Bayesian modelling and analysis of spatio-temporal neuronal networks

Fabio Rigat [†]
Mathisca de Gunst [‡]
Jaap van Pelt [§]

November 2005

Abstract: This paper illustrates a novel dynamic Bayesian network for stochastic integrate-and-fire neuronal systems. We adopt a Bayesian hierarchical perspective to introduce at different model stages the parameters characterizing the neuronal spiking process over a discrete time grid, the unknown structure of functional connectivities and its dependence on the spatial arrangement of the neurons. We compute estimates for all the model parameters and predictions for future spiking states through the standard Gibbs sampler using a shrinkage prior. The paper includes the analyses of a simulated dataset and of experimental in vitro multi-electrode spike trains. In the latter case, we find that the estimates of the neuronal parameters are consistent with their biological interpretation. Furthermore, estimation of the network structure reveals a complex pattern of functional relationships which significantly depend on the spatial distribution of the neurons. Evaluation of the goodness of fit and of the prediction residuals indicates that the model can adequately explain the observed spiking patterns.

Keywords: Bayesian model selection, Bayesian dynamic networks, integrate-and-fire neuronal networks, multi-electrode arrays.

[†]Postdoctoral fellow, EURANDOM, The Netherlands; rigat@eurandom.tue.nl
[‡]Professor of Statistics, Free University, The Netherlands; degunst@cs.vu.nl
[§]Senior researcher, Netherlands Institute for Brain Research, The Netherlands; j.van.pelt@nih.knaw.nl

Report 2005-040

Bayesian modelling and analysis of spatio-temporal neuronal networks

ISSN 1389-2355

1. Introduction

High-throughput technology currently generates measurements of the activity of the nervous system from the level of single neurons (Van Pelt et al. [2004], Morin et al. [2005]) up to the whole brain (Woolrich et al. [2004]). In this context, stochastic modeling provides a formal framework to evaluate the scientific hypotheses explaining specific aspects of the neuronal dynamics. This paper presents a novel dynamic Bayesian network (DBN) model for the statistical analysis of the functional connectivity of stochastic neuronal networks based on measurements of the action potential of each cell over time. Seminal papers in the area of Bayesian networks and DBNs are Spirtes [1995], Heckerman [1996], Ghahramani [1998], Friedman et al. [1998] and Murphy [2001]. In particular, Murphy and Mian [1999] and Friedman [2004] employed DBNs to model genetic and cellular networks. Moreover, while static Bayesian networks are represented by directed acyclic graphs (DAGs), DBNs can be represented by directed cyclic graphs (DCGs). The latter feature of DBNs is in fact relevant in the present context because stochastic neuronal networks can be thought of as DCGs where the directed edges identify pairwise functional connections and the cycles represent each neuron's self dependence over time. The network model described in this work incorporates a Markov dependence of varying order over a discrete time frame and a regression term, both of which are not standard features of DBNs. Furthermore, we adopt a Bayesian hierarchical perspective to jointly estimate from the data all the model parameters, namely the membrane polarization parameters, the functional connectivities, their baseline standard deviation, the unknown network structure and the covariate effects. By letting the network relationships be unknown parameters, we face a problem of model uncertainty (Gelfand and Dey [1994], Draper [2002], Spiegelhalter et al. [2002], Clyde and George [2004]). Dellaportas et al. [2000] and George and McCulloch [1997] provide extensive reviews of several Bayesian model selection methods based on the Gibbs sampler. Godsill [2001] compares these methods with the reversible jump Markov chain Monte Carlo (Green [1995]). In order to incorporate in our model the unknown network structure, we employ a prior mixture along the lines of George and McCulloch [1993]. In this framework, inference can be carried out within a parameter space of fixed dimension using the standard Gibbs sampler. We also show that, in the context of multiple spike trains analysis, adopting such a mixture prior yields robust inferences with respect to random spike sorting errors which, if not properly accounted for, might lead to the identification of spurious network relationships.

The remainder of this paper is organized as follows. Section 2 reviews the statistical literature relevant for our work. Section 3 presents our model for the joint distribution of the network spiking activity and the hierarchical prior. Section 4 reports the posterior distributions of the model parameters and the predictive probabilities of future spike states. Section 5 describes the assessment of the model's goodness of fit and of its predictive power. Sections 6 and 7 report the analyses of simulated data and of in vitro spike data. Section 8 discusses some of the main implications of our modelling approach and it describes some directions for future developments.

2. A review of the statistical modeling of spike trains

Neurons are complex input-output systems whose dynamics have been investigated by experimental neurophysiologists during the last sixty years. For a comprehensive review of the literature in this area, the reader might be referred to Dayan and Abbott [2001] or Gerstner and Kistler [2002] and to the fundamental works by Hodgkin and Huxley [1952], FitzHugh [1961], Nagumo et al. [1962] and Izhikevitch [2001]. Fienberg [1974] describes the essential physiology behind the neuronal spike process and the early literature on the statistical analysis of single neuron spike trains. Brillinger and coauthors (Brillinger [1975, 1976, 1988a, 1988b, 1992, 1996, 2003], Brillinger et al. [1976], Brillinger and Segundo [1979]) developed several point process models characterizing the spiking activity of single neurons and the interactions among small numbers of neurons over time. Doss [1989] proposed a nonparametric method based on point processes to study the pair-wise dependence of the neuronal rates of activity in the time domain. West and Turner [1992], Turner and West [1993] and West [1997] employed a Dirichlet process mixture of Gaussian densities (Ferguson [1983], Escobar and West [1995]) to model the distribution of the response to excitatory post synaptic potentials of single neurons. In the context of single spike trains analysis Kass and Ventura [2001] introduced the inhomogeneous Markov interval point process, which varying memory structure is determined by the inter-arrival times of successive spikes. Kass et al. [2003] used the BARS method (Di Matteo et al. [2001]) as a model based smoother of the instantaneous firing rate function. As emphasized in Iyengar [2001], the focus of these models is the behaviour of either single neurons or of small numbers of neurons. Brown et al. [2004] and Kass et al. [2005] offer two perspectives on the state of the art of multiple spike trains analysis, which is concerned with the development of statistical models for the joint firing activity of many neurons over time. These papers indicate some open challenges in this area that current statistical methods do not adequately address yet, namely modelling the noise induced by the occurrence of false positive and false negative spiking events (i.e. the spike sorting problem) and the estimation of high order interactions among neurons. Along these lines, Okatan et al. [2005] introduced a maximum likelihood method to estimate the functional connectivity

of stochastic neuronal networks based on a discretization of the approach of Chornoboy et al. [1988], whereas Rao [2005] proposed a Bayesian hierarchical model for integrate-and-fire networks in continuous time based on measurements of the membrane potential.

The model presented in this work differs from the previous literature in three essential aspects. First, we distinguish the set of parameters defining the observed spiking process from those identifying the unknown network relationships. Second, we explain the network connectivity through a regression term at the top of the model hierarchy. Third, we assess model adequacy from two perspectives, evaluating the goodness of the model fit for a training sample and its predictive power for an independent validation sample of spike states.

3. A binary stochastic neuronal network

At any time point t a neuron is spiking if its membrane potential exhibits a characteristic large fluctuation called an action potential. Because action potentials are all-or-nothing events, the spiking state of a neuron at time t has only two values, spiking and not-spiking. We assume that experimental measurements of the spiking states of a fixed set of $i = 1, ..., K$ neurons are available over a fixed discrete time grid $t \in (1, ..., T)$. In what follows, the time grid is chosen so that the time points t are equally spaced and we do not observe spikes within and of the intervals $(t-1, t)$.

Let Y with elements Y_{it}, dimensions $K \times T$ and columns Y_t be the binary matrix of random spiking states with $Y_{it} = 1$ if neuron i is firing at time t and $Y_{it} = 0$ otherwise, so that $S_i(t) = \sum_{w=1}^{t} Y_{iw}$ represents the cumulative spike process for neuron i at time t. Let P denote the joint distribution of the data Y, indexed by the parameter vector θ with length K, the parameter matrix β with dimensions $K \times K$ and a fixed initial state $Y_0 = y_0$. Here the parameters θ_i represent neuron specific effects whereas each coefficient β_{ij} represents the pair-wise functional connectivity with j as the transmitting neuron and i as the receiving neuron. Conditionally on (θ, β, y_0) the joint data sampling distribution can always be factored as

$$\begin{aligned} P(Y \mid \theta, \beta, y_0) &= P(Y_1, ..., Y_T \mid \theta, \beta, y_0) \\ &= \prod_{t=1}^{T} P_t(Y_t \mid \theta, \beta, y_0, ..., Y_{t-1}), \end{aligned} \tag{1}$$

where P_t is the conditional distribution of the K dimensional vector Y_t. Moreover, let the distribution of the firing state Y_{it} depend only on (θ, β) and on the past network history between neuron i's last firing time, τ_{it}, and time $(t-1)$. Formally τ_{it} is defined as

$$\tau_{it} = \begin{cases} 0 & \text{if } \sum_{\tau=0}^{t} Y_{i\tau} = 0, \\ \max\{0 \le \tau < t : Y_{i\tau} = 1\} & \text{otherwise.} \end{cases}$$

Under the latter assumption, the distribution P can be factored as

$$P(Y \mid \theta, \beta, y_0) = \prod_{t=1}^{T} \prod_{i=1}^{K} P_{it}(Y_{it} \mid \theta_i, \beta_{i1}, ..., \beta_{iK}, \{Y_w\}_{w=\tau_{it}}^{t-1}), \tag{2}$$

being P_{it} the conditional distribution of Y_{it}. The spiking pattern of each neuron is modeled in equation (2) as a discrete time renewal process which renewals take place after each firing. This assumption reflects an essential feature of the underlying biological process, namely that at each point in time the action potential of a neuron can be thought of as a function of the post synaptic potentials taking place between its successive firings (Dayan and Abbott [2001], Brillinger [1988b], Kass and Ventura [2001]). Moreover, equation (2) states that at each point in time t, conditionally on their parameters $(\theta_i, \beta_{i1}, ...\beta_{iK})$ and on the relevant history of the network, the spiking states $(Y_{1t}, ..., Y_{Kt})$ are independent random variables. We also note that the factorization (2) of the joint distribution P would not hold if the random spiking states would be allowed to simultaneously depend on each other. In such a case, to define consistently the joint distribution of the K spike trains we would face the issues illustrated for instance in Lauritzen and Spiegelhalter [1988] and Lauritzen [1996].

