A Bayesian Nonlinear Source Separation Method for Smart Ion-Selective Electrode Arrays

Leonardo Tomazeli Duarte, Christian Jutten, Saïd Moussaoui

▶ To cite this version:

HAL Id: hal-00429900

https://hal.archives-ouvertes.fr/hal-00429900

Submitted on 5 Nov 2009

A Bayesian Nonlinear Source Separation Method for Smart Ion-selective Electrode Arrays

Leonardo T. Duarte, *Student Member, IEEE*, Christian Jutten, *Fellow, IEEE*, and Saïd Moussaoui

Abstract—Potentiometry with ion-selective electrodes (ISEs) provides a simple and cheap approach for estimating ionic activities. However, a well-known shortcoming of ISEs regards their lack of selectivity. Recent works have suggested that smart sensor arrays equipped with a blind source separation (BSS) algorithm offer a promising solution to the interference problem. In fact, the use of blind methods eases the time-demanding calibration stages needed in the typical approaches. In this work, we develop a Bayesian source separation method for processing the outputs of an ISE array. The major benefit brought by the Bayesian framework is the possibility of taking into account some prior information, which can result in more realistic solutions. Concerning the inference stage, it is conducted by means of Markov chain Monte Carlo (MCMC) methods. The validity of our approach is supported by experiments with artificial data and also in a scenario with real data.

Index Terms—Blind source separation, ion-selective electrode, chemical sensor array, Bayesian approach.

I. INTRODUCTION

ION-selective electrodes (ISEs) [2], [3], [4] are devices used for estimating the ionic activity, a measure of effective concentration of an ion in aqueous solution. The transducer mechanism in an ISE is based on a sensitive membrane in which the electrochemical potential depends on the concentration of a given charged chemical specie. A typical example of ISE is the glass electrode used for measuring the pH value. Moreover, ISEs for ions such as ammonium, potassium, and calcium are of great interest in applications like water quality control [5] and biomedical monitoring [6], [7].

The most relevant drawback of an ISE concerns its lack of selectivity, that is, an ISE may respond to interfering ions other than the target one. While in some situations the interference process can be negligible, there are other cases in which this problem seems to be more accentuated, especially if the ions under analysis have similar physicochemical properties. If no care is taken in these cases, the measurements provided by the ISE may become uncertain.

A possible way to overcome the interference problem relies on the concept of smart sensor array (SSA). In a SSA, the data acquisition is done through several ISEs. Then, the desired information is extracted by a data processing block that tries to exploit the diversity provided by the array. Typically, the development of a SSA follows a supervised paradigm, i.e. a set of training (or calibration) samples is used to adjust the parameters of the SSA data processing block. Despite the good results obtained in both quantitative and qualitative analysis (for example, see the electronic tongues [8]), supervised methods suffers from two practical problems: 1) the acquisition of a sufficient number of training samples may be time-demanding, and 2) the calibration procedure must be performed from time to time due to the sensor's drift. These limitations are the main motivations for the use of unsupervised (or blind) methods. In this new case, the SSA data processing block only uses the array responses and some parametric knowledge on the interference process. Therefore, the calibration stage may be skipped or, at least, quite reduced.

In this work, we are interested in the problem of unsupervised quantitative analysis via an ISE array, which can be expressed as follows: there is a set of source signals, each one representing the activity of a given ion. Due to the interference phenomena in the transduction stage, the signals recorded by the ISE array are mixed versions of the sources. Then, our goal is to retrieve the sources using only the ISE array outputs. This description is an example of a Blind Source Separation problem (BSS) [9], [10], to which the signal processing community has been devoting a great deal of attention. The main difficulty in the application of BSS methods to the problem treated in this work stems from the nonlinear behavior of an ISE. While linear BSS is now a well established matter, the more general problem of nonlinear BSS is still a very challenging one (see [11]).

The first studies [12], [13], [14] on the design of BSS methods for chemical sensor arrays were based on Independent Component Analysis (ICA) methods [9]. Despite the encouraging results presented in these works, the developed methods have several limitations. Firstly, they do not take into account some prior information that would be useful for improving the separation quality. For instance, the sources in our problem are always non-negative, since they represent ionic activities. Secondly, given that ICA methods works with the assumption that the sources are statistically independent, they cannot be used when the sources are correlated. Finally, all these works were evaluated with synthetic data in noiseless situations.

In this paper, we propose a Bayesian source separation method for processing the outputs of an ISE array. The Bayesian approach has been proved to be attractive in several domains including audio source separation [15] and spectroscopy data analysis [16]. Moreover, one of the main advantages of this approach is exactly the possibility of incorporating prior information into our method and of working in noisy scenarios. Finally, the advent of sampling techniques [17] such

Leonardo T. Duarte and C. Jutten are with the GIPSA-lab (UMR CNRS 5216), Institut Polytechnique de Grenoble, Grenoble, France (e-mail: leonardo.duarte@gipsa-lab.inpg.fr, christian.jutten@gipsa-lab.inpg.fr). C. Jutten is also with Institut Universitaire de France.

Saïd Moussaoui is with IRCCyN (UMR CNRS 6597), Ecole Centrale Nantes, Nantes, France (e-mail:said.moussaoui@irccyn.ec-nantes.fr).

as Monte Carlo Markov Chain (MCMC) methods has provided an efficient solution to the calculation of the complex integrals that are usually associated with a Bayesian framework.

This work is an extended version of the conference paper [1], which only contained the basics of our method and some preliminary results. In the present version, we detail important aspects of our method and widely expand the results section, including a comparison with a supervised technique. Concerning the paper's organization, we start with a description of the mixing model associated with the application in mind. After that, in Section III, we describe the building blocks of our Bayesian source separation scheme. In Section IV, we conduct some experiments with artificial and real data to assess the performance of our proposal. Finally, our conclusions are presented in Section V.

II. PROBLEM STATEMENT

Let us consider a solution containing n_s ions and an array composed of n_c electrodes. Due to the interference problem, the response of the i-th ISE within the array at the instant t, represented by y_{it}, is dependent not only on the activity of its target ion, s_{it}, but also on the activities of the other ions in the solution, s_{jt}. According to the Nicolsky-Eisenman (NE) equation [3], this phenomenon can be modeled as[1]

$$y_{it} = e_i + d_i \log \left(s_{it} + \sum_{j=1, j \neq i}^{n_s} a_{ij} s_{jt}^{z_i/z_j} \right), \qquad (1)$$

where e_i is a constant, z_j is the valence of the j-th ion and a_{ij} denotes the selectivity coefficients. Finally, $d_i = RT/z_iF$, where R is the gas constant, T the temperature in Kelvin, and F the Faraday constant. When the valences of the ions under analysis are the same, the model (1) becomes a particular case of the class of post-nonlinear (PNL) models [13]. Nonetheless, when the valences are different and the ISEs within the array have different ions as targets, the resulting model becomes tougher due to the appearance of a nonlinearity inside the logarithm term [14].

