

Une approche distribuée asynchrone pour la factorisation en matrices non-négatives – Application au démixage hyperspectral

Pierre-Antoine Thouvenin, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIH



XXVIe Colloque GRETSI, Juan-les-Pins, France

8 Septembre 2017

Contexte

Image hyperspectrale

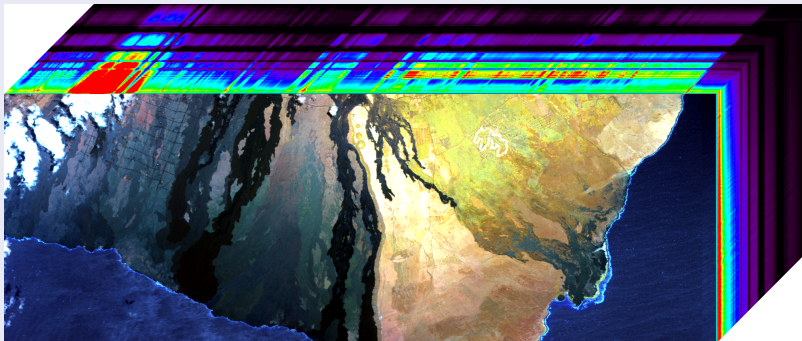
- même scène observée à plusieurs longueurs d'onde,

Contexte

Image hyperspectrale

- même scène observée à plusieurs longueurs d'onde,

Cube hyperspectral



Contexte

Image hyperspectrale

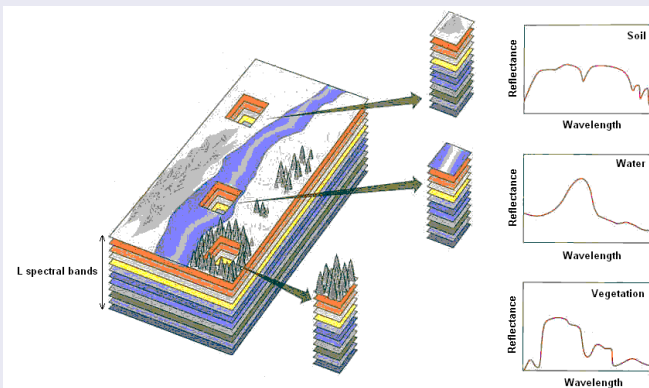
- même scène observée à plusieurs longueurs d'onde,
- pixel représenté par un vecteur de plusieurs centaines de mesures.

Contexte

Image hyperspectrale

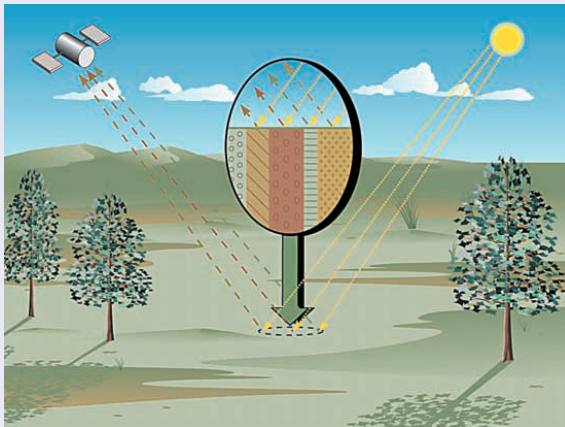
- même scène observée à plusieurs longueurs d'onde,
- pixel représenté par un vecteur de plusieurs centaines de mesures.

Cube hyperspectral



Démélange spectral

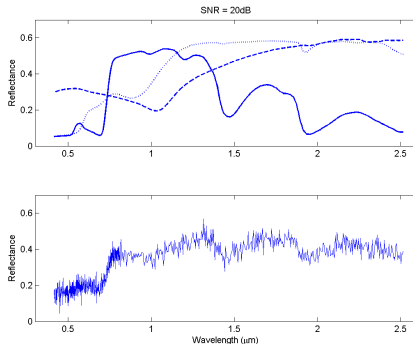
Modèle de mélange linéaire : $\mathbf{y}_n = \sum_{r=1}^R a_{n,r} \mathbf{m}_r + \mathbf{b}_n$



Référence : IEEE Signal Proc. Magazine, Jan. 2002.