In order to give an explicit form to the likelihood function (2), at any time t let $\{Y_{it}\}_{i=1}^{K}$ be conditionally independent Bernoulli random variables given their success probabilities $\{\pi_{it}(\theta_i, \beta_{i1}, ..., \beta_{iK}, \{Y_w\}_{w=\tau_{it}}^{t-1})\}_{i=1}^{K}$. For notational convenience, in what follows we will drop the dependence of π_{it} on the array of parameters and network history. Under this assumption equation (2) can be rewritten as

$$P(Y \mid \theta, \beta, y_0) = \prod_{t=1}^{T} \prod_{i=1}^{K} (\pi_{it})^{Y_{it}} (1 - \pi_{it})^{1 - Y_{it}}. \tag{3}$$

We note that Brillinger [1988a] proposed a Bernoulli specification of the conditional spike probabilities similar to equation (3) in the context of a random threshold spike train model. However, in our formulation we do not assume that a spike occurs at time t when the membrane potential of neuron i exceeds an unknown threshold, but we define the firing probability π_{it} as the expectation of the random spiking state Y_{it}. In order to define the firing probabilities, we assume a linear integration of the input signal and a logistic link and we write π_{it} as

$$\pi_{it} = \left(1 + e^{-\theta_i - \sum_{j=1}^{K} \beta_{ij} \left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_{jw}}{t - \tau_{it}} \right)} \right)^{-1}. \tag{4}$$

Each parameter β_{ii} in equation (4) measures the slope of the linear relationship between the time lag $(t - \tau_{it})$ and the firing log odds $log\left(\frac{\pi_{it}}{1-\pi_{it}}\right)$. These parameters represent the overall refractory effect of each neuron,

i.e. the propensity of the firing probability to decrease after a spike as a consequence of the underlying ion channel dynamics. Each extra diagonal term β_{ij} represents the constant contribution of neuron j's firing proportions $\frac{\sum_{w=\tau_{it}}^{t-1} Y_{jw}}{t-\tau_{it}}$ to the spiking probability π_{it}. Positive values of β_{ij} represent excitatory functional relationships, whereas negative values represent inhibitory functional relationships. These parameters in fact do not correspond to individual synapses between pairs of interconnected neurons but they represent the average effect of the activity of neuron j on the firing rate of neuron i under the input integration process assumed in equation (4). The neuron-specific frailty parameters θ_i are unknown baseline membrane polarization. In absence of external inputs and when the neuron's own refractoriness becomes negligible, equation (4) implies that neuron i's firing probability is $1/(1+e^{-\theta_i})$, constant over time.

Although the exponent of equation (4) is linear in the model parameters (θ, β), the logit link is a nonlinear mapping of the integrated input onto the interval $(0, 1)$. In other terms, the logit link defines a symmetric saturation of the spiking probability with respect to the input process. Since the derivative of the logistic function is a symmetric bell-shaped curve, under (4) the fluctuations of the network firing activity produce small changes of a neuron's firing probability when its integrated input is either very large or very small and larger changes when the integrated input is closer to zero.

Under (3) and (4) the joint probability of the data Y conditionally on the membrane potential parameters θ, the functional connectivities β and the fixed initial condition y_0 can be written explicitly as

$$P(Y \mid \theta, \beta, y_0) = \prod_{t,i} \frac{e^{Y_{it}\left(\theta_i + \sum_j \beta_{ij}\left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_w^j}{t-\tau_{it}}\right)\right)}}{1+e^{\theta_i + \sum_j \beta_{ij}\left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_w^j}{t-\tau_{it}}\right)}}. \tag{5}$$

Finally, let X with elements x_{ij} be a $K \times K$ matrix of predictors fixed over time, representing the available neuron specific characteristics which may influence the probability that a neuron develops network functional connections. Potential covariates of interest are the neuron types (e.g. pyramidal, interneuron, sensory, motor), the spatial coordinates of the somas, the location of a neuron within a particular brain section, past exposure of the neurons to chemicals and genetic covariates such as indicator variables for knock-out genes and so forth. The predictors'influese on the network connectivity is modeled in this work via a regression term at the top of the model hierarchy, as illustrated in the next subsection.

3.1. **Prior distributions.** The frailty coefficients θ are assigned a K dimensional Gaussian prior density with zero mean and covariance matrix $I_K s_\theta^2$, where the scalar s_θ^2 is the fixed common prior variance and I_K is the $K \times K$ identity matrix.

In order to introduce the unknown network structure, for any couple (i, j) let the connectivities β_{ij} be a priori Gaussian with zero mean, zero prior correlation with the other connectivity parameters and standard deviations

$$\sigma_{ij} = \sigma(\upsilon_{ij} + \epsilon(1 - \upsilon_{ij})). \tag{6}$$

In equation (6), the coefficient σ represent the unknown common baseline prior standard deviation, $\epsilon \in (0, 1)$ is a fixed common shrinkage factor and $\upsilon_{ij} = 1$ indicates the existence of a statistically significant functional relationship with direction $j \to i$. Moreover, let υ be the $K \times K$ binary matrix which (i, j)th element υ_{ij} is one with probability $(1 + e^{-\alpha x_{ij}})^{-1}$ and zero otherwise. The parameters υ_{ij} are assumed a priori conditionally independent, so that the prior probability of the matrix υ is a product of Bernoulli random variables with probability mass function

$$P(\upsilon \mid \alpha, X) = \prod_{ij} e^{\alpha x_{ij} \upsilon_{ij}} (1 + e^{\alpha x_{ij}})^{-1}. \tag{7}$$

The prior hierarchy (6)-(7) implies that, conditionally on α and σ, when the shrinkage factor ϵ is close to zero with probability $(1 + e^{-\alpha x_{ij}})^{-1}$ the parameter β_{ij} is a priori very small, although it is not constrained to be exactly zero. Thus, the value of ϵ defines the size of a negligible network effect relative to the prior spread of the connectivities β. The prior mixture (6) has been successfully employed in the context of covariate selection for the linear regression model by George and McCulloch [1993] and by George and Foster [1997], who label their method stochastic search variable selection. Equations (6) and (7) extend their shrinkage prior in two directions. First, in (7) we introduced a second level logistic model explaining the probability that the size of each coefficient β_{ij} is not negligible as a function of the covariates X and of the regression coefficient α. In what follows, the elements x_{ij} represent the log Euclidean distance between the somas of neurons (i, j). Therefore, the scalar α represents the spatial dependence of the network structure so that $\alpha < 0$ implies that the further apart are the neurons, the lower is the prior expectation of a direct functional connection being established between them and vice versa. The parameter α is assigned a Gaussian prior with mean zero and with fixed standard deviation s_α. Second, we place a conjugate inverse gamma prior $IG(a, b)$ on the variance σ^2 of the connectivity parameters. Moreover, when both a and b tend to zero the inverse gamma density approaches $1/\sigma^2$, which is the Jeffrey's prior for the variance of the Gaussian distribution. Given a fixed value of the shrinkage factor ϵ, we estimate σ to fit the shrinkage mechanism to the data, thus avoiding the fine tuning issues encountered when both ϵ and σ are fixed, as in George and McCulloch [1993]. Given the above prior structure and the data joint

probability function, the full model can be written as

$$
\begin{aligned}
\alpha &\sim Norm(0, s_\alpha^2), \\
\theta &\sim Norm(0, I_K s_\theta^2), \\
\sigma^2 &\sim IG(a, b), \\
\upsilon_{ij} \mid \alpha, x_{ij} &\overset{ind}{\sim} Bernoulli\left(\frac{1}{1+e^{-\alpha x_{ij}}}\right), \\
\beta_{ij} \mid \upsilon_{ij}, \sigma, \epsilon &\overset{ind}{\sim} Norm(0, \sigma_{ij}^2), \\
Y \mid \theta, \beta, y_0 &\sim \text{ equation (5).}
\end{aligned} \tag{8}
$$

4. Posterior distributions and inference

Estimates of the parameters $(\alpha, \upsilon, \sigma, \theta, \beta)$ of model (8) can be computed by posterior simulation via the Gibbs sampler (Gelfand and Smith [1990], Smith and Roberts [1993] Tierney [1994]). The latter proceeds by iteratively sampling realizations of all the parameters from their respective full conditional posterior densities (FCPD), which can be obtained up to a multiplicative constant by dropping from the joint posterior $f(\alpha, \upsilon, \sigma, \theta, \beta \mid Y)$ all terms which do not depend on the parameter of interest.

The FCPD of each frailty parameter θ_i, up to a multiplicative constant, can be written as

$$
f(\theta_i \mid Y, \beta, \theta_{-i}, s_\theta) \propto \phi(\theta_i \mid s_\theta) \prod_{t,i} \frac{e^{Y_{it}\theta_i}}{1 + e^{\theta_i + \sum_j \beta_{ij}\left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_w^j}{t-\tau_{it}}\right)}}, \tag{9}
$$

where $\phi(. \mid s_\theta)$ is the Gaussian density with zero mean and standard deviation s_θ and θ_{-i} stands for the vector of all the membrane polarization parameters but its ith term. These conditional posterior densities do not have a closed form so that updating of each θ_i within the Gibbs sampler will be carried out via the Metropolis-Hastings algorithm (Metropolis et al. [1953], Hastings [1970]).

For the FCPD of each parameter β_{ij} we have

$$
\begin{aligned}
f(\beta_{ij} \mid Y, \upsilon_{ij}, \beta_{i,-j}, \sigma, \epsilon, \theta_i) &\propto \phi(\beta_{ij} \mid \sigma_{ij}) \prod_t e^{Y_{it}\beta_{ij}\left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_w^j}{t-\tau_{it}}\right)} \times \\
&\times \left(1 + e^{\theta_i + \sum_k \beta_{ik}\left(\frac{\sum_{w=\tau_{it}}^{t-1} Y_w^k}{t-\tau_{it}}\right)}\right)^{-1},
\end{aligned} \tag{10}
$$

where σ_{ij} is given in equation (6) and $\beta_{i,-j}$ represents the array of coefficients $(\beta_{i,1}, ..., \beta_{i,j-1}, \beta_{i,j+1}, ..., \beta_{i,K})$. Updating of the parameters β will be carried out via a Metropolis within Gibbs step as for the membrane potentials θ.

Conditionally on $(\epsilon, \beta, \upsilon)$ and under the inverse gamma prior adopted in model (8), the FCPD of σ^2 is inverse gamma with parameters (a^*, b^*) defined

by

$$a^* = a + \frac{K^2}{2}, \tag{11}$$

$$b^* = b + 0.5 \sum_{ij} \frac{\beta_{ij}^2}{\upsilon_{ij} + \epsilon^2(1 - \upsilon_{ij})}, \tag{12}$$

so that updating for σ will be carried out in closed form.

Conditionally on (s_α, X, υ), the spatial dependence coefficient α is a posteriori independent of $(\theta, \beta, Y, \sigma)$. Up to a multiplicative constant its FCPD is

$$f(\alpha \mid s_\alpha, X, \upsilon) \propto \phi(\alpha \mid s_\alpha) \prod_{ij} e^{\alpha x_{ij} \upsilon_{ij}} (1 + e^{\alpha x_{ij}})^{-1}, \tag{13}$$

which is not available for exact sampling, so that the Metropolis-Hastings algorithm will be employed to produce approximate posterior inferences also for α.

