

A Bayesian Blind Source Separation Method for a Linear-quadratic Model

Stéphanie MADROLLE
Univ. Grenoble Alpes, CEA,
LETI, MINATEC Campus,
F-38054 Grenoble, France.
stephanie.madrolle@cea.fr

Leonardo T. DUARTE
School of Applied Sciences,
University of Campinas
(UNICAMP), Brazil
leonardo.duarte@fca.unicamp.br

Pierre GRANGEAT
Univ. Grenoble Alpes, CEA,
LETI, MINATEC Campus,
F-38054 Grenoble, France.
pierre.grangeat@cea.fr

Christian JUTTEN
Univ. Grenoble Alpes,
GIPSA-lab, UMR CNRS 5216,
F-38400 Saint-Martin-d'Hères,
France.
christian.jutten@grenoble-inp.fr

Abstract— We propose a blind source separation method based on Bayesian inference in order to separate two sources in a linear-quadratic mixing model. This nonlinear model describes for example the response of a non-selective metal-oxide gas sensor (MOX) in presence of two gases such as acetone and ethanol diluted in an air buffer. In order to quantify the gas components, it is necessary to inverse the linear-quadratic model. In addition, we look for reducing the number of samples for the calibration step. Therefore, we here propose a Bayesian blind source separation method, with only few points of calibration and which is based on Monte Carlo Markov Chain (MCMC) sampling methods to estimate the mean of the posterior distribution. We analyze the performance on a set of simulated samples. We use a cross-validation approach, with three steps: first, we blindly estimate the mixing coefficients and sources; second, we correct the scale factors thanks to few calibration samples; and third, we validate the method on validation samples, estimating sources thanks to mixing coefficients estimated before on the calibration samples. We compare this unsupervised nonlinear method with a supervised method to evaluate the performance with respect to the number of calibration points: with 10 calibration points instead of 160, the performance achieves 11 dB, with a loss limited to 1.5 dB.

Keywords—Linear-quadratic model – Inverse problems – Statistical signal processing - Bayesian method – Blind source separation – Monte Carlo Markov Chain (MCMC)- Gas mixture – MOX sensor.

I. INTRODUCTION

Blind source separation (BSS) aims at estimating source signals by only using the measurements of these sources. For example, BSS can be applied in audio processing to separate the speech signals of different people speaking simultaneously in a room. Here, we apply it to quantify gases in a mixture with a non-specific sensor.

The problem of BSS is ill-posed, but while simple priors lead to identifiability and uniqueness in linear mixture problems, there are no such theoretical results for general nonlinear mixture problems. Theoretical results and algorithms have been proposed for particular mixtures, like post-nonlinear models or linear-quadratic models, using both priors on the mixture model and on the sources [1] (like independence, non-negativity, sparsity, etc.). In this paper, we propose to consider a Bayesian method, which allows to add priors to help the convergence to a

relevant solution. Also, Bayesian approach is robust to noise for solving this ill-posed nonlinear inversion problem. We expand the method proposed in [2] to a linear-quadratic model. As an application, we consider the quantification of a two gas mixture such as acetone and ethanol diluted in an air buffer with a non-specific metal-oxide (MOX) gas sensor used in a triple temperature mode.

The paper is organized as follow. In Section II, we present the ill-posed problem, and we introduce the nonlinear mixing model (linear-quadratic model), which is a mixing model adapted to MOX sensors for gas analysis. In Section III, the Bayesian source separation method is presented, including the motivations, the definition of the prior distributions, the Bayesian estimation strategy and the different algorithms to set-up. In Section IV, we present the results on simulated data to validate the nonlinear blind method and we compare the blind method with a supervised one.

