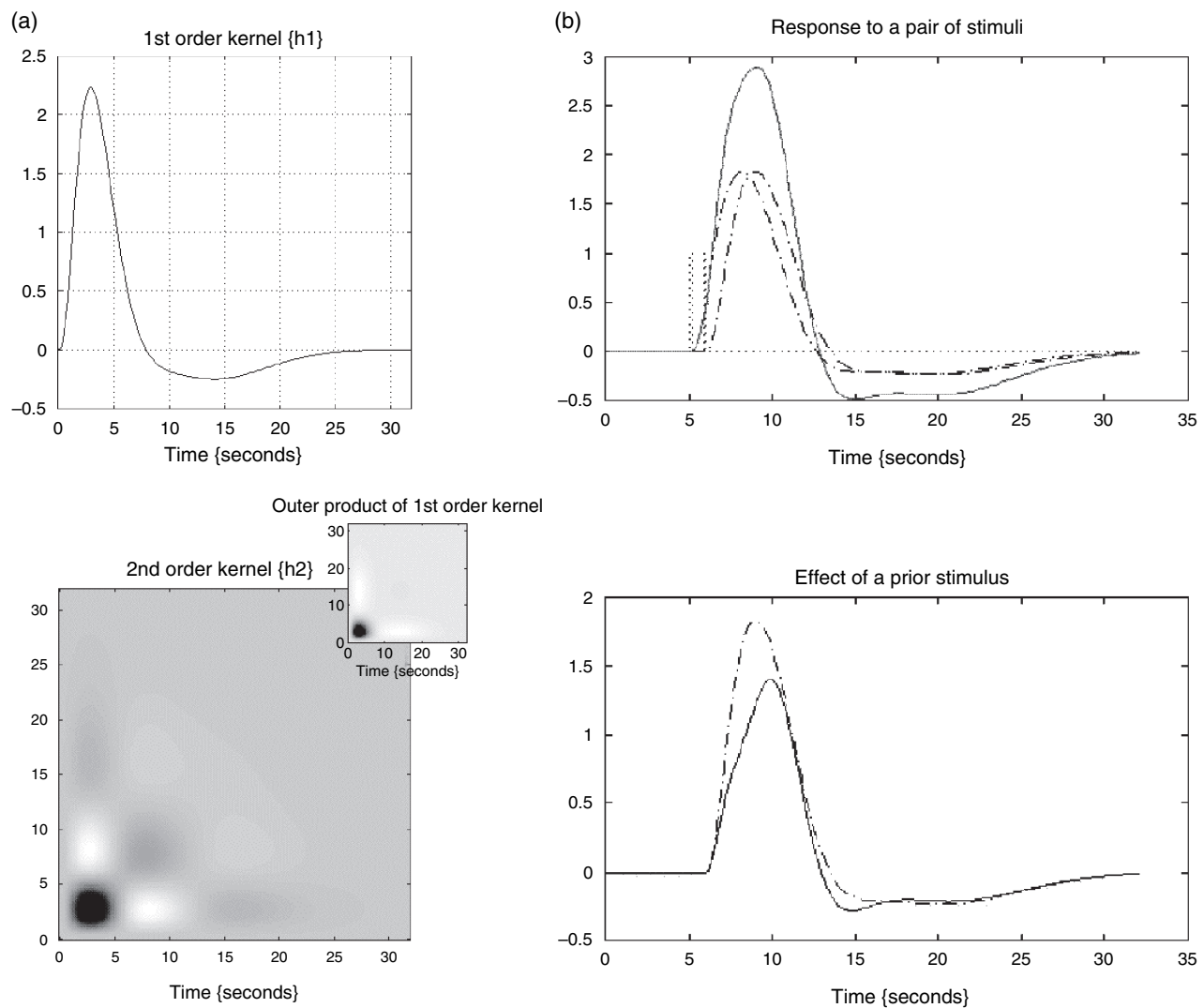
$$h_0 = f_0$$

$$h(\tau_1) = \sum_{k=1}^K \beta_k^{(1)} f_k(\tau_1) \quad 14.20$$

$$h(\tau_1, \tau_2) = \sum_{k=1}^K \sum_{l=1}^K \beta_{kl}^{(2)} f_k(\tau_1) f_l(\tau_2)$$