Conditionally on $(\beta_{ij}, \alpha, x_{ij}, \sigma, \epsilon)$ each υ_{ij} is a Bernoulli random variable independent of the remaining parameters. Its conditional posterior success probability is given by

$$P(\upsilon_{ij} = 1 \mid \beta_{ij}, \alpha, x_{ij}, \sigma, \epsilon) = \frac{e^{\alpha x_{ij}} \phi(\beta_{ij} \mid \sigma)}{e^{\alpha x_{ij}} \phi(\beta_{ij} \mid \sigma) + \phi(\beta_{ij} \mid \sigma\epsilon)}. \tag{14}$$

4.1. **Parameter estimation.** In this work we adopt the sample means and the sample 95% equal tails intervals computed from the Gibbs sampler output as approximate estimates of the marginal posterior means and marginal posterior probability intervals of the model parameters. The posterior mean network $\hat{\upsilon} = \{\hat{\upsilon}_{ij}\}_{i,j=1}^K$ can be obtained by letting $\hat{\upsilon}_{ij} = 1$ if $\frac{\sum_{m=B+1}^M \upsilon_{ij}^m}{M-B} \geq 0.5$ and zero otherwise, where $\{\upsilon_{ij}^m\}$ represents the sequence of the Gibbs sampler draws for the parameter υ_{ij}, B is a fixed burn-in period and M is the total number of draws. An alternative point estimate of the network structure $\hat{\upsilon}$ is the posterior modal network, which is the configuration of υ most visited during the posterior sampling. From a decision theoretical perspective, the mean is the posterior summary which minimizes the posterior risk under a quadratic loss function whereas the mode minimizes the posterior risk under an absolute value loss function (Berger [1985]). In the examples of sections 6 and 7 the mean estimate is preferred to the mode because the latter might not be unique. In order to visualize the network point estimate, if $\hat{\upsilon}_{ij} = 1$, an arrow will be drawn from neuron j to neuron i. The DCG of the posterior mean network will be constructed by matching the set of arrows between all pairs of neurons together with their spatial coordinates.

4.2. **Spike prediction.** In what follows, the data Y will be divided in two parts: a training sample with all the spike states recorded over $[0, ..., t]$ with $t < T$ and a validation sample including the spikes recorded after time t. The former will be used to estimate the posterior distribution of the parameters

$(\alpha, \upsilon, \sigma, \theta, \beta)$. Such estimates will be then used to sequentially predict the spiking states over the validation sample. Within the Bayesian framework, predicting the spiking state of neuron i at time $t+1$ conditionally on the past data up to time t can be carried out by computing the neuron's one step ahead marginal posterior predictive spiking probability, which for model (8) is

$$\begin{aligned} p_{i,t+1} &= P(Y_{i,t+1} = 1 \mid \{Y_w\}_{w=1}^t), \\ &= \int P(Y_{i,t+1} = 1, \theta_i, \beta_{i1}, ..., \beta_{iK} \mid \{Y_w\}_{w=1}^t) d(\theta_i, \beta_{i1}, ..., \beta_{iK}), \\ &= \int P(Y_{i,t+1} = 1 \mid \theta_i, \beta_{i1}, ..., \beta_{iK}, \{Y_w\}_{w=\tau_{i,t+1}}^t) \times \\ &\times \quad f(\theta_i, \beta_{i1}, ..., \beta_{iK} \mid \{Y_w\}_{w=1}^t) d(\theta_i, \beta_{i1}, ..., \beta_{iK}). \end{aligned} \tag{15}$$

The right hand side of equation (15) is the expectation of the spiking probability for neuron i at time $t+1$ with respect to the joint posterior density of the parameters $(\theta_i, \beta_{i1}, ..., \beta_{iK})$. The one step ahead predictive probabilities for model (8) cannot be computed analytically. However, the left hand side of equation (15) can be approximated by Monte Carlo integration:

$$p_{i,t+1} \approx \frac{\sum_{m=B+1}^{M} \pi_{i,t+1}(\theta_i^m, \beta_{i1}^m, ..., \beta_{iK}^m, \{Y_w\}_{w=\tau_{i\ t+1}}^t)}{M - B}, \tag{16}$$

where the summands on the right hand side of equation (16) are defined by (4) and the sequence $\{\theta_i^m, \beta_{i1}^m, ..., \beta_{iK}^m\}_{m=1}^M$ includes the Gibbs sampler draws of the parameters $(\theta_i, \beta_{i1}, ..., \beta_{iK})$. Given the approximation (16), we let the predicted spiking status be $\hat{Y}_{i\ t+1} = 1$ if $p_{i\ t+1} \geq 0.5$ and zero otherwise.

The motivation for using (16) to estimate $\hat{Y}_{i,t+1}$ is that the marginal posterior predictive probabilities incorporate the posterior uncertainty of the parameters $(\theta_i, \beta_{i1}, ..., \beta_{iK})$ as reflected by their posterior samples, whereas the fitted spiking probabilities $\pi_{i,t+1}(\hat{\theta}, \hat{\beta}_{i1}, ..., \hat{\beta}_{iK}, \{Y_w\}_{w=\tau_{i,t+1}}^t)$ take into account only their point estimates irrespectively of their uncertainty. Therefore, if the 95% posterior intervals of the parameters (θ, β) are wide, the fitted spike probabilities can be misleadingly close to either 0 or 1 whereas the marginal predictive spike probabilities will be more conservative.

5. Model assessment

Evaluating statistically the goodness of fit of model (8) is not a straightforward task because, conditionally on the training data, for any $t^* \in [0, t]$ the sampling distribution of the raw fitting residuals

$$r_{it^*}^f = y_{it^*} - \pi_{it^*} \tag{17}$$

is not known exactly. As illustrated by Pregibon [1981], Landwehr et al. [1984] and Albert and Chib [1996], similar difficulties are encountered in the evaluation of the goodness of fit for generalized linear regression models for binary outcomes. Following Albert and Chib we note that, given the

spike data, the distribution of the fitting residuals $r^f_{it^*}$ only depends on the unknown parameters (θ, β) through equation (4). Therefore, the Gibbs sampler draws for (θ, β) provide also an estimate of the distribution of $r^f_{it^*}$ which can be used to evaluate the goodness of the model fit. If $Y_{it} = 1$, then $r^f_{it^*} \in [0, 1]$ and the closer the residuals are to 1, the more the fit is at odds with the observed data. If $Y_{it} = 0$, then $r^f_{it^*} \in [-1, 0]$ and lack of fit is reflected by residuals close to -1. In the following examples we will display the estimated mean and 95% probability intervals for the fitting residuals. Moreover, as a summary indicator of the goodness of fit we will compute for each neuron the proportion of raw residuals significantly crossing the thresholds 0.5 or -0.5. In the former case, we do not observe a spike but the upper endpoint of the estimated 95% probability interval of the fitting residual exceeds one half whereas in the latter case we do observe a spike but the lower endpoint of the 95% interval crosses the value minus one half.

After observing the spike data up to time T, we can also assess the model's predictive power through the prediction residuals

$$r^p_{it'} = y_{it'} - p_{it'} \tag{18}$$

where t' belongs to the validation period $[t + 1, ..., T]$. Unlike the fitting residuals $r^f_{it^*}$, the prediction residuals $r^p_{it'}$ are fixed values because the randomness induced by the parameters (θ, β) has been eliminated in equation (16) through Monte Carlo integration. Therefore, the only indicator of the model's predictive accuracy is the number of prediction residuals exceeding a given threshold. As for the fitting residuals, we will report the proportions of prediction residuals crossing the threshold 0.5 in absolute value.

6. Example 1: analysis of a simulated dataset

In this section, first we illustrate the analysis of a set of simulated data which was performed by employing our implementation of model (8). Second, we use the same set of simulated data to evaluate the robustness of the Bayesian estimates with respect to moderate random spike sorting errors.

The simulated network includes nine neurons, so that the number of possible distinct matrices υ is $2^{9^2} = 2.4179e + 024$. The true value of the connectivity matrix υ, of the network parameters β and of the membrane polarization coefficients θ were sampled from the hierarchical prior introduced in section 3, having fixed $s_\theta = 3.00$, $\alpha = -1.50$, the prior standard deviation $\sigma = 3.00$ and the shrinkage factor $\epsilon = 0.01$. Figure 1 shows the true network structure within its spatial layout. Table 1 reports the true values of the parameters β. Bold figures in this table are associated with existing pairwise functional connections, thus with the nonzero elements of the connectivity matrix υ and with an edge in figure 1. The second column in table 2 shows the true value of the membrane polarization coefficients θ.

Four thousand data points were generated for each neuron by implementing equations (3) and (4). The sample was then divided in a training batch

including the first two thousand time points and a validation batch composed of the last two thousand points. Figure 2 shows the spike intensity functions (SIF) for the nine simulated neurons. The SIFs were obtained by computing the proportion of spikes over windows of constant width 50 time points. In order to produce a smoother curve, successive windows overlap by 45 time points. The firing rates displayed by the nine neurons reproduce a wide range of possible biological conditions, from the extremely frequent firings of neuron 5 to the low activity of neuron 2. The high firing rates of neuron 5 arise from the combination of a strong excitatory input from neuron 9 and a high membrane polarization parameter (3.48). Neurons $(1, 4, 7)$ display rather low firing rates whereas neurons $(3, 6, 8, 9)$ have intermediate firing rates over time. Although neuron 2 receives a strong excitatory input from 7, the combination of its own refractoriness (-1.24) with its low membrane coefficient (-4.43) decreases its SIF level to an average of 5% over the whole period.

In order to analyze the simulated data through model (8), the prior standard deviation of the spatial dependence parameter was set at $s_\alpha = 3.00$ whereas that of the membrane polarization coefficients was set at $s_\theta = 5.00$. We used the Jeffrey's prior for the synaptic variance σ^2, letting the inverse Gamma parameters be $a = 0$ and $b = 0$. The starting configuration for the posterior sampling of the network structure was the fully connected network whereas the starting value for the other coefficients is zero. Simulation studies showed that posterior estimation is not significantly affected by the starting values of the model parameters υ. The posterior estimates were computed using a posterior simulation of twenty thousand iterations using a burn-in period of five thousand iterations. Posterior sampling for (θ, β) was accomplished via a component-wise random scan random walk Metropolis within Gibbs step with independent Gaussian proposals having standard deviation 0.30, yielding acceptance ratios between 20 and 60 percent.

