Spike sorting by stochastic simulation.

Di Ge, Eric Le Carpentier, Jérôme Idier, Dario Farina

Spike Sorting by Stochastic Simulation

Di Ge, Eric Le Carpentier, Jérôme Idier (*Member, IEEE*) and Dario Farina (*Senior Member, IEEE*)

***Abstract*—The decomposition of multiunit signals consists of the restoration of spike trains and action potentials in neural or muscular recordings. Because of the complexity of automatic decomposition, semiautomatic procedures are sometimes chosen. The main difficulty in automatic decomposition is the resolution of temporally overlapped potentials. In a previous study [1], we proposed a Bayesian model coupled with a maximum *a posteriori* (MAP) estimator for fully automatic decomposition of multiunit recordings and we showed applications to intramuscular EMG signals. In this study, we propose a more complex signal model that includes the variability in amplitude of each unit potential. Moreover, we propose the Markov Chain Monte Carlo (MCMC) simulation and a Bayesian minimum mean square error (MMSE) estimator by averaging on samples that converge in distribution to the joint posterior law. We prove the convergence property of this approach mathematically and we test the method representatively on intramuscular multiunit recordings. The results showed that its average accuracy in spike identification is greater than 90% for intramuscular signals with up to 8 concurrently active units. In addition to intramuscular signals, the method can be applied for spike sorting of other types of multiunit recordings.**

***Index Terms*—Bayesian model, MMSE estimation, Markov chain Monte Carlo, intramuscular EMG decomposition**

I. Introduction

Invasive electrodes inserted into the brain, nerves, or muscles provide multiunit recordings consisting of the activity of neural cells or muscle fibers that respond to the activity of motor neurons. In several applications, it is necessary to decompose these multiunit recordings into the individual sources, i.e. to identify the individual unit spike trains from the interference signal. For example, the decomposition of intramuscular recordings provides information on the behavior of spinal motor neurons. The decomposition problem of multiunit recordings, often referred to as spike sorting, is usually solved with template matching approaches [2]. In some of these approaches, the action potentials that are overlapped in time are not classified and are considered outliers [3]. When the number of sources is small, this limitation may provide an acceptable amount of information on the sources under study. However, when there are many sources active concurrently, overlapped action potentials constitute the majority of the potentials in the recorded signals. Therefore, in some template matching methods, the superpositions are resolved by iterative subtraction of all possible template combinations from unidentified waveforms [4], [5]. As an alternative, neural network classifiers are applied to resolve the superposition problem by introducing overlapped spikes into the training data [6].

The main challenge underlying the resolution of superimposed spikes is that the global optimization problem is non-deterministic polynomial-type (NP) hard, i.e., it cannot be solved by polynomial complexity algorithms [1]. Therefore, the existing methods either perform on the restrained search spaces [4], [7], which reduces the complexity, or are based on recursive algorithms [8], [9] with specific trial strategies and residual threshold estimations. For example, Atiya [7] proposed a robust approach for decomposing overlaps of action potentials in neural recordings by comparing all possible combinations of up to two action potentials (restrained search space). Within the family of spike sorting algorithms, the dynamic programming method [4], which uses the fast exploration technique of a k-d tree, is also limited by the memory space necessary to generate such data structure, resulting in practice in an equally restrained search space of up to two overlapping action potentials. This constraint is not justified in several applications. Similar issues arise when determining a representative training set of overlapped spikes and then using neural networks to identify overlapped sources [6].

Other approaches are not limited in the number of overlapping sources but require an interaction with an operator. For example, a recent algorithm for decomposing intramuscular electromyographic (EMG) signals into the constituent motor unit spike trains recursively matches the templates and reevaluates the residual errors,

D. Ge is with the Glaizer groupe, 32 rue Guy Moquet - 92240 Malakoff France (email: ge.di@glaizer.com)

E. Le Carpentier and J. Idier are with the IRCCyN UMR-CNRS 6597, École Centrale de Nantes, 1 rue de la Noë, 44321 Nantes Cedex 03, France (email: Eric.Le-Carpentier|Jerome.Idier@irccyn.ec-nantes.fr)

D. Farina is with the Department of Neurorehabilitation Engineering, Bernstein Center for Computational Neuroscience, University Medical Center Göttingen, Georg-August University, Göttingen, Germany (email: dario.farina@bccn.uni-goettingen.de)

until it is able to prove that it has found the global optimum [8]. To avoid too long recursions in cases of difficult spike superpositions, the algorithm stops and requires an expert intervention after a number of trials without reaching the lower bound value.

Other methods make use of independent component analysis (ICA), of clustering and ICA [10], [11], [12] to avoid the limitation on the number of units concurrently active and the need for the intervention of an operator. However, these methods are applied to multi-channel recordings. Similarly, blind source separation (BSS) approaches such as the Convolution Kernel Compensation (CKC) used for surface EMG decomposition [13], [14], are usually applied to multi-channel recordings for which at least as many channels as sources are needed.

A Bayesian model coupled with a maximum *a posteriori* (MAP) estimator was recently proposed for fully automatic decomposition of multiunit recordings in single-channel signals [1]. The method was tested on intramuscular EMG recordings to identify individual motor unit spike trains and proved to have similar performance as obtained with a semi-automatic decomposition by expert operators [1]. Furthermore, the number of concurrent action potentials was not limited although the search space augmented exponentially.

One assumption of the Bayesian model in [1] is that action potentials discharged by a unit have the same amplitude and shape for the entire recording. However, although the shape of action potentials may not change substantially, the amplitude may be variable in some conditions. For example, the amplitude of single motor unit action potentials in intramuscular EMG may be influenced by the rate of discharge because of the velocity-recovery function of muscle fibers [15]. Moreover, small displacements of the recording electrodes with respect to the sources may influence action potential amplitude.

In this study, we address the spike sorting problem in single-channel recordings by proposing a more complex signal model than in [1]. This model includes the variability in magnitudes of each unit potential to provide a better fit with the experimental signals in a least-squared-residual sense. Since action potentials may also change in shape, which is not modeled, this new model is not expected to provide a perfect fit. However, it is hypothesized that it would provide better overall performance compared with the invariant amplitude and shape model. Moreover, instead of the MAP estimator [1], we propose a new approach for the solution of the spike sorting problem. The approach is based on the Markov Chain Monte Carlo (MCMC) simulation tool to build a minimum mean square error (MMSE) estimator for continuous parameters and a marginal MAP estimator for discrete parameters (discharge instants) using samples that converge in distribution to the joint posterior law. The study also provides the mathematical proof of the convergence property of this approach. As in [1], the proposed method is representatively tested on intramuscular multiunit recordings, although it can also be directly applied to other multiunit signals. The test on intramuscular EMG signals will allow a direct performance comparison with the MAP estimator previously proposed [1]. The main part of the validation is performed on experimental data, however we also present results on EMG signals simulated with very irregular spike trains as a proof of its robustness.

II. Model with variable impulse magnitudes

The generation of a multiunit recording is schematically represented in Fig. 1. Each impulse train is convolved by a characteristic action potential whose shape is assumed to vary minimally during the recording.

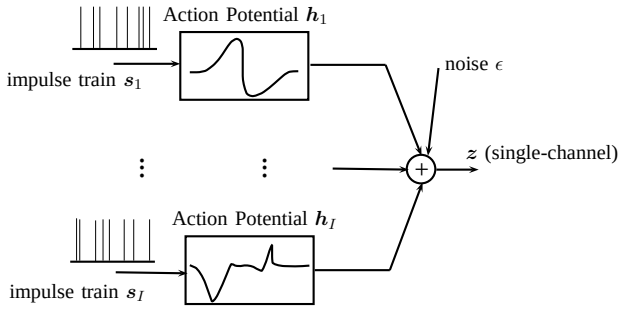

Fig. 1. Multiunit direct model.

The direct problem model can thus be expressed mathematically as a convolution product:

$$\boldsymbol{z} = \sum_{i=1}^{I} \boldsymbol{h}_i * \boldsymbol{s}_i + \epsilon \tag{1}$$

where $\boldsymbol{z}$ is the recorded signal of length N, $\boldsymbol{s}_i, i = 1, \ldots, I$ and $\boldsymbol{h}_i, i = 1, \ldots, I$ are the impulse trains (the discharge patterns) and their linear filters (action potentials) respectively. Note that in the present study, the impulses in $\boldsymbol{s}_i$ can have different amplitudes reflecting the amplitude variation of the different action potentials in the train. $(\bullet)_i$ denotes the set $\{\bullet, i = 1 \ldots, I\}$, *e.g.*, $(\boldsymbol{s}_i)_i = \{\boldsymbol{s}_i, i = 1 \ldots, I\}$ gathers the discharge patterns of all impulse trains.

Assuming that ϵ is an independent, identically distributed (i.i.d.) Gaussian noise with unknown variance σ_ϵ^2, the likelihood of the data, given the source parameters $(\boldsymbol{s}_i, \boldsymbol{h}_i)_i, \sigma_\epsilon^2$ can be written:

$$P(\boldsymbol{z} \,|\, (\boldsymbol{s}_i, \boldsymbol{h}_i)_i, \sigma_\epsilon^2) = \frac{1}{(2\pi\sigma_\epsilon^2)^{\frac{N}{2}}} e^{-\frac{|\,\boldsymbol{z} - \sum_i \boldsymbol{s}_i * \boldsymbol{h}_i\,|^2}{2\sigma_\epsilon^2}} \tag{2}$$

A. Prior laws on parameters

The modeling of the impulse trains $(\boldsymbol{s}_i)_i$ is based on the following assumptions and notations:

A1) $(\boldsymbol{s}_i)_i$ are supposed mutually independent discrete time impulse trains;

A2) the coordinates of non-zero elements (discrete discharge instants) in each $\boldsymbol{s}_i$ are contained in a vector noted as $\boldsymbol{x}_i$;

A3) each $\boldsymbol{s}_i$ conditional to $\boldsymbol{x}_i$ and $\sigma^2_{\boldsymbol{s}_i}$ is supposed to be a truncated Gaussian:

$$\boldsymbol{s}_i \,|\, \boldsymbol{x}_i, \sigma^2_{\boldsymbol{s}_i} \sim \mathcal{N}(\mathbf{1}_{\boldsymbol{x}_i}, \sigma^2_{\boldsymbol{s}_i} \mathrm{diag}(\mathbf{1}_{\boldsymbol{x}_i})) \cdot \mathbf{1}_{\boldsymbol{s}_i \geq 0} \quad (3)$$

where $\mathbf{1}_{\boldsymbol{x}_i}$ is binary vector with ones at positions $\boldsymbol{x}_i$, and $\mathrm{diag}(\mathbf{1}_{\boldsymbol{x}_i})$ is the diagonal matrix whose diagonal is $\mathbf{1}_{\boldsymbol{x}_i}$.

