



Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration

Clément Dorffer, Matthieu Puigt, Gilles Delmaire, Gilles Roussel

► To cite this version:

Clément Dorffer, Matthieu Puigt, Gilles Delmaire, Gilles Roussel. Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration. 2017. hal-01580604v1

HAL Id: hal-01580604

<https://hal.science/hal-01580604v1>

Preprint submitted on 1 Sep 2017 (v1), last revised 20 Feb 2018 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Informed Nonnegative Matrix Factorization Methods for Mobile Sensor Network Calibration

Clément Dorffer, Matthieu Puigt, Gilles Delmaire, and Gilles Roussel

Abstract

In this paper, we consider the problem of blindly calibrating a mobile sensor network—i.e., determining the gain and the offset of each sensor—from heterogeneous observations on a defined spatial area over time. For that purpose, we propose to revisit blind sensor calibration as an informed Nonnegative Matrix Factorization (NMF) problem with missing entries. In the considered framework, one matrix factor contains the calibration structure of the sensors—and especially the values of the sensed phenomenon—while the other one contains the calibration parameters of the whole sensor network. The available information is taken into account by using a specific parameterization of the NMF problem. Moreover, we also consider additional NMF constraints which can be independently taken into account, i.e., an approximate constraint over the mean calibration parameters and a sparse approximation of the sensed phenomenon over a known dictionary. The enhancement of our proposed approaches is investigated through more than 5000 simulations and is shown to be accurate for the considered application and to outperform a multi-hop micro-calibration technique as well as a method based on low-rank matrix completion and nonnegative least squares.

Index Terms

Nonnegative matrix factorization, mobile sensor calibration, matrix completion, wireless sensor network, mobile crowdsensing

C. Dorffer, M. Puigt, G. Delmaire and G. Roussel are with Univ. Littoral Côte d'Opale, LISIC – EA 4491, F-62228 Calais, France, e-mail: {clement.dorffer,matthieu.puigt,gilles.delmaire,gilles.roussel}@univ-littoral.fr.

This work was funded by the "OSCAR" project within the Région Hauts-de-France "Chercheurs Citoyens" Program and the CNRS MASTODONS project under the grant "DoMaSQ'Air 2017".

I. INTRODUCTION

Mobile crowdsensing appears to be a novel way to get measurements from large-scale low-cost sensor networks [1], [2], [3]. It consists of acquiring geolocated and timestamped data aggregated by mobile devices—e.g., smartphones—from a crowd of volunteers along their daily moves. Mobile crowdsensing is increasingly used in participatory science—using the smartphone communication facilities to connect to additional sensor devices—and some specific challenges emerged. The first issue with mobile crowdsensing is the loss of privacy within the acquisition of geolocated data. Such an issue has been well-studied in the literature [4] and some solutions met in participatory science consist of using a trusted intermediate server—located between the mobile sensors and the recipient of the experiment—where the collected data are processed and anonymized [5].

Among the other challenges inherent to mobile crowdsensing, the sensor network calibration is a key step to make sense of the collected data. A calibrated sensor allows recovering a physical input y from the sensor output x . Calibrating a sensor consists of estimating some *calibration parameters*, provided a sensor *calibration model*. Classical calibration models are offset [6], gain [7], [8], gain/phase [9], [10], [11], gain/offset [12], [13], [14], [15], [16], [17], [18], nonlinear [19], [20], [21], or multilinear [22] response models. Such models can also be extended to deal with sensor drifts [22]. In most cases, the calibration is performed in a laboratory by inferring some known physical inputs to the sensor outputs. When applied to a large-scale network—and in particular with mobile crowdsensing where volunteers cannot be requested to regularly go to a laboratory—the calibration must be performed “on the fly” and some so-called *blind* or *self-calibration* techniques were proposed to that end. The latter can be classified into two categories, i.e., macro- and micro-calibration. On the one hand, macro-calibration methods aim to calibrate the whole sensor network in order to provide consistent sensor readings [11], [12], [13], [18], [17], [19], [20]. They usually require strong assumptions such as knowledge of the signal subspace [12], [13], [17], [19], sparse assumptions [11], or long integration time [18], [20]. On the other hand, micro-calibration methods aim to iteratively calibrate one unique sensor from the network at a given time [23], [24], [14]. To that end, they are usually exploiting measurements provided by calibrated reference sensors and thus require the sensor network to be dense enough to ensure cross-measurements between sensors.

In this paper, we consider the above crowdsensing framework with an intermediate server

where we have access to all the collected information before anonymization [5]. We revisit the self-calibration of a mobile sensor network as an informed matrix factorization problem, which may be seen as a macro-calibration technique using micro-calibration assumptions. The information provided in the factorization is due to both the considered application—which provides a specific structure in one of the factor matrices—and the fact that a few observed data are obtained using calibrated sensors. As the matrix to factorize is low rank and has missing entries, the calibration problem meets some similarities with matrix completion [25], [26] or collaborative filtering/recommender systems [27]. In particular, we aim to calibrate environmental mobile sensors providing nonnegative outputs (e.g., voltages) which are due to nonnegative inputs (e.g., gas concentrations) linked to the outputs through nonnegative calibration parameters. As a consequence, the blind calibration problem can be tackled by informed Nonnegative Matrix Factorization (NMF). We propose to that end four informed NMF techniques which were partially and preliminary proposed in [15], [16]. In [15], we were showing how to write the calibration problem in an informed NMF framework while we extended it in [16] by adding some sparse assumptions in one column of one matrix factor. In this paper, we extend our previous work as follows: (i) we incorporate a novel information in the NMF, which can be independently applied to [15] or [16], and (ii) we provide an in-depth performance analysis of the proposed methods, with a comparison with one state-of-the-art method [14].

The remainder of the paper thus reads as follows: in Section II we introduce the calibration problem together with the definitions and assumptions that we use in this paper. In Section III we propose four blind calibration methods named "IN-cal", "ACIN-cal", "SpIN-cal", and "SpAIN-cal", which are informed NMF techniques using information such as the sensor response model, the known average calibration parameter values or/and a sparse approximation of the observed phenomenon, respectively. In Section IV, we investigate the performance of the above methods and a comparison with a state-of-the-art micro-calibration technique [14]. We conclude and discuss about future work in Section V while we provide the proof of some update rules and investigate the estimation of the sensed field in Appendices A and B, respectively.

II. PROBLEM STATEMENT, DEFINITIONS AND ASSUMPTIONS

In this paper, we aim to calibrate a network composed of m geolocalized and timestamped mobile sensors. We assume that each sensor of the network provides a reading y linked to an

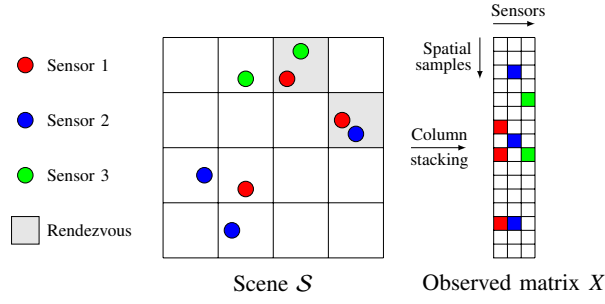


Fig. 1. From a scene S (with $n = 16$ spatial samples, $m + 1 = 3$ sensors and 2 rendezvous) to the data matrix X (white pixels mean no observed value).

input phenomenon x through a calibration function $\mathcal{F}(\cdot)$ which is assumed to be affine, i.e.,

$$x \approx \mathcal{F}(y) \approx f_1 + f_2 \cdot y, \quad (1)$$

where f_1 and f_2 are the unknown sensor parameters, i.e., its offset and gain, respectively. Calibrating the sensor network then consists of estimating the gain and the offset of each sensor. To that end, our proposed methods are exploiting the following definitions.

Definition 1 ([24]): A *rendezvous* is a temporal and spatial vicinity between two sensors.

A rendezvous is then characterized by a spatial distance Δ_d and a temporal duration Δ_t . When two sensors are in rendezvous, the fluctuations of the phenomenon between two locations closest than Δ_d during a time interval $[t, t + \Delta_t)$ are negligible. However, both highly depend on the sensed physical phenomenon [24]. As an example, if one observe the variations of temperature and of carbon monoxide concentrations, the values of Δ_d and Δ_t for the latter will be much lower than for the former [24].

Definition 2 ([15]): A *scene* S is a discretized area observed during a time interval $[t, t + \Delta_t)$. The size of the spatial pixels is set so that any couple of points inside the same pixel have a distance below Δ_d .

As shown in Fig. 1, two sensors sharing the same location of the scene are in rendezvous and then should be exposed to the same physical input. However, defining a scene with an appropriate distance Δ_d might be quite complex and we will see in this paper how to relax it.

Denoting $x_{i,j}$ as the output provided by Sensor j at Location i of S , and assuming that each of the considered $(m + 1)$ sensors of the network has sensed the whole scene, we define the $n \times (m + 1)$ data matrix $X \triangleq [x_{i,j}]_{i,j}$ such that each of its column contains the measurements

of one sensor at each location and each line contains the measurements of each sensor at one location. The affine response model (1) then yields

$$X \approx G \cdot F, \quad (2)$$

where

$$G = \begin{bmatrix} 1 & y_1 \\ \vdots & \vdots \\ 1 & y_n \end{bmatrix} \quad \text{and} \quad F = \begin{bmatrix} f_{1,1} & \cdots & f_{1,m+1} \\ f_{2,1} & \cdots & f_{2,m+1} \end{bmatrix}, \quad (3)$$

where $\forall j = 1, \dots, m+1$, $f_{1,j}$ and $f_{2,j}$ are the unknown offset and gain associated with the j -th sensor, respectively. Both factor matrices G and F thus contain the calibration model structure—hence the column of ones in G to handle the offset in the calibration function of the sensors—and the calibration parameters, respectively. Calibrating the network using factorization then consists of estimating the matrices G and F which provide the best low-rank estimation of X , while keeping the constrained structure in G .

In practical crowdsensing campaigns, the mobile sensors are usually unconstrained in their movements—as they are carried by a crowd of volunteers—and thus sparsely sense the scene. If Location i has been sensed by Sensor j then Eq. (1) is verified, otherwise the data point is unavailable, i.e., $x_{i,j}$ is unknown. Denoting Ω_X the domain where X is observed and $\mathcal{P}_{\Omega_X}(\cdot)$ the sampling operator of X , the considered calibration problem able to handle missing values reads

$$\begin{aligned} \{\tilde{G}, \tilde{F}\} &= \arg \min_{G, F} \frac{1}{2} \cdot \|\mathcal{P}_{\Omega_X}(X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t.} \quad \mathbf{g}_1 &= \mathbb{1}_{[n \times 1]}, \end{aligned} \quad (4)$$

where \mathbf{g}_1 is the first column of G , $\mathbb{1}_{[n \times 1]}$ is a column vector of ones, and $\|\cdot\|_{\mathcal{F}}$ stands for the Frobenius norm. Please note that Eq. (4) without the constraint on \mathbf{g}_1 is the formulation of matrix factorization for the low-rank matrix completion problem [25]. In practice, the latter can be solved using a weighted matrix factorization technique for example. The weighted extension of Eq. (4) then reads

$$\begin{aligned} \{\tilde{G}, \tilde{F}\} &= \arg \min_{G, F} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t.} \quad \mathbf{g}_1 &= \mathbb{1}_{[n \times 1]}, \end{aligned} \quad (5)$$

where \circ is the Hadamard product, and $W \triangleq (w_{i,j})_{i,j}$ is a properly chosen weight matrix, e.g.,

$$w_{i,j} = \begin{cases} \rho_j & \text{if } x_{i,j} \text{ is observed,} \\ 0 & \text{otherwise.} \end{cases} \quad (6)$$

The weight ρ_j can be proportional to the confidence in the accuracy of Sensor j . If all the sensors share the same confidence, one can set $\rho_j = 1$, i.e., we set W to a binary matrix. At this stage, we thus show that blindly calibrating a sensor network may be seen as a specific matrix factorization and completion problem, for which we add some additional assumptions hereafter.