Table 2 shows the inferences for the frailty parameters θ and table 3 displays the posterior inferences for the parameters β associated to the statistically significant pair-wise connections. The tables show that the true values of the parameters (θ, β) lie well within their 95% posterior intervals and that the posterior means are remarkably close to their respective true values. Figure 3 shows the sample frequencies of the distinct configurations of the matrix υ visited along the posterior sampling. The number of edges of the sampled networks ranged from 81 to 22 and the acceptance ratio for the matrix υ stabilized at 11.40 percent. Figure 3 reveals one global mode in the space of visited network configurations and a large number of network structures scarcely supported by the data. The estimated posterior modal network coincides with the true configuration of the matrix. Table 4 reports the estimated posterior mean inclusion probabilities for all the neuronal pairwise interactions. The bold figures in the table correspond to the estimates exceeding 0.5, which define the posterior mean network. By comparing tables 1 and 4 it can be noted that also the latter point estimate

(i,j)	1	2	3	4	5	6	7	8	9
1	$\mathbf{-2.97}$	0.04	-0.02	0.04	0.00	-0.02	0.02	0.05	0.02
2	0.00	$\mathbf{-1.24}$	0.05	-0.02	0.00	0.02	**5.76**	-0.03	0.02
3	-0.04	0.01	$\mathbf{-1.94}$	-0.03	0.00	$\mathbf{-1.20}$	**2.71**	-0.01	**2.01**
4	0.00	-0.03	0.04	$\mathbf{-1.24}$	-0.05	0.04	-0.03	**1.52**	$\mathbf{-3.81}$
5	0.04	0.00	0.02	0.00	$\mathbf{-1.67}$	0.00	0.00	0.02	**2.14**
6	0.02	-0.03	$\mathbf{-3.46}$	-0.05	-0.02	$\mathbf{-1.90}$	0.00	0.06	-0.04
7	0.00	-0.05	$\mathbf{-1.42}$	-0.01	-0.03	0.00	$\mathbf{-1.52}$	0.00	-0.03
8	0.02	0.02	**7.03**	0.01	0.02	**2.29**	0.03	$\mathbf{-1.75}$	0.02
9	0.03	**5.29**	-0.02	-0.07	0.03	0.03	-0.04	**4.00**	$\mathbf{-2.43}$

TABLE 1. true values of the functional connectivity parameters β for the simulated data. Bold figures in the table correspond to the edges in figure 1 and with nonzero entries in the corresponding matrix υ.

coincides with the true configuration of υ. Since the data was generated from model (8), these results are not unexpected. Figures 4 and 5 show the Gibbs sampler output for the standard deviation σ and for the spatial dependence parameter α. After discarding the burn-in period, their estimated posterior means are respectively 3.45 and -2.09, and their estimated 95% posterior interval are $(2.48, 4.92)$ and $(-3.29, -1.31)$. As for the likelihood parameters, the posterior estimation of (σ, α) for this set of simulated data is remarkably close to their underlying true values.

Table 5 shows the proportions of fitting residuals $r^f_{it^*}$ which estimated 95% posterior intervals cross the thresholds 0.5 or -0.5 and the proportions of prediction residuals $r^p_{it'}$ crossing the same threshold values. Figure 6 displays the fitting residuals for neurons 2 and 3, which are respectively associated to the best and to the worst proportions in table 5. Figure 6 and tables 2, 3 and 5 emphasize that although the precision of the posterior estimates for all the model parameters is comparable, the goodness of fit for different neurons is remarkably different. In particular, table 5 suggests that the proportions of large fitting residuals for neurons with either very high or a very low firing rates are lower than those of the neurons with intermediate firing rates. Moreover, figure 7 suggests that the relationship between the number of spikes produced by the nine neurons during the validation sample and the proportions of large prediction residuals is roughly quadratic, being worse for neurons displaying intermediate firing activity and best for extreme neurons with either very low or very high firing rates. This is due to the fact that, as pointed out in section 3, under the logit link adopted in equation (4) the variability of the spiking patterns of neurons with intermediate firing probabilities is higher than that of the neurons with either large or small probabilities.

Parameter	True value	Post. mean	95% post.interval
θ_1	-0.29	-0.24	$(-0.47, -0.01)$
θ_2	-4.43	-4.11	$(-4.50, -3.72)$
θ_3	0.90	0.84	$(0.53, 1.19)$
θ_4	0.29	0.30	$(-0.40, 0.41)$
θ_5	3.48	3.76	$(2.65, 4.97)$
θ_6	3.68	3.68	$(3.34, 4.02)$
θ_7	0.25	0.36	$(0.01, 0.78)$
θ_8	-3.70	-3.58	$(-4.10, -2.88)$
θ_9	0.45	0.39	$(0.12, 0.67)$

TABLE 2. true values and posterior inferences for the membrane polarization parameters θ of the simulated data.

Parameter	True value	Post. mean	95% post.interval
$\beta_{1,1}$	-2.97	-3.68	$(-4.93, -3.11)$
$\beta_{2,2}$	-1.24	-1.16	$(-2.50, 0.08)$
$\beta_{2,7}$	5.76	5.67	$(4.26, 7.15)$
$\beta_{3,3}$	-1.94	-1.87	$(-2.21, -1.54)$
$\beta_{3,6}$	-1.20	-1.12	$(-1.50, -0.93)$
$\beta_{3,7}$	2.71	2.69	$(2.32, 3.05)$
$\beta_{3,9}$	2.00	2.13	$(1.86, 2.39)$
$\beta_{4,4}$	-1.24	-0.76	$(-1.29, -0.26)$
$\beta_{4,8}$	1.52	1.39	$(0.94, 1.86)$
$\beta_{4,9}$	-3.81	-3.67	$(-4.30, -3.10)$
$\beta_{5,5}$	-1.67	-1.97	$(-3.18, -0.83)$
$\beta_{5,9}$	2.14	2.45	$(1.95, 3.05)$
$\beta_{6,3}$	-3.46	-3.47	$(-3.77, -3.18)$
$\beta_{6,6}$	-1.90	-1.97	$(-2.34, -1.62)$
$\beta_{7,3}$	-1.42	-1.46	$(-1.90, -1.05)$
$\beta_{7,7}$	-1.52	-1.88	$(-2.37, -1.47)$
$\beta_{8,3}$	7.03	6.58	$(5.88, 7.27)$
$\beta_{8,6}$	2.29	2.10	$(1.64, 2.56)$
$\beta_{8,8}$	-1.76	-1.31	$(-1.89, -0.77)$
$\beta_{9,2}$	5.30	6.45	$(4.83, 8.40)$
$\beta_{9,8}$	4.00	4.00	$(3.67, 4.34)$
$\beta_{9,9}$	-2.43	-2.39	$(-2.78, -2.03)$

TABLE 3. true values, estimated posterior means and 95% posterior probability intervals of the functional connectivity parameters β for the simulated data.

(i,j)	1	2	3	4	5	6	7	8	9
1	**1.00**	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01
2	0.01	**1.00**	0.01	0.01	0.01	0.02	**1.00**	0.01	0.03
3	0.01	0.02	**1.00**	0.02	0.02	**1.00**	**1.00**	0.06	**1.00**
4	0.01	0.01	0.01	**1.00**	0.02	0.04	0.01	**1.00**	**1.00**
5	0.01	0.01	0.01	0.01	**1.00**	0.01	0.05	0.01	**1.00**
6	0.01	0.01	**1.00**	0.01	0.01	**1.00**	0.03	0.02	0.02
7	0.05	0.02	**1.00**	0.01	0.03	0.04	**1.00**	0.03	0.06
8	0.01	0.01	**1.00**	0.02	0.11	**1.00**	0.09	**1.00**	0.03
9	0.01	**1.00**	0.06	0.01	0.02	0.01	0.03	**1.00**	**1.00**

TABLE 4. estimated posterior mean inclusion probabilities of each pairwise network connection υ_{ij} for the simulated data.

Neuron	Prop. large $r^f_{it^*}$	Prop. large $r^p_{it'}$
1	0.19	0.19
2	0.06	0.04
3	0.27	0.26
4	0.18	0.17
5	0.07	0.07
6	0.23	0.24
7	0.23	0.22
8	0.24	0.19
9	0.21	0.19

TABLE 5. proportions of residuals significantly crossing the thresholds 0.5 or -0.5 for the nine simulated neurons.

6.1. **Spike sorting and robustness.** The artificial data generated from model (8) can also be used to demonstrate that the shrinkage prior (6) can effectively prevent the posterior inferences from being significantly affected by random spike sorting errors. In this section we consider the same set of simulated data analyzed in the previous example, but we reversed the spiking states of 5% of the training data chosen at random across neurons to reproduce the noise generated by a moderate spike sorting problem. Then we analyzed the noisy data using the same priors and simulation strategy as above. The posterior means for the spatial dependence parameter α and for the prior standard deviation σ are respectively -2.13 and 3.43 with 95% posterior intervals $(-3.12, -1.27)$ and $(2.50, 4.94)$. Tables 6 and 7 report the true values and the posterior inferences for the membrane coefficients θ and for the network connectivities β. The large overlap of all the posterior intervals obtained in the two analyses indicate that the inferences for $(\alpha, \sigma, \beta, \theta)$ are robust with respect to the spike sorting noise. Figure 8 represents the

Parameter	True value	Post. mean	95% post.interval
θ_1	-0.29	-0.25	$(-0.51, -0.02)$
θ_2	-4.43	-4.13	$(-4.62, -3.66)$
θ_3	0.90	0.84	$(0.54, 1.16)$
θ_4	0.29	0.20	$(-0.40, 0.47)$
θ_5	3.48	3.73	$(2.20, 5.10)$
θ_6	3.68	3.67	$(3.34, 4.03)$
θ_7	0.25	0.35	$(0.03, 0.74)$
θ_8	-3.70	-3.57	$(-4.11, -2.78)$
θ_9	0.45	0.40	$(0.10, 0.69)$

TABLE 6. true values and posterior inferences of the membrane polarization parameters θ for the simulated data with spike sorting errors.

differences between the estimates of the marginal posterior probabilities of each network pair-wise connection $P(\upsilon_{ij} = 1 \mid Y)$ obtained respectively before and after adding the noise. Their maximum absolute difference is 0.06 and both the estimated modal network and the estimated mean network coincide with the true network structure. The latter result implies that, under model (8), a moderate random spike sorting noise only marginally affects the posterior estimation of the network structure.

7. Example 2: in vitro spike data

In this section we use model (8) to analyze a dataset of spike trains originating from a neuronal network cultured in vitro at the Netherlands Institute for Brain Research (`http://www.nih.knaw.nl/`). A detailed description of the material and methods employed to generate the data can be found in Van Pelt et al. [2004]. The neurons recorded during the experiment are part of a pool of about 150 thousand dissociated rat cortical cells. After extraction from the rat embryo and centrifugation, the cells were plated as a monolayer in a culture chamber, the bottom of which consisted of a multielectrode array (MEA), i.e. a glass plate in which 61 conductive lanes where etched, ending in a hexagonal pattern of electrode tips with diameter of $12\mu m$, represented in figure 9. The electrode diameters roughly match those of cell bodies, ensuring that the detected spikes mostly originate from single cells. Since the Euclidean distance between neighboring electrodes is $70\mu m$, the space between electrodes is filled with many cells from which no activity is recorded. The data consist of the spiking times of the 61 electrodes during a period of more than 40 days in vitro (DIV). For the whole recording period, no external stimulation was given so that the resulting firing patterns and network dynamics are entirely spontaneous. We will analyze in this section the spikes recorded during the first four minutes of the twelfth hour of day 14 of the experiment for electrodes $(7, 11, 22, 29, 52, 53)$.