This allows us to express (linearize) the Volterra expansion within the GLM, with each basis function coefficient associated with a column of the design matrix. The regressors for the first-order coefficients  $\beta_k^{(1)}$  are simply the input convolved with each basis function in the usual way. The second-order coefficients  $\beta_{kl}^{(2)}$  have regressors that are the [Hadamard] products of the first-order regressors. Friston *et al.* (1998a) used three gamma functions, leading to three columns for the first-order kernel plus a further nine columns for the second-order kernel (to model quadratic non-linearities). Using fMRI data from an experiment in which words were presented at different rates,  $F$ -tests on the non-linear partition showed reliable effects in bilateral superior temporal regions. The estimated first and second-order kernels are shown in Figure 14.3(a). The first-order kernel (a linear combination of the three gamma functions) closely resembles the canonical HRF. The second-order kernel shows evidence of under-additivity (e.g. saturation) at short SOAs below 5 s (the dark region in the lower left), consistent with other studies (see above). Interestingly, evidence of super-additivity was also found for SOAs of approximately 8 s (the light regions between 5 and 10 s; the kernel is necessarily symmetric).

Using these first- and second-order kernels, the response to any temporal pattern of word presentations can be simulated. Using only the first-order kernel (i.e. a linear convolution model), the response to two words presented one second apart is simply the sum of the BOLD responses to each word alone (Figure 14.3(b), top panel). However, adding the second-order kernel shows the expected effect of saturation, whereby the response to the pair of events is less than the sum of their responses when presented alone (Figure 14.3(b), bottom panel). In



**FIGURE 14.3** Volterra kernels. (a) shows the first-order (upper) and second-order (lower) Volterra kernels from superior temporal cortex in an experiment in which auditory words were presented at different rates (see Friston *et al.*, 1998a for more details). The second-order kernel shows non-linearities, resulting in both underadditivity (dark regions) and superadditivity (light regions). (b) shows the response predicted for two stimuli 1 s apart when using a linear convolution model – i.e. the first-order kernel only (upper) – and when adding the second-order kernel from (a), resulting in a predicted saturation of the response relative to the linear case.

principle, this saturation could be caused by neuronal factors, blood-flow factors or even blood-oxygenation factors. However, the fact that a PET experiment, using the same paradigm, showed that blood-flow increased linearly with word presentation rate suggests that the dominant source of saturation in these fMRI data arose in the mapping between perfusion and BOLD signal. Indeed, using a detailed biophysical, ‘balloon’ model of the BOLD response, Friston *et al.* (2000a) proposed that the reason the second stimulus is compromised, in terms of elaborating a BOLD signal, is because of the venous pooling, and consequent dilution of deoxyhaemoglobin, incurred by the first stimulus. This means

that less deoxyhaemoglobin can be cleared for a given increase in flow. The second type of non-linearity – the superadditivity for events presented approximately 8 s apart – was attributed to the fact that, during the flow undershoot following a first stimulus, deoxyhaemoglobin concentration is greater than normal, thereby facilitating clearance of deoxyhaemoglobin following a second stimulus.

Although these non-linearities may be specific to this particular paradigm and auditory cortex, they do suggest caution in using event-related designs with very short SOAs. The saturation in particular provides important (and intuitive) limits on the statistical efficiency

of event-related designs as a function of SOA (see next chapter). Even if the significant non-linearities are small enough that SOAs below 5 s (but above 1 s) are still more efficient from the statistical perspective, one could consider adding a second-order Volterra kernel (linearized via a number of basis functions) in order to capture systematic, event-related variability in the residuals.