II. PROBLEM DEFINITION

A. Mixing model

The mixing model is a mathematical representation connecting the sources (i.e. the gas concentrations) and the sensors responses. For breath application, we consider a MOX sensor to detect volatile organic compounds (VOCs). We have introduced a linear quadratic model which coefficients are changing according the temperature of the sensitive layer defined by the heater. This model was experimentally validated in [3]. By considering 3 different temperatures of the sensitive layer, we obtain three virtual sensors with sufficient diversity [3], [4] for using source separation methods. As our goal is to estimate 2 sources, the availability of 3 (virtual) sensors allows to have more information. In mathematical terms, the resulting mixing model for this case is given by:

$$\begin{cases} x|_{T_1} \approx a_{11}s_1 + a_{12}s_2 + d_1s_1s_2 + b_{11}s_1^2 + b_{12}s_2^2 + e_1 \\ x|_{T_2} \approx a_{21}s_1 + a_{22}s_2 + d_2s_1s_2 + b_{21}s_1^2 + b_{22}s_2^2 + e_2 \\ x|_{T_3} \approx a_{31}s_1 + a_{32}s_2 + d_3s_1s_2 + b_{31}s_1^2 + b_{32}s_2^2 + e_3 \end{cases} \quad (1)$$

where a_{ij} , b_{ij} , d_i and e_i are the model coefficients at temperature T_i , s_j are the sources, linked to the gas

concentrations c_j by a power law thanks to a calibrated exponent r_j :

$$s_j = c_j^{r_j} \quad (2)$$

For N_d measures, the model (1) can be written in matrix form as follows, where “o” is the Hadamard (i.e. entrywise) product:

$$\mathbf{X} - \mathbf{E} = \mathbf{A}\mathbf{S} + \mathbf{d}(\mathbf{s}_1 \circ \mathbf{s}_2) + \mathbf{B}\mathbf{S}^2 \quad (3)$$

This model is known as linear-quadratic model [5].

B. Calibration

In a standard protocol, gas analysis in MOX sensor is done using supervised learning [5], which requires many samples in the learning step and is a costly procedure. Here, to lighten this step, we propose to consider a blind source separation approach. In this method, sources and mixing coefficients are estimated blindly, but up to a scale factor. Therefore, to estimate the exact value of the sources, few points of calibration are required. Nevertheless, the number of calibration points are limited compared to the one required for a supervised method. Thus, the main interest to work is to show that blind source separation method, with a few additional calibration points, is able to achieve accurate gas.

III. A BAYESIAN BLIND SOURCE SEPARATION METHOD

We here develop a Bayesian blind source separation method for the linear-quadratic model related to MOX-sensors. This unsupervised probabilistic method allows to estimate sources linked to gas concentrations and to introduce priors on the solution for regularization.

A. Assumptions

In order to implement the Bayesian method, we consider the following assumptions:

- the sources are statistically independent
- the sources are positive and included in the region $[s_j^{min}, s_j^{max}]$,
- the expression of the mixing model is known (i.e. linear-quadratic), but not the coefficients,
- the mixing model is invertible for the studied sources,
- the mixing coefficients are included in the region $[u_{im}^{min}, u_{im}^{max}]$.

Based on these assumptions and the associated priors, we will now describe the method to estimate the model coefficients and the sources.

B. Bayes' theorem

The Bayes' theorem is the framework of our processing method. We denote the following indexes $i \in [1, N_c]$, $j \in [1, N_s]$, $t \in [1, N_d]$, where N_c is the number of sensors, N_s the number of sources and N_d the number of samples. We note also $\boldsymbol{\theta}$ the vector of all unknown parameters including unknown source values for all samples t and all coefficients we want to estimate, so, for all (i, j, t) , $\boldsymbol{\theta} = \{\theta_{ijt}\}$, with $\theta_{ijt} = [s_{jt}, a_{ij}, d_i, b_{ij}, \sigma_i^2, \mu_j, p_j]$. s_{jt} are the sources j for the sample t ,

a_{ij}, d_i, b_{ij} are the mixing coefficients, σ_i^2 is the noise variance and μ_j and p_j are respectively the mean and the variance of the prior source distribution. The principle of the Bayesian method is to determine the posterior distribution $p(\boldsymbol{\theta}|\mathbf{X})$, thanks to the likelihood function $p(\mathbf{X}|\boldsymbol{\theta})$ and the prior distributions $p(\boldsymbol{\theta})$, using the Bayes' theorem:

$$p(\boldsymbol{\theta}|\mathbf{X}) \propto p(\mathbf{X}|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (4)$$

