

Méthodes de séparation de sources non linéaires pour des capteurs gaz à oxyde métallique

STEPHANIE MADROLLE¹, RAFAEL A. ANDO^{2,3}, LEONARDO T. DUARTE⁴, ROMIS ATTUX³, CHRISTIAN JUTTEN²,
PIERRE GRANGEAT¹

¹ Univ. Grenoble Alpes, CEA, LETI, Minatec Campus, F-38054 Grenoble Cedex, France.

² Université Grenoble Alpes, GIPSA-lab, UMR CNRS 5216, France

³ School of Electrical and Computer Engineering, University of Campinas (UNICAMP), Brazil

⁴ School of Applied Sciences, University of Campinas (UNICAMP), Brazil

¹ stephanie.madrolle@cea.fr, pierre.grangeat@cea.fr,

² rafael.ando@gipsa-lab.grenoble-inp.fr, christian.jutten@gipsa-lab.grenoble-inp.fr

Résumé – L’analyse quantitative de l’air expiré présente un fort intérêt médical, que ce soit pour le diagnostic ou le suivi de traitement. Dans ce cadre, nous proposons d’analyser les observations faites par un capteur à oxyde métallique (MOX), en réponse à des mélanges de deux gaz (acétone et éthanol). Une diversité dans les mesures est obtenue en combinant plusieurs températures de fonctionnement appliquées sur un seul capteur, ceci étant équivalent à plusieurs capteurs virtuels. Nous utilisons des méthodes de séparation de sources non linéaires, aveugle et non aveugle, afin d’estimer les concentrations de gaz présents dans l’échantillon, de valider et de calibrer le modèle de mélange non linéaire des capteurs MOX. L’inversion et le principe sont validés sur données simulées et expérimentales.

Abstract - Quantitative analysis of exhaled breath has a strong medical interest, whether for diagnosis or treatment follow-up. In this purpose, we propose to analyze the observations made by a metal oxide (MOX) sensor in response to mixtures of two gases (acetone and ethanol). A diversity in the measurements is obtained by combining several temperature modes applied to a single sensor, this being equivalent to several virtual sensors. We use nonlinear, blind and non-blind source separation methods to estimate gas concentrations, to validate and calibrate the nonlinear mixing model of MOX sensors. The analysis methods and the principle are validated on simulated and experimental data.

1 Introduction

Les capteurs gaz, notamment les capteurs à oxyde métallique (MOX), sont de plus en plus utilisés pour détecter les composés organiques volatils (COVs), notamment dans les dispositifs de type nez électronique [1,2]. Ces capteurs résistifs ne sont pas très spécifiques et détectent alors plusieurs COVs présents. Afin de séparer les différents composés détectés par le capteur, l’approche par séparation de sources est donc une étape essentielle. Elle va permettre de différencier les gaz, et même de les quantifier en estimant leur concentration. Cette quantification de gaz n’est pas encore exploitée dans la littérature, puisque la plupart des auteurs regardent plutôt la classification des mélanges [3,4].

Dans ce travail, nous proposons de valider notre modèle de capteur et l’inversion du système mathématique obtenu, en utilisant des approches de séparation de sources, qui sont dans un premier temps supervisées. Puis, des méthodes non supervisées seront explorées pour éviter de disposer de nombreuses mesures de calibration, qui sont coûteuses, longues et difficiles à réaliser. Cependant, la mise en œuvre de ces méthodes requiert (1) une diversité, généralement obtenue à l’aide de plusieurs capteurs [2,3], et (2) la caractérisation préalable du modèle qui relie les mesures aux sources, ici aux concentrations des gaz du mélange.

Dans ce papier, la diversité sera obtenue en utilisant un seul capteur physique, mais avec des températures de

fonctionnement différents, permettant ainsi de créer un réseau de capteurs virtuels. La caractérisation du modèle mesures/concentrations sera obtenue par régression à partir de données expérimentales.

Dans la partie 2, nous poserons le problème, et détaillerons les modes de fonctionnement du capteur ainsi que le modèle de mélange. Puis, dans la partie 3, nous expliquerons les méthodes de séparation de sources utilisées pour calibrer le modèle et enfin dans la partie 4 nous montrerons les résultats obtenus avec données simulées et données expérimentales

2 Caractérisation expérimentale du modèle

2.1 Diversité due à la température

Nous montrons dans ce travail que la diversité peut être obtenue, avec un unique capteur physique, en contrôlant la température du capteur. En effet, la couche sensible au gaz est aussi sensible aux variations de température de l’élément chauffant situé sous elle. La détection des particules gazeuses par le capteur MOX variera donc suivant la température appliquée sur ce capteur, et ce de manière différente pour chaque gaz. Ceci nous donnera alors la diversité recherchée.