## A WORKED EXAMPLE

In this section, the concepts of this chapter are illustrated in a single-session event-related fMRI dataset from one of the 12 subjects reported in Henson *et al.* (2002b), and freely available from the SPM website <http://www.fil.ion.ucl.ac.uk/spm/data>. Events comprised 500 ms presentations of faces, to which the subject made a famous/non-famous decision with the index and middle fingers of their right hand. One half of the faces were famous, one half were novel (unfamiliar), and each face was presented twice during the session. This corresponds to a  $2 \times 2$  factorial design consisting of first and second presentations of novel and famous faces (conditions N1, N2, F1 and F2 respectively, each containing  $J = 26$  events). To these 104 events, 52 null events were added and the whole sequence permuted. This meant that the order of novel/famous faces was pseudo-randomized (given the finite sequence), though the order of first and second presentations, while inter-mixed, was constrained by the fact that second presentations were necessarily later than first presentations on average. The minimum SOA ( $SOA_{\min}$ ) was 4.5 s, but varied near-exponentially over multiples of  $SOA_{\min}$  due to the null events (see next chapter). The time series comprised 351 images acquired continuously with a  $T_R$  of 2 s. The images were realigned spatially, slice-time corrected to the middle slice, normalized with a bilinear interpolation to  $3 \times 3 \times 3$  mm voxels and smoothed with an isotropic Gaussian FWHM of 8 mm. The ratio of  $SOA_{\min}$  to  $T_R$  ensured an effective peristimulus sampling rate of 2 Hz.

Events were modelled with  $K = 3$  basis functions consisting of the canonical HRF, its temporal derivative and its dispersion derivative. The resolution of the simulated BOLD signal was set to 83 ms ( $N = 24$ ) and the event onsets synchronized with the middle slice ( $T_0 = 12$ ). Six user-specified regressors, derived from the rigid-body realignment parameters (3 translations and 3 rotations) were included to model residual (linear) movement

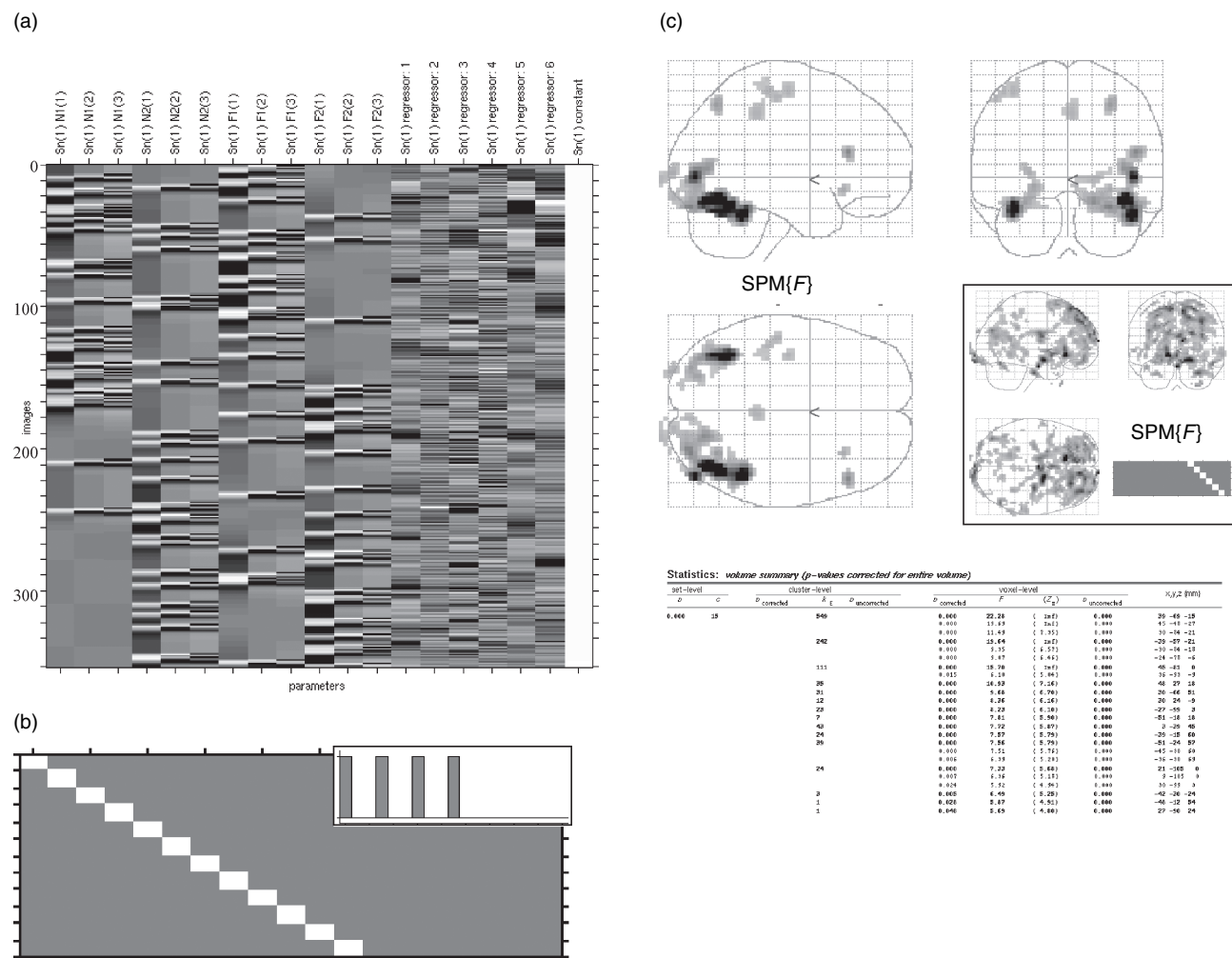
effects.<sup>7</sup> A highpass filter with cut-off period of 120 s was applied to both model and data, with an AR(1) model for temporal autocorrelations. No global scaling was used. Two different models are considered below: a 'categorical' one and a 'parametric' one. In the categorical model, each event-type is modelled separately. In the parametric model, a single event-type representing all face-trials is modulated by their familiarity and the 'lag' since their last presentation.