C. Likelihood

The likelihood is the distribution describing the match between an observed sample measurement \mathbf{x} and the mixing model associated with unknown parameters vector $\boldsymbol{\theta}$. As we suppose the noise model to be white Gaussian, the likelihood is Gaussian, the likelihood term is given by:

$$p(\mathbf{X}|\boldsymbol{\theta}) = \prod_{t=1}^{N_d} \prod_{i=1}^{N_c} \mathcal{N}_{x_{it}}(\mu_{it}^{likelihood}, \sigma_i^2) \quad (5)$$

where $\mathcal{N}_{x_{it}}(\cdot)$ represents a normal distribution of variance σ_i^2 and mean $\mu_{it}^{likelihood}$, which, based on model (1), is given by:

$$\mu_{it}^{likelihood} = \sum_{j=1}^{N_s} a_{ij}s_{jt} + \sum_{1 \leq j \leq k \leq N_s} d_i s_{jt} s_{kt} + \sum_{j=1}^{N_s} b_{ij}s_{jt}^2 \quad (6)$$

D. Prior distributions

The prior distributions are chosen by the user for modeling the initial knowledge on the parameters according to the assumptions described in section III.A. In the sequel, we explain how the priors were selected in our work.

1) Sources

For the sources we consider the interval, in which sources are living is known. Such an information can be represented, for instance, by a truncated Gaussian distribution:

$$p(s_{jt} | \mu_j, p_j, s_j^{min}, s_j^{max}) = \frac{\sqrt{\frac{p_j}{2}} \exp\left(-\frac{p_j}{2}(s_{jt} - \mu_j)^2\right) \mathbb{1}_{[s_j^{min}, s_j^{max}]}(s_{jt})}{\Phi\left(\sqrt{p_j}(s_j^{max} - \mu_j)\right) - \Phi\left(\sqrt{p_j}(s_j^{min} - \mu_j)\right)} \quad (7)$$

where $\mathbb{1}(\cdot)$ is the indicator function. $\mathbb{1}_{[s_j^{min}, s_j^{max}]}$ is equal to 1 in the interval $[s_j^{min}, s_j^{max}]$ and 0 outside. $\Phi(\cdot)$ is the cumulative density function of the Gaussian distribution, μ_j and p_j are respectively the mean and the inverse variance of the Gaussian distribution. The hyperparameters μ_j and p_j are also unknown. We suppose they are uniform in an interval, and we estimate them in the algorithm.

2) Mixing coefficients

For the mixing coefficients, we denote \mathbf{u}_i all the model parameters corresponding to the i sensor, that is:

$$\mathbf{u}_i = [a_{i1}, a_{i2}, d_i, b_{i1}, b_{i2}] \quad (8)$$

We denote u_{im} , the m^{th} parameter of \mathbf{u}_i , and $m \in [1, N_m]$. We assume that these parameters lie in a known interval. Thus, we

choose a uniform distribution as prior, with bounds $[u_{im}^{min}, u_{im}^{max}]$:

$$p(u_{im}) = \mathcal{U}_{[u_{im}^{min}, u_{im}^{max}]}(u_{im}) \quad (9)$$

3) Noise variance

For the noise variance, it is usual [6] to choose an inverse gamma prior:

$$p(\lambda_i) = \lambda_i^{\alpha\lambda_i - 1} \exp\left(-\frac{\lambda_i}{\beta\lambda_i}\right) \mathbb{1}_{[0, +\infty]}(\lambda_i) \quad (10)$$

where λ_i is the inverse variance.

E. Bayesian Inference

To obtain the posterior distribution, we assume that almost all parameters of θ are statistically independent (except μ_j, p_j with s_{jt}), which allows to factorize the prior distributions:

$$p(\theta|\mathbf{X}) \propto p(\mathbf{X}|\theta) \cdot \prod_{i=1}^{N_c} \prod_{m=1}^{N_m} p(u_{im}) \cdot \prod_{j=1}^{N_s} p(s_{jt}|\mu_j, p_j) \cdot \prod_{j=1}^{N_s} p(\mu_j) \cdot \prod_{j=1}^{N_s} p(p_j) \cdot \prod_{i=1}^{N_c} p(\sigma_i^2) \quad (11)$$

As the second order term affects the likelihood (6), and the posterior distribution depends on the likelihood, this posterior will be affected by the quadratic term.