Cette idée, partiellement exploitée en classification [5,6] est totalement inédite en séparation de sources. Cependant, il n’existe pas dans la littérature de modèle physique formel de capteur MOX utilisé dans ce contexte. Il est donc indispensable de valider cette idée en modélisant la réponse du capteur à différentes

températures, et pour différents mélanges de concentrations de gaz.

Dans notre cas, la réponse du capteur à la température T_i , notée x_i , est le rapport de résistances entre la mesure avec gaz cible (R_{gaz}) et sans, uniquement dans l'air (R_0) :

$$x_i = \frac{R_0}{R_{gaz}} \Big|_{T_i} \quad (1)$$

2.2 Dispositif expérimental et mesures

De manière à travailler avec un modèle correspondant au mieux aux données expérimentales, nous avons acquis des données à l'aide d'un capteur MOX venant du commerce (SB-30 de FIS). Ces capteurs sont sensibles aux composés organiques volatils (COVs), mais également à l'environnement. Il faut alors s'assurer que nos mesures ne sont pas erronées. Nous travaillons donc à paramètres extérieurs constants, avec une température de $28,4 \text{ °C} \pm 0,4 \text{ °C}$, une humidité relative (rh) de $44 \% \pm 1 \%$ et un taux d'oxygène de $17,80 \% \pm 0,05 \%$. Des mélanges de deux gaz, ici l'acétone et l'éthanol, dilués dans de l'air, sont étudiés, avec des concentrations variant de 0 à 650 ppm. Chaque mesure est faite à deux températures correspondant aux deux capteurs virtuels et répétée deux fois, pour obtenir plus de robustesse.

2.3 Modèle de mélange

La littérature propose des modèles contradictoires de capteurs MOX (cités dans [7]), mais aucun prenant en compte la température du capteur. Aussi, par régression à partir des mesures expérimentales présentées au paragraphe précédent, nous avons validé le modèle de capteur [7]. A une température T_i , la réponse x_i du capteur lorsque les concentrations des deux gaz sont c_1 et c_2 suit alors le modèle non linéaire suivant :

$$x_i = a_{i1}c_1^{r_1} + a_{i2}c_2^{r_2} + b_{i1}(c_1^{r_1})^2 + b_{i2}(c_2^{r_2})^2 + d_i c_1^{r_1} c_2^{r_2} \quad (2)$$

où les a_{ij} , b_{ij} , d_i sont des paramètres dépendants de la température T_i , alors que r_1 et r_2 n'en dépendent pas. En notant $s_1 = c_1^{r_1}$, $s_2 = c_2^{r_2}$, $\mathbf{x} = (x_1, x_2)^T$, $\mathbf{s} = (s_1, s_2)^T$, $\mathbf{s}^2 = (s_1^2, s_2^2)^T$, \mathbf{M} la matrice 2 x 2 de terme général m_{ij} et $\mathbf{d} = (d_1, d_2)^T$, pour des mesures à deux températures, cette équation du second degré (2) peut s'écrire sous forme compacte :

$$\mathbf{x} \approx \mathbf{A}\mathbf{s} + \mathbf{B}\mathbf{s}^2 + \mathbf{d}s_1s_2 \quad (3)$$

Avec les paramètres estimés à partir de nos données expérimentales, pour deux températures de capteur, $T_1 = 0,5 \text{ V}$ et $T_2 = 0,9 \text{ V}$, la diversité pour des concentrations de gaz uniformément réparties est illustrée sur la Figure 1, qui montre les points de mesures dans le plan. Ainsi pour deux températures de capteur, les deux capteurs virtuels fournissent un système non linéaire de deux équations dont les inconnues sont les sources $s_1 = c_1^{r_1}$ et $s_2 = c_2^{r_2}$. Grâce à l'estimation des paramètres à partir de mesures réelles, nous disposons alors d'un modèle permettant de simuler des données réalistes \mathbf{x} .