Parameter	True value	Post. mean	95% post.interval
$\beta_{1,1}$	-2.97	-3.66	$(-4.31, -3.03)$
$\beta_{2,2}$	-1.24	-1.29	$(-2.90, -0.03)$
$\beta_{2,7}$	5.76	5.76	$(4.16, 0.07)$
$\beta_{3,3}$	-1.94	-1.87	$(-2.24, -1.51)$
$\beta_{3,6}$	-1.20	-1.20	$(-1.47, -0.94)$
$\beta_{3,7}$	2.71	2.66	$(2.29, 3.06)$
$\beta_{3,9}$	2.00	2.12	$(1.85, 2.39)$
$\beta_{4,4}$	-1.24	-0.78	$(-1.23, -0.32)$
$\beta_{4,8}$	1.52	1.40	$(0.91, 1.87)$
$\beta_{4,9}$	-3.81	-3.70	$(-4.33, -3.11)$
$\beta_{5,5}$	-1.67	-1.94	$(-3.32, -0.37)$
$\beta_{5,9}$	2.14	2.47	$(1.94, 3.03)$
$\beta_{6,3}$	-3.46	-3.47	$(-3.79, -3.19)$
$\beta_{6,6}$	-1.90	-1.96	$(-2.32, -1.62)$
$\beta_{7,3}$	-1.42	-1.46	$(-1.85, -1.08)$
$\beta_{7,7}$	-1.52	-1.90	$(-2.32, -1.49)$
$\beta_{8,3}$	7.03	6.57	$(5.98, 7.22)$
$\beta_{8,6}$	2.29	2.09	$(1.63, 2.58)$
$\beta_{8,8}$	-1.76	-1.32	$(-1.92, -0.77)$
$\beta_{9,2}$	5.30	6.37	$(4.84, 8.60)$
$\beta_{9,8}$	4.00	4.01	$(3.70, 4.35)$
$\beta_{9,9}$	-2.43	-2.40	$(-2.85, -1.98)$

TABLE 7. true values, estimated posterior means and 95% posterior probability intervals of the functional connectivity parameters β for the simulated data with spike sorting errors.

The first three minutes of recordings will be used to estimate the model parameters and the last minute will be used to compute the prediction residuals. We focus the analysis on this subset of neurons because during day 14 the remaining 53 electrodes do not record enough activity to estimate their parameters with acceptable precision. Since we consider a millisecond-by-millisecond time scale to obtain biologically meaningful estimates, being able to estimate the model parameters from a few minutes of recordings is a key factor to contain the computational burden required by the Gibbs sampler. Moreover, although the cultured network at 14 DIV is still in a developing state, the period over which the data is being analyzed is much smaller than the time scale of neurite outgrowth and synapse formation. Therefore we may reasonably assume that the functional connectivity within this short time frame is stable. Figure 10 shows the spike intensity function for the six neurons. The neurons display very low levels of activity over these four minutes of recording and, as noted in Van Pelt et al. [2005], they exhibit a markedly periodical firing pattern consistently over time. Furthermore, the

figure shows that neuron 29 displays a higher firing intensity with respect to all the other neurons. Specifically, for each cycle of network activity the firing rate of 29 increases over time until a burst of all the other neurons is triggered, which in turn inhibits firing of number 29. Thus, neuron 29 seems to have a pivotal role in the regulation of the network firing pattern.

In the analysis of this dataset, simulation based inferences for all the model parameters were obtained by summarizing Gibbs sampler runs of twenty thousand iterations after discarding a burn-in period of five thousand iterations. As in section 6, in the Gibbs sampler the parameters (σ, υ) were updated in closed form, whereas updating for (α, θ, β) was performed by a component-wise random walk Metropolis Hastings step with independent Gaussian proposals. The prior standard deviation of the parameter α was set at $s = 5$, whereas the parameters of the inverse gamma prior for the standard deviation σ were set at $a = 2$ and $b = 0.5$, resulting in are rather vague prior densities.

The posterior mean for α is -0.16 and its estimated 95% posterior probability interval is $(-0.26, -0.01)$. These estimates confirm that the network in culture developed in such a way that distant neurons are statistically less likely to establish direct functional connections respect to close pairs of cells. Figure 11 shows the posterior frequencies of each pair-wise functional connection (left) and the DCG representation of the estimated posterior mean network. The latter was obtained by thresholding the posterior frequencies to the value 0.5 and then by mapping the estimated functional connections on the layout of the MEA used for the experiment. The high probabilities of the connections on the main diagonal of the left plot indicate that all neurons exhibit a statistically significant time dependent refractoriness. The most isolated neuron in the network is number 53, which firing activity depends only on its own refractoriness and on a two-way relationship with neuron 22. The most connected neuron in the network corresponds to electrode 22, which provides input to all the other cells in the network and receives input from all but number 7. The estimated posterior mean for σ is 13.84 and its estimated 95% posterior interval $(10.36, 18.71)$. Since σ is the common component in the posterior variance of the parameters β, these large estimates also drive upwards the posterior variability of the functional connectivities, as shown in figure 12. Table 8 reports the posterior inferences for the parameters β associated to the statistically significant pair-wise functional connections and table 9 displays the inferences for the membrane polarization coefficients. The parameters θ are highly significant and negative, indicating that the model captures a form of time independent refractoriness for all neurons. In particular, the frailty parameter for neuron 29 is significantly higher than that of the other cells, explaining the partly network-independent activity displayed by the neuron. Moreover, the significant functional connections directed into neuron 29 are excitatory, except

Parameter	Post. mean	95% post.interval
$\beta_{7,7}$	-1.96	$(-8.55, 6.69)$
$\beta_{7,11}$	9.60	$(-0.08, 18.30)$
$\beta_{7,22}$	9.32	$(-0.08, 17.07)$
$\beta_{7,29}$	5.61	$(-6.06, 15.25)$
$\beta_{11,7}$	5.28	$(-0.20, 17.66)$
$\beta_{11,11}$	-4.94	$(-10.86, 2.43)$
$\beta_{11,22}$	25.64	$(17.76, 33.30)$
$\beta_{11,29}$	13.30	$(2.17, 18.62)$
$\beta_{11,52}$	11.50	$(-0.08, 22.99)$
$\beta_{22,11}$	11.74	$(4.64, 19.64)$
$\beta_{22,22}$	2.90	$(1.54, 3.93)$
$\beta_{22,29}$	2.61	$(-0.22, 8.64)$
$\beta_{22,52}$	1.89	$(-1.68, 7.19)$
$\beta_{22,53}$	12.35	$(8.53, 16.03)$
$\beta_{29,11}$	13.28	$(4.99, 18.77)$
$\beta_{29,22}$	13.87	$(8.70, 17.85)$
$\beta_{29,29}$	-15.82	$(-23.61, -9.17)$
$\beta_{29,52}$	16.97	$(-0.01, 34.19)$
$\beta_{52,11}$	-8.23	$(-20.20, 0.21)$
$\beta_{52,22}$	-5.20	$(-16.27, 0.25)$
$\beta_{52,52}$	-14.67	$(-26.94, 4.88)$
$\beta_{53,22}$	14.09	$(8.74, 19.80)$
$\beta_{53,53}$	5.32	$(2.59, 7.10)$

TABLE 8. estimated posterior means and 95% posterior intervals of the functional connectivity parameters β for the in vitro spike data.

for its own refractoriness $\beta_{29,29}$, and its outgoing connections are all excitatory. These results confirm that neuron 29 has a key role in the initiation and in sustaining the network activity over time.

Table 10 reports the number of large fitting and prediction residuals for the six neurons over the four minutes of recording considered in this analysis. The best fit and the most accurate predictions are associated to the relatively high activity neurons 22 and 29 whereas the worst are obtained for neuron 53, which is the most isolated neuron in the network. Figure 13 shows the estimated posterior means and the 95% probability intervals of the fitting residuals for the six analyzed neurons. The figure shows that all the residuals corresponding to the spikes of neuron 52 are very close to 1, suggesting that its firing probabilities might be underestimated.

Parameter	Post. mean	95% post.interval
θ_7	-9.25	$(-9.84, -8.75)$
θ_{11}	-8.95	$(-9.46, -8.53)$
θ_{22}	-6.80	$(-6.98, -6.64)$
θ_{29}	-5.54	$(-5.66, -5.43)$
θ_{52}	-9.90	$(-10.83, -9.14)$
θ_{53}	-8.74	$(-9.20, -8.33)$

TABLE 9. posterior inferences of the membrane polarization parameters θ for the in vitro spike data.

Neuron	Num. large $r^f_{it^*}$	Num. large $r^p_{it'}$
7	6	1
11	19	3
22	9	0
29	2	0
52	6	1
53	19	8

TABLE 10. number of residuals significantly crossing the thresholds 0.5 or -0.5 for the six analyzed neurons respectively over the training period and over the validation minute.

8. Summary and discussion

Section 3 introduced a Bayesian hierarchical network model explaining the neuronal dynamics in a coherent probabilistic framework. We adopted a Markovian dependence structure of varying order over time in order to mimic the features of the spiking process, along the lines of Kass and Ventura [2001]. As in Brillinger [1988a] and Okatan et al. [2005], we defined the spiking probabilities over a discrete time grid . The main reason to adopt a discrete time model is that the neuronal firing process is characterized by an absolute refractory period, which is a small time interval following a spike during which the neuron is not able to fire (Dayan and Abbott [2001]). If the spike train is modeled as a random process, absolute refractoriness implies that conditionally on a spike occurring at time t, the firing probability in the small time interval $(t, t+\delta)$ is zero. On the other hand, there are several analogies between model (8) and a multivariate homogeneous Poisson process. First, under a homogeneous Poisson process the probability of observing more than one spike over small time intervals is very low. Second, simulation studies not reported in this paper showed that the mean spike intensity implied by model (8), as measured by the average number of spikes over fixed time windows, is constant as for the homogeneous Poisson process.

A distinctive feature of model (8) is that the network structure υ is one of its unknown parameters. Since the number of possible network configurations for any set of K neurons is 2^{K^2}, when the network structure is unknown exhaustive search is computationally infeasible even for small scale networks. The Markov chain Monte Carlo (MCMC) estimation for υ proposed in section 4 is an appealing alternative to exhaustive model search because rather than systematically scanning the state space, the Gibbs sampler tends to move with higher probability towards network configurations with higher posterior probability and disregards those which could not have generated the observed spiking patterns. Furthermore, when the posterior distribution of υ is characterized by several peaks and pits, MCMC estimation is more informative than mode finding algorithms. In this work we used the standard Gibbs sampler to produce approximate marginal posterior inferences for all the model parameters. Recently Nott and Green [2004] proposed a method for Bayesian model uncertainty assessment based on the Swendsen-Wang algorithm (Swendsen and Wang [1987]). In the context of model (8), Nott and Green's method prescribes the introduction of a further layer of auxiliary variables which would conditionally remove the mutual dependence across clusters of functional connectivity parameters, allowing the efficient block-updating of the matrix β within the Gibbs sampler. Nott and Green showed that, in the case of covariates selection for Gaussian linear regression models, their method increases the efficiency of the stochastic exploration of the model space with respect to component-wise Metropolis updates. However, the computational cost of Nott and Green's method increases proportionally to the complexity of the network structure. Since the hierarchy (8) is rich and the typical size of the data Y is large, in this work we did not use Nott and Green's method because it would significantly increase the computational burden involved in fitting our model.

As shown by example 2, experimental multiple spike trains typically report the activity of a small fraction of the neurons in culture and can be affected by several sources of noise. In section 4 we showed that the shrinkage prior (6) yields robust estimates with respect to a moderate random spike sorting error. From this perspective, an interesting topic for further research is the assessment of the robustness of the Bayesian estimates with respect to missing data. For instance, in section 7 we noticed that the firing probabilities for neuron 52 appear to be underestimated. Considering that during the analyzed recording period the firing intensity of the remaining 55 available neurons is lower that that of the six analyzed cells, this bias is likely due to the existence of significant excitatory functional relationships between number 52 and one or more neurons not in contact with the MEA electrodes.