Démélange spectral

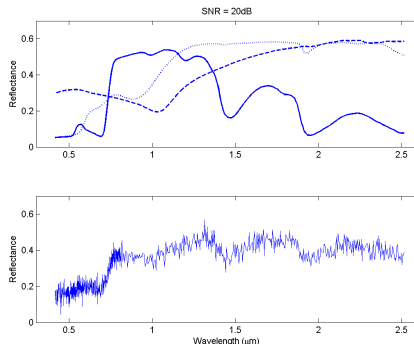
Modèle de mélange linéaire : $\mathbf{y}_n = \sum_{r=1}^R \mathbf{a}_{n,r} \mathbf{m}_r + \mathbf{b}_n$



- $L = 825$ ($0.4\mu\text{m} \rightarrow 2.5\mu\text{m}$),
- $R = 3$:
 - herbe verte (trait plein),
 - métal galvanisé (tirets),
 - brique rouge (pointillés),
- $\mathbf{a}_n = [0.3, 0.6, 0.1]^T$,
- $\text{SNR} \approx 20\text{dB}$.

Démélange spectral

Modèle de mélange linéaire : $\mathbf{y}_n = \sum_{r=1}^R \mathbf{a}_{n,r} \mathbf{m}_r + \mathbf{b}_n$



- $L = 825$ ($0.4\mu\text{m} \rightarrow 2.5\mu\text{m}$),
- $R = 3$:
 - herbe verte (trait plein),
 - métal galvanisé (tirets),
 - brique rouge (pointillés),
- $\mathbf{a}_n = [0.3, 0.6, 0.1]^T$,
- $\text{SNR} \approx 20\text{dB}$.

Formulation du problème

Estimation de \mathbf{a}_n sous des contraintes de **positivité** et **additivité** et $\mathbf{m}_1, \dots, \mathbf{m}_R$ sous des contraintes **positivité**.

Motivations

- Images composées d'un grand nombre de pixels \Rightarrow mise en œuvre de procédures d'estimation distribuées

Contexte de distribution des données

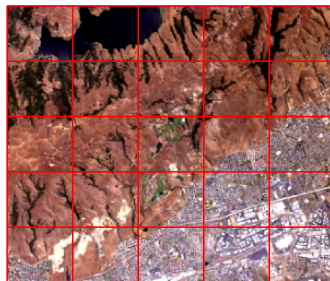


FIGURE 2 – Image hyperspectrale scindée en plusieurs sous-blocs, assignés chacun à un nœud esclave



FIGURE 3 – Collection d'images hyperspectrales, dont chacune est assignée à un nœud esclave

Motivations

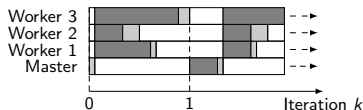
- Images composées d'un grand nombre de pixels \Rightarrow mise en œuvre de procédures d'estimation distribuées

Motivations

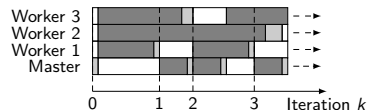
- Images composées d'un grand nombre de pixels \Rightarrow mise en œuvre de procédures d'estimation distribuées
 - ▷ architecture de type maître-esclave ;

Motivations

- Images composées d'un grand nombre de pixels \Rightarrow mise en œuvre de procédures d'estimation distribuées
 - ▷ architecture de type maître-esclave ;
- Asynchronicité : permet d'exploiter les différences de performance entre machines/processeurs utilisés, réduisant le temps de calcul nécessaire par rapport à un algo. synchrone ;



(a) Système synchrone.

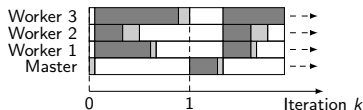


(b) Système asynchrone.

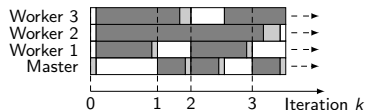
FIGURE 4 – Exemples de comportement d'un algorithme synchrone et asynchrone (temps d'attente en blanc, transmission en gris clair, traitement en gris).