By arranging the response of each ISE within the array, the mixing process can be expressed as follows:

$$\mathbf{Y} = \mathbf{e} \cdot \mathbf{1}_{1 \times n_d} + \text{diag}(\mathbf{d}) \log \left(\mathbf{A} \otimes^{\mathbf{z}} \mathbf{S} \right), \qquad (2)$$

where n_d corresponds to the number of available samples, $\mathbf{Y} \in \mathbb{R}^{n_c \times n_d}$, $\mathbf{e} = [e_1, \ldots, e_{n_c}]^T$, $\mathbf{d} = [d_1, \ldots, d_{n_c}]^T$, $\mathbf{A} \in \mathbb{R}_+^{n_c \times n_s}$ and $\mathbf{S} \in \mathbb{R}_+^{n_s \times n_d}$. The element it of $\mathbf{Y}$ corresponds to y_{it} and therefore the i-th row of $\mathbf{Y}$ corresponds to the time response of the i-th ISE within the array. Analogously, the element jt of $\mathbf{S}$ denotes the activity of the j-th ion at the instant t, i.e. s_{jt}. Matrix $\mathbf{A}$ contains the selectivity coefficients. The vector of valences is denoted by $\mathbf{z} = [z_1, \ldots, z_{n_s}]^T$, and the operator $\otimes^{\mathbf{z}}$ describes the nonlinear transformation inside the logarithm function present in the NE model (see Eq. (1)). If the valences z_i are equal, then $\otimes^{\mathbf{z}}$ results in a simple matrix multiplication. Finally, $\mathbf{1}_{1 \times n_d}$ corresponds to a vector (dimension n_d) with all elements equal to one.

In view of a possible model inaccuracy and/or of the errors introduced by the measurement system, a more realistic description for the array outputs is given by:

$$\mathbf{X} = \mathbf{Y} + \mathbf{N}, \qquad (3)$$

where $\mathbf{N} \in \mathbb{R}^{n_c \times n_d}$ represents the noise terms. We assume a zero mean additive white Gaussian noise (AWGN) with covariance matrix $\mathbf{C_n} = \text{diag}(\boldsymbol{\sigma}^2)$, where the elements of the vector $\boldsymbol{\sigma}^2$, the variances $\sigma_1^2, \ldots, \sigma_{n_c}^2$, are unknown.

With equations (2) and (3) in mind, we may formulate the source separation problem treated in this work: given the array response $\mathbf{X}$ and assuming that the valences z_i are known, our goal is to estimate the elements of $\mathbf{S}$ (ionic activities). Since we envisage a blind method, the other parameters related to the mixing model (except $\mathbf{z}$) and the noise variance at each electrode are also unknown and, thus, should be estimated. Furthermore, as it will become clear later, there are other unknown parameters, denoted by $\boldsymbol{\phi}$, which are related to the prior distributions assigned to the sources. Henceforth, all these unknown parameters will be represented by the vector $\boldsymbol{\theta} = [\mathbf{S}, \mathbf{A}, \mathbf{d}, \mathbf{e}, \boldsymbol{\sigma}, \boldsymbol{\phi}]$ and the following notation will be adopted: $\boldsymbol{\theta}_{-\theta_q}$ represents the vector containing all elements of $\boldsymbol{\theta}$ except θ_q.

III. BAYESIAN SOURCE SEPARATION METHOD

The first step in Bayesian estimation is to assign an *a priori* probability distribution function, or simply prior, for each unknown parameter. This is done based on some available knowledge or constraint related to the unknown parameters; in Section III-A, we will discuss how this step can be accomplished in our problem. Later, in Section III-B, we will present the likelihood function $p(\mathbf{X}|\boldsymbol{\theta})$ resulting from the mixing model (Eqs. (2) and (3)). Then, in Section III-C, we shall derive the posterior distribution, which can be seen as an update of the priors using the information brought by the likelihood function. According to the Bayes' rule, the posterior is given by $p(\boldsymbol{\theta}|\mathbf{X}) = p(\mathbf{X}|\boldsymbol{\theta})p(\boldsymbol{\theta})/p(\mathbf{X})$. Given that $p(\mathbf{X})$ is not a function of the unknown parameters $\boldsymbol{\theta}$, we are mainly interested in the following expression

$$p(\boldsymbol{\theta}|\mathbf{X}) \propto p(\mathbf{X}|\boldsymbol{\theta}) \, p(\boldsymbol{\theta}). \qquad (4)$$

Finally, in Section III-D, we shall discuss how an inference scheme based on (4) can be set to resolve our nonlinear source separation problem.

A. Defining the prior distributions

1) Prior distribution of the sources $\mathbf{S}$: We adopted log-normal distributions to model the sources. This distribution is always non-negative which seems natural in our problem as the sources represent ionic activities. Moreover, there are two other aspects that motivate our choice. Firstly, the estimation of the parameters of a log-normal distribution is not complicated. As it will be checked later, it is possible to define conjugate priors in this case[2]. Also, there is a practical argument behind

---

[1]In this work, log and ln stands for the logarithm of base 10 and the natural logarithm, respectively.

[2]From the Bayes' rule $p(X/Y) \propto p(Y/X)p(X)$. If $p(X/Y)$ and $p(X)$ belong to the same family, then $p(X)$ is said to be conjugate with respect to the likelihood $p(Y/X)$. When the prior is conjugate, the Bayesian inference only needs the update of the posterior distribution (see [18], for instance).

this choice. Ionic activities are expected to have a small variation in the logarithmic scale. This can be taken into account by the log-normal distribution, since such a distribution is nothing but a Gaussian distribution in the logarithmic scale. In mathematical terms, the prior assigned to s_{jt} is given by

$$p(s_{jt}) \propto \frac{1}{s_{jt}} \exp\left(-\frac{(\ln(s_{jt}) - \mu_j)^2}{2\sigma_j^2}\right) \mathbb{1}_{[0,+\infty[}(s_{jt}), \quad (5)$$

where μ_j, σ_j are the distribution parameters and $\mathbb{1}_{[0,+\infty[}(s_{jt})$ is the indicator function. We assume that the samples are independent and identically-distributed (i.i.d) and also that the sources are mutually statistically independent[3], that is

$$p(\mathbf{S}) = \prod_{j=1}^{n_s} \prod_{t=1}^{n_d} p(s_{jt}).$$

2) Prior distribution of the sources hyperparameters ϕ: The parameters $\phi_j = [\mu_j \ \sigma_j]$ in Eq. (5) are also unknown in our problem and, thus, priors must be defined for them. We adopted a Gaussian prior distribution for μ_j, that is:

$$p(\mu_j) = \frac{1}{\sqrt{2\pi\sigma_{\mu_j}^2}} \exp\left(-\frac{(\mu_j - \mu_{\mu_j})^2}{2\sigma_{\mu_j}^2}\right), \quad (6)$$

where μ_{μ_j} and σ_{μ_j} correspond to the hyperparameters. For $r_j = 1/\sigma_j^2$, a Gamma distribution is considered:

$$p(r_j) = \frac{r_j^{\alpha_{r_j}-1}}{\Gamma(\alpha_{r_j})\beta_{r_j}^{\alpha_{r_j}}} \exp\left(\frac{-r_j}{\beta_{r_j}}\right) \mathbb{1}_{[0,+\infty[}(r_j), \quad (7)$$

where α_{r_j} and β_{r_j} are the hyperparameters. In Section III-D, we will show that these two priors lead to conjugate pairs in the estimation of μ_j and $r_j = 1/\sigma_j^2$.

3) Prior distribution of the selectivity coefficients $\mathbf{A}$: According to the literature in potentiometric sensors (see [3], [19]), the selectivity coefficients are also non-negative. Moreover, it is rare to find a sensor whose response depends more on the interfering ion than on the target one, that is, a_{ij} usually lies in the interval $[0, A_{max}]$ where $A_{max} \in [0, 1]$. Thus, an uniform distribution can be assumed for each[4] a_{ij}, that is

$$p(a_{ij}) \propto \mathbb{1}_{[0,A_{max}]}(a_{ij}). \quad (8)$$

If no additional information is available, one can set $A_{max} = 1$. However, it is possible to refine this information by using databases such as [19] or by incorporating some information acquired during the fabrication process. Finally, the coefficients of $\mathbf{A}$ are assumed mutually independent, i.e.

$$p(\mathbf{A}) = \prod_{i=1}^{n_c} \prod_{j=1}^{n_s} p(a_{ij}).$$

[3]If some information concerning a possible dependency between the sources is available, it can be used for defining the priors. In this case, however, one goes toward a less general approach that is useful only to the modeled situation. Besides, the inference problem becomes more difficult in this case.

