

Elimination of response latency variability in neuronal spike trains

Martin P. Nawrot, Ad Aertsen, Stefan Rotter

Neurobiology and Biophysics, Institute of Biology III, Albert-Ludwigs-University, 79104 Freiburg, Germany

Received: 8 April 2002 / Accepted: 26 November 2002 / Published online: 7 April 2003

Abstract. Neuronal activity in the mammalian cortex exhibits a considerable amount of trial-by-trial variability. This may be reflected by the magnitude of the activity as well as by the response latency with respect to an external event, such as the onset of a sensory stimulus, or a behavioral event. Here we present a novel nonparametric method for estimating trial-by-trial differences in response latency from neuronal spike trains. The method makes use of the dynamic rate profile for each single trial and maximizes their total pairwise correlation by appropriately shifting all trials in time. The result is a new alignment of trials that largely eliminates the variability in response latency and provides a new internal trigger that is independent of experiment time. To calibrate the method, we simulated spike trains based on stochastic point processes using a parametric model for phasic response profiles. We illustrate the method by an application to simultaneous recordings from a pair of neurons in the motor cortex of a behaving monkey. It is demonstrated how the method can be used to study the temporal relation of the neuronal response to the experiment, to investigate whether neurons share the same dynamics, and to improve spike correlation analysis. Differences between this and other, previously published methods are discussed.

1 Introduction

Neural activity in the cortex measured during repeated experimental trials exhibits two distinct aspects of variability. The first aspect is expressed in the variable shape of the neural activation that follows a stimulus (Bach and Krüger 1986; Vogels et al. 1989; Arieli et al. 1996; Shadlen and Newsome 1998) or accompanies a

motor action (Vaadia et al. 1988; Lee et al. 1998; Oram et al. 2001; MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted). The second aspect concerns the temporal relation of the neuronal activation profile with respect to the time frame of the experiment. This type of variability is more distinctly expressed in recordings from prefrontal (Radons et al. 1994; Seidemann et al. 1996), parietal (Seal et al. 1983; Requin et al. 1988), and motor areas (Vaadia et al. 1988), where both the processing of sensory inputs and the planning and execution of motor behavior is reflected in the neuronal activation patterns. In sensory areas, by contrast, the neuronal response is often tightly locked to the sensory stimulus, and thus latency variability is small (Richmond et al. 1990).

We describe here a novel nonparametric method for estimating the trial-by-trial differences in the temporal latency of dynamic spike responses. The majority of previously published methods (Ellaway 1978; Seal et al. 1983; Commenges and Seal 1985; Davey et al. 1986; Churchward et al. 1997; Baker and Gerstein 2001) are based on parametric models and rely on the estimation of one or more parameters during an initial period of stationary firing (see also Discussion). A nonparametric method related to ours has been described by Sanderson (1980).

In contrast to most other methods, the procedure described here does not attempt to estimate the absolute latencies of changes in neural firing rate as measured with respect to some external trigger event. Rather, for a set of N trials our method yields $N - 1$ relative latencies corresponding to an optimal alignment of all N trials such that latency variation is minimized. The improved trial-averaged response may then, if desired, be used to determine the absolute mean latency with a different method. The previously estimated relative latencies provide a quantification for the variance of the response latency.

We used surrogate data for the calibration of our method. The data were simulated by means of rate-modulated gamma processes. The underlying dynamic rate followed a phasic-tonic response profile, paramet-

Correspondence to: Stefan Rotter
(e-mail: rotter@biologie.uni-freiburg.de),
Tel.: +49-761-2032862, Fax: +49-761-2032860

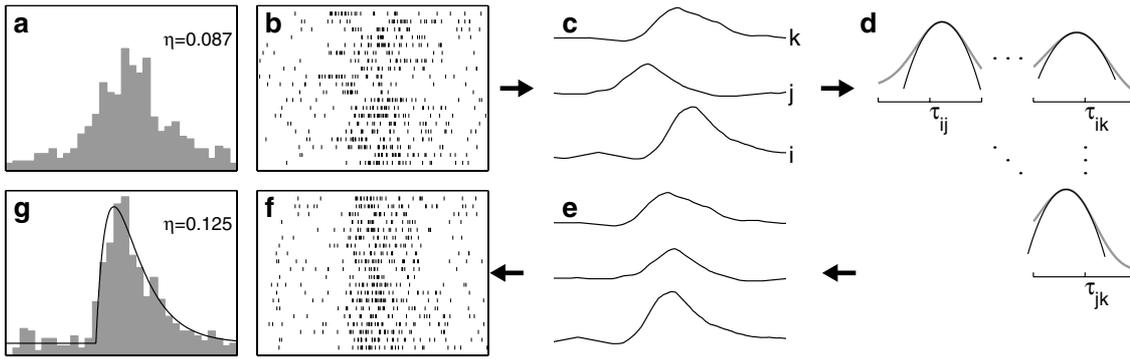


Fig. 1. Method for estimation and elimination of trial-by-trial response latency variations. **a** Peristimulus time histogram (PSTH) and **b** raster display of underlying spike trains for 20 trials in an observation interval of 800 ms. **c** Estimated single-trial rate profiles for three selected trials. **d** For each pair (i, j) of trials we calculate the crosscorrelation of rates as a function of the time lag τ_{ij} . Fitting of a parabolic function to each of the correlogram peaks allows us to determine the set of $N - 1$ relative latencies that maximize the total

pairwise crosscorrelation (see text). Compensating for these latencies results in realigned response profiles **e, f** and a much improved PSTH estimate **g** that matches the underlying rate function (*black line*) used for the simulation of the single-trial responses. The saliency of the response peak in the PSTH was quantified before and after realignment by the modulation index η that measures the relative difference in entropy of the area-normalized PSTH from a completely flat PSTH with the same number of bins (see Methods)

alized by response strength, response width, and background rate.

The application of our method is illustrated for a pair of neurons that were simultaneously recorded from the primary motor cortex of a behaving monkey. It allowed us to determine the temporal relation of the neural activity to various external trigger events. The analysis of neuronal interaction by means of crosscorrelation and the joint peristimulus time histogram (JPSTH) is improved through elimination of latency variability, allowing for a better estimate of the trial-averaged predictor. Preliminary results were presented in abstract form (Nawrot et al. 1999b).

2 Methods

2.1 Estimation of trial-by-trial latency differences

We assume an ensemble of N spike trains as the outcome of N experimental trials. All trials are aligned in time with respect to some external trigger event (Fig. 1b), e.g., the onset of a sensory stimulus or a behavioral event, such as pushing a button. We will first consider the case of one single unit, and then extend the method to the combined analysis of several simultaneously recorded units.

2.1.1 Estimation of single-trial rate profiles. In the first step, we obtain single-trial rate estimates by applying the method of kernel estimation (Parzen 1962; Silverman 1986; Nawrot et al. 1999a; Paulin and Hoffman 2001). For each of the N trials we obtain a rate profile

$$\lambda(t) := \sum_{i=1}^l K(t - t_i)$$

where t_1, \dots, t_l are the corresponding spike times. As the kernel function $K(t)$ we chose a symmetric triangular kernel according to

$$K(t) = \begin{cases} \frac{1}{6\kappa^2}(\sqrt{6\kappa} - |t|) & \text{for } |t| \leq \sqrt{6\kappa} \\ 0 & \text{otherwise} \end{cases}$$

with standard width κ and unit area.

The performance of single-trial rate estimation depends essentially on kernel width κ , which determines the time resolution. Here, we used an unsupervised method as described in Nawrot et al. (1999a) to find the optimal value for κ . We applied this method to each trial separately and then used the average width to finally estimate each of the N single-trial rate profiles using a single kernel. Figure 1c shows three exemplary estimates from simulated spike data. An alternative method for determining the optimal kernel width is described in Paulin and Hoffman (2001).

2.1.2 Maximization of total pairwise crosscorrelation.

Our goal is to determine the trial-by-trial differences in latency such that their compensation leads to an “optimal” temporal alignment of trials. A suitable alignment criterion for pairs of trials was found to be the maximal overlap of the single-trial rate functions. The N -fold crosscorrelation of all individual rate profiles seems to be a natural generalization of this approach. However, the corresponding optimization problem is computationally very costly, and impractical for typical trial numbers in the range of tens to hundreds. To circumvent this problem we propose here an alternative method involving three steps of analysis.

For N trials our optimization problem has $N - 1$ degrees of freedom. In other words, for the optimal alignment we may choose the time course of trial 1, say, as a reference ($\tau_1 := 0$) and then apply the $N - 1$ shifts $\tau_2^{\text{opt}}, \dots, \tau_N^{\text{opt}}$ to the remaining trials. For each pair (i, j) of trials we calculate the crosscorrelation separately:

$$C_{ij}(\tau_j - \tau_i) = \int \lambda_i(\tau_i + s)\lambda_j(\tau_j - s) ds$$

depending on the relative time lag $\tau_j - \tau_i$ between the i th and j th trial, as shown in Fig. 1d. For practical reasons, the maximum time lag m was restricted to half the length of the observation interval. In the second step, we fit a parabolic function $p_{ij}(\tau)$ to each of the $N(N-1)/2$ crosscorrelation functions (Fig. 1d) in a small neighborhood around τ_{ij}^{\max} where the correlation is maximal. Finally, we compute the sum P of all parabolas. This function is quadratic in the $N-1$ variables $\tau = (\tau_2, \dots, \tau_N)$ and possesses a unique global maximum. This maximum defines the shifts for the optimal alignment of all trials.