As X only contains sensor outputs—which are usually voltages—it is a nonnegative matrix. Moreover, in our considered application where we aim to calibrate some environmental sensors, the vector $\mathbf{y} \triangleq [y_1, \dots, y_n]^T$ corresponds to a nonnegative physical phenomenon, e.g., a mass, a proportion (gas, dust, humidity), or a temperature (which can be expressed in Kelvin degree). As a consequence, G is also nonnegative. Lastly, many environmental sensors with affine response model—e.g., [28] for temperature/humidity or [29] for particulate matter concentration—get nonnegative calibration parameters and F can also be assumed to be nonnegative¹. At this stage, the calibration problem can then be tackled within the scope of informed NMF, i.e.,

$$\begin{aligned} \{\tilde{G}, \tilde{F}\} &= \arg \min_{G, F \geq 0} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t. } \mathbf{g}_1 &= \mathbb{1}_{[n \times 1]}. \end{aligned} \tag{7}$$

Lastly, large-scale networks composed of low-cost sensors are usually deployed in order to reinforce existing networks composed of few highly accurate calibrated sensors. Indeed, fusing the large amount of data provided by the low-cost sensors with a few accurate measurements may enhance the spatial sampling of the observed phenomenon—due to the dense deployment of the low-cost sensors over a large area—while keeping the good accuracy of the precise measurements [30]. In air quality monitoring, the measurements collected by low-cost sensors can be melt with accurate measurements provided by authoritative air quality monitoring stations for example. Such accurate sensor readings directly provide a fine estimation—considered as ground truth—of the physical phenomenon \mathbf{y} at the sensed locations. These punctually known values of the physical phenomenon can then be modelled as the measurements provided by one unique sensor—say the $(m + 1)$ -th one—with gain and offset equal to 1 and 0, respectively.

¹Please note that if F has no sign constraint, blind calibration can be revisited as a semi-NMF problem [21].

Taking such an information into Eq. (7) leads to

$$\begin{aligned}
\{\tilde{G}, \tilde{F}\} &= \arg \min_{G, F \geq 0} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\
\text{s.t. } \mathbf{g}_1 &= \mathbb{1}_{[n \times 1]}, \\
g_{i,2} &= x_{i,m+1} \quad \forall i \in \mathcal{I}, \\
\mathbf{f}_{m+1} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix},
\end{aligned} \tag{8}$$

where \mathcal{I} is the subset of locations that have been sensed by the reference sensors, $g_{i,2}$ is the i -th element of second column of G and \mathbf{f}_{m+1} is the $(m+1)$ -th column of F . Equation (8) shows that mobile sensor calibration can be tackled by NMF which (i) is using a specific structure of the matrix factors and (ii) is able to handle missing data. Let us stress that contrary to other applications of NMF such as source separation, we cannot afford to estimate F up to a scale indeterminacy. Indeed, such an indeterminacy yields relative calibration—allowing the consistency of the sensor readings only—but not the full sensor calibration. Moreover, Eq. (8) can be solved if the available data points in X are “diverse” enough and in “sufficient quantities”. As an extreme example, let us imagine that the sensed phenomenon is constant over the whole scene. Then, the columns of X are constant and the estimation of G and F is not possible, even if all the data points are available in X . As a consequence and as for any latent variable analysis technique, we need some diversity in the data points in order to perform the factorization. Similarly, let us assume that one sensor only senses once the observed phenomenon. Consequently, we cannot estimate the corresponding column of F without additional assumptions. Moreover, let us assume that the sensor network may be decomposed into two disjoint sets which do not share any sensor rendezvous between both groups. Then, one can perform the full calibration of the set which has rendezvous with authoritative sensors but only some relative calibration for the sensors of the other group.

Lastly, please note that one do not need to perfectly recover G in order to perform calibration. Indeed, the calibration parameters are only contained in F . This means that we can afford some lines of X without any available entry, provided the above assumptions in terms of matrix sampling and diversity are met. Such a scenario implies that G cannot be fully recovered while F can.

III. PROPOSED CALIBRATION METHODS

NMF finds applications in many domains such as blind source separation [31], low-rank matrix completion [26], [32], recommendation system/collaborative filtering [33], classification/clustering [34], or dictionary learning [35]. It is also increasingly used in environmental monitoring [36]. It thus has been widely studied and many methods have been proposed to perform it [37]. Some extensions also include weights [38] and/or known entries in the factor matrices [39], or use some parametric cost functions [40]. However, to the best of the authors' knowledge, using NMF as a blind calibration method has never been investigated before. In this paper, we thus propose several NMF-based calibration methods which take into account the above constraints to calibrate a mobile sensor network.

A. A natural technique based on the low-rankness of X

We here propose a natural approach to solve blind calibration, based on the low-rankness of X . Let us first assume that X is complete. According to Eq. (8), the last columns of X and G are equal. As a consequence, F can be derived from X and G , by nonnegative least squares for example. However, since many entries of X may be missing, the above approach cannot be applied directly. As a first stage, the missing values in X must be estimated. As X is low-rank, the imputation of its missing entries using low-rank matrix completion techniques provides a matrix denoted X^{comp} which reads

$$X^{\text{comp}} = \arg \min_{\tilde{X}} \text{rank}(\tilde{X}) \text{ s.t. } \mathcal{P}_{\Omega_X}(\tilde{X}) = \mathcal{P}_{\Omega_X}(X), \quad (9)$$

where $\mathcal{P}_{\Omega_X}(\cdot)$ is the sampling operator of X . As such a problem is NP-hard, it may be relaxed as

$$X^{\text{comp}} = \arg \min_{\tilde{X}} \|\tilde{X}\|_* \text{ s.t. } \mathcal{P}_{\Omega_X}(\tilde{X}) = \mathcal{P}_{\Omega_X}(X), \quad (10)$$

where $\|\cdot\|_*$ denotes the nuclear norm. In practice, many methods have been proposed to solve Eq. (10), e.g., truncated SVD [41], [42]. However, by construction, X is known to be rank-2 and can be written as the product of two nonnegative matrices, as shown in Eq. (2). Equation (9) can

thus be solved by Nonnegative Matrix Factorization/Completion (NMF/C) methods [32], [33], [26] which consist of solving²

$$\{\tilde{A}, \tilde{S}\} = \arg \min_{A, S \geq 0} \|\mathcal{P}_{\Omega_X} (X - A \cdot S)\|_{\mathcal{F}}. \quad (11)$$

using an Alternating Direction Method (ADM) [32], multiplicative updates [33], or a Nesterov gradient technique [26], respectively³. The product of the estimated matrices $\tilde{A} \cdot \tilde{S}$ then provides an estimation X^{comp} of the completed version of X which satisfies both the known rank and the nonnegative decomposition of X .

To summarize, a natural approach to perform blind calibration consists of

- 1) Estimating X^{comp} using a low-rank matrix completion technique, e.g., based on truncated SVD [42] or NMF/C [32].
- 2) Constructing \tilde{G} as

$$\tilde{G} \triangleq \begin{bmatrix} \mathbb{1}_{[n \times 1]} & \mathbf{x}_{m+1}^{\text{comp}} \end{bmatrix}, \quad (12)$$

where $\mathbf{x}_{m+1}^{\text{comp}}$ is the $(m+1)^{\text{th}}$ column—i.e., the last—of X^{comp} which corresponds to reference measurements.

- 3) Estimating \tilde{F} using nonnegative least squares by solving

$$\tilde{F} = \arg \min_{F \geq 0} \|X^{\text{comp}} - \tilde{G} \cdot F\|_{\mathcal{F}}. \quad (13)$$

Despite its nice properties, this approach—using [42] in the completion stage—was shown not to provide a good performance in our preliminary work [15], [16]. Indeed, its enhancement is subject to the completion accuracy in the very first stage. Using NMF/C techniques slightly improves its calibration abilities but the whole approach does not reach the performance of the methods proposed below.

B. Informed NMF for Sensor Calibration

In order to include the whole constraints into a new informed NMF method, we reformulate the optimization problem (8) by using the parameterization proposed in [39]. The latter consists of decomposing both the matrices G and F into the sum of their free parts—corresponding to

²It should be noticed that Eq. (11) is not similar to the considered calibration problem. Indeed, as in blind source separation, A (respectively S) can be equal to G (respectively F) up to a permutation and a scale ambiguity.

³The initialization of such methods within the framework of the considered paper is described in Section III-E.

the elements which are not subject to any constraint—and their known parts containing the fixed known values in each matrix, i.e.,

$$G = \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \quad (14)$$

and

$$F = \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F, \quad (15)$$

where

- Ω_G and Ω_F (respectively $\bar{\Omega}_G$ and $\bar{\Omega}_F$) are the binary matrices informing the locations of the known (respectively unknown) values in G and F , respectively,
- Φ_G and Φ_F contain the constrained values in G and F , respectively,
- and Δ_G and Δ_F contain the unknown values in G and F , respectively.

The binary masks Ω_G , $\bar{\Omega}_G$, Ω_F and $\bar{\Omega}_F$ are deployed in order to ensure the null intersection of the known and free parts of each matrix, i.e., $\Omega_G \circ \bar{\Omega}_G = 0$ and $\Omega_F \circ \bar{\Omega}_F = 0$ [39].

Using Paramerizations (14) and (15), the optimization problem (8) then reads

$$\begin{aligned} \{\tilde{G}, \tilde{F}\} &= \arg \min_{G, F \geq 0} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t. } G &= \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \\ F &= \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F. \end{aligned} \quad (16)$$

As Eq. (16) is non-convex with respect to both G and F , one classical strategy consists of splitting it into the following sub-problems

$$\begin{aligned} \tilde{G} &= \arg \min_{G \geq 0} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t. } G &= \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \end{aligned} \quad (17)$$

and

$$\begin{aligned} \tilde{F} &= \arg \min_{F \geq 0} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2, \\ \text{s.t. } F &= \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F, \end{aligned} \quad (18)$$

that we aim to alternately and iteratively solve. For the update of F , the authors in [39] proposed the following multiplicative update rule

$$F \leftarrow \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \circ \left[\frac{G^T \cdot (W \circ (X - G \cdot \Phi_F)^+)}{G^T \cdot (W \circ (G \cdot \Delta_F))} \right], \quad (19)$$

where the superscript $^+$ denotes the function defined as $(z)^+ \triangleq \max(\epsilon, z)$, where ϵ is a small user-defined threshold. One can note that if F has no constrained part—i.e., if Ω_F is null and if

$\bar{\Omega}_F = \mathbb{1}_{[2 \times m+1]}$ —then the whole matrix F is equal to its free part Δ_F . The multiplicative update rule (19) is then equivalent to the weighted version [38] of the well-known Lee and Seung multiplicative update rule [43].

Similarly to F , we derive the update rules associated to Eq. (17) as

$$G \leftarrow \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G \circ \left[\frac{(W \circ (X - \Phi_G \cdot F)^+) \cdot F^T}{(W \circ (\Delta_G \cdot F)) \cdot F^T} \right]. \quad (20)$$

One can note that in matrix factorization, a scale ambiguity between G and F is usually observed. It is classically solved by normalizing the lines of F or the columns of G . However, in the calibration framework that we consider in this paper, we aim to exactly recover the lines of F (and the columns of G). As a consequence, we can not arbitrarily normalize them. In our proposed calibration method, the scale ambiguity is actually tackled thanks to the calibrated sensor readings in X . The constrained column in F then scales the whole factorization problem as long as there are enough rendezvous into the data matrix X to link each mobile sensor—possibly using a path between them—to at least two distinct reference measurements⁴. If there is not enough reference measurements and/or not enough rendezvous between mobile sensors to be calibrated and references ones, then the above Informed NMF-based calibration method—denoted IN-cal below—performs the relative calibration of the network, i.e., sensors are consistent but a scale ambiguity remains. The IN-cal algorithm is summarized in Algorithm 1.

Algorithm 1 Informed NMF-based calibration (IN-cal)

Require: $G^{(0)}, F^{(0)}$

for iter = 1 **to** ITERMAX **do**

 Updating G using Eq. (20)

 Updating F using Eq. (19)

end for

C. Average-constrained extension of IN-cal

Most sensor manufacturers provide average calibration parameters associated with their sensors. While being useful, this information is not used in the above IN-cal method. Moreover, it

⁴Please note that the scale ambiguity can be solved if the number of distinct sensed entries in the last column of X is at least equal to the rank of X , which is 2 in this paper. When the rank of X is higher—as met in, e.g., [21] for blindly calibrating a sensor with a polynomial model—this number increases as well.

allows regularizing the calibration problem if the data sampling requirements are not satisfied for some sensors, i.e., if they do not have enough rendezvous, or if no reference sensor readings are available. Lastly, it should be noticed that such an information has been used in [11] for a compressed-sensing calibration problem, or in [18], [6] for moment-based mobile sensor calibration methods. However, we use it in a slightly different way, as we explain hereafter.