[4]In view of Eq. (1), we consider $a_{ij} = 1$ when $i = j$.

4) Prior distribution of the Nernstian slopes $\mathbf{d}$: As already discussed, d_i is related to physical parameters. For a room temperature, it takes approximately $0.059/z_i$, and the electrodes with such sensibility are said to have a Nernstian response. However, due to the sensor fabrication process and aging, a deviation from this theoretical value is usually observed. Furthermore, even the injection scheme can play a role on the value of this parameter [20]. This possible deviation from the Nernstian value can be taken into account by setting a Gaussian prior of mean $\mu_{d_i} = 0.059/z_i$V, i.e.:

$$p(d_i) \propto \exp\left(-\frac{(d_i - \mu_{d_i})^2}{2\sigma_{d_i}^2}\right). \quad (9)$$

$\sigma_{d_i}^2$ must be high enough to correctly model the derivations from the theoretical value. For example, for an electrode whose target ion is monovalent, one obtains a coherent distribution by setting $\sigma_{d_i} = 0.01$. Again, this knowledge could be refined if additional information is available, or if the measurements are conducted in different temperatures. Finally, we assume that the elements of $\mathbf{d}$ are statistically independent.

5) Prior distribution of $\mathbf{e}$: In contrast to the parameters d_i, there is no theoretical value for e_i. Still, as can be observed in [2], [21], e_i usually lies in the interval $[0.05, 0.35]$V. Hence, the following Gaussian prior may be adopted

$$p(e_i) \propto \exp\left(-\frac{(e_i - \mu_{e_i})^2}{2\sigma_{e_i}^2}\right), \quad (10)$$

where the mean is given by $\mu_{e_i} = 0.20$V and $\sigma_{e_i}^2$ must be defined so (10) goes toward a flat prior in the interval $[0.05, 0.35]$. The elements of $\mathbf{e}$ are assumed mutually independent.

6) Prior distribution of the noise variance σ_i^2: A very common approach [16], [22] is to assign inverse Gamma priors for the noise variances; that is, $\gamma_i = 1/\sigma_i^2$ is modeled through a Gamma distribution with parameters α_{σ_i} and β_{σ_i}

$$p(\gamma_i) = \frac{\gamma_i^{\alpha_{\sigma_i}-1}}{\Gamma(\alpha_{\sigma_i})\beta_{\sigma_i}^{\alpha_{\sigma_i}}} \exp\left(\frac{-\gamma_i}{\beta_{\sigma_i}}\right) \mathbb{1}_{[0,+\infty[}(\gamma_i). \quad (11)$$

The motivation behind this modeling comes from the fact that it results in a conjugate pair. Moreover, it is possible to set α_{σ_i} and β_{σ_i} to obtain a non-informative prior [22].

B. Probabilistic modeling of the mixing process

Based on the mixing model defined in Section II and on the assumption of i.i.d. Gaussian noise which is also spatially uncorrelated, the likelihood is defined as

$$p(\mathbf{X}|\boldsymbol{\theta}) = \prod_{t=1}^{n_d} \prod_{i=1}^{n_c} \mathcal{N}_{x_{it}}\left(e_i + d_i \log\left(\sum_{j=1}^{n_s} a_{ij} s_{jt}^{z_i/z_j}\right), \sigma_i^2\right), \quad (12)$$

where $\mathcal{N}_{x_{ik}}(\mu, \sigma^2)$ denotes a Gaussian random variable in x_{ik} with mean μ and variance σ^2.

C. Bayesian inference

Having defined the priors $p(\boldsymbol{\theta})$ and the likelihood $p(\mathbf{X}|\boldsymbol{\theta})$, we may now apply the Bayes' rule to obtain the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$. Given that the unknown variables of our problem are mutually independent (except $\mathbf{S}$ and ϕ), the prior distribution factorizes and, as a consequence, the posterior distribution becomes

$$p(\boldsymbol{\theta}|\mathbf{X}) \propto p(\mathbf{X}|\boldsymbol{\theta})\,p(\mathbf{S}|\phi)\,p(\phi)\,p(\mathbf{A})\,p(\mathbf{e})\,p(\mathbf{d})\,p(\boldsymbol{\sigma}). \quad (13)$$

This *a posteriori* distribution can be used in the derivation of an inference scheme to estimate the parameters of interest. For instance, one may adopt the minimum mean square error (MMSE) estimator [23], which is given by

$$\boldsymbol{\theta}_{MMSE} = \int \boldsymbol{\theta}\, p(\boldsymbol{\theta}|\mathbf{X})d\boldsymbol{\theta}. \quad (14)$$

The exact implementation of this estimator is difficult because the analytical evaluation of this integral is hard in the problem considered here. Nonetheless, a good approximation of (14) can be provided by sampling methods. The idea here is to approximate the MMSE estimator using samples obtained from the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$. If, for instance, the generated samples are represented by $\boldsymbol{\theta}^{(1)}, \boldsymbol{\theta}^{(2)}, \ldots, \boldsymbol{\theta}^{(M)}$, then expression (14) can be approximated by

$$\widetilde{\boldsymbol{\theta}}_{MMSE} = \frac{1}{M} \sum_{m=1}^{M} \boldsymbol{\theta}^{(m)}. \quad (15)$$

According to the law of large numbers, $\widetilde{\boldsymbol{\theta}}_{MMSE} = \boldsymbol{\theta}_{MMSE}$ as $M \to +\infty$. This important result gives the theoretical foundation for the above-described methodology, which is referred as Monte Carlo integration [17].

D. Gibbs sampling scheme

The implementation of the MMSE estimator boils down to the task of generating samples from $p(\boldsymbol{\theta}|\mathbf{X})$. In this work, this is done by the Gibbs' sampler, a Markov Chain Monte Carlo (MCMC) method tailored for simulating joint distributions. The idea in MCMC methods is to generate a Markov chain that admits the desired distribution ($p(\boldsymbol{\theta}|\mathbf{X})$ in our case) as stationary distribution. Assuming that $x \sim p(x)$ stands for the sampling operation, i.e. x is a sample from the distribution $p(x)$, then the Gibbs' sampler can be summarized as follows:

1) Set initial samples $\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_N^{(0)}$;
2) For $m = 1$ to M, do

$$\begin{aligned}
\theta_1^{(m)} &\sim p(\theta_1|\theta_2^{(m-1)}, \theta_3^{(m-1)}, , \ldots, \theta_N^{(m-1)}, \mathbf{X}) \\
\theta_2^{(m)} &\sim p(\theta_2|\theta_1^{(m)}, \theta_3^{(m-1)}, \ldots, \theta_N^{(m-1)}, \mathbf{X}) \\
&\vdots \\
\theta_N^{(m)} &\sim p(\theta_N|\theta_1^{(m)}, \theta_2^{(m)}, \ldots, \theta_{N-1}^{(m)}, \mathbf{X})
\end{aligned}$$

end.

Note that the Gibbs' sampler simulates a high-dimensional joint distribution by sequentially sampling from the conditional distribution of each variable[5]. Therefore, aiming at the

implementation of the Gibbs sampler for our problem, we should obtain the conditional density for each element of $\boldsymbol{\theta}$. In the sequel, this will be accomplished by observing that

$$p(\theta_q|\boldsymbol{\theta}_{-\theta_q}, \mathbf{X}) \propto p(\mathbf{X}|\boldsymbol{\theta})p(\theta_q) \quad (16)$$

and by considering the likelihood function (12) and the prior distributions (defined in Section III-A).