The sparsity in each spike train is modeled by the vector $\boldsymbol{x}_i$, whereas the variability in the amplitude of spikes generated by the same source is taken into account via the conditional Gaussian law (Eq.(3)). We note that negative magnitudes are not relevant in spike sorting, thus the law of the magnitudes should be a Gaussian truncated to the positive hyperoctant, *i.e.*, $\mathbf{1}_{\boldsymbol{s}_i \geq 0}$ imposes a non-negative constraint on each element of $\boldsymbol{s}_i$.

In the literature, the Bernoulli-Gaussian (BG) model has been adopted for the restoration of a sparse and magnitude variable train [16], [17], [18], [19]. The BG model first defines the discharge instants $\boldsymbol{x}_i$ using the Bernoulli law:

$$P(\boldsymbol{x}_i) = \lambda^{n_i}(1-\lambda)^{N-n_i}$$

where $n_i = \dim(\boldsymbol{x}_i)$ denotes the unknown number of discharges, and $\lambda \in (0,1)$ the Bernoulli parameter; and then defines $\boldsymbol{s}_i$ conditionally to $\boldsymbol{x}_i$ as a Gaussian variable, in the same manner as in Eq. (3). Let us remark that a Gaussian random variable with zero-variance corresponds to a variable of constant value (the Gaussian mean m), for which the distribution can also be defined as a Dirac.

On the other hand, the Bernoulli model needs also adaptations when used for neural spike trains, in order to account for physiological constraints. In this study, the model of the impulse instants $(\boldsymbol{x}_i)_i$ is based on the following assumptions:

A4) the *inter-spike interval* (ISI) $T_{ij} = \boldsymbol{x}_{i,j+1} - \boldsymbol{x}_{i,j}$, or the temporal distance between two consecutive impulses of the same source, is assumed to be larger than a threshold value T_R. This value represents the *refractory period* needed to discharge the spikes and thus depends on the specific application; *e.g.*, for intramuscular EMG decomposition, it is the absolute refractory period of muscle fibers;

A5) among all the admissible solutions satisfying the refractory period condition, spike trains that are more regularly spaced are favored. Following [1], this is achieved by a Gaussian-shaped distribution on the variables $T_{ij} - T_R$.

The resulting law of $\boldsymbol{x}_i$ (in a discrete, regularly sampled time framework) for each unit can be written:

$$P\left(\boldsymbol{x}_i \,|\; m_i, \sigma_i^2\right) \propto 1_{\mathcal{C}}(\boldsymbol{x}_i)\, f_{m_i,\sigma_i^2}^{-(n_i-1)} \exp\left\{ -\frac{1}{2\sigma_i^2} \sum_{j=1}^{n_i-1} (\boldsymbol{x}_{i,j+1} - \boldsymbol{x}_{i,j} - m_i - T_R)^2 \right\} \quad (4)$$

where

$$f_{m_i,\sigma_i^2} = \sum_{k=-m_i}^{+\infty} \exp\left(-\frac{k^2}{2\sigma_i^2}\right),$$

and $\mathcal{C} = \{\boldsymbol{x}_i,\ \boldsymbol{x}_{i,j+1} - \boldsymbol{x}_{i,j} > T_R \text{ for all } j\}$ is the set of admissible $\boldsymbol{x}_i$ satisfying the refractory period condition, $1_{\mathcal{C}}$ is the indicator function of $\mathcal{C}$:

$$1_{\mathcal{C}}(\boldsymbol{x}_i) = \prod_{j=1}^{n_i-1} 1_{\{\boldsymbol{x}_{i,j+1} - \boldsymbol{x}_{i,j} > T_R\}}.$$

Contrarily to the BG case, a spike train $\boldsymbol{s}_i$ is no more an i.i.d. sequence, because the discharge instants $\boldsymbol{x}_i$ are not independent. In the BG deconvolution framework, a comparable model called *modified Bernoulli process* has been proposed in [20] to impose the minimum distance constraint on the impulse train. The integration constant f_{m_i,σ_i^2} depends on both m_i and σ_i^2. In practice, the following approximation can be adopted:

$$\begin{aligned} f_{m_i,\sigma_i^2} &\approx \int_{-m_i}^{+\infty} \exp\left(-\frac{k^2}{2\sigma_i^2}\right) \mathrm{d}k \\ &\approx \int_{-\infty}^{+\infty} \exp\left(-\frac{k^2}{2\sigma_i^2}\right) \mathrm{d}k = \sqrt{2\pi\sigma_i^2}, \end{aligned} \quad (5)$$

under the assumptions that

- σ_i is large enough to ignore the discretization error; *e.g.*, for $\sigma_i = 10$ms and a sampling frequency of 10kHz, the 3σ Gaussian lobe is discretized by 600 samples;
- σ_i/m_i is small enough to ignore the integral truncation at $-m_i$, typically if $\sigma_i/m_i < 1/3$, then the truncation error is controlled at 0.135%.

Finally, in order to define a proper probability, Eq. (4) should be modified to incorporate a prior law on the first discharge instant x_{i1}. Here we suppose it is uniform for the sake of simplicity.

Using the independence assumption A1, we obtain:

$$P\left((\boldsymbol{x}_i)_i \,|\; (m_i, \sigma_i^2)_i\right) = \prod_{i=1}^{I} P\left(\boldsymbol{x}_i \,|\, m_i, \sigma_i^2\right). \quad (6)$$

B. Joint posterior law

In the Bayesian framework, the posterior distribution $P(\Theta \mid \boldsymbol{z})$ for $\Theta = \left\{(\boldsymbol{x}_i, \boldsymbol{s}_i, m_i, \sigma_i^2, \boldsymbol{h}_i, \sigma_{\boldsymbol{s}_i}^2)_i, \sigma_\epsilon^2\right\}$ can be expressed:

$$P(\Theta \mid \boldsymbol{z}) \propto P(\boldsymbol{z} \mid (\boldsymbol{s}_i, \boldsymbol{h}_i)_i, \sigma_\epsilon^2) P(\sigma_\epsilon^2) \prod_i P(\boldsymbol{s}_i \mid \boldsymbol{x}_i, \sigma_{\boldsymbol{s}_i}^2) P(\sigma_{\boldsymbol{s}_i}^2) P(\boldsymbol{x}_i \mid m_i, \sigma_i^2) P(m_i) P(\sigma_i^2) P(\boldsymbol{h}_i) \tag{7}$$

Fig. 2 illustrates the hierarchical Bayesian model with all intervening parameters, among which unknown parameters (in Θ) are circled. The decomposition task consists of deriving an estimator $\hat{\Theta}$ from Eq. (7). We note that Θ contains both continuous and discrete parameters and that the combinatorial nature of $(\boldsymbol{x}_i)_i$, precludes the exhaustive exploration method even in one segment [7].

Eqs. (2), (3) and (4) provide the likelihood term and the prior laws on $(\boldsymbol{s}_i, \boldsymbol{x}_i)_i$ that enter the joint posterior (7). For the remaining terms, we choose conjugate priors with non-informative hyper-parameters $(\alpha_i, \beta_i)_i$, $(\alpha_s, \beta_s, \mu_0, \sigma_0^2, \sigma_{\boldsymbol{h}}^2)$ (see [1] for a discussion on this choice):

$$\begin{aligned} m_i &\sim \mathcal{N}(\mu_0, \sigma_0^2), \quad \boldsymbol{h}_i \sim \mathcal{N}(\boldsymbol{h}_i^{(0)}, \sigma_{\boldsymbol{h}}^2 \mathbf{I}), \\ \sigma_i^2 &\sim \mathcal{IG}(\alpha_i, \beta_i), \quad \sigma_\epsilon^2 \sim \mathcal{IG}(\alpha_\epsilon, \beta_\epsilon), \\ \sigma_{\boldsymbol{s}_i}^2 &\sim \mathcal{IG}(\alpha_{\boldsymbol{s}}, \beta_{\boldsymbol{s}}), \end{aligned}$$

where $\mathcal{IG}$ stands for the inverse Gamma distribution and $\mathbf{I}$ is an identity matrix of appropriate size. $(\boldsymbol{h}_i^{(0)})_i$ denote approximate spike shapes that are determined by the preprocessing step described in the following section.

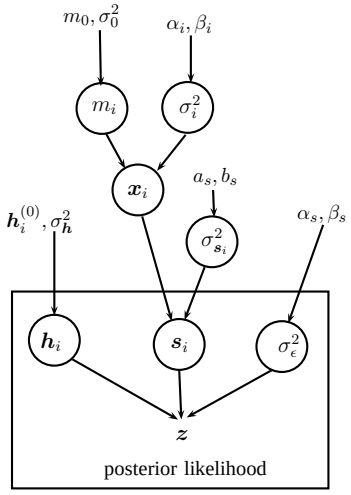

Fig. 2. Directed acyclic graph representing the hierarchical Bayesian model. Unknown parameters are circled in the graph.