## Categorical model

The design matrix for the categorical model is shown in Figure 14.4(a). A (modified) effects-of-interest  $F$ -contrast, corresponding to a reduced  $F$ -test on the first 12 columns of the design matrix (i.e. removing linear movement effects), is shown in Figure 14.4(b) and the resulting SPM{ $F$ } in Figure 14.4(c). Several regions, most notably in bilateral posterior inferior temporal, lateral occipital, left motor and right prefrontal cortices, show some form of significant response to the events (versus baseline) at  $p < 0.05$ , corrected for whole brain. Note that these responses could be activations (positive amplitude) or deactivations (negative amplitude), and may differ across the event-types. A  $t$ -contrast like that inset in Figure 14.4(b) would test a more constrained hypothesis, namely that the response is positive when averaged across all event-types, and is a more powerful test for such responses (producing more suprathreshold voxels in this dataset). Also inset in Figure 14.4(c) is the SPM{ $F$ } from an  $F$ -contrast on the realignment parameters, in which movement effects can be seen at the edge of the brain.

The parameter estimates (plotting the modified effects-of-interest contrast) and best-fitting event-related responses for a right fusiform voxel (close to what has been called the 'Fusiform Face Area', Kanwisher *et al.*, 1997) are shown in Plate 13(a) and 13(b). First presentations of famous faces produced the greatest response (green fitted response). Furthermore, responses in this region appear to be slightly earlier and narrower than the canonical response (indicated by the positive

<sup>7</sup> One might also include the temporal derivatives of the realignment parameters, and higher-order interactions between them, in a Volterra approximation to residual movement effects (regardless of their cause). Note also that the (rare) events, for which the fame decision was erroneous, could be modelled as a separate event-type (since they may involve physiological changes that are not typical of face recognition). This was performed in the demonstration on the website, but is ignored here for simplicity.



**FIGURE 14.4** Categorical model: effects of interest. (a) Design matrix. (b)  $F$ -contrast for effects of interest (inset is  $t$ -contrast that tests for positive mean parameter estimate for canonical HRF). (c) SPM{ $F$ } MIP for effects of interest  $F$ -contrast, thresholded at  $p < 0.05$  whole-brain corrected, together with SPM tabulated output (inset is SPM{ $F$ } for contrast on movement parameters, also at  $p < 0.05$  corrected).

