Informed nonnegative matrix factorization for mobile sensor calibration

Matthieu PUIGT

Univ. Littoral Côte d'Opale, EA 4491 – LISIC, F-62228 Calais, France
matthieu.puigt@univ-littoral.fr
http://www-lisic.univ-littoral.fr/~puigt

March 3, 2017

Introducing ULCO & LISIC (1)

Université du Littoral Côte d'Opale (ULCO)

- `http://www.univ-littoral.fr/`
- "Proximity" University born in 1993
- 4 sites ($\simeq$ 40 km between Calais and the other sites):
  - Boulogne (fishing industry)
  - Calais (chemistry, ferry port, shuttle)
  - Dunkerque (industry port—metallurgy, petrol, etc—and energy—nuclear plants)
  - Saint-Omer (marshes used for agriculture, industry—crystal, paper, cardboard)
- Research activities with applications in environment (air, ground, water) and seas

Introducing ULCO & LISIC (2)

Laboratoire d'Informatique, Signal, Image de la Côte d'Opale (LISIC)

- Created in Calais in 2010 (fusion of 2 ULCO labs) with currently: 40 permanent faculty members, 16 Ph.D. students, 4 post-docs
- Four research teams with theoretical computer science and signal processing researchers:
  1. IMAP (Images and Learning)
  2. OSMOSE (Evolutionnary modelization, optimization, simulation)
  3. MODEL (Multi-Modelisation et Software Evolution)
  4. SPeciFI (Peception systems and Information Fusion)
- Several research projects in collaboration with industry (ArcelorMittal, Innocold, etc), research institutes (CNRS, IFREMER, etc) or public institutions (DREAL, Région Hauts-de-France, etc)

The OSCAR project

- Funded by the Région Hauts-de-France within the "Chercheurs-citoyens" program (2015–2017)
- Consortium with associations (ATMO HdF, BES) and research institutions (LISIC & Spirals [Inria,U. Lille, CNRS, IUF])
- Goal: provide fine-grained yet accurate air quality maps by combining precise measurements from ATMO with mobile sensor readings

The OSCAR project

Issues:

1. Designing a low-cost sensing device and launching a mass production involving (high school) students;
2. Sending data from the devices to a server using smartphones within the APISENSE *mobile-crowdsensing* platform
3. Performing sensor calibration from the sensor readings and ATMO measurements
4. Deriving precise air quality maps

The OSCAR project

Internet of Thing (IoT)

"One of the myths about the IoT is that companies have all the data they need, but ***their real challenge is making sense of it****. In reality, [...] the quality of the data isn't always good enough, and it remains difficult to integrate multiple data sources."* – Chris Murphy

The OSCAR project

Issues:

1. Designing a low-cost sensing device and launching a mass production involving (high school) students;
2. Sending data from the devices to a server using smartphones within the APISENSE *mobile-crowdsensing* platform
3. Performing sensor calibration from the sensor readings and ATMO measurements
4. Deriving precise air quality maps

Our publications within OSCAR

- One ongoing Ph.D. thesis (Clément Dorffer, since Nov. 2014)
- Publications (+ communications):

- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Blind calibration of mobile sensors using informed nonnegative matrix factorization*, Proc. of LVA/ICA, vol. LNCS 9237, pp. 497-505, 2015.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Blind mobile sensor calibration using an informed nonnegative matrix factorization with a relaxed rendezvous model*, Proc. of ICASSP, pp. 2941-2945, 2016.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Nonlinear mobile sensor calibration using informed semi-nonnegative matrix factorization with a Vandermonde factor*, Proc. of SAM, 2016.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Fast nonnegative matrix factorization and completion using Nesterov iterations*, Proc. of LVA/ICA, vol. LNCS 10179, pp. 26–35, 2017.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Outlier-robust calibration method for sensor network*, submitted to ECMSM
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Informed nonnegative matrix factorization methods for mobile sensor network calibration*, in preparation

Our publications within OSCAR

- One ongoing Ph.D. thesis (Clément Dorffer, since Nov. 2014)
- Publications (+ communications):

- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Blind calibration of mobile sensors using informed nonnegative matrix factorization*, Proc. of LVA/ICA, vol. LNCS 9237, pp. 497-505, 2015.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Blind mobile sensor calibration using an informed nonnegative matrix factorization with a relaxed rendezvous model*, Proc. of ICASSP, pp. 2941-2945, 2016.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Nonlinear mobile sensor calibration using informed semi-nonnegative matrix factorization with a Vandermonde factor*, Proc. of SAM, 2016.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Fast nonnegative matrix factorization and completion using Nesterov iterations*, Proc. of LVA/ICA, vol. LNCS 10179, pp. 26–35, 2017.
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Outlier-robust calibration method for sensor network*, submitted to ECMSM
- C. Dorffer, M. Puigt, G. Delmaire, G. Roussel, *Informed nonnegative matrix factorization methods for mobile sensor network calibration*, in preparation

Outline of the talk

1. Problem statement, definitions and assumptions
   - Problem statement
   - Definitions
   - Assumptions
2. Blind calibration as an NMF problem
   - Parameterization
   - Update rules
   - Initialization
3. Adding more information into NMF
   - Sparse assumptions to inform G
   - Sensor manufacturer information to inform F
4. The big picture
5. Experimental validation
   - Simulation of a scene
   - Interest of the parameterization
   - Performance achieved by all the methods on similar tests
   - Comparison with state-of-the-art methods
   - Physical field estimation
6. Conclusion and future work

Outline of the talk

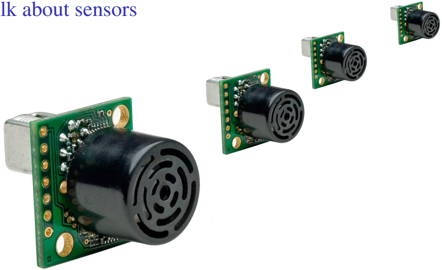
Problem statement

Let's talk about sensors

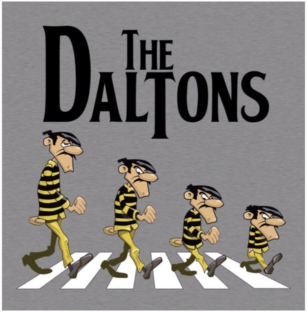
- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?