Motivations

- Images composées d'un grand nombre de pixels \Rightarrow mise en œuvre de procédures d'estimation distribuées
 - ▷ architecture de type maître-esclave ;
- Asynchronicité : permet d'exploiter les différences de performance entre machines/processeurs utilisés, réduisant le temps de calcul nécessaire par rapport à un algo. synchrone ;



(a) Système synchrone.



(b) Système asynchrone.

FIGURE 4 – Exemples de comportement d'un algorithme synchrone et asynchrone (temps d'attente en blanc, transmission en gris clair, traitement en gris).

- Grand nombre de méthodes asynchrones récemment proposées dans la littérature [1]–[11]
 - ▷ nombre limité de méthodes d'estimation proposées pour le démixage hyperspectral, toutes synchrones [12]–[14].

Objectif

Mise en œuvre d'un algorithme de démelange asynchrone inspiré de l'algorithme PALM [15], [16]

- cadre adapté au démelange, pour lequel des garanties de convergence sont disponibles [15]–[17] ;
- la version standard de l'algo. se prête facilement à la mise en place d'une procédure distribuée synchrone ;
- cadre approprié pour évaluer l'intérêt de l'asynchronicité.

Plan de la présentation

1 Introduction

- Introduction
- Motivations

2 Une approche distribuée asynchrone

- Modèle de mélange
- Formulation du problème
- Démélange en présence de variabilité spectrale
- Algorithme de démélange distribué asynchrone

3 Simulations sur données synthétiques

- Simulations en l'absence de variabilité
- Simulations en présence de variabilité

4 Conclusion et perspectives

Plan de la présentation

- 1 Introduction
 - Introduction
 - Motivations
- 2 Une approche distribuée asynchrone
 - Modèle de mélange
 - Formulation du problème
 - Démélange en présence de variabilité spectrale
 - Algorithme de démélange distribué asynchrone
- 3 Simulations sur données synthétiques
 - Simulations en l'absence de variabilité
 - Simulations en présence de variabilité
- 4 Conclusion et perspectives

Modèle de mélange

Modèle de mélange linéaire (MML)

- Les N pixels \mathbf{y}_n sont représentés par une combinaison convexe de R spectres purs \mathbf{m}_r [18] ;
- Données regroupées en $I < N$ groupes disjoints de J pixels ;

$$\forall i \in \{1, \dots, I\}, \mathbf{Y}_i = \mathbf{M}\mathbf{A}_i + \mathbf{B}_i \quad (1)$$

Contraintes :

$$\mathbf{A}_i^T \mathbf{1}_R = \mathbf{1}_J, \quad \mathbf{A}_i \succeq \mathbf{0}_{R,J}, \quad \mathbf{M} \succeq \mathbf{0}_{L,R} \quad (2)$$

Remarques :

- analyse d'une image scindée en différents blocs, ou d'une collections d'images composées des mêmes spectres purs (e.g., données multi-temporelles) ;
- l'estimation des coefficients d'abondance peut s'effectuer en parallèle.

Formulation du problème

Problème de démixage

$$(\mathbf{A}^*, \mathbf{M}^*) \in \arg \min_{\mathbf{A}, \mathbf{M}} \left\{ F(\mathbf{A}, \mathbf{M}) + g(\mathbf{A}) + \iota_{\mathcal{A}_{R,N}}(\mathbf{A}) + h(\mathbf{M}) + \iota_{\{\cdot \succeq \mathbf{0}\}}(\mathbf{M}) \right\} \quad \text{avec} \quad (3)$$

$$F(\mathbf{A}, \mathbf{M}) = \sum_{i=1}^I f_i(\mathbf{M}, \mathbf{A}_i) = \frac{1}{2} \sum_{i=1}^I \|\mathbf{Y}_i - \mathbf{M}\mathbf{A}_i\|_{\mathbb{F}}^2$$

$$\mathcal{A}_{R,J} = \{ \mathbf{A} \in \mathbb{R}^{R \times J} : \mathbf{a}_n \in \mathcal{S}_R, \forall n \in \{1, \dots, J\} \}$$

$$\mathcal{S}_R = \{ \mathbf{x} \in \mathbb{R}^R : x_r \geq 0 \text{ et } \mathbf{x}^T \mathbf{1}_R = 1 \}$$