III. Preprocessing

The resolution of superimposed spikes and complete decomposition is preceded by a preprocessing phase. The preprocessing consists of: 1) filtering the signal to enhance spike train activities, 2) segmenting the filtered signal into temporal intervals containing spikes [21], [9], 3) classifying isolated individual spikes (those not overlapped temporally with others) to determine the number of units I and the approximate spike shape for each class $\boldsymbol{h}_i^{(0)}$.

The preprocessing steps described above have robust solutions described in the literature. The thresholding method proposed in [21] was adopted in this study for the segmentation. The classification of isolated action potentials can be performed with several techniques. For example, the canonically registered discrete Fourier transform (CRDFT) [22] or non-parametric Bayesian estimation (NPB) [23], have been proven to be effective methods for this purpose [24]. These solutions for segmentation and classification of isolated spikes have excellent performance and do not need substantial improvements. Therefore, the methods described in [21] and [22] have been used in this study for preprocessing. On the contrary, the decomposition of temporally overlapped spikes poses more challenges and is the main contribution of this study, as described in the following.

IV. Decomposition of overlapped potentials using MCMC

Markov chain Monte Carlo (MCMC) methods are a class of stochastic simulation algorithms used to construct a Markov chain whose stationary distribution is invariant and converges to the desired distribution after a number of iterations (burn-in period). The estimator is then calculated from valid samples (*i.e.*, those after the burn-in period) and its quality improves as a function of the sample population according to the Monte Carlo principle. The *Metropolis-Hastings* (MH) and *Gibbs* algorithms are among the most classic algorithms in the MCMC family, while *Reversible Jump MCMC* (RJMCMC) [25] further extends the application field by including variable dimension problems. These algorithms constitute widely used numerical tools in the field of Bayesian statistics and computational physics [26].

In the multiunit spike sorting context, the joint posterior law in Eq. (7) is considered as the distribution of interest. The Markov chain is generated using a *Gibbs* sampler, by re-sampling iteratively each parameter in Θ according to its posterior conditional law derived from Eq. (7) while fixing the other parameters. The choice of conjugate prior laws on $\{(m_i, \sigma_i^2, \boldsymbol{h}_i, \sigma_\epsilon^2)_i, \sigma_\epsilon^2\}$ facilitates

the re-sampling on continuous parameters (steps (6)-(10) of Tab. I). The conditional laws to sample steps (6)-(10) are detailed in [1]. In particular, the conditional law for m_i and σ_i^2 are, respectively, Gaussian and Inverse Gamma if the two approximations (5) are satisfied.

The difficulty lies in the sampling of the discrete parameters $(\boldsymbol{x}_i)_i^{(k)}$[1]. We denote $\mathcal{C}_k$ as the set of firing instants $(\boldsymbol{x}_i)_i^{(k)}$ of all units in the k-th segment that satisfy the refractory period condition with respect to $(\boldsymbol{x}_i)_i^{(-k)}$. Thus, each element in $\mathcal{C}_k$ yields a non-zero joint posterior probability according to Eq. (4). Let us roughly evaluate the cardinality of $\mathcal{C}_k$ to stress the computational burden of sampling $(\boldsymbol{x}_i)_i^{(k)}$. Even in a very favorable case for which at most 4 units are active in a segment of duration of 8ms sampled at 10kHz, we would have $81^4 < |\mathcal{C}_k| < 2^{500}$. The inferior bound corresponds to a subspace of $\mathcal{C}_k$ containing at most one impulse for each unit in the segment. For segment lengths that vary between 8 and 60ms and number of units that varies between 4 and 8, the cardinal of $|\mathcal{C}_k|$ is at least 4.30×10^7 and 1.68×10^{22} in each case. Thus, sampling the conditional probabilities of $(\boldsymbol{x}_i)_i^{(k)}$ by evaluating all probabilities in the set $\mathcal{C}_k$ for each segment yields unrealistic computational load.

To solve this problem, we propose a Metropolis-Hastings algorithm summarized in Tab. I. The algorithm explores a subspace of $\mathcal{C}_k$ at each iteration Its validity is shown by verifying that the corresponding Markov chain remains irreducible: all configurations of non-zero probabilities ($\in \mathcal{C}_k$) are explored with non-zero probabilities regardless of the initial state. This proof is provided in appendix. Consequently, the MCMC algorithm in Tab. I is a valid stochastic simulation algorithm, of *Metropolis-Hastings within Gibbs* type.

TABLE I
STOCHASTIC SIMULATION ALGORITHM BY MCMC.

repeat
  for $k = 1 \dots K$ **do** % Integrate $(\boldsymbol{s}_i)_i$ analytically
    Sample $(\boldsymbol{x}_i)_i^{(k)} \sim P((\boldsymbol{x}_i)_i^{(k)} \,|\, \Theta \setminus \{(\boldsymbol{s}_i)_i, (\boldsymbol{x}_i)_i^{(k)}\}, \boldsymbol{z})$ % MH
  end for
  Sample $(\boldsymbol{s}_i)_i \sim P((\boldsymbol{s}_i)_i \,|\, \Theta \setminus (\boldsymbol{s}_i)_i, \boldsymbol{z})$
  Sample $(\boldsymbol{h}_i)_i \sim P((\boldsymbol{h}_i)_i \,|\, \Theta \setminus (\boldsymbol{h}_i)_i, \boldsymbol{z})$
  Sample $(m_i)_i \sim P((m_i)_i \,|\, \Theta \setminus (m_i)_i)$
  Sample $(\sigma_i^2)_i \sim P((\sigma_i^2)_i \,|\, \Theta \setminus (\sigma_i^2)_i)$
  Sample $(\sigma_{\boldsymbol{s}_i}^2)_i \sim P((\sigma_{\boldsymbol{s}_i}^2)_i \,|\, \Theta \setminus (\sigma_{\boldsymbol{s}_i}^2)_i)$
  Sample $\sigma_\epsilon^2 \sim P(\sigma_\epsilon^2 \,|\, \Theta \setminus \sigma_\epsilon^2, \boldsymbol{z})$
until number of iteration reached

In what follows, Sections IV-A and IV-B enter into more details about the sampling of $(\boldsymbol{x}_i)_i^{(k)}$ (first step in Tab. I): they respectively focus on its marginal conditional law and on the implementation of the Metropolis-Hastings step.

[1]The superscript (k) and $(-k)$ denote parameters in the k-th segment and parameters in all other segments, respectively.

A. Marginal conditional law

We first rewrite the convolution sum in the following matrix form:

$$\sum_{i=1}^{I} \boldsymbol{h}_i * \boldsymbol{s}_i = \mathbf{H}\boldsymbol{S}$$

where $\mathbf{H} = [\mathbf{H}_1, \dots, \mathbf{H}_I]$ is composed of convolution matrices such that $\boldsymbol{h}_i * \boldsymbol{s}_i = \mathbf{H}_i \boldsymbol{s}_i$, and $\boldsymbol{S} = [\boldsymbol{s}_1; \dots; \boldsymbol{s}_I]$ represents a $NI \times 1$ column vector by vertical concatenation. In the k-th segment, the data generation can be written:

$$\boldsymbol{z}^{(k)} = \mathbf{H}^{(k)} \boldsymbol{S}^{(k)} + \epsilon^{(k)}$$

where $\boldsymbol{z}^{(k)}$ is the signal in the k-th segment; $\mathbf{H}^{(k)} = [\mathbf{H}_1^{(k)}, \dots, \mathbf{H}_I^{(k)}]$ and $\boldsymbol{S}^{(k)} = [\boldsymbol{s}_1^{(k)}; \dots; \boldsymbol{s}_I^{(k)}]$ denote the corresponding submatrix of $\mathbf{H}$ and subvector of $\boldsymbol{S}$.

From Eq. (7), we obtain for each segment k the conditional law of $(\boldsymbol{s}_i, \boldsymbol{x}_i)_i^{(k)} \,|\, \text{rest}$. In the following, the term **rest** stands for the set $\{\Theta, \boldsymbol{z}\}$, except the concerned parameters.

$$\begin{aligned}
&P\Big((\boldsymbol{s}_i, \boldsymbol{x}_i)_i^{(k)} \,|\, \text{rest}\Big) \propto\; P\Big(\boldsymbol{z}^{(k)} \,|\, (\boldsymbol{s}_i^{(k)}, \boldsymbol{h}_i)_i, \sigma_\epsilon^2\Big) \\
&\prod_i P\Big(\boldsymbol{s}_i^{(k)} \,|\, \boldsymbol{x}_i^{(k)}, \sigma_{\boldsymbol{s}_i}^2\Big)\, P(\boldsymbol{x}_i^{(k)} \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)}\Big) \\
&\propto\; \exp\left(-\frac{1}{2\sigma_\epsilon^2}\big\|\boldsymbol{z}^{(k)} - \mathbf{G}\boldsymbol{a}_1\big\|^2\right) |\mathbf{V}|^{\frac{1}{2}}\, \delta(\boldsymbol{a}_0) \\
&\exp\Big(-\frac{1}{2}(\boldsymbol{a}_1 - \mathbf{1})^{\mathrm{t}}\, \mathbf{V}(\boldsymbol{a}_1 - \mathbf{1})\Big) \prod_i P\Big(\boldsymbol{x}_i^{(k)} \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)}\Big)
\end{aligned} \tag{8}$$

where the k-th segment specific matrices are noted without the superscript (k) for simplicity:

$$\begin{cases}
\mathbf{G} &= \Big[(\mathbf{H}_1^{(k)})_{\boldsymbol{x}_1^{(k)}}, \dots, (\mathbf{H}_I^{(k)})_{\boldsymbol{x}_I^{(k)}}\Big] \\
\mathbf{V} &= \operatorname{diag}\Big([\underbrace{\sigma_{\boldsymbol{s}_1}^{-2}, \dots, \sigma_{\boldsymbol{s}_1}^{-2}}_{n_1^{(k)}}, \dots, \underbrace{\sigma_{\boldsymbol{s}_I}^{-2}, \dots, \sigma_{\boldsymbol{s}_I}^{-2}}_{n_I^{(k)}}]\Big) \\
\boldsymbol{a}_1 &= \Big[(\boldsymbol{s}_1^{(k)})_{\boldsymbol{x}_1^{(k)}}, \dots, (\boldsymbol{s}_I^{(k)})_{\boldsymbol{x}_I^{(k)}}\Big] \\
\boldsymbol{a}_0 &= \boldsymbol{S}^{(k)} \setminus \boldsymbol{a}_1
\end{cases}$$

$\delta(\cdot)$ denotes the multi-dimensional Dirac function, and $n_i^{(k)} = \dim(\boldsymbol{x}_i^{(k)})$ is the number of impulses of unit $\#i$ in the segment k. The vector $\mathbf{1}$ is a column vector of size $\sum_i n_i^{(k)}$ while $\boldsymbol{a}_1$ and $\boldsymbol{a}_0$ are, respectively, the non-zero spike magnitudes of all units in the segment and its complement. $\{\boldsymbol{a}_1, \boldsymbol{a}_0\}$ constitutes thus a permutation of spike trains $\boldsymbol{S}^{(k)}$.