Problem statement

Let's talk about sensors

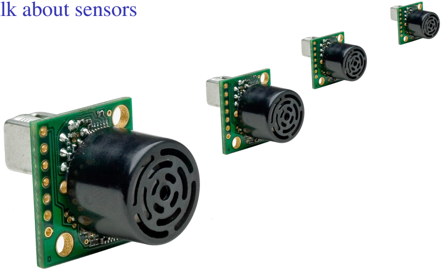


- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?

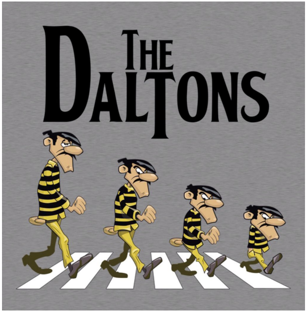


Problem statement

Let's talk about sensors

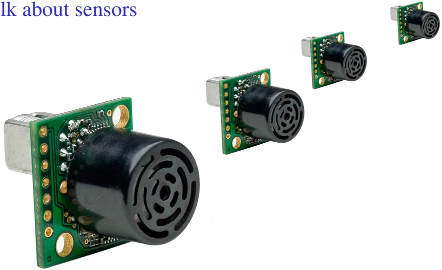


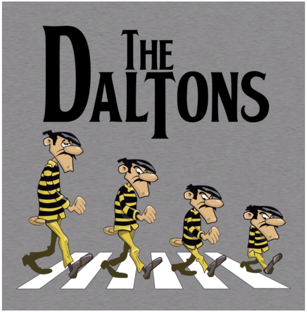
- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?
  - Sensor calibration needed
  - Not always physically possible
  - ➪ Blind sensor calibration



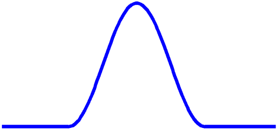
Problem statement

Let's talk about sensors





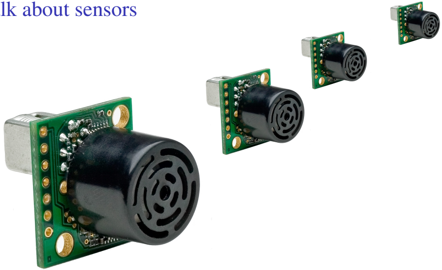
- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?
  - Sensor calibration needed
  - Not always physically possible
  - ➩ Blind sensor calibration

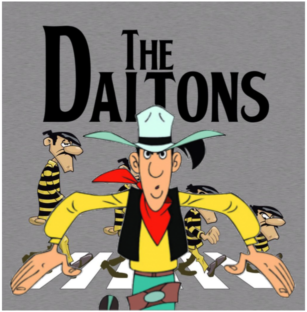


- Mobile sensors (***rendezvous***)
  - Averaging-based calibration (Lee *et al.*, 2014)

Problem statement

Let's talk about sensors



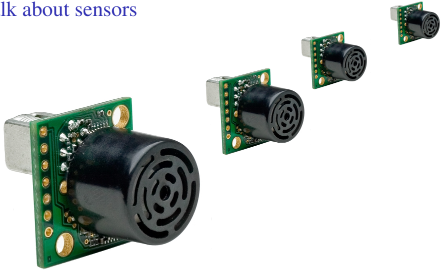


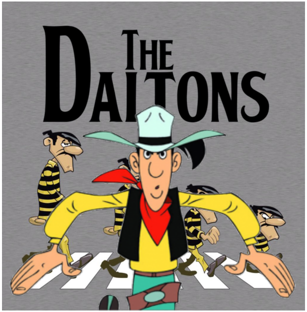
- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?
  - Sensor calibration needed
  - Not always physically possible
  - ➪ Blind sensor calibration

- Mobile sensors (***rendezvous***)
  - Averaging-based calibration (Lee *et al.*, 2014)
  - Calibration using a reference (multi-hop calibration, Saukh *et al.*, 2015)

Problem statement

Let's talk about sensors



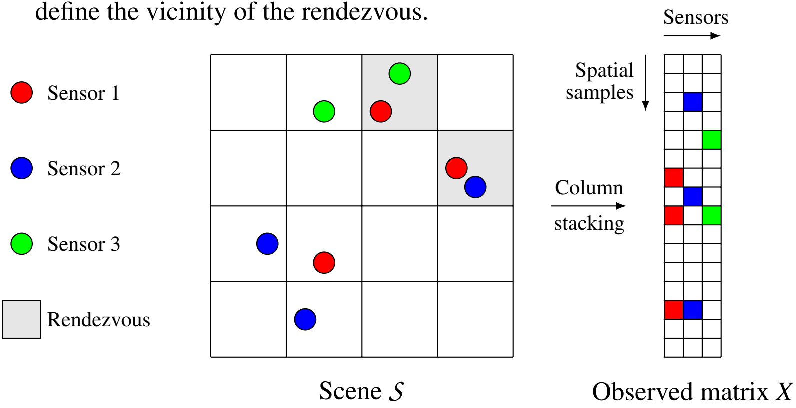


- Sensed phenomenon $\Longrightarrow$ voltage
- Voltage $\Longrightarrow$ phenomenon?
  - Sensor calibration needed
  - Not always physically possible
  - ➪ Blind sensor calibration

- Mobile sensors (***rendezvous***)
  - Averaging-based calibration (Lee *et al.*, 2014)
  - Calibration using a reference (multi-hop calibration, Saukh *et al.*, 2015)
- ➪ **Blind mobile sensor calibration without multi-hops?**

Definitions

- A rendezvous is a temporal and spatial vicinity between two sensors (Saukh *et al.*, 2013).
- A scene $\mathcal{S}$ is a discretized area observed during a time interval $[t, t+\Delta t)$. A spatial pixel has a size lower than Δd, where Δt and Δd define the vicinity of the rendezvous.