2.1.3 Multiple simultaneous recordings. In case we want to study higher-order interactions of two or more simultaneously recorded channels of spike activity, we must assume that all observed neuronal processes operate on the same time axis. Thus, when eliminating relative latencies, we must preserve the joint time axis. We can then either rely on the estimation of the relative latencies for one preferred neuron (or channel) with particular salient response profiles or take into account the data of all n neurons. To implement the latter, we chose the following procedure: we performed the single-trial rate estimation and the calculation of all pairwise trial crosscorrelations independently for each neuron. In the next step, we summed the correlation functions across all neurons, each with equal weight. All subsequent steps are the same as for one neuron. This approach is by no means the only way of jointly estimating the latency variations in multiple neurons (see, for example, Baker and Gerstein 2001). A more detailed investigation of alternative methods, however, is beyond the scope of this paper.

2.2 Calibration

To calibrate our method for estimating trial-by-trial differences in latency, we used surrogate data from stochastic point process simulations. We repeatedly generated spike trains according to a fixed dynamic rate profile. For all trials of an ensemble we introduced latencies that were independently drawn from a Gaussian distribution. An error measure was used to quantify the deviation between actual and estimated latencies, and its dependence on various parameters.

2.2.1 Simulation. We used stochastic point process models for simulating spike train data. We mostly employed nonhomogenous Poisson processes (Cox and Isham 1980; Teich et al. 1997; Nawrot et al. 1999a) where all points, or spike times, occurred independently of each other with an average rate intensity that changed in time according to a given rate profile $\rho(t)$. In additional simulations we used rate-modulated gamma processes (Reich et al. 1998; Baker and Lemon 2000; Baker and Gerstein 2000; Barbieri et al. 2001) of integer order k . Realizations for this model can be achieved by first simulating the nonhomogenous Poisson process

with k -fold rate and then eliminating all but every k th spike. Thus, the Poisson process is a special gamma process with order $k = 1$.

2.2.2 Dynamic rate functions. We employed a family of rate profiles $\rho(t)$ that model realistic peristimulus time histograms (PSTHs) typically encountered *in vivo*, as used earlier in Nawrot et al. (1999a). The rate function is comprised of a stationary background rate b , which accounts for spontaneous activity, and an additional phasic response profile with the time course of a beta function:

$$\beta_{\tau}(t) := \begin{cases} \frac{1}{\tau}(e^{-\frac{t}{\tau}} - e^{-\frac{t}{\tau}}) & \text{for } t \geq 0 \\ 0 & \text{for } t < 0. \end{cases}$$

This function is normalized to unit area and has a decay time constant that is twice the rise time constant τ . The standard width of the beta function is

$$w = \sqrt{5}\tau$$

which is used for parametrization. Finally, we define the following family of rate functions:

$$\rho_{b,A,w}(t) := b + A \cdot \beta_w(t - t_0)$$

where t_0 denotes the response onset. The factor A corresponds to the area under the phasic response profile. It determines the number of excess spikes that can be expected during the response in addition to the background spikes. We chose $A > 0$ to model an excitatory response. Figure 1g shows an example with background rate $b = 10$ Hz, an expected number of $A = 20$ response spikes, and a response width of $w = 100$ ms. The simulations were performed for a variety of realistic choices of parameter combinations (b, A, w).

2.2.3 Measure of performance. For each ensemble of trials we first repeated the simulation N times using exactly the same rate profile. We then randomly drew $N-1$ relative latencies from a normal probability distribution with standard deviation σ , restricted to the symmetric interval that accounts for 99% of its area. From this ensemble of simulated spike trains we then estimated the relative latencies. The error ϵ of estimation was quantified by the standard deviation of the differences between the actual latencies (which are known here) and the estimated latencies. We simulated 500 ensembles with N trials each for several combinations of simulation parameters. Based on such ensembles (see example in Fig. 2) we obtained a good estimate of the mean error $\bar{\epsilon}$, which we used as the measure for mean performance. Moreover, we determined the reliability of the latency estimation by the standard deviation of errors $SD(\epsilon)$ across all ensembles. Clearly, the variance of estimation depended on the number of trials per ensemble. We chose $N = 20$ trials, unless stated otherwise. This is a realistic number that applies even in the case of complex experimental paradigms.

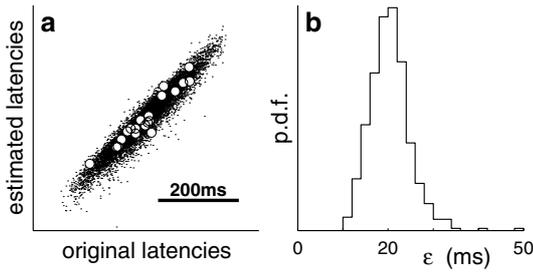


Fig. 2. Performance of latency estimation for simulated data. **a** Original latencies versus estimated latencies. The black dots show the results of an application of our method for 500 ensembles of 20 trials each, with simulation parameters $k = 1$, $b = 10$ Hz, $A = 20$, $w = 100$ ms. The open circles illustrate the latencies of the trials in the example of Fig. 1. **b** Distribution of the error measure ϵ for all 500 ensembles. On average, the error was $\bar{\epsilon} = 20.4$ ms, the standard width of the original latency distribution was 75 ms. The reliability of the error estimation was $\text{SD}(\epsilon) = 4.5$ ms

2.3 Modulation index

The error measure ϵ can be applied only in the case of surrogate data, where the “true” latencies are known. In the case of real data we need other indicators that allow one to assess the goodness of the estimate. Typically, the realignment of trials according to the estimated latencies sharpens the trial-averaged response profile calculated with the PSTH.

Here we describe one way of quantifying this effect. To this end we define the modulation index as a measure that quantifies the deviation of any given histogram from the uniform histogram with the same number of bins, but with all bins having the same value. Consider a PSTH comprised of l bins with entries (h_1, \dots, h_l) . We first normalize it to unit area, i.e., we divide each of the entries by the sum over all entries. The resulting histogram may then be interpreted as a probability density function. For each particular bin i the respective entry then gives the probability p_i of finding a spike within that bin. We can now calculate the entropy of this distribution and divide it by the maximum entropy that would be assumed by the uniform histogram with the same number of bins. Subtracting this number from unity leaves us with the definition of the modulation index as

$$\eta := 1 - \left(\sum_{i=1}^l p_i \log_2 p_i \right) / \log_2 l$$

which is a number in $[0, 1]$. For a uniform histogram with l bins we have $p_i = 1/l$ for all i and, therefore, $\eta = 0$. By contrast, for a histogram with a single nonzero entry, we obtain $\eta = 1$.

2.4 Time-resolved Fano factor

The number of observed spikes within a fixed observation interval I varies across repeated trials. The Fano factor is an established measure for this variability across trials. Here, we define its empirical estimate

$$\hat{F}_I = \frac{\text{Var}(\text{count})}{\text{Mean}(\text{count})}$$

as the variance of counts observed across individual trials of duration I , normalized by the mean count in that interval. In the framework of point process theory, the Fano factor is defined in the limit of infinitely long observations. For a gamma process, the Fano factor is then equal to the inverse of the order parameter k (Tuckwell 1988). Thus, the special case of the Poisson process has a Fano factor equal to unity. However, for intervals I of finite length, the empirical estimate \hat{F} as defined above exhibits a strong dependency on the expected count within this interval (Cox 1967; Teich et al. 1997; Ratnam and Nelson 2000; Chacron et al. 2001; Nawrot et al., in preparation). An analytic expression for this dependence can be found for a gamma process with integer order (Cox 1967). For $k > 1$, the empirical measure \hat{F}_I lies in the range $1/k < \hat{F}_I < 1$, with a stronger bias toward unity for lower expected counts in I (Nawrot et al., in preparation). In the special case of the Poisson process, the Fano factor is 1, irrespective of the length of the interval.