1) Conditional distributions for the sources: Substituting expressions (12) and (5) into (16), one has

$$p(s_{jt}|\boldsymbol{\theta}_{-s_{jt}}, \mathbf{X}) \propto \exp\left[\sum_{i=1}^{n_c} -\frac{1}{2\sigma_i^2}\left(x_{it} - e_i \right.\right.$$
$$\left. - d_i \log\left(a_{ij}s_{jt}^{z_i/z_j} + \sum_{b=1, b \neq j}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)^2$$
$$\left. - \frac{(\ln(s_{jt}) - \mu_j)^2}{2\sigma_j^2}\right]\frac{1}{s_{jt}}\mathbb{1}_{[0,+\infty[}(s_{jt}). \quad (17)$$

2) Conditional distributions for the sources hyperparameters ϕ: Because $p(\mathbf{X}|\boldsymbol{\theta})$ is not a function of the parameters $\phi_j = [\mu_j\ \sigma_j]$, the conditional density of μ_j is given by

$$p\left(\mu_j|\sigma_j, \mathbf{S}_{(j,:)}\right) \propto p\left(\mathbf{S}_{(j,:)}|\mu_j, \sigma_j\right)p(\mu_j), \quad (18)$$

where $\mathbf{S}_{(j,:)}$ denotes all the elements of the j-th row. By substituting (5) and (6) into (18), we obtain:

$$p\left(\mu_j|\sigma_j, \mathbf{S}_{(j,:)}\right) \propto \prod_{t=1}^{n_d}\left[\exp\left(-\frac{(\ln(s_{jt}) - \mu_j)^2}{2\sigma_j^2}\right)\right]$$
$$\exp\left(-\frac{(\mu_j - \mu_{\mu_j})^2}{2\sigma_{\mu_j}^2}\right). \quad (19)$$

The first term of (19) can be rewritten as a Gaussian function in μ_j, with mean $\mu_{L_{\mu_j}} = 1/n_d \sum_{t=1}^{n_d} \ln(s_{jt})$ and variance $\sigma_{L_{\mu_j}}^2 = \sigma_j^2/n_d$. Thus, (19) becomes a product of two Gaussian distributions which is also a Gaussian distribution, i.e.:

$$p\left(\mu_j|\sigma_j, \mathbf{S}_{(j,:)}\right) \propto \exp\left(-\frac{\left(\mu_j - \mu_{Post_{\mu_j}}\right)^2}{2\sigma_{Post_{\mu_j}}^2}\right), \quad (20)$$

where[6] $\sigma_{Post_{\mu_j}} = \sigma_{L_{\mu_j}}^2 \sigma_{\mu_j}^2/(\sigma_{L_{\mu_j}}^2 + \sigma_{\mu_j}^2)$ and $\mu_{Post_{\mu_j}} = (\mu_{L_{\mu_j}}\sigma_{\mu_j}^2 + \mu_{\mu_j}\sigma_{L_{\mu_j}}^2)/(\sigma_{L_{\mu_j}}^2 + \sigma_{\mu_j}^2)$.

Similarly to the case of μ_j, the conditional density of $r_j = 1/\sigma_j^2$ is given by $p\left(r_j|\mu_j, \mathbf{S}_{(j,:)}\right) \propto p\left(\mathbf{S}_{(j,:)}|\mu_j, r_j\right)p(r_j)$. Thus, by considering (5) and (7), one has

$$p\left(r_j|\mu_j, \mathbf{S}_{(j,:)}\right) \propto \prod_{t=1}^{n_d}\left[\sqrt{r_j}\exp\left(-0.5r_j\left(\ln(s_{jt}) - \mu_j\right)^2\right)\right]$$
$$r_j^{\alpha_{r_j}-1}\exp\left(\frac{-r_j}{\beta_{r_j}}\right)\mathbb{1}_{[0,+\infty[}(r_j). \quad (21)$$

---

[5]Actually, the Gibbs' sampler requires only the conditional distributions up to a proportional gain.

[6]The derivation of the mean and variance of a product of two Gaussian can be found in [23].

We can rewrite $p(r_j|\mu_j, \mathbf{S}_{(j,:)})$ as a Gamma distribution, i.e.

$$p\left(r_j|\mu_j, \mathbf{S}_{(j,:)}\right) \propto r_j^{\alpha_{Post_{r_j}}} \exp\left(-\frac{r_j}{\beta_{Post_{r_j}}}\right) \mathbb{1}_{[0,+\infty[}(r_j),$$

(22)

where $\alpha_{Post_{r_j}} = \alpha_{r_j} + n_d/2$ and $\beta_{Post_{r_j}}^{-1} = \left(\sum_{t=1}^{n_d} (\ln(s_{jt}) - \mu_j)^2\right)/2\beta_{r_j}$ It is now clear that we obtain a conjugate pair and, thus, we can sample from $p\left(r_j|\mu_j, \mathbf{S}_{(j,:)}\right)$ by sampling from a Gamma distribution.

3) Conditional distribution of a_{ij}: The derivation of the conditional distribution of $p(a_{ij}|\boldsymbol{\theta}_{-a_{ij}}, \mathbf{X})$ is close to the one conducted for s_{jt}. Indeed, by considering expression (12), (16), and (8), it turns out that

$$p(a_{ij}|\boldsymbol{\theta}_{-a_{ij}}, \mathbf{X}) \propto \exp\left[-\frac{1}{2\sigma_i^2}\sum_{t=1}^{n_d}\left(x_{it} - e_i - d_i \log\left(a_{ij}s_{jt}^{z_i/z_j} + \sum_{b=1,b\neq j}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)^2\right]\mathbb{1}_{[0,1]}(a_{ij}).$$

(23)

4) Conditional distribution of d_i: Using Eqs. (12) and (9), and after some straightforward calculations, $p(d_i|\boldsymbol{\theta}_{-d_i}, \mathbf{x})$ can be written as the following Gaussian distribution

$$p(d_i|\boldsymbol{\theta}_{-d_i}, \mathbf{X}) \propto \exp\left(-\frac{\left(d_i - \mu_{Post_{d_i}}\right)^2}{2\sigma_{Post_{d_i}}^2}\right),$$

(24)

where $\sigma_{Post_{d_i}} = \sigma_{L_{d_i}}^2 \sigma_{\mu_{d_i}}^2/(\sigma_{L_{d_i}}^2 + \sigma_{\mu_{d_i}}^2)$, $\mu_{Post_{d_i}} = (\mu_{L_{d_i}}\sigma_{\mu_{d_i}}^2 + \mu_{\mu_{d_i}}\sigma_{L_{d_i}}^2)/(\sigma_{L_{d_i}}^2 + \sigma_{\mu_{d_i}}^2)$, and

$$\mu_{L_{d_i}} = \frac{\left(\sum_{t=1}^{n_d}(x_{it} - e_i)\right)\log\left(\sum_{b=1}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)}{\left(\log\left(\sum_{b=1}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)^2},$$ (25)

$$\sigma_{L_{d_i}}^2 = \frac{\sigma_{\mu_{d_i}}^2}{\left(\log\left(\sum_{b=1}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)^2}.$$

(26)

5) Conditional distribution of e_i: The development of this expression is similar to the one performed for d_i. After some calculations, it is not difficult to show that

$$p(e_i|\boldsymbol{\theta}_{-e_i}, \mathbf{X}) \propto \exp\left(-\frac{\left(e_i - \mu_{Post_{e_i}}\right)^2}{2\sigma_{Post_{e_i}}^2}\right),$$

