1 Testing competing hypotheses about single trial fMRI

Lars Kai Hansen
Department of Mathematical Modelling
Technical University of Denmark B321
DK-2800 Lyngby, Denmark
lkhansen@imm.dtu.dk

Archana Purushotham and Seong-Gi Kim
Center for Magnetic Resonance Research,
Dept. of Radiology and Dept. of Biomedical Engineering,
University of Minnesota, USA.

We use a Bayesian framework to compute probalities of competing hypotheses about functional activation based on single trial fMRI measurements. Within the Bayesian framework we obtain a complete probabilistic picture of competing hypotheses, hence control of both type I and type II errors.

1.1 Introduction

In single trial fMRI experiments we may want to analyse the local activation with respect to several competing hypotheses. As a specific example we have designed an experiment which consists of a cued, delayed motor action with a delay between cue and go-signal that varies from trial to trial. In each pixel we then face three competing hypotheses: no activation, motor preparation (the activation last for the delay between cue and action), and motor execution (the activation lasts only for the time interval of the actual execution). In figure 1.1 we show the two activation reference functions corresponding to motor preparation and motor execution hypotheses. The reference functions are binary time series, the actual fMRI signal will be modeled below as the reference function convolved with a linear filter (the hemodynamic response) degraded by additive white noise.

We will present a Bayesian framework below that allows calculation of relative

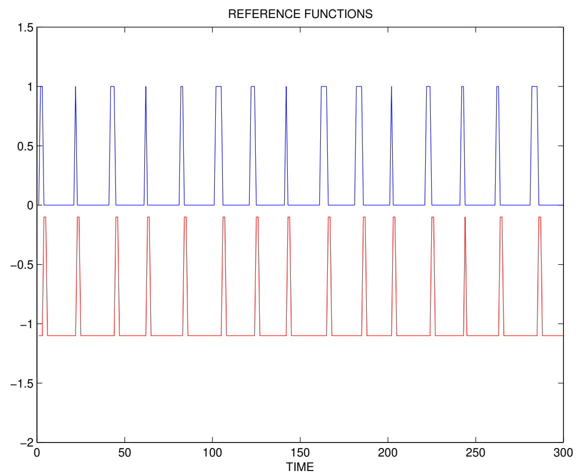

Figure 1.1 The two reference functions for a sequence of single trial cued, delayed motor activation experiments. The two reference functions have been offset vertically for illustration, the upper function is "on" during the preparation phases, while the lower reference function is on during the execution phases.

probabilities of such competing hypotheses.

The Bayesian framework is of interest in this context because it gives a more complete picture of the interplay between null hypotheses and alternatives and the framework has embedded a quantitative statement of the a priori knowledge that enters the formulation of hypotheses.

Frank et al. (1998) recently reviewed a Bayesian framework for signal detection in fMRI data. Here we expand on the application of the Bayesian framework based on so-called *conjugate priors* and we include an explicit treatment of linear hemodynamic response effects.

1.2 Bayes' Theory

We will be concerned with models of the local activation in a region or a single pixel. Let $\mathbf{y}$ be a fMRI signal measured at times $t = 1, ..., T$, and represented as a $T \times 1$ vector with components $\mathbf{y}(t)$. The experiment is characterized in terms of one or more activation reference functions, representing alternative hypotheses about the local activation. Consider linear fMRI signal models of the form,

$$\widehat{\mathbf{y}}(t) = \sum_{\tau=0}^{l} \mathbf{x}(t-\tau)\mathbf{b}(\tau) \tag{1.1}$$

where $\mathbf{x}(t)$ is the average activation in the region or pixel under consideration at time t. $\mathbf{b}$ is a set of $(l+1)$ linear coefficients describing the local hemodynamic response to the activation. Introducing the $T \times (l+1)$ matrix with components $\mathbf{X}(t,\tau) \equiv \mathbf{x}(t-\tau)$ the linear model can be written in matrix form

$$\widehat{\mathbf{y}} = \mathbf{X}\mathbf{b}.. \tag{1.2}$$

In an fMRI experiment we expect that the actual measurement deviates from the "ideal" model output by various noise contributions that we will represent by a random white noise process so that $\mathbf{y}(t) = \widehat{\mathbf{y}}(t) + \mathbf{n}(t)$ where $\mathbf{n}(t)$ is assumed zero mean normal with unknown variance (σ^2). Such models have been investigated e.g. by Lange and Zeger (1997) based on parameterized hemodynamic response functions $\mathbf{b}(\tau)$ of the gamma density form. In et al. (1997) more general finite impulse response models (FIR models) were invoked. FIR models fit the coefficients $\mathbf{b}(\tau)$ in the hemodynamic response function individually. A fundamental problem with Eq. (1.2) is that the neither the model "order" l nor the noise variance σ^2 are known.