Here we employed a time-resolved version of the empirical Fano factor $\hat{F}(t)$ computed for a moving time window of fixed length (Oram et al. 2001). For nonstationary rates in the case of misaligned trials due to variable trial-by-trial latencies, the time course of $\hat{F}(t)$ will falsely indicate too high values in periods of strong rate changes. This is demonstrated for an ensemble of gamma data in Fig. 3a. After realignment of all N trials, this effect is much reduced (Fig. 3b). Closer inspection

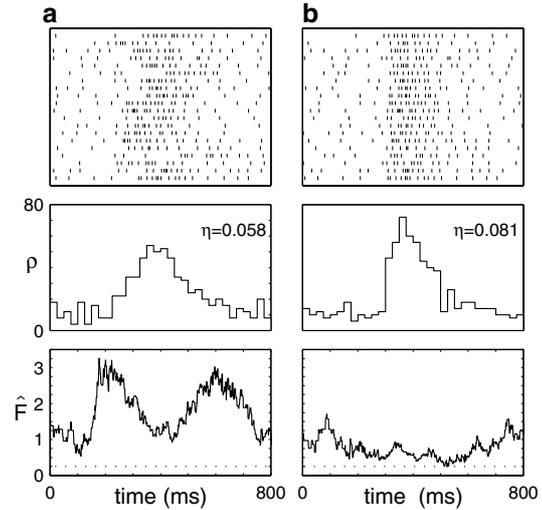


Fig. 3. Time-resolved Fano factor. **a** Before compensating for latency differences, the empirical Fano factor estimated in a moving window 300 ms wide exhibits strong positive excursions due to the misalignment of trials around times where the rate increases or decreases. **b** The realigned ensemble yields a more stationary time course of $\hat{F}(t)$. The total time of observation is 800 ms, as in Fig. 1. The simulated point process was a gamma process with order parameter $k = 4$, which generated more regular spike patterns than in the Poisson case (Fig. 1); further simulation parameters were $b = 10$ Hz, $A = 10$, $w = 100$ ms, $\sigma = 75$ ms. The dotted line indicates the limiting case for the Fano factor given by $1/k = 0.25$ (see text)

reveals that after realignment the Fano factor is increased on either edge of the observation interval. This is due to the low (background) rate in these periods. The low expected count of three spikes per window introduces a bias of the Fano factor estimation toward 1. Only during the period of high rate is this bias reduced, and the curve approaches the limiting value of $F = 0.25$ for a gamma process of order $k = 4$ (dotted line). Occasional excursions beyond the Poisson mark of $F = 1$ are primarily caused by the variance of the Fano estimator, due to the low number of trials and the small window size, and possibly to residual latency fluctuations that could not be eliminated.

In the case of neuronal data, for which we do not know the “true” latencies, an overall reduction in $\hat{F}(t)$ due to elimination of latency variations can be expected. Baker and Gerstein (2001) demonstrated that minimizing the time-averaged variance of single-trial rate profiles can serve as an a posteriori criterion in an alternative algorithm to estimate cross-trial latencies (see Discussion). Here we suggest the time-resolved Fano factor $\hat{F}(t)$ as a post hoc indicator for successful trial alignment, which is independent of the criterion used for the latency estimation itself.

2.5 Computational costs

For practical purposes we need to know the computation time T of the proposed algorithm, depending on parameters like the number of trials per ensemble and the temporal resolution of the experimental data. Our method comprises two separate operations. First, we determine the optimal kernel width κ . Here, the computational cost increases linearly with the number of trials, the number of tested kernel widths, and the number of neurons. Far more demanding is the second operation of estimating the relative latencies. Here, the pairwise calculation of the crosscorrelations for all possible pair combinations of trials leads to an increase in computation time proportional to the squared number of trials, N^2 . For a sample set of 50 trials, an observation interval of 1 s in length and a temporal resolution of 1 ms, our algorithm required 2.2 s for first estimating κ on the base of 20 tested kernel widths and then 21 s for determining all 49 per-trial latencies on a 1-GHz Pentium III.

All numerical calculations and point process simulations were performed with Matlab (Mathworks Inc.). Crosscorrelations were calculated using the “xcorr” function of the Signal Processing Toolbox of Matlab, which employs a fast Fourier transform.

3 Results

3.1 Calibration results

In this section we investigate the performance of the proposed method for latency estimation in terms of the error measure ϵ . First, on the basis of Poisson simula-

tions we will determine how mean performance and reliability depend on the parameters (b, A, w) of the rate profile. Second, we will see how these results change when we alter the order k of the gamma process, which controls its statistical properties. We then explore the dependence on the number of trials N within each ensemble and the scaling behavior in the case of n simultaneously recorded neurons. Finally, we test the performance in the presence of additional trial-by-trial variability, introduced by fluctuations of rate amplitude across trials. Unless stated otherwise, the following results are based on point process simulations of 500 ensembles of 20 trials each for each of the parameter combinations.

3.1.1 Dependence on rate parameters. For simplicity we first investigate the case of zero background firing ($b = 0$) corresponding to a neuron that has no spontaneous activity and generates spikes only in response to a stimulus. In this case, the performance of the algorithm shows a rather simple dependence on the remaining two rate parameters A and w . Keeping the response strength A fixed, the mean error of estimation $\bar{\epsilon}$ increases linearly with the increase in response width w , as demonstrated in Fig. 4b. Conversely, if w is kept fixed and the number of expected response spikes A is increased (Fig. 4a), the mean error decreases according to a power law $\bar{\epsilon} \propto A^{-m}$, with m ranging between 0.5 and 0.6 for all tested parameter pairs.

When a constant background firing rate ($b > 0$) is added to the response profile, the picture becomes more complicated. Generally, the Poissonian background adds noise to the single-trial rate estimates and, hence, leads to a decreased performance. In Figs. 4c and 4d we repeat the analysis of Figs. 4a and 4b for a relatively high background rate of $b = 20 \text{ s}^{-1}$. In the case of a weak response of $A = 5$ spikes, the algorithm produces equally large errors for all four response widths in Fig. 4c. Similarly, in Fig. 4d and for $A = 5$ (top curve) the error approaches the width $\sigma = 75 \text{ ms}$ of the initial latency distribution (dashed line) when increasing the response width. This saturation is a desired effect of the algorithm in cases where the single-trial rate estimate cannot faithfully capture the weak and broad response on a high background. Then, many trial-by-trial latencies are effectively estimated to be zero, i.e., the original alignment remains unchanged in many cases, and thus the mean error $\bar{\epsilon}$ across many ensembles approaches the standard error of the original latency distribution.

Generally, the width of the residual latency variations after trial realignment quantified by ϵ is essentially independent of the distribution of the original latencies (not shown), i.e., the performance of the algorithm is independent of the width σ .

Note that the calibration results presented in Figs. 4 through 6 demonstrate parameter dependencies that are universal in their shape. The absolute scaling of errors $\bar{\epsilon}$, however, is influenced by additional parameters of the method, in particular the length of the observation interval I and the maximal time lag of the crosscorrelation function. Based on prior knowledge about the temporal

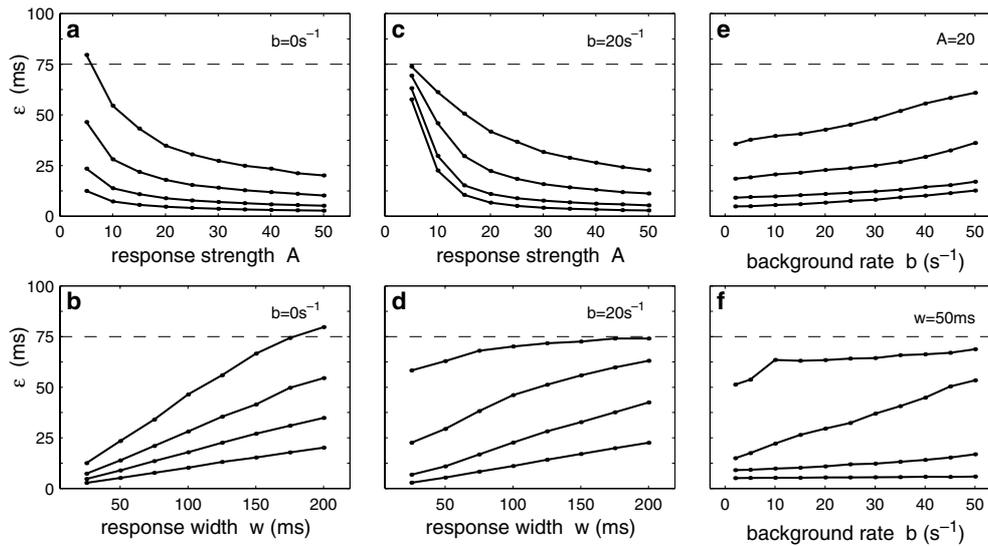


Fig. 4. Performance of latency estimation for different combinations of rate parameters. **a, b** Zero background firing ($b = 0$). **a** The mean error of the estimate decreases rapidly with increasing response strength for fixed response widths of $w = 25, 50, 100, 200$ ms (bottom to top). **b** Conversely, keeping the expected number of response spikes fixed to $A = 5, 10, 20, 50$ (top to bottom), the error increases approximately linearly with the response width. **c, d** Fixed background firing at 20 s^{-1} , other parameters as in **a, b**. On the high

Poisson background, a faithful estimation of latencies is only feasible for large ($A \gtrsim 10$) or narrow ($w \lesssim 100$ ms) responses. **(e, f)** Performance as a function of background rate. **e** Response equals $A = 20$ spikes on average, response widths as in **a, c**. **f** Responses with a fixed width of $w = 50$ ms yield a mean error of estimation that increases linearly with the background rate, only for a weak response ($A = 5$, upper graph) the error is always high, irrespective of the background rate

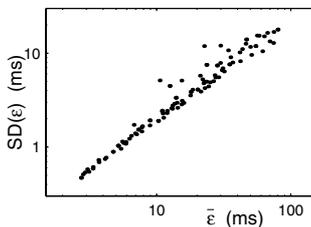


Fig. 5. Reliability of latency estimation. Standard deviation of errors vs. the mean error across the ensemble of 500 trials for all parameter combinations (b, A, w) shown in Fig. 4a–d. There exists an approximately linear relation between both parameters (see text)

structure of the data, the observer can make an educated guess regarding the position of the responses within trials as well as on the range of per-trial latencies that can be expected. This allows one to confine the observation interval. The shorter this interval, the weaker will be the distorting influence of background firing on the optimized single-trial rate estimation and on the detection of the peaks in the pairwise crosscorrelation functions. To ensure comparable calibration results, we chose a fixed observation interval of 1 s throughout this paper.