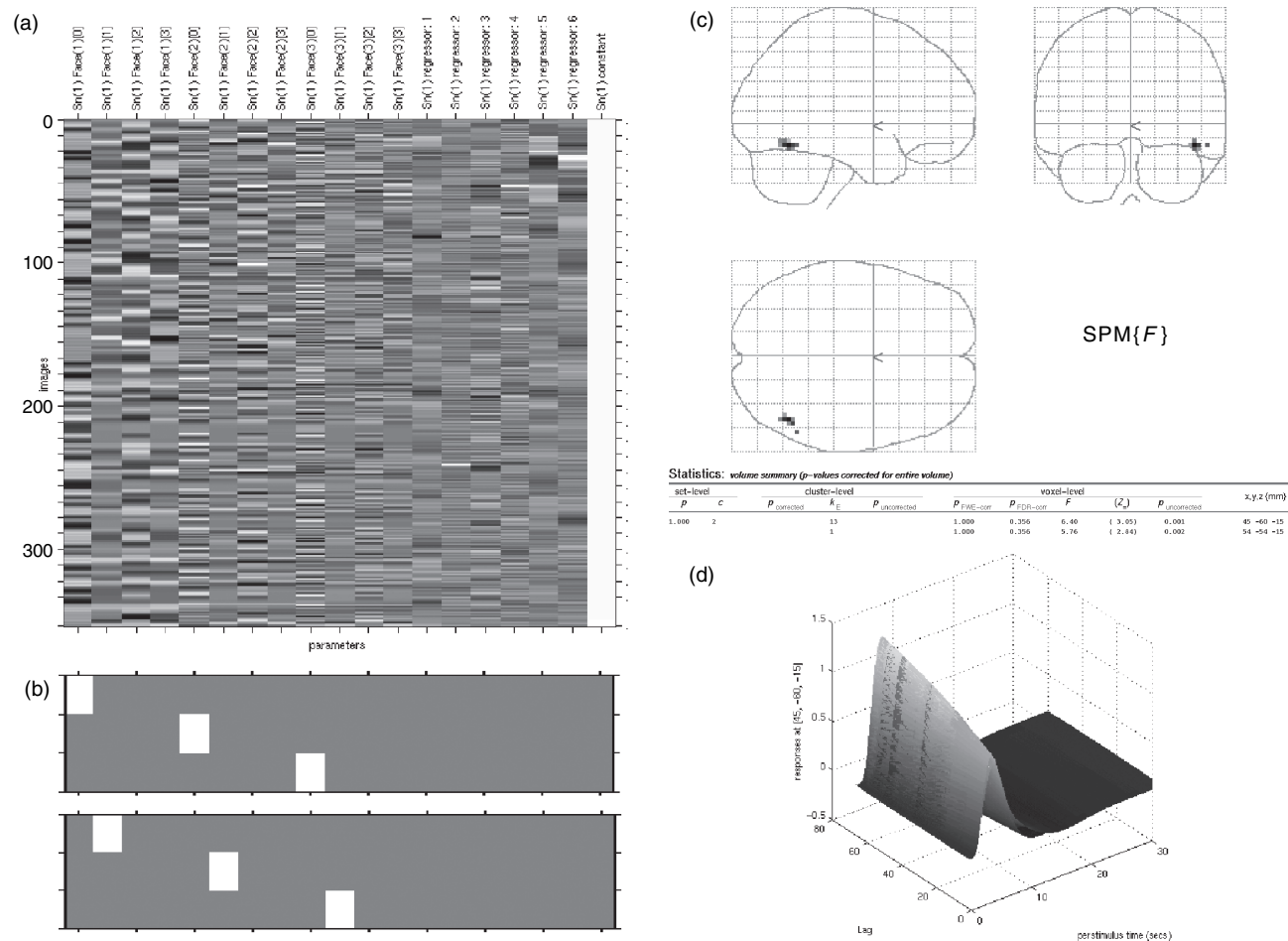
parameter estimates for the temporal and dispersion derivatives).

There are three obvious further effects of interest: the main effects of familiarity and repetition, and their interaction. The results from an  $F$ -contrast for the repetition effect are shown in Plate 13(c), after inclusive masking with the effects-of-interest  $F$ -contrast in Figure 14.4(c). This mask restricts analysis to regions that are generally responsive to faces (without needing a separate face-localiser scan, cf. Kanwisher *et al.*, 1997), and could be used for a small-volume correction (see Chapter 17). Note that this masking is facilitated by the inclusion of null events (otherwise the main effect of faces versus baseline could not be estimated efficiently, see Chapter 15). The contrast of parameter estimates and fitted responses for the single right posterior occipitotemporal region identified by the repetition contrast are shown in Plate 13(d).