A main objective of neuroimaging is to answer questions about the form of activation in regions or locations in the brain. When testing competing hypotheses about the activation of a certain region (or a single pixel) we consider several possible activation patterns $\mathbf{X}_m$ indexed by m. We will show below how we can compute the probability for a given model order, and the probability of a given type of activation (m) given the measurement $\mathbf{y}$ is then

$$P(m|\mathbf{y}) = \sum_{l=1}^{l_{\max}} P(m,l|\mathbf{y}) = \sum_{l=1}^{l_{\max}} P(\mathbf{y}|m,l)P(m,l)/P(\mathbf{y}), \tag{1.3}$$

where we have used Bayes' relation and where $P(m,l)$ is a *prior* probability of the combination of reference function and model dimension. In the following we will let the prior probabilies be a given by a uniform distribution over combinations of (m,l), in the set of reference functions and the corresponding model orders $(l = 1, ..., l_{\max})$, hence, we express no *a priori* preference for any of the alternatives.

For a fixed set of parameters $\mathbf{b}, \sigma^2$ we can use Eq. (1.2) to establish the likelihood function, i.e., the probability density of the observation s given the parameters,

$$P(y|\sigma^2, \mathbf{b}, \mathbf{X}, l) = \left(\frac{1}{2\pi\sigma^2}\right)^{T/2} \exp\left(-\frac{1}{2\sigma^2}(\mathbf{y} - \mathbf{X}\mathbf{b})^2\right). \tag{1.4}$$

Since, however, these parameters too are unknown we need to eliminate them using a *prior distribution* $P(\mathbf{b}, \sigma^2|m,l)$ which quantifies the general knowledge we have on the domain and which potentially depends on the given activation function and model order,

$$\begin{aligned} P(\mathbf{y}|m,l) &= \int d\sigma^2 \int d\mathbf{b} P(\mathbf{b}, \sigma^2|m,l) P(\mathbf{y}|\sigma^2, \mathbf{b}, \mathbf{X}_m, l) \\ &= \int d\sigma^2 \int d\mathbf{b} P(\mathbf{b}, \sigma^2|m,l) \left(\frac{1}{2\pi\sigma^2}\right)^{T/2} \exp\left(-\frac{(\mathbf{y} - \mathbf{X_m}\mathbf{b})^2}{2\sigma^2}\right) \end{aligned} \tag{1.5}$$

see Ohagan (1994).

We will use the principle of *congugate priors* to establish a convenient prior $P(\mathbf{b}, \sigma^2|m,l)$ Ohagan (1994). The conjugate prior for the above linear model with additive gaussian noise is the so-called *normal-inverse-gamma* or $\mathrm{NIG}(a,d,\mathbf{m},\mathbf{V})$, distribution,

$$P(\mathbf{b}, \sigma^2|m,l) = \frac{(a/2)^{d/2}(\sigma^2)^{-(d+l+2)/2}}{(2\pi)^{l/2}|\mathbf{V}|^{1/2}\Gamma(d/2)} \exp\left(-(\mathbf{b}-\mathbf{m})'(2\sigma^2\mathbf{V})^{-1}(\mathbf{b}-\mathbf{m}) - \frac{a}{2\sigma^2}\right). \quad (1.6)$$

The new (hyper-) parameters $d, a, \mathbf{m}, \mathbf{V}$ have the following meaning. The marginal prior distribution of $\mathbf{b}$,

$$\begin{aligned} P(\mathbf{b}|m,l) &= \int d\sigma^2 P((\mathbf{b},\sigma^2|m,l) \\ &= \frac{(a/2)^{-l/2}\Gamma((d+l)/2)}{(2\pi)^{l/2}|\mathbf{V}|^{1/2}\Gamma(d/2)} \left(1 + (\mathbf{b}-\mathbf{m})'(a\mathbf{V})^{-1}(\mathbf{b}-\mathbf{m})\right)^{-(d+l+2)/2} \end{aligned} \quad (1.7)$$

is multivariate t-distribution with mean $\mathbf{m}$ and covariance determined by $(a/(d-2))\,\mathbf{V}$. This distribution is unimodally centered at $\mathbf{m}$, with heavier "tails" than a normal distribution, see figure 1.2. The marginal prior distribution of σ^2 is given by

$$P(\sigma^2|m,l) = \frac{(a/2)^{-d/2}(\sigma^2)^{-(d+2)/2}}{\Gamma(d/2)} \exp\left(-a/(2\sigma^2)\right). \quad (1.8)$$

Hence an inverse gamma distribution (meaning: $1/\sigma^2$ is gamma distributed) of mean $a/(d-2),\ d > 2$.