(27)

where $\sigma_{Post_{e_i}} = \sigma_{L_{e_i}}^2 \sigma_{\mu_{e_i}}^2/(\sigma_{L_{e_i}}^2 + \sigma_{\mu_{e_i}}^2)$, $\mu_{Post_{e_i}} = (\mu_{L_{e_i}}\sigma_{\mu_{e_i}}^2 + \mu_{\mu_{e_i}}\sigma_{L_{e_i}}^2)/(\sigma_{L_{e_i}}^2 + \sigma_{\mu_{e_i}}^2)$, and

$$\mu_{L_{e_i}} = \frac{\sum_{t=1}^{n_d}\left(x_{it} - d_i \log\left(\sum_{b=1}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)}{n_d}$$ (28)

$$\sigma_{L_{e_i}}^2 = \sigma_{e_i}/n_d.$$

(29)

6) Conditional distribution of the noise variance σ_i: As discussed before, the attribution of a Gamma prior for $\gamma_i = 1/\sigma_i^2$ culminates in a conjugate pair. Indeed, by considering expressions (12) and (11), one can show that $p(\gamma_i|\boldsymbol{\theta}_{-\gamma_i}, \mathbf{X})$ can be reduced to the following Gamma distribution:

$$p(\gamma_i|\boldsymbol{\theta}_{-\gamma_i}, \mathbf{X}) \propto \gamma_i^{\alpha_{Post_{\sigma_i}}-1} \exp\left(\frac{-\gamma_i}{\beta_{Post_{\sigma_i}}}\right) \mathbb{1}_{[0,+\infty[}, \quad (30)$$

where $\alpha_{Post_{\sigma_i}} = \alpha_{\sigma_i} + n_d/2$ and $\beta_{Post_{\sigma_i}}^{-1} = \frac{1}{2}\sum_{t=1}^{n_d}\left(x_{it} - e_i - d_i \log\left(\sum_{b=1}^{n_s} a_{ib}s_{bt}^{z_i/z_b}\right)\right)^2 + \beta_{\sigma_i}^{-1}$.

E. Algorithm Description

Let us make some remarks on the final algorithm, which is summarized in Tab. I. In a first step, we must define hyperparameters that lead to non-informative distributions [18]. For instance, in the experiments described in Section IV, this strategy was implemented by setting the following values for the variances of the Gaussian priors: $\sigma_{d_i}^2 = 0.01$ and $\sigma_{e_i}^2 = 0.03$. Moreover, a high value of the parameter $\sigma_{\mu_j}^2$, which is related to the sources prior, was defined ($\sigma_{\mu_j}^2 = 100$).

Concerning the Gibbs' sampler, the resulting conditionals for almost all parameters are given by standard distributions (normal and Gamma) and, thus, sampling is straightforward in these cases. However, $p(s_{jt}|\boldsymbol{\theta}_{-s_{jt}}, \mathbf{X})$ and $p(a_{ij}|\boldsymbol{\theta}_{-a_{ij}}, \mathbf{X})$ result in non-standard distributions, thus requiring more sophisticated sampling methods. In this work, this task is accomplished by the Metropolis-Hasting (MH) algorithm [17]. Based on an instrumental distribution $g(x; y)$, the MH algorithm simulates a given distribution $p(x)$ in an iterative fashion; given $x^{(t)}$, the current sample of $p(x)$, the following iteration is done for obtaining the next sample $x^{(t+1)}$:

1) $x^* \sim g(x; x^{(t)})$ (generation of a candidate sample);
2) Calculate $a = \min\left(1, \frac{p(x^*)g(x^{(t)};x^*)}{p(x^{(t)})g(x^*;x^{(t)})}\right)$;
3) $u \leftarrow$ sample from a uniform distribution in $[0,1]$;
4) if $u \leq a$, $x^{(t+1)} = x^*$ (accept the proposed sample) else $x^{(t+1)} = x^{(t)}$ (reject the proposed sample).

In this work, truncated Gaussian are considered as instrumental distributions; in order to obtain good acceptance rate, we conducted, for each situation, preliminary simulations to adjust the variances of these distributions.

As the resulting Markov chain of the Gibbs' sampler takes some time to converge to its stationary distribution, the samples obtained in a first period (burn-in period) are not taken into account by the estimator (see Eq. (31)). A simple visual inspection of the evolution of some channel states was done for determining the burn-in period. However, more sophisticated strategies for convergence monitoring can be envisaged.

A last step in our algorithm concerns a post-processing stage for dealing with the scale and translation ambiguities inherent in blind methods [10]. For example, in PNL mixtures, the best we can do is to obtain an estimation $\tilde{s}_{jt}$ given by $\tilde{s}_{jt} = a_j s_{jt} + b_j$, where a_j and b_j are unknown, and s_{jt} is the actual source. Given that this ambiguity in not acceptable in a sensing problem, we are forced to use at least two calibration points in order to retrieve the scale and translation parameters. This can be done by a simple linear regression, i.e. we need to find a_j and b_j that minimize the mean squared error $J = 1/N_{cal}\sum_{n=1}^{N_{cal}}(\tilde{s}_{jn}^{(c)} - s_{jn}^{(c)})^2$, where $s_{jn}^{(c)}$ denotes the calibration points and $\tilde{s}_{jt}^{(c)}$ the corresponding estimations.

IV. EXPERIMENTAL RESULTS

To assess the proposed solution, we conducted in a first moment a set of experiments considering artificially generated

TABLE I
PROPOSED BAYESIAN SOURCE SEPARATION ALGORITHM

1) Define hyperparameters $\mu_{\mu_j}, \sigma^2_{\mu_j}, \alpha_{p_j}, \beta_{p_j}, \alpha_{\sigma_i}, \beta_{\sigma_i}$;
2) Random initialization of the current samples $\boldsymbol{\theta}^0$;
3) Run Gibbs sampler
   For $m = 1$ to M do
   - For $j = 1, \cdots, n_s, t = 1, \cdots, n_d$
     $s^m_{jt} \sim p(s_{jt}|\boldsymbol{\theta}_{-s_{jt}}, \mathbf{X})$ (Eq. (17))
   - For $j = 1, \cdots, n_s$
     $\mu^m_j \sim p(\mu_j|\boldsymbol{\theta}_{-\mu_j}, \mathbf{X})$ (Eq. (20))
     $\sigma^m_j \sim p(\sigma_j|\boldsymbol{\theta}_{-\sigma_j}, \mathbf{X})$ (Eq. (22))
   - For $i = 1, \cdots, n_c, j = 1, \cdots, n_s$
     $a^m_{ij} \sim p(a_{ij}|\boldsymbol{\theta}_{-a_{ij}}, \mathbf{X})$ (Eq. (23))
   - For $i = 1, \cdots, n_c$
     $d^m_i \sim p(d_i|\boldsymbol{\theta}_{-d_i}, \mathbf{X})$ (Eq. (24))
     $e^m_i \sim p(e_i|\boldsymbol{\theta}_{-e_i}, \mathbf{X})$ (Eq. (27))
     $\sigma^m_i \sim p(\sigma_i|\boldsymbol{\theta}_{-\sigma_i}, \mathbf{X})$ (Eq. (30))
   end
4) Infer the sources through the Bayesian MMSE estimation

$$\tilde{s}_{jt} = \frac{1}{M-B}\sum_{m=B+1}^{M} \mathbf{s}^m_{jt}, \forall j, t, \qquad (31)$$

   where B denotes the number of iterations of the burn-in period.
5) Retrieve the source scale using N_{cal} calibration points ($N_{cal} \geq 2$).

data. Then, in subsection IV-B, we test our proposal in a real situation. In both cases, the performance of each estimated source was inferred according to the following index:

$$\text{SIR}_i = 10\log\left(\frac{E\{s_i^2\}}{E\{(s_i - \hat{s}_i)^2\}}\right), \qquad (32)$$

where s_i denotes the actual source and $\hat{s}_i$ its respective estimation (after correct scaling). A global index is defined as $\text{SIR} = 1/n_s \sum_{i=1}^{n_s} \text{SIR}_i$.

