Biol Cybern (2007) 96:99–112
DOI 10.1007/s00422-006-0106-4

ORIGINAL PAPER

Biological Cybernetics

Decoding spike train ensembles: tracking a moving stimulus

Enrico Rossoni · Jianfeng Feng

Received: 8 November 2005 / Accepted: 26 July 2006 / Published online: 16 September 2006

Abstract We consider the issue of how to read out the information from nonstationary spike train ensembles. Based on the theory of censored data in statistics, we propose a 'censored' maximum-likelihood estimator (CMLE) for decoding the input in an unbiased way when the spike activity is observed over time windows of finite length. Compared with a rate-based, moment estimator, the CMLE is proved consistently more efficient, particularly with nonstationary inputs. Using our approach, we show that a dynamical input to a group of neurons can be inferred accurately and with high temporal resolution (50 ms) using as few as about one spike per neuron within each decoding window. By applying our theoretical results to a population coding setting, we then demonstrate that a spiking neural network can encode spatial information in such a way to allow fast and precise tracking of a moving target.

1 Introduction

How to read out information from neural spike activity is a longstanding issue in Neuroscience (Feng 2004; Gerstner and Kistler 2002; Sanger 2003). Answering this question has more than a purely theoretical interest in view of potential applications of neural decoding techniques, for instance in neuroprosthetics (see Brown et al. 2004 and references therein). Assuming that sensory inputs are encoded into noisy, nonstationary spike train ensembles (Harris 2004; Raffi and Siegel 2005; Schaette 2005; Shadlen and Newsome 1998), the brain must be able to read out this information quickly enough if a prompt response is required. Indeed, as the time available for decoding gets shorter, any internal representation of a stimulus becomes less accurate in the presence of noise. This inevitably affects any further processing and eventually degrades the final response. It is important, therefore, to provide theoretical bounds for the *accuracy* and the *rapidity* with which the information can be extracted from a neural system.

From a theoretical viewpoint, the chance of achieving a more or less accurate readout depends on both the statistical features of the spike activity given the stimulus (the encoding model), and the inference method used to extract the information (the decoding strategy). It is well known that maximum-likelihood estimation provides an *optimal* decoding strategy given exact knowledge of the encoding model, e.g. a parametric expression for the interspike interval probability density. Unfortunately the latter may be obtained analytically only for simple neuronal models, whereas in the general case one has to resort to a semi- or to a nonparametric approach.

Notably, in Feng and Ding (2004) the authors derived a maximum-likelihood estimator (MLE) for the input to a leaky integrate-and-fire model. However, such estimator is *biased* when interspike intervals are collected during trials of finite length, reflecting the so-called *censoring* of the data. Since in a real-world setting the time available for decoding the stimulus is always limited, a similar problem would generally affect any decoding scheme based on interspike intervals. On the other hand, being able to estimate a rapidly shifting stimulus with the most accuracy, clearly represents an advantage for survival, and it is reasonable to assume that real nervous systems have developed efficient strategies to solve this

E. Rossoni (✉) · J. Feng
Department of Computer Science,
Warwick University,
Coventry CV4 7AL, UK
e-mail: enrico.rossoni@gmail.com

problem in the course of evolution. Indeed, psychophysical experiments have shown that efficient computations can be performed in the cortex during time windows so small that each neuron may fire only once (Rolls and Tovee 1994; Thorpe 1996).

Methods for the analysis of *censored* data have long been considered in statistics, particularly in survival analysis (see Klein and Moeschberger 1997), but applications to neuroscience have only been limited to date. Here, we present a 'censored' maximum-likelihood estimator (CMLE) for decoding spike train ensembles within finite time windows. The CMLE largely removes the bias from the standard MLE, and is proved more efficient than the simpler moment, or rate-based, estimator (ME) which is commonly employed in the artificial neural network theory. This means that using CMLE one can retrieve a signal more accurately using the same amount of resources, i.e. number of neurons and time. We illustrate this approach in Sect. 3.2, showing that a time-dependent input signal can be decoded accurately from an ensemble of 100 spiking neurons, using time windows that comprise only about one spike per neuron on average.

Next, we extend our approach to a population coding setting. In particular, we consider a tracking task where the position of a moving target has to be determined based on the spike activity elicited in a population of spatially tuned neurons. To evaluate how efficiently spatial information can be encoded in such a spiking neural network, we consider an optimal decoding strategy based upon CMLE, and compare it with the one based on ME. A movie showing the behavior of our model is available at http://www.dcs.warwick.ac.uk/~feng/papers/tracking.avi.

2 Methods

We consider a system of leaky integrate-and-fire (LIF) neurons of parameters V_{thre} (threshold potential, relative to rest), and γ (membrane time constant) (Feng 2004). The synaptic input is modeled by the Gaussian process

$$\mathrm{d}I_t = \mu \mathrm{d}t + \sigma \mathrm{d}B_t \tag{1}$$

where

$$\mu = a\lambda(1-r), \quad \sigma = a\sqrt{\lambda(1+r)} \tag{2}$$

with B_t a standard Brownian motion, independent for each neuron. According to the usual approximation (Tuckwell 1988), this is equivalent to considering a set of independent Poissonian excitatory and inhibitory inputs (EPSPs, IPSPs) of individual size a, and total rates λ and $r\lambda$.

The excitatory input rate λ, is assumed to encode a generic time-dependent stimulus. Additionally we require that the total input be *balanced* at all times, i.e. that its mean is constant and equal to $\mu = V_{\text{thre}}/\gamma$. The balancing condition is enforced by adjusting the inhibitory input rate dynamically (push–pull effect) according to

$$r(t) = 1 - \frac{V_{\text{thre}}}{\lambda(t)a\gamma} \tag{3}$$

In the sequel, we will refer to the model just described as the LIF model. Note that, for the input to be balanced, the condition $\lambda(t) \geq \lambda_0 = V_{\text{thre}}/a\gamma$ must hold at all times, i.e. all neurons are subject to a noisy background input even in absence of any specific stimulus.

2.1 Censored maximum-likelihood estimate

According to Feng and Ding (2004), the probability density of the interspike intervals generated by the LIF model is

$$p_\lambda(\xi) = \frac{2\sigma^2 V_{\text{thre}} \exp(-\xi/\gamma)}{\sqrt{\pi[\sigma^2\gamma(1-\exp(-2\xi/\gamma))]^3}} \cdot \exp\left[-\frac{(V_{\text{thre}})^2 \exp(-2\xi/\gamma)}{\sigma^2\gamma(1-\exp(-2\xi/\gamma))}\right] \tag{4}$$

where σ is intended to depend on the λ as in Eq. (2). It is worth noting that the probability distribution of the ISI generated by the LIF model can be given in closed form *only if* the balancing condition, Eq. (3), holds. A plot of the distribution (4) is shown in Fig. 1. Given the observed interspike intervals $\{\xi_i : i = 1, \ldots, N\}$, the log-likelihood of the data is simply

$$\mathcal{L} = \sum_{i=1}^{N} \log p_\lambda(\xi_i) \tag{5}$$

hence the MLE of λ is

$$\lambda_{\text{MLE}} = \text{argmax}_\lambda \sum_{i=1}^{N} \log p_\lambda(\xi_i) \tag{6}$$

Notably, a solution λ_{MLE} of Eq. (6) can be given in closed form as

$$\lambda_{\text{MLE}} = \frac{\sum_{i=1}^{N} f_a(\xi_i)}{N} + \frac{V_{\text{thre}}}{2a\gamma} \tag{7}$$

where

$$f_a(x) = \frac{V_{\text{thre}} \exp(-2x/\gamma)}{a^2\gamma(1-\exp(-2x/\gamma))}. \tag{8}$$

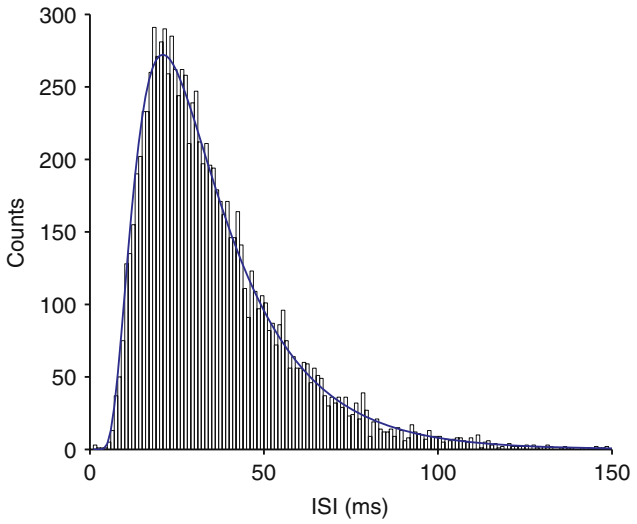

Fig. 1 The histogram of the ISIs generated by the leaky integrate-and-fire (LIF) model ($a = 0.5\,\text{mV}$, $\gamma = 20\,\text{ms}$, $V = 20\,\text{mV}$) is compared with the expected frequencies obtained by the theoretical distribution Eq. 4 (*blue*). The histogram was constructed from 10,000 events; bin size is 1 ms