3.1.2 Reliability of latency estimation. So far we have investigated the average performance in terms of $\bar{\epsilon}$, calculated as the mean error across 500 ensembles. Considering the variance of the errors enables us to assess the reliability of the estimator. It turns out that the standard deviation exhibits an approximately linear dependence on the mean error $\bar{\epsilon}$, independently of the

rate parameters. This is demonstrated in Fig. 5 for a large variety of rate parameter combinations (b, A, w). A linear fit on double-logarithmic scale yields a slope of 1.06.

3.1.3 Dependence on gamma-order k . When increasing the order parameter k of the gamma process starting from 1 (Poisson), we know that the number of spikes per trial becomes less variable (see Methods). Thus, we can expect that, due to the associated reduction in variability of the single-trial rate estimates, the performance in estimating the relative latencies will improve. This effect is indeed observed in Fig. 6a for three example response profiles (b, A, w). The mean error $\bar{\epsilon}$ decreases with increasing gamma order roughly as $1/\sqrt{k}$. Even more substantial is the reduction of the standard deviation of estimation $SD(\epsilon)$ shown in Fig. 6b. Generally, for each value of k the relation between mean error and standard deviation of errors remains approximately linear (not shown), independent of the rate parameter combinations, as found earlier for the Poisson process (Fig. 5).

3.1.4 Multiple simultaneously recorded neurons. How does the joint estimation of latency variations in several simultaneously recorded neurons affect the quality of the estimation? To investigate this question we first calibrated our method for an increasing number of up to ten neurons, where all neurons were simulated with identical rate functions. The results are shown in Fig. 6c, again for the same three response profiles used throughout the subplots of this figure. The mean error decreased approximately reciprocally to the square root of the number of neurons n . Similarly, the reliability increases

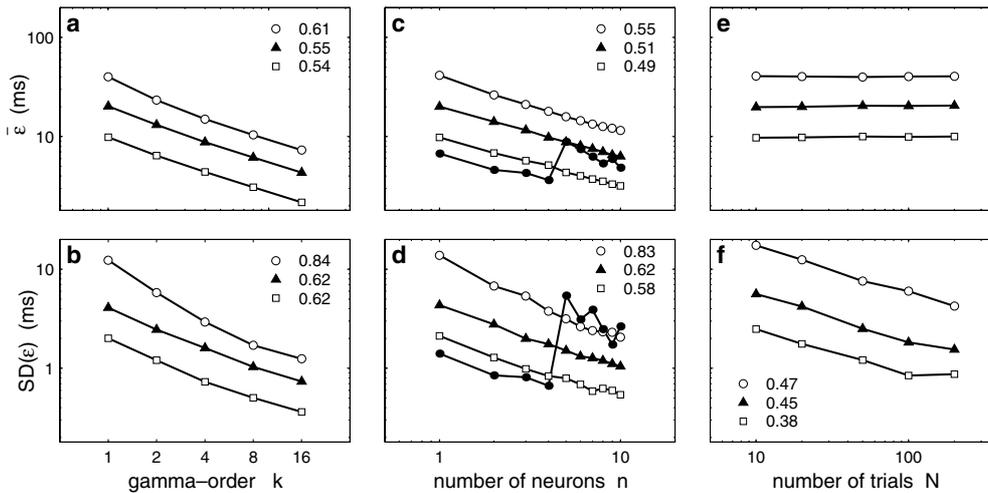


Fig. 6. (a) Mean performance of latency estimation is improved for gamma processes less variable than the Poisson process. The mean error decreases with increasing order parameter k according to a power law approximately as $1/\sqrt{k}$. The legend gives the values of the negative exponent for a linear fit on a double-logarithmic scale. The same rate parameters are used for the three examples (\circ , \blacktriangle , \square) in all plots a–d. b The reliability of the error estimation improves along with the mean performance as k is increased. c Increasing the number n of simultaneously recorded neurons in a combined analysis reduces the mean error, again approximately as $1/\sqrt{n}$, and d increases reliability if

all neurons respond with identical rate parameters (b, A, w). If these parameters are randomly assigned (\bullet), both mean and variance of the error may increase again when including neurons with latency variations that are particularly hard to estimate. e As expected, the mean performance is independent of the number of trials N . f However, a larger sample of spike trains reduces the variance of the estimation. Rate parameters of examples: $\circ b = 10 \text{ s}^{-1}$, $A = 10$, $w = 100 \text{ ms}$; $\blacktriangle b = 10 \text{ s}^{-1}$, $A = 20$, $w = 100 \text{ ms}$; $\square b = 10 \text{ s}^{-1}$, $A = 20$, $w = 50 \text{ ms}$. c–f based on Poisson simulations

while $SD(\epsilon)$ decreases with n . If, however, the rate parameters of the neurons differ (simulated here by random assignment of values), then the inclusion of a single neuron for which an estimation of latencies is particularly difficult may confound both performance and reliability of the estimation procedure. In Fig. 6c, d, this is shown for one selected ensemble of neurons where the inclusion of one such “problematic” neuron (here: neuron no. 5) spoils both the mean and the standard deviation of errors. This particular neuron had been simulated with a small response of $A = 5$ spikes, superimposed on a high background level of $b = 20$ spikes per second.

3.1.5 Dependence on number of trials. From any method of estimating variations in response latency we would require that its performance be independent of the number of trials, whereas we can expect a higher reliability for a larger sample size. Both effects are indeed demonstrated in Fig. 6e, f.

3.1.6 Performance in the presence of trial-by-trial nonstationarities of firing rate. So far we have simulated neuronal spike train data by means of stochastic point processes with a rate function that was fixed across trials. In this case, and in the absence of variations in response latency, the resulting degree of trial-by-trial variability in the number of spikes depended solely on the stochastic nature of the point process model (see Methods).

To test our method under more stringent conditions, with additional sources contributing to a larger trial-by-trial variability in spike count, we introduced two separate types of cross-trial nonstationarity in firing rate.

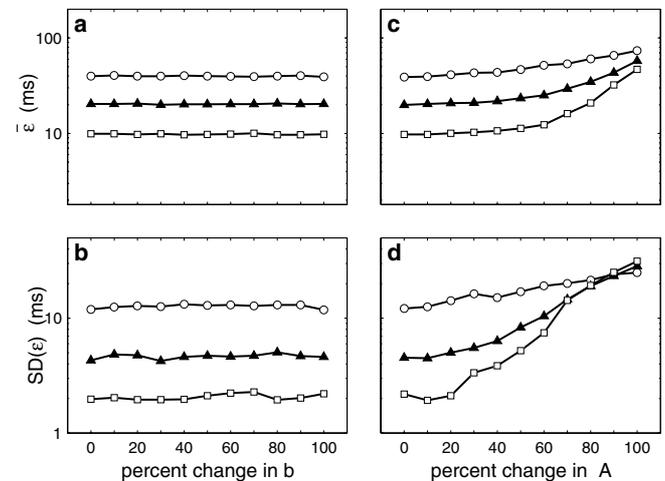


Fig. 7. Performance in the presence of additional trial-by-trial nonstationarity in rate. a Mean performance of latency estimation and b reliability are independent of the degree of nonstationarity in background rate b . c Nonstationarity of the response strength A increases the mean error. This effect becomes significant for large trial-by-trial variations. d The reliability of the error estimation decreases strongly with larger variations in response amplitude. The total range of possible values for the background rate in a, b and for the response strength in c, d is given in percent of the mean rate b or the mean response strength A , respectively (see 3.1.6). Symbols as in Fig. 6

First, we varied the background rate on a trial-by-trial basis while keeping the response strength fixed. In each trial i , the actual background rate b_i was drawn randomly from a uniform distribution on $[b - \Delta b, b + \Delta b]$ around its mean b , where Δb is given in percent of b . In effect, the variance of spike count, as measured in repeated trials, increases with increasing width of the

uniform distribution. The results of our calibration are summarized in Fig. 7a, b. Both the mean performance and the reliability of the estimation were unaffected by the additional nonstationarity of the background firing rate, even for variations up to $\pm 100\%$ of the mean value b . Second, we varied the response strength A , and thus the amplitude of the superimposed phasic response, while keeping the background fixed. To this end the response strength A_i of trial i was drawn randomly from a uniform distribution $[A - \Delta A, A + \Delta A]$ around the mean A , with ΔA given in percent of A . Such nonstationarity impairs the mean performance as well as the reliability of the latency estimation for response amplitude fluctuations larger than 50% of the mean response strength, as shown in Fig. 7c, d.