Let us remark that matrices $\mathbf{G}$ and $\mathbf{V}$ are functions of $(\boldsymbol{x}_i)_i^{(k)}, (\boldsymbol{h}_i, \sigma^2_{\boldsymbol{s}_i})_i$, but not of $(\boldsymbol{s}_i)_i^{(k)}$. It is thus possible to integrate out first $\boldsymbol{a}_0$, and then $\boldsymbol{a}_1$ in Eq. (8) since it is merely the product of two exponential terms and thus remains Gaussian.

The conditional probability of $(\boldsymbol{x}_i)_i^{(k)}$ after marginalization of $(\boldsymbol{s}_i)_i^{(k)}$ from Eq. (8) can be written:

$$P\Big((\boldsymbol{x}_i)_i^{(k)} \,|\, \text{rest} \setminus (\boldsymbol{s}_i)_i^{(k)}\Big) \propto \\ \big|\Sigma\big|^{\frac{1}{2}} \exp\Big(\frac{1}{2}\big(\boldsymbol{m}\big)^{\mathrm{t}}\big(\Sigma\big)^{-1}\boldsymbol{m} - \frac{1}{2}\mathbf{1}^{\mathrm{t}}\mathbf{V}\mathbf{1}\Big) \\ \prod_i \sigma_{\boldsymbol{s}_i}^{-n_i^{(k)}} P\left(\boldsymbol{x}_i^{(k)} \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)}\right) \tag{9}$$

where

$$\begin{cases} \Sigma^{-1} = \frac{1}{\sigma_\epsilon^2}\left(\mathbf{G}\right)^{\mathrm{t}}\mathbf{G} + \mathbf{V} \\ \boldsymbol{m} \quad = \Sigma\left(\frac{1}{\sigma_\epsilon^2}\left(\mathbf{G}\right)^{\mathrm{t}}\boldsymbol{z}^{(k)} + \mathbf{V}\mathbf{1}\right) \end{cases}$$

Σ^{-1} and $\boldsymbol{m}$ are, respectively, the variance and mean of the Gaussian vector $\boldsymbol{a}_1$. Let us remark that the marginal law in Eq. (9) involves Σ^{-1} and $\boldsymbol{m}$. In the special case where $(\boldsymbol{x}_i^{(k)})_i = \emptyset$, Σ^{-1} and $\boldsymbol{m}$ cannot be defined. To obtain the marginal distribution, we first rewrite Eq. (8):

$$P\Big((\boldsymbol{s}_i^{(k)})_i = \mathbf{0}, (\boldsymbol{x}_i)_i^{(k)} = \emptyset \,|\, \text{rest}\Big) \propto \exp\left(-\frac{\|\boldsymbol{z}^{(k)}\|^2}{2\sigma_\epsilon}\right) \\ \delta\big((\boldsymbol{s}_i^{(k)})_i\big) \prod_i P\left(\boldsymbol{x}_i^{(k)} = \emptyset \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)}\right).$$

Integrating out $(\boldsymbol{s}_i^{(k)})_i$ and considering the common exponential factor w.r.t. cases where $(\boldsymbol{x}_i)_i^{(k)} \neq \emptyset$, its marginal probability then can be written:

$$P\Big((\boldsymbol{x}_i)_i^{(k)} = \emptyset \,|\, \text{rest} \setminus (\boldsymbol{s}_i)_i^{(k)}\Big) \\ \propto \prod_i P\left(\boldsymbol{x}_i^{(k)} = \emptyset \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)}\right) \tag{10}$$

Eqs. (9) and (10) describe the marginal conditional probability of $(\boldsymbol{x}_i)_i^{(k)}$ up to a normalization factor.

The term $P(\boldsymbol{x}_i^{(k)} \,|\, m_i, \sigma_i^2, \boldsymbol{x}_i^{(-k)})$ can be directly derived from Eq. (4) and measures the regularity of the binary sequences $(\mathbf{1}_{\boldsymbol{x}_i})_i$ for each spike train while the Gaussian variability of magnitudes of the spikes is expressed in the first two terms in Eq. (9).

We recall that though the marginal conditional probability of $(\boldsymbol{x}_i)_i^{(k)}$ can be analytically expressed, its combinatorial space makes it a hard problem either to maximize (as proposed in [1] using a Tabu search) or to simulate the distribution using the MCMC approach. The next section proposes an MCMC algorithm that does not sample directly according to the conditional probability, but rather iteratively on local subspaces.

B. Metropolis-Hastings step

We propose here a reversible *Metropolis-Hastings* step to avoid the evaluation of all probabilities of $(\boldsymbol{x}_i)_i^{(k)} \in \mathcal{C}_k$. The strategy is to explore a *reasonable* subspace of $\mathcal{C}_k$ per iteration while the exploration of the whole space $\mathcal{C}_k$ is mathematically guaranteed in the long run. We note that this strategy can be adapted to the full model that includes Gaussian variability in spike magnitudes.

Let $\omega(\boldsymbol{u}) \subset \mathcal{C}_k$ be the local subspace (to be specified hereafter) to explore for each iteration and thus contains accessible configurations from the current configuration, noted by $\boldsymbol{u} = [\mathbf{1}_{\boldsymbol{x}_1^{(k)}}, \ldots, \mathbf{1}_{\boldsymbol{x}_I^{(k)}}]$ of length $I \cdot \dim(\text{Seg}_k)$. It is constructed by concatenating spike trains of all units in the segment k. In Tab. II, $P(\boldsymbol{u})$ denotes the conditional probability as specified by Eq. (9) and (10) for a particular configuration $\boldsymbol{u}$, $F(\boldsymbol{u})$ the sum of probabilities of the configurations in $\omega(\boldsymbol{u})$, *i.e.*, $F(\boldsymbol{u}) = \sum_{\boldsymbol{y}\in\omega(\boldsymbol{u})} P(\boldsymbol{y})$, and $q(\boldsymbol{u} \mapsto \boldsymbol{u}^+)$ the instrumental law or the probability of proposing $\boldsymbol{u}^+$ given the current state $\boldsymbol{u}$. Note that $\sum_{\boldsymbol{u}^+\in\omega(\boldsymbol{u})} q(\boldsymbol{u} \mapsto \boldsymbol{u}^+) = 1$.

TABLE II
METROPOLIS-HASTINGS STEP TO SAMPLE $(\boldsymbol{x}_i)_i^{(k)}$ IN TAB. I.

1: Propose $\boldsymbol{u} \mapsto \boldsymbol{u}^+$ using the instrumental law:

$$q(\boldsymbol{u} \mapsto \boldsymbol{u}^+) = \begin{cases} 0 & \text{if } \boldsymbol{u}^+ \notin \omega(\boldsymbol{u}) \\ P(\boldsymbol{u}^+)/F(\boldsymbol{u}) & \text{otherwise} \end{cases}$$

2: Accept $\boldsymbol{u}^+$ with probability

$$\rho(\boldsymbol{u} \mapsto \boldsymbol{u}^+) = \min\{1, F(\boldsymbol{u})/F(\boldsymbol{u}^+)\}$$

In this application, we define $\omega(\boldsymbol{u}) = \{\boldsymbol{v} \mid |\boldsymbol{u}-\boldsymbol{v}|_1 \leq 2 \text{ and } \big||\boldsymbol{u}|_1 - |\boldsymbol{v}|_1\big| \leq 1\}$ with $|\cdot|_1$ the L1-norm. The following conditions are verified:

1) $|\boldsymbol{u} - \boldsymbol{v}|_1 = 0 \Leftrightarrow \boldsymbol{u} = \boldsymbol{v}$, s.t. $\boldsymbol{u} \in \omega(\boldsymbol{u})$;
2) $|\boldsymbol{u} - \boldsymbol{v}|_1 = 1$, adding or removing a spike;
3) $|\boldsymbol{u} - \boldsymbol{v}|_1 = 2$ and $\big||\boldsymbol{u}|_1 - |\boldsymbol{v}|_1\big| \leq 1$, shifting an existing spike in the same spike train or being replaced by a spike in another train;
4) $\boldsymbol{u} \in \omega(\boldsymbol{v}) \Leftrightarrow \boldsymbol{v} \in \omega(\boldsymbol{u})$, together with condition 1), assures the reversibility of the chain: for all pairs $(\boldsymbol{u}, \boldsymbol{u}^+) \in \mathcal{C}_k$,

$$P(\boldsymbol{u})K(\boldsymbol{u} \mapsto \boldsymbol{u}^+) = P(\boldsymbol{u}^+)K(\boldsymbol{u}^+ \mapsto \boldsymbol{u}),$$

where $K(\boldsymbol{u} \mapsto \boldsymbol{u}^+)$ denotes the transition kernel derived from the Metropolis-Hastings step of Tab. II and equals to:

$$\begin{cases} \frac{P(\boldsymbol{u}^+)}{F(\boldsymbol{u})}\rho(\boldsymbol{u} \mapsto \boldsymbol{u}^+) + r(\boldsymbol{u})\delta_{\boldsymbol{u}}(\boldsymbol{u}^+) & \boldsymbol{u}^+ \in \omega(\boldsymbol{u}) \\ 0 & \text{otherwise} \end{cases} \tag{11}$$

$$\text{where} \quad r(\boldsymbol{u}) = \sum_{\boldsymbol{u}^+ \in \omega(\boldsymbol{u})} \frac{P(\boldsymbol{u}^+)}{F(\boldsymbol{u})} \left(1 - \rho(\boldsymbol{u} \mapsto \boldsymbol{u}^+)\right)$$