The next step of the inference is then to set the parameters of the prior. In general we prefer to give the parameters values so that they have minimal influence on results. In particular, we should check that for long time series their effects should vanish completely. The prior mean of the noise variance can, e.g., be set to the observed signal variance, $a/(d-2) = \sigma_y^2 \equiv \mathbf{y}'\mathbf{y}/T$, meaning that we do not expect much larger noise variances than the total observed variance. Further we will let $d = 3$ leading to a prior as shown in figure 1.2. We express no prior knowledge about the mean hemodynamic response, hence, $\mathbf{m} = \mathbf{0}$. The form of the prior covariance structure will be $\mathbf{V} = v\mathbf{1}$, where $\mathbf{1}$ is a unit matrix. The parameter v will be determined essentially by data. The prior variance of the fitted signal $\widehat{\mathbf{y}}$, is given by

$$\langle \widehat{\mathbf{y}}'\widehat{\mathbf{y}}\rangle_{\mathrm{prior}}/T = \mathsf{Tr}[\mathbf{X}\mathbf{X}'\langle \mathbf{b}\mathbf{b}'\rangle_{\mathrm{prior}}/T] = (va/(d-2))\mathsf{Tr}[\mathbf{X}\mathbf{X}']/T \quad (1.9)$$

We can let this be some fraction Q of the variance of the measured signal, i.e., let $v = Q/\mathsf{Tr}[\mathbf{X}\mathbf{X}']/T \sim Q/((l+1)\sigma_{x,m}^2)$, where $\sigma_{x,m}^2$ is the variance of the m'th reference function. The role of the parameter Q will be investigated below; we find that there is a rather wide window of values of this parameter in which the Bayes decisions are stable.

Comparing Eqs. (1.2) and (1.6) we see that by conjugacy they are of the same exponential form, so when we multiply them together the integrand in Eq. (1.5) is again an (unnormalized) NIG distribution, hence the integral is simply the NIG

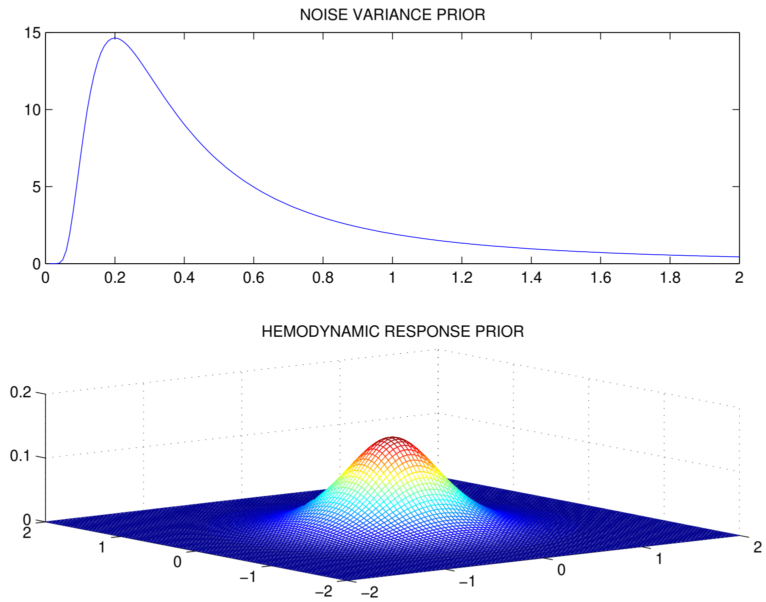

Figure 1.2 Visualization of the priors on the noise variance and the coefficients of the hemodynamic impulse response. The upper panel shows the invers gamma distribution of the noise variance parameter, while the lower panel show the joint distribution of two hemodynamic impulse response parameters. The parameters of both distributions are matched to the actual paradigm and expected signal-to-noise levels.

normalization integral Ohagan (1994), we find

$$P(\mathbf{y}|m,l) = \left(\frac{|\mathbf{V}_{\mathrm{P}}|a^d}{|\mathbf{V}|(a_P)^{d_P}\pi^T}\right)^{1/2}\frac{\Gamma(d_P/2)}{\Gamma(d/2)}, \tag{1.10}$$

with the following definitions

$$\mathbf{V}_P^{-1} = \mathbf{V}^{-1} + \mathbf{X}'\mathbf{X}, \tag{1.11}$$

$$\mathbf{m}_P = \mathbf{V}_P(\mathbf{V}^{-1}\mathbf{m} + \mathbf{X}'\mathbf{y}), \tag{1.12}$$

$$a_P = a + \mathbf{m}'\mathbf{V}^{-1}\mathbf{m} + \mathbf{y}'\mathbf{y} - \mathbf{m}'_P\mathbf{V}_P^{-1}\mathbf{m}_P, \tag{1.13}$$

$$d_P = d + T. \tag{1.14}$$

Using our specifications of the prior parameters we obtain the simplification

$$\mathbf{V}_P^{-1} = \frac{(l+1)\sigma_{x,m}^2}{Q}\mathbf{1} + \mathbf{X}'\mathbf{X}, \tag{1.15}$$

$$a_P = (T+1)\sigma_y^2 - \mathbf{y}'\mathbf{X}\mathbf{V}_P\mathbf{X}'\mathbf{y}, \tag{1.16}$$

$$d_P = 3 + T. \tag{1.17}$$

We can see explicitly that the influence of the prior is weak for $T \gg 1$.