Assumptions (1)

- Sensor response (calibration function $\mathcal{F}(.)$ of Sensor j)

$$\underbrace{x(i,j)}_{\text{sensor-output voltage}} \simeq \mathcal{F}_j(y(i))$$

$$\simeq \underbrace{(y(i)}_{\text{physical phenomenom}} \cdot \overbrace{\alpha_j) + \beta_j}^{\text{unknown gain and offset}}$$

➪ Matrix form (if **each** of the m sensor senses **all** the scene)

$$\underbrace{\begin{bmatrix} x(1,1) & \cdots & x(1,m) \\ \vdots & & \vdots \\ x(n,1) & \cdots & x(n,m) \end{bmatrix}}_{X} \simeq \underbrace{\begin{bmatrix} y(1) & 1 \\ \vdots & \vdots \\ y(n) & 1 \end{bmatrix}}_{G} \cdot \underbrace{\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_m \\ \beta_1 & \beta_2 & \cdots & \beta_m \end{bmatrix}}_{F}$$

- In practice, irregular sampling: $W \circ X$ with

$$W(i,j) \triangleq \begin{cases} 0 & \text{if } x(i,j) \text{ is not available,} \\ \rho_j & \text{otherwise,} \end{cases}$$

where ρ_j is a weight coefficient associated with Sensor j

Assumptions (2)

- X, G, and F are nonnegative (air quality application)
- A known reference

➪ $\forall i = 1, \ldots, n, \quad x(i,m) = y(i)$ (i.e., $\alpha_m = 1$, $\beta_m = 0$)

➪ Blind calibration revisited as a weighted nonnegative matrix factorization problem

$$W \circ \underbrace{\begin{bmatrix} x(1,1) & \cdots & x(1,m-1) & y(1) \\ \vdots & & \vdots & \vdots \\ x(n,1) & \cdots & x(n,m-1) & y(n) \end{bmatrix}}_{X} \simeq W \circ \left(\underbrace{\begin{bmatrix} y(1) & 1 \\ \vdots & \vdots \\ y(n) & 1 \end{bmatrix}}_{G} \cdot \underbrace{\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{m-1} & 1 \\ \beta_1 & \beta_2 & \cdots & \beta_{m-1} & 0 \end{bmatrix}}_{F} \right)$$

Considered blind calibration problem

Estimating

$$\min_{G \geq 0, F \geq 0} \| W \circ (X - G \cdot F) \|_f^2$$

Assumptions to perform the blind calibration

Sensor network must be **dense enough** to guarantee enough rendezvous and X must be **well-conditioned** and with “enough” **diversity** on $W \circ X$

Outline of the talk

Proposed informed NMF method (1)

Parameterization

- Recent informed NMF methods with known entries:
  - penalization term (Choo *et al.*, 2015)
  - specific parameterization (Limem *et al.*, 2014)
- We generalize the parameterization in (Limem *et al.*, 2014):
  - Example ($n = 4, m = 3$, $k = 2$ calibrated measurements)

$$W \circ \begin{bmatrix} x(1,1) & x(1,2) & y(1) \\ x(2,1) & x(2,2) & y(2) \\ x(3,1) & x(3,2) & y(3) \\ x(4,1) & x(4,2) & y(4) \end{bmatrix} \simeq W \circ \left(\begin{bmatrix} y(1) & \mathbf{1} \\ \mathbf{y(2)} & \mathbf{1} \\ \mathbf{y(3)} & \mathbf{1} \\ y(4) & \mathbf{1} \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 & \alpha_2 & \mathbf{1} \\ \beta_1 & \beta_2 & \mathbf{0} \end{bmatrix} \right)$$

  - Parameterization: separate set and free parts in G and F

$$G = \underbrace{\begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 1 & 1 \\ 0 & 1 \end{bmatrix}}_{\Omega_G} \circ \underbrace{\begin{bmatrix} 0 & \mathbf{1} \\ \mathbf{y(2)} & \mathbf{1} \\ \mathbf{y(3)} & \mathbf{1} \\ 0 & \mathbf{1} \end{bmatrix}}_{\Phi_G} + \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 1 & 0 \end{bmatrix}}_{\bar{\Omega}_G} \circ \underbrace{\begin{bmatrix} y(1) & 0 \\ 0 & 0 \\ 0 & 0 \\ y(4) & 0 \end{bmatrix}}_{\Delta G} = \underbrace{\Omega_G \circ \Phi_G}_{\text{set}} + \underbrace{\bar{\Omega}_G \circ \Delta_G}_{\text{free}}$$

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

Proposed informed NMF method (2)

Update rules

- Informed NMF subproblems:
  1. $F = \arg\min_{\tilde{F}} ||W \circ \left(X - G \cdot \tilde{F}\right)||_f^2$ s.t. $\tilde{F} \geq 0, \quad \tilde{F} = \Phi_F + \Delta_F$
  2. $G = \arg\min_{\tilde{G}} ||W \circ \left(X - \tilde{G} \cdot F\right)||_f^2$ s.t. $\tilde{G} \geq 0, \quad \tilde{G} = \Phi_G + \Delta_G$
- Proposed method using multiplicative updates (Lee & Seung, 1999)
- Using an MM strategy, we derive