5) the Markov chain is irreducible: it is capable of exploring the entire space $\mathcal{C}_k$ (see demonstration in Appendix);
6) the Markov chain is also aperiodic since the kernel (11) satisfies:

$$\forall \boldsymbol{u}, P(\boldsymbol{u}) > 0 \Rightarrow K(\boldsymbol{u} \mapsto \boldsymbol{u}) > 0.$$

by verifying that $\frac{P(\boldsymbol{u})}{F(\boldsymbol{u})} > 0$ and $\rho(\boldsymbol{u} \mapsto \boldsymbol{u}) = 1$. It is thus possible to have two consecutive samples that are identical, and such a Markov chain cannot be periodic.
7) the complexity per iteration (measured by $|\omega(\boldsymbol{u})|$) remains linear with respect to to the segment length $\dim(\text{Seg}_k)$ and the number of units I.

In conclusion, the Markov chain generated from the proposed algorithm is irreducible, reversible and aperiodic. Therefore the only equilibrium distribution is $P(\boldsymbol{u})$ (which means from any initial state $\boldsymbol{u}$, $K^n(\boldsymbol{u}, \bullet) \xrightarrow{\mathcal{D}} P(\bullet)$ [27, Theorem 1]).

The numerical implementation of the spike decomposition algorithm is summarized in Tab. III.

TABLE III
IMPLEMENTATION OF THE MCMC ALGORITHM.

repeat
  for $k = 1, \ldots, K$ **do** % for each segment k
    % --- Metropolis-Hastings step---------
    $\boldsymbol{u} \leftarrow [\mathbf{1}_{\boldsymbol{x}_1^{(k)}}, \ldots, \mathbf{1}_{\boldsymbol{x}_I^{(k)}}]$
    Evaluate $P(\boldsymbol{u}^+), \boldsymbol{u}^+ \in \omega(\boldsymbol{u})$ in Eq. (9) (10)
    $F_{\boldsymbol{u}} \leftarrow \sum_{\boldsymbol{u}^+ \in \omega(\boldsymbol{u})} P(\boldsymbol{u}^+)$
    Propose $\boldsymbol{u}^* \sim P(\boldsymbol{u}^+)/F_{\boldsymbol{u}}, \boldsymbol{u}^+ \in \omega(\boldsymbol{u})$
    $F_{\boldsymbol{u}^*} \leftarrow \sum_{\boldsymbol{u}^+ \in \omega(\boldsymbol{u}^*)} P(\boldsymbol{u}^+)$
    Accept $\boldsymbol{u}^*$ with probability $\rho = \min\{1, F_{\boldsymbol{u}}/F_{\boldsymbol{u}^*}\}$
    % ---- Magnitudes sampling------
    $[\boldsymbol{s}_1^{(k)}, \ldots, \boldsymbol{s}_I^{(k)}] \leftarrow \mathbf{0}$
    if $(\boldsymbol{x}_i^{(k)})_i \neq \emptyset$ **then**
      $\mathbf{G} \leftarrow \left[(\mathbf{H}_1^{(k)})_{\boldsymbol{x}_1^{(k)}}, \ldots, (\mathbf{H}_I^{(k)})_{\boldsymbol{x}_I^{(k)}}\right]$
      $\boldsymbol{d} \leftarrow [\sigma_{\boldsymbol{s}_1}^{-2}\texttt{ones}(n_1^{(k)}, 1), \ldots, \sigma_{\boldsymbol{s}_I}^{-2}\texttt{ones}(n_I^{(k)}, 1)]$
      $\mathbf{V} \leftarrow \text{diag}(\boldsymbol{d})$
      $\Sigma^{-1} \leftarrow \frac{1}{\sigma_\epsilon^2}\mathbf{G}^{\text{t}}\mathbf{G} + \mathbf{V}$
      $\mathbf{Q} \leftarrow \texttt{chol}(\Sigma^{-1})$ % Cholesky factor of Σ^{-1}
      $\boldsymbol{b} \leftarrow \texttt{randn}(\sum_i n_i^{(k)}, 1)$ % iid Gaussian vector
      % $\boldsymbol{a}_1 \sim \mathcal{N}(\boldsymbol{m}, \Sigma)$, Eq. (9)
      $\boldsymbol{a}_1 \leftarrow \mathbf{Q}^{-1}(\mathbf{Q}^{-\text{t}}(\mathbf{G}^{\text{t}}\boldsymbol{z}^{(k)}/\sigma_\epsilon^2 + \mathbf{V} \cdot \mathbf{1}) + \boldsymbol{b})$
      $[(\boldsymbol{s}_1^{(k)})_{\boldsymbol{x}_1^{(k)}}, \ldots, (\boldsymbol{s}_I^{(k)})_{\boldsymbol{x}_I^{(k)}}] \leftarrow \boldsymbol{a}_1$
    end if
  end for
  % -Sampling of remaining parameters [1] -
  $\boldsymbol{h} = [\boldsymbol{h}_1; \ldots; \boldsymbol{h}_I], \boldsymbol{h} \sim P((\boldsymbol{h}_i)_i \,|\, \text{rest})$
  For each $m_i, m_i \sim P((m_i)_i \,|\, \text{rest})$
  For each $\sigma_i^2, \sigma_i^2 \sim P((\sigma_i^2)_i \,|\, \text{rest})$
  For each $\sigma_{\boldsymbol{s}_i}^2, \sigma_{\boldsymbol{s}_i}^2 \sim P((\sigma_{\boldsymbol{s}_i}^2)_i \,|\, \text{rest})$
  $\sigma_\epsilon^2 \sim P(\sigma_\epsilon^2 \,|\, \text{rest})$
until number of iteration reached

C. Bayesian estimators

With the convergence property of the Markov chain, we thus obtain for each parameter in Θ a population of random samples distributed according to its marginal posterior law $P(\cdot \,|\, \boldsymbol{z})$. For continuous parameters, *i.e.*, $\Theta \setminus (\boldsymbol{x}_i, \boldsymbol{s}_i)_i$, a *minimum mean square error* estimator (MMSE) optimizes the following criterion:

$$\begin{aligned}\widehat{\boldsymbol{y}} &= \min_{\boldsymbol{y}_0} \int \|\boldsymbol{y} - \boldsymbol{y}_0\|^2 P(\boldsymbol{y} \,|\, \boldsymbol{z}) d\boldsymbol{y} \\ &= \min_{\boldsymbol{y}_0} E_{P(\boldsymbol{y} \,|\, \boldsymbol{z})}\left(\|\boldsymbol{y} - \boldsymbol{y}_0\|^2\right) \\ &= E_{P(\boldsymbol{y} \,|\, \boldsymbol{z})}(\boldsymbol{y})\end{aligned}$$

According to the Monte Carlo principle, this estimator is approximated by:

$$\widehat{\boldsymbol{y}} = E_{P(\boldsymbol{y} \,|\, \boldsymbol{z})}(\boldsymbol{y}) = \int \boldsymbol{y} P(\boldsymbol{y} \,|\, \boldsymbol{z}) d\boldsymbol{y} \approx \sum_{j=N_0+1}^{N} \boldsymbol{y}_j,$$

where $\{\boldsymbol{y}_j\}_{j=1,\ldots,N}$ are simulated samples, N_0 and N represent, respectively, the number of burn-in iterations and the total number of iterations. In our tests, we fixed $N = 2N_0 = 200$.

The same estimator, however cannot be applied on the discrete parameters $(\boldsymbol{x}_i)_i$ since averaged firing instants do not have any physical interpretation. Note that different samples may contain vectors of $(\boldsymbol{x}_i)_i$ of different dimensions since the number of detected spikes are not fixed. An alternative is to estimate binary sequences $(\mathbf{1}_{\boldsymbol{x}_i})_i$ using a marginal component MAP detector, *i.e.*, the marginal majority vote for each instant (component).

However, considering only one component marginally at a time may result in counter-intuitive estimators. Suppose for example that in an interval $\mathcal{J}, 2 < \dim(\mathcal{J}) < T_R$, the probability that there is one and only one detected event is 1 ($\sum_{j \in \mathcal{J}} P(\boldsymbol{s}_{i,j} \neq 0) = 1$) for a particular sample population. Further suppose that no one event is more likely than the sum of the others (in terms of probability), and thus $P(\boldsymbol{s}_{i,j} \neq 0) < 0.5$ for all $j \in \mathcal{J}$. From these assumptions, it would then follow that $\widehat{\boldsymbol{s}}_{i,j} = 0$ for all $j \in \mathcal{J}$, using the marginal component MAP detector. The result is clearly counter-intuitive because the probability that there is no event in $\mathcal{J}$ is zero.

Recently, Kail *et al.* [20] proposed a MAP block detector. The main idea is to perform the majority vote on a block of the binary variables in $\mathbf{1}_{\boldsymbol{x}_i}$. The number of binary combinations in each block should be much smaller than the number of available samples (the Markov chain length) while the length of the block is

big enough to avoid false negatives using the marginal component MAP detector due to spike shifts in the sample population. The same strategy is adopted in this study for the estimation of the discrete parameters $(\boldsymbol{x}_i)_i$.