The final minimum mean square estimator can be written as:

$$\theta_{MMSE} = \int \theta p(\theta|\mathbf{X}) d\theta \quad (12)$$

F. Correction of estimated parameters

This blind method estimates the parameters up to a scale factor. In order to find the exact sources, a calibration with mixtures of known sources is needed for estimating the scale factor. As calibration points, we take the four points at the edges of source domain (minimum and maximum for each source), which is the minimum number in order to avoid error of interpolation, and we add some others points drawn randomly within this source domain, to be more robust. Thus, we find the correction factors and we correct the initial estimation according to the following relationship:

$$\hat{s}_j = \alpha_j \hat{s}_{j, \text{before calib}} + \beta_j \quad (13)$$

Thus, we denote \hat{s}_j as the estimated sources.

G. Algorithm

To implement this method, we use the Gibbs' sampler [7] which is a Markov Chain Monte Carlo (MCMC) method based on a stochastic sampling algorithm. The Gibbs' sampler randomly draws a sample according to the conditional distributions of each variable, with respect to the others, in

order to find the posterior distributions (Fig. 1). These conditional distributions are calculated from each prior law:

$$p(\theta_q | \theta_{-q}, \mathbf{X}) \propto p(\mathbf{X}, \theta) p(\theta_q) \quad (14)$$

θ_q is the q-th parameter of θ and θ_{-q} represents the vector with all parameters except θ_q . Thus, the mixing parameters conditional law is an exponential distribution:

$$p(u_{im} | \theta_{-u_{im}}, \mathbf{X}) \propto \mathcal{U}_{[u_{im}^{min}, u_{im}^{max}]}(u_{im}) \cdot \exp\left(-\sum_{t=1}^{N_d} \frac{1}{2\sigma_i^2} (x_{it} - \mu_{it}^{\text{likelihood}}(u_{im}))^2\right) \quad (15)$$

The source conditional distribution is also an exponential distribution:

$$p(s_{jt} | \theta_{-s_{jt}}, \mathbf{X}) \propto \mathbb{1}_{[s_j^{min}, s_j^{max}]}(s_{jt}) \cdot \exp\left(-\sum_{i=1}^{N_c} \frac{1}{2\sigma_i^2} (v_{ij} s_{jt}^2 + \psi_{ijt} s_{jt} + \Omega_{ijt})^2 - 0,5 \cdot p_j (s_{jt} - \mu_j)^2\right) \quad (16)$$

With:

$$v_{ij} = -d_{ij} \quad (17)$$

$$\psi_{ijt} = -\left(a_{ij} + \sum_{g=1}^{N_s} b_{ijg} s_{gt}\right) \quad (18)$$

$$\Omega_{ijt} = x_{ijt} - \sum_{g=1}^{N_s} a_{ig} s_{gt} - \sum_{g=1}^{N_s} d_{ig} s_{gt}^2 - \sum_{\substack{1 \leq g < k \leq N_s \\ g \neq j}} b_{igk} s_{gt} s_{kt} \quad (19)$$

Nevertheless, as $\mu_{it}^{\text{likelihood}}$ contains a quadratic term, we do not have conjugated distributions and so to implement this conditional distribution, we need to use a Metropolis Hasting (MH) algorithm [8]. The principle of MH algorithm is to propose a new sample drawn according to an instrumental distribution. The new sample is accepted only if its acceptance probability is higher than a random number. The instrumental law must be simple to implement and close to the posterior distribution. We choose Gaussian laws.

Finally, we compute as final estimator the mean of the posterior distribution, which is the Bayesian minimum mean square error estimator (MMSE) over the drawn samples:

$$\tilde{\theta}_{MMSE} = \frac{1}{M} \sum_{i=1}^M \theta^i \quad (20)$$

where M is the number of θ samples of the Markov chain drawn after a "heating period" after which the distribution of the samples is supposed to have converged towards the posterior distribution.