Testing the above linear system hypotheses, a natural nul hypothesis is formed by $\mathbf{x}_0 \equiv 0$, corresponding to a *no fitted signal* model. The corresponding probability density is given by the $\mathbf{X} = \mathbf{0}$ limit of the above expressions.

Assuming that all models are equally probable *a priori*, we get the appropriately normalized posterior probabilities,

$$\begin{aligned} P(m|\mathbf{y}) &= \frac{\sum_l P(\mathbf{y}|m,l)}{P(\mathbf{y}|0) + \sum_{m,l} P(\mathbf{y}|m,l)} \\ P(0|\mathbf{y}) &= \frac{P(\mathbf{y}|0)}{P(\mathbf{y}|0) + \sum_{m,l} P(\mathbf{y}|m,l)} \end{aligned} \quad (1.18)$$

In our particular cued, delayed motor action experiment there will be three such probabilities.

1.3 Evaluation on simulated and fMRI data

1.3.1 Simulation experiments

In order to illustrate the viability of the Bayesian approach for testing hypotheses about time series data we have set up a simulation experiment. The experiment involves a stimulus reference function, here taken to be a simple on/off block design for simplicity. The temporal extent of the experiment is $T = 200$ time units (TR's). In figure 1.3 we show in three panels the stimulus reference function, the hemodynamic response obtained by convolving the impulse response (here a *boxcar* of duration $l_0 = 10$ TR's, see figure 1.5) and the reference function, and finally we show the noise degraded (simulated) fMRI signal. In the first experiment we test the ability of the Bayesian approach to find the correct model, i.e., the correct duration of the response. In figure 1.4 we show the Bayes probabilities $P(l|\text{data})$, for $l = 0, ..., l_{\max}$, with $l_{\max} = 20$. As seen the probabilities are indeed centered on the correct value $l \sim 10$.

The boxcar filter is, of course, not a very realistic model of the hemodynamic response, so the above experiment was redone in figures 1.6-1.7 for the *Gaussian* impulse response shown figure 1.5. Since the Gaussian shape of the impuse response has only little weight for the larger delays, the probabilities are centered on somewhat smaller values of l.

We next investigate the role of the parameter Q. In figure 1.8 we show the probability of activation (summed on all filter lenghts) for both data with activation (ACTIVATION) and without activation (CONTROL) for four different Q-values. It is seen that for Q's in the range $10 - 100$ we get good detection for reasonable relative noise levels, hence, low type-II error rate, while at the same time suppression of the activation probability for signals without activation, hence low type-I error rate.

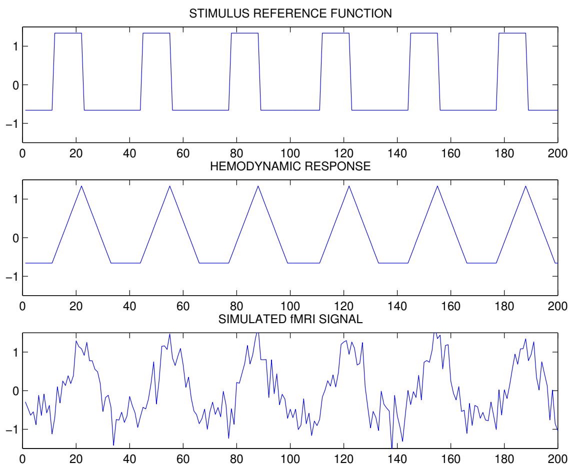

Figure 1.3 Simulation experiment to show the ability of the Bayesian approach factor for picking up the "correct" model order. The upper panel shows the stimulus reference function for a simple block design. The hemodynamic response function is here modeled as a boxcar of length 10 (TR's). The resulting hemodynamic response is shown in the second panel, and finally we show the simulated fMRI signal in the lower panel. The noise level is 0.3 (ratio of standard deviations – noise/signal).

1.3.2 fMRI experiment

We now return to the experiment briefly introduced in the introduction. We analyse a small subset of an experiment designed to determine the location of two different aspects of motor control (see Purushotham et al. (2001) for details and analysis of the complete data set). The experiment involves a cued, delayed, motor response and in the motor areas we can expect different behavior in regions involved in planning versus execution of the response. In the context of the new method we consider here three competing hypotheses. A pixel can either be non-activated (nul hypothesis), be activated during preparation or be activated during execution.