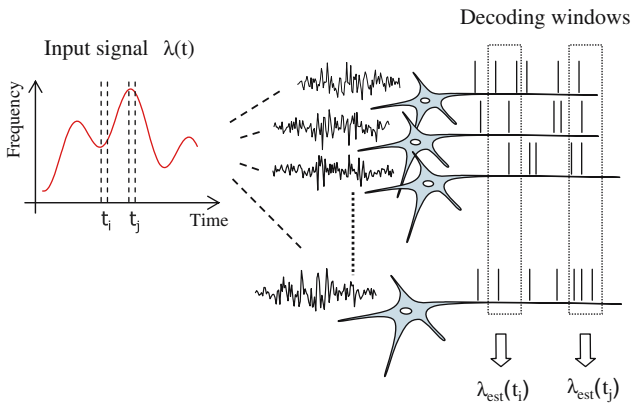

Fig. 2 Decoding a dynamical signal from a population of spiking neurons. A signal is encoded in the time-dependent rate $\lambda(t)$ of the synaptic inputs to a population of N integrate-and-fire neurons. For each decoding time window (*dotted line boxes*), a 'censored' maximum-likelihood estimate of the input, $\lambda_{\text{est}}(t)$, is obtained in terms of the regular and truncated interspike intervals pooled across the whole neuronal population

When spike trains are collected during time windows of length $T < \infty$, the estimator (6) is biased, as clearly shown in Feng and Ding (2004), Fig. 5.

For a given estimator $\hat{\lambda}$, we define the *relative bias* as $\epsilon(\hat{\lambda}) = E((\hat{\lambda} - \lambda)/\lambda)$. A lower bound for $\epsilon(\lambda_{\text{MLE}})$ can then be obtained from (8). As shown in Fig. 3, λ_{MLE} dramatically overestimates inputs of small λ with short decoding windows.

In order to include the effect of censoring and remove the bias, the likelihood model Eq. (5) must be modified. Consider a set of N neurons, and let $\{t_i^k\}$, $i = 1, 2, \ldots$, be

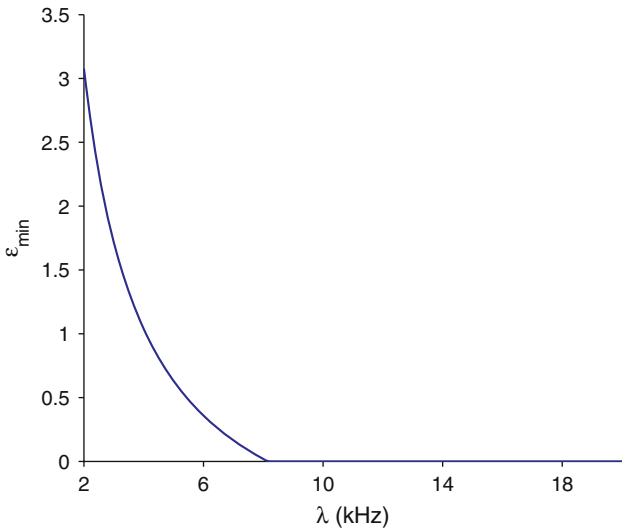

Fig. 3 A lower bound for the relative bias of the standard maximum-likelihood estimator in presence of censored observation ($T_{\text{w}} = 25\,\text{ms}$). Notice the strong bias with small signals

the spike times recorded in the interval $[t, t+T_{\text{w}}]$ for neuron $k = 1, \ldots, N$. For simplicity of notation we shift the time axis so that $t = 0$ in the following. Two sets of intervals are introduced: the 'regular' interspike intervals, for which both the start and the endpoint are observed, and the 'censored' interspike intervals, for which the endpoints exceed the recording window and are replaced by T. We indicate these two sets by $\{\xi_i\}_{i\in R}$, and $\{\xi_i\}_{i\in T}$, respectively.

To illustrate the censoring approach, we consider first the case where (at most) the first two spikes are recorded for each neuron. This parallels the typical setting in survival analysis, where a study is allotted a fixed start and end time, say $[0, T]$, and each patient may enter the study at a random time $t_1 \in [0, T]$. In our case, the finite time of observation and the random occurrence of the first spike within the interval determine what is known in statistics as type I random right-censoring, meaning that only the minimum between the true interspike interval and the random time $T - t_1$ can actually be observed. It is important to note that, in the particular setup considered here, censoring is *noninformative* about the distribution of the true interspike interval, whereas in a survival study censoring of the patient's lifetime, e.g. due to the patient dropping out of the study, may be due to causes that are not independent from the latent process under observation and bring some information about it. Under the hypothesis of noninformative censoring, the log-likelihood of the data conditional to the censoring times can be written as (Andersen et al. 1993)

$$\mathcal{L} = \sum_{i\in R} \log p_\lambda(\xi_i) + \sum_{i\in C} \log S_\lambda(\xi_i) \tag{9}$$

where we have introduced the *survival* function

$$S_\lambda(t) = \int_t^{+\infty} p_\lambda(t')\mathrm{d}t'.$$

Accordingly, we define the *censored maximum-likelihood estimator* (CMLE) of λ as

$$\lambda_{\mathrm{CMLE}} = \mathrm{argmax}_\lambda \left[\sum_{i\in R} \log p_\lambda(t_i) + \sum_{i\in C} \log S_\lambda(t_i) \right] \quad (10)$$

Since the survival is a monotonically decreasing function of λ, the additional 'censored' term in the likelihood shifts the estimates toward zero, partly removing the positive bias. In fact, such censored estimator is known to be asymptotically unbiased in the limit of large sample size.

A problem with the application of the interval-based decoding methods arises when all observations are censored. Indeed, the probability of this event is finite for samples of finite size, and increases with short decoding windows and/or small signal rates. In this case, the censored log-likelihood is maximized for $\lambda = \lambda_0/2$, which corresponds to a degenerate distribution of interspike intervals ($\sigma = 0$). Such "degenerate" estimates can be avoided by introducing a suitable penalising term in the likelihood; see e.g. Pettitt et al. (1998). However, in this study we simply discarded these trials for analysis whenever they occurred.

The censored approach described above is clearly inefficient, since it generally ignores a great deal of information contained in the spike train ensemble. In order to improve efficiency, one could include the distribution of the censoring time in the likelihood. However, a closed form for the distribution of t_1 in terms of interspike interval density is obtainable only in the stationary case (Gnedenko et al. 1969), which makes such approach unsuitable in a time-dependent setting.

Ideally, we would like to use *all* the information available in the spike train ensemble, including all the regular and censored intervals. Studies that involve similar *recurrent event data* are common in reliability and survival analysis studies, e.g. when monitoring the status of a repairable system, or keeping track of hospitalization visit of a patient with a chronic disease.

We suggest that an unbiased estimator can be obtained by maximizing the censored likelihood Eq. (10) including the whole set of interspike intervals. Arguably, such approach would neglect the informative censoring resulting from the particular data accrual scheme (Hollander and Pena 2004). However, we conjecture that in our case this would simply affect the variance of the censored estimator, not on its bias. In the Appendix we report a simple proof of the unbiasedness of the CMLE in the Poisson case. A rigorous approach to this problem demands a more delicate treatment and will constitute the object of a further study.

2.2 Moment estimate

Consider an ensemble of q neurons and let $\{t_i^k\} i = 1,\ldots,q, k = 1,2,\ldots$ be their spike times. Assuming that the interspike intervals $\xi_i^k = t_i^{k+1} - t_i^k$ are i.i.d. random variables, with probability density p_λ, the *spike counts* $N_{i,t} = \sum_{k=1}^{\infty} \chi(t_i^k < t)$ form a set of independent renewal processes, where we defined, for a statement S, $\chi(S) = 1$ if S is true, $\chi(S) = 0$ otherwise. Thus, from the Renewal Theorem we have for $T > 0$

$$E\left[\frac{N_{i,t+T} - N_{i,t}}{T}\right] \overset{t\to\infty}{\longrightarrow} \frac{1}{E[\xi|\lambda]} \quad (11)$$

where

$$E[\xi|\lambda] = \int_0^{\infty} \xi p_\lambda(\xi)\mathrm{d}\xi. \quad (12)$$

That is, when averaging across many trials, the spike rate becomes *asymptotically* stationary and converges to the inverse of the expected ISI. Using the ensemble average of the spike rate to estimate the expectation in the right-hand side of (11), and assuming that t is large enough for the ensemble to be considered 'at the equilibrium', we can then write the following equation for the Moment Estimate of the input (λ_{ME}),

$$\frac{1}{q}\sum_{i=1}^{q} \frac{N_{i,t+T} - N_{i,t}}{T} = \frac{1}{E[\xi|\lambda_{\mathrm{ME}}]} \quad (13)$$

The function

$$\lambda_{\mathrm{out}} = F(\lambda_{\mathrm{in}}) := \frac{1}{E[\xi|\lambda_{\mathrm{in}}]} \quad (14)$$

defines the *input–output relationship* of our system; see Fig. 4. Notice that, whereas in artificial neuronal networks the output firing rate is given as an heuristic (sigmoidal) function of the input, for the LIF model the input–output relationship can be obtained *exactly* from the probability distribution of the interspike intervals.