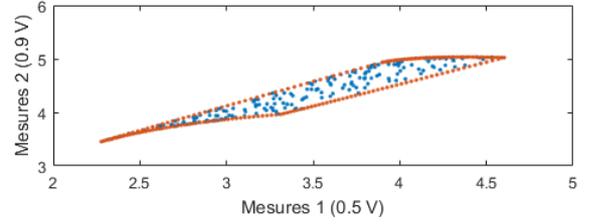


Figure 1 : Diversité des mesures x_1, x_2 , obtenue ici avec des données simulées utilisant deux températures (associées à 0.5 V et 0.9 V) pour des concentrations de gaz uniformément réparties

3 Méthode de séparation de source

Dans cette partie, nous nous proposons de vérifier que le modèle estimé permet de déduire les concentrations de gaz à partir des mesures. Pour cela, il faut vérifier (1) que le système non linéaire est inversible (cas non aveugle). Dans ce cas, la calibration a été réalisée au préalable avec des mélanges de concentrations connues. Il faut également vérifier (2) que les sources peuvent être estimées même lorsque seule la nature du mélange est connue, les paramètres étant inconnus (cas aveugle, méthode non supervisée). La calibration est alors réalisée avec des mesures respectant une diversité suffisante.

3.1 Séparation de sources non aveugle

Nous nous plaçons tout d'abord dans le cas supervisé, et considérons que les paramètres \mathbf{A} , \mathbf{B} et \mathbf{d} du système d'équations sont parfaitement connus. Nous vérifions dans ce paragraphe que nous pouvons retrouver les sources \mathbf{s} à partir des mesures \mathbf{x} . Ceci revient à chercher les zéros de l'équation $G(\mathbf{s}) = 0$ avec :

$$G(\mathbf{s}) = \mathbf{A}\mathbf{s} + \mathbf{B}\mathbf{s}^2 + \mathbf{d}s_1s_2 - \mathbf{x} = 0 \quad (4)$$

Ce problème classique peut être résolu itérativement par un algorithme de Newton-Raphson :

$$\hat{\mathbf{s}}_{k+1} = \hat{\mathbf{s}}_k - \mathbf{J}_G^{-1}(\hat{\mathbf{s}}_k)G(\hat{\mathbf{s}}_k) \quad (5)$$

où $\hat{\mathbf{s}}_k$ est la solution calculée à l'itération k et \mathbf{J}_G la matrice jacobienne de la fonction G .

Cependant, pour converger vers les valeurs souhaitées, la méthode de Newton-Raphson nécessite une bonne initialisation. Si les bornes des sources sont connues, nous pouvons par exemple initialiser $\hat{\mathbf{s}}_0$ de manière uniforme dans cet intervalle ou avec une valeur fixe au milieu de cet intervalle. Si cependant les bornes des sources sont inconnues, nous pouvons initialiser $\hat{\mathbf{s}}_0$ avec la solution obtenue en supposant que le mélange est linéaire. Cette valeur peut être calculée à l'aide d'un algorithme de séparation de sources linéaire aveugle tel que l'algorithme SOBI (Second Order Blind Identification) [8]. Nous comparerons ces différentes initialisations.

3.2 Séparation de sources aveugle

Au lieu d'étalonner chaque capteur à l'aide de mélanges étalon, ce qui requiert une bonne précision, difficile à obtenir pour des petites concentrations (ppm), nous proposons de calibrer le système à l'aide d'une série de mélanges dont les concentrations sont inconnues mais en garantissant une diversité suffisante sur la composition des mélanges. Pour cela, supposant le

modèle (3) dont les paramètres sont maintenant inconnus, nous allons utiliser une méthode de séparation de sources aveugle pour estimer \mathbf{s} et les paramètres \mathbf{A} , \mathbf{B} et \mathbf{d} à partir des mesures. Ando et al. [9] ont développé une méthode de séparation de sources aveugle non-linéaire, appliquée à des capteurs ioniques. Nous allons étudier ici comment l'appliquer sur nos capteurs gaz. Cette méthode itérative consiste à alterner deux algorithmes. Le premier, de type Newton-Raphson (paragraphe 3.1), résout l'équation (6) et permet d'estimer les sources s_j , à l'itération j , avec les paramètres A_j , B_j , et d_j fixés :

$$G(\mathbf{s}; \mathbf{A}, \mathbf{B}, \mathbf{d}) = 0 \quad (6)$$

Le second algorithme estime les coefficients du mélange, en calculant A_{j+1} , B_{j+1} et d_{j+1} par une méthode d'analyse en composantes indépendantes (ICA), supposant que les sources sont indépendantes et en minimisant l'information mutuelle de sources s_j , à l'itération j .