The paradigm consisted of a delayed, cued joystick movement task performed by normal, right-handed adult human subjects. The subject was asked to move a cursor from the centre of the screen to the memorized location of a target using a joystick. The cursor was positioned in the centre of the screen at the beginning of the task. A yellow target appeared for 200 ms at one of 8 possible locations along the circumference of a circle centred on the screen (view angle: < 5 degrees). Then the target disappeared. Following a pseudo-randomized variable (0-3 seconds) delay period, the colour of the central circular home zone changed from red to green, the *go* cue. The subject was to then move the cursor radially to the memorized location of the target, as accurately and quickly as possible, and then move it back to

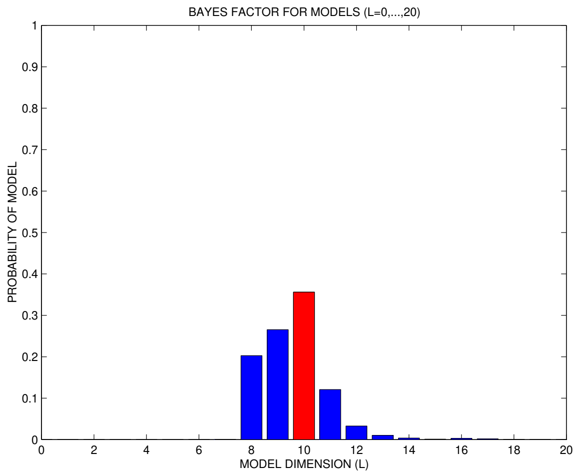

Figure 1.4 Simulation experiment to show the ability of the Bayes approach for picking up the "correct" model order. The stimulus reference function for a simple block design as shown in figure 1. The hemodynamic response function is here modeled as a boxcar of length 10 (TR's) see figure 3. The model probability was computed for models with lags ranging from 0 to $l_{\max}, l_{\max} = 20$. The probability of the zero lag model (no fitted signal) is the nul hypothesis. The figure shows the probability of each of the 21 models considered and is as expected focused on the "correct" model $l = 10$

the home circle. This signalled the end of a single trial, and a return to the control state. During the control period between tasks, the targets continued to appear and disappear, one every 3 seconds, to control for visual stimulation and eye movement. The home circle remained black during this period, and the subject was instructed not to make any joystick movements.

A modified joystick (Measurement Systems, Model 521) from which all ferromagnetic components were removed, was used. The maximum range of motion is an arc of $+/-30$ degrees. The joystick was positioned and secured to the side of the subjects right thigh, at a position that was comfortable for the subject to reach with the right hand while lying down. Subjects moved the joystick using their right thumbs and forefingers in a pincer grasp. The output of the joystick was sampled at 100 Hz, and controlled the movement of the cursor on the screen.

16 single trials as described above, were performed, with an interval of 20 sec between the beginning of one trial and the next. 20 volumes (at a rate of 1 volume per sec) were acquired per trial, and alongwith 20 initial pre-steady-state volumes, made up a total of 340 volumes per run. Response accuracy, reaction and movement times and trajectories were recorded for each trial but were not used in the present analysis.

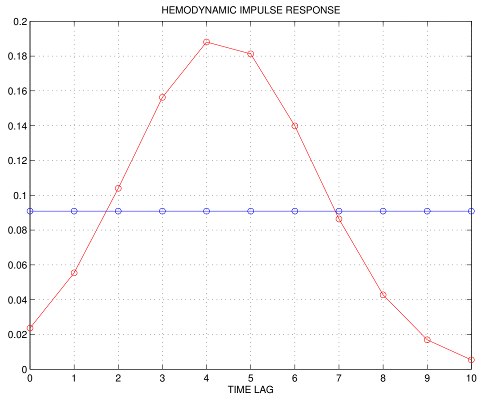

Figure 1.5 Two model hemodynamic response functions used in the initial experiments. The hemodynamic response function is here modeled either as a boxcar or a Gaussian of both length 10 (TR's).

1.3.2.1 Data acquisition

Magnetic resonance imaging experiments were performed on a 4 Tesla whole body MRI system (Varian, Palo Alto, CA) with a homogenous birdcage coil. Multi-slice T1-weighted scout images were acquired with a Turbo FLASH sequence (inversion time of 1.2 s) to select the slices for functional imaging. From these anatomic images, ten coronal 5mm slices, including the primary, pre- and supplementary motor areas, were selected for the joystick task. For the mental rotation task, four axial 10 mm slices were chosen, including the primary and secondary motor areas. For functional MRI studies, a gradient-echo echo-planar imaging (EPI) technique was used. Typical fMRI parameters were matrix size of 64 x 64, field of view (FOV) of 24 x 24 cm2, echo time (TE) of 20 ms (for joystick experiment) or 15ms (for mental rotation experiment), and repetition time of 1000 (for joystick experiment) or 480 ms (for mental rotation experiment).