The results of our analysis summarized in tables 8 and 9 suggest that neuron 29 triggers the response of neurons $(7, 11, 22)$ through its own higher

propensity to fire, revealing over time a complex pattern of functional network relationships among the six analysed cells. Neurons $(11, 22, 52)$ stimulate in turn the activity of neuron 29 until the neuron's refractoriness produces an abrupt shut-down of its firing activity, which in turn deprives the connected cells of an important source of excitatory inputs, leading to a temporary silencing of the whole network. We note that these findings are consistent with the exploratory analysis over longer time periods reported in section $III - E$ and figures $7 - 8$ of Van Pelt et al. [2004].

Further developments of model (8) currently under study aim at better capturing the complexity of the underlying biological process. In particular, the issue of synaptic plasticity, which is the propensity of the functional connectivity to change over time, has not been directly addressed in the present work. Synaptic efficacy might change, for instance, as the result of external stimulation through activity-dependent mechanisms or by experimental pharmacological conditions. Modelling synaptic plasticity will thus enable model (8) to effectively capture the fine scale structural dynamics of the network arising from the generation of new synapses, from the pruning of inefficient connections and induced by experimental treatments applied to the neurons in culture.

Acknowledgements

The development of this paper has benefited from numerous conversations with the members of the CASPAN group Arjen B. Brussaard, Ronald van Elburg and Arjen van Ooyen. All the DCGs represented in this paper have been generated by the graphical wizard of the BNT toolbox, which can be downloaded at the URL: `http://bnt.sourceforge.net/`. The implementation of model (8) employed in sections 6 and 7 can be obtained upon request to the first author in the form of an independent `MATLAB` toolbox for non-commercial use only.

References

J. Albert and S. Chib. Bayesian residual analysis for binary response regression models. *Biometrika*, 82:747–759, 1996.

J.O. Berger. *Statistical Decision Theory*. Springer-Verlag, second edition, 1985.

D.R. Brillinger. Maximum likelihood analysis of spike trains of interacting nerve cells. *Biological Cybernetics*, 59:189–200, 1988a.

D.R. Brillinger. Some statistical methods for random processes data from seismology and neurophysiology. *The Annals of Statistics*, 16:1–54, 1988b.

D.R. Brillinger, H.L. Bryant, and J.P. Segundo. Identification of synaptic interactions. *Biological Cybernetics*, 22:213—228, 1976.

D.R. Brillinger and J.P. Segundo. Empirical estimation of the threshold model of neuron firing. *Biological Cybernetics*, 35:213—220, 1979.

E. Brown, R.E. Kass, and P.P. Mitra. Multiple neural spike train data analysis: state-of-the-art and future challenges. *Nature Neuroscience*, 7:456–461, 2004.

E.S. Chornoboy, L.P. Schramm, and A.F. Karr. Maximum likelihood identification of neuronal point process systems. *Biological cybernetics*, 59:265–275, 1988.

M. Clyde and E.I. George. Model uncertainty. *Statistical science*, 19:81–94, 2004.

P. Dayan and L.F. Abbott. *Theoretical neuroscience*. MIT press, 2001.

P. Dellaportas, J. Forster, and I. Ntzoufras. Bayesian variable selection using the gibbs sampler. *In: Generalized linear models: A Bayesian perspective (D.K.Dey, S.Ghosh and B.Mallick, eds.). New York: Marcel Dekker*, 2000.

I. Di Matteo, C.R. Genovese, and R.E. Kass. Bayesian curve-fitting with free-knot splines. *Biometrika*, 88:1055–1071, 2001.

H. Doss. On estimating the dependence between two point processes. *Annals of Statistics*, 17:749–963, 1989.

D. Draper. Assessment and propagation of model uncertainty. *Journal of the Royal Statistical Society B*, 57:45–97, 2002.

M.D. Escobar and M. West. Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90:577–588, 1995.

T.S. Ferguson. Bayesian density estimation by mixtures of Normal distributions. *Recent advances in Statistics (M. Rizvi, J. Rustagi, and D. Siegmund, eds.); Academic Press, New York*, pages 287–302, 1983.

S. E. Fienberg. Stochastic models for single neuron firing trains: a survey. *Biometrics*, 30:399–427, 1974.

R.A. FitzHugh. Impulse and physiological states in models of nerve membrane. *Biolphysics Journal*, 1:445–466, 1961.

N. Friedman. Inferring cellular networks using probabilistic graphical models. *Science*, 303:799–805, 2004.

N. Friedman, K. Murphy, and S. Russell. Learning the structure of dynamic probabilistic networks. *In: Adaptive Processing of Temporal Information; Lecture notes in Artificial Intelligence; Springer-Verlag*, 1998.

A.E. Gelfand and D.K. Dey. Bayesian model choice: Asymptotics and exact calculations. *Journal of the Royal Statistical Society B*, 56:501–514, 1994.

A.E. Gelfand and A. F. M. Smith. Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association*, 85:398–409, 1990.

E. George and D. Foster. Calibration and empirical Bayes variable selection. *Technical Report, University of Texas at Austin and University of Pennsylvania, USA*, 1997.

E.I. George and R.E. McCulloch. Variable selection via Gibbs sampling. *Journal of the American Statistical Association*, 88:882–889, 1993.

E.I. George and R.E. McCulloch. Approaches fro Bayesian variable selection. *Statistica Sinica*, 7:339–373, 1997.

W. Gerstner and W.M. Kistler. *Spiking Neuron Models*. Cambridge University Press, 2002.

Z. Ghahramani. Learning dynamic Bayesian networks. *Adaptive processing of sequences and data structures. Lecture notes in artificial intelligence. C.L.Giles and M.Gori editors. Springer-Verlag*, 1998.

S. Godsill. On the relationship between MCMC model uncertainty methods. *Journal of Computational and Graphical Statistics*, 10:230–248, 2001.

P.J. Green. Reversible jump MCMC computation and Bayesian model determination. *Biometrika*, 82:711–732, 1995.

W.K. Hastings. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57-1:97–109, 1970.

D. Heckerman. A tutorial on learning with Bayesian networks. *Technical Report MSR-TR-95-06, Microsoft Research, Redmond, Washington*, 1996.

A.L. Hodgkin and A.F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *Journal of Physiology*, 117:500–544, 1952.

S. Iyengar. The analysis of multiple neural spike train analysis. *In: Advances in Methodological and Applied Aspects of probability and Statistics; N. Bolakrishnan editor; Gordon and Breack*, pages 507–524, 2001.

E.M. Izhikevitch. Resonate and fire neurons. *Neural networks*, 10:1171–1266, 2001.

R.E. Kass and V. Ventura. A spike train probability model. *Neural computation*, 13:1713–1720, 2001.

R.E. Kass, V. Ventura, and E. Brown. Statistical issues in the analysis of neuronal data. *Journal of Neurophysiology; in press*, 2005.

R.E. Kass, V. Ventura, and C. Cai. Statistical smoothing of neuronal data. *NETWORK: Computation in Neural Systems*, 14:5–15, 2003.

J.M. Landwehr, D. Pregibon, and A.C. Shoemaker. Graphical methods for assessing logistic regression models. *Journal of the American Statistical Association*, 385: 61–71, 1984.

S.L. Lauritzen. *Graphical models*. Oxford University press, 1996.

S.L. Lauritzen and D.J. Spiegelhalter. Local computations with probabilities on graphical structures and their applications to expert systems. *Journal of the Royal Statistical Society B*, 50:157–224, 1988.

N. Metropolis, A. Rosenbluth, M. Rosenbluth, M. Teller, and E. Teller. Equations of state calculations by fast computing machines. *J. Chem. Phys.*, 21:1087–2092, 1953.

F.O. Morin, Y. Takamura, and E. Tamiya. Investigating neuronal activity with planar microelectrode arrays: Achievements and new perspectives. *Journal of Bioscience and Bioengineering*, 100:131–143, 2005.

K. Murphy. An introduction to graphical models. *Intel Research Technical Report*, 2001.

K. Murphy and S. Mian. Modelling gene expression data using dynamic bayesian networks. *Technical report, Computer Science Division, University of California, Berkeley, CA.*, 1999.

J.S. Nagumo, S. Arimoto, and S. Yoshizawa. An active pulse transmission line simulating nerve axon. *Proceedings of the IRE*, 50:2061–2072, 1962.

D.J. Nott and P.J. Green. Bayesian veriable selection and the Swendsen-Wang algorithm. *Journal of computational and Graphical Statistics*, 13:141–157, 2004.

M. Okatan, M.A. Wilson, and E.N. Brown. Analyzing functional connectivity using a network likelihood model of ensemble neural spiking activity. *Neural computation*, 17:1927–1961, 2005.

D. Pregibon. Logistic regression diagnostics. *The Annals of Statistics*, 9:705–724, 1981.

R.P.N. Rao. Hierarchical Bayesian inference in networks of spiking neurons. *To appear in: Advances in NIPS; MIT press*, 17, 2005.

A. F.M. Smith and G.O. Roberts. Bayesian computations via the gibbs sampler and related markov chain monte carlo methods. *Journal of the Royal Statistical*

Society B, 55:3–23, 1993.

D.J. Spiegelhalter, N.G. Best, P.B. Carlin, and A. Van der Linde. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society B*, 64:583–639, 2002.

Peter Spirtes. Directed cyclic graphical representations of feedback models. *Proceedings of the eleventh conference on uncertainty in artificial intelligence. Philippe Besnard and Steve Hanks editors. Morgan Kauffmann Publishing Incorporation. San Mateo*, pages 491–498, 1995.

R.H. Swendsen and J.S. Wang. Nonuniversal critical dynamics in Monte Carlo simulations. *Physics review letters*, 58:86–88, 1987.

L. Tierney. Markov chains for exploring posterior distributions. *The Annals of Statistics*, 22:1701–1762, 1994.

D.A. Turner and M. West. Statistical analysis of mixtures applied to postsynaptic potential fluctuations. *Journal of neuroscience methods*, 47:1–23, 1993.

J. Van Pelt, I. Vajda, P.S. Wolters, M.A. Corner, and G.J.A. Ramakers. Dynamics and plasticity in developing neuronal networks in vitro. *Progress in brain research*, 147:173–187, 2005.

J. Van Pelt, P.S. Wolters, M.A. Corner, W.L.C. Rutten, and G.J.A. Ramakers. Long-term characterization of firing dynamics of spontaneous bursts in cultured neural networks. *IEEE Trans. BioMed. Eng.*, 51:2051–2062, 2004.

M. West. Hierarchical mixture models in neurological transmission analysis. *Journal of the american statistical association*, 92:587–606, 1997.

M. West and D.A. Turner. Deconvolution of mixtures in analysis of neural synaptic transmission. *The Statistician*, 43:31–43, 1992.

M.W. Woolrich, M. Jenkinson, M.J. Brady, and S.M. Smith. Fully Bayesian spatio-temporal modelling of FMRI data. *IEEE transcripts on medical imaging*, 23: 213–231, 2004.