A. Experiments with artificial data

We made use of the database of selectivity coefficients presented in [19] to define the following testing scenarios:

- First scenario ($n_s = 3$ and $n_c = 3$): array of three electrodes (one K^+-ISE, one NH_4^+-ISE and one Na^+-ISE) to estimate K^+, NH_4^+ and Na^+. Mixing system parameters: $\mathbf{A} = [1\ 0.16\ 0.40; 0.25\ 1\ 0.19; 0.40\ 0.13\ 1]$, $\mathbf{d} = [0.059\ 0.050\ 0.055]^T$ and $\mathbf{e} = [0.095\ 0.105\ 0.110]^T$.
- Second scenario ($n_s = 2$ and $n_c = 2$): array of two electrodes (one Ca^{2+}-ISE and one Na^+-ISE) to estimate the activities of Ca^{2+} and Na^+. Mixing system parameters: $\mathbf{A} = [1\ 0.39; 1\ 0.20]$, $\mathbf{d} = [0.026\ 0.046]^T$ and $\mathbf{e} = [0.100\ 0.090]^T$.

In the first scenario we have a PNL model as mixing system (see Eq. (1)). In the second one, the mixing system is composed of a nonlinear mixing mapping followed by component-wise logarithm functions. Finally, we consider noisy mixtures with a signal-to-noise ratio of $\text{SNR} = 18$ dB in both cases.

We tested our method in a situation where the sources are given by log-normal distributions (the number of samples was $n_d = 500$). For the first scenario, we considered $P = 50000$ iterations for the Gibbs sampler with a burn-in period of $B = 30000$. The number of calibration points used in the post-processing stage was $N_{cal} = 5$. The performance indices in this situation were $\text{SIR}_1 = 12.83$ dB, $\text{SIR}_2 = 19.0$ dB, $\text{SIR}_3 =$
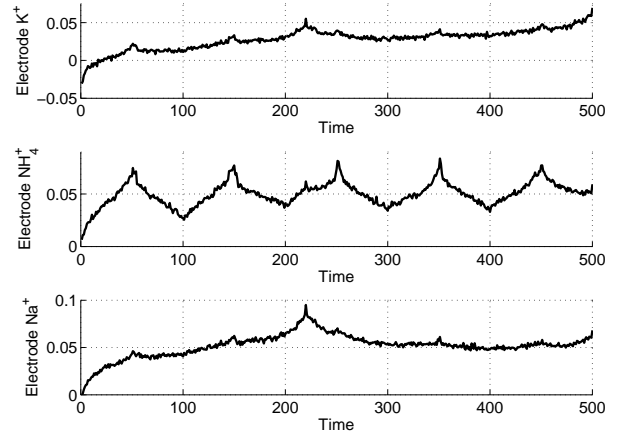

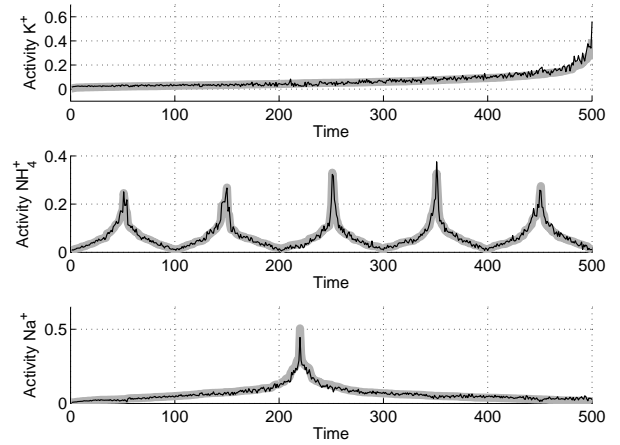Fig. 1. Artificial data (first scenario): ISE array outputs (mixtures).

Fig. 2. Artificial data (first scenario): actual sources (gray) and their estimation (black).

17.4 dB, and $\text{SIR} = 16.4$ dB, which indicates that our proposal was able to achieve a good source separation. To illustrate that, we show in Fig. 1 the mixed signals, and in Fig. 2, the actual sources and their respective estimations.

In the second scenario, we considered $P = 25000$ iterations for the Gibbs sampler with a burn-in period of $B = 18000$, and $N_{cal} = 3$ calibration points were used. The performance indices in this case were $\text{SIR}_1 = 17.4$ dB, $\text{SIR}_2 = 16.2$ dB, and $\text{SIR} = 16.8$ dB. Again, our method was able to provide fair estimations of the sources.

B. Experiments with real data

A well-known example of interference in ion sensing concerns the estimation of ammonium and potassium. These ions play an important role in applications such as water quality monitoring and food industry [3]. A brief look at [19] confirms that in many ammonium and potassium electrodes one has selectivity coefficients that are not negligible. Moreover, in some situations, these coefficients can reach quite high values, which represents a strong interference process.

We conducted a set of experiments where a solution containing K^+ and NH_4^+ was analyzed through an ISE array composed of one K^+-ISE and one NH_4^+-ISE. The activities of
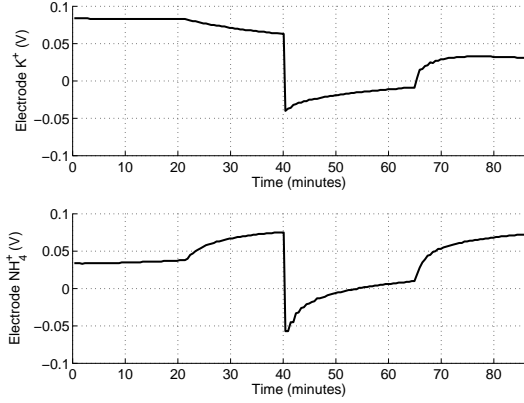
Fig. 3. Real data: responses provided by the ISE array.

Fig. 5. Real data: Retrieved signals (black) and actual sources (gray).

both ions were varied according to the following procedure. Initially, a solution containing a high concentration of potassium (10^{-1}M) was prepared. Then, by injecting a solution containing ammonium chloride (NH_4^+Cl), the concentration of NH_4^+ was increased whereas the concentration of K^+ was decreased. We repeat the same experiment but starting with a different initial concentration of K^+ (almost 10^{-4}M). After joining these two experiments in a single dataset, we obtained the sources (the variation in time of ionic activities[7]) shown in Fig. 5. The total number of samples of the two experiments altogether was $n_d = 170$.

In Fig. 3, we present the responses of the ISE array. Since we have access to the inputs and to the outputs of the electrode array, it is possible to analyze the fitness of the NE model for this case, that is, we can have an idea about the amount of noise in the mixing model. Concerning the potassium electrode, we measured a signal-to-noise ratio (SNR) of $SNR_{K^+} = 24$ dB. For the ammonium electrode, this value was given by $SNR_{NH_4^+} = 20$ dB. We assumed a Gaussian modeling for the noise in each electrode, and, to have an insight into the pertinence of such assumption, the distribution of the regression errors is plotted in Fig. 4.

(a) Potassium electrode.
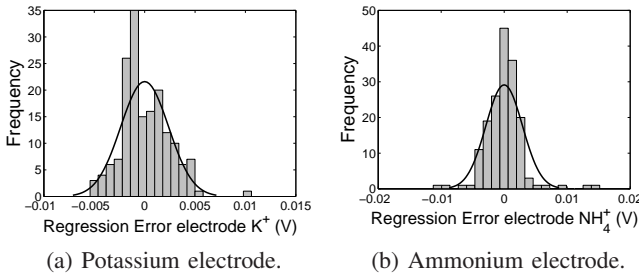
(b) Ammonium electrode.