It is evident that λ_{ME} is biased for $q < \infty$. Indeed, since $\phi = F^{-1}$ is nonlinear we have

$$E[\phi(f_{T,p})] \neq \phi(E[f_{T,p}]) \quad (15)$$

where we denoted with $f_{T,p}$ the ensemble average of the spike rate and we dropped the dependency on t. Using a Taylor expansion to approximate $\phi(f_{T,p})$, it can be seen that the leading term in the bias (of order $1/p$) is positive and decreasing with T.

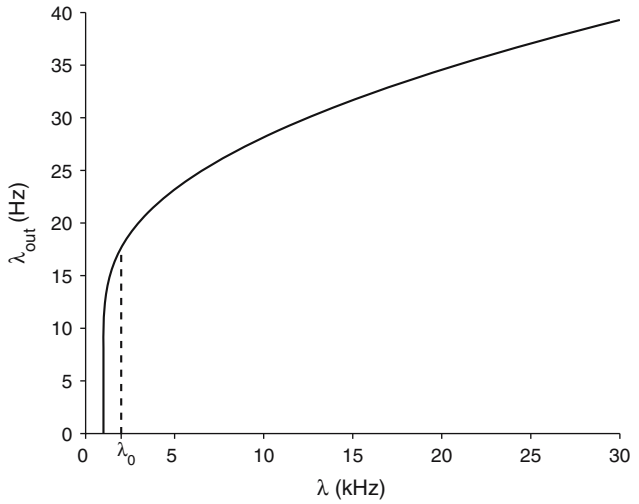

Fig. 4 The input–output relationship for a leaky integrate-and-fire neuron with balanced input, Eq. (14). Values are obtained for $a = 0.5\,\text{mV}$, $\gamma = 20\,\text{ms}$, $V = 20\,\text{mV}$. The *dashed line* at $\lambda = V_{\text{thre}}/a\gamma$ marks the minimum value of the excitatory input compatible with the balancing condition

In theory, a suitable hierarchy of neural observers may learn the nonlinear relationship between λ and $E[\lambda_{\text{ME}}]$, in order to obtain as output an unbiased estimate of the input. However, as we will see in next section, such approach would become substantially more involved in the time-dependent case where a static mapping between λ and $E[\lambda_{\text{ME}}]$ may no longer be definable.

3 Results

To first illustrate the behavior of the CMLE estimator in a renewal case, we report the results of a simple Monte Carlo study. Here, we generated ensembles of N spike trains, with inter-event times drawn from a Gamma distribution of mean $\mu = 42\,\text{ms}$, and standard deviation $\sigma = 22\,\text{ms}$ (these values were obtained by fitting a sample of interspike intervals generated by the LIF model with $\lambda = 6\,\text{kHz}$). The series of event times were parsed into time windows of $T = 25, 50, 100\,\text{ms}$ for analysis. Assuming σ known, we estimated the parameter μ using three different models: a standard likelihood model based exclusively on the regular intervals (Model A); a censored likelihood model including only the first interval (Model B); and a censored likelihood model including all observed intervals (Model C). In Table 1 we reported the mean and SD of the estimates computed over a set of 1,000 trials. It is apparent the strong bias obtained with the standard model (A) is substantially reduced by the censored models (B and C). Both censored models tend to become unbiased for large samples, although with the shorter time window, the full censored model (C) still retained a statistically significant bias at $N = 1{,}000$.

Table 1 Compared maximum-likelihood estimates of the expected inter-event time μ for a renewal process of inter-event distribution Gamma(μ,σ), with $\mu = 42\,\text{ms}$, $\sigma = 22\,\text{ms}$

T_{w}	N	Model A	Model B	Model C
100	10	36.30 ± 4.61	42.62 ± 6.92	42.93 ± 5.16
100	100	35.73 ± 1.36	42.10 ± 2.06	42.14 ± 1.52
100	1,000	35.62 ± 0.43	42.01 ± 0.66	42.04 ± 0.47
50	10	29.20 ± 7.36	43.63 ± 9.11	45.63 ± 8.68
50	100	27.53 ± 1.95	42.13 ± 2.75	43.02 ± 2.59
50	1,000	27.43 ± 0.61	42.00 ± 0.86	42.29 ± 0.80
25	10	34.10 ± 12.48	31.57 ± 5.45	35.89 ± 4.82
25	100	19.17 ± 4.25	43.17 ± 5.51	45.15 ± 5.40
25	1,000	18.55 ± 0.89	42.11 ± 1.59	42.59 ± 1.57

Reported values are mean $\pm$ SD of the estimates calculate over 1,000 trials. Individual estimates were computed after parsing an ensemble of N independent realizations of the process, into 1,000 consecutive windows of length T_{w}. For Model A, the estimates are computed using only the regular intervals observed within each window; Models B and C use a censoring correction to account for the finite window size, but B uses only the first interval in each window. The parameter σ is assumed known for estimation

3.1 Time-independent signal

We now consider the case of the LIF model with a constant input. Spike data were generated by simulating the activity of $N = 100$ leaky integrate-and-fire neurons in response to the input (1) with λ constant. The parameters used for simulation are $V_{\text{thre}} = 20\,\text{mV}$, $\gamma = 20\,\text{ms}$, and $a = 0.5\,\text{mV}$; the membrane potentials at time zero were drawn from a uniform distribution in $[0, V_{\text{thre}}]$. The time step for numerical integration was $1\,\mu\text{s}$. The amplitude of the input λ was varied from 2 to 20 kHz in successive trials. For each trial, the spike train ensemble $\{t_i^j\}$, $i = 1, \ldots, N$, $j = 1, \ldots, n_i$, was parsed into 1,000 consecutive windows of length T_{w}. For each window (for ME), or whenever at least one interspike interval was observed (for CMLE), we calculated the estimates of the input and their relative bias.

The box plots of Fig. 5 summarize the observed distribution of the relative bias for CMLE (left) and ME (right) with varying stimulus amplitude and window lengths. The CMLE appears to largely overestimate inputs of low amplitude with short decoding windows. Notice that for $\lambda = 2\,\text{kHz}$ the CMLE could be calculated only on 2% of the trials, given the low mean spike count per window ($\bar{n} = 0.440 \pm 0.0492$). For $\lambda = 4\,\text{kHz}$ the percentage of accepted trials increased to 67%, and was above 99% for $\lambda \geq 6\,\text{kHz}$. For $T = 50, 100\,\text{ms}$, all the trials were accepted. With increasing input level the bias