The study necessitated the provision of visual instruction, and stimuli were generated by a PC. The PC was synchronized with the MRI data collection. Images projected onto a screen inside the magnet room via a projector were viewed by subjects through a mirror. The screen subtended less than 5 degrees of the field of view.

1.3.2.2 Testing competing hypotheses

In each voxel we test three hypotheses: a nul, and two different reference activation functions, demarcating the preparation and the execution phases of the response

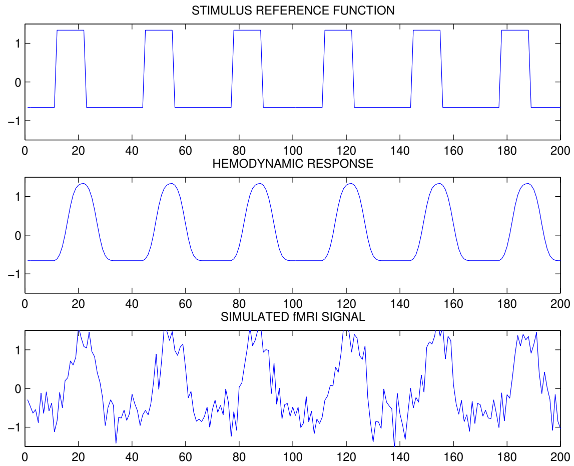

Figure 1.6 Simulation experiment to show the ability of the Bayes approach for picking up the "correct" model order. The upper panel shows the stimulus reference function for a simple block design. The hemodynamic response function is here modeled as a Gaussian of length 10 (TR's). The resulting hemodynamic response is shown in the second panel, and finally we show the simulated fMRI signal in the lower panel. The noise level is 0.3 (ratio of standard deviations – noise/signal).

respectively, as shown in figure 1.1. For each voxel in a 5-by-5 slab of pixels exspected to cover motor and pre-motor areas aggregate probabilities (averaged over both response values and response lenghts) for the where calculated and presented in figure 1.11. The probabilities suggest that all three types of response (nul, prepraration and execution) are present in the slab. For illustration of the Bayesian analysis we also present the probability distributions over response lenghts in figures 1.9-1.10 for both the preparation hypothesis and the execution hypothesis. The two pixels have strongly peaked probability distributions for the two different hypotheses respectively at response durations of 15 and 20 TR's.

1.4 Conclusion

We have outlined a Bayesian framework for signal detection in noisy linear systems. We used weak conjugate priors and as a result we obtain closed form expressions for the relative probabilities over competing hypotheses, depending on only one free parameter. The value of this parameter can be estimated from simulations and the system appears to quite insensitive to its precise value. We used the Bayesian framework to estimate the probability of three alternatives in an fMRI experiment involving planning and execution of motion. In a small slap covering both motor and

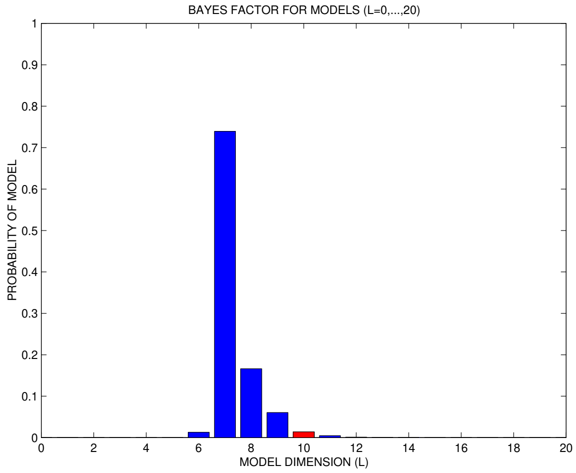

Figure 1.7 Simulation experiment to show the ability of the Bayes approach for picking up the "correct" model order. The stimulus reference function for a simple block design as shown in figure 1. The hemodynamic response function is here modeled as a boxcar of length 10 (TR's). The probability $P(\mathrm{model}|\mathrm{data})$ was computed for models with lags ranging from 0 to $l_{\max}, l_{\max} = 20$. The probability of the zero lag model is the nul-hypothesis probability (no fitted signal). The figure shows the probability of each of the 21 models considered, since the gaussian model has small (relative to the noise level) contributions to the fitted signal for large delays , the probality is focussed on somewhat smaller models than the "correct" model $l = 10$

premotor pixels we found contiguous regions designated to the nul, the execution and to the preparation hypotheses.