Update rules:

$$F_{k+1} = \Phi_F + \Delta_{F_k} \circ \bar{\Omega}_F \circ \left[\frac{G_k^T \left(W^{\circ 2} \circ \left(X - G_k \cdot \Phi_F\right)^+\right)}{G_k^T \left(W^{\circ 2} \circ \left(G_k \cdot \Delta_{F_k}\right)\right)}\right]$$

$$G_{k+1} = \Phi_G + \Delta_{G_k} \circ \bar{\Omega}_G \circ \left[\frac{\left(W^{\circ 2} \circ \left(X - \Phi_G \cdot F_{k+1}\right)^+\right) F_{k+1}^T}{\left(W^{\circ 2} \circ \left(\Delta_{G_k} \cdot F_{k+1}\right)\right) F_{k+1}^T}\right]$$

Proposed informed NMF method (3)

Initialization

- Initialization is known to be tricky
- Classical strategies: random, expert's one (Limem *et al.*, 2014) or output of another approach (Benachir *et al.*, 2013)
- We take into account the low-rank structure of X to initialize G and F

$$W \circ \underbrace{\begin{bmatrix} x(1,1) & \cdots & x(1,m-1) & y(1) \\ \vdots & & \vdots & \vdots \\ x(n,1) & \cdots & x(n,m-1) & y(n) \end{bmatrix}}_{X} \simeq W \circ \left(\underbrace{\begin{bmatrix} y(1) & 1 \\ \vdots & \vdots \\ y(n) & 1 \end{bmatrix}}_{G} \cdot \underbrace{\begin{bmatrix} \alpha_1 & \alpha_2 & \cdots & \alpha_{m-1} & 1 \\ \beta_1 & \beta_2 & \cdots & \beta_{m-1} & 0 \end{bmatrix}}_{F} \right)$$

1. We apply matrix completion (Becker *et al.*, 2011) to $W \circ X$ and obtain $\tilde{X}$ (also possible through weighted NMF)
2. We replace the missing entries in the first column of G by those of the last of $\tilde{X}$
3. We estimate F from $\tilde{X}$ and G using nonnegative least-squares

Outline of the talk

ADding more information into NMF

- We aim to relax the above constraints on the proposed calibration methods
- Two extensions which might be applied together (or not!) to inform the estimation of G or F, respectively
  1. Adding sparse priors
  2. Adding average calibration parameters

A new informed NMF approach

Adding sparse priors to inform NMF

- What if the sensor network is not dense enough?

$$W \circ X = \begin{bmatrix} x_{1,1} & - & - \\ - & - & y_2 \\ - & - & y_3 \\ - & x_{n,2} & - \end{bmatrix}$$

A new informed NMF approach

Adding sparse priors to inform NMF

- What if the sensor network is not dense enough?

$$W \circ X = \begin{bmatrix} x_{1,1} & - & - \\ - & - & y_2 \\ - & - & y_3 \\ - & x_{n,2} & - \end{bmatrix}$$

- Sparse decomposition of the scene, i.e., $y \approx \tilde{y} = \mathcal{D} \cdot a$ with $\mathcal{D} \in \mathbb{R}^{n,l}$ known dictionary, a **sparse** vector of contributions

A new informed NMF approach

Adding sparse priors to inform NMF

- What if the sensor network is not dense enough?

$$W \circ X = \begin{bmatrix} x_{1,1} & - & \tilde{y}_1 \\ - & - & y_2 \\ - & - & y_3 \\ - & x_{n,2} & \tilde{y}_4 \end{bmatrix} \text{ with } \tilde{y} = \mathcal{D} \cdot a$$

- Sparse decomposition of the scene, i.e., $y \approx \tilde{y} = \mathcal{D} \cdot a$ with $\mathcal{D} \in \mathbb{R}^{n,l}$ known dictionary, a **sparse** vector of contributions

A new informed NMF approach

Adding sparse priors to inform NMF

- What if the sensor network is not dense enough?

$$W \circ X = \begin{bmatrix} x_{1,1} & - & \tilde{y}_1 \\ - & - & y_2 \\ - & - & y_3 \\ - & x_{n,2} & \tilde{y}_4 \end{bmatrix} \text{ with } \tilde{y} = \mathcal{D} \cdot a$$

- Sparse decomposition of the scene, i.e., $y \approx \tilde{y} = \mathcal{D} \cdot a$ with $\mathcal{D} \in \mathbb{R}^{n,l}$ known dictionary, a **sparse** vector of contributions

$\Rightarrow$ Global optimization problem

$$\begin{aligned} \hat{G}, \hat{F} \quad &= \quad \arg\min_{G,F} ||W \circ (X - G \cdot F)||_f^2 \\ \text{s.t.} \quad & G, F \geq 0 \\ & G = \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G \\ & F = \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \\ & \exists a \in \mathbb{R}^l \text{ s.t. } g_1 \approx \mathcal{D} \cdot a \end{aligned}$$

Structure of informed NMF with sparse priors

$$\min_{G,F,a} ||W \circ (X - G \cdot F)||_f^2 + \lambda \cdot ||g_1 - \mathcal{D} \cdot a||_f^2 + \mu \cdot ||a||_1^2$$

$$\text{s.t.} \begin{cases} G, F & \geq & 0 \\ G & = & \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G \\ F & = & \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \end{cases}$$

- Non-convex problem w.r.t all variables
- ➩ Alternating strategy
  1. Updating F (as before)
  2. Updating G
  3. Estimating a

Structure of informed NMF with sparse priors

Updating G

$$G = \arg\min_{\tilde{G}\geq 0} \left|\left| W \circ (X - \tilde{G} \cdot F) \right|\right|_f^2 + \lambda \cdot ||g_1 - \mathcal{D} \cdot a||_f^2$$

$$\text{s.t. } \tilde{G} = \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G$$

- Defining

$$\mathbf{X} = [X, \mathcal{D} \cdot a] \qquad \mathbf{W} = \left[W, \sqrt{\lambda} \cdot \mathbb{1}_{x\times 1}\right] \qquad \mathbf{F} = \left[F, \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right],$$