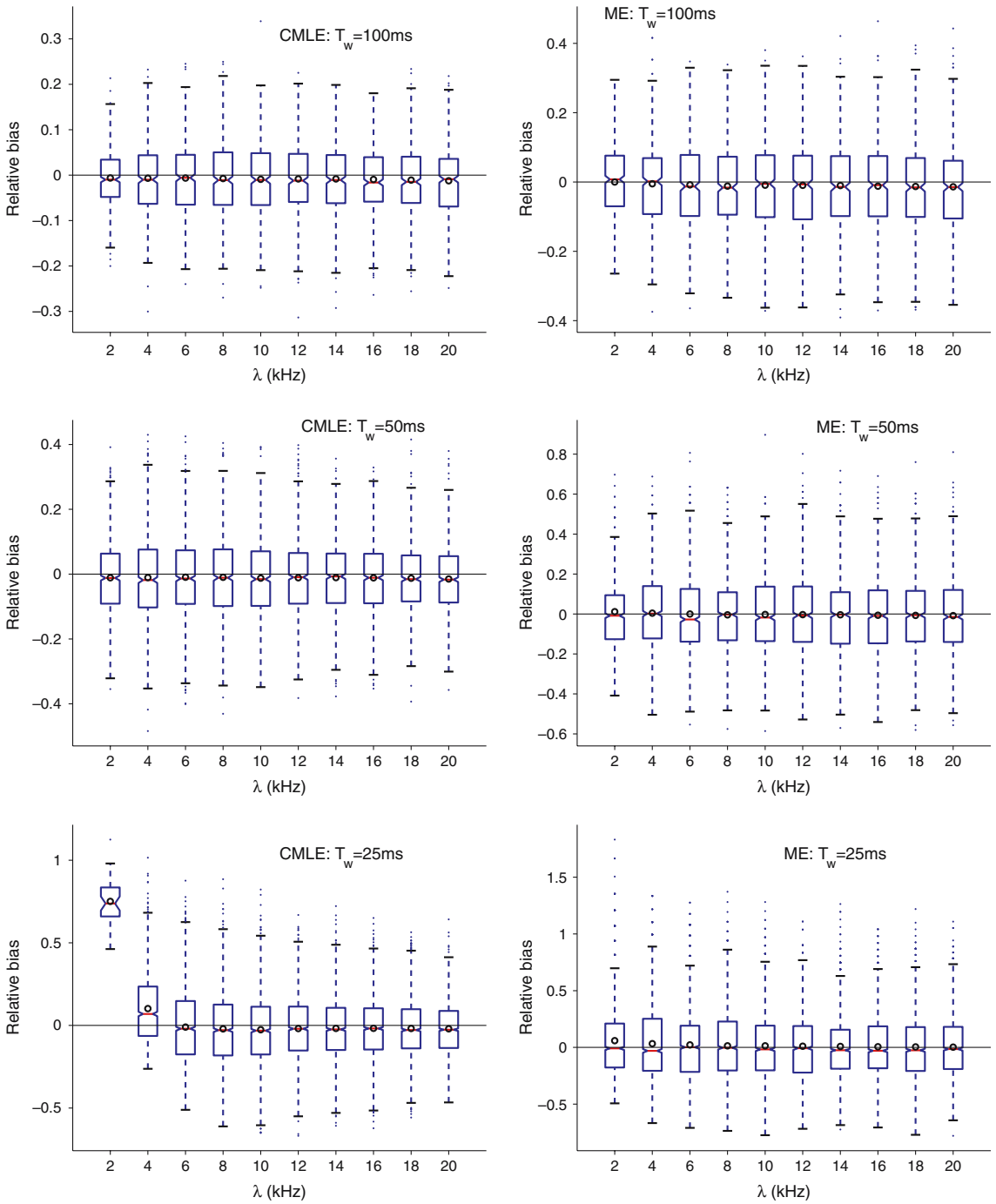

Fig. 5 Decoding a constant stimulus: the relative bias of censored maximum-likelihood estimator (*CMLE*; *left*) and ME (*right*). The box plots have *lines* at the lower quartile, median, and upper quartile values, while *circles* indicate the mean. The *whiskers* extend to the most extreme data value within 1.5 of the interquartile range of the box; outliers shown as *dots*. Results are obtained for decoding windows of lengths 100 ms (*top*), 50 ms (*middle*) and 25 ms (*bottom*)

of CMLE decreases rapidly, and it becomes negative at high input levels for all the window lenghts considered. Comparing with the example at the end of Sect. 2.1, we note that such small negative bias seems to be characteristic of the censored estimator at finite sample sizes.

The average bias for ME is always positive with $T = 25$ ms, in agreement with what anticipated in Sect. 2.2, and decreases monotonously from about 6 to 0.25% in the range considered. For $T = 50, 100$ ms, the bias tends to become negative (about 1%) at high input levels.

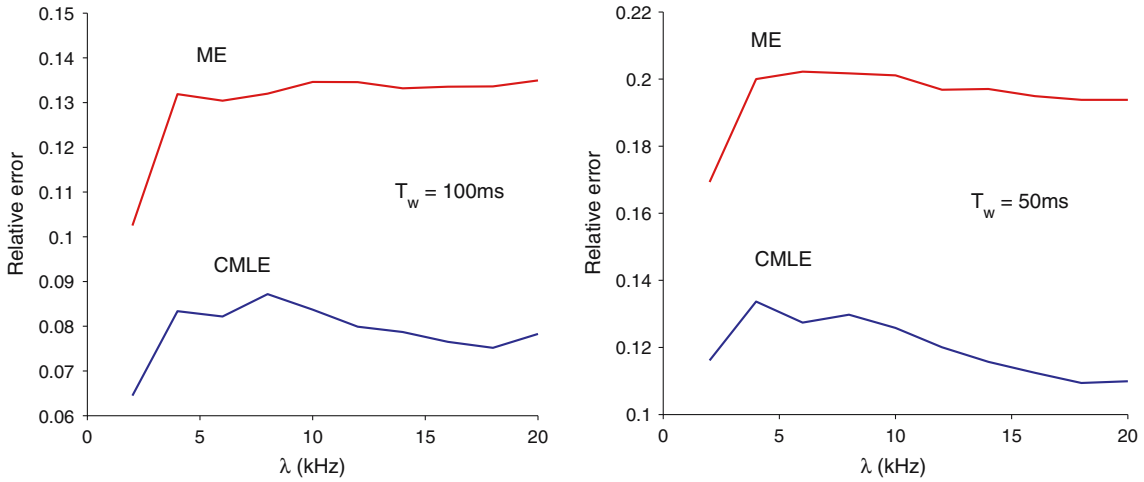

Fig. 6 Decoding a constant stimulus. The relative error of CMLE and ME, defined as $\mathrm{SD}(\hat{\lambda})/\lambda$. Results obtained for $T_w = 100$ ms (*left*) and $T_w = 50$ ms (*right*)

In Fig. 6 we reported the relative error of the two estimators measured by the ratio $\mathrm{SD}(\hat{\lambda})/\lambda$ for $T = 50$, 100 ms. As expected, the CMLE is more efficient than ME in the whole stimulus range. Also notice the relatively weak dependency of the error on the input. This result is due to the balancing of the input which partially rectifies the input–output relationship of the system.

We stress the fact that, in the definition of the Moment Estimator, Eq. (13), the ensemble average of the spike rate is assumed to be stationary, i.e. the renewal process is considered at the equilibrium. Only under this condition, the ME is asymptotically unbiased. In the nonstationary case, estimates obtained by ME may show even a large bias depending on time and on the distribution of membrane potentials in the ensemble. Assume, for instance, that all neurons are at their resting potential at time $t = 0$. Measuring the ensemble spike rate in $[0, T]$, with T much smaller than the 'typical' ISI, we would obtain a much smaller value than what expected from the Renewal Theorem. This reflects the fact that, for the LIF model, the *hazard* of emitting a spike goes to zero at small intervals. Therefore, the ME would dramatically underestimate the input under such conditions. On the other hand, if the initial distribution for the membrane potentials is skewed toward the threshold, there would be an excess of spikes in the early phase of the dynamics compared with the stationary case, thus the ME would overestimate the input. Since the effect of the initial conditions on the dynamics is eventually dampened by the noisy input, the distribution of the estimates, hence the expected value and the bias of the estimator, converge in the long run. In the language of statical physics, we could say that the ME cannot provide any consistent estimate of the stimulus unless the system is 'thermalized'.

To further illustrate this issue, we generated new spike train ensembles by holding the stimulus constant, but randomizing the membrane potentials every 25 ms to uniformly distributed random values in $[0, V_{\mathrm{thre}}]$. In this way we prevent the system to relax to the equilibrium state. The box plots of Fig. 7 summarize the distribution of the relative bias for CMLE and ME obtained under this conditions. While for CMLE we observed no significant change compared with the equilibrium case, for ME we obtained a strongly input-dependent bias, with small signals being underestimated and large signals being overestimated. This behavior is predictable on the basis of the equilibrium distribution of membrane potentials, which is shown in Fig. 7 (top). Indeed, for $\lambda \sim \lambda_0$ the uniform distribution has an excess of neurons near the resting potential and a lack of neurons close to threshold compared with the equilibrium distribution, hence it would lead to a reduction in the number of spikes. With signals of increasing amplitude, instead, the situation is reversed, as neurons are found farther from the threshold in a thermalized rather than in a uniformly distributed ensemble.