In the affine response model that we are considering in this paper, the average calibration parameters are provided by the sensor manufacturers as a mean offset \bar{f}_1 and a mean gain \bar{f}_2 . In practice, these values are approximately equal to the average calibration parameters to be estimated, i.e.,

$$\frac{1}{m} \cdot \sum_{i=1}^m f_{1,i} \approx \bar{f}_1 \quad \text{and} \quad \frac{1}{m} \cdot \sum_{i=1}^m f_{2,i} \approx \bar{f}_2 \quad . \quad (21)$$

By defining $\bar{\mathbf{f}} \triangleq [\bar{f}_1, \bar{f}_2]^T$, Eq. (21) then reads

$$\frac{1}{m} \cdot F \cdot \begin{bmatrix} \mathbb{1}_{[m \times 1]} \\ 0 \end{bmatrix} \approx \bar{\mathbf{f}} \quad , \quad (22)$$

where $\begin{bmatrix} \mathbb{1}_{[m \times 1]} \\ 0 \end{bmatrix}$ is the $(m+1) \times 1$ vector composed of a column of ones concatenated with a zero, in order not to include the last column of F to the mean computation. One can note that by using the parameterization (15)—and by noticing that the last column of $\bar{\Omega}_F$ is null—this approximation reads

$$\frac{1}{m} \cdot (\bar{\Omega}_F \circ F) \cdot \mathbb{1}_{[(m+1) \times 1]} \approx \bar{\mathbf{f}} \quad , \quad (23)$$

which can then be included into our previous calibration approach. Indeed, it appears as an extra-penalization term in Eq. (16), i.e.,

$$\begin{aligned} \{\tilde{G}, \tilde{F}\} = \arg \min_{G, F \geq 0} & \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2 + \frac{\mu}{2} \cdot \left\| \frac{1}{m} \cdot (\bar{\Omega}_F \circ F) \cdot \mathbb{1}_{[(m+1) \times 1]} - \bar{\mathbf{f}} \right\|_{\mathcal{F}}^2, \\ \text{s.t.} \quad & G = \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \end{aligned} \quad (24)$$

$$F = \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F,$$

where μ is a user-defined weight. As for IN-cal, Eq. (24) may be decomposed into two sub-problems. Actually, the sub-problem in G is the same as IN-cal and Eq. (20) still applies. However, the average constraint affects the sub-problem relative to F , i.e.,

$$\begin{aligned} \tilde{F} = \arg \min_{F \geq 0} & \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2 + \frac{\mu}{2} \cdot \left\| \frac{1}{m} \cdot (\bar{\Omega}_F \circ F) \cdot \mathbb{1}_{[(m+1) \times 1]} - \bar{\mathbf{f}} \right\|_{\mathcal{F}}^2, \\ \text{s.t.} \quad & F = \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F. \end{aligned} \quad (25)$$

It should be noticed that normalizing the rows of F is classical in matrix factorization and is usually performed after each update of F [34] or within the optimization problem [44]. Informed NMF with normalization was also investigated in [45], [46]. However, such a normalization implies rows of F to be exactly equal to the chosen means, which may lead to calibration errors in the case of an inexact knowledge of the mean parameters in the considered application. Such an estimation error is also a limitation of any blind calibration technique not using reference sensors as ground truth, e.g., [11]. By considering the norm constraint as a penalization term such as Eq. (25), then one can change the penalization weight depending whether or not the sampling and diversity assumptions in X are satisfied to ensure calibration with the above IN-cal method.

We derive in Appendix A the multiplicative update rule from Eq. (25), which reads

$$\tilde{F} \leftarrow \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \circ B, \quad (26)$$

where

$$B = \frac{\bar{\Omega}_F \circ \left(G^T \left(W^2 \circ X_F \right) \right) + \frac{\mu}{m} \cdot \text{diag} \left(\bar{\mathbf{f}} \right) \cdot \bar{\Omega}_F}{\bar{\Omega}_F \circ \left(G^T \left(W^{\circ 2} \circ \left(G \cdot \left(\bar{\Omega}_F \circ \Delta_F \right) \right) \right) + \frac{\mu}{m^2} \cdot \text{diag} \left(\left(\bar{\Omega}_F \circ \Delta_F \right) \cdot \mathbb{1}_{[(m+1) \times 1]} \right) \cdot \bar{\Omega}_F \right)}. \quad (27)$$

Equations (26) and (27) take into account the Average Constraint in IN-cal, hence the name denoted ACIN-cal of this extension in the remainder of the paper. Its structure is summarized in Algorithm 2.

Algorithm 2 Average Constrained extension of IN-cal (ACIN-cal)

Require: $G^{(0)}$, $F^{(0)}$, $\bar{\mathbf{f}}$, μ

for iter = 1 **to** ITERMAX **do**

 Updating G using Eq. (20)

 Updating F using Eqs. (26) and (27)

end for

D. Sparsity-constrained extension of IN-cal

In order to provide a good performance, the IN-cal method needs enough rendezvous between sensors, which is not necessarily satisfied in practice. Moreover, properly discretizing the scene \mathcal{S} may be an issue, even for an expert.

In this section, we thus propose to regularize the IN-cal method by including one more constraint on the physical input phenomenon \mathbf{y} , i.e., the second column of G . Indeed, we assume that \mathbf{y} allows a sparse decomposition according to a known dictionary \mathcal{D} , i.e.,

$$\exists \mathbf{a} \in \text{Sp}(\mathbb{R}^p), \text{ s.t. } \mathbf{y} \approx \mathcal{D} \cdot \mathbf{a}, \quad (28)$$

where p is the number of atoms in \mathcal{D} and $\text{Sp}(\mathbb{R}^p)$ denotes the space of sparse vectors from \mathbb{R}^p , i.e., $\text{Sp}(\mathbb{R}^p) \triangleq \{\mathbf{u} \in \mathbb{R}^p, \text{ s.t. } \|\mathbf{u}\|_0 \ll p\}$, where $\|\mathbf{u}\|_0$ stands for the number of non-zeros elements in \mathbf{u} . This extra-constraint then leads to a new Sparsity-constrained extension of IN-cal, denoted SpIN-cal. In environment monitoring, the dictionary \mathcal{D} can be learned from realizations of a physical model—e.g., a Gaussian plume diffusion model [47]—which provides patterns of the observed signal \mathbf{y} . The sparse decomposition of the observed phenomenon according to the dictionary means that the observed physical phenomenon has to be a combination of a few of these known patterns.

Including the constraint (28) into the factorization problem (16) leads to a novel problem with a third unknown, i.e.,

$$\begin{aligned} \{\tilde{G}, \tilde{F}, \mathbf{a}\} &= \arg \min_{G \geq 0, F \geq 0, \mathbf{a}} \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2 + \frac{\lambda}{2} \cdot \|\mathbf{g}_2 - \mathcal{D} \cdot \mathbf{a}\|_{\mathcal{F}}^2, \\ \text{s.t. } G &= \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \\ F &= \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F, \\ \|\mathbf{a}\|_0 &\ll p, \end{aligned} \quad (29)$$

where λ is the positive parameter which controls the strength of the sparsity of the contribution vector \mathbf{a} . This problem differs from classical sparse NMF [48], [49], [50], [51] which consists of estimating a sparse matrix factor. On the contrary, we here look for a dense matrix with a sparse decomposition, which meets some similarities with [52]. So as for the IN-cal and ACIN-cal optimization problems, the cost function in Eq. (29) is non-convex with respect to both G , F , and \mathbf{a} . Moreover, as the sparse decomposition constraint in Eq. (29) only influence the update of G , the subproblems that we have to care about read,

$$\begin{aligned} \tilde{G} &= \arg \min_G \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2 + \frac{\lambda}{2} \cdot \|\mathbf{g}_2 - \mathcal{D} \cdot \mathbf{a}\|_{\mathcal{F}}^2, \\ \text{s.t. } G &= \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G, \end{aligned} \quad (30)$$

and

$$\tilde{\mathbf{a}} = \arg \min_{\mathbf{a}} \frac{\lambda}{2} \cdot \|\mathbf{g}_2 - \mathcal{D} \cdot \mathbf{a}\|_{\mathcal{F}}^2 \quad \text{s.t.} \quad \|\mathbf{a}\|_0 \ll p. \quad (31)$$

One can note that the update rule (19) still applies here, as the sparsity constraint does not impact F . Noticing that

$$\mathbf{g}_2 = G \cdot [0, 1]^T, \quad (32)$$

the penalization term in Eq. (30) can be rewritten as

$$\frac{\lambda}{2} \cdot \|\mathbf{g}_2 - \mathcal{D} \cdot \mathbf{a}\|_{\mathcal{F}}^2 = \frac{1}{2} \cdot \left\| \left(\sqrt{\lambda} \cdot \mathbb{1}_{[n \times 1]} \right) \circ \left(\mathcal{D} \cdot \mathbf{a} - G \cdot [0, 1]^T \right) \right\|_{\mathcal{F}}^2. \quad (33)$$

By defining

$$\mathbf{W} \triangleq [W, \sqrt{\lambda} \cdot \mathbb{1}_{[n \times 1]}], \mathbf{X} \triangleq [X, \mathcal{D} \cdot \mathbf{a}], \text{ and } \mathbf{F} \triangleq \left[F, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right], \quad (34)$$

it is straightforward to show that

$$\begin{aligned} & \frac{1}{2} \cdot \|\mathbf{W} \circ (\mathbf{X} - G \cdot \mathbf{F})\|_{\mathcal{F}}^2 \\ &= \frac{1}{2} \cdot \|W \circ (X - G \cdot F)\|_{\mathcal{F}}^2 + \frac{\lambda}{2} \cdot \|\mathbf{g}_2 - \mathcal{D} \cdot \mathbf{a}\|_{\mathcal{F}}^2, \end{aligned} \quad (35)$$

and the optimization problem (30) then reads

$$\begin{aligned} \tilde{G} &\approx \arg \min_{G \geq 0} \frac{1}{2} \cdot \|\mathbf{W} \circ (\mathbf{X} - G \cdot \mathbf{F})\|_{\mathcal{F}}^2, \\ \text{s.t. } G &= \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G. \end{aligned} \quad (36)$$

The update rule (20) thus applies by replacing W , X , and F by \mathbf{W} , \mathbf{X} , and \mathbf{F} , respectively.

Lastly, for the update of \mathbf{a} , the optimization problem (31) provides a sparse approximation of \mathbf{g}_2 and can be tackled in several ways. As is, Problem (31) is NP-hard and one can consider its convex relaxation instead, i.e.,

$$\mathbf{a} = \arg \min_{\tilde{\mathbf{a}}} \|\mathbf{g}_2 - \mathcal{D} \cdot \tilde{\mathbf{a}}\|_{\mathcal{F}}^2 + \nu \cdot \|\tilde{\mathbf{a}}\|_1, \quad (37)$$

where ν is a user-defined threshold balancing the sparsity level. The above equation is the exact formulation of the LASSO—or equivalently Basis Pursuit Denoising—problem which has been extensively investigated since [53] and for which many algorithms may be found in the literature [50]. However, the LASSO algorithms are usually sensitive to the choice of ν that has to be carefully set [50]. As an alternative, such a problem can also be tackled using greedy algorithms like Orthogonal Matching Pursuit (OMP) [54] for example. OMP iteratively incorporates atoms into the sparse decomposition up to a maximum number p of selected atoms or a given approximation error. Since we found OMP to provide a good performance in preliminary

tests, we stick with this method in this paper but SpIN-cal might use a more recent sparse approximation technique if needed.

One main interest of SpIN-cal is that it relaxes the rendezvous requirement needed in IN-cal and ACIN-cal. Indeed, by using the dictionary, sensors are in *virtual rendezvous* as long as they share the same atom support. The *exact* rendezvous assumption in IN-cal can thus be relaxed provided some *virtual* rendezvous still exist. This implies that the Scene discretization can be refined—i.e., Δ_d can be set to a smaller value—without affecting the calibration enhancement. Moreover, this allows estimating the whole scene even when no sensor senses a given location, as shown in Appendix B. Moreover, we noticed in our preliminary work [16] that SpIN-cal needed much less iterations than IN-cal to converge. The structure of SpIN-cal may be summarized in Algorithm 3.