Fig. 4. Histograms of the regression errors resulting from a fitting with the NE model. The black curves correspond to the fitted Gaussian distribution.

After scale and translation normalization with $N_{cal} = 4$ calibration points, the performance indices (average of 30 experiments) obtained by the Bayesian method in this scenario were $SIR_1 = 24.0$ dB, $SIR_2 = 22.5$ and $SIR = 23.2$.

Concerning the parameters of the Gibbs sampler, the number of performed iterations was $P = 10000$ and the burn-in period was $B = 7000$. The obtained signals are shown in Fig. 5. Despite a small residual interference, mainly for the K^+ activity, the method was able to provide good estimations of the sources. On the other hand, the application of a PNL source separation method based on ICA [24] provided poor approximations of the sources ($SIR_1 = 7.6$ dB, $SIR_2 = -0.3$ and $SIR = 3.6$).

C. On the number of calibration points used in the post-processing stage

We argued in the introductory section that the main benefit brought by BSS methods is the reduction of the calibration step. However, as mentioned before, our BSS method needs at least two calibration points for retrieving the correct sources' scales. From this apparent paradox, we may ask ourselves why not simply use the available training points in a supervised context. For example, we could define a separating system (inverse of NE model) where the estimate $\hat{s}_{jt}$ is given by

$$\hat{s}_{jt} = \sum_{i=1}^{n_c} a_{ij}^* 10^{\frac{x_{it} - e_i^*}{d_i^*}}. \tag{33}$$

Then, we could estimate a_{ji}^*, d_i^*, and e_i^* based on a supervised approach, i.e. by minimizing the MSE error $J = \sum_{j=1}^{n_s} \sum_{t=1}^{n_d} (\hat{s}_{jt} - s_{jt})^2$. We will refer to such strategy as the supervised-NE algorithm.

In fact, the argument presented in the last paragraph is somewhat naive since it implicitly assumes that the performance of a supervised method does not depend on the number of training samples, which is not the case. Indeed, in Fig. 6, we present the evolution of the training error of the supervised-NE method as the number of calibrations points N_{cal} grows. In this figure, which represents an average of 100 runs, one can note that supervised-NE method requires, at least, $N_{cal} = 25$ for providing good estimations. Conversely, our proposal requires only $N_{cal} = 5$ to provide equivalent estimations[8]. This is achieved because the "real" data processing in our

---

[7]The ionic activities were estimated by using the actual concentrations and the Debye-Hückel formalism [2].

[8]In Fig. 6, if the test error were plotted, then the difference between the supervised NE method and the Bayesian method would become even smaller.
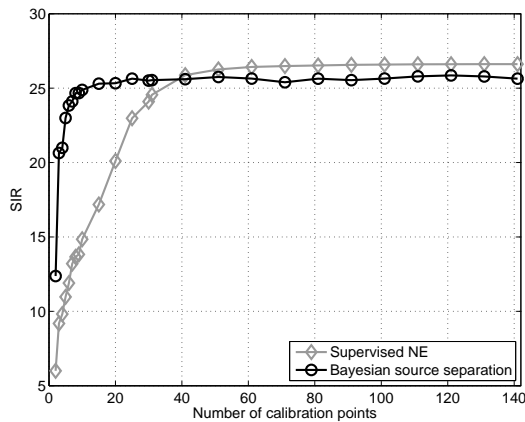
Fig. 6. Performance index as the number of calibration points grows.

proposal is done by the Bayesian BSS algorithm, whereas the calibration points are only used to circumvent the signal ambiguities. Hence, although the complete suppression of the calibration stage is not possible, our method can work even if only a few calibration samples are available.

D. Discussion

The experiments with artificial data attested that the proposal can achieve good estimations even in the difficult case where the valences are different. In the scenario with real data, the Bayesian algorithm achieved a much better performance than the ICA-based PNL algorithm. In fact, in this case the sources were clearly dependent and, thus, they violated the fundamental assumption of any ICA method. Evidently, a scenario with dependent sources also poses a problem to a Bayesian method, since there is no guarantee that the provided data representation is unique in such a case. Nonetheless, in contrast to ICA, a Bayesian method does not optimize a functional associated with the statistical independence. Rather, it searches for a "good" data representation given a set of prior information. Thus, in the Bayesian approach, the independence should be seen as a simplifying assumption, that is, we are just omitting an additional prior information to the inference machine, which may still work.

Concerning the algorithm's convergence, we observed that the Gibbs sampler may get trapped in local minima, thus leading to poor estimations. For the situations with two sources the percentage of poor convergences[9] was 3% for artificial data and 7% for the real case; in the case with three sources, this value attained 17%. Although we do not have access to the source, it is still possible to identify a bad solution: we observed that poor convergence usually implies in a significant mismatch between the representation provided by the Bayesian algorithm and the actual array response.

Another important point concerns the execution time of our method. For the experiments with real data, the MCMC algorithm took about[10] 26s to perform 10000 iterations. Nev-

---

[9]These percentages were obtained through a visual inspection after 30 executions.

[10]The method was implemented in Matlab (Windows XP) and the simulations were performed in a Intel Core 2 duo 3GHz, 2048MB RAM.

ertheless, the algorithm took 380s to perform 50000 iterations in the situation with three sources. This points out a well-known drawback of MCMC-based Bayesian methods: their computational burden can become quite large as the number of sources and samples grows.

Finally, let us make a remark concerning the choice of the priors. In this work, we tried to define priors 1) that ease the resulting inference problem and 2) that, based on the available information, limit the range of the unknown parameters. Evidently, there is no guarantee that our choices are optimum and, as we mentioned before, a large dataset of measurements would permit to refine the priors' definition. An interesting issue in this context would be to compare the selected priors with alternative models. Because the possibilities are non-exhaustive, we consider a single example where, instead of log-normal priors for the sources, Gamma distributions are considered. This distribution also describes non-negative variables and, moreover, provides a flexible solution as it can model from sparse to almost uniform sources [16].

After applying the solution with the Gamma modeling to process the real data, the obtained performance (average of 30 experiments with $N_{cal} = 4$ calibration points) —$SIR_1 = 22.4\,dB$ and $SIR_2 = 17.5\,dB$ —was inferior to that of the log-normal prior. Furthermore, unlike the log-normal case, it is not possible to find a conjugate prior in the estimation of the Gamma distribution parameters. Thus, it becomes necessary to incorporate an additional MH algorithm into the Gibbs' sampler, which increases the algorithm's complexity and has the inconvenient of requiring the definition of instrumental distributions, which is not an easy task.

V. CONCLUSIONS

In this work, we proposed a Bayesian source separation method for unsupervised quantitative analysis via an ISE array. A MCMC algorithm, the Gibbs sampler, was used in the inference stage. We defined the prior distributions based on informations that are usually available in a chemical sensing problem, such as the non-negativity of the sources. The resulting separation method did well in different scenarios, including in a real problem where the estimation of the activities of ammonium and potassium was desired. The results shown that ISE arrays equipped with blind source separation can operate even if only a reduced number of calibration points is available. This nice feature can pave the way for alternatives to supervised methods. For example, a less demanding calibration step could ease analysis in the field.

Despite the encouraging results, there are still some questions that demand future work. A first one concerns the possibility of increasing the precision of the developed source separation method. Indeed, in our experiments with real data, the remaining small interference would not be acceptable in very-high-precision applications. We believe that a more precise mixing model could increase the estimation quality. In this spirit, a first step is to search for alternatives to the NE model that take into account, for example, dynamics aspects of the interference modeling. Furthermore, it could be equally interesting to develop a Bayesian framework based on

a temporal modeling for the sources, since chemical sources do have a time-structure.