3.2 Time-dependent signal

We now turn to the time-dependent case. For simplicity we consider a piecewise constant random signal of the form,

$$\lambda(t) = \sum_{j=1}^{N} \lambda_j \chi(t \in I_j) \tag{16}$$

where λ_j, $j = 1, 2, \ldots$ are independent, identically distributed random variables, with uniform distribution in

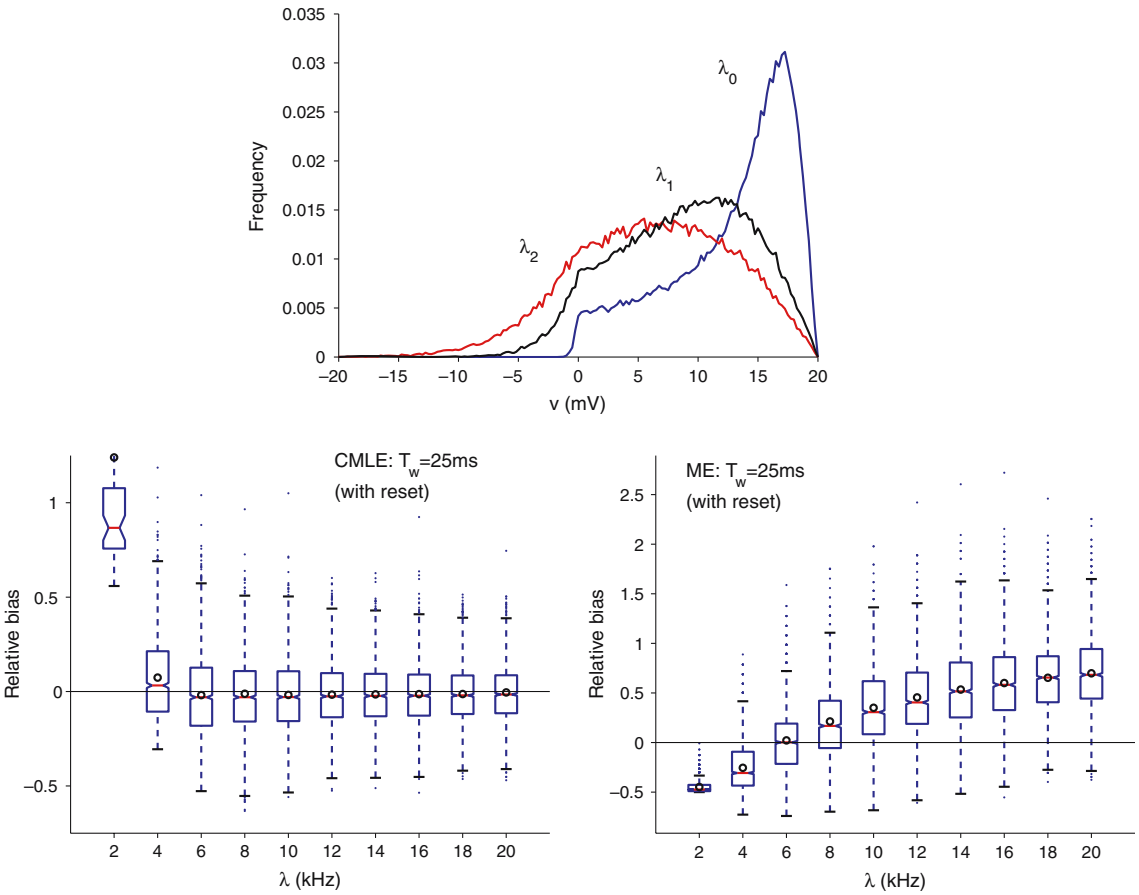

Fig. 7 *Top* The asymptotic distributions of membrane potentials for the LIF model at various input rates $\lambda_0 = 2\,\text{kHz}$, $\lambda_1 = 10\,\text{kHz}$ and $\lambda_2 = 20\,\text{kHz}$. The histogram is constructed on a sample of 10^5 neurons; bin size is 0.25 mV. The convergence to equilibrium is evaluated by comparing the transformed distributions after 25 ms of dynamics by means of a nonparametric rank sum test. *Bottom* The relative bias of CMLE (*left*) and ME (*right*) obtained with a 25 ms-decoding window under conditions of nonequilibrium. Here the membrane potentials of all neurons in the ensemble were reset to uniformly distributed random values in the interval $[0, V_{\text{thre}}]$ at the beginning of each window

$[\lambda_0, \lambda_1]$ with $\lambda_0 = 2\,\text{kHz}$, $\lambda_1 = 5\lambda_0$, and $I_j = \{t : (j-1)\,T_\text{w} \le t < jT_\text{w}\}$. Spike time data were generated for T_w= 25, 50, and 100 ms, during trials of length 2.5, 5, and 10 s, respectively, in order to compare the same number of intervals. In the decoding phase, the input is estimated at times $t_j = jT_\text{w}$ by applying CMLE on the interspike intervals observed during the preceding window I_j. Finally, we computed the average estimating error over the trial, $E = \frac{1}{N}\sum_{j=1}^{N} \frac{|\hat{\lambda}_j - \lambda_j|}{\lambda_j}$.

Figure shows snippets of the actual and reconstructed signals for $T_\text{w} = 50, 100\,\text{ms}$. In Fig. 8 (bottom left) we reported the obtained reconstruction errors (mean ± SEM) at different window lengths. Our results indicate that, already with a 50 ms window, the random signal can be reconstructed with about 90% accuracy, in spite of the few spikes observed on average within each decoding window ($\bar{n} = 1.21$; see Fig. 8, bottom right). Interestingly, although with a 50 ms window we observed less than one regular interval per neuron on average, a decoding scheme that employed all the interspike intervals led to 20% reduction in the error compared with a scheme that uses only the first interspike interval. The CMLE was found to about 50% more accurate than ME in such dynamical setting. Also, while the error for CMLE was comparable with the average (absolute) bias obtained in the stationary case (0.100 ± 0.077 vs. 0.101 ± 0.034, mean ± SD, $T_\text{w} = 50\,\text{ms}$), for ME the former was significantly larger than latter (0.20 ± 0.14 vs. 0.138 ± 0.053). This result reflects the increased variability of the moment estimator when it is applied to neuronal ensembles that are not in an equilibrium state. Finally we note that, while the method used in Feng and Ding (2004) requires an offline step to correct the bias, the present method takes into account the effect of the finite-time sampling already in the expression of the likelihood, and therefore allow decoding on line.

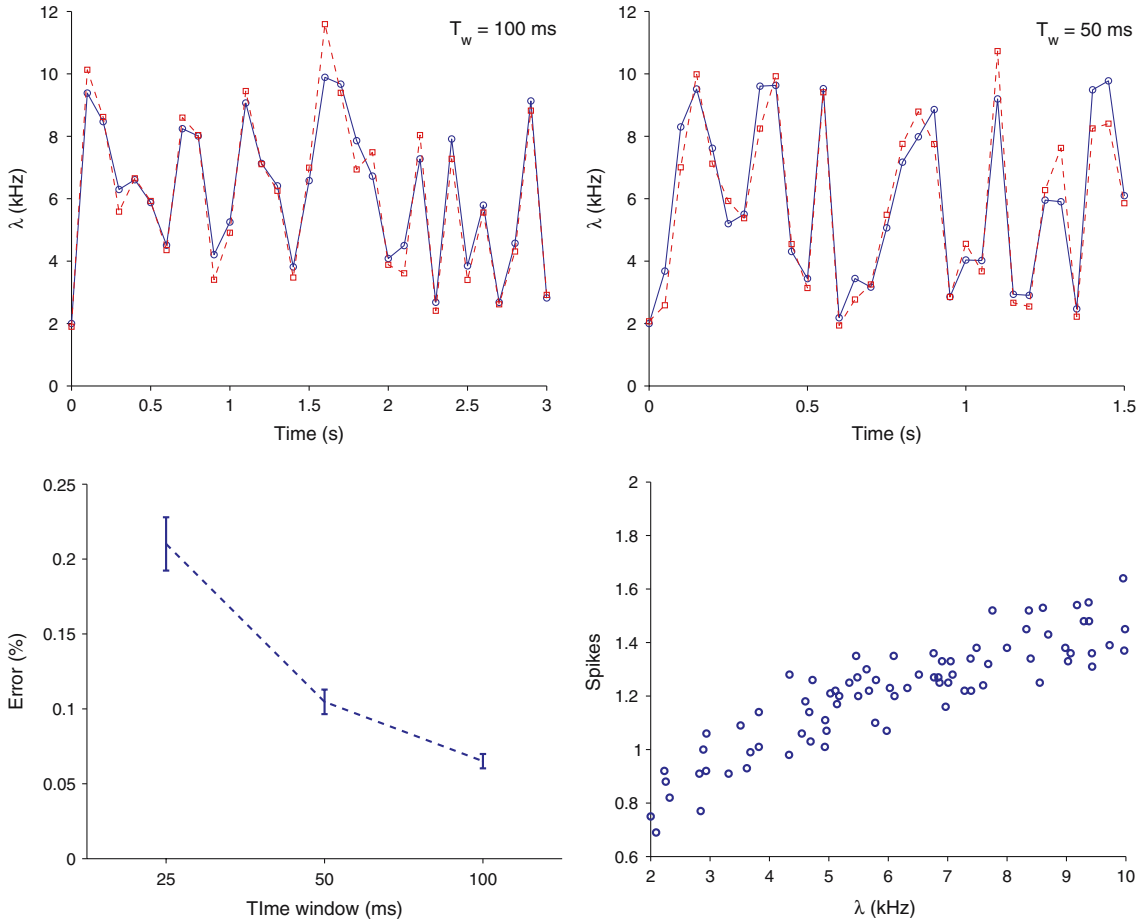

Fig. 8 Decoding a time-dependent stimulus. *Top* The actual (*circles*) and estimated (*squares*) stimulus values, obtained for $T_w = 100$ ms (*left*), and $T_w = 50$ ms (*right*); *lines* have been added for visual clarity. *Bottom left* The reconstruction error (average over 100 points) obtained for different window length; *bar* are SEM. *Bottom right* The mean spike counts obtained for a 50 ms-time window

3.3 Tracking a moving stimulus

In the previous examples, we assumed that all neurons shared an identical response to the same input. However, the brain is known to make a wide use of 'population codes', where neurons with different responses are involved collectively in the representation of the same stimulus. This is reflected for instance in the columnar organization of the visual cortex. Neurons in V1 are organized in *position columns*, with small receptive fields localized at specific locations in the visual space. The structural arrangement of these columns of neurons is such that moving horizontally through the cortex there is a smooth progression of visual field position versus cortical position, so to form a topographic map of the visual space.