Algorithm 3 Sparsity constrained extension of IN-cal (SpIN-cal)

Require: $G^{(0)}$, $F^{(0)}$, \mathcal{D} , λ

for iter = 1 **to** ITERMAX **do**

 Updating \mathbf{a} using OMP [54] to solve Eq. (31)

 Constructing \mathbf{W} , \mathbf{X} , and \mathbf{F} defined in Eq. (34)

 Updating G using Eq. (20) with \mathbf{W} , \mathbf{X} , and \mathbf{F} instead of W , X , and F , respectively

 Updating F using Eq. (19)

end for

Please note that, as SpIN-cal and ACIN-cal extend IN-cal by adding extra-information in G and F , respectively, it is straightforward to derive an extension of IN-cal, denoted SpAIN-cal, which combines the sparse assumption on G and the average constraint on F , as shown in Algorithm 4.

In the remainder of this paper, we investigate the performance of the low-rank-based approach using NMF/C [32], and of the IN-cal, ACIN-cal, and SpIN-cal methods with respect to different parameter values.

E. Discussion

In this subsection, we discuss some potential issues met with the above methods. First of all, it should be noticed that these approaches require the matrices G and F to be initialized in order to perform the updates. NMF initialization is known to be tricky, and classical strategies consist of

Algorithm 4 Sparsity and Average constrained extension of IN-cal (SpAIN-cal)

Require: $G^{(0)}$, $F^{(0)}$, \mathcal{D} , λ , $\bar{\mathbf{f}}$, μ
for iter = 1 **to** ITERMAX **do**

 Updating \mathbf{a} using OMP [54] to solve Eq. (31)

 Constructing \mathbf{W} , \mathbf{X} , and \mathbf{F} defined in Eq. (34)

 Updating G using Eq. (20) with \mathbf{W} , \mathbf{X} , and \mathbf{F} instead of W , X , and F , respectively

 Updating F using Eqs. (26) and (27)

end for

a random initialization while some authors propose an initialization provided by experts [39], the output of another factorization method [55], or of a physical model [56]. In our previous work [15], [16], we used the natural method proposed in Section III-A as the input of the proposed IN-cal and SpIN-cal methods, respectively. The low-rank completion stage was using the method proposed in [42]. Unfortunately, such an initialization was time consuming and we replaced the truncated SVD-based completion [42] by NMF/C [32]. However, as any NMF technique, NMF/C must be initialized. This is done as follows: we randomly initialize F around the theoretical values of the calibration parameters $\bar{\mathbf{f}}$ while we initialize G as follows. Its first column only contains some ones while the second column \mathbf{g}_2 is randomly generated. The eventual negative entries of both matrices G and F are then projected to a small non-negative threshold, in order to ensure the nonnegativity of the whole matrices. The outputs of NMF/C with such an initialization were then used to initialize the proposed informed NMF-based calibration methods. We finally noticed in some preliminary tests that initializing the latter with the above random strategy provided the same performance than using the output of the natural calibration method using NMF/C. As a consequence, in this paper, we stick with the above random initialization strategy for the blind calibration techniques proposed in this paper.

Another issue is the presence of the user-defined thresholds λ and μ in the proposed ACIN-cal, SpIN-cal, and SpAIN-cal methods. Indeed, the calibration performance of these methods might be linked to the choice of the values of such parameters. Even if we investigate their influence on the enhancement of the proposed methods in Section IV, their optimization is out of the scope of the paper. However, we here discuss some strategies possibly well-suited to that purpose. A first technique might consist of (i) deriving multiple factor matrices with different values of μ and λ ,

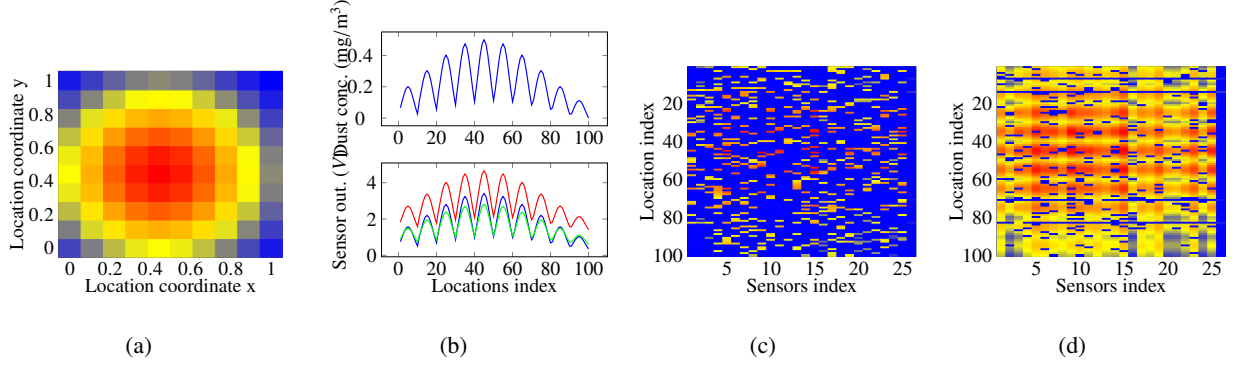


Fig. 2. (a) The 10×10 simulated scene \mathcal{S} (pollution field); (b) (top) the vectorized scene \mathbf{y} , (bottom) three vectorized sensor readings of \mathbf{y} ; (c-d): simulated dataset with 25 mobile sensor and 4 reference measurements (in the last column of \mathbf{X}) with (c) $\rho_{MV} = 0.8$ and $\rho_{RV} = 0.9$; (d) $\rho_{MV} = 0.1$ and $\rho_{RV} = 0.1$.

and (ii) providing actual solutions using *late fusion*, e.g., clustering the outputs of the multiple runs. Late fusion is classically met with source separation [57] or multiple source localization [58], when the same parameters are estimated several times. However, such a strategy may degrade the calibration performance if some of the tested values of the user-defined thresholds yield poor estimates of F . As an alternative, one can notice that the different terms of the cost functions in Eqs. (25) and (29) might not be simultaneously optimized. As a consequence, no unique solution might exist to solve them and one might consider *multi-objective optimization* techniques to find the set of the best parameter values (in the sense of a Pareto front) instead [59]. Combining multi-objective optimization with NMF was recently proposed in [60], [61] and similar strategies might be used with the proposed techniques as well.

IV. EXPERIMENTAL VALIDATION

In this section, we investigate the performance of the proposed calibration methods. For that purpose, we simulate a crowdsensing-like particulate matter sensing during a time interval $[t, t + \Delta_t)$, which satisfies the assumptions in Section II. The scene \mathcal{S} is a 10×10 discretized area—as shown in Fig. 2(a)—which is observed by $m+1 = 26$ sensors, i.e., m uncalibrated and mobile dust sensors [29] connected to mobile devices and one calibrated, high quality, and mobile sensor⁵

Once the scene is unfolded, \mathbf{y} is a vector of length $n = 100$. The observed concentrations in \mathbf{y} range between 0 and 0.5 mg/m^3 , for which the sensor response is assumed to be affine

⁵Actually and as discussed above, we get k fixed, calibrated, and accurate sensors whose obtained values are modeled as those of the $(m+1)$ -th sensor in the mobile sensor calibration problem.

[29]. For each uncalibrated sensor, each observed data point represents a nonnegative voltage linked to the corresponding ground truth point in \mathbf{y} according to Eq. (1). In particular, following the datasheet in [29], the offset and gain coefficients are randomly set according to a Gaussian distribution centered around $\tilde{f}_1 = 0.9$ V and $\tilde{f}_2 = 5$ V/(mg/m³), respectively, and then projected onto their respective interval of admissible values—provided by the manufacturer [29]—i.e., $0 < f_{1,j} < 1.5$ and $3.5 < f_{2,j} < 6.5$, $\forall j = 1, \dots, m$. We then get a 26×100 theoretical observation matrix for which we randomly keep $k + l$ samples in X only, where k (respectively, l) is the number of calibrated (respectively, uncalibrated) sensor samples—with $k \ll l$ —hence providing the irregular spatial sampling over the scene. Moreover, we remove data in X so that each uncalibrated sensor is in rendezvous at most one time with the reference sensor, which is a strong condition as discussed in Section IV-A. Indeed, the performance of the proposed blind calibration methods not only depends on the number of sensed values in X but also on the number of rendezvous between uncalibrated and calibrated sensors. We denote ρ_{MV} and ρ_{RV} the proportions of missing values in X , and of rendezvous between the uncalibrated sensors and the calibrated one, respectively. The resulting partially observed matrix X can then be plotted, according to the values of these parameters, as shown in Figs. 2(c) and 2(d) for example. Moreover, Gaussian noise realizations may be added to the observed uncalibrated sensor data and the weight coefficients ρ_j defined in Eq. (6) are set to $\rho_j = 1$. Lastly, we construct a dictionary \mathcal{D} with $p = 62$ atoms, such that \mathcal{S} can be decomposed using only 2 atoms of \mathcal{D} , i.e., \mathbf{y} is 2-sparse.

Let us stress that in all the tests performed in this paper, we use the same fixed scene \mathcal{S} . However, for each value of each tested parameter, i.e., the simulation of the sensor calibration parameters, the proportions ρ_{MV} and ρ_{RV} , and the input signal-to-noise ratio (SNR), we repeat 40 times the experiment for the sake of statistical significance. Moreover, except when we provide another criterion, the calibration accuracy is measured using the root mean-square error between the m estimated $\tilde{f}_{2,j}$ and actual $f_{2,j}$ uncalibrated gains⁶, i.e.,

$$\text{RMSE}(\mathbf{f}_2, \tilde{\mathbf{f}}_2) = \sqrt{\frac{\sum_{j=1}^m (f_{2,j} - \tilde{f}_{2,j})^2}{m}}. \quad (38)$$

⁶RMSEs computed over the estimated offsets, not plotted in the paper for space considerations, provide a similar performance as the gain RMSEs.

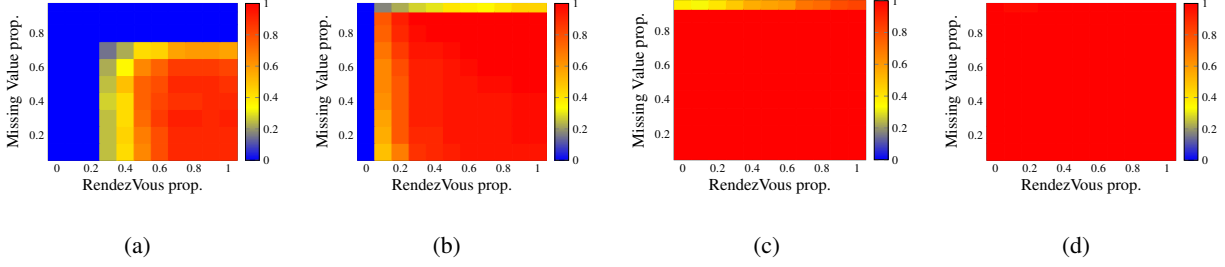


Fig. 3. Efficient calibration proportion according to the rendezvous proportion and the missing value proportion, blue color means no efficient calibration within the trials while red means perfect calibration for all trials with Fig. 3(a): NMFC-cal; Fig. 3(b): IN-cal; Fig. 3(c) ACIN-cal; Fig. 3(d): SpIN-cal.

A. Global performance of the proposed methods

In this subsection, we aim to characterize the enhancement provided by the proposed calibration methods, i.e., the “natural” approach based on the low-rankness of X using the NMF/C method to perform the completion of X , and the IN-cal, ACIN-cal, and SpIN-cal methods.

In a first series of tests, we make ρ_{MV} and ρ_{RV} vary from 0% to 95%, and from 0% to 90%, respectively. For each pair of parameters $(\rho_{MV}; \rho_{RV})$, we generate 40 simulations for which we run the above methods during $5 \cdot 10^5$ iterations after the same initialization. In this experiment, we do not add additive noise, the values of μ and λ in Eqs. (25) and (29) are set to 10 and 10, respectively. Their influence on the calibration performance is investigated in Sect IV-C. It should be noticed that the approximate mean calibrations parameters are set to the exact mean parameters, which provides an upper bound of the expected performance of ACIN-cal. Actually, such a perfect knowledge was also used in [11] for a compressed sensing calibration problem. As explained above, we discuss the effects of the error of estimation of the mean calibration values in Section IV-C. Inspired by the phase transition diagrams that are commonly used in matrix completion, we count the proportions of cases when the achieved RMSEs are below 10^{-7} for each tested value of $(\rho_{MV}; \rho_{RV})$ and for each tested methods. The obtained diagrams are plot in Figs. 3(a) to 3(d).