V. Validation on experimental and simulated signals

A. Experimental signals

The method proposed can be applied to several spike sorting applications, such as intraneural or intramuscular recordings. The assumptions of the model include a refractory period, which can be set depending on the application, and the cost associated to the firing irregularity. Note that the latter does not imply that irregular firings cannot be identified but simply that a solution with very irregular firings is less favored than one with more regular firings.

The validation of the method was performed on a set of experimental intramuscular EMG signals, where the spikes represent the activation of spinal motor neurons. These signals were chosen for direct comparison between the proposed method with the Tabu search method proposed in [1]. The same experimental signals as described in [1] were thus used for this study.

Briefly, the experimental signals were recorded from the abductor digiti minimi muscle of five healthy men (age, mean $\pm$ SD = 25.3 ± 4.5 yr) with a pair of wire electrodes made of Teflon coated stainless steel (A-M Systems, Carlsborg, WA, USA; diameter 50μm) inserted into the muscle with a 25 G needle. The intramuscular EMG signals were amplified and provided one bipolar recording (Counterpoint EMG, DANTEC Medical, Skovlunde, Denmark) that was band-pass filtered (500 Hz-5 kHz) and sampled at 10 kHz. A reference electrode was placed around the wrist. The decomposition method was evaluated on intervals of 20 s of signals recorded during isometric contractions at 10% of the maximal voluntary contraction (MVC) force. The validation was performed by comparing the results of the proposed methods with those provided as reference results by manual decomposition of an expert operator using the EMGLAB tool. The proposed method was applied in a fully automatic way, without any intervention by the operator.

B. Simulated signals

The experimental signals were recorded in conditions where the firing patterns were expected to be rather regular (isometric contractions). In addition to experimental signals, the method was also applied to a set of simulated signals. This set of simulations was used to test the performance of the algorithm in cases of very irregular spike firings. The simulated signals were generated with the intramuscular EMG model proposed in [28] and also previously applied for the validation of the method proposed in [1]. The simulated signals were obtained by an imposed coefficient of ISI variation of 60%, corresponding to very irregular firings. The other parameters of the model were set as in a previous study [1], including the level of noise that corresponded to a SNR of 10dB). In total, five signals were simulated with the same set of parameters. These signals differed from each other because the shape of the action potentials were varied within the library of shapes of the model [28] and because the spike trains were generated randomly, according to the imposed statistics.

C. Performance index

The performance index for each spike source was defined as [29]:

$$A(i) = \frac{N_{\text{Dis}}(i) - N_{\text{FP}}(i) - N_{\text{FN}}(i)}{N_{\text{Dis}}(i)} \times 100\%$$

where $N_{\text{Dis}}(i)$ are spike numbers of the unit $\#i$ detected by EMGLAB for the experimental signals or simulated by the model for the synthetic signals. $N_{\text{FN}}(i)$ and $N_{\text{FP}}(i)$ are respectively the false negatives and the false positives with respect to the reference (EMGLAB or the model). A detected impulse was considered a correct one if within a 1ms-window centered on the instant of reference for the same unit. The global performance of the decomposition was then measured by the following index [29]:

$$A = \frac{1}{I} \sum_{i=1}^{I} A(i)$$

For the experimental signals, discharge patterns for which $\sigma_i/m_i > 0.3$ were excluded from the analysis, since the ISI variability of motor unit spike trains is usually lower than this limit in the experimental conditions [30], [31]. Therefore, among the discharge patterns obtained by the application of the proposed method to experimental signals, only those satisfying the condition $\sigma_i/m_i < 0.3$ were selected for analysis and the global performance criterion A was thus averaged on the $A(i)$ of the validated units. This *a-posteriori* selection of sources is not necessary in other applications in which the firing instants are less regular. The decomposition results of simulated firings provide a test of the method in conditions of very irregular spike firing. The a-posteriori selection of sources was not applied to the simulated signals and results are reported as average over all motor units. Note that the ISI variability of the simulated spike

trains is doubled compared with the upper limit imposed to the experimental spike train decomposition.

D. Results of decomposition

The constant parameters of the Bayesian model (Fig. 2) and the Markov chain initializations are listed in Tab. IV. $(\boldsymbol{h}_i^{(0)})_i, \hat{\sigma}_\epsilon$ are results of the preprocessing and we fixed $\sigma_{\boldsymbol{h}} = 0.1 \max(\boldsymbol{h})$ in the tests.

TABLE IV
CONSTANT PARAMETERS IN FIG. 2 AND MARKOV CHAIN INITIALIZATION.

m_0	σ_0^2	α_i	β_i	α_s	β_s	a_s	b_s
100ms	$(30\text{ms})^2$	1	1	1	1	1	1

m_i	σ_i^2	$\boldsymbol{s}_i$	$\boldsymbol{x}_i$	$\sigma_{\boldsymbol{s}_i}^2$	$\boldsymbol{h}_i$	σ_ϵ
100ms	$(30\text{ms})^2$	$\mathbf{0}$	$\emptyset$	$(0.15)^2$	$\boldsymbol{h}_i^{(0)}$	$\hat{\sigma}_\epsilon$

Fig. 3 shows an example of raw experimental signal and the decomposition of two adjacent segments with overlapped potentials using the proposed method. In this example, the fully automatic method proposed provided the same result as that obtained by the reference decomposition tool used by an expert operator.

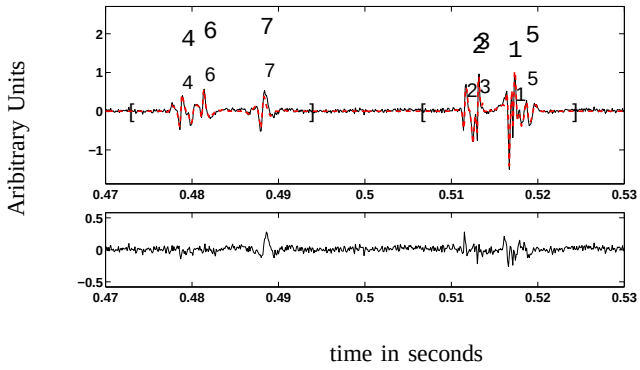

Fig. 3. Example of EMG decomposition of two adjacent segments using the proposed approrach in comparison with the reference decomposition. In these two segments a total of 7 sources are active. The numbers on top of the raw signal represent the source labels as identified by the automatic decomposition with the proposed method (top numbers) and the decomposition with interaction of an expert operator using EMGLAB (lower numbers). The labelling of sources in this example led to perfect matching with the reference result. The raw (solid line) and the reconstructed signal (dashed line) are shown in the upper panel whereas the residual error is plotted in the lower panel.

Fig. 4 shows an example of decomposition of simulated signals to better represent the ability of the method to identify complex superimpositions of action potentials. In this example, four overlapping action potentials are present in two segments. The shape of action potential $\#4$, which is isolated in the second segment, is fully cancelled by overlaps of other action potentials in the first segment. Similarly, action potential $\#1$ was relatively isolated in the first segment and cancelled in the second. Despite these complex overlaps, the automatic method was able to identify all action potentials in the two segments correctly. In this example, the reference decomposition result was provided by the simulation tool and is thus exact.

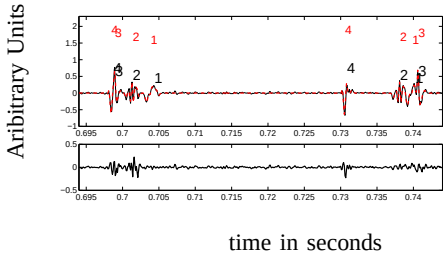

Fig. 4. Illustration of decomposition result of one segment containing up to four overlapped action potentials and the neighboring segment with three overlapped action potentials. Upper labels represent the source numbers identified by the proposed automatic method while lower labels those from the simulation tool. The residual error is also shown in the lower panel.

The decomposition results for the experimental signals are shown in Tab. V. # Sources and # Val are respectively the number of sources identified by EMGLAB (reference results) and the number of validated spike trains according to the *a-posteriori* criterion on the regularity. Because the main challenge in the decomposition is dealing with overlapped action potentials[2], especially when more than two action potentials are superimposed, in Tab.V we also report the percentage of action potentials that are overlapped with others. Therefore, in Tab. V, *% overlap* and *% overlap* ≥ 3 indicates the percentage of overlapped action potentials (over the total number of action potentials) and the percentage of action potentials in superposition with at least other two action potentials (superposition of at least 3 action potentials), respectively. The algorithm runtime complexity was measured by computational time in seconds per iteration for the decomposition of each second of the experimental signal $\boldsymbol{z}$. In order to decompose EMG signals of sufficient length ($\approx$ 20 s to get reliable statistics) and upon several experimental signals, the Markov chain was systematically run for 200 iterations. The estimation of the discharge patterns was then obtained by a majority vote on each temporal window (of length $= T_R$), whose advantages are discussed in [20].

The decomposition accuracy of the proposed method, as evaluated with the global performance index A, was improved (by approximately 3%) compared to that obtained by the joint maximization by Tabu search [1]

(see Tab. V). Most importantly, more spike trains were validated with the proposed method than with the Tabu search. For all tested methods in Tab. V, the detection errors (both false positives and false negatives) occurred less often in segments with single action potentials, while most errors were found in segments with overlaps, as it was expected. Therefore, the accuracy of the method on identifying overlapped action potentials was also computed. Over the entire number of overlapped action potentials, the accuracy was 89% and over all action potentials in superimpositions of at least three action potentials, the accuracy was 85%. Note that by using the MCMC method, the decomposition results would not vary for the same signal even if the latter is analyzed several times with different initialization values. This convergence property (independent of initialization values) derives directly from the convergence in the distribution sense of the MCMC method (see discussion in Section IV).