1) Initialization
 2) For $p=1:P$
 a) $u_{im}^p \sim p(u_{im} | s_{jt}^{p-1}, \sigma_i^{2p-1}, \mu_j^{p-1}, p_j^{p-1}, \mathbf{X})$
 b) $s_{jt}^p \sim p(s_{jt} | u_{im}^p, \sigma_i^{2p-1}, \mu_j^{p-1}, p_j^{p-1}, \mathbf{X})$
 c) $1/\sigma_i^{2p} \sim p(1/\sigma_i^2 | u_{im}^p, s_{jt}^p, \mu_j^{p-1}, p_j^{p-1}, \mathbf{X})$
 d) $\mu_j^p \sim p(\mu_j | u_{im}^p, s_{jt}^p, \sigma_i^{2p}, p_j^{p-1}, \mathbf{X})$
 e) $p_j^p \sim p(\sigma_i^2 | u_{im}^p, s_{jt}^p, \sigma_i^{2p}, \mu_j^p, \mathbf{X})$
 3) end
 with the conditional distributions :
 $p(\theta_q | \theta_{-q}, \mathbf{X}) \propto p(\mathbf{X}, \theta) p(\theta_q)$

Fig. 1 : Gibbs sampling for the Bayesian algorithm

IV. SIMULATION RESULTS AND DISCUSSION

A. Validation of the method

In order to validate the method, we propose to divide randomly the total samples in 5 subsets and we proceed in three steps:

- A blind source separation (BSS), as described, knowing only the measures. This step allows to estimate mixing coefficients and sources up to a scale factor. For this step, 4 subsets out of 5 are used.
- A calibration with a few points. It allows to estimate the scale factors. One-sixteenth of points are used. For these calibration points, we know exactly the values of the sources. We choose the edges of the source domain and few points selected randomly in the region. At the end of these two steps, the values of coefficients and sources are estimated, for 4 segments.
- A validation step, using the remaining subset. We fix the mixing coefficients to the corrected values taking into account the scale factors obtained at the last step, and we estimate the sources.

If there are $N=200$ data points, 160 points from 4 subsets are used for the BSS, among which 10 points are selected for the calibration, and the remaining 40 points of the last subset are used for validation. These three steps are repeated until each sample is used one time during the validation step. As there are 5 subsets, it is repeated 5 times.

Then, we reiterate the process ten times, for ten different segmentations drawn randomly and we keep the mean value over these ten estimations.

B. Quality factors

To quantify the performance of source estimations, we use the signal to interference ratio (SIR) and the correlation coefficient ρ_j . Using all sample t of one source j that we note \mathbf{s}_j , we calculate the SIR thanks to the following equation:

$$\text{SIR}(s_j) = 10 \cdot \log_{10} \left(\frac{\|\mathbf{s}_j\|_2}{\|\mathbf{s}_j - \hat{\mathbf{s}}_j\|_2} \right) \quad (21)$$

C. Simulated data and diversity

To validate the method, we use simulated data. First, we simulate $N=200$ sources according to a truncated lognormal law. The lognormal law is close to the normal law which is the prior we included in the algorithm, but we keep it different to insert an error between the prior and the actual sources. The first source is sampled using a lognormal distribution with a mean of 8.3 and a variance of 0.1 and the second source using a lognormal distribution with a mean of 1.7 and a variance of 0.06. Then, we simulate the measures thanks to the linear-quadratic model, using the mixing coefficients corresponding to measures on two gases, acetone and ethanol, acquired with a MOX sensor at three sensor temperatures. We add a zero-mean Gaussian noise, with a standard deviation of 0.01. The diversity of this simulated data is presented in Fig. 2.

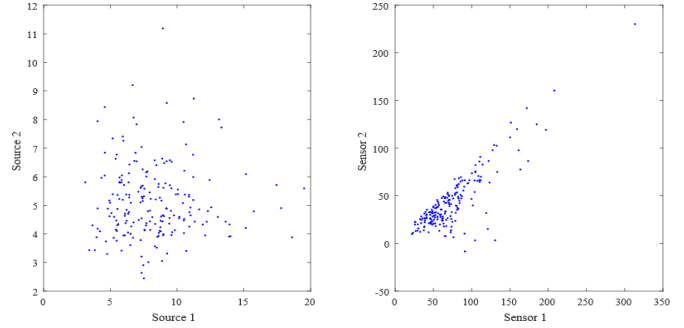


Fig. 2 : Diversity of the simulated data. On the left graph, the diversity on sources (source 2 as a function of source 1). On the right, diversity of sensor measurement (sensor 2 as a function of sensor 1).