- $\iota_{\mathcal{S}}$: indicatrice de l'ensemble \mathcal{S} ($\iota_{\mathcal{S}} = 0$ si $\mathbf{x} \in \mathcal{S}$, $+\infty$ sinon)
- g et h : pénalités convexes appropriées
 - ▷ de façon générale : g et h séparables pour distribuer la résolution de (3).
 - ▷ par la suite : $g = 0$, h distance mutuelle entre les spectres purs [19].

Démélange en présence de variabilité spectrale (I)

- Les spectres purs peuvent présenter une variabilité spectrale significative en fonction des conditions d'acquisition (illumination, ...)

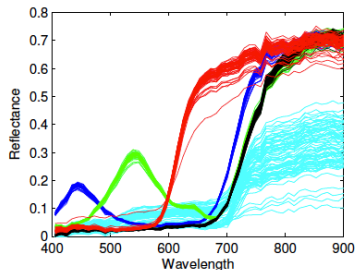


FIGURE 5 – Illustration de la variabilité des spectres purs (extraite de [20]).

Démélange en présence de variabilité spectrale (I)

- Les spectres purs peuvent présenter une variabilité spectrale significative en fonction des conditions d'acquisition (illumination, ...)

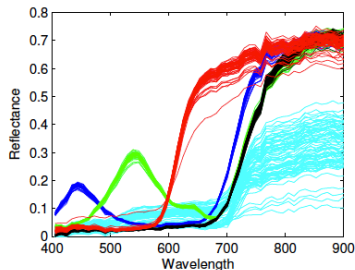


FIGURE 5 – Illustration de la variabilité des spectres purs (extraite de [20]).

- ⇒ possibles erreurs d'estimations, propagées dans le cadre de procédures non-supervisées ;
- ⇒ variabilité prise en compte à l'aide d'un modèle de mélange linéaire perturbé [21] ;
- Estimation des coefficients d'abondance et des termes de variabilité en parallèle.

Démélange en présence de variabilité spatiale (II)

Modèle de mélange linéaire perturbé (MMLP)

$$\mathbf{y}_n = \sum_{r=1}^R (\mathbf{m}_r + \mathbf{d}\mathbf{m}_{r,n}) a_{r,n} + \mathbf{b}_n \quad (4)$$

Contrainte additionnelle :

$$\|\mathbf{d}\mathbf{M}_n\|_{\mathbb{F}}^2 \leq \varepsilon, \text{ où } \mathbf{d}\mathbf{M}_n = [\mathbf{d}\mathbf{m}_{1,n}, \dots, \mathbf{d}\mathbf{m}_{R,n}] \quad (5)$$

Remarques :

- distribution des données similaire à celle adoptée dans le cas du modèle linéaire ;
- l'estimation des coefficients d'abondance et des termes de variabilité peut s'effectuer en parallèle ;
- pas de contrainte de positivité sur les spectres perturbés (limitation lié à l'asynchronie...)
- distribution des données : indifféremment applicable à une image scindée en plusieurs sous-ensembles ou à une collection d'images hyperspectrales.

Distribution des données

- Chaque nœud de calcul responsable de la mise à jour d'un sous-ensemble des coefficients d'abondance ;
- Le nœud maître chargé d'agréger les données des différents nœuds pour la mise à jour des spectres purs.

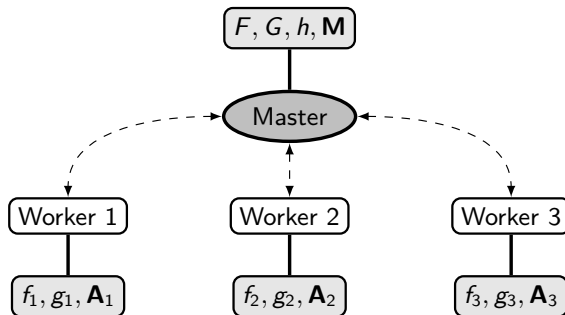


FIGURE 6 – Illustration de la distribution des variables pour le démixage basé sur le MML (pour $I = 3$ blocs).