- Updating G as previously (except that X, W and F are replaced by $\mathbf{X}$, $\mathbf{W}$ and $\mathbf{F}$)

Structure of informed NMF with sparse priors

$$\min_{G,F,a} ||W \circ (X - G \cdot F)||_f^2 + \lambda \cdot ||g_1 - \mathcal{D} \cdot a||_f^2 + \mu \cdot ||a||_1^2$$

$$\text{s.t.} \begin{cases} G, F & \geq & 0 \\ G & = & \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G \\ F & = & \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \end{cases}$$

- Non-convex problem w.r.t all variables

➩ Alternating strategy

1. Updating F (as before)
2. Updating G
3. Estimating a

Structure of informed NMF with sparse priors

Estimating a

$$\hat{a} = \arg\min_{a} \lambda \cdot ||g_1 - \mathcal{D} \cdot a||_f^2 + \mu \cdot ||a||_1^2$$

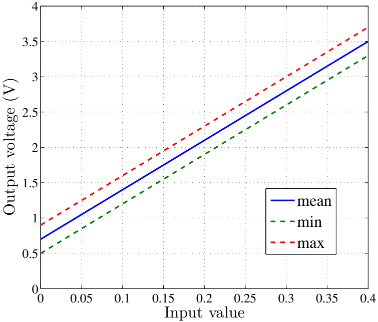
➯ Best sparse decomposition of g_1

Solved using the Orthogonal Matching Pursuit (OMP) greedy algorithm

- Iteratively selecting the best atom in $\mathcal{D}$,
- Defining the orthogonal projection on the selected atom,
- Projecting the residual,
- Repeating until a given number of atoms or a given residual error.

Sensor manufacturer information to inform F

- Calibration parameters are distributed around known values
- **Typical manufacturer data**



➭ Sum constraint on α_j and β_j

! Sum constraints are classical
  - in hyperspectral unmixing (e.g., Heinz, 2001)
  - chemical source apportionment (see, e.g., Chreiky *et al.*, 2015)

- Here, they are approximately satisfied
  ➭ Defining $\bar{f} \triangleq [\bar{f}_1, \bar{f}_2]^T$
  - if $m \rightarrow +\infty$ then $\frac{1}{m-1} \cdot F \cdot \begin{bmatrix} \mathbb{1}_{m-1\times 1} \\ 0 \end{bmatrix} \rightarrow \bar{f}$ or equivalently $\frac{1}{m} \cdot \left(\bar{\Omega}_F \circ \Phi_F\right) \cdot \mathbb{1}_{m,1} \rightarrow \bar{f}$

New cost function

- New penalization term in the optimization of F

$$\min_{G,F} ||W \circ (X - G \cdot F)||_f^2 + \gamma \left|\left| \frac{1}{m} \cdot \left(\bar{\Omega}_F \circ \Phi_F\right) \cdot \mathbb{1}_{m,1} - \bar{f} \right|\right|_f^2$$

$$\text{s.t.} \begin{cases} G, F & \geq & 0 \\ G & = & \Omega_G \circ \Phi_G + \bar{\Omega}_G \circ \Delta_G \\ F & = & \Omega_F \circ \Phi_F + \bar{\Omega}_F \circ \Delta_F \end{cases}$$

- Using the heuristic MU, we derive the update rule

$$F \leftarrow \bar{\Omega}_F \cdot \Phi_F + \Omega_F \Delta_F \circ \frac{\bar{\Omega}_F \circ \left(G^T \left(W^2 \circ (X - G \cdot \Delta_F)\right)\right) + \frac{\gamma}{m} \cdot \mathrm{diag}(\bar{f}) \cdot \bar{\Omega}_F}{\bar{\Omega}_F \circ \left(G^T \left(W^2 \circ \left(G \cdot \left(\bar{\Omega}_F \circ \Delta_F\right)\right)\right)\right) + \frac{\gamma}{m(m-1)} \cdot \mathrm{diag}\left(\bar{\Omega}_F \circ \Delta_F \cdot \mathbb{1}_{m,1}\right) \cdot \bar{\Omega}_F}$$

Outline of the talk

The big picture

- Blind sensor calibartion ≈ matrix factorization with missing entries and
  - Structure on G: $\begin{bmatrix} y(1) & \mathbf{1} \\ \mathbf{y}(\mathbf{2}) & \mathbf{1} \\ \mathbf{y}(\mathbf{3}) & \mathbf{1} \\ y(4) & \mathbf{1} \end{bmatrix}$
  - Structure on F: $\begin{bmatrix} \alpha_1 & \alpha_2 & \mathbf{1} \\ \beta_1 & \beta_2 & \mathbf{0} \end{bmatrix}$

➪ Informed matrix factorization with specific parameterization

1. $F = \arg\min_{\tilde{F}} ||W \circ \left(X - G \cdot \tilde{F}\right)||_f^2$ s.t. $\tilde{F} \geq 0, \quad \tilde{F} = \Phi_F + \Delta_F$
2. $G = \arg\min_{\tilde{G}} ||W \circ \left(X - \tilde{G} \cdot F\right)||_f^2$ s.t. $\tilde{G} \geq 0, \quad \tilde{G} = \Phi_G + \Delta_G$

\+ Information on G (sparse priors)

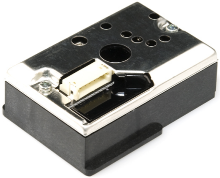
$$G = \arg\min_{\tilde{G} \geq 0} ||W \circ (X - \tilde{G} \cdot F)||_f^2 + \lambda \cdot ||g_1 - \mathcal{D} \cdot a||_f^2 \text{ s.t. } \tilde{G} = \Phi_G + \Delta_G$$