D. Results of estimation

Thus, with the simulated data, we apply our algorithm to estimate the two sources (Fig. 3). The bounds for the sources are: $[s_1^{min}, s_1^{max}] = [2, 20]$ and $[s_2^{min}, s_2^{max}] = [2, 12]$. The SIR is good: 11.1 dB for both sources. Also, the correlation between exact and estimated sources is very good: superior to 0.96.

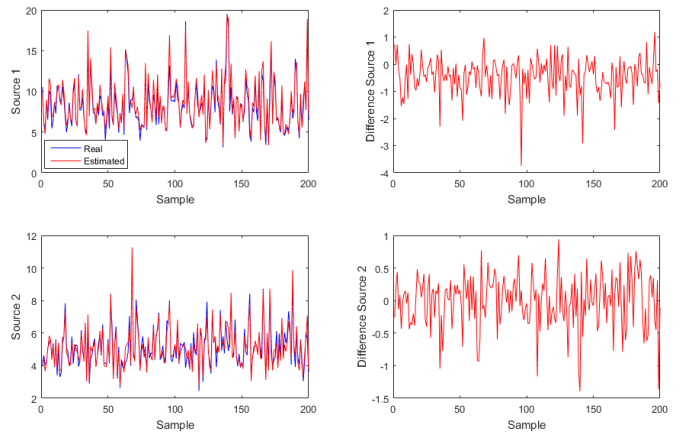


Fig. 3 : Source estimation thanks to the unsupervised Bayesian method. On the top, the source 1 and on the bottom the source 2. On the left, the source values, exact source values in blue and estimated source values in red. On the right, the difference between the exact and the corrected estimated source values.

E. Comparison with the supervised method

Now, we compare (Fig. 4) the results of this unsupervised method to the results of a Bayesian supervised method [9]. For this the mixing coefficients are estimated assuming we know the source concentration for all the samples. Then, we use these mixing coefficients to estimate the source values, on the validation samples. The aim of this comparison is to see the difference of source quantification error while reducing the number of calibration samples. On Fig 4, the remaining estimation error on source 2 with the unsupervised method is probably due to the choice of calibration points that needs to be optimized.

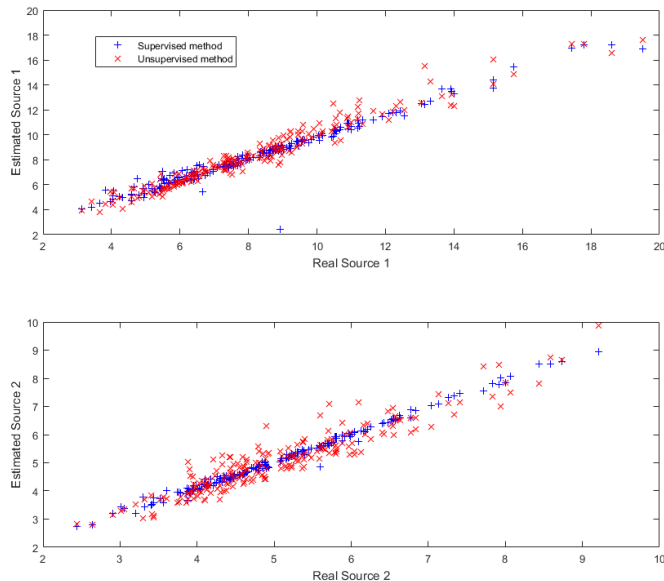


Fig. 4 : Comparison between the supervised method (blue) and the unsupervised method (red). The estimated source values are presented here in function to the exact source values. The correlation is close to 1.