Acknowledgments

This work was funded by the Danish Research Councils through the THOR Center for Neuroinformatics, and by the National Institutes of Health, USA, grants MH57180 and RR08079.

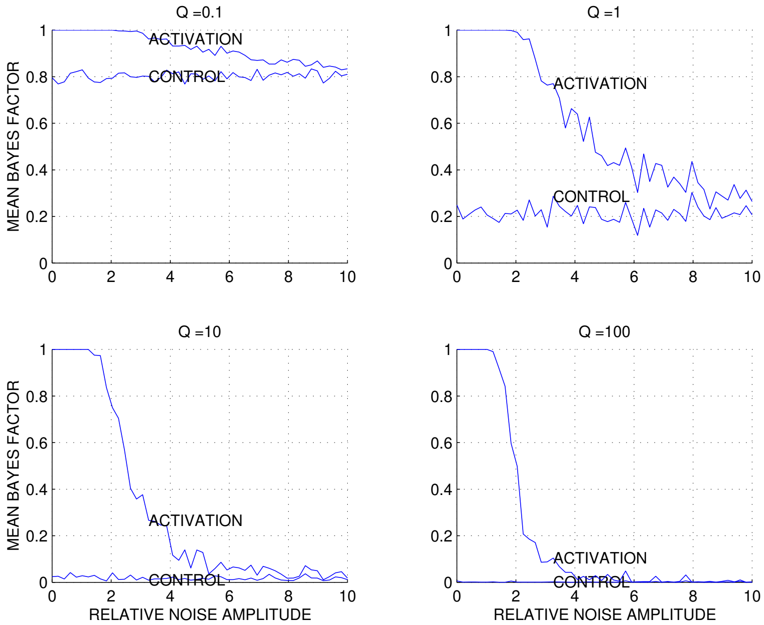

Figure 1.8 Simulation experiment to investigate the role of the prior expected signal-to-noise level. The stimulus reference function for a simple block design as shown in figure 1. The hemodynamic response function is here modeled as a Gaussian of length 10 (TR's). The model probabilities $P(\mathrm{model}|\mathrm{data})$ were computed for models with lags ranging from 0 to $l_{\max}, l_{\max} = 20$. The zero lag model is the $\mathbf{X} = \mathbf{0}$ nul-hypothesis (the measured signal is modeled as a white noise gaussian signal with unknown variance). We calculated the probability of the activation (summing again over all model "lags"), for increasing noise levels. In the plot we show the probability for four different Q's and in each case we show the probability as function of the relative noise amplitude for a situation where the signal actually was activated (ACTIVATION) and for the case when the reference function is zero (CONTROL). We note that for Q's in the range $10 - 100$ we find good detection of signal for relative noise levels less than 1 (low type-II error rate), and at the same time good suppression of the activation probability in case of no activation (CONTROL), i.e., low type-1 error rate.

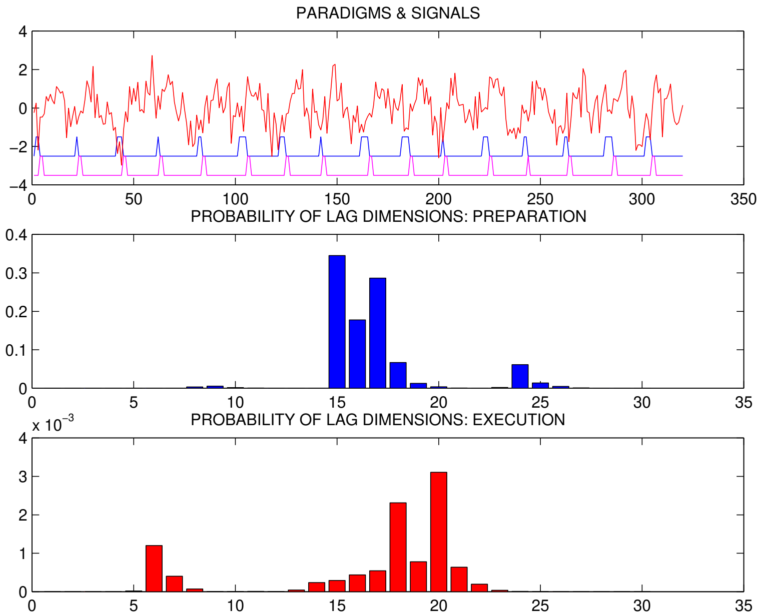

Figure 1.9 Distribution of probabilities accross different model lags ($l = 1, .., 30$) for two models based on different reference functions in the region of a single pixel. In the upper panel we show the measured fMRI signal and the two reference functions for a cued-delay motor activation single trial experiment (from above: fMRI signal, motor preparation, and motor execution). Hypothesis $m = 1$ is a reference function corresponding to motor preparation, while $m = 2$ corresponds to motor execution. This pixel shows high probability of the preparation model.