Differential effects were seen on all three basis functions, and represent decreased responses to repeated faces.<sup>8</sup>

Plate 14(a) shows the design matrix using a more general FIR basis set of  $K = 6$  2s time bins. The effects-of-interest contrast (see Plate 14(b)) reveals a subset of the regions identified with the canonical basis set (cf. Plate 14(c) and Figure 14.4(c)). The absence of additional suprathreshold voxels when using the FIR model is likely to reflect the reduced statistical power for this  $F$ -test to detect BOLD responses with a canonical form (and the

<sup>8</sup> Note that this difference in the temporal derivative parameter estimates does not imply a difference in latency, given the concurrent difference in canonical parameter estimates: i.e. larger canonical responses require larger temporal derivatives to shift them in time (Henson *et al.*, 2002); as mentioned previously, it is the ratio of the two parameter estimates that estimates latency.



**FIGURE 14.5** Parametric model (a) Design matrix, columns ordered by basis function – canonical HRF, temporal derivative, dispersion derivative – and within each basis function by parametric effect – main effect, lag, familiarity, lag-x-familiarity. (b) *F*-contrasts for main effect (top) and lag effect (bottom). (c) SPM{*F*} MIP for lag effect, together with SPM tabulated output, thresholded at  $p < 0.005$  uncorrected, after inclusive masking with main effect at  $p < 0.05$  corrected. (d) Parametric plot of fitted response from right occipitotemporal region (+45, -60, -15), close to that in Plate 14(c), in terms of percentage signal change versus PST and lag (infinite lag values for first presentations not shown).

likely absence of non-canonical responses). Plate 14(d) shows the parameter estimates from a right fusiform voxel for each of the event-types (concatenated), which clearly demonstrate canonical-like impulse responses in all four cases. No right occipitotemporal region was identified by an *F*-contrast testing for the repetition effect (inset in Plate 14(c)) when using the FIR basis set. This reflects the reduced power of this unconstrained contrast. Note that assumptions about the shape of the HRF can be imposed via appropriate contrasts within this FIR model, as illustrated by the *t*-contrast inset in Plate 14(b), which corresponds to a canonical HRF.

### Parametric model

In this model, a single event-type was defined (collapsing the onsets for the four event-types above), which

was modulated by three parametric modulations. The first modelled how the response varied according to the recency with which a face had been seen. This was achieved by an exponential parametric modulation of the form:

$$\alpha_j = \exp(-L_j/50) \quad 14.21$$

where  $L_j$  is the 'lag' for the  $j$ -th face presentation, defined as the number of stimuli between that presentation and the previous presentation of that face. The choice of an exponential function (rather than, say, a polynomial expansion) was based simply on the observation that many biological processes have exponential time-dependency, and the half-life of the function (50 scans) was somewhat arbitrary (ideally it would be derived empirically from separate data). Thus, as lag increases, the modulation decreases. For first presentations of faces,

## REFERENCES

191

$L_j = \infty$  and the modulation is zero (i.e. there is no possible adaptation or repetition suppression).

The second parametric modulation had a binary value of 1 or -1, indicating whether the face was famous or novel; the third modulation was the interaction between face familiarity and lag (i.e. the product of the first and second modulations, after mean-correction). Each modulation was applied to the three temporal basis functions, producing the design matrix in Figure 14.5(a). The  $F$ -contrast for the main effect of faces versus baseline (upper contrast in Figure 14.5(b)) identified regions similar to those identified by the effects-of-interest contrast in the categorical model above (since the models span similar spaces). As expected, the  $F$ -contrast for the lag effect (lower contrast in Figure 14.5(b)), after masking with the main effect, revealed the same right occipitotemporal region (Figure 14.5(c)) that showed a main effect of repetition in the categorical model. The best-fitting event-related parametric response in Figure 14.5(d) shows that the response increases with lag, suggesting that the repetition-related decrease observed in the categorical model may be transient.

These examples illustrate the use of basis functions and the convolution model for detecting non-stationary (adapting) haemodynamic responses of unknown form in the brain. The experimental design in this instance was as efficient as possible, under the psychological constraints imposed by our question. In the next chapter, we use the basic principles behind the convolution model to look at the design of efficient experiments.