Here, we consider a spiking neuronal network which resembles this structural arrangement. In particular, we are interested in assessing how accurately a tracking task can be performed when the spatial information is encoded in such a network.

Our model system is composed of 1,000 neurons grouped into $p = 10$ columns of 100 neurons each. The parameters of the neuron are as above. The target is represented by a point object whose position is restricted to the interval $[0, L]$ of the x-axis. Given the target's position $x \in [0, L]$, we assume that neurons in the ith column receive independent synaptic inputs of rate

$$\lambda(x_i; x) = \lambda_0 + c \cdot \lambda_0 \exp\left(-\frac{(x - x_i)^2}{2\sigma^2}\right), \tag{17}$$

where $x_i = \left(i + \frac{1}{2}\right) \cdot L/p$ are the columns' *field centers*, σ is the *tuning width*, assumed constant for simplicity, and c is a scaling factor. This function defines the topographical map for our system. The setup is illustrated schematically in Fig. 9. Note that the term λ_0 in Eq. (17)

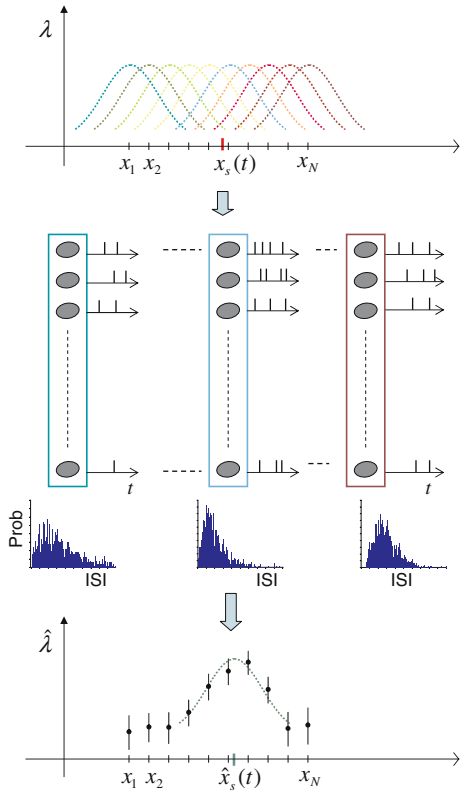

Fig. 9 Tracking a moving target: schematic drawing of the setup. Neurons are grouped into columns (*boxes*) and their response tuned to specific locations along a segment (*tuning curves*, *colored dotted lines*). The position of a moving target $x_s(t)$ is inferred from the ensemble spike activity sampled during a short time interval. An estimate of the input to each column, $\hat{\lambda}$, is obtained from the observed distribution of the interspike intervals (CMLE), or from the mean spike count (ME). Finally, the reconstructed input profile (*bars*) is fitted by a template curve (*green dotted line*, *bottom*) to provide an estimate of the target location, $\hat{x}_s(t)$

ensures that each neuron maintain a background activity even in absence of a target-related stimulus.

We further assume that the target position change in time, according to

$$x(t) = \sum_{j=1}^{N} \xi_j \chi(t \in I_j) \tag{18}$$

where ξ_j are independent random variables uniformly distributed in $[0, L]$, and $I_j = \{t : (j-1)T_w < t \leq jT_w\}$. Thus, in this model the target 'hops' to a new random position every T_w ms and remains still otherwise.

The tracking algorithm consists in two steps:

1. at time t_j, the value of the input for each column $\lambda_i(t_j) = \lambda(x_i; x(t_j))$, i.e. the stimulus' profile, is estimated using CMLE/ME based on the spike activity in the the preceding time window I_j,

$$\hat{\lambda}_i(t_j) = \hat{\lambda}(x_i | \{t\} \cap I_j), \tag{19}$$

2. the target's position is estimated by least-squares fitting the tuning curve Eq. (17) to the estimated input profile. That is, the final estimate of $x(t_j)$ is given by

$$\hat{x}(t_j) = \text{argmin}_x \sum_{i=1}^{p} \left[\hat{\lambda}_i(t_j) - \lambda(x_i; x)\right]^2. \tag{20}$$

The last step corresponds to a Bayesian approach and allows us to provide a smooth estimate of the target's position. We also considered a modified version of (20) where less weight was put on more variable estimates, i.e.

$$\hat{x}(t_j) = \text{argmin}_x \sum_{i=1}^{p} \frac{\left[\hat{\lambda}_i(t_j) - \lambda(x_i; x)\right]^2}{\text{Var}(\hat{\lambda}_i)} \tag{21}$$

However, preliminary results indicated that such modification had no significant effect on the final position estimate so it was not considered any further. Finally, the tracking accuracy was assessed by the root-mean-squared error,

$$E = \left[\frac{1}{N}\sum_{j=1}^{N}(\hat{x}(t_j) - x(t_j))^2\right]^{\frac{1}{2}}. \tag{22}$$

In Fig. 10 we show examples of actual and reconstructed target trajectories for different time windows and tuning widths. In agreement with the results of Sect. 3.1, tracking accuracy was improved sensibly by using CMLE for decoding instead of ME. In particular, using CMLE the target could be tracked accurately already with a 50 ms time window.

Among the factors affecting the performance of such a system, the role of the spatial sensitivity of the single neurons has been widely discussed in the literature in the past few years (Amari et al. 2005; Peng and Van Essen 2005; Paz et al. 2004). When the reconstruction error was calculated as a function of the tuning curve width σ, we obtained that the best performance was achieved when the tuning width was set at about half the field centers' spacing; see Fig. 11.

Finally, we calculated the tracking accuracy as a function of the number of columns used for decoding. More specifically, at each step only the stimulus estimates from the q columns with the highest firing rates were used in Eq. (20), and we allowed q to vary from 1 to p. Figure 11 shows the results obtained for a 100 ms-window. It is found that tracking accuracy always *increased* by enlarging the pool of neurons involved in the readout. This