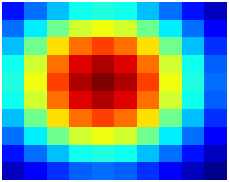
& Information on F (known average calibration parameter values)

$$F = \arg\min_{\tilde{F} \geq 0} ||W \circ (X - G \cdot \tilde{F})||_f^2 + \gamma || \frac{1}{m} \cdot \left(\bar{\Omega}_F \circ \Phi_F\right) \cdot \mathbb{1}_{m,1} - \bar{f}||_f^2 \text{ s.t. } \tilde{F} = \Phi_F + \Delta_F$$

Outline of the talk

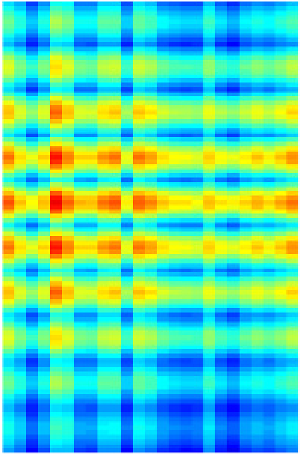
Simulation of a scene





- Crowdsensing-like particulate matter sensing
- Compact Optical Dust Sensor (Sharp GP2Y1010AU0F)
- Scene = 10×10 discretized area ($n = 100$) observed by $m = 26$ sensors
- The values in $\underline{y}$ range between 0 and 0.5 mg/m^3 (no sensor saturation)
- Dictionary with 62 atoms
- $\underline{y}$ is 2-sparse

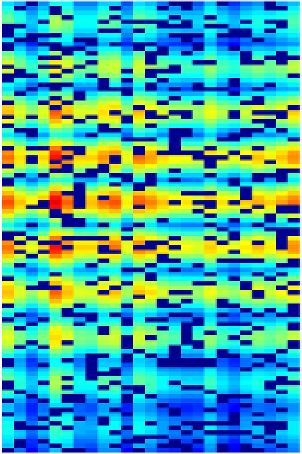
Sensors and observations



No missing samples
($l = 2500$)

- Calibration parameters randomly chosen according to truncated Gaussian distribution, $\forall j = 1, \ldots, m-1$:
  - α_j centered around 5 V/(mg/m^3) (variance 5e-2) with $3.5 < \alpha_j < 6.5$
  - β_j centered around 0.9 V (variance 6e-2) with $0 < \beta_j < 1.5$
- ➯ X is a 26×100 matrix for which we randomly keep $k + l$ samples (with $k \ll l$)
  - $k = 4$ high-quality & calibrated measurements in the m-th column of X
  - Non-null entries of W are set to $\rho_j = 1, \forall j = 1, \ldots, m-1$ and $\rho_m = l$.
- + Gaussian noise realizations (input SNR varying from 11 dB to 100 dB)

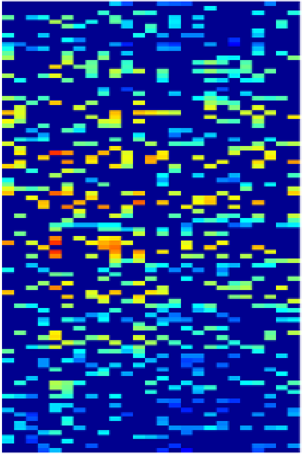
Sensors and observations



25% missing samples
($l = 1875$)

- Calibration parameters randomly chosen according to truncated Gaussian distribution, $\forall j = 1, \ldots, m-1$:
  - α_j centered around 5 V/(mg/m^3) (variance 5e-2) with $3.5 < \alpha_j < 6.5$
  - β_j centered around 0.9 V (variance 6e-2) with $0 < \beta_j < 1.5$
- ➪ X is a 26×100 matrix for which we randomly keep $k + l$ samples (with $k \ll l$)
  - $k = 4$ high-quality & calibrated measurements in the m-th column of X
  - Non-null entries of W are set to $\rho_j = 1, \forall j = 1, \ldots, m-1$ and $\rho_m = l$.
- + Gaussian noise realizations (input SNR varying from 11 dB to 100 dB)

Sensors and observations



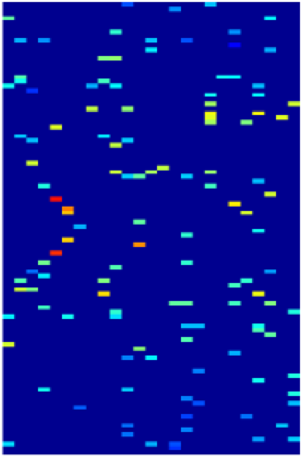
75% missing samples
($l = 625$)

- Calibration parameters randomly chosen according to truncated Gaussian distribution, $\forall j = 1, \ldots, m-1$:
  - α_j centered around 5 V/(mg/m^3) (variance 5e-2) with $3.5 < \alpha_j < 6.5$
  - β_j centered around 0.9 V (variance 6e-2) with $0 < \beta_j < 1.5$

➩ X is a 26×100 matrix for which we randomly keep $k + l$ samples (with $k \ll l$)
  - $k = 4$ high-quality & calibrated measurements in the m-th column of X
  - Non-null entries of W are set to $\rho_j = 1, \forall j = 1, \ldots, m-1$ and $\rho_m = l$.