Tab. V also shows the decomposition results obtained by the proposed MCMC algorithm with the constant spike magnitude model for each source, *i.e.*, without including the variable amplitude model. The same expert results were used in each case and the "*#Sources*" row reports the number of detected sources. The decomposition performance degraded using this simplified model compared to the model with variable magnitudes: the latter have achieved better results in both extracting more valid trains and better accuracy calculated upon these valid trains.

TABLE V
COMPARISON OF DECOMPOSITION RESULTS ON THE 5 EXPERIMENTAL EMG SIGNALS (10%MVC, 20 S RECORDINGS).

EMG #	1	2	3	4	5
# Sources	5	5	8	8	4
% overlap	81	84	91	88	82
% overlap ≥ 3	63	61	74	69	65
	MCMC on extended model				
# Val	5	4	6	5	2
A	91.1%	90.6%	90.5%	89.3%	94.4%
Computing time (s.)	13.2	14.9	15.2	13.8	15.0
	MCMC on uniform magnitude model				
# Val	4	4	4	4	2
A	87.1%	90.4%	92.3%	89.2%	90.7%
	Optimization using Tabu search [1]				
# Val	3	4	4	4	2
A	89.4%	85.1%	87.7%	87.2%	92.5%

The background noise variance estimation $\widehat{\sigma}_\epsilon^2$ can also be considered a good indicator of decomposition quality, since it measures the difference between the observed signal and its reconstruction from the detected discharge patterns and the estimated spikes. Indeed, σ_ϵ^2 is sampled according to its conditional law $\text{IG}(\bar{\alpha}_s, \bar{\beta}_s)$, where

$$\overline{\alpha}_s = \alpha_s + \frac{N}{2}, \quad \overline{\beta}_s = \beta_s + \frac{1}{2}\Big\|\boldsymbol{z} - \sum_i \boldsymbol{s}_i * \boldsymbol{h}_i\Big\|^2$$

and the mean value decreases as $\|\boldsymbol{z} - \sum_i \boldsymbol{s}_i * \boldsymbol{h}_i\|^2$ decreases.

Tab. VI shows $\widehat{\sigma}_\epsilon^2$ from the preprocessing (shown as reference), after full decomposition with optimization by Tabu search [1] and with the proposed MCMC approach with and without spike magnitude variability. The lower value of residual noise level with the proposed approach (Tab. VI) is in agreement with the improvement of A by the MCMC method and the magnitude variability model that adjusts the energy for each detected spike. The lower error likely represents a better fit of the model w.r.t. the data rather than overfitting the noises since the reduction in residual error is also accompanied by superior accuracy, as shown in Tab. V.

TABLE VI
COMPARISON OF $\widehat{\sigma}_\epsilon^2$ BY OPTIMIZATION WITH TABU SEARCH AND BY MCMC ON THE 5 EXPERIMENTAL EMG SIGNALS (10% MVC, 20 S RECORDINGS).

EMG #	1	2	3	4	5
preprocessing	0.0012	0.0014	0.0012	0.0019	0.0011
Tabu search	0.0023	0.0029	0.0030	0.0034	0.0026
MCMC (full model)	0.0016	0.0027	0.0024	0.0030	0.0016
MCMC ($\boldsymbol{s}_i = \mathbf{1}_{\boldsymbol{x}_i}$)	0.0018	0.0027	0.0026	0.0033	0.0019

Finally, simulated signals with high variability in firing were used to evaluate the proposed decomposition method. One example of decomposition of simulated signals is presented in Fig. 4, as commented above. The average accuracy (over all motor units and over the 5 signals simulated) obtained for these signals was $83 \pm 5\%$. The accuracy calculated for overlapped action potentials in the simulations was $80 \pm 3\%$.

VI. DISCUSSION AND CONCLUSION

We have proposed a novel method for fully automatic spike sorting in multiunit recordings. The main difficulty in the problem of spike sorting is the resolution of superimposed spikes. The isolated spikes can indeed be detected and clustered with robust techniques, that have high performance. On the contrary, the global optimization of overlapped action potentials is a non-deterministic polynomial-type (NP) hard problem, and thus cannot be solved by polynomial complexity algorithms. Therefore, methods for full spike sorting differ mainly for their performance when separating overlapped action potentials. We have previously proposed a method for the solution of overlapped spikes in neural or intramuscular recordings based on a Bayesian model coupled

with a MAP estimator and the Tabu search technique [1]. In the present study, we have expanded the signal model to allow for a variability in the spike magnitudes in the same source. Moreover, we have proposed a new method for the solution of the spike sorting problem based on MCMC simulation. The convergence property of this method was proven theoretically (Appendix), whereas the Tabu search algorithm described in [1] is heuristic. The performance of this approach is further improved by including variable magnitudes in the spike model.

The model is based on the assumption of a refractory period in neural firing, which implies to set a minimum time lapse between two consecutive discharges of the same source. Further, the model also favors regular firings with respect to irregular firings, but without constraining the solution space to regular firings. Accordingly, the method was tested on simulated signals with very large firing variability and proved to be relatively accurate also in those conditions. As it was expected, the accuracy was lower in case of irregular than for regular firings but the method could still be successfully applied in a fully automatic way with accuracy greater than 80% for an ISI variability of 60% (simulation results). It has to be noted that the imposed variability in the simulations was so large that the spike train regularity could not be used as a relevant information for reducing the solution space in the automatic method. The model is also based on the assumption of independence of the spike trains. We note however that this assumption is not always met since neural cells often fire more synchronously than by mere chance. However, the spike train independence assumption should not be viewed as a limitation of the approach since it does not imply that dependent spike trains cannot be identified, similarly to the discussion above for regular firings. The assumption only implies that the potential correlation between firing patterns is not taken into account in the prior laws, thus the correlation property is simply not exploited by the model. Accordingly, the method was proved successful in decomposing motor unit activities in a hand muscle, characterized by a high degree of motor unit synchronization [32].

A limitation of the method is that, although it includes changes in amplitude, it does not track changes in the shapes of the action potentials, which often occur over time. Some previous methods (*e.g.,* [33], [34]) have included this possibility, at least for slow changes over time. Inclusion of slow changes in shape in the current approach could be incorporated in the first pre-processing phase in which the individual motor unit action potential shapes are estimated from segments containing only one potential. However, this extension has not been tested in the current study.

In conclusion, a new fully automatic method for spike sorting of multiunit recordings has been derived and compared to a previously proposed solution on experimental intramuscular recordings. The method was proven to provide good accuracy in signal decomposition with relatively limited computational time even in the case of several spikes overlapped in time.

APPENDIX

Proof: To show the irreducibility of the chain in Tab. I, it is sufficient to show that for all $\boldsymbol{u} \in \mathcal{C}_k$, such that $P(\boldsymbol{u}) > 0$, the probability of a particular trajectory composed of $\{\boldsymbol{u}^{(0)} = \boldsymbol{0}, \ldots, \boldsymbol{u}^{(i)}, \ldots \boldsymbol{u}^{(L)} = \boldsymbol{u}\}$, where $L = |\boldsymbol{u}|_1$ and $|\boldsymbol{u}^{(i)} - \boldsymbol{u}^{(i-1)}|_1 = 1$, is non-zero.

Such a chain is constructed by adding a spike of a certain source in each iteration. It is then sufficient to show that for all i, $K(\boldsymbol{u}^{(i-1)} \mapsto \boldsymbol{u}^{(i)}) > 0$, or equivalently,

$$q\big(\boldsymbol{u}^{(i-1)} \mapsto \boldsymbol{u}^{(i)}\big)\, \rho(\boldsymbol{u}^{(i-1)} \mapsto \boldsymbol{u}^{(i)}\big) > 0. \qquad \text{(A-1)}$$

It is evident that $\boldsymbol{u}^{(i)} \in \omega(\boldsymbol{u}^{(i-1)})$, and it follows that:

$$\begin{aligned} q\big(\boldsymbol{u}^{(i-1)} \mapsto \boldsymbol{u}^{(i)}\big)\, &\rho(\boldsymbol{u}^{(i-1)} \mapsto \boldsymbol{u}^{(i)}\big) \\ &= \frac{P(\boldsymbol{u}^{(i)})}{F(\boldsymbol{u}^{(i-1)})} \min\left\{1, \frac{F(\boldsymbol{u}^{(i-1)})}{F(\boldsymbol{u}^{(i)})}\right\} \\ &= \min\left\{\frac{P(\boldsymbol{u}^{(i)})}{F(\boldsymbol{u}^{(i-1)})}, \frac{P(\boldsymbol{u}^{(i)})}{F(\boldsymbol{u}^{(i)})}\right\} \end{aligned}$$

The inequality (A-1) is thus guaranteed by the fact that $P(\boldsymbol{u}^{(i)}) > 0$ for all i. Since $P(\boldsymbol{u}) > 0$ implies that $P(\boldsymbol{u}^{(i)}) > 0$ for all i ($\boldsymbol{u} \in \mathcal{C}_k \Rightarrow \boldsymbol{u}^{(i)} \in \mathcal{C}_k$ for all i). ■

REFERENCES

[1] D. Ge, E. Le Carpentier, and D. Farina, "Unsupervised Bayesian Decomposition of Multi-Unit EMG Recordings using Tabu Search", *IEEE Trans. Biomed Eng.*, vol. 57, no. 3, pp. 561–517, mar 2010.

[2] M. S. Lewicki, "A review of methods for spike sorting: the detection and classification of neural action potentials", *Network: Computation in Neural Systems*, vol. 9, no. 4, pp. R53–R78, 1998.

[3] R. Quian Quiroga, Z. Nadasdy, and Y. Ben-Shaul, "Unsupervised spike detection and sorting with wavelets and superparamagnetic clustering", *Neural Computation*, vol. 16, pp. 1661–1687, 2004.

[4] M. S. Lewicki, "Bayesian modeling and classification of neural signals", *Neural Computation*, vol. 6, pp. 1005–1030, 1994.

[5] I. Bankman, K. Johnson, and W. Schneider, "Optimal detection, classification, and superposition resolution in neural waveform recordings", *IEEE Trans. Biomed Eng.*, vol. 8, no. 40, pp. 836–841, Aug. 1993.