The natural approach often succeeds in calibrating the network when the network is dense, i.e., when the proportion ρ_{MV} of missing values is below 50% and the proportion ρ_{RV} of rendezvous between calibrated and uncalibrated sensors is above 60%. Within these conditions, the proportion of calibration success is between 80 and 100%, as shown in Fig. 3(a). When ρ_{MV} increases and/or when ρ_{RV} decreases, the method completely fails as it never reaches an RMSE below 10^{-7} when

$$\rho_{MV} \geq 0.8 \text{ or } \rho_{RV} \leq 0.2.$$

As we can see in Fig. 3(b), the IN-cal method is much more versatile than the previous technique as it almost always succeed in calibrating the network, except when there is no rendezvous between uncalibrated and calibrated sensors or when the proportion ρ_{RV} of missing values is equal to 95%. This shows the relevance of the proposed approach with its specific parameterization. The loss of performance for the above “extreme” values of ρ_{MV} and ρ_{RV} was expected, as the assumptions needed by the method were not satisfied anymore.

The plot in Fig. 3(c) shows that ACIN-cal takes advantage of the average values of the calibration parameters to regularize the calibration. In particular, normalizing the rows of F provides some robustness to the calibration performance when there is no rendezvous between calibrated and uncalibrated sensors. Let us stress again that these diagrams are obtained when the average calibration parameters are perfectly estimated. Their poor estimation should provide a much lower performance, as discussed in Section IV-C.

Lastly, as shown in Fig. 3(d), the SpIN-cal method outperforms all the above techniques as it always succeeds in performing calibration in all the tests. This was expected as the dictionary provides a relaxed condition in terms of rendezvous which regularizes the problem. The SpAIN-cal technique provides the same performance as SpIN-cal on these tests and the corresponding diagram is not shown for the sake of readability.

We now deeper investigate the sensitivity of our proposed methods to (i) the missing value proportion in the data matrix X —due to the network sampling—(ii) the proportion of mobile sensors having a rendezvous with a reference, and (iii) the input SNR. Except when we make these parameters vary, we set the proportions ρ_{MV} and ρ_{RV} of missing values and of rendezvous between calibrated and uncalibrated sensors to 90% and to 30%, respectively. Moreover, except when we test the influence of noise, we do not add noise in the simulated matrices. Each experiment is repeated 40 times and all the methods are run after the same initialization—as explained in Section III-E—and stopped after $5 \cdot 10^5$ iterations. We evaluate their performance with the RMSE criterion (38).

1) Influence of the missing value proportion: In order to better characterize the impact of the missing value proportion on our calibration methods, we make the missing value proportion ρ_{MV} range between 0 and 95%.

Figures 4(a) to 4(e) show the achieved enhancement they provide. The plain curves and the blue areas represent the median RMSEs and the spread of the reached RMSEs obtained for each

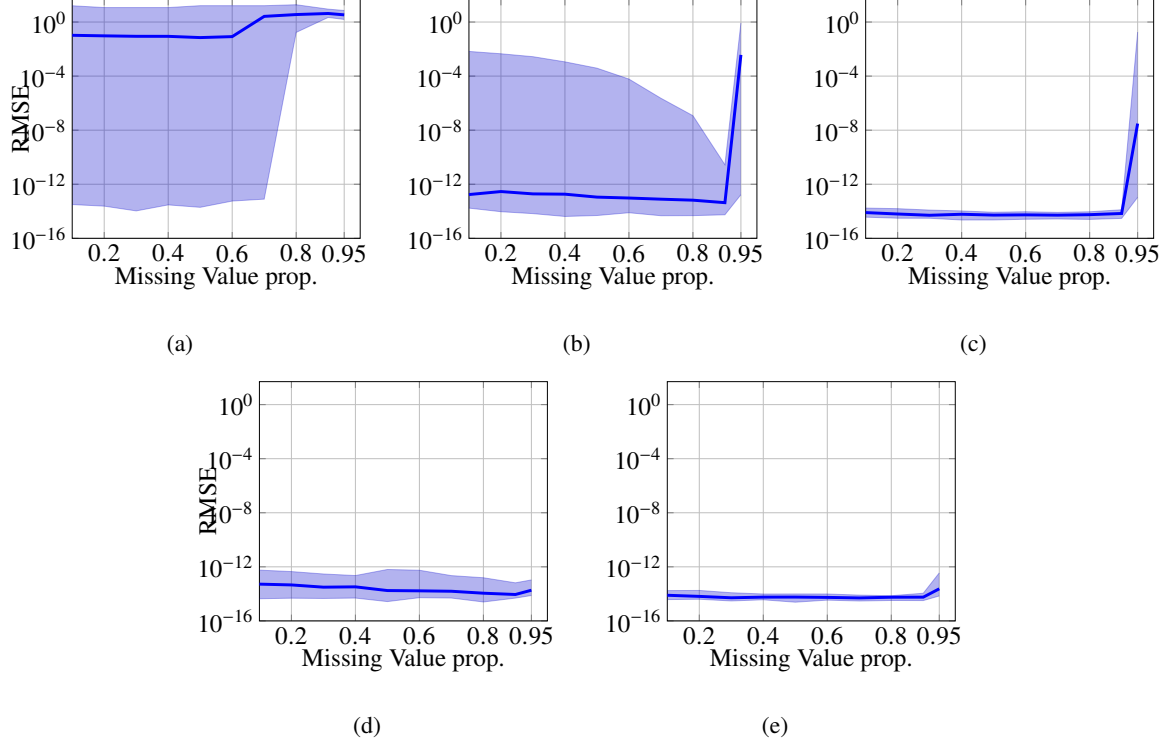


Fig. 4. Evolution of the RMSE with respect to the missing value proportion: (a) Natural approach; (b) IN-cal; (c) ACIN-cal; (d) SpIN-cal; (e) SpAIN-cal.

parameter value. One can note that the completion-based calibration is not providing satisfying results—neither with the median, nor with the envelope—and is almost always outperformed by all the informed NMF-based calibration methods. The IN-cal technique is performing well from $\rho_{MV} = 0\%$ to $\rho_{MV} = 90\%$ with median RMSEs below 10^{-12} . The large spread of the envelope is due to the low convergence rate of the method in some cases. It should be noticed that increasing the number of iteration highly reduces the spread, as shown in [15] where IN-cal is run with 10^6 iterations instead of $5 \cdot 10^5$ as in this paper. For $\rho_{MV} = 95\%$, the poor calibration performance is due to the lack of data in X , which makes some uncalibrated sensors not being connected with calibrated ones, even through a multiple path. As a consequence, the estimation of F is done up to a scale ambiguity, as discussed in Section II.

ACIN-cal provides a slightly lower median calibration error than IN-cal and a much lower spread. When the proportion of missing values is high, ACIN-cal provides a much higher calibration error than with lower values of ρ_{MV} (but still much lower than IN-cal in the same configuration). These results correspond to our expectations, as the average-based regularization

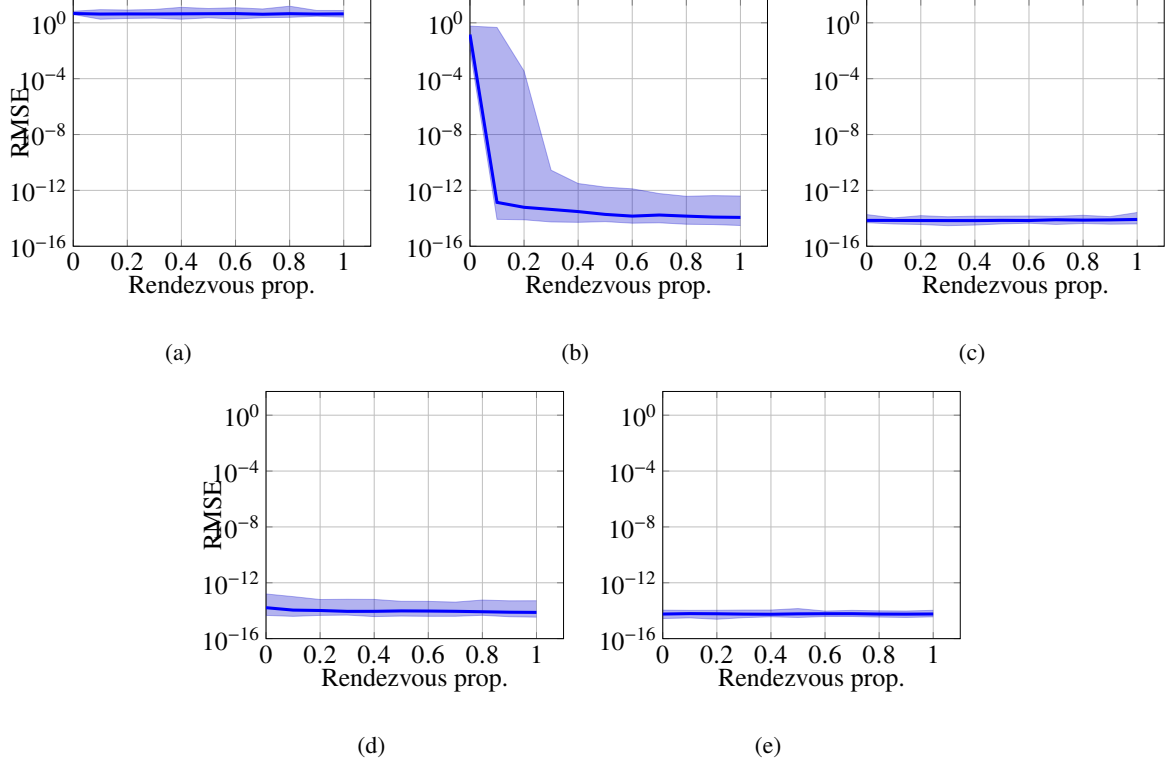


Fig. 5. Evolution of the RMSE with respect to the input rendezvous proportion: (a) Natural approach; (b) IN-cal; (c) ACIN-cal; (d) SpIN-cal; (e) SpAIN-cal.

in ACIN-cal was designed to cure calibration problems due to unavailable reference measurements but not those due to some high proportions of missing values. However, one can note that this regularization also speeds up the NMF convergence, as the envelope is very close to the median RMSE for parameters ρ_{MV} below 95%.

As expected, the SpIN-cal technique is providing better results for all tested proportions of missing value. The upper limit of the envelope is much lower than the one associated with IN-cal, thanks to the regularization which speeds up the NMF convergence, as shown in [16]. Lastly, Fig. 4(e) shows that SpAIN-cal combines the "best of" SpIN-cal and ACIN-cal, as the median performance is almost similar to SpIN-cal while the envelope is much reduced, as in ACIN-cal.

2) *Effect of the proportion of rendezvous between calibrated and uncalibrated sensors:* In this experiment, we make ρ_{RV} range between 0 to 100%. The obtained performance with respect to the rendezvous proportion ρ_{RV} is shown in Figs. 5(a) to 5(e). As for the missing value proportion, all proposed informed NMF methods almost always outperform the natural approach

based on low-rankness using NMF/C. IN-cal provides a high performance as long as there exist rendezvous between calibrated and uncalibrated sensors. When the proportion ρ_{RV} reduces to zero, the classical NMF scale ambiguity cannot be fixed, thus leading to a drastic drop of calibration enhancement. However, let us stress again that even if there is no rendezvous between uncalibrated and reference sensors, IN-cal still performs relative calibration, which means that the parameters of each sensor are well estimated up to a common scale ambiguity.

ACIN-cal thus provides a better performance than IN-cal, as it uses the knowledge of the average calibration parameters to reduce the calibration error when relative calibration is performed with IN-cal. While the RMSEs shown on Fig. 5(c) are almost always the lowest achieved by all the tested methods, let us stress again that the reached performance may be lowered in the case of a coarser estimation of the mean parameters, as studied in Section IV-C.

Lastly, the performance achieved by both SpIN-cal and SpAIN-cal is constant regardless of the rendezvous proportion. Indeed and as explained before, the strong rendezvous assumption needed in IN-cal is not needed anymore by SpIN-cal and SpAIN-cal, since two sensors are in rendezvous if they share the same atom support. The dictionary thus makes these methods robust to the proportion of rendezvous between mobile and reference sensors. However and as noticed above, the average information allows to reduce the spread of the reached RMSEs in SpAIN-cal with respect to SpIN-cal.