3.3 Correction des sources estimées

Cet algorithme fournit des sources avec une indétermination d'échelle et de translation:

$$\hat{\mathbf{s}}_{corr} = \alpha \hat{\mathbf{s}} + \beta \quad (7)$$

La restauration de la source \mathbf{s} nécessite alors au moins deux mesures de calibration. Les coefficients α et β sont trouvés grâce à une simple régression du premier ordre sur la courbe présentant les sources estimées en fonction des points de calibration.

4 Résultats

De manière à valider cette méthode de séparation de sources aveugle, l'algorithme est testé sur des données simulées et des données réelles, acquises expérimentalement. Nous proposons de comparer les résultats obtenus sur les approches aveugles et non-aveugles. La qualité des estimations est mesurée grâce au coefficient de corrélation entre les sources estimées et les sources initiales, et grâce au rapport signal à bruit (RSB) défini en décibel, tel que :

$$RSB = 10 \cdot \log_{10} \left(\frac{\|\mathbf{s}\|_2}{\|\mathbf{s} - \hat{\mathbf{s}}\|_2} \right) \quad (8)$$

4.1 Estimation non aveugle

4.1.1 Données simulées

Les données sont simulées suivant le modèle présenté dans l'équation (3), avec $N = 200$ points. Nous choisissons des concentrations d'acétone (source 1), variant de manière sinusoïdale dans le temps, et des concentrations d'éthanol (source 2), aléatoires à distribution uniforme. Afin de rester dans les mêmes gammes de concentrations que celles utilisées pour établir le modèle (3) ci-dessus, nous travaillons avec des sources variant de 2,5 à 3,5 ppm^{r1} pour la source 1 et de 4 à 7 ppm^{r2} pour la source 2, ce qui est respectivement l'équivalent de 124 à 730 ppm et de 47 à 223 ppm.

Tout d'abord, nous travaillons sans ajout de bruit, pour vérifier que le système est inversé correctement, puis

avec du bruit ajouté, suivant une loi normale, centrée en 0 et d'écart type la MSE obtenue lors de la régression.

Nous appliquons donc l'algorithme de Newton décrit précédemment sur ces données simulées (figure 2).

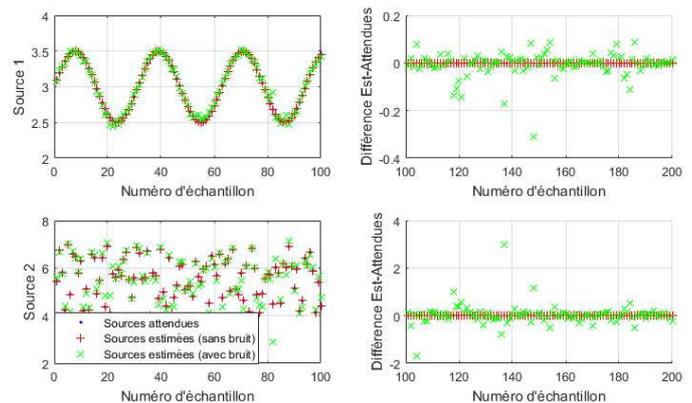


Figure 2 : Estimation des sources avec Newton-Raphson, sur des données simulées avec et sans bruit. L'initialisation ici est le milieu de l'intervalle attendu. A gauche, les sources sur les 100 premiers échantillons, à droite la différence entre les sources estimées et les sources attendues, sur les 100 suivants.

En l'absence de bruit, les sources se retrouvent alors sans problème (les sources initiales et estimées sans bruit sont totalement superposées), avec un coefficient de corrélation de 1 et un RSB de 124 dB. En présence de bruit, l'erreur est un peu plus grande (RSB de 15 dB coefficient de corrélation de 0.98) mais reste encore suffisamment faible pour estimer les sources de manière précise. Par rapport aux autres initialisations, celle-ci donne un résultat de qualité supérieure (Tab 1) :

Tab 1 : Comparaison des différentes initialisations et leur influence sur la qualité de l'estimation dans le cas non aveugle

| Initialisation | Corrélation | | RSB (dB) | |
|----------------|-------------|------------|----------|------------|
| | Sans | Avec bruit | Sans | Avec bruit |
| Uniforme | 0,99 | 0,96 | 19 | 10 |
| Milieu | 1 | 0,98 | 124 | 15 |
| SOBI | 1 | 0,77 | 95 | 9 |

4.1.2 Données expérimentales

Nous appliquons maintenant le même algorithme sur des données expérimentales, comprenant alors des erreurs de mesures et de modélisation (figure 3).