Algorithme de démélange distribué asynchrone (MML)

Algorithme 1 : Algorithme du i^{e} nœud esclave.

Données : \mathbf{M} , $\tilde{\mathbf{A}}_i$.

début

 Réception des données mises à jour par le processus maître, $(\mathbf{M}, \tilde{\mathbf{A}}_i)$;

$\hat{\mathbf{A}}_i \in \text{prox}_{\ell_{\mathcal{A}_{R,J}}} \left(\tilde{\mathbf{A}}_i - \frac{1}{c_{\mathbf{A}_i}} \nabla_{\mathbf{A}_i} f_i(\tilde{\mathbf{A}}_i, \mathbf{M}) \right)$;

 Transmission de $\hat{\mathbf{A}}_i$ au processus maître;

Résultats : $\hat{\mathbf{A}}_i$.

Algorithme de démelange distribué asynchrone (MML)

Algorithme 2 : Algorithme du maître.

Données : \mathbf{A}^0 , \mathbf{M}^0 , $\mu \in]0, 1[$ ($\mu = 10^{-6}$).

$\gamma^0 \leftarrow 1$, $k \leftarrow 0$;

tant que critère d'arrêt non satisfait, **faire**

 Attente de $\tilde{\mathbf{A}}_{i^k}$ provenant de l'un des nœuds ;

 // Mise à jour des abondances

$$\mathbf{A}_i^{k+1} = \begin{cases} \mathbf{A}_i^k + \gamma^k (\tilde{\mathbf{A}}_i^k - \mathbf{A}_i^k), & i = i^k \\ \mathbf{A}_i^k, & i \neq i^k \end{cases} ;$$

 // Mise à jour des spectres purs

$$\tilde{\mathbf{M}}^k = \text{prox}_{\ell_{\{\cdot, \succeq 0\}}} \left(\mathbf{M}^k - \frac{1}{\nu^k} \nabla_{\mathbf{M}} [F(\mathbf{A}^{k+1}, \mathbf{M}^k) + h(\mathbf{M}^k)] \right);$$

$$\mathbf{M}^{k+1} = \mathbf{M}^k + \gamma^k (\tilde{\mathbf{M}}^k - \mathbf{M}^k);$$

 // Mise à jour du pas de relaxation

$$\gamma^{k+1} = \gamma^k (1 - \mu \gamma^k);$$

 // Transmission des résultats au processus i^k

 Transmettre $(\mathbf{M}^{k+1}, \mathbf{A}_{i^k}^{k+1})$ au processus i^k ;

$k \leftarrow k + 1$;

Résultats : \mathbf{A}^k , \mathbf{M}^k .

Algorithme similaire pour le MMLP

Plan de la présentation

- 1 Introduction
 - Introduction
 - Motivations
- 2 Une approche distribuée asynchrone
 - Modèle de mélange
 - Formulation du problème
 - Démélange en présence de variabilité spectrale
 - Algorithme de démélange distribué asynchrone
- 3 Simulations sur données synthétiques
 - Simulations en l'absence de variabilité
 - Simulations en présence de variabilité
- 4 Conclusion et perspectives

Simulations sur données synthétiques (MML) (I)

- Données : 3 images hyperspectrales, mélanges linéaires de $R = 9$ spectres purs, $L = 413$ bandes ;
- Images : 100×100 pixels, bruit blanc additif gaussien de variance telle que $RSB = 30$ dB ;
- Évaluation des performances pour $I = 3$ processus ;
- Initialisation : VCA [22] / FCLS [23] ;
- Comparaison de l'algo. avec son analogue synchrone, implantés en Julia [24].

TABLE 1 – Résultats de simulation sur données synthétiques.