3) *Influence of the additive noise:* To study the sensitivity of our proposed methods against the input SNR, we make the input SNR vary from ∞ (no noise) to 0 dB by adding truncated realizations of Gaussian noise to the data matrix X . The noise realizations are truncated in order to keep nonnegative the entries of X .

The results obtained with respect to the input SNR are shown in Figs. 6(a) to 6(e) and are quite similar for all the proposed informed NMF-based calibration methods, except for the lowest tested input SNRs. Indeed, their median performance linearly decreases with the input SNR. One might expect SpIN-cal and SpAIN-cal to provide better results than the reached ones. However, it should be noticed that only G is estimated using the dictionary. This implies that the noise present in the data still affects the estimation of F , which is mainly investigated in this paper. Moreover, when the input SNR is 0 dB, the maximum RMSE reached with SpIN-cal is above 10^5 , while all the other methods have a spread (much) below this value. In a few cases, the atoms selected in the sparse decomposition did not correspond to the actual ones, which provide a poor calibration enhancement. On the contrary, the additional average information allows to

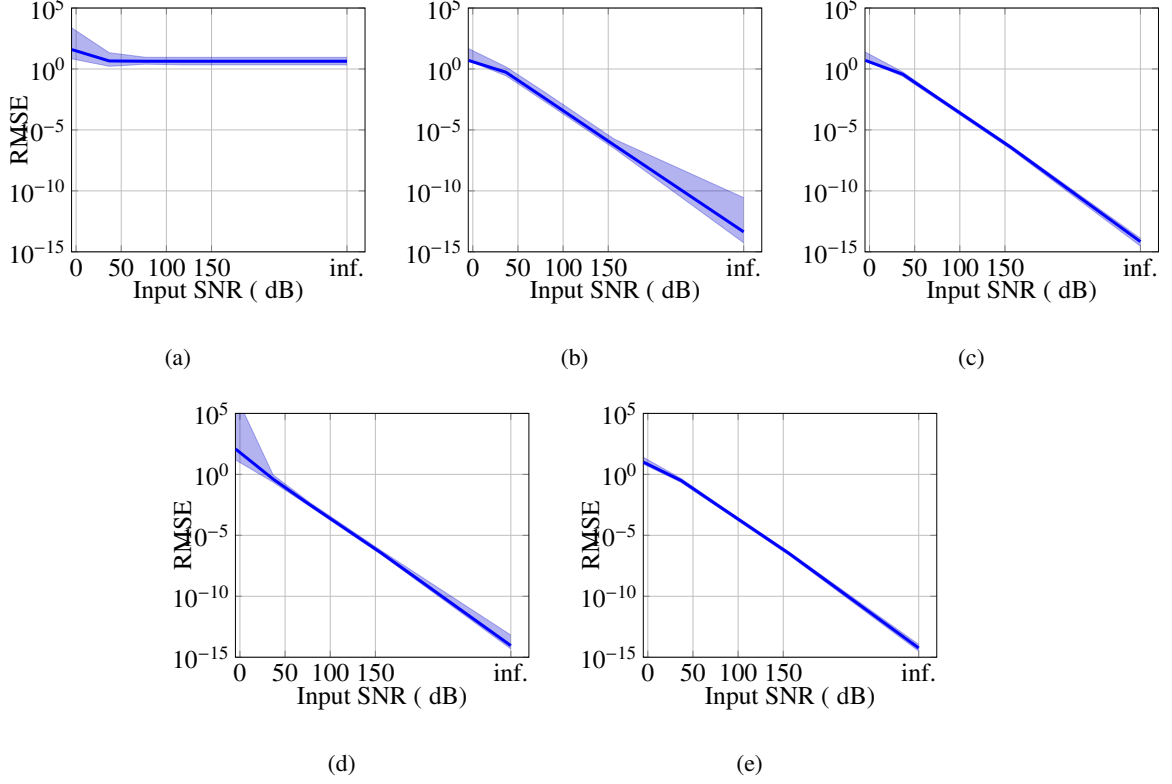


Fig. 6. Evolution of the RMSE with respect to the input SNR: (a) Natural approach; (b) IN-cal; (c) ACIN-cal; (d) SpIN-cal; (e) SpAIN-cal.

reduce this issue in SpAIN-cal. Denoising the data in X using the dictionary is out of the scope of this paper and is let for future work.

B. Comparison with a multi-hop calibration method

In this section, we propose to compare our calibration methods with the multi-hop technique proposed in [14] which is a micro-calibration technique exploiting the rendezvous definition used in our proposed IN-cal method, on which are based our methods.

This state-of-the-art approach consists of sequentially calibrating each mobile sensor of the network, using their rendezvous with previously calibrated sensors. This method basically consists of

- 1) selecting a mobile sensor to be calibrated,
- 2) calibrating this sensor by inferring its measurements in rendezvous with readings provided by the previously calibrated mobile sensor or a reference sensor,
- 3) repeating this operation until each mobile sensor is calibrated.

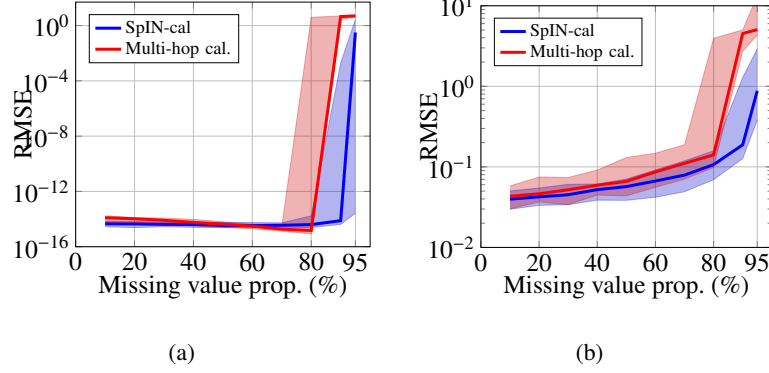


Fig. 7. Comparison of SpIN-cal with multi-hop calibration [14] with respect to the proportion of missing values in (a) noiseless and (b) noisy scenarii.

This "multi-hop" calibration was designed for a mobile crowdsensing application when the sensors are carried by tramways, thus always following the same paths and ensuring a very large number of rendezvous, even with authoritative sensors. In such a configuration, this method was shown to provide a good performance. However, when carried by volunteers, the sensor network might not be that dense and the multiple rendezvous of a mobile sensor with a reference one is not guaranteed. This is for example the case in the tests reported above, where each mobile sensor is in rendezvous at most once with the calibrated sensor. We thus propose to compare the performance of both the multi-hop and our proposed SpIN-cal methods on new simulated data satisfying the requirements of [14]. These new simulations consist of randomly removing data in X to match a desired missing value proportion ρ_{MV} that we make vary from 0 to 95%. We consider both noiseless and noisy scenarios and we do not pay attention to the rendezvous proportion and the number of reference measurements in the data. Such an experiment is repeated 40 times. Figs. 7(a) and 7(b) show the median and the spread of the RMSEs—as defined in Eq. (38)—reached by the multi-hop (in red) and the SpIN-cal (in blue) methods, without and with additive noise in the observed data matrix X (with an input SNR around 30 dB), respectively. In Fig. 7(a), both methods provide a similar performance when the number of missing values in X is low, i.e., $\rho_{MV} \leq 0.8$. However, when ρ_{MV} is higher, our proposed calibration approach outperforms [14], which is expected as the minimum number of rendezvous between the calibrated sensor and any uncalibrated one might be below 2 in some experiments. For more than 90% of missing values, both methods fail, due to the cases where at least one sensor is providing less than two distinct measurements. The plots in Fig. 7(b) show

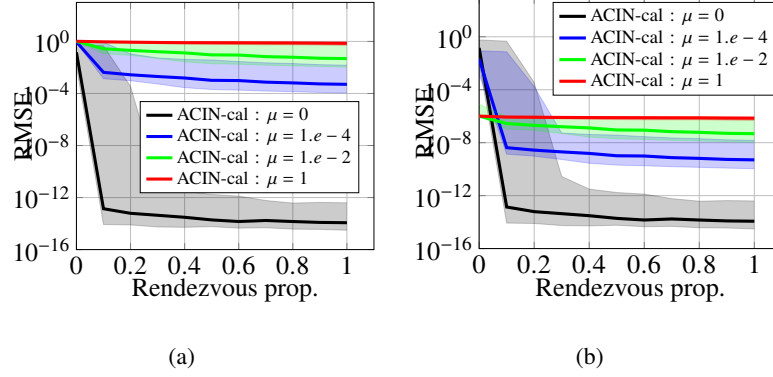


Fig. 8. Influence of the value of μ on the calibration performance of ACIN-cal when the estimation error on $\bar{\mathbf{f}}$ is (a) equal to 1, and (b) equal to $1.e-6$.

that the median RMSEs reached by SpIN-cal are always below the one reached by the multi-hop technique, even when the missing value proportion is low. When the number of missing entries in X is high, the multi-hop method provides a poor calibration estimation accuracy while the SpIN-cal method yields much lower RMSEs. Actually, when $\rho_{MV} = 0.95$, the performance reached by SpIN-cal is almost the same in both the noiseless and the noisy configurations.

It should be noticed that when all the assumptions required by the calibration methods are satisfied—i.e., at least two "diverse-enough" measurements per sensor, at least two rendezvous with a reference for each sensor—the approach in [14] works faster than our proposed ones which requires multiple iterations. However, let us stress again that these required conditions—especially the number of rendezvous—are difficult to hold in real crowdsensing applications, hence the need for new approaches able to handle a large proportion of missing data and a low number of rendezvous, as proposed in this paper.

C. Influence of the values of λ and μ

In the proposed calibration methods, different parameters have been introduced and the choice of their values have been partially discussed in Section III-E.

In the SpIN-cal method (Section III-D), two parameters have to be set, i.e., on the one hand, the number p of atoms which decompose the signal using the dictionary \mathcal{D} , and on the other hand, the weight λ associated to the regularization term in Eq. (29). In the above simulations, as the dictionary is constructed in order to perfectly infer the signal with two atoms, we set $p = 2$ while the parameter λ is set to 10. We made vary λ from 0.5 to 10 in preliminary tests

but we did not notice any significant differences between the obtained results. However, in some real data, as the dictionary will not exactly fit to the observed signal, and as the number of needed atoms is unknown *a priori*, these parameters have both to be set w.r.t. the confidence of the dictionary. It should be noticed that the selection of both these parameters is tricky—and highly subject to the construction of the dictionary—which is out of the scope of this paper, as explained in Section III-E.

In ACIN-cal method proposed in Section III-C, as $\bar{\mathbf{f}}$ is provided by the sensor manufacturer, the weight μ associated to the regularization term in Eq. (25) must be carefully chosen. In our simulations, we set $\bar{\mathbf{f}}$ to the actual means of the calibration parameters—an optimistic configuration which was also used in [11]—and $\mu = 10$. However, as briefly explained in Section III-C, in real situations the true mean parameters are not perfectly known, and Eq. (21) is approximately true only. Setting a too high value for μ then constrains the means of the rows of F to a wrong average, which thus provides a poor calibration accuracy. On the contrary, if one select a small weight μ , then both the ACIN-cal and IN-cal methods might perform favorably. The influence of μ is highlighted in Figs. 8(a) and 8(b) where we show the performance obtained with ACIN-cal when we set the missing value proportion to 0.9, we add no noise, and we make the rendezvous proportion vary from 0 to 100%. The tested values of μ are set to 0—in that case, ACIN-cal reduces to IN-cal—1E-4, 1E-2, and 1. In Fig. 8(a), $\bar{\mathbf{f}}$ is chosen such that the distance between $\bar{\mathbf{f}}$ and the true mean parameters is equal to 1 while in Fig. 8(b), $\bar{\mathbf{f}}$ is chosen with an error of $1.e - 6$.

Let us first focus on the case when $\mu = 0$ on both figures. Indeed, in that case, ACIN-cal is reduced to IN-cal, i.e., the average constraint is not taken into account. The calibration performance is thus similar to the one studied above, i.e., IN-cal is able to perform calibration provided there are enough rendezvous between calibrated and uncalibrated sensors. In the tests in Section IV-A, a perfect estimation of the average calibration parameters was used and the average constraint allowed to regularize the calibration problem and to reduce the spread. We now detail the behaviour of ACIN-cal when the average constraint does not perfectly fit to the theoretical average values.