+ Gaussian noise realizations (input SNR varying from 11 dB to 100 dB)

Sensors and observations

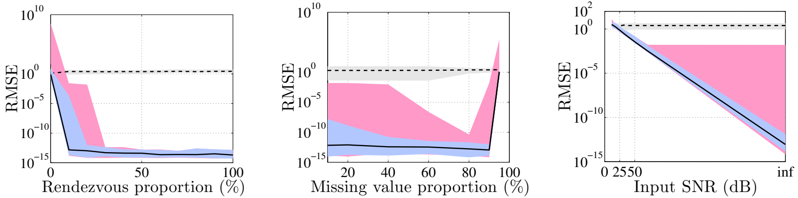


95% missing samples
($l = 125$)

- Calibration parameters randomly chosen according to truncated Gaussian distribution, $\forall j = 1, \ldots, m-1$:
  - α_j centered around 5 V/(mg/m^3) (variance 5e-2) with $3.5 < \alpha_j < 6.5$
  - β_j centered around 0.9 V (variance 6e-2) with $0 < \beta_j < 1.5$
- ➪ X is a 26×100 matrix for which we randomly keep $k + l$ samples (with $k \ll l$)
  - $k = 4$ high-quality & calibrated measurements in the m-th column of X
  - Non-null entries of W are set to $\rho_j = 1, \forall j = 1, \ldots, m-1$ and $\rho_m = l$.
- + Gaussian noise realizations (input SNR varying from 11 dB to 100 dB)

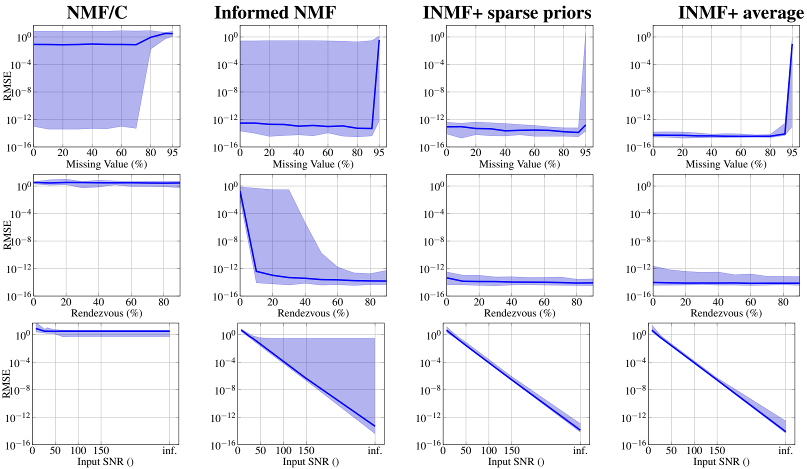
Interest of the parameterization

- We aim to test the influence of:
  1. the proportion of uncalibrated sensors to have rendezvous with calibrated ones (default 20%),
  2. the proportion of missing entries in X (default 90%)
  3. the input SNR (default noiseless)
- For each test condition, we launch 25 random runs with **2e6** iterations
- Perf. criterion: Root Mean Square Error (RMSE) over rows of F
- Comparison with a "naive" informed NMF method (classical NMF followed by a projection step to replace the known entries by their actual values) in pink vs proposed method in blue



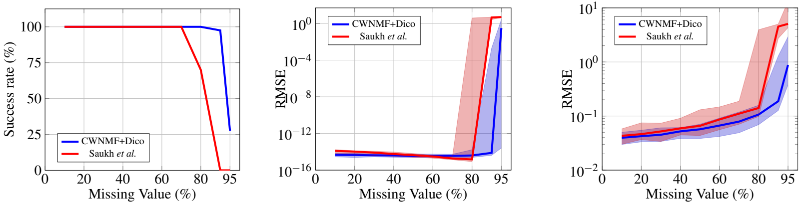
Performance achieved by all the methods

- Slightly different initialization of the methods:
  1. Random initialization of F around theoretical values of $\bar{f}$
  2. NMF/C used for completion
- We aim to test the influence of:
  1. the proportion of uncalibrated sensors to have rendezvous with calibrated ones (default 30%),
  2. the proportion of missing entries in X (default 90%)
  3. the input SNR (default noiseless)
- For each test condition, we launch 25 random runs with **5e5** iterations
- Perf. criterion: RMSE over rows of F


NMF/C
Informed NMF
INMF+ sparse priors
INMF+ average
RMSE
Missing Value (%)
Rendezvous (%)
Input SNR ()
0 20 40 60 80 95
0 20 40 60 80
0 50 100 150 inf.
10^{0} 10^{-4} 10^{-8} 10^{-12} 10^{-16}

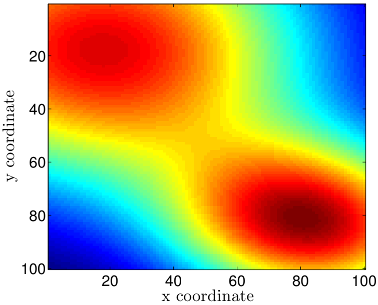
Comparison with state-of-the-art method

- Saukh *et al.* proposed a multi-hop calibration technique
  1. Perform regression between readings of calibrated and uncalibrated sensors in rendezvous
  2. Estimate the calibration parameters of the uncalibrated sensor
  3. Use it as a new calibrated reference and go back to step 1.
- The multi-hop approach does not work on the previous tests (conditions not satisfied)
- We perform new simulations with more rendezvous to see how the approaches perform
  - We randomly remove a proportion of missing entries in X (much more calibrated data than in the previous tests)
  - Perf criterion: RMSE, success rate (RMSE below 1e-10) in noiseless and noisy cases (SNR $\approx$ 30 dB)



Physical field estimation

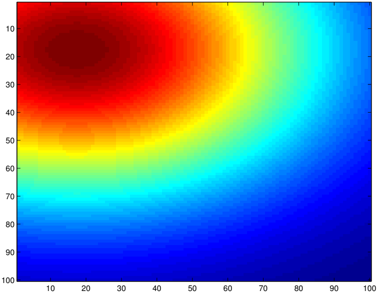
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

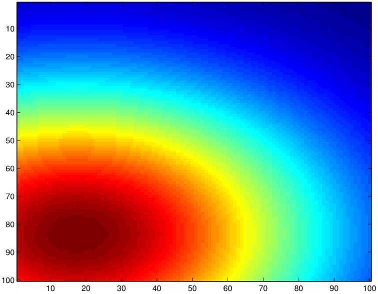
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