3.2 Application to experimental data

In this section we demonstrate the application of our method to experimental data. We investigated a pair of neurons recorded simultaneously with two separate electrodes in the hand area of the primary motor cortex of a behaving monkey (*Macaca mulatta*). Details of the experimental paradigm have been described elsewhere (Grammont and Riehle 2002). In short, the monkey performed a delayed hand-reaching task from the central resting position to one out of six targets arranged equidistally in a circle. During each trial, two signals were presented to the animal. The first, the preparatory signal (PS), provided prior information about the target to be pointed after the second, the GO signal. Two delay durations (600 and 1200 ms) were presented at random with equal probability. Only successfully completed trials were used for further analysis.

In the subsequent analysis, all trials of equal direction and equal delay were grouped together, yielding 2×6 trial ensembles for each neuron, with an average trial number of $N \approx 20$. As an initial trigger for the alignment of all trials within an ensemble we chose the appearance of the GO signal at time $t = 0$.

3.2.1 Elimination of trial-by-trial differences in latency. To estimate the relative latencies of these neurons, we chose an observation interval of 800 ms duration, starting 100 ms prior to the appearance of the GO signal. The upper raster plot in Fig. 8a shows the single spike trains for neuron 1 as recorded during $N = 19$ experimental trials with the same movement direction and for short delays. The simultaneous recording of neuron 2 is shown in Fig. 8b. The observation interval of 800 ms is indicated by the gray shadings on either side. For each direction and each delay period the application of our algorithm resulted in $N - 1$ latencies relative to the first trial. Compensating for these relative latencies resulted in the realigned ensembles of the lower raster plots in Fig. 8a, b. Accordingly, the gray boundary is now jagged, indicating the shifts applied to each trial.

How can we know whether our algorithm produced meaningful results? Since the “true” underlying latencies

in the activity relative to external time are not measurable, there is no way to directly assess the performance of our estimation. Thus, we must rely on indirect evidence for a successful estimation. The visual comparison of the raster displays may serve as a first subjective judgement of the success of the realignment procedure (Fig. 8a, b). Similarly, we can observe that the rate profile in the PSTH has sharpened. We can quantify the sharpness of the measured response by calculating the modulation index η . As indicated in Fig. 8c, d, for both neurons the realignment of trials led to a substantially higher value η_2 than the original alignment with η_1 . The insets in the lower panel show the distribution of the modulation index η as calculated for 1000 random permutations of the trial order of the estimated relative latencies. As expected, already the original alignment to the GO signal yielded a fairly high value η_1 , compared to the values obtained for random shufflings of the estimated latencies. Nevertheless, the value for the realigned trials η_2 is much higher, demonstrating the significance of the improvement as measured by the modulation index.

Further, we may investigate the changes in trial-by-trial variability by means of the time-resolved Fano factor. As seen in Fig. 8e, f, the strong excursions in $\hat{F}(t)$ around times of strong rate modulations are mostly abolished as a result of the realignment. Calculation of the Fano factor for the complete observation interval would yield only a negligible change since the overall spike counts change only minimally due to minor increases or decreases at the borders of observation. It is thus important to choose a short enough sliding window so as not to average out the rate modulations. In other words, the width of the counting window should roughly equal the time constant of rate modulation κ , estimated in the first step of our algorithm (see Methods). To avoid a time-consuming visual inspection of the $\hat{F}(t)$ curves, we can calculate the reduction in mean Fano factor over time (61% and 45% in Figs. 8e and 7f, respectively) as a characteristic parameter to assess the quality of the latency estimation. We note in passing that, while neuron 1 with a Fano factor of about 1 satisfies the expected variability of a Poisson process, neuron 2 has a distinctly lower Fano factor, indicating considerably less variability than expected for Poisson (MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted).

3.2.2 Simultaneously recorded neurons. It is well known that the changes in activity of neurons in the motor cortex may be related to either stimulus presentation and movement preparation or to movement execution, or they may show an intermediate behavior (e.g., in parietal area 5: Seal et al. 1983; Requin et al. 1988; in motor cortex: Riehle and Requin 1989; Riehle 1991; Miller et al. 1992; Riehle et al. 1994, 1997; Zhang et al. 1997; in prefrontal cortex: Radons et al. 1994; Seidemann et al. 1996; in frontal eye fields: Bruce and Goldberg 1985). This presumably reflects the differential involvement of individual neurons in different stages of processing, from extracting relevant information about the stimulus to preparing a movement and initiating and controlling its

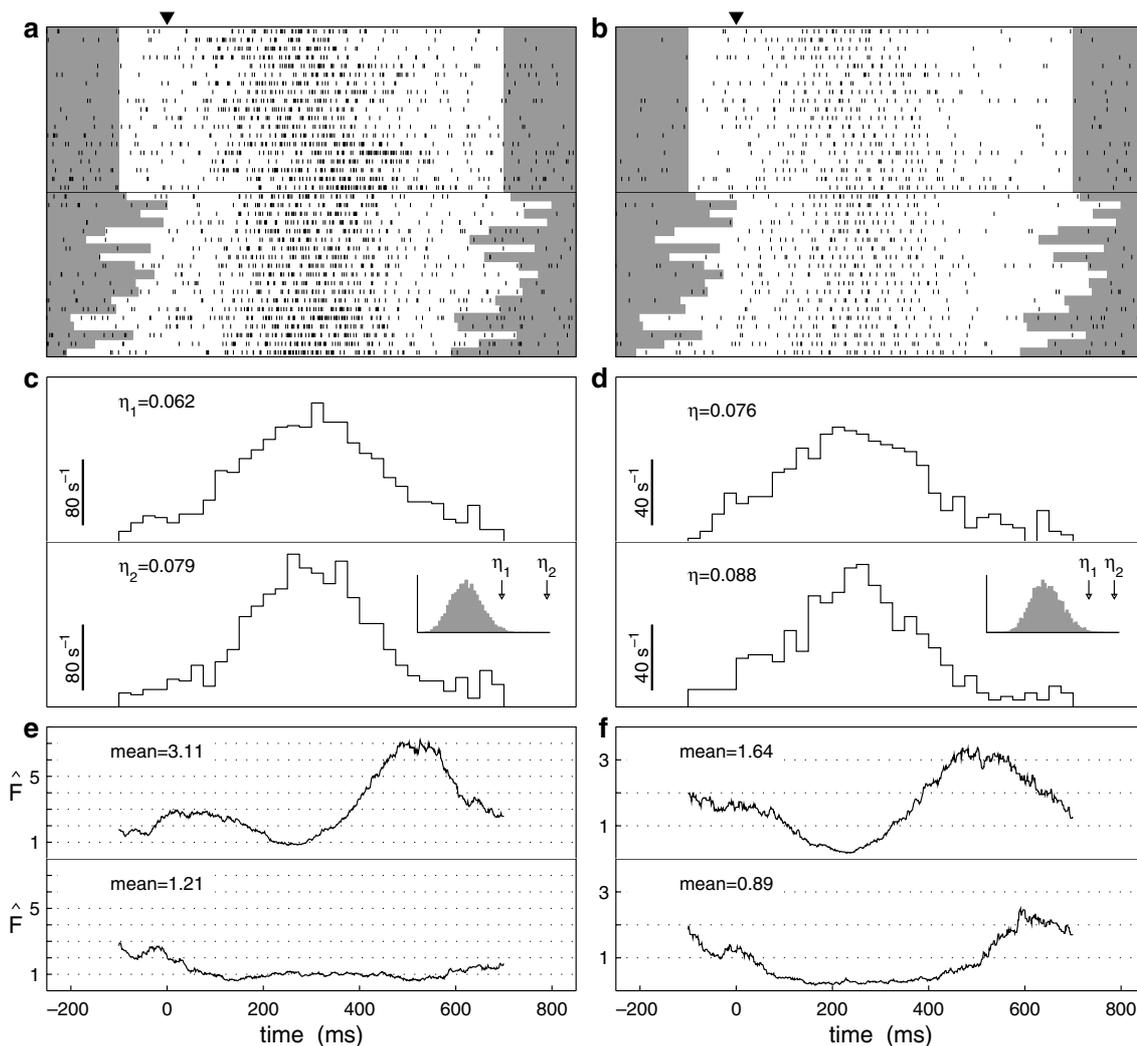


Fig. 8. Elimination of estimated latency variations sharpens the PSTH and reduces the time-resolved Fano factor in two simultaneously recorded neurons. **a, b** Spike raster diagrams with trials aligned to the GO signal (\blacktriangledown) at $t = 0$ (*upper panels*) and realigned to compensate for the estimated per-trial latency differences (*lower panel*) for neurons 1 and 2. The *gray shadings*, which in the upper diagram restrict the observation interval used by the algorithm, in the lower diagram indicate the individual shifts. The sum of all shifts is set to zero. **c, d** PSTH estimates of trial-averaged rate function for the alignment to the GO signal (*upper panel*) and after trials have been