Vu le faible nombre de points, nous ajoutons de fortes contraintes sur les sources, supposant qu'on les connaît à $\pm 20\%$. C'est alors de cette manière que nous les initialisons. Le RSB est de 7 dB et le coefficient de corrélation de 0.94. Les données simulées étaient très indépendantes et un plus grand nombre de points était utilisé, ce qui explique que ces résultats sur données expérimentales sont moins performants que ceux obtenus précédemment.

4.2 Estimation aveugle

Passons maintenant au cas non supervisé, avec les méthodes de séparation de sources aveugles, appliquées sur données simulées (figure 4). Les différentes initialisations sont également testées (Tab 2) :

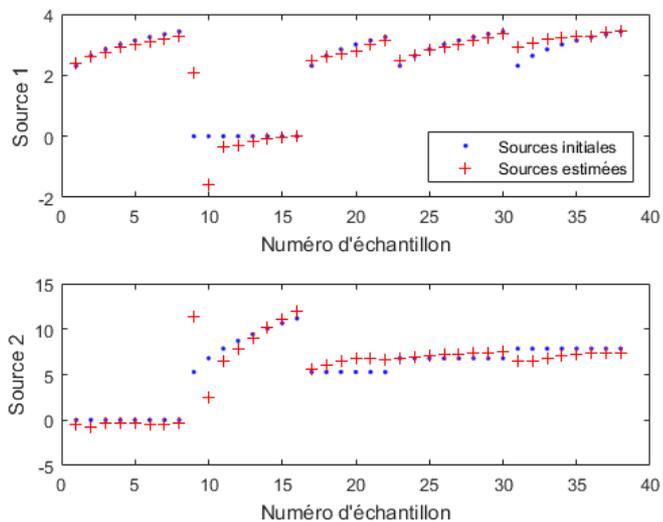


Figure 3 : Estimation des sources avec l'algorithme de Newton-Raphson, sur des données expérimentales, la source 1 étant liée à l'acétone et la source 2 à l'éthanol. L'initialisation des sources a ici été faite à $\pm 20\%$ des sources initiales

Tab 2 : Comparaison des différentes initialisations et leur influence sur la qualité de l'estimation dans le cas non aveugle

| Initialisation | Corrélation | | RSB (dB) | |
|----------------|-----------------|-----------------|-----------------|-----------------|
| | Sans/Avec bruit | Sans/Avec bruit | Sans/Avec bruit | Sans/Avec bruit |
| Uniforme | 0.99 / 0.99 | 16 / 15 | 16 / 15 | 15 / 15 |
| Milieu | 0.98 / 0.99 | 12 / 15 | 12 / 15 | 15 / 15 |
| SOBI | 0.97 / 0.98 | 13 / 13 | 13 / 13 | 13 / 13 |

Nous gardons donc dans ce cas l'initialisation où \hat{s}_0 est tiré au hasard de manière uniforme dans l'intervalle attendu, car cela donne le meilleur résultat, même si l'initialisation au milieu de l'intervalle reste très correcte.

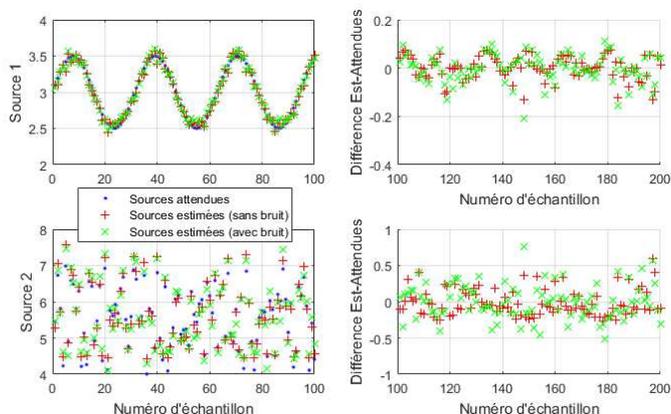


Figure 4 : Estimation aveugle des sources, sur des données simulées avec et sans bruit. L'initialisation ici est tirée uniformément. A gauche, les sources sur les 100 premiers échantillons, à droite la différence entre les sources estimées et les sources attendues, sur les 100 suivants.