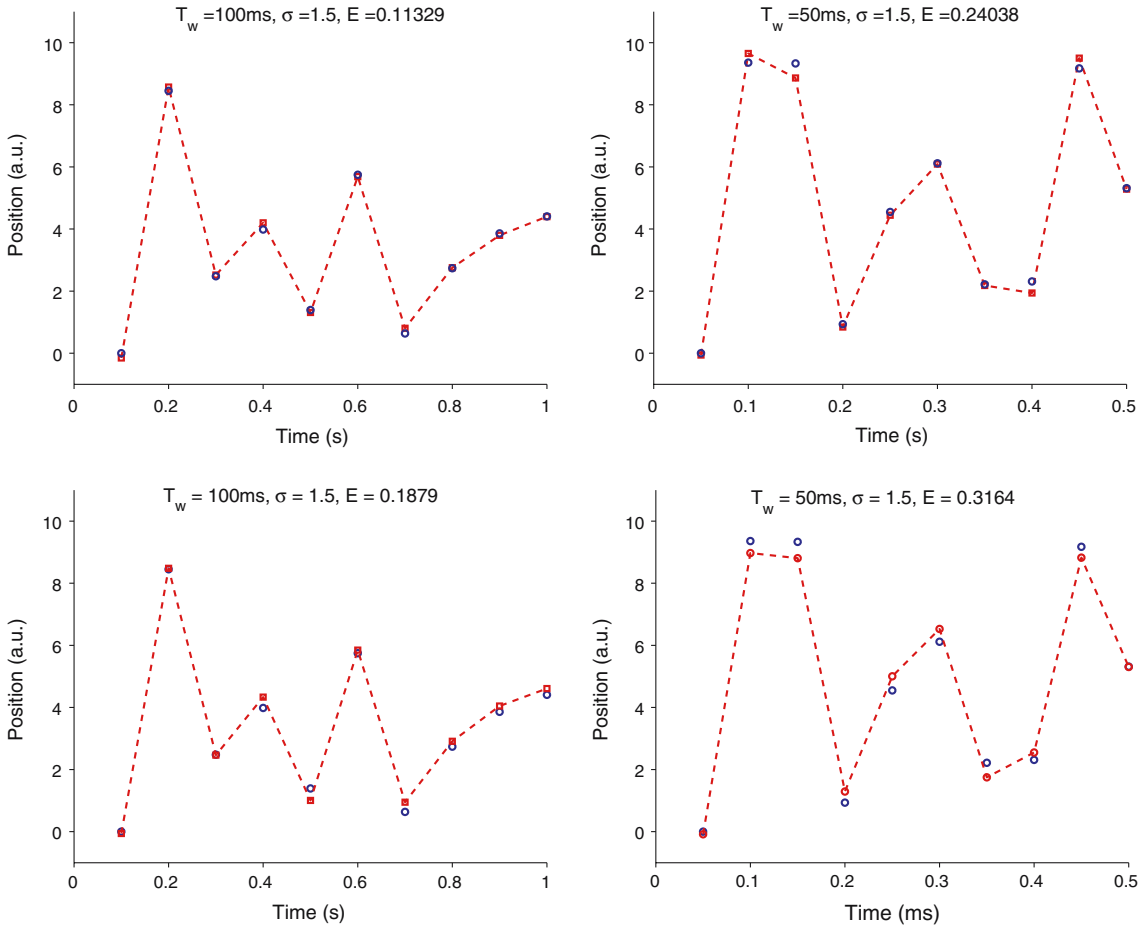

Fig. 10 Tracking a moving stimulus. Snippets of the actual (*blue circles*) and reconstructed trajectories (*red squares*), as estimated using CMLE *top*, or ME *bottom*. Results are obtained with a tuning width of $\sigma = 1.5$, and decoding windows of $T_w = 100$ ms (*left*), $T_w = 50$ ms (*right*). The reconstruction error (E) is averaged over 40 trajectory points. All columns are used in Eq. (20)

result was independent on the estimator used for decoding and the tuning curve width.

Finally, we modify our setup to allow for inhomogeneity in the neuronal population. In particular, we assume that the thresholds, θ^i, are distributed randomly within the interval $\left[V_{\text{thre}}(1-\delta), V_{\text{thre}}(1+\delta)\right]$, with $\delta > 0$. In a 'mean-field' approach, any estimate based upon the output of such a population can be approximated by considering a *homogeneous* population of threshold $\theta = \langle\theta^i\rangle$. Therefore, we expect that the previous findings will not change in presence of small inhomogeneity. Note, however, that the CMLE is no longer optimal for $\delta > 0$, since the actual distribution of the ISI collected from an inhomogeneous population would differ from that given by Eq. (4).

In Fig. 12 we plot the reconstruction error (relative to the homogeneous case) as a function of the inhomogeneity parameter δ. Results are calculated for networks of increasing inhomogeneity, set to the same tracking task ($\sigma = 1.5$, $T_w = 50$ ms). The algorithm appears to be robust with respect to the system's inhomogeneity, confirming the validity of the mean-field approach. However, while the CMLE-based algorithm is progressively less accurate as the inhomogeneity increases, the reconstruction error tends to decrease with the ME-based approach.

4 Discussion

Although the problem of decoding neural activity has received much attention in the literature, this is possibly the first paper where a truly statistical approach, based on maximum-likelihood estimation, is employed to evaluate how much information can be extracted from spike train ensembles. So far, the only application of such approach has been to cases where neurons were described as following a Poisson spike statistics, an approximation which is far from being satisfactory for any real neuron.

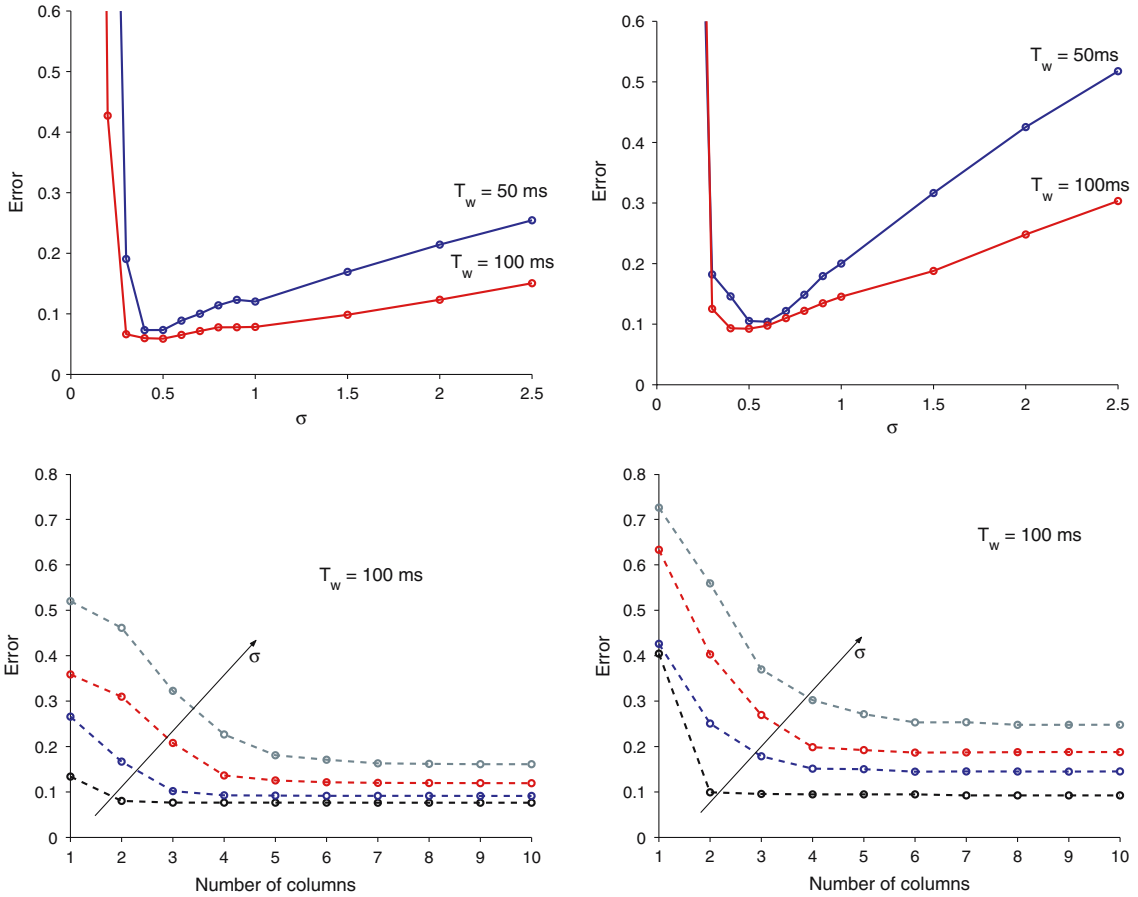

Fig. 11 Tracking a moving stimulus: comparison between the CMLE (*left*) and the ME (*right*) approach. *Top* The reconstruction error as a function of the number of columns used in the least square method, Eq. (20). The different *curves* plotted are obtained for $\sigma = 0.5, 1.0, 1.5, 2.0$ (*bottom to top*) and a 100 ms-decoding window. *Bottom* The reconstruction error as a function of the place field width (σ). Results are obtained for $T_w = 50, 100$ ms

What has hindered the application of this approach to more realistic settings, is the difficulty of deriving the exact probability distribution of the interspike intervals generated by even the simplest spiking neuronal models. A further obstacle is represented by the presence of *censoring* in the data, due to the short time windows during which spike trains can be reasonably approximated as stationary when they are produced in response to time-dependent stimuli.

Here, we have succesfully tackled the latter problem by introducing a 'censored' maximum-likelihood estimator. This allows us to obtain unbiased estimates of the input even when short windows are used for decoding.

We have applied our approach to decode the spike data generated by the leaky integrate-and-fire model with balanced inputs, a model for which we could obtain the interspike interval probability density in closed form. Thanks to the higher efficiency of CMLE compared with a rate-based estimator, the decoding accuracy can be increased towards its theoretical limit. In particular, we found that, with a population of 100 neurons and a time window of only 50 ms for decoding, a dynamical stimulus can be read out with 90% accuracy, using as few as about 1.2 spikes per neuron.

The potential of CMLE for decoding applications has been illustrated in a simple tracking task. We found that a rapidly moving target could be tracked accurately using no other information than the spike activity elicited in the network during 50 ms-time windows. This confirms that, in spite of the noisy character of the individual spike trains, the ensemble firing pattern in a spiking network may contain enough information to allow efficient computation.

It is arguable that real neurons would generate interspike intervals that conform to the probability distribution used in this paper, although in many cases the latter may provide a better approximation than other models used so far in the literature, such as the exponential, the gamma or the inverse Gaussian distribution.