	Sync.	Async.
aSAM(M) (°)	9.74e-01	1.04e+00
GMSE(A)	3.48e-04	5.25e-04
RE	1.05e-04	1.07e-04
aSAM(Y) (°)	2.23e-02	2.24e-02
temps (s)	1.39e+03	3.33e+02

Simulations sur données synthétiques (MML) (II)

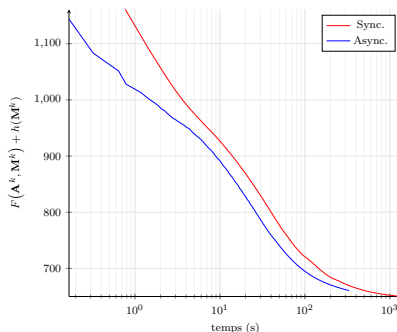


FIGURE 7 – Évolution de la fonction objectif au cours du temps, obtenue avec les différentes versions de l'algorithme PALM (jusqu'à convergence).

Remarques :

- intérêt de l'asynchronie en terme de temps de calcul (pour parvenir à convergence) ;
- légère dégradation des performances d'estimation.

Simulations sur données synthétiques (MMLP) (I)

- Données : 3 images hyperspectrales, mélanges linéaires perturbés de $R = 3$ spectres purs, $L = 413$ bandes ;
- Images : 100×100 pixels, bruit blanc additif gaussien de variance telle que $RSB = 30$ dB ;
- Génération de la variabilité : produit (terme à terme) de chacun des spectres purs de référence avec des fonctions affines par morceaux générées aléatoirement (fonctions différentes pour chacun des pixels et chaque matériau)
 - ▷ obtention de variabilité spatialement et spectralement variable.

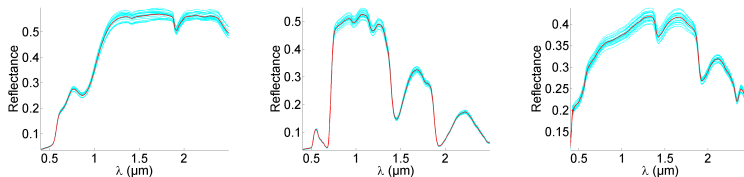


FIGURE 8 – Spectres purs de référence (en rouge) et exemples de spectres affectés par la variabilité générée (en cyan).

Simulations sur données synthétiques (MMLP) (II)

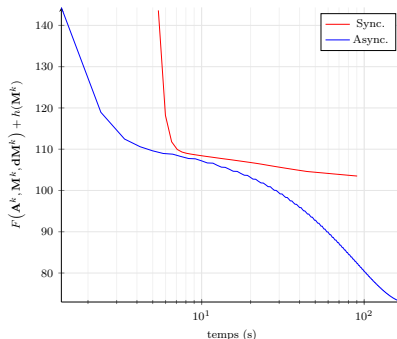


FIGURE 9 – Évolution de la fonction objectif au cours du temps pour les versions synchrone et asynchrone de l'algorithme PALM (jusqu'à convergence).

Remarques :

- intérêt de l'asynchronie en terme de temps de calcul pour ce problème (pour ce jeu de données certes... mais ce n'est pas toujours le cas) ;
- en fonction des données, l'algo. peut converger vers un point stationnaire moins intéressant que la version synchrone.

Simulations sur données synthétiques (MMLP) (III)

	Sync.	Async.
aSAM(M) (°)	1.28e+00	1.47e+00
GMSE(A)	6.73e-04	8.52e-04
GMSE(dM)	2.17e-05	2.56e-05
RE	4.08e-05	2.88e-05
aSAM(Y) (°)	1.92e-02	1.57e-02
temps (s)	9.05e+01	0.73

TABLE 2 – Résultats de simulation sur données synthétiques (présence de variabilité).

Plan de la présentation

- 1 Introduction
 - Introduction
 - Motivations
- 2 Une approche distribuée asynchrone
 - Modèle de mélange
 - Formulation du problème
 - Démélange en présence de variabilité spectrale
 - Algorithme de démélange distribué asynchrone
- 3 Simulations sur données synthétiques
 - Simulations en l'absence de variabilité
 - Simulations en présence de variabilité
- 4 Conclusion et perspectives

Conclusion et perspectives

Travail effectué

- ▶ apport d'une procédure de démelange asynchrone [17] basée sur PALM [15], [16] par rapport à sa version synchrone ;
- ▶ intérêt pour le démelange d'images hyperspectrales multi-temporelles.