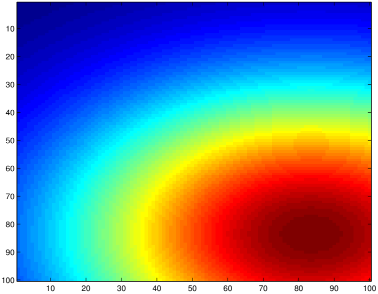
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

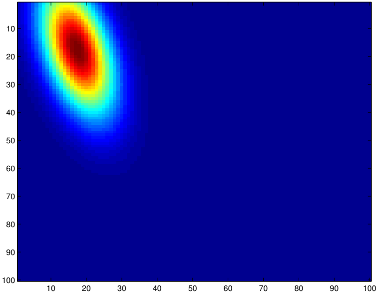
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

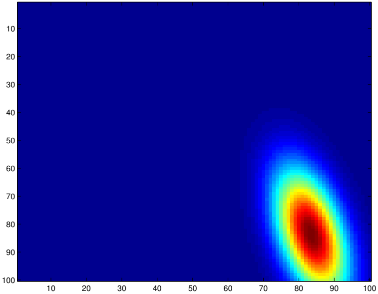
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

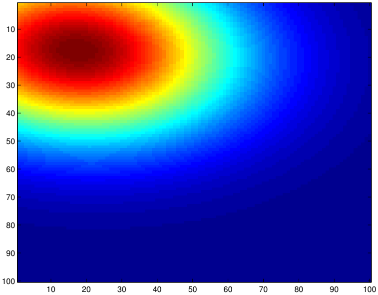
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

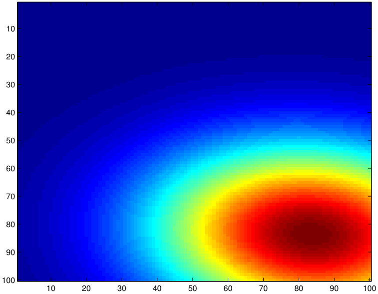
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Physical field estimation

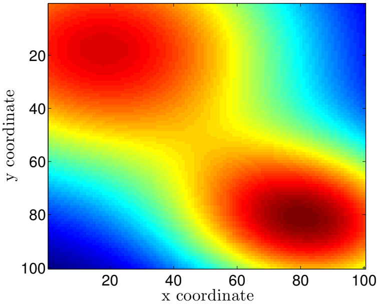
- So far, we focused on the estimation of F
- G provides an estimation of the sensed physical field
- Let us see an example



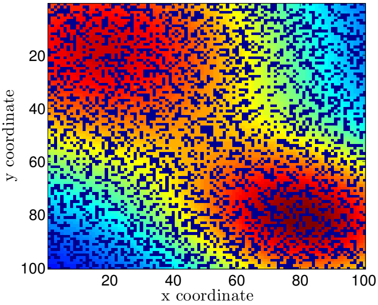
- 100×100 scene
- Sensed by 50 sensors
- $k = 10$ calibrated sensor readings
- 98% of missing data
- No rendezvous between uncalibrated and calibrated sensors
- 100 atoms in the dictionnary (Gaussian-like atoms)

Reconstruction accuracy

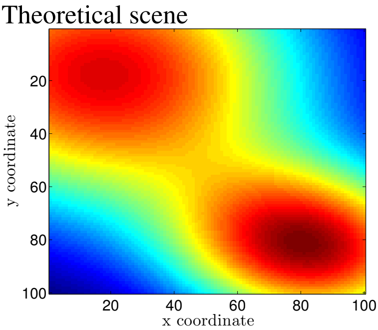
Theoretical scene

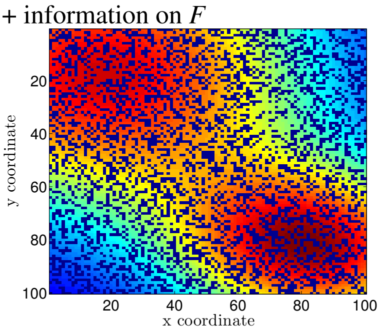


First informed NMF

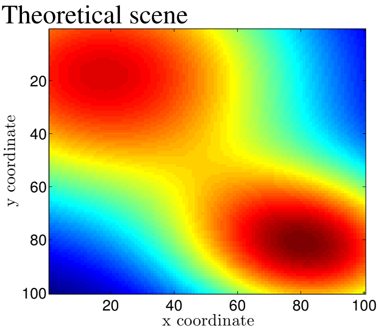


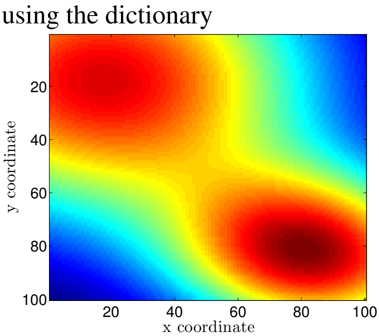
Reconstruction accuracy





Reconstruction accuracy





Conclusion and future work

Conclusion

- Blind mobile sensor calibration revisited as informed matrix factorization
  - specific constraints due to the problem (parameterization, sparse priors, known average calibration parameters)
- Proposed approaches robust to the number of missing entries and of rendezvous (no spatial discretization is required with the sparse approximation)
- well-conditioned X restrictive and diversity in $W \circ X$?
  - More likely to be satisfied in a multiple-scene configuration

We also proposed approaches for

- nonlinear calibration models (SAM'16)
- fast factorization using Nesterov iterations (LVA-ICA'17)

Conclusion and future work

Future work

- Replacing the dictionnary by a geostatistical physical model
- Case of sensors whose readings also depend on humidity and temperature
- Blind calibration through privacy-preserving crowdsensing?

Thank you for your attention !