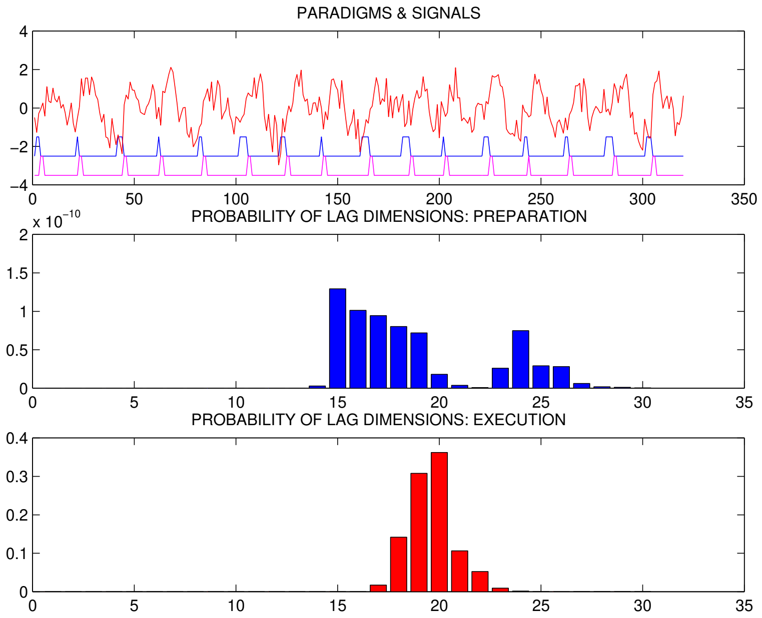

Figure 1.10 Distribution of probabilities accross different model lags ($l = 1, .., 30$) for two models based on different reference functions. In the upper panel we show the measured fMRI signal, and the two reference functions for a cued-delay motor activation single trial experiment (from above: fMRI signal, motor preparation, and motor execution). Hypothesis $m = 1$ is a reference function corresponding to motor preparation, while $m = 2$ corresponds to motor execution. This pixel shows high probability of the execution model.

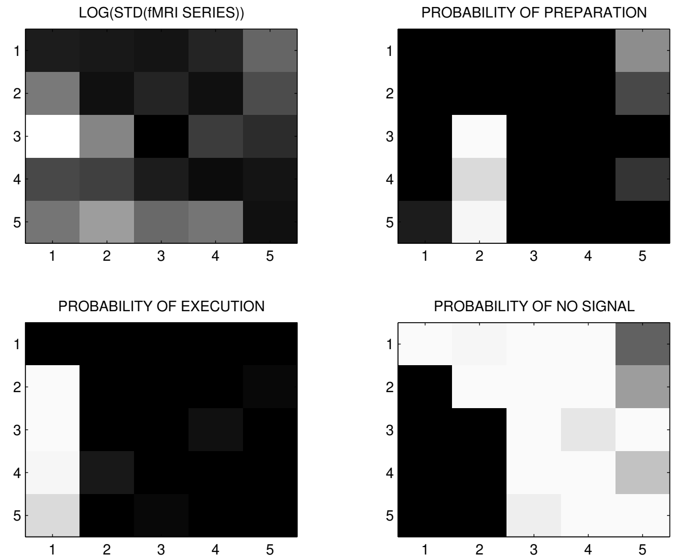

Figure 1.11 Distribution of probabilities in the 5×5 slap of pixels covering pre-motor and motor areas. In the upper left corner we show the signal standard deviation σ_y for the 25 pixels. The three other panels shows the spatial distribution of the probability for the two activation hypotheses and for the nul hypothesis. We use a linear gray map so that bright pixels have probability ~ 1, dark pixels have probability ~ 0).

References

F. Å. Nielsen et al. Comparison of two convolution models for fmri time series. *NeuroImage*, 5:473, 1997.

L.R. Frank, R.B. Buxton, and E.C. Wong. Probabilistic analysis of functional magnetic resonance imaging data. *Magnetic Resonance in Medicine*, 39:132–148, 1998.

N. Lange and S.L. Zeger. Non-linear fourier time series analysis for human brain mapping by functional magnetic resonance imaging. *J. Royal Stat. Soc, Appl. Stat.}, 46:1– 26, 1997.*

A. Ohagan. Bayesian Inference. *The University Press, Cambridge, 1994.*

A. Purushotham, F.Å. Nielsen, L.K. Hansen, and S-G. Kim. Resolving activity due to different functional components of a cognitive task in fmri single trial data using meta-k-means clustering. Unpubllished, *2001.*