## REFERENCES

- Aguirre G K, Zarahn E, D'Esposito M (1998) The variability of human, BOLD hemodynamic responses. *NeuroImage* 8: 360-69
- Birn RM, Saad ZS, Bandettini PA (2001) Spatial heterogeneity of the nonlinear dynamics in the fMRI bold response. *NeuroImage* 14: 817-26
- Boynton GM, Engel SA, Glover GH *et al.* (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16: 4207-21
- Bullmore ET, Brammer MJ, Williams SCR *et al.* (1996) Statistical methods of estimation and inference for functional MR images. *Mag Res Med* 35: 261-77
- Burock MA, Dale AM (2000) Estimation and detection of event-related fMRI signals with temporally correlated noise: a statistically efficient and unbiased approach. *Hum Brain Mapp* 11: 249-60
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8: 109-14
- Dale A, Buckner R (1997) Selective averaging of rapidly presented individual trials using fMRI. *Hum Brain Mapp* 5: 329-40
- Fransson P, Kruger G, Merboldt KD *et al.* (1999) MRI of functional deactivation: temporal and spatial characteristics of oxygenation-sensitive responses in human visual cortex. *NeuroImage* 9: 611-18
- Friston KJ (2002) Bayesian estimation of Dynamical systems: an application to fMRI. *NeuroImage* 16: 513-30
- Friston KJ, Jezzard PJ, Turner R (1994) Analysis of functional MRI time-series. *Hum Brain Mapp* 1: 153-71
- Friston KJ, Josephs O, Rees G *et al.* (1998a) Non-linear event-related responses in fMRI. *Mag Res Med* 39: 41-52
- Friston KJ, Fletcher P, Josephs O *et al.* (1998b) Event-related fMRI: characterizing differential responses. *NeuroImage* 7: 30-40
- Friston KJ, Mechelli A, Turner R *et al.* (2000a) Nonlinear responses in fMRI: the Balloon model, Volterra kernels, and other hemodynamics. *NeuroImage* 12: 466-77
- Friston KJ, Josephs O, Zarahn E *et al.* (2000b) To smooth or not to smooth? Bias and efficiency in fMRI time-series analysis. *NeuroImage* 12: 196-208
- Friston KJ, Glaser DE, Henson RNA *et al.* (2002) Classical and Bayesian inference in neuroimaging: applications. *NeuroImage* 16: 484-512
- Glover GH (1999) Deconvolution of impulse response in event-related BOLD fMRI. *NeuroImage* 9: 416-29
- Gossel C, Fahrmeir L, Auer DP (2001) Bayesian modeling of the hemodynamic response function in BOLD fMRI. *NeuroImage* 14: 140-48
- Henson R, Andersson J, Friston K (2000) Multivariate SPM: application to basis function characterisations of event-related fMRI responses. *NeuroImage* 11: 468
- Henson RNA, Buechel C, Josephs O *et al.* (1999) The slice-timing problem in event-related fMRI. *NeuroImage* 9: 125
- Henson RNA, Price C, Rugg MD *et al.* (2002a) Detecting latency differences in event-related BOLD responses: application to words versus nonwords, and initial versus repeated face presentations. *NeuroImage* 15: 83-97
- Henson RNA, Rugg MD (2001) Effects of stimulus repetition on latency of the BOLD impulse response. *NeuroImage* 13: 683
- Henson RNA, Rugg MD, Friston KJ (2001) The choice of basis functions in event-related fMRI. *NeuroImage* 13: 149
- Henson RNA, Shallice T, Gorno-Tempini M-L *et al.* (2002b) Face repetition effects in implicit and explicit memory tests as measured by fMRI. *Cerebr Cortex* 12: 178-86
- Huettel SA, McCarthy G (2001) Regional differences in the refractory period of the hemodynamic response: an event-related fMRI study. *NeuroImage* 14: 967-76
- Josephs O, Turner R, Friston KJ (1997) Event-related fMRI. *Hum Brain Mapp* 5: 243-48
- Josephs O, Henson RNA (1999) Event-related fMRI: modelling, inference and optimisation. *Phil Trans Roy Soc London* 354: 1215-28
- Kanwisher N, McDermott J, Chun MM (1997) The fusiform face area: a module in human extrastriate cortex specialised for face perception. *J Neurosci* 17: 4302-11
- Kiebel SJ, Glaser DE, Friston KJ (2003) A heuristic for the degrees of freedom of statistics based on multiple variance parameters. *NeuroImage* 20: 466-78[\*\*14.2]
- Kiebel SJ, Friston KJ (2004) Statistical parametric mapping for event-related potentials. I: Generic considerations. *NeuroImage* 22: 492-502
- Krugel F, von Cramon DY (1999) Temporal properties of the hemodynamic response in functional MRI. *Hum Brain Mapp* 8: 259-71
- Krugel F, Zysset S, von Cramon DY (2000) Nonlinear regression of functional MRI data: an item recognition task study. *NeuroImage* 12: 173-83[\*\*14.3]
- Lee AT, Glover GH, Meyer CH (1995) Discrimination of large venous vessels in time-course spiral blood-oxygenation-level-