Perspectives

- ▷ envisager différentes configurations de distribution des données [4], [25]
⇒ améliorations de l'algorithme à envisager en présence de variabilité spatiale/temporelle ;
- ▷ possibilité d'inclure des opérateurs proximaux approchés [16] ;
- ▷ généralisations à d'autres problèmes de NMF : possibilité de traiter d'autres termes d'attaches aux données (e.g., s'affranchir en partie de l'hypothèse gradient Lipschitz pour le terme différentiable ? [26], [27]).

Merci pour votre attention !

Une approche distribuée asynchrone pour la factorisation en matrices non-négatives – Application au démixage hyperspectral

Pierre-Antoine Thouvenin, Nicolas Dobigeon and Jean-Yves Tourneret

University of Toulouse, IRIT/INP-ENSEEIH



XXVIe Colloque GRETSI, Juan-les-Pins, France

8 Septembre 2017

- [1] P. L. COMBETTES et J. ECKSTEIN, “Asynchronous block-iterative primal-dual decomposition methods for monotone inclusions”, *Math. Program., Ser. B*, p. 1–28, 2016.
- [2] Z. PENG, Y. XU, M. YAN et W. YIN, “Arock : an algorithmic framework for asynchronous parallel coordinate updates”, *SIAM J. Sci. Comput.*, t. 38, n° 5, p. 2851–2879, sept. 2016.
- [3] P. BIANCHI et J. JAKUBOWICZ, “Convergence of a multi-agent projected stochastic gradient algorithm for non-convex optimization”, *IEEE Trans. Autom. Control*, t. 58, n° 2, p. 391–405, fév. 2013.
- [4] J.-C. PESQUET et A. REPETTI, “A Class of Randomized Primal-Dual Algorithms for Distributed Optimization”, *Journal of nonlinear and convex analysis*, t. 16, n° 12, p. 2453–2490, nov. 2015.
- [5] S. SRA, A. W. YU, M. LI et A. SMOLA, “Adadelay : delay adaptive distributed stochastic optimization”, in *Proc. Int. Conf. Artificial Intelligence and Statistics (AISTATS)*, Cadiz, Spain, 2016, p. 957–965.
- [6] M. LI, D. G. ANDERSEN, A. SMOLA et K. YUY, “Communication efficient distributed machine learning with the parameter server”, in *Adv. in Neural Information Processing Systems*, 2014, p. 91–27.
- [7] D. DAVIS, “The asynchronous PALM algorithm for nonsmooth nonconvex problems”, , avr. 2016, submitted. adresse : <https://arxiv.org/abs/1604.00526>.

- [8] T.-H. CHANG, M. HONG, W.-C. LIAO et X. WANG, "Asynchronous Distributed ADMM for Large-Scale Optimization–Part I : Algorithm and Convergence Analysis", *IEEE Trans. Signal Process.*, t. 62, n° 12, p. 3118–3130, juin 2016.
- [9] F. FACCHINEI, G. SCUTARI et S. SAGRATELLA, "Parallel selective algorithms for nonconvex big data optimization", *IEEE Trans. Signal Process.*, t. 63, n° 7, p. 1874–1889, avr. 2015.
- [10] Z. PENG, Y. XU, M. YAN et W. YIN, "On the convergence of asynchronous parallel iteration with arbitrary delays", , déc. 2016, arXiv preprint. adresse : <https://arxiv.org/abs/1612.04425>.
- [11] G. SCUTARI, F. FACCHINEI, L. LAMPARIELLO et P. SONG, "Parallel and distributed methods for nonconvex optimization-part i : theory", *IEEE Trans. Signal Process.*, t. 65, n° 8, p. 1929–2944, avr. 2017.
- [12] S. A. ROBILA et D. RICART, "Distributed algorithms for unmixing hyperspectral data using nonnegative matrix factorization with sparsity constraints", in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, Melbourne, Australia, juil. 2013, p. 2156–2159.
- [13] J. SIGURDSSON, M. O. ULFARSSON, J. R. SVEINSSON et J. M. BIOUSCAS-DIAS, "Sparse distributed hyperspectral unmixing", in *Proc. IEEE Int. Conf. Geosci. Remote Sens. (IGARSS)*, Beijing, China, juil. 2016, p. 6994–6997.