Ces résultats restent préliminaires mais montrent que malgré le nombre d'inconnus qui a fortement augmenté (12 avec les coefficients au lieu de 2 uniquement pour les sources), l'estimation des sources reste du bon ordre de grandeur. Des améliorations, que nous discuterons en conclusion, sont cependant à prévoir, pour diminuer ces erreurs d'estimation.

5 Conclusion et perspectives

Dans cette communication, nous avons montré que dans le cas non aveugle, nous pouvons estimer correctement les sources, même s'il reste pour l'instant des erreurs non négligeables sur données expérimentales. Cela valide donc l'inversion de notre modèle non linéaire mais également le protocole d'expérimentation, puisque un seul capteur permet bien de détecter deux gaz, grâce au travail à deux températures. Il restera à passer aux concentrations, par un changement de variables, en ayant préalablement effectué des calibrations pour estimer les coefficients r_i .

Le cas aveugle est pour l'instant dans un état exploratoire mais reste prometteur. L'algorithme itératif utilisé étant fortement sensible à l'initialisation des sources, ce point devra être davantage exploré, notamment en introduisant des contraintes de régularisation sur les solutions recherchées. Afin d'améliorer l'estimation des sources, notamment la robustesse au bruit et aux erreurs de modèle, des méthodes de descente, de type Gauss ou Quasi-Newton, ou alors des méthodes exploitant plus d'a priori, comme la non-négativité (NMF [10]), ou encore des méthodes plus robustes au faible nombre d'échantillons, comme des méthodes récursives [11] ou des méthodes bayésiennes [12], qui ont déjà fait leurs preuves sur des capteurs ioniques avec peu d'échantillons, seront également testées.

6 Références

- [1] W. H. van Geffen, et al., "Diagnosing viral and bacterial respiratory infections in acute COPD exacerbations by an electronic nose: a pilot study," *J. Breath Res.*, vol. 10, no. 3, p. 036001, Jun. 2016.
- [2] T. Toyooka, et al., "A prototype portable breath acetone analyzer for monitoring fat loss," *J. Breath Res.*, vol. 7, no. 3, p. 036005, Sep. 2013.
- [3] K. Yan, et al., "Design of a Breath Analysis System for Diabetes Screening and Blood Glucose Level Prediction," *IEEE Trans. Biomed. Eng.*, vol. 61, no. 11, pp. 2787–2795, Nov. 2014.
- [4] F. Hossein-Babaei, A. Amini, "Recognition of complex odors with a single generic tin oxide gas sensor," *Sens. Actuators B Chem.*, vol. 194, pp. 156–163, Apr. 2014.
- [5] F. Hossein-Babaei, A. Amini, "A breakthrough in gas diagnosis with a temperature-modulated generic metal oxide gas sensor," *Sens. Actuators B Chem.*, vol. 166–167, pp. 419–425, May 2012.
- [6] R. Gosangi and R. Gutierrez-Osuna, "Active temperature modulation of metal-oxide sensors for quantitative analysis of gas mixtures," *Sens. Actuators B Chem.*, vol. 185, pp. 201–210, Aug. 2013.
- [7] S. Madrolle, et al., "Dual-temperature mode for quantitative analysis of gas mixtures with MOX sensor," in *2017 ISOCS/IEEE International Symposium on Olfaction and Electronic Nose (ISOEN 2017)*, Montreal, Canada, 2017.
- [8] A. Belouchrani, et al., "A blind source separation technique using second-order statistics," *IEEE Trans. Signal Process.*, vol. 45, no. 2, pp. 434–444, 1997.
- [9] R. A. Ando, et al., "A blind source separation method for chemical sensor arrays based on a second order mixing model," in *Signal Processing Conference (EUSIPCO), 2015 23rd European*, 2015, pp. 933–937.
- [10] S. Hosseini, et al., "Extending NMF to blindly separate linear-quadratic mixtures of uncorrelated sources," in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, 2016, pp. 1–6.
- [11] S. Hosseini, Y. Deville, "Recurrent networks for separating extractable-target nonlinear mixtures. Part II. Blind configurations," *Signal Process.*, vol. 93, no. 4, pp. 671–683, Apr. 2013.
- [12] L. T. Duarte, et al., "Bayesian source separation of linear and linear-quadratic mixtures using truncated priors," *J. Signal Process. Syst.*, vol. 65, no. 3, pp. 311–323, 2011.