In Fig. 8(a), the error due to the average calibration parameter is high and affects the calibration performance for any non-null value of μ . Indeed, when $\mu = 1$, the median RMSE remains around 1 even when the rendezvous proportion ρ_{RV} is equal to 1, i.e., when all the data are available. This shows that the average information has a major impact on the factorization. When μ decreases,

the RMSEs are slightly lower—as the impact of the average constraint is more balanced by the factorization—except when $\rho_{RV} = 0$ where the RMSE obtained with IN-cal and the error due to the average constraint are similar.

Fig. 8(b) shows a slightly different behaviour. In that case, the error due to the average calibration is rather low. As a consequence, when $\rho_{RV} = 0$, the average constraint in ACIN-cal allows to improve the RMSE with respect to IN-cal. However, when $\mu = 10^{-4}$, the weight due to the average constraint is too low and ACIN-cal almost behaves as IN-cal. When ρ_{RV} increases—i.e., when the assumptions needed by IN-cal to perform favorably are satisfied—we find the same behaviour as above, i.e., the highest tested value of μ makes ACIN-cal stick to the average calibration estimation error while lower values of μ provide lower RMSEs.

These tests show the importance of the error due to the average constraint on the calibration performance, hence the interest to take it into consideration as a penalized optimization term in ACIN-cal. As discussed in Section III-E, the automatic selection of the best value of μ is out of the scope of this paper but could be investigated, e.g., using late fusion or multi-objective optimization.

V. CONCLUSION

In this paper, we proposed four informed NMF techniques for mobile sensor calibration. These methods can take care of the constraints met in crowdsensing applications such as the being able to handle missing data and few rendezvous between uncalibrated and calibrated sensors. We deeply investigated their performance on 4800 data simulations and showed they outperformed a natural technique based on low-rank matrix completion. In particular, our sparsity-based and average-constrained SpIN-cal and SpAIN-cal methods were shown to be robust to both the missing value proportion ρ_{MV} and the rendezvous proportion ρ_{RV} . We then compared the performance of SpIN-cal with a state-of-the-art multi-hop calibration technique over 400 simulations and showed our approach to be more versatile in noiseless and noisy configurations. Our work shows that matrix factorization is a promising way to perform mobile sensor calibration and opens the way to many problems.

In our future work, we will investigate other sensor responses, e.g., nonlinear and multi-linear models for which specific constraints and update rules must be derived. Moreover, we will aim to learn a representation of the sensed phenomenon from the irregularly observed signals, instead of using a fixed dictionary as in this paper. In addition, we will investigate the enhancement

provided by data denoising using sparse approximation on the calibration performance. Lastly, we will investigate the performance of the tested methods on real mobile crowdsensing experiments.

APPENDIX A

PROOF OF ACIN-CAL UPDATE RULE FOR F

In this appendix, we aim to derive the update rule for F in the ACIN-cal method. We here denote J the cost function used in Eq. (25). J can be written as a sum $J \triangleq J_1 + J_2$ where

$$\begin{aligned} J_1 &= \|X - G \cdot F\|_{\mathcal{F}}^2, \\ \text{s.t. } F &= \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F, \end{aligned} \quad (39)$$

and

$$\begin{aligned} J_2 &= \frac{\mu}{2} \cdot \left\| \frac{1}{m} \cdot (\bar{\Omega}_F \circ F) \cdot \mathbb{1}_{[(m+1) \times 1]} - \bar{\mathbf{f}} \right\|_{\mathcal{F}}^2, \\ \text{s.t. } F &= \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F. \end{aligned} \quad (40)$$

Then, introducing the constraint of Eq. (40) in J_2 leads to

$$J_2 = \frac{\mu}{2} \cdot \left\| \frac{1}{m} \cdot (\bar{\Omega}_F \circ (\Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F)) \cdot \mathbb{1}_{[(m+1) \times 1]} - \bar{\mathbf{f}} \right\|_{\mathcal{F}}^2. \quad (41)$$

Noticing that $\bar{\Omega}_F \circ \Omega_F = 0$ and $\bar{\Omega}_F \circ \bar{\Omega}_F = \bar{\Omega}_F$, we derive

$$\begin{aligned} J_2 &= \frac{\mu}{2} \cdot \left\| \frac{1}{m} \cdot (\bar{\Omega}_F \circ \Delta_F) \cdot \mathbb{1}_{[(m+1) \times 1]} - \bar{\mathbf{f}} \right\|_{\mathcal{F}}^2, \\ &= \frac{\mu}{2} \cdot \sum_{i=1}^2 \left[\frac{1}{m} \cdot \sum_{j=1}^{m+1} (\bar{\Omega}_F \circ \Delta_F)_{ij} - \bar{f}_i \right]^2. \end{aligned} \quad (42)$$

The derivative of Eq. (42) with respect to the (i, j) -th element of Δ_F then reads

$$\frac{\partial J_2}{\partial \Delta_{F_{ij}}} = \mu \cdot \left[\frac{1}{m} \left(\sum_{k=1}^{m+1} \bar{\Omega}_{F_{ik}} \cdot \Delta_{F_{ik}} \right) - \bar{f}_i \right] \cdot \frac{1}{m} \cdot \bar{\Omega}_{F_{ij}}, \quad (43)$$

whose matrix extension reads

$$\frac{\partial J_2}{\partial \Delta_F} = \frac{\mu}{m^2} \cdot \text{diag} \left((\bar{\Omega}_F \circ \Delta_F) \cdot \mathbb{1}_{[(m+1) \times 1]} \right) \cdot \bar{\Omega}_F - \frac{\mu}{m} \cdot \text{diag}(\bar{\mathbf{f}}) \cdot \bar{\Omega}_F. \quad (44)$$

Following the heuristic optimization procedure, we associate the positive part of Eq. (44) with the positive part of the derivative of J_1 and the negative part of Eq. (44) to the negative part of the derivative of J_1 , which leads to

$$\partial J^+ = \bar{\Omega}_F \circ \left(G^T \left(W^{\circ 2} \circ \left(G \cdot (\bar{\Omega}_F \circ \Delta_F) \right) \right) \right) + \frac{\mu}{m^2} \cdot \text{diag} \left((\bar{\Omega}_F \circ \Delta_F) \cdot \mathbb{1}_{[(m+1) \times 1]} \right) \cdot \bar{\Omega}_F, \quad (45)$$

and

$$\partial J^- = \bar{\Omega}_F \circ \left(G^T (W^2 \circ X_F) \right) + \frac{\mu}{m} \cdot \text{diag}(\tilde{\mathbf{f}}) \cdot \bar{\Omega}_F. \quad (46)$$

The heuristic update rule for Δ_F then reads

$$\Delta_F \leftarrow \Delta_F \circ \frac{\partial J^-}{\partial J^+}, \quad (47)$$

and the global update for F thus reads

$$F \leftarrow \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \circ \frac{\partial J^-}{\partial J^+}, \quad (48)$$

which turns out to be the update rule in Section III-C.

APPENDIX B

ESTIMATION OF THE SENSED FIELD

In this paper, we aim to calibrate the responses of a sensor network, i.e., we aim to accurately estimate F in the NMF problem. However, the estimate of G and more precisely of its second column provides an estimation of the sensed field \mathbf{y} , as shown in 3. In this appendix, we aim to investigate its estimation accuracy with the proposed calibration methods. For that purpose, we consider again the experiments in Section IV-A.

Actually, depending on the chosen calibration approach, not all the entries of \mathbf{y} can be estimated. Indeed, we assumed the existence of rendezvous between sensors but we did not assume that each location of the Scene was sensed by at least one sensor. This implies that some entries of \mathbf{y} are not *accessible* through the sensed data in X . Moreover, as one calibrated sensor provides some values, their corresponding entries in \mathbf{g}_2 are known. In the experiments reported below, we thus consider two performance criteria denoted $\text{RMSE}_1(\mathbf{y}, \mathbf{g}_2)$ and $\text{RMSE}_2(\mathbf{y}, \mathbf{g}_2)$. $\text{RMSE}_1(\mathbf{y}, \mathbf{g}_2)$ (respectively $\text{RMSE}_2(\mathbf{y}, \mathbf{g}_2)$) consists of a RMSE computed between \mathbf{y} and \mathbf{g}_2 on the accessible data in \mathbf{g}_2 (the whole vector \mathbf{y} , respectively) without taking into account the known entries.

Fig. 9 shows the evolution of these performance criteria with respect to the rendezvous proportion, in the same experiment as in Section IV-A2 for the calibration performance. In all these plots, the envelope and the median of $\text{RMSE}_1(\mathbf{y}, \mathbf{g}_2)$ and $\text{RMSE}_2(\mathbf{y}, \mathbf{g}_2)$ are plotted in blue and red, respectively.

Fig. 9(a) put on light the fact that the natural approach based on low-rankness of X completely fails to estimate the sensed field as both measures are very high. On the contrary—and up to the

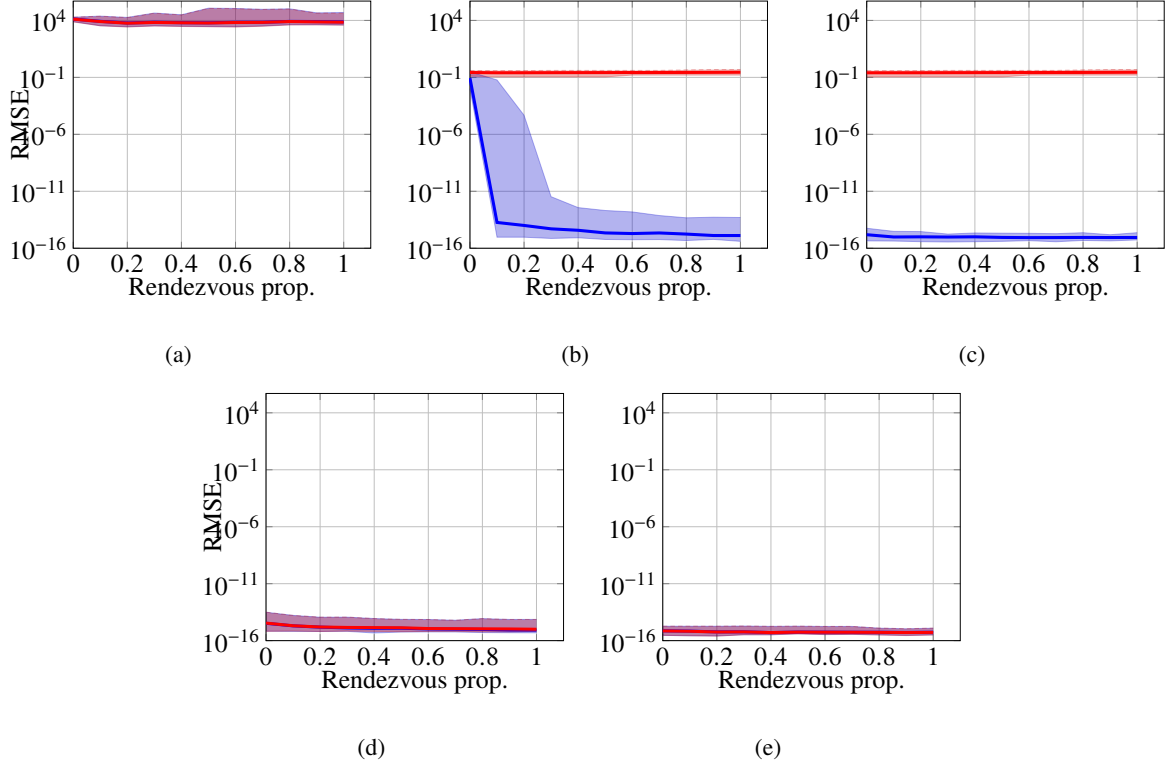


Fig. 9. Evolution of the RMSE₁ (in blue) and RMSE₂ (in red) with respect to the rendezvous proportion: (a) Natural approach; (b) IN-cal; (c) ACIN-cal; (d) SpIN-cal; (e) SpAIN-cal.

calibration limits of the methods which were already discussed—the four proposed methods are able to derive accurate estimations of \mathbf{y} in its accessible locations. However, for the locations when no sensor provides readings, only the sparsity-based techniques, i.e., SpIN-cal and SpAIN-cal, provide a consistent performance. On the contrary, IN-cal and ACIN-cal fail in estimating these missing entries.