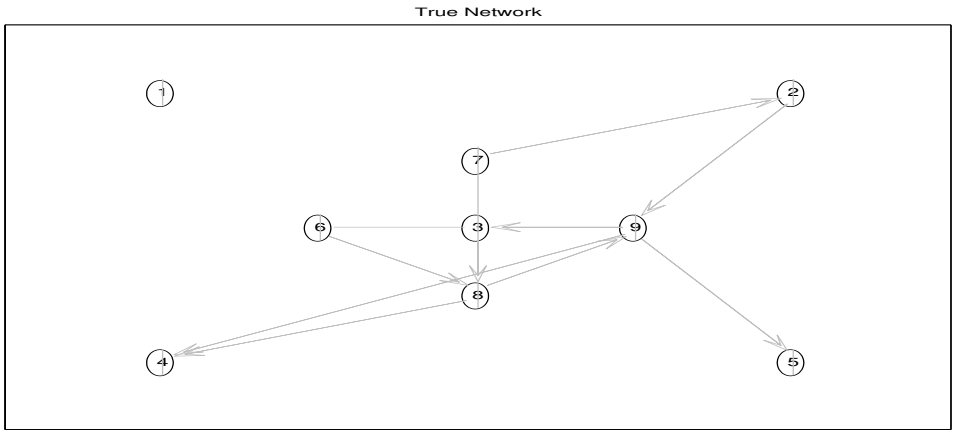

FIGURE 1. simulated network with 9 neurons. Undirected edges represent two way connections whereas directed edges represent one way connections. Vertical bars within neurons represent their refractory effects.

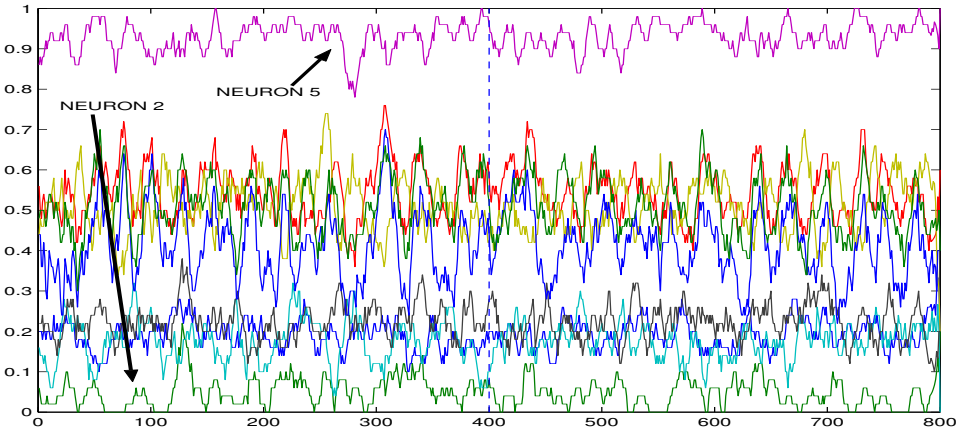

FIGURE 2. firing intensity function for the nine simulated neurons. Each point in the figure represents the proportion of spikes for the respective neuron over a time window of fifty time points; successive windows overlap by 45 time points in order to provide a smoother representation. A vertical dashed line separates the training sample (left) from the validation sample (right). The two arrows identify respectively the highest and the lowest intensity neurons, which are number 5 and number 2.

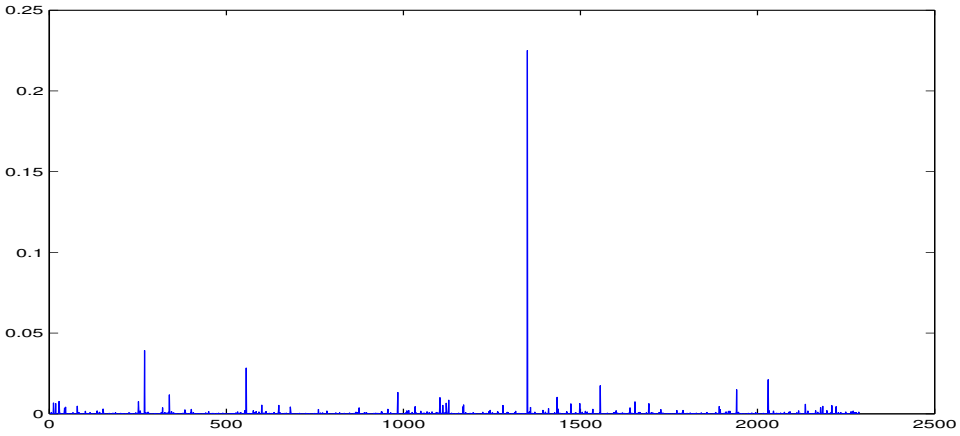

FIGURE 3. sample frequencies of the unique configurations of the connectivity matrix v visited along the posterior sampling. The Gibbs sampler reveals one global mode in the space of networks, which coincides with the underlying true network represented in figure 1.

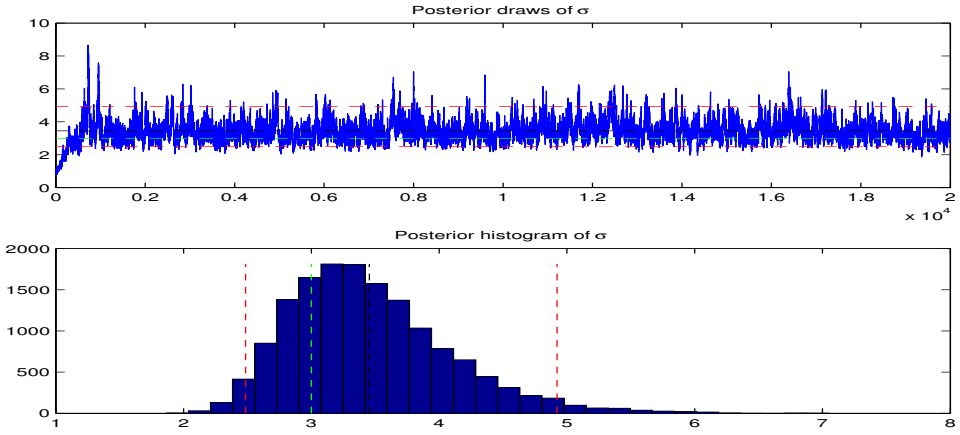

FIGURE 4. Gibbs sampler draws and corresponding histogram for the standard deviation σ. The true value of σ is 3; its posterior mean (marked by the central dashed lines in the two plots) is 3.45 and its equal tails 95% posterior probability interval (marked by the outer dashed lines) is $(2.48, 4.92)$.

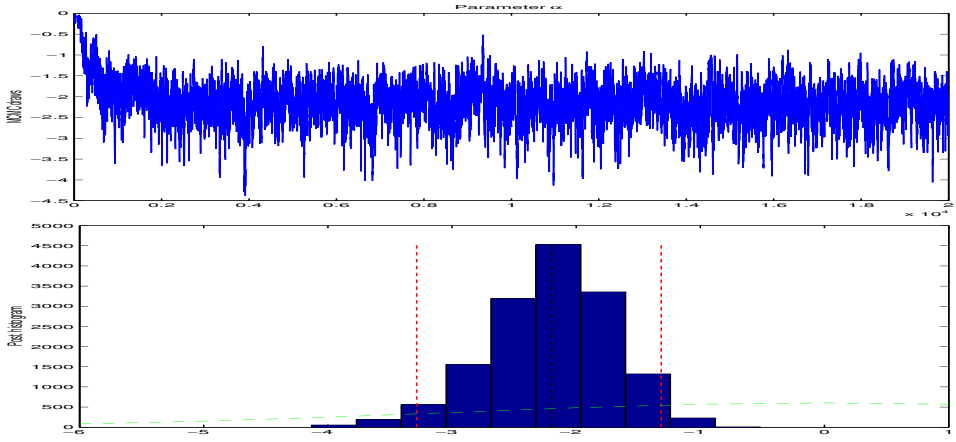

FIGURE 5. Gibbs sampler draws of the spatial dependence parameter α and posterior histogram. The true value of α is -1.50; the posterior mean is -2.09. The vertical dashed segments in the bottom plot represent the endpoints of the 95% posterior interval, which are $(-3.29, -1.31)$. The horizontal dashed curve represents the Gaussian prior density for α on the scale of the histogram.

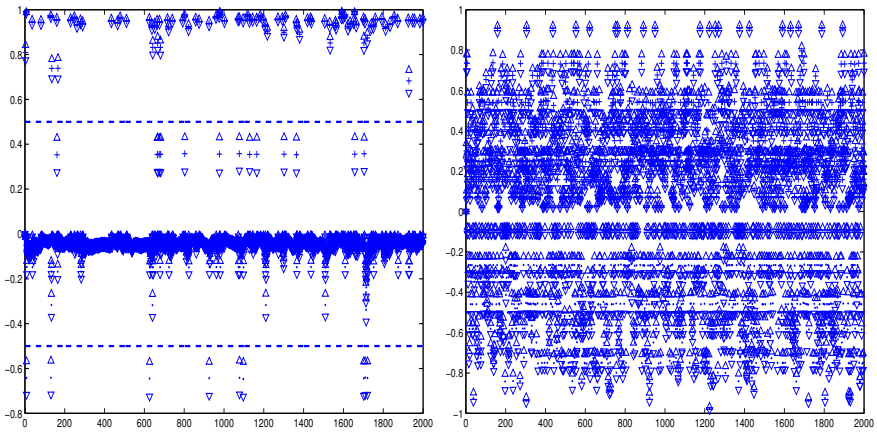

FIGURE 6. estimated posterior means and 95% posterior intervals for the fitting residuals of neurons 2 (left) and 3. In both figures, dots mark the estimated posterior means of the residuals when no spike is observed and plus signs mark their posterior means when a spike takes place. The estimated 95% posterior intervals are marked in the figure by two triangles.

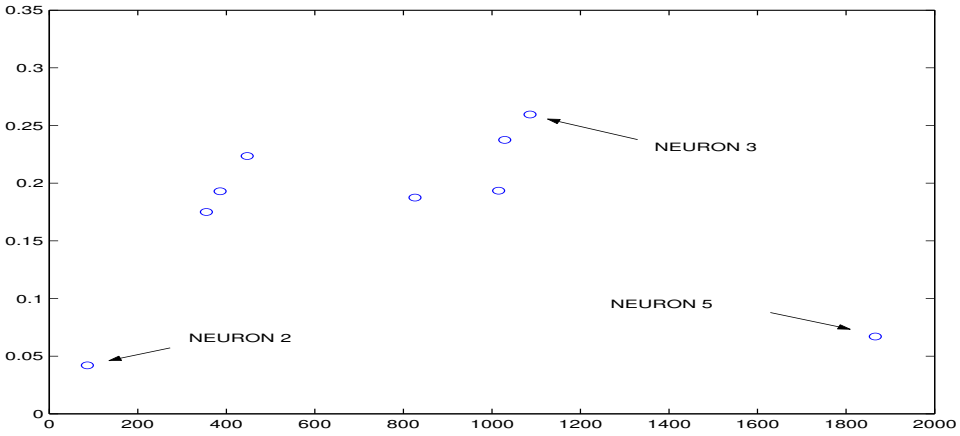

FIGURE 7. the number of spikes observed during the validation period (horizontal axis) is plotted versus the proportion of large prediction residuals for the nine simulated neurons. The predictive performance for the extreme neurons 2 and 5 appears significantly better than that of the neurons with intermediate firing rates.