realigned to eliminate response variability (*lower panel*). The saliency of the response peak in the PSTH was quantified by the modulation index η , which is increased through trial realignment by about 27% and 16% in neuron 1 and 2, respectively. Inset: chance distribution of modulation index η_2 calculated from randomly assigning the relative latencies to trial numbers. **e, f** The time-resolved Fano factor before (*upper*) and after (*lower*) realignment, measured in a moving window of 300 ms in length with 1 ms time resolution. In both neurons the elimination of the estimated latency variations clearly reduced the Fano factor throughout the 800 ms of observation

execution. Thus, when observing two or more cells at a time, it is not at all obvious whether and to what degree such neurons are temporally related. In the first application of our method, we addressed this question by making use of the relative latencies estimated separately for each neuron. These latencies provide an “internal” trigger for each neuron, independent of external cues. Figure 9a shows a scatter diagram of the individual per-trial latencies of neuron 1 vs. those of neuron 2 for all 12 ensembles of trials. The strong correlation of the two variables ($r = 0.94$) in this case reflects a tight temporal coupling in the occurrence of the respective neuronal responses and, thus, of the two neurons’ time frames. To further quantify this similarity, we aligned all trials according to the internal trigger

events of one neuron and then measured the standard deviation of latency differences in the other one. For our example and across all trials from all 12 ensembles, this yields a relatively small number of 20.1 ms compared to the estimated time constant of rate modulations of $\kappa \approx 40$ ms. This indeed strongly suggests that these two spatially separate neurons were involved in a single, task-dependent dynamic process with a common temporal frame of reference.

3.2.3 Distinction between preparation and execution-related neural activation. Here we illustrate the application of our method to the question of whether the neurons’ activity is temporally related to the visual GO signal or, rather, to movement onset (MO). Having established

that both our example neurons share the same time frame, we can now estimate the latency differences jointly for both neurons (see Methods). Figure 9b shows the scatter diagram of the joint trial-by-trial latencies against the reaction times (MO-GO) for all 12 trial ensembles. The high correlation between these two variables indicates that the two neurons are related to movement execution, rather than to movement preparation, i.e., their activity is better locked to movement onset than to the GO signal or the preparatory signal (PS), respectively. Thus, MO would have been a more adequate external alignment trigger than GO (Fig. 8a, b). We may quantify the relatedness to any available external trigger by measuring the standard deviation of per-trial latencies relative to that trigger (Commenges and Seal 1985). In the example of Fig. 8a, this yields values of 65.6 ms for GO and 28.8 ms for MO. For the jointly estimated latencies, the average across all 12 ensembles yields 61.2 ms for GO and 25.5 ms for MO. Again, this clearly indicates that the changes in activity are better related to movement execution than to signal presentation.

In the next step, we consider whether the “residual” latencies relative to the behavioral event MO are due solely to the statistical properties of the observed spiking process and, hence, can be considered as random fluctuations. Alternatively, they might signify that the neural activity is temporally locked to another external (or internal) event that has not yet been considered. If we assume again that both neurons share the same time frame, then the differences in per-trial latencies estimated separately for each neuron (Fig. 9a) should be of stochastic origin. For our example pair of neurons, these differences of cross-neuron latencies yield an average standard deviation of 20.1 ms across all trials and all ensembles. The spread of residual latencies (25.5 ms) is of comparable size and might, thus, indeed represent a variability due to the random properties of the spiking process. Therefore, in this case it is unlikely that we could find a trigger event that matches the observed latencies better than movement onset.

3.2.4 Improved correlation analysis. The classical shuffle-corrected crosscorrelogram (Perkel et al. 1967) and other methods for investigating precise temporal correlations among multiple neurons (e.g., Gerstein et al. 1985; Aertsen et al. 1987, 1989) are all based on the general assumption that for each neuron the spike count in each bin is identically distributed and independent across trials.

One possible violation of the null-hypothesis is, as Brody (1999) correctly pointed out, brought about by variations in trial-by-trial latency with respect to the chosen trigger event (see also Baker and Gerstein 2001; Grün et al. 2002). In that case, the PSTH obtained by averaging across trials is an incorrect (usually blurred) descriptor of the rate dynamics within a trial. Consequently, the correlation dynamics predicted from such incorrect rate descriptors must also be incorrect. We demonstrate here how the estimation and subsequent elimination of trial-by-trial latency variability can

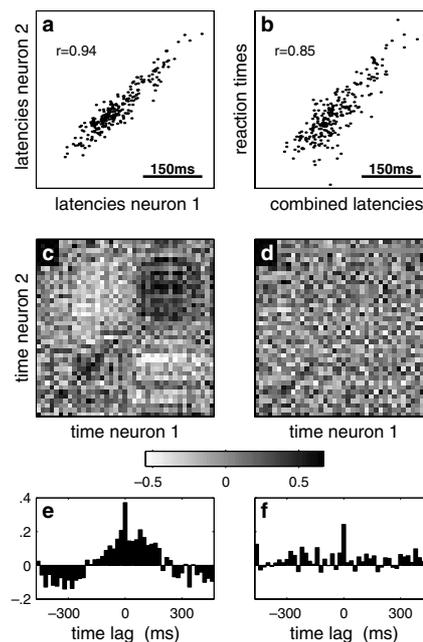


Fig. 9. **a** Separately estimated trial-by-trial latencies for neuron 1 vs. those of neuron 2, which were simultaneously recorded. Strong correlation indicates processing in the same time frame. **b** Strong correlation of jointly estimated latencies and reaction times indicates that the activity of both neurons is related to movement execution. **c** When trials are aligned to the visual trigger, the normalized JPSTH shows broad features of low and high correlation that are symmetric with respect to the diagonal. These artifacts are reduced in **d** through a realignment of trials according to the jointly estimated latencies. **e** The shuffle-corrected crosscorrelogram shows a broad positive central peak and negative side peaks due to the covariation in latencies. These effects disappeared after the realignment of trials in **f**. *Gray scale in c, d and vertical axis in e, f denotes linear correlation coefficient; time resolution is 20 ms; same 19 trials as in Fig. 8*

improve the measurement of spike correlation by the normalized JPSTH (Aertsen et al. 1989). First, using the two trial ensembles shown in Fig. 8a, b with trials aligned to GO (upper raster plots), we computed the JPSTH matrix in Fig. 9c for 800 ms of observation, starting 100 ms prior to the GO signal. Each bin was 20 ms wide. The matrix shows a broad pattern, suggesting a modulation in the pairwise correlation of the two neurons. Particularly salient is the feature of high correlation (dark patch) toward the end of the trial, roughly symmetrical with respect to the main diagonal, leading to a broad peak in the normalized shuffle-corrected crosscorrelogram shown in Fig. 9e. Likewise, the negative dips in the crosscorrelogram would suggest an anticorrelation for large time lags, also apparent in the two off-diagonal regions (bright patches) of the JPSTH matrix.

Note, however, that the clear horizontal-vertical (instead of diagonal) layout of the features in the JPSTH matrix is a signature of incorrect normalization and, hence, an indicator that excess correlation and/or modulations thereof are probably delusive (Aertsen et al. 1989). Instead, the temporal width of the correlogram features (Fig. 9c, e) in this case most likely reflects the temporal spread of latency fluctuations that we estimated earlier (Fig. 8a, b).

Indeed, after realigning trials such that the estimated trial-by-trial variations in latency are eliminated, all broad features in the recalculated JPSTH disappeared (Figs. 9d, f). The sharp peak in the central bin of the crosscorrelogram (Fig. 9f) remained, however, its height being reduced approximately by the height of the broad peak in Fig. 9e. It predominantly reflects the sequence of dark bins along the early part of the main diagonal of the JPSTH, indicating a correlation at zero time lag and with a precision of 20 ms in the early part of the observation interval that clearly exceeds expectation.

4 Discussion

4.1 Calibration results

We presented a nonparametric method for eliminating response latency variability across trials. For a given ensemble of spike trains, it determines the alignment of trials that maximizes the total pairwise correlation of single-trial rate profiles.

Our method was calibrated by means of numerical simulations of stochastic point processes. The resulting calibration plots (Figs. 4–6) describe how its performance depends on the relevant features of phasic rate response profiles and how the stochastic nature of the underlying point process influences its capability to eliminate latency variations.

We demonstrated the application of our method for a sample data set from a pair of neurons simultaneously recorded in monkey primary motor cortex. We found that compensation for the estimated trial-by-trial latencies clearly reduced the time-resolved Fano factor in both neurons, and eliminated broad features initially present in the crosscorrelogram and the JPSTH, which therefore must be attributed to the covariation in latencies (Brody 1999).

The most important control parameter of the proposed method is the length of the observation interval I . To ensure comparability in our various calibrations we used a fixed observation interval of 1 s. This was found to be suboptimal in cases of narrow or weak responses. Generally, this parameter should be chosen as small as possible to minimize the influence of “noisy” background activity that is not part of the response. For a given data set, an educated guess on location and length of the optimal window can be inferred, for instance, from the initial PSTH or from the estimated single-trial rate profiles. In cases where strong latency variations are expected, the method may be iterated using successively shorter intervals.