- [14] —, “Sparse distributed multitemporal hyperspectral unmixing”, *IEEE Trans. Geosci. Remote Sens.*, 2017, accepted.
- [15] J. BOLTE, S. SABACH et M. TEBoulLE, “Proximal alternating linearized minimization for nonconvex and nonsmooth problems”, *Mathematical Programming*, t. 1-2, n° 146, p. 459–494, juil. 2013.
- [16] E. CHOUZENOUX, J.-C. PESQUET et A. REPETTI, “A block coordinate variable metric forward-backward algorithm”, *J. Glob. Optim.*, t. 66, n° 3, p. 547–485, 2016.
- [17] L. CANNELLI, F. FACCHINEL, V. KUNGURTSEV et G. SCUTARI, “Asynchronous Parallel Algorithms for Nonconvex Big-Data Optimization. Part I : Model and Convergence”, , juil. 2016, arXiv preprint. adresse : <https://arxiv.org/abs/1607.04818>.
- [18] J. M. BIOUCAS-DIAS, A. PLAZA, N. DOBIGEON, M. PARENTE, Q. DU, P. GADER et J. CHANUSSOT, “Hyperspectral unmixing overview : geometrical, statistical, and sparse regression-based approaches”, *IEEE J. Sel. Topics Appl. Earth Observ. in Remote Sens.*, t. 5, n° 2, p. 354–379, avr. 2012.
- [19] M. BERMAN, H. KIIVERI, R. LAGERSTROM, A. ERNST, R. DUNNE et J. F. HUNTINGTON, “ICE : a statistical approach to identifying endmembers in hyperspectral images”, *IEEE Trans. Geosci. Remote Sens.*, t. 42, n° 10, p. 2085–2095, oct. 2004.

- [20] P. GADER, A. ZARE, R. CLOSE, J. AITKEN et G. TUELL, "MUUFL gulfport hyperspectral and LiDAR airborne data set", University of Florida, Gainesville, FL, rapp. tech. REP-2013-570, oct. 2013.
- [21] P.-A. THOUVENIN, N. DOBIGEON et J.-Y. TOURNERET, "Hyperspectral unmixing with spectral variability using a perturbed linear mixing model", *IEEE Trans. Signal Process.*, t. 64, n° 2, p. 525–538, jan. 2016.
- [22] J. M. NASCIMENTO et J. M. BIOUCAS-DIAS, "Vertex component analysis : a fast algorithm to unmix hyperspectral data", *IEEE Trans. Geosci. Remote Sens.*, t. 43, n° 4, p. 898–910, avr. 2005.
- [23] D. C. HEINZ et C. -I CHANG, "Fully constrained least-squares linear spectral mixture analysis method for material quantification in hyperspectral imagery", *IEEE Trans. Geosci. Remote Sens.*, t. 29, n° 3, p. 529–545, mar. 2001.
- [24] J. BEZANSON, A. EDELMAN, S. KARPINSKI et V. B. SHAH, "Julia : a fresh approach to numerical computing", *SIAM Review*, t. 59, n° 1, p. 65–98, 2017. DOI : 10.1137/141000671.
- [25] P. BIANCHI, W. HACHEM et F. IUTZELER, "A Coordinate Descent Primal-Dual Algorithm and Application to Distributed Asynchronous Optimization", *IEEE Trans. Autom. Control*, t. 61, n° 10, p. 2947, oct. 2016.

- [26] H. H. BAUSCHKE, J. BOLTE et M. TEBoulLE, "A descent lemma beyond Lipschitz gradient continuity : first-order methods revisited and applications", *Mathematics of Operations Research*, t. 2, n° 2, p. 330–348, 2016.
- [27] J. BOLTE, S. SABACH, M. TEBoulLE et Y. VAISBOURD, "First order methods beyond convexity and Lipschitz gradient continuity with applications to quadratic inverse problems", , 2017, arXiv preprint. adresse : <https://arxiv.org/abs/1706.06461v1>.