References

[1] L. T. Duarte, C. Jutten, and S. Moussaoui, "Ion-selective electrode array based on a bayesian nonlinear source separation method," in *Proceedings of the 8th ICA Conference (LNCS)*, 2009, pp. 662–669.

[2] P. Gründler, *Chemical sensors: an introduction for scientists and engineers*. Springer, 2007.

[3] P. Fabry and J. Fouletier, Eds., *Microcapteurs chimiques et biologiques*. Lavoisier, 2003, in French.

[4] E. Bakker and E. Pretsch, "Modern potentiometry," *Angewandte Chemie International Edition*, vol. 46, pp. 5660–5668, 2007.

[5] H. Sakai, S. Iiyama, and K. Toko, "Evaluation of water quality and pollution using multichannel sensors," *Sensons and Actuators B*, vol. 66, pp. 25–255, 2000.

[6] U. Oesch, D. Ammann, and W. Simon, "Ion-selective membrane electrodes for clinical use," *Clinical Chemistry*, vol. 32, pp. 1448–1459, 1986.

[7] O. T. Guenat, S. Generelli, N. F. de Rooij, M. Koudelka-Hep, F. Berthiaume, and M. L. Yarmush, "Development of an array of ion-selective microelectrodes aimed for the monitoring of extracellular ionic activities," *Analytical Chemistry*, vol. 78, pp. 7453–7460, 2006.

[8] Y. G. Vlasov, A. V. Legin, and A. M. Rudnitskaya, "Electronic tongue: chemical sensor systems for analysis of aquatic media," *Russian Journal of General Chemistry*, vol. 78, pp. 2532–2544, 2008.

[9] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis*. John Wiley & Sons, 2001.

[10] P. Comon and C. Jutten, Eds., *Séparation de sources 1 : concepts de base et analyse en composantes indépendantes*. Hermes Science Publications, 2007, in French.

[11] C. Jutten and J. Karhunen, "Advances in blind source separation (BSS) and independent component analysis (ICA) for nonlinear mixtures," *International Journal of Neural Systems*, vol. 14, pp. 267–292, 2004.

[12] S. Bermejo, C. Jutten, and J. Cabestany, "ISFET source separation: foundations and techniques," *Sensors and Actuators B*, vol. 113, pp. 222–233, 2006.

[13] G. Bedoya, C. Jutten, S. Bermejo, and J. Cabestany, "Improving semiconductor-based chemical sensor arrays using advanced algorithms for blind source separation," in *Proceedings of Sensors for Industry Conference, 2004*, 2004, pp. 149–154.

[14] L. T. Duarte and C. Jutten, "A mutual information minimization approach for a class of nonlinear recurrent separating systems," in *Proceedings of the IEEE MLSP 2007*, 2007.

[15] C. Févotte and S. J. Godsill, "A Bayesian approach for blind separation of sparse sources," *IEEE Transactions on Audio, Speech and Language Processing*, vol. 14, pp. 2174–2188, 2006.

[16] S. Moussaoui, D. Brie, A. Mohammad-Djafari, and C. Carteret, "Separation of non-negative mixture of non-negative sources using a bayesian approach and mcmc sampling," *IEEE Transactions on Signal Processing*, vol. 54, pp. 4133–4145, 2006.

[17] C. P. Robert and G. Casela, *Monte Carlo Statistical Methods*, 2nd ed. Springer, 2004.

[18] C. P. Robert, *The Bayesian Choice*. Springer, 2007.

[19] Y. Umezawa, P. Bühlmann, K. Umezawa, K. Tohda, and S. Amemiya, "Potentiometric selectivity coefficients of ion-selective electrodes," *Pure and Applied Chemistry*, vol. 72, pp. 1851–2082, 2000.

[20] S. A. Dorneanu, V. Coman, I. C. Popescu, and P. Fabry, "Computer-controlled system for ISEs automatic calibration," *Sensors and Actuators B*, vol. 105, p. 521531, 2005.

[21] M. Hartnett and D. Diamond, "Potentometric nonlinear multivariate calibration with genetic algorithm and simplex optmization," *Analytical Chemistry*, vol. 69, pp. 1909–1918, 1997.

[22] A. Doucet and X. Wang, "Monte Carlo methods for signal processing," *IEEE Signal Processing Magazine*, vol. 22, pp. 152–170, 2005.

[23] S. M. Kay, *Fundamentals of statistical signal processing: estimation theory*. Prentice-Hall, 1993.

[24] L. T. Duarte, R. Suyama, R. R. F. Attux, F. J. V. Zuben, and J. M. T. Romano, "Blind source separation of post-nonlinear mixtures using evolutionary computation and order statistics," in *Proceedings of the 6th ICA Conference (LNCS)*, 2006, pp. 66–73.
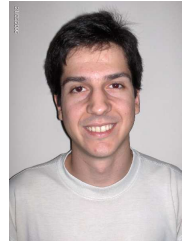
Leonardo Tomazeli Duarte received the B.S and the M.S. degrees in electrical engineering from the University of Campinas (Unicamp), Campinas, Brazil, in 2004 and 2006 respectively. He is currently a PhD student at the Grenoble Institute of Technology (Grenoble INP), working at the Grenoble images, speech, signal and control laboratory (GIPSA-lab), France. His research interests include blind source separation, independent component analysis, smart chemical sensor arrays and nonlinear models.

Christian Jutten received the PhD degree in 1981 and the Docteur ès Sciences degree in 1987 from the Institut National Polytechnique of Grenoble (France). He taught as associate professor in the Electrical Engineering Department from 1982 to 1989. He was visiting professor in Swiss Federal Polytechnic Institute in Lausanne in 1989, before to become full professor in University Joseph Fourier of Grenoble. He is currently associate director of the Grenoble images, speech, signal and control laboratory (GIPSA, 300 people) and head of the Department Images-Signal (DIS) of this laboratory. For 25 years, his research interests are blind source separation, independent component analysis and learning in neural networks, including theoretical aspects and applications in signal processing (biomedical, seismic, speech). He is author or co-author of more than 45 papers in international journals, 3 books, 16 invited papers and 140 communications in international conferences. He has been associate editor of IEEE Trans. on Circuits and Systems (1994-95), and co-organizer the 1st International Conference on Blind Signal Separation and Independent Component Analysis (Aussois, France, 1999). He has been a scientific advisor for signal and images processing at the French Ministry of Research from 1996 to 1998 and for the French National Research Center from 2003 to 2006. He has been associate editor of IEEE Trans. CAS from 1992 to 1994. He is a member of the technical committee Blind signal Processing of the IEEE CAS society and of the technical committee Machine Learning for signal Processing of the IEEE SP society. He is a reviewer of main international journals (IEEE Trans. on Signal Processing, IEEE Signal Processing Letters, IEEE Trans. on Neural Networks, Signal Processing, Neural Computation, Neurocomputing, etc.) and conferences in signal processing and neural networks (ICASSP, ISCASS, EUSIPCO, etc.). He received the EURASIP best paper award in 1992 and Medal Blondel in 1997 from SEE (French Electrical Engineering society) for his contributions in source separation and independent component analysis, and has been elevated as a Fellow IEEE and a senior Member of Institut Universitaire de France in 2008.

Saïd Moussaoui received the State engineering degree from Ecole Nationale Polytechnique, Algiers, Algeria, in 2001, and the Ph.D. degree in 2005 from Universit Henri Poincar, Nancy, France. He is currently Associate Professor at Ecole Centrale de Nantes. Since september 2006, he is with the Institut de Recherche en Communications et Cybrntiques de Nantes (IRCCYN, UMR CNRS 6597). His research interests are in statistical signal processing methods and their applications.