4.2 Application to experimental data

The method presented is not limited to the analysis of spike trains. In fact, the single-trial rate functions can be replaced by any type of analog signal. Knott (2001) successfully used our method to determine the spatio-temporal spread of local field potentials in the hippo-

campus recorded with planar multi-electrode arrays in organotypical slice cultures. The method may be further adapted to meet specific needs. Instead of calculating and maximizing the pairwise correlation, it might be advantageous to use a different criterion, e.g., to seek minimization of the mean squared difference in rate calculated for each pair of trials. Similarly, it might be desirable to align the first or second temporal derivative of the signal rather than the signal itself (Fries et al. 2001).

Often it is of interest to estimate the absolute latency of a response with respect to a sensory stimulus or a behavioral event. In this case, one may first align all N trials based on pairwise trial correlations using our method and then, in a second step, apply an independent method (e.g., Ellaway 1978; Commenges and Seal 1985; Maunsell and Gibson 1992; Friedman and Priebe 1998, 1999) to estimate the absolute latency from the trial-averaged response. The estimated trial-by-trial differences in latency then readily quantify the jitter in absolute latency.

4.3 Relation to other methods

Recently, Baker and Gerstein (2001) described three alternative parametric methods to estimate trial-by-trial latency differences. All three methods differ from our nonparametric method in that they rely on the prior estimate of some specific parameter that characterizes the stochasticity of the underlying spiking process.

The first of the three, the *firing rate variance method*, minimizes the mean squared difference between the empirical estimate of the variance across single-trial rate estimates and the rate variance as predicted from the trial-averaged rate function, assuming a fixed ratio α of variance and mean rate (as is the case for a gamma process; see Methods). α is assumed to be a fixed parameter that can be estimated from the data over some 50 ms of “background activity”. Two problems may arise when estimating this parameter. First, for a finite observation interval and finite rate, the parameter α (like the Fano factor) exhibits a bias for low average spike counts (see Methods). In fact, α is directly related to the Fano factor, with a relation depending on the kernel used for the rate estimation and on the time course of the rate itself. Generally, it is advantageous to use count-based measures such as the Fano factor, since for stochastic point processes they permit a more direct analytical treatment (Cox 1967; Teich et al. 1997; Gütig et al. 2002; Nawrot et al., in preparation). In the case of gamma processes, we found that reliable estimates for the Fano factor are obtained for an average of seven to ten (or more) spikes per trial, depending on k (Nawrot et al., in preparation). In the case of a moderate background rate of ten spikes per second, this already requires a rather long observation interval of about 1 s of activity. A similarly long observation is required to reliably estimate α . Second, the degree of rate-normalized trial-by-trial variability and, thus, α may vary during the trial (Fig. 8f). We recently found that neurons in the

primary motor cortex of behaving monkeys exhibit considerable context-related changes in Fano factor with time, with values highest during phases of background activity and lowest during the actual response (e.g., Fig. 8f, Nawrot et al. 2000; MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted). Therefore, a modification of the algorithm of Baker and Gerstein (2001), such that, rather than estimating α , one iteratively reduces the Fano factor using a nonparametric threshold criterion to establish convergence, could be advantageous.

The second, *Bayesian method* (Baker and Gerstein 2001), which is similar to the method described by Seal et al. (1983) and Commenges and Seal (1985), employs a Bayesian framework to estimate the absolute response latency for each trial separately. It is based on the explicit assumption that the interspike intervals are gamma-distributed. To this end it uses the gamma parameter k , which is estimated from the interspike interval distribution rather than from the empirical Fano factor. Again, estimating k is critical in two ways. First, the empirical coefficient of variation of interspike intervals yields a direct estimate of $k = 1/CV^2$, but it exhibits a negative bias for low average spike counts within the observation interval (Nawrot et al., in preparation). This bias is even stronger than that for the Fano factor. Thus, again, a relatively long interval of constant background activity prior to the actual response must be available. Second, to increase the number of intervals for a reliable estimation of k , one must pool intervals over multiple trials. This implicitly assumes a gamma process with an identical, stationary rate for each trial. This is not necessarily the case in reality, however. In motor cortical cells we found that, presumably due to nonstationary network activity, the trial-by-trial variability in terms of spike counts is much higher than expected from the interspike interval distribution (Nawrot et al. 2000, 2001; MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted). Thus, the firing rate can be quite different from trial to trial, which – even in the case of a fixed gamma parameter – will lead to an erroneous estimate of k . For the Bayesian method, Baker and Gerstein (Fig. 3D, 2001) showed that, in the tested case of gamma simulations with $k = 8$ and for a steplike response profile, an erroneous choice for the order parameter k led to a relatively weak increase of the error in the estimated absolute latencies. However, it is to be expected that more stringent conditions, i.e., a more irregular process (smaller value of k) and/or a nonstationary background firing rate, will pose serious problems for the Bayesian method of latency estimation.

The third, the *rate change method* (Baker and Gerstein 2001), first estimates the trial-averaged rate and standard deviation over an initial baseline region. It then measures the individual trial latencies by detecting the crossing of a certain threshold above the base rate, defined as the n -fold initial standard deviation. Two problems can arise with this method. First, the trial-averaged standard deviation in rate, like the Fano factor, exhibits a bias for small average count numbers, necessitating relatively long intervals of baseline activity.

Second, if the degree of trial-by-trial variability in the baseline region varies considerably across different neurons, e.g., due to sources of variability other than the neuron's spiking statistics (MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted), then the factor n for determining the threshold criteria must be adjusted for each neuron separately, requiring additional user intervention. Otherwise, the method may lose sensitivity in detecting single-trial responses in some neurons. In contrast, the performance of the nonparametric method presented here is essentially independent of the degree of trial-by-trial variability in background rate (Fig. 7).

A direct quantitative comparison of our calibration results with those of Baker and Gerstein (2001) is unfortunately not possible since they used steplike responses while we focused on more realistic phasic response profiles (Fig. 1). Moreover, they only used simulations with gamma processes of order $k = 8$ (equivalent to an expected Fano factor of $1/8$). This number is consistent with estimates of k from some motor cortical neurons (Baker and Lemon 2000; Pauluis and Baker 2000), where the authors employed a bootstrap method to reproduce the empirical interspike interval distribution using gamma simulations on the basis of the kernel-estimated single-trial rates. In contrast, however, the empirical estimates of trial-by-trial variability in the cortex in terms of spike counts suggest a much lower gamma order k . The literature on variability in the visual cortex reports typical Fano factor values of around $\hat{F} = 1$, implying Poisson statistics with $k = 1$ (Vogels et al. 1989; Snowden et al. 1992; Gur et al. 1997; Shadlen and Newsome 1998). In the motor cortex, however, we found a broad spectrum of Fano factors, with averages around or slightly above 1, but individual values covering a large range between 0.1 and 10 (Nawrot et al. 2000; MP Nawrot, A Riehle, A Aertsen, S Rotter 2002, submitted), depending on the behavioral context. Therefore, we used predominantly Poisson simulations ($k = 1$) for calibrating our method.

Generally, we have found it useful to apply different methods in parallel when analyzing experimental data. This not only allows for cross validation of results, but may also help to interpret differing results in case any of the methods fail due to their individual weaknesses in the face of some particular property of the data under investigation.

4.4 Simultaneously recorded neurons

In general, we cannot assume that two or more simultaneously recorded neurons operate in the same temporal frame, as is the case for the pair of neurons presented in Figs. 8 and 9. It has been shown that in the motor cortex (Riehle and Requin 1989; Riehle 1991; Miller et al. 1992; Riehle et al. 1994, 1997; Zhang et al. 1997), as well as in other areas, e.g., the prefrontal cortex (Seidemann et al. 1996) and area 5 of parietal cortex (Seal et al. 1983; Commenges and Seal 1985; Requin et al. 1988), responses of single neurons may be locked to either stimulus events or to the animal's movements. From this it is clear that in

many cases the rate dynamics of different neurons, even within the same cortical area, do not share the same temporal frame of reference. Instead, separate cells may be involved in different stages of the task-specific computational processing. Methods that quantify the trial-by-trial differences in response latency between multiple single neurons offer the opportunity to investigate the relevant temporal relations in more detail and to study their possible functional role in higher brain function.

Acknowledgements. We are grateful to Alexa Riehle for providing us with the monkey data and for valuable discussions. We also thank Felix Kümmell, Hiroyuki Nakahara, and Shun-ichi Amari for helpful discussions. Partial funding was received by the Deutsche Forschungsgemeinschaft (DFG, SFB 505) and the German-Israeli Foundation (GIF). Additional support was provided by the RIKEN Brain Science Institute.