- dependent magnetic-resonance functional imaging. *Mag Res Med* **33**: 745–54
- Liao CH, Worsley KJ, Poline J-B *et al.* (2002) Estimating the delay of the hemodynamic response in fMRI data. *NeuroImage* **16**: 593–606
- Malonek D, Grinvald A (1996) Interactions between electrical activity and cortical microcirculation revealed by imaging spectroscopy: implications for functional brain mapping. *Science* **272**: 551–54
- Miezin FM, Maccotta L, Ollinger JM *et al.* (2000) Characterizing the hemodynamic response: effects of presentation rate, sampling procedure, and the possibility of ordering brain activity based on relative timing. *NeuroImage* **11**: 735–59
- Ollinger JM, Shulman GL, Corbetta M (2001) Separating processes within a trial in event-related functional MRI. *NeuroImage* **13**: 210–17
- Penny WD, Trujillo-Barreto N, Friston KJ (2005) Bayesian fMRI time series analysis with spatial priors. *NeuroImage* **24**: 350–62
- Penny WD, Flandin G, Trujillo-Barreto N (in press) Bayesian comparison of spatially regularised general linear models. *Hum Brain Mapp* [**\*\*14.4**]
- Pollmann S, Wiggins CJ, Norris DG *et al.* (1998) Use of short intertrial intervals in single-trial experiments: a 3T fMRI-study. *NeuroImage* **8**: 327–39
- Purdon PL, Weisskoff RM (1998) Effect of temporal autocorrelation due to physiological noise and stimulus paradigm on voxel-level false-positive rates in fMRI. *Hum Brain Mapp* **6**: 239–95
- Rajapakse JC, Kruggel F, Maisog JM *et al.* (1998) Modeling hemodynamic response for analysis of functional MRI time-series. *Hum Brain Mapp* **6**: 283–300
- Saad ZS, Ropella KM, Cox RW *et al.* (2001) Analysis and use of fMRI response delays. *Hum Brain Mapp* **13**: 74–93[**\*\*14.5**]
- Schacter DL, Buckner RL, Koutstaal W *et al.* (1997) Late onset of anterior prefrontal activity during true and false recognition: an event-related fMRI study. *NeuroImage* **6**: 259–69
- Turner R, Howseman A, Rees GE *et al.* (1998) Functional magnetic resonance imaging of the human brain: data acquisition and analysis. *Exp Brain Res* **123**: 5–12
- Vazquez AL, Noll CD (1998) Nonlinear aspects of the BOLD response in functional MRI. *NeuroImage* **7**: 108–18
- Woolrich MW, Behrens TE, Smith SM (2004) Constrained linear basis sets for HRF modelling using variational Bayes. *NeuroImage* **21**: 1748–61
- Woolrich MW, Ripley BD, Brady M *et al.* (2001) Temporal autocorrelation in univariate linear modeling of fMRI data. *NeuroImage* **14**: 1370–86
- Worsley KJ, Liao, CH, Aston, J, Petre, V, Duncan, GH, Morales, F & Evans, AC (2002). A general statistical analysis for fMRI data. *NeuroImage* **15**: 1–15
- Worsley KJ, Friston KJ (1995) Analysis of fMRI time-series revisited – again. *NeuroImage* **2**: 173–81
- Zarahn E (2000) Testing for neural responses during temporal components of trials with BOLD fMRI. *NeuroImage* **11**: 783–96
- Zarahn E, Aguirre G, D'Esposito M (1997) A trial-based experimental design for fMRI. *NeuroImage* **6**: 122–38
- Zarahn E, Aguirre GK, D'Esposito M (1997) Empirical analyses of BOLD fMRI statistics: I Spatially unsmoothed data collected under null-hypothesis conditions. *NeuroImage* **5**: 179–97