[6] R. Chandra and L. Optican, "Detection, classification, and superposition resolution of action potentials in multiunit single-channel recordings by an on-line real-time neural network", *IEEE Trans. Biomed Eng.*, vol. 44, pp. 403–412, Aug. 1997.

[7] F. A. Atiya, "Recognition of multiunit neural signals", *IEEE Trans. Biomed Eng.*, vol. 39, pp. 723–729, Jul 1992.

[8] K. C. McGill, "Optimal resolution of superimposed action potentials", *IEEE Trans. Biomed Eng.*, vol. 49, no. 7, pp. 640–650, Jul 2002.

[9] K. C. McGill, Z. C. Lateva, and H. R. Marateb, "EMGLAB: An interactive EMG decomposition program", *J. Neurosci. Methods*, vol. 149, no. 2, pp. 121–133, Dec 2005.

[10] S. Takahashi, Y. Anzai, and Y. Sakurai, "Automatic sorting for multineuronal activity recorded with tetrodes in the presence of overlapping spikes", *J. Neurophysiol*, vol. 89, pp. 2245–2258, 2003.

[11] S. Takahashi and Y. Sakurai, "Real-time and automatic sorting of multi-neuronal activity for sub-millisecond interactions in vivo", *Neuroscience*, vol. 134, no. 1, pp. 301 – 315, 2005.

[12] S. Takahashi, Y. Sakurai, M. Tsukada, and Y. Anzai, "Classification of neuronal activities from tetrode recordings using independent component analysis", *Neurocomputing*, vol. 49, pp. 289–298(10), 2002.

[13] A. Holobar and D. Zazula, "Correlation-based decomposition of surface electromyograms at low contraction forces", *Med. biol. eng. comput.*, vol. 42, no. 4, pp. 487–495, 2004.

[14] A. Holobar, D. Farina, M. Gazzonib, R. Merlettib, and D. Zazula, "Estimating motor unit discharge patterns from high-density surface electromyogram", *Clin neurophysiol.*, vol. 120, pp. 551–562, 2009.

[15] D. Farina and D. Falla, "Effect of muscle-fiber velocity recovery function on motor unit action potential properties in voluntary contractions.", *Muscle and nerve*, vol. 37, no. 5, pp. 650–658, May 2008.

[16] J. J. Kormylo and J. M. Mendel, "Maximum-likelihood seismic deconvolution", *IEEE Trans. on Geoscience and Remote sensing*, vol. GE-21, no. 1, pp. 72–82, Jan. 1983.

[17] Q. Cheng, R. Chen, and T.-H. Li, "Simultaneous wavelet estimation and deconvolution of reflection seismic signals", *IEEE Trans. on Geoscience and Remote sensing*, vol. 34, pp. 377–384, Mar. 1996.

[18] F. Champagnat, Y. Goussard, and J. Idier, "Unsupervised deconvolution of sparse spike trains using stochastic approximation", *IEEE Trans. on Signal Processing*, vol. 44, no. 12, pp. 2988–2998, Dec. 1996.

[19] S. Bourguignon and H. Carfantan, "Bernoulli-Gaussian spectral analysis of unevenly spaced astrophysical data", in *IEEE Workshop Stat. Sig. Proc.*, Bordeaux, France, July 2005, pp. 811–816.

[20] G. Kail, C. Novak, B. Hofer, and F. Hlawatsch, "A blind Monte Carlo detection-estimation method for optical coherence tomography", in *ICASSP*, Taipei, Taiwan, R.O.C., Apr. 2009, pp. 493–496.

[21] K. C. McGill, K. L. Cummins, and L. J. Dorfman, "Automatic decomposition of the clinical electromyogram", *IEEE Trans. Biomed Eng.*, vol. BME-32, no. 7, pp. 470–477, Jul 1985.

[22] K. C. McGill and L. J. Dorfman, "High-resolution alignment of sampled waveforms", *IEEE Trans. Biomed Eng.*, vol. BME-31, no. 6, pp. 462–468, Jun 1984.

[23] F. Wood and M. J. Black, "A nonparametric bayesian alternative to spike sorting", *J. Neurosci. Methods*, vol. 173, no. 1, pp. 1 – 12, 2008.

[24] D. Ge, *Déconvolution impulsionnelle multi-source. Application aux signaux électromyographiques*, PhD thesis, Ecole centrale de Nantes, 1 rue de la Noë, December 2009.

[25] P. J. Green, "Reversible jump MCMC computation and Bayesian model determinaion", *Biometrika*, vol. 82, pp. 711–732, 1995.

[26] C. P. Robert and G. Casella, *Monte Carlo Statistical Methods*, Springer Texts in Statistics. Springer Verlag, New York, NY, USA, deuxieme edition, 2004.

[27] L. Tierney, "Markov Chains for Exploring Posterior Distributions", *Ann. Statist.*, vol. 22, no. 4, pp. 1701–1728, 1994.

[28] D. Farina, A. Crosetti, and R. Merletti, "A model for the generation of synthetic intramuscular EMG signals to test decomposition algorithms", *IEEE Trans. Biomed. Eng.*, vol. 48, pp. 66–77, 2001.

[29] C. J. D. Luca, A. Adam, R. Wotiz, L. D. Gilmore, and S. H. Nawab, "Decomposition of Surface EMG Signals", *J. Neurophysiol.*, vol. 96, pp. 1646–1657, May 2006.

[30] C. T. Moritz, B. K. Barry, M. A. Pascoe, and R. M. Enoka, "Discharge Rate Variability Influences the Variation in Force Fluctuations Across the Working Range of a Hand Muscle", *J. Neurophysiol.*, vol. 93, no. 5, pp. 2449–2459, may 2005.

[31] A. J. Fuglevand, D. A. Winter, and A. E. Patla, "Models of recruitment and rate coding organization in motor-unit pools", *J. Neurophysiol.*, vol. 70, no. 6, pp. 2470–2488, dec 1993.

[32] J. L. Dideriksen, D. Falla, M. Baekgaard, M. L. Mogensen, K. L. Steimle, and D. Farina, "Comparison between the degree of motor unit short-term synchronization and recurrence quantification analysis of the surface EMG in two human muscles", *Clin Neurophysiol.*, vol. 120, no. 12, pp. 2086–2092, may 2009.

[33] R. LeFever and C. De Luca, "A procedure for decomposing the myoelectric signal into its constituent action potentials-part I: technique, theory, and implementation", *IEEE Trans. Biomed. Eng.*, vol. 29, pp. 149–157, 1982.

[34] R. M. Studer, R. J. P. De Figueiredo, and G. S. Moschytz, "An Algorithm for Sequential Signal Estimation and System Identification for EMG Signals", *IEEE Trans. Biomed. Eng.*, vol. 31, no. 3, pp. 285–295, 1984.

Di Ge received the B.S. degree in telecommunication from Jiaotong University, Shanghai, China, in 2003, the Diplôme d'Ingénieur from Ecole Centrale de Nantes, France, in 2006, and the Ph.D. degree in signal processing at the Institut de Recherche en Communications et Cybernétique de Nantes, UMR-CRNS 6597, France in 2009. He is currently a research consultant in Glaizer group. His research interests include Bayesian statistical inference and stochastic processes in solving inverse problems.

Eric Le Carpentier received the Ph.D. degree in automatic control from the Nantes University, Nantes, France, in 1990. Since 1992, he has been an Assistant Professor at Ecole Centrale de Nantes, Nantes. His current research interests include signal processing, stochastic simulation for estimation, detection, tracking, and control, with applications in biomedical signals, robotics, and precise geo-location.

Jérôme Idier was born in France in 1966. He received the diploma degree in electrical engineering from École Supérieure d'Électricité, Gif-sur-Yvette, France, in 1988 and the Ph.D. degree in physics from University of Paris-Sud, Orsay, France, in 1991.
Since 1991, he joined the Centre National de la Recherche Scientifique. He is currently a Senior Researcher at the Institut de Recherche en Communications et Cybernétique in Nantes. His major scientific interests are in probabilistic approaches to inverse problems for signal and image processing. He is serving as an Associate Editor for the IEEE Transactions on Signal Processing and for the Journal of Electronic Imaging, copublished by SPIE and IS&T.

Dario Farina (M'01-SM'09) obtained the MSc degree in Electronics Engineering from Politecnico di Torino, Torino, Italy, in 1998, and the PhD degrees in Automatic Control and Computer Science and in Electronics and Communications Engineering from the Ecole Centrale de Nantes, Nantes, France, and Politecnico di Torino, respectively, in 2002. In 2002-2004 he has been Research Assistant Professor at Politecnico di Torino and in 2004-2008 Associate Professor in Biomedical Engineering at Aalborg University, Aalborg, Denmark. From 2008 to 2010 he has been Full Professor in Motor Control and Biomedical Signal Processing and Head of the Research Group on Neural Engineering and Neurophysiology of Movement at Aalborg University. In 2010 he has been appointed Full Professor and Founding Chair of the Department of Neurorehabilitation Engineering at the University Medical Center Göttingen, Georg-August University, Germany, within the Bernstein Center for Computational Neuroscience. Since 2010 he is the Vice-President of the International Society of Electrophysiology and Kinesiology (ISEK) and a Fellow of the American Institute for Medical and Biological Engineering. He is the recipient of the 2010 IEEE Engineering in Medicine and Biology Society Early Career Achievement Award. He is an Associate Editor of Medical & Biological Engineering & Computing and member of the Editorial Boards of the Journal of Electromyography and Kinesiology and of the Journal of Neuroscience Methods. His research focuses on biomedical signal processing and modeling, neurorehabilitation, and neural control of movement. Within these areas, he has (co)-authored approximately 200 papers in peer-reviewed Journals and about 300 among conference papers/abstracts, book chapters and encyclopedia contributions. He is an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL ENGINEERING.