References

- Aertsen A, Bonhoeffer T, Krüger J (1987) Coherent activity in neuronal populations: analysis and interpretation. In: Caianiello ER (ed) *Physics of cognitive processes*. World Scientific Publishing, Singapore, pp 1–34
- Aertsen A, Gerstein G, Habib M, Palm G (1989) Dynamics of neuronal firing correlation: modulation of 'effective connectivity'. *J Neurophysiol* 61: 900–917
- Arieli A, Sterkin A, Grinvald A, Aertsen A (1996) Dynamics of ongoing activity: explanation of the large variability in evoked cortical responses. *Science* 273: 1868–1871
- Bach M, Krüger J (1986) Correlated neuronal variability in monkey visual cortex revealed by a multi-microelectrode. *Exp Brain Res* 61: 451–456
- Baker SN, Gerstein GL (2000) Improvements to the sensitivity of gravitational clustering for multiple neuron recordings. *Neural Comput* 12: 2597–2620
- Baker SN, Gerstein GL (2001) Determination of response latency and its application to normalization of cross-correlation measures. *Neural Comput* 13: 1351–1377
- Baker SN, Lemon R (2000) Precise spatiotemporal repeating patterns in monkey primary and supplementary motor areas occur at chance level. *J Neurophysiol* 84: 1770–1780
- Barbieri R, Quirk MC, Frank LM, Wilson MA, Brown EN (2001) Construction and analysis of non-poisson stimulus-response models of neural spiking activity. *J Neurosci Meth* 105: 25–37
- Brody CD (1999) Disambiguating different covariation types. *Neural Comput* 11: 1527–1535
- Bruce CJ, Goldberg ME (1985) Primate frontal eye fields. I. Single neurons discharging before saccades. *J Neurophysiol* 53: 603–635
- Chacron MJ, Longtin A, Maler L (2001) Negative interspike interval correlations increase the neuronal capacity for encoding time-dependent stimuli. *J Neurosci* 21: 5328–5343
- Churchward PR, Butler EG, Finkelstein DI, Aumann TD, Sudbury A, Horne MK (1997) A comparison of methods used to detect changes in neuronal discharge patterns. *J Neurosci Meth* 76: 203–210
- Commenges D, Seal J (1985) The analysis of neuronal discharge sequences: change-point estimation and comparison of variances. *Stat Med* 4: 91–104
- Cox DR (1967) *Renewal theory*. Science Paperbacks, Chapman & Hall, London
- Cox DR, Isham V (1980) *Point processes*. Monographs on applied probability and statistics, Chapman & Hall
- Davey NJ, Ellaway PH, Stein RB (1986) Statistical limits for detecting change in the cumulative sum derivative of the peristimulus time histogram. *J Neurosci Meth* 17: 153–166
- Ellaway P (1978) Cumulative sum technique and its application to the analysis of peristimulus time histograms. *EEG Clin Neurophysiol* 45: 302–204
- Friedman HS, Priebe CE (1998) Estimating stimulus response latency. *J Neurosci Meth* 83: 185–194
- Friedman HS, Priebe CE (1999) Smoothing bandwidth selection for response latency estimation. *J Neurosci Meth* 87: 13–16
- Fries P, Neunenschwander S, Engel AK, Goebel R, Singer W (2001) Rapid feature selective neuronal synchronization through correlated latency shifting. *Nat Neurosci* 4: 194–200
- Gerstein GL, Perkel DH, Dayhoff JE (1985) Cooperative firing activity in simultaneously recorded populations of neurons: detection and measurement. *J Neurosci* 5: 881–889
- Grammont F, Riehle A (2003) Spike synchronization and firing rate in a population of motor cortical neurons in relation to movement direction and reaction time. *Biol Cybern* DOI 10.1007/s00422-002-0385-3
- Grün S, Riehle A, Diesmann M (2003) Effects of across trial non-stationarity on joint-spike events. *Biol Cybern* DOI 10.1007/s00422-002-0386-2
- Gur M, Beylin A, Snodderly M (1997) Response variability of neurons in primary visual cortex (V1) of alert monkeys. *J Neurosci* 17: 2914–2920
- Gütig R, Aertsen A, Rotter S (2002) Statistical significance of coincident spikes: count-based versus rate-based statistics. *Neural Comput* 14: 121–253
- Knott T (2001) Population synchronization during propagation of epileptiform activity in organotypic hippocampal slices: a microelectrode array study. *Der Andere Verlag, Osnabrück, Germany*
- Lee D, Port NL, Kruse W, Georgopoulos AP (1998) Variability and correlated noise in the discharge of neurons in motor and parietal areas of the primate cortex. *J Neurosci* 18: 1161–1170
- Maunsell JHR, Gibson JR (1992) Visual response latencies in striate cortex of the macaque monkey. *J Neurophysiol* 68: 1332–1344
- Miller J, Riehle A, Requin J (1992) Effects of preliminary perceptual output on neuronal activity of the primary motor cortex. *J Exp Psychol* 18: 1121–1138
- Nawrot MP, Aertsen A, Rotter S (1999a) Single-trial estimation of neuronal firing rates: from single-neuron spike trains to population activity. *J Neurosci Meth* 94: 81–92
- Nawrot MP, Riehle A, Aertsen A, Rotter S (2000) Spike count variability in motor cortical neurons. *Eur J Neurosci* 12(Suppl 11): 506
- Nawrot MP, Rodriguez V, Heck D, Riehle A, Aertsen A, Rotter S (2001) Trial-by-trial variability of spike trains *in vivo* and *in vitro*. *Abstr Soc Neurosci* 27: 64.9
- Nawrot MP, Rotter S, Riehle A, Aertsen A (1999b) Variability of neuronal activity in relation to behaviour. In: Elsner N, Eysel U (eds) *Proceedings of the 1st Göttingen neurobiology conference of the German Neuroscience Society 1999*, 1: 101
- Oram MW, Hatsopoulos NG, Richmond BJ, Donoghue JP (2001) Excess synchrony in motor cortical neurons provides redundant direction information with that from coarse temporal measures. *J Neurophysiol* 86: 1700–1716
- Parzen E (1962) On estimation of a probability density function and mode. *Ann Math Stat* 33: 1065–1076
- Paulin MG, Hoffman LF (2001) Optimal firing rate estimation. *Neural Netw* 14: 877–881
- Pauluis Q, Baker SN (2000) An accurate measure of the instantaneous discharge probability, with application to unitary joint-event analysis. *Neural Comput* 12: 647–669
- Perkel DH, Gerstein GL, Moore GP (1967) Neuronal spike trains and stochastic point processes. II. Simultaneous spike trains. *Biophys J* 7: 419–440
- Radons G, Becker DJ, Dülfer B, Krüger J (1994) Analysis, classification and coding of multielectrode spike trains with hidden markov models. *Biol Cybern* 71: 359–373

- Ratnam R, Nelson ME (2000) Nonrenewal statistics of electro-sensory afferent spike trains: implications for the detection of weak sensory signals. *J Neurosci* 20: 6672–6683
- Reich DS, Victor JD, Knight BW (1998) The power ratio and the interval map: spiking models and extracellular recordings. *J Neurosci* 18: 10090–10104
- Requin J, Riehle A, Seal J (1988) Neuronal activity and information processing in motor control: from stages to continuous flow. *Biol Psychol* 26: 179–198
- Richmond BJ, Optican LM, Spitzer H (1990) Temporal encoding of two-dimensional patterns by single units in primate primary visual cortex. I. Stimulus-response relations. *J Neurophysiol* 64: 351–369
- Riehle A (1991) Visually induced signal-locked neuronal activity changes in precentral motor areas of the monkey: hierarchical progression of signal processing. *Brain Res* 540: 131–137
- Riehle A, Kornblum S, Requin J (1997) Neuronal correlates of sensorimotor association in stimulus-response compatibility. *J Exp Psychol* 23: 1708–1726
- Riehle A, MacKay WA, Requin J (1994) Are extent and force independent movement parameters? Preparation- and movement-related neuronal activity in the monkey cortex. *Exp Brain Res* 99: 56–74
- Riehle A, Requin J (1989) Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement. *J Neurophysiol* 61: 534–549
- Sanderson AC (1980) Adaptive filtering of neuronal spike train data. *IEEE Trans Biomed Eng BME-27*: 271–274
- Seal J, Commenges D, Salamon R, Bioulac B (1983) A statistical method for the estimation of neuronal response latency and its functional interpretation. *Brain Res* 278: 382–386
- Seidemann E, Meilijson I, Abeles M, Bergman H, Vaadia E (1996) Simultaneously recorded single units in the frontal cortex go through sequences of discrete and stable states in monkeys performing a delayed localization task. *J Neurosci* 16: 752–768
- Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18: 3870–3896
- Silverman BW (1986) *Density Estimation for Statistics and Data Analysis*. Chapman & Hall
- Snowden R, Treue S, Anderson R (1992) The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Exp Brain Res* 88: 389–400
- Teich MC, Heneghan C, Lowen SB, Ozaki T, Kaplan E (1997) Fractal character of the neural spike train in the visual system of the cat. *J Opt Soc Am A* 14: 529–546
- Tuckwell HC (1988) *Introduction to theoretical neurobiology*, vol 2. Cambridge University Press, Cambridge
- Vaadia E, Kurata K, Wise SP (1988) Neuronal activity preceding directional and nondirectional cues in the premotor cortex of rhesus monkeys. *Somatosens Mot Res* 6: 207–230
- Vogels R, Spileers W, Orban G (1989) The response variability of striate cortical neurons in the behaving monkey. *Exp Brain Res* 77: 432–436
- Zhang J, Riehle A, Requin J, Kornblum S (1997) Dynamics of single neuron activity in monkey primary motor cortex related to sensorimotor transformation. *J Neurosci* 17: 2227–2246