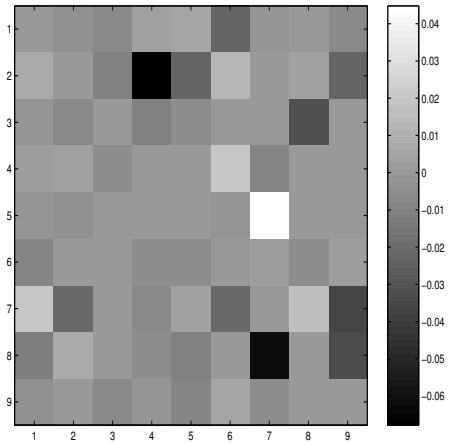

FIGURE 8. differences between the estimated marginal posterior probabilities of each network pair-wise connection obtained from the simulated data before and after adding the spike sorting noise. The largest absolute difference is 0.06, indicating that under model (8) the estimation of the network structure is only marginally affected by the introduction of a moderate random spike sorting error.

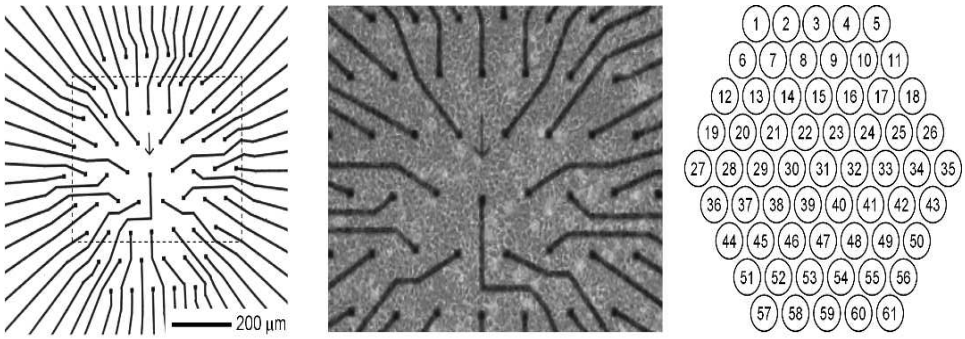

FIGURE 9. wiring of the hexagonal MEA used to record 61 spike trains in vitro for over 40 days (left plot); picture of the MEA plate containing the neurons (central plot); numbering of the MEA electrodes (right plot).

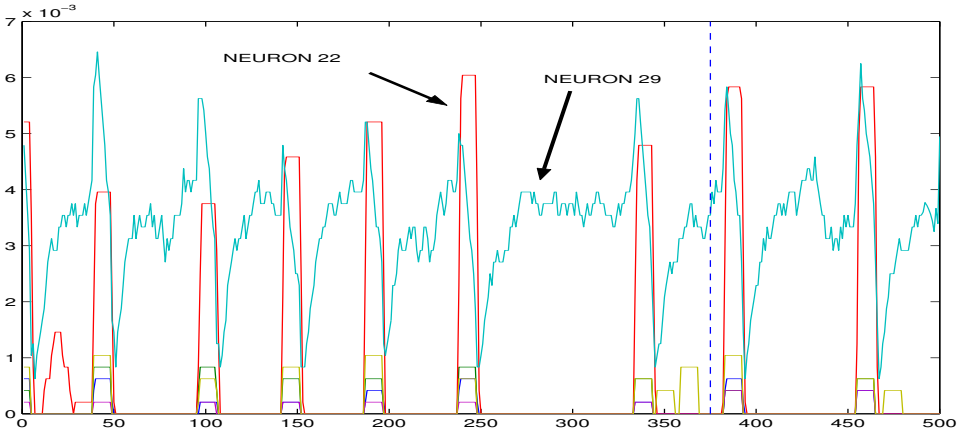

FIGURE 10. SIF of the first four minutes of the twelfth hour of recording for day 14. Each point of the spike intensities represents the proportion of spikes for the corresponding neuron over 500 windows of width 4.8 seconds. Successive time windows overlap by 0.48 seconds. The vertical dashed line separates the training sample (left) from the validation sample (right). Neuron 29 exhibits the highest spike intensity, followed by neuron 22.

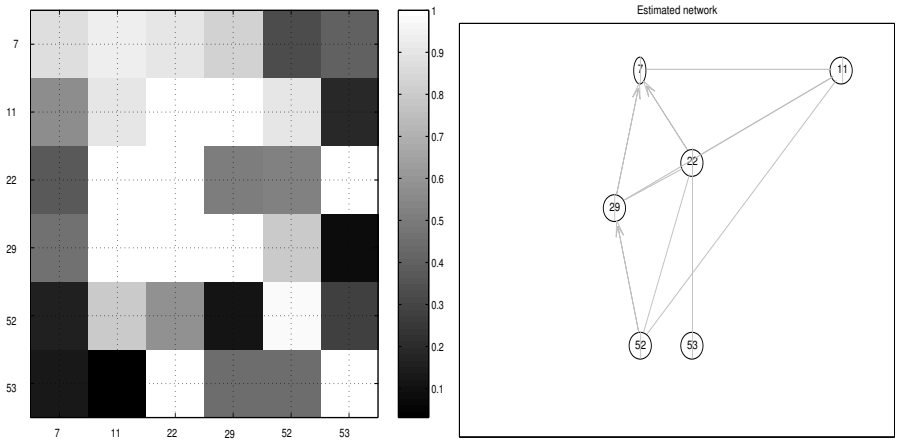

FIGURE 11. posterior frequencies for each pair-wise network functional connection (left) and estimated median posterior network (right) projected on the MEA layout.

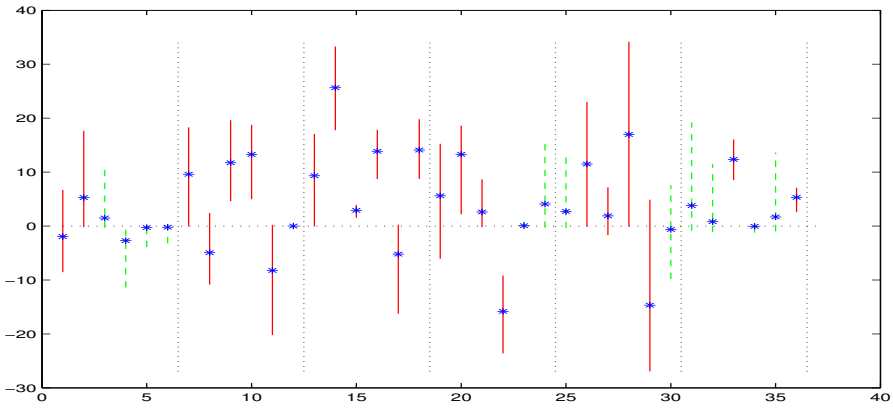

FIGURE 12. estimated posterior means and 95% intervals for the functional connectivities β. The posterior intervals of parameters associated to the statistically significant network connections are marked by whole vertical segments whereas those of the insignificant connections are marked by dashed segments. The vertical dotted lines separate groups of parameters associated to different transmitting neurons. For instance, the first six bars from the left indicate that neuron 7 transmits statistically significant input only to itself and to neuron 11.

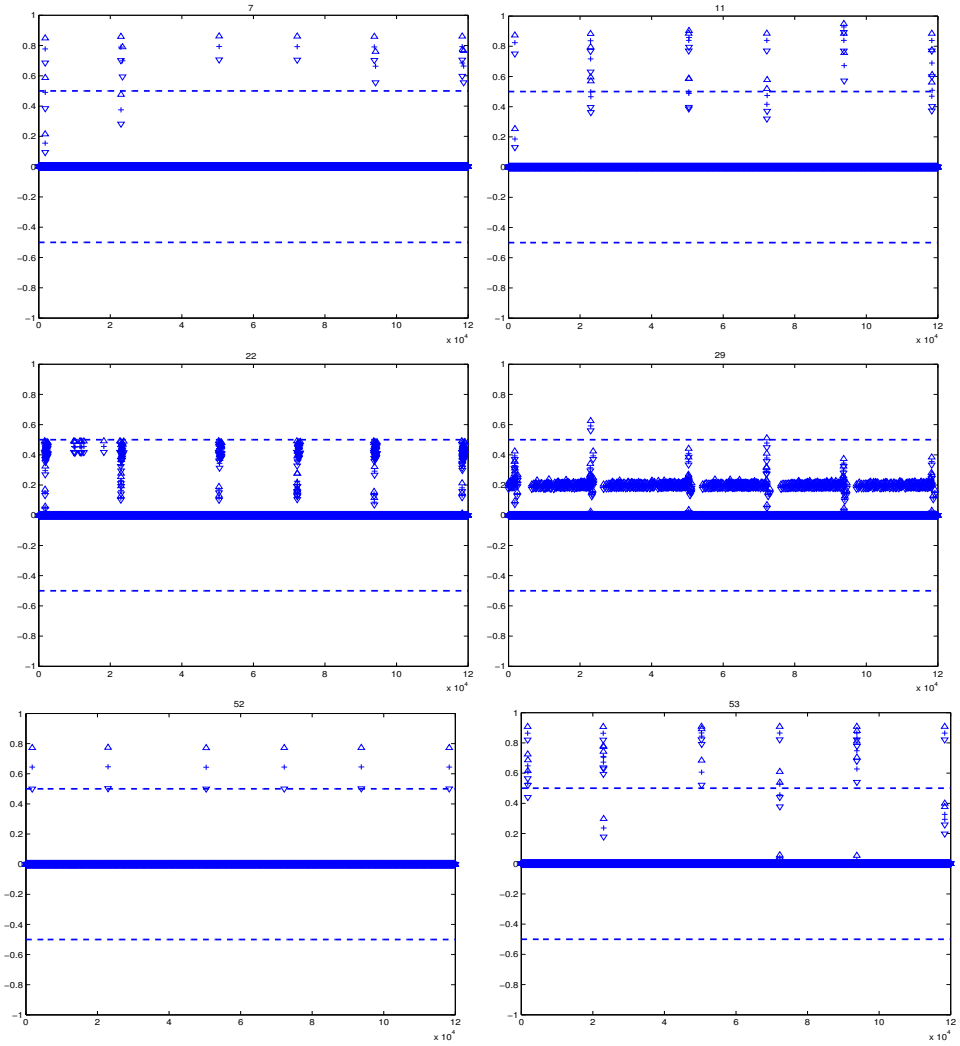

FIGURE 13. estimated posterior means and 95% posterior intervals of the fitting residuals for the six analyzed neurons over the training sample. In the six plots, dots mark the estimated posterior means of the residuals when no spike is observed and plus signs mark their posterior means when a spike takes place. The estimated 95% posterior intervals are marked by two triangles. The small proportion of large residuals indicates that the model adequately fits most of the observed spike patterns. However, the spiking probability for number 52 appears to be underestimated.