TABLE I. QUALITY FACTORS OF SOURCE ESTIMATIONS WITH SIMULATED DATA, USING A BAYESIAN SUPERVISED ALGORITHM AND A BAYESIAN UNSUPERVISED ALGORITHM

	SIR	Correlation	Number of calibration points
Supervised Method	12.5 dB	0.98	160
Unsupervised Method	11.1 dB	0.96	10
Unsupervised Method	9.8 dB	0.93	4

Thus, here, only 10 calibration points are sufficient, and the quality of sources estimation is minimally affected by this diminution. Indeed, the correlation is slightly below and we lost about 1.5 dB on the signal to interference ratio. This loss is illustrated on Fig.4 where red points are more dispersed. With 4 calibration mixtures, the SIR is also interesting, close to 10 dB. Reducing the number of calibration points while losing only a little on the quality of estimation, is of high practical interest. This is in particular the case for MOX gas sensors.

V. CONCLUSIONS

To conclude, we have presented in this work a new method to inverse a linear-quadratic model. The Bayesian method allows to add statistical priors to obtain a good estimation of the sources. This advantage is particularly interesting as the problem is ill-posed. In addition, the blind source separation method presents the interest to only require a few number of calibration points, only for estimating the scale factors, which is an inherent indeterminacy of source separation method. Only using 10 calibration points impacts only a little the source estimation quality. Thus, this new unsupervised method applied on a linear-quadratic model allows to estimate for example two gases in a gas mixture with only 10 calibration samples. It's very advantageous to alleviate the step of sensors calibration, which is usually long, fastidious and costly, especially since the estimation error remains low and acceptable.

The next step is to optimize the choice of calibration points. Others methods like non-negative matrix factor (NMF) [10] also could be interested to study. Then, we will increase the number of sources and apply this nonlinear method on experimental data, acquired by a MOX sensor, to distinguish several gases in a mixture and to quantify the gas concentrations, which is useful particularly for medical applications such as breath gas analysis.

VI. ACKNOWLEDGEMENT

This work was partly supported by the project 2012-ERC-AdG-320684 CHESS.

REFERENCES

- [1] Y. Deville and L. T. Duarte, "An Overview of Blind Source Separation Methods for Linear-Quadratic and Post-nonlinear Mixtures," in *Latent Variable Analysis and Signal Separation*, 2015, pp. 155–167.
- [2] L. T. Duarte, C. Jutten, and S. Moussaoui, "Bayesian source separation of linear and linear-quadratic mixtures using truncated priors," *J. Signal Process. Syst.*, vol. 65, no. 3, pp. 311–323, 2011.
- [3] S. Madrolle, P. Grangeat, and C. Jutten, "A Linear-Quadratic Model for the Quantification of a Mixture of Two Diluted Gases with a Single Metal Oxide Sensor," *Sensors*, vol. 18, no. 6, p. 1785, Jun. 2018.
- [4] S. Madrolle, P. Grangeat, and C. Jutten, "Dual-temperature mode for quantitative analysis of gas mixtures with MOX sensor," in *2017 ISOC/IEEE ISOEN*, Montreal, Canada, 2017.
- [5] K. Yan, D. Zhang, D. Wu, H. Wei, and G. Lu, "Design of a Breath Analysis System for Diabetes Screening and Blood Glucose Level Prediction," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 11, pp. 2787–2795, Nov. 2014.
- [6] 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 Trans. Signal Process.*, vol. 54, no. 11, pp. 4133–4145, Nov. 2006.
- [7] C. P. Robert, *The Bayesian Choice*. New York: Springer-Verlag France, Paris, 2006.
- [8] W. K. Hastings, "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, vol. 57, no. 1, pp. 97–109, Apr. 1970.
- [9] S. Madrolle, L. T. Duarte, P. Grangeat, and C. Jutten, "Supervised Bayesian source separation of nonlinear mixtures for quantitative analysis of gas mixtures," in *40th International Conference of the IEEE EMBC*, Honolulu, US, 2018, p. Accepted.
- [10] S. Hosseini, Y. Deville, L. T. Duarte, and A. Selloum, "Extending NMF to blindly separate linear-quadratic mixtures of uncorrelated sources," in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2016, pp. 1–6.