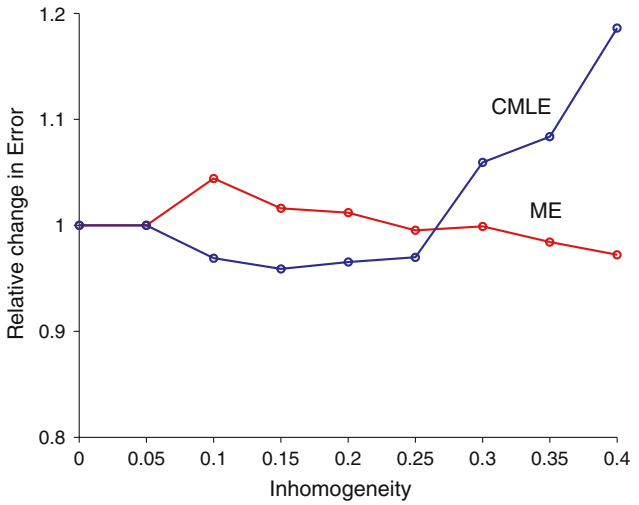

Fig. 12 Tracking a moving stimulus: inhomogeneous case. The reconstruction error (relative to the homogeneous case) is plotted against the population's inhomogeneity parameter, δ. For any value of δ, the neuronal thresholds are randomly distributed within the interval $\left[V_{\text{thre}}(1-\delta), V_{\text{thre}}(1+\delta)\right]$. Results are obtained for $\sigma = 1.5$, $T_{\text{w}} = 50\,\text{ms}$, using CMLE (*blue*) and ME (*red*). Reconstruction errors are averaged over 40 trajectory points for each trial

However, regardless of how good our model is in approximating real neurons, the results presented in the paper clearly show the advantage of employing an optimal decoding strategy based on the distribution of the interspike intervals. To this regard, we speculate that brains may have evolved similar strategies for decoding population codes. This might have pushed the decoding accuracy beyond what is achievable using simple rate codes, towards the theoretical limit illustrated here for maximum-likelihood estimation.

As for the pratical applications of our decoding approach, we note that the censored maximum-likelihood method does not depend per se on the actual distribution of the interspike intervals. Obviously, when the latter is unknown, one must resort to a model that approximates the data, which may sensibly reduce the efficiency of the estimator. In general, this requires fitting a parametric family of probability distributions, or a mixture of them, to the interspike interval distributions observed at varying values of the stimulus, taking into account the presence of censoring as discussed in Rossoni and Feng (2006). As a result, one obtains a map between the parameters of the stimulus and that of the approximating distribution. Using such approximated encoding model in the expression of the censored likelihood, the time course of a stimulus can then be reconstructed from the ensemble spike activity.

Finally, we list here two issues that we intend to explore in further developments of this work.

Negative correlations inside each column The analysis of multi-electrode spike data has revealed a small negative cross-correlation of activity between neurons, for instance in the olfactory bulb (Horton et al. 2005). By virtue of the central limit theorem, similar negative correlations would lead to a partial noise cancellation in the neural response. Unfortunately, our CMLE approach is no longer applicable in presence of interactions. On the other hand, a rate-based estimator may become significatively more efficient under such circumstances than in the independent case discussed here. As a result, we expect to see an improvement in the performance of a tracking algorithm based on moment estimation.

Positive correlations between columns A positive correlation between the activity of neurons in neighboring columns, would cause stimulus' estimates in different columns to vary coherently. Since our decoding scheme does not rely on the absolute values of the estimated stimulus, but rather on the *relative* response of each column, we expect this kind of correlation may even lead to an improvement in the performance.

Acknowledgements J.F. was partially supported by grants from UK EPSRC(GR/R54569), (GR/S20574), and (GR/S30443).

Appendix: Poisson case

Denote t_i^k as the kth interspike interval of the ith neuron of the population and $t_i^k = \tau_i^{k,s} - \tau_i^{k,e}$ where $\tau_i^{k,s}(\tau_i^{k,e})$ is the starting (end) spike time. According to Eq. (10) we then have

$$\begin{aligned}\hat{t} &= \frac{\sum_{i=1}^{N}\sum_{k=1}^{N_R^i} t_i^k + \sum_{i=1}^{N_T}\left(T - \tau_i^{1,s} - \sum_{k=1}^{N_R^i} t_i^k\right)}{\sum_{i=1}^{N} N_R^i} \\ &= \frac{\sum_{i=1}^{N_T}\left(T-\tau_i^{1,s}\right)}{\sum_{i=1}^{N} N_R^i} \\ &= \frac{\sum_{i=1}^{N_T}\left(T-\tau_i^{1,s}\right)}{N} \cdot \frac{1}{\sum_{i=1}^{N} N_R^i/N}\end{aligned} \tag{23}$$

where $\hat{t}$ is the CMLE for the interspike interval, and T is the time window. From the law of large numbers we obtain

$$\begin{aligned}\frac{\sum_{i=1}^{N_T}\left(T-\tau_i^{1,s}\right)}{N} &= \left\langle\left(T-\tau_1^{1,s}\right) I_{\{N_1(T)\geq 1\}}\right\rangle \\ &= T(1-P(N_1(T)=0)) - \int_0^T t\lambda\exp(-\lambda t)\mathrm{d}t \\ &= T - \frac{1}{\lambda}(1-\exp(-\lambda T))\end{aligned} \tag{24}$$

and

$$\frac{\sum_{i=1}^{N} N_R^i}{N} = \left\langle (N_1(T)-1) I_{\{N_1(T))\geq 2\}} \right\rangle = \lambda T - 1 + \exp(-\lambda T) \quad (25)$$

where $N_i(T)$ is the number of spikes of the ith neuron within the time window T. Combining equations above, we conclude that

$$\hat{t} = \frac{1}{\lambda}, \quad (26)$$

i.e. the CMLE is an unbiased estimator for the Poisson process and it is independent of the time window.

References

Amari S, Nakahara H (2005) Difficulty of singularity in population coding. Neural Comput 17:839–858

Andersen PK, Borgan O, Gill RD, Keiding N (1993) Statistical models based on counting processes. Springer, Berlin Heidelberg New York

Brown, EN, Kaas RE, Mitra PP (2004) Multiple neural spike train data analysis: state-of-the-art and future challenges. Nat Neurosci 7:456–461

Feng JF (2004) Computational neuroscience: a comprehensive approach. Chapman & Hall/CRC, Boca Raton

Feng J, Ding M (2004) Decoding spikes in a spiking neuronal network. J Phys A Math Gen 37:5713–5727

Gerstner W, Kistler WM (2002) Spiking neuron models: single neurons, populations, plasticity. Cambridge University Press, Cambridge, UK

Gnedenko BV, Belyayev YK, Solovyev AD (1969) Mathematical methods of reliability theory. Academic, New York

Harris KD (2004) Hallucinations and nonsensory correlates of neural activity. Behav Brain Sci 27:796–796

Hollander M, Pena EA (2004) Nonparametric methods in reliability. Stat Sci 19:644–651

Horton P, Nicol A, Kendrick K, Feng JF (2005) Applications of multi-variate analysis of variance (MANOVA) to multi-electrode array electrophysiology data. J Neurosci Methods 166:22–41

Klein JP, Moeschberger ML (1997) Survival analysis: techniques for censored and truncated data. Springer, Berlin Heidelberg New York

Paz R, Wise SP, Vaadia E (2004) Viewing and doing: similar cortical mechanisms for perceptual and motor learning. Trends Neurosci 27:496–503

Peng XM, Van Essen DC (2005) Peaked encoding of relative luminance in macaque areas V1 and V2. J Neurophysiol 93:1620–1632

Pettitt AN, Kelly JM, Gao JT (1998) Bias correction for censored data with exponential lifetimes. Statist Sinica 8:941–963

Raffi M, Siegel RM (2005) Functional architecture of spatial attention in the parietal cortex of the behaving monkey. J Neurosci 25:5171–5186

Rolls ET, Tovee MJ (1994) Processing speed in the cerebral cortex and the neurophysiology of visual masking. Proc R Soc Lond 257:9–15

Rossoni E, Feng J (2006) Non-parametric approach to extract information from interspike intervals. J Neurosci Methods 150:30–40

Sanger TD (2003) Neural population codes. Curr Opin Neurobiol 13:238–249

Schaette R, Gollisch T, Herz AVM (2005) Spike-train variability of auditory neurons in vivo: dynamic responses follow predictions from constant stimuli. J Neurophysiol 93:3270–3281

Shadlen MN, Newsome WT (1998) The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J Neurosci 18:3870–3896

Thorpe S, Fize D, Marlot C (1996) Speed of processing in the human visual system. Nature 381:520–522

Tuckwell HC (1988) Introduction to theoretical neurobiology: nonlinear and stochastic theories. Cambridge University Press, New York

Verbeek A (1989) The compactification of generalised linear models. In: Decorli A, Frances B, Gilchrist R and Seebeer G (eds) Statistical modelling. Springer, New York, pp 314–327