These tests highlight the interest of matrix factorization techniques which not only allow to perform sensor calibration but also derive the sensed physical phenomenon.

REFERENCES

- [1] R. Ganti, F. Ye, and H. Lei, “Mobile crowdsensing: current state and future challenges,” *IEEE Communications Magazine*, vol. 49, no. 11, pp. 32–39, November 2011.
- [2] M. Talasila, R. Curtmola, and C. Borcea, “Mobile crowd sensing,” 2015.
- [3] L. Wang, D. Zhang, Y. Wang, C. Chen, X. Han, and A. M’hamed, “Sparse mobile crowdsensing: challenges and opportunities,” *IEEE Communications Magazine*, vol. 54, no. 7, pp. 161–167, July 2016.
- [4] D. Christin, A. Reinhardt, S. Kanhere, and M. Hollick, “A survey on privacy in mobile participatory sensing applications,” *Journal of systems and software*, vol. 84, no. 11, pp. 1928–1946, 2011.

- [5] N. Haderer, R. Rouvoy, and L. Seinturier, "Dynamic deployment of sensing experiments in the wild using smartphones," in *IFIP Intl. Conf. on Distributed Applications and Interoperable Systems*, 2013, pp. 43–56.
- [6] B.-T. Lee, S.-C. Son, and K. Kang, "A blind calibration scheme exploiting mutual calibration relationships for a dense mobile sensor network," *IEEE Sensors Journal*, vol. 14, no. 5, pp. 1518–1526, May 2014.
- [7] H. Carfantan and J. Idier, "Statistical linear destriping of satellite-based pushbroom-type images," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 1860–1871, Apr. 2010.
- [8] C. Schulke, F. Caltagirone, F. Krzakala, and L. Zdeborova, "Blind calibration in compressed sensing using message passing algorithms," in *NIPS 26*, 2013, pp. 566–574.
- [9] P. H. Schönemann, "On metric multidimensional unfolding," *Psychometrika*, vol. 35, no. 1, pp. 349–366, Mar. 1970.
- [10] M. Pollefeys and D. Nister, "Direct computation of sound and microphone locations from time-difference-of-arrival data," in *Proc. of ICASSP*, March 2008, pp. 2445–2448.
- [11] C. Bilen, G. Puy, R. Gribonval, and L. Daudet, "Convex optimization approaches for blind sensor calibration using sparsity," *IEEE Trans. Signal Processing*, vol. 62, no. 18, pp. 4847–4856, Sept. 2014.
- [12] L. Balzano and R. Nowak, "Blind calibration of sensor networks," in *Proc. of IPSN*, 2007, pp. 79–88.
- [13] J. Lipor and L. Balzano, "Robust blind calibration via total least squares," in *Proc. of ICASSP*, May 2014, pp. 4244–4248.
- [14] O. Saukh, D. Hasenfratz, and L. Thiele, "Reducing multi-hop calibration errors in large-scale mobile sensor networks," in *Proc. of IPSN*, 2015.
- [15] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Blind calibration of mobile sensors using informed nonnegative matrix factorization," in *Proc. of LVA-ICA*, ser. LNCS, 2015, vol. 9237, pp. 497–505.
- [16] —, "Blind mobile sensor calibration using a nonnegative matrix factorization with a relaxed rendezvous model," in *Proc. of ICASSP*, Mar. 2016, pp. 2941–2945.
- [17] —, "Outlier-robust calibration method for sensor networks," in *Proc. of ECMSM*, 2017.
- [18] C. Wang, P. Ramanathan, and K. Saluja, "Moments based blind calibration in mobile sensor networks," in *Proc. of ICC '08*, May 2008, pp. 896–900.
- [19] C. Wang, P. Ramanathan, and K. K. Saluja, "Calibrating nonlinear mobile sensors," in *Proc. of SECON*, Jun. 2008, pp. 533–541.
- [20] —, "Blindly calibrating mobile sensors using piecewise linear functions," in *Proc. of SECON*, Jun. 2009, pp. 1–9.
- [21] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Nonlinear mobile sensor calibration using informed semi-nonnegative matrix factorization with a vandermonde factor," in *Proc. of SAM*, Jul. 2016.
- [22] R. Piedrahita, Y. Xiang, N. Masson, J. Ortega, A. Collier, Y. Jiang, K. Li, R. Dick, Q. Lv, M. Hannigan, and L. Shang, "The next generation of low-cost personal air quality sensors for quantitative exposure monitoring," *Atmospheric Measurement Techniques*, vol. 7, no. 10, pp. 3325–3336, 2014.
- [23] E. Miluzzo, N. D. Lane, A. T. Campbell, and R. Olfati-Saber, "CaliBree: A self-calibration system for mobile sensor networks," in *Proc. of DCOSS*, ser. LNCS, vol. 5067, 2008, pp. 314–331.
- [24] O. Saukh, D. Hasenfratz, C. Walser, and L. Thiele, "On rendezvous in mobile sensing networks," in *Proc. of REALWSN*, ser. LNCS, vol. 281, 2014, pp. 29–42.
- [25] J. P. Haldar and D. Hernando, "Rank-constrained solutions to linear matrix equations using PowerFactorization," *IEEE Sig. Proc. Letters*, vol. 16, no. 7, pp. 584–587, July 2009.
- [26] C. Dorffer, M. Puigt, G. Delmaire, and G. Roussel, "Fast nonnegative matrix factorization and completion using Nesterov iterations," in *Proc. of LVA/ICA*, vol. LNCS 10179, 2017, pp. 26–35.
- [27] Y. Koren, R. Bell, and C. Volinsky, "Matrix factorization techniques for recommender systems," *Computer*, vol. 42, no. 8, pp. 30–37, Aug 2009.

- [28] Sensirion Corp., “SHT1x humidity and temperature sensor,” 2008, datasheet.
- [29] Sharp Corp., “GP2Y1010AU0F compact optical dust sensor,” 2006, datasheet.
- [30] E. Cochran, J. Lawrence, A. Kaiser, B. Fry, A. Chung, and C. Christensen, “Comparison between low-cost and traditional MEMS accelerometers: a case study from the M7.1 Darfield, New Zealand, aftershock deployment,” *Annals of Geophysics*, vol. 54, no. 6, 2012.
- [31] A. Cichocki, R. Zdunek, A. H. Phan, and S.-I. Amari, *Nonnegative matrix and tensor factorizations: applications to exploratory multi-way data analysis and blind source separation*. John Wiley & Sons, 2009.
- [32] Y. Xu, W. Yin, Z. Wen, and Y. Zhang, “An alternating direction algorithm for matrix completion with nonnegative factors,” *Frontiers of Mathematics in China*, vol. 7, no. 2, pp. 365–384, 2012.
- [33] S. Zhang, W. Wang, J. Ford, and F. Makedon, “Learning from incomplete ratings using non-negative matrix factorization,” in *Proceedings of the 2006 SIAM International Conference on Data Mining*. SIAM, 2006, pp. 549–553.
- [34] W. Xu, X. Liu, and Y. Gong, “Document clustering based on non-negative matrix factorization,” in *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*. ACM, 2003, pp. 267–273.
- [35] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, “Online learning for matrix factorization and sparse coding,” *Journal of Machine Learning Research*, vol. 11, no. Jan, pp. 19–60, 2010.
- [36] M. Puigt, G. Delmaire, and G. Roussel, “Environmental signal processing: new trends and applications,” in *Proc. of ESANN*, 2017, pp. 205–214.
- [37] N. Gillis, “The why and how of nonnegative matrix factorization,” in *Regularization, Optimization, Kernels, and Support Vector Machines*. Chapman and Hall/CRC, 2014, pp. 257–291.
- [38] D. Guillaumet, J. Vitria, and B. Schiele, “Introducing a weighted non-negative matrix factorization for image classification,” *Pattern Recognition Letters*, vol. 24, no. 14, pp. 2447–2454, 2003.
- [39] A. Limem, G. Delmaire, M. Puigt, G. Roussel, and D. Courcot, “Non-negative matrix factorization under equality constraints—a study of industrial source identification,” *Applied Numerical Mathematics*, vol. 85, pp. 1–15, Nov. 2014.
- [40] A. Cichocki, S. Cruces, and S. Amari, “Generalized alpha-beta divergences and their application to robust nonnegative matrix factorization,” *Entropy*, vol. 13, pp. 134–170, 2011.
- [41] E. J. Candès and B. Recht, “Exact matrix completion via convex optimization,” *Foundations of Computational Mathematics*, vol. 9, no. 6, pp. 717–772, 2009.
- [42] S. R. Becker, E. J. Candès, and M. C. Grant, “Templates for convex cone problems with applications to sparse signal recovery,” *Math. Progr. Comp.*, vol. 3, no. 3, pp. 165–218, 2011.
- [43] D. D. Lee and H. S. Seung, “Algorithms for non-negative matrix factorization,” in *Advances in Neural Information Processing Systems 13*, T. K. Leen, T. G. Dietterich, and V. Tresp, Eds. MIT Press, 2001, pp. 556–562.
- [44] H. Lantéri, C. Theys, C. Richard, and C. Févotte, “Split gradient method for nonnegative matrix factorization,” in *Proc. of EUSIPCO*, 2010.
- [45] A. Limem, M. Puigt, G. Delmaire, G. Roussel, and D. Courcot, “Bound constrained weighted NMF for industrial source apportionment,” in *Proc. of MLSP*, 2014.
- [46] R. Chreiky, G. Delmaire, M. Puigt, G. Roussel, D. Courcot, and A. Abche, “Split gradient method for informed non-negative matrix factorization,” in *Proc. of LVA/ICA*, vol. LNCS 9237, 2015, pp. 376–383.
- [47] J. M. Stockie, “The mathematics of atmospheric dispersion modeling,” *Siam Review*, vol. 53, no. 2, pp. 349–372, 2011.
- [48] P. O. Hoyer, “Non-negative sparse coding,” in *Proc. of NNSP*, 2002, pp. 557–565.
- [49] —, “Non-negative matrix factorization with sparseness constraint,” *Journal of Machine Learning Research*, vol. 5, pp. 1457–1469, November 2004.

- [50] Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: Algorithms and applications," *IEEE Access*, vol. 3, pp. 490–530, 2015.
- [51] J. Le Roux, F. J. Weninger, and J. R. Hershey, "Sparse NMF—half-baked or well done?" Mitsubishi Electric Research Labs, Tech. Rep. TR2015-023, 2015.
- [52] R. Hamon, V. Emiya, and C. Févotte, "Convex nonnegative matrix factorization with missing data," in *Proc. of MLSP*, 2016.
- [53] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 58, no. 1, pp. 267–288, 1996.
- [54] Y. Pati, R. Rezaifar, and P. Krishnaprasad, "Orthogonal Matching Pursuit: recursive function approximation with application to wavelet decomposition," in *Proc. of Asilomar Conf. on Signals, Systems and Comput.*, 1993.
- [55] D. Benachir, Y. Deville, S. Hosseini, M. S. Karoui, and A. Hameurlain, "Hyperspectral image unmixing by non-negative matrix factorization initialized with modified independent component analysis," in *Proc. of WHISPERS*, 2013.
- [56] M. Plouvin, A. Limem, M. Puigt, G. Delmaire, G. Roussel, and D. Courcot, "Enhanced NMF initialization using a physical model for pollution source apportionment," in *Proc. of ESANN*, 2014, pp. 261–266.
- [57] I. Meganem, Y. Deville, and M. Puigt, "Blind separation methods based on correlation for sparse possibly-correlated images," in *Proc. of ICASSP*, 2010, pp. 1334–1337.
- [58] D. Pavlidi, A. Griffin, M. Puigt, and A. Mouchtaris, "Real-time multiple sound source localization and counting using a circular microphone array," *IEEE Trans. on Audio, Speech, and Language Processing*, vol. 21, no. 10, pp. 2193–2206, October 2013.
- [59] C. A. Coello Coello, G. B. Lamont, and D. A. Van Veldhuizen, *Evolutionary algorithms for solving multi-objective problems*, 2nd ed. Springer, 2007.
- [60] K. Foley, D. Greene, and P. Cunningham, "Optimizing conflicting objectives in NMF using Pareto simulated annealing," in *Proc. of AICS*, 2010.
- [61] F. Zhu and P. Honeine, "Biobjective nonnegative matrix factorization: Linear versus kernel-based models," *IEEE Trans. on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 4012–4022, 2016.