

Comment "robustifier" les estimateurs par dictionnaire face aux erreurs de grille ?

Stéphanie BERNHARDT¹, Rémy BOYER², Sylvie MARCOS², Pascal LARZABAL³

¹Thales Systèmes Aéroportés, Elancourt, France

²Laboratoire des Signaux et Systèmes (L2S), Université Paris Sud-CNRS-CentraleSupélec, Gif-Sur-Yvette, France

³SATIE, Université Paris Sud, Cachan, France

stephanie.gourdin@fr.thalesgroup.com, remy.boyer@l2s.centralesupelec.fr

Résumé – Considérons le problème très général de l’estimation d’un paramètre d’intérêt tel que $\theta \in \mathbb{P}$ où \mathbb{P} est un ensemble à valeurs continues. La stratégie des estimateurs à contrainte de parcimonie repose sur la recherche d’un représentant issu d’une discrétisation régulière de l’ensemble \mathbb{P} , souvent appelée grille. Dans un contexte réaliste, le problème d’erreur de grille est donc intrinsèque et limite de manière drastique les performances en estimation des algorithmes standards dans le régime des hauts RSBs où l’erreur de grille devient prépondérante au regard de l’erreur due au bruit. Dans ce contexte, il existe dans la littérature de nombreuses contributions traitant de ce problème. Selon nos connaissances du domaine, les approches proposées peuvent être qualifiées de techniques *ad-hoc* dédiées à un estimateur en particulier. Dans ce travail, notre objectif est de proposer un post-traitement générique, nommé OGEC (Off-Grid Error Correction), dans le sens où celui-ci pourra être exploité pour "robustifier" aux erreurs de grille tout estimateur à contrainte de parcimonie existant dans la littérature. De plus, nous souhaitons que cette sur-couche soit la moins couteuse possible en terme de temps de calcul. Après avoir présenté l’estimateur OGEC, nous analysons son biais et son EQM de manière théorique. Enfin, les performances de l’OGEC sont illustrées dans le contexte de l’échantillonnage comprimé de signaux à bande non-limitée.

Abstract – Let us consider the very general problem of the estimation of a parameter of interest $\theta \in \mathbb{P}$, where \mathbb{P} is a set of continuous values. Sparse based estimation is based on the matching of the parameters of interest to a regular discretization of \mathbb{P} , often referred to as the grid. In realistic scenarios, the off-grid error problem is thus inherent to the presence of the grid and limits in a drastic way the estimation performances of the standard sparse estimation algorithms in the high SNR regime where the off-grid error is predominant with respect to the error due to the noise. In this context, numerous contributions in the literature deals with this problem, but according to our knowledge, the proposed approaches can be defined as ad hoc techniques dedicated to a particular estimator. In this work, our objective is to propose a generic post-treatment, called OGEC (Off-Grid Error Correction), in the sense that it can be used to "robustify" against the off-grid problem any existing sparse based estimator. Furthermore, we design this post-treatment to be cheap in term of calculation cost. OGEC being introduced, we derive theoretically its bias and MSE and finally its performances are illustrated in the context of the compressed sampling of non-bandlimited signals.

1 Contexte du problème

1.1 Modèles à contrainte de parcimonie

Estimer un vecteur parcimonieux à partir d’un ensemble de mesures bruitées est un problème fondamental et d’actualité. Considérons le modèle paramétrique suivant :

$$y_n = \sum_{l=1}^L x_l h(\theta_l, t_n) + w_n, \quad 1 \leq n \leq N$$

dans lequel la fonction $h(\theta, t)$, supposée de classe C^1 , propre au contexte applicatif (radar, communication sans fil, ...) dépend du l -ième paramètre d’intérêt θ_l à valeur réelle sur un ensemble \mathbb{P} et de l’instant de mesure t_n . L’amplitude non-nulle de la l -ième composante est notée x_l et w_n modélise le n -ième échantillon de bruit suivant un processus i.i.d. Gaussien de moyenne nulle et de variance σ^2 . Ce modèle peut être écrit en discrétisant régulièrement l’ensemble \mathbb{P} selon une grille $\bar{\theta} =$

$[\bar{\theta}_1, \dots, \bar{\theta}_K]^T$ où $K \gg N > L$. Il est alors possible de construire un dictionnaire redondant (plus de colonnes que de lignes) noté $\mathbf{H}_{\mathcal{K}}$ à partir de la discrétisation de la fonction $h(\bar{\theta}_k, t_n)$ sur l’ensemble $\mathcal{K} = \{1, \dots, K\}$. Le vecteur de mesures peut alors s’écrire sous la forme d’un modèle à amplitudes parcimonieuses, selon

$$\mathbf{y} = \mathbf{H}_{\mathcal{K}} \mathbf{s} + \mathbf{w}$$

où le vecteur de bruit $\mathbf{w} = [w_1, \dots, w_N]^T$ et le vecteur L -parcimonieux sur l’ensemble \mathcal{S} , \mathbf{s} , contient les amplitudes non-nulles x_l quand $\bar{\theta}_k = \theta_l$, c’est-à-dire quand le paramètre d’intérêt θ_l appartient à la grille. En pratique, ce scénario est irréaliste et le problème d’erreur de grille où $\bar{\theta}_k \neq \theta_l$ ne peut être ignoré.

1.2 Prise en compte de l’erreur de grille

Notons $h(\bar{\theta}_{k_l} + e_l, \mathbf{t})$ la fonctionnelle $h(\theta, t)$ où $\bar{\theta}_{k_l}$ est l’élément de la grille le plus proche de θ_l et $\mathbf{t} = [t_1, \dots, t_N]^T$. On

a alors $\theta_l = \bar{\theta}_{k_l} + e_l$ où e_l est l'erreur de grille. Pour une erreur de grille petite devant $O(e_l^2)$, l'approximation de Taylor à l'ordre un de la fonction, nous permet de formuler le modèle pour N mesures entachées de l'erreur de grille selon

$$\begin{aligned} \mathbf{y} &= \sum_{l=1}^L s_l \cdot h(\bar{\theta}_{k_l} + e_l, \mathbf{t}) + \mathbf{w} \\ &= \sum_{l=1}^L s_l \cdot h(\bar{\theta}_{k_l}, \mathbf{t}) + e_l \cdot s_l \left. \frac{\partial h(\theta, \mathbf{t})}{\partial \theta} \right|_{\theta=\bar{\theta}_{k_l}} + \mathbf{w} \\ &= \underbrace{\left(\mathbf{H}_S + \dot{\mathbf{H}}_S \text{diag}(\mathbf{e}) \right)}_{\tilde{\mathbf{H}}_S} \mathbf{x} + \mathbf{w} \end{aligned}$$

où $\mathbf{e} = [e_1, \dots, e_L]^T$ et

$$\dot{\mathbf{H}}_S = \left[\left. \frac{\partial h(\theta, \mathbf{t})}{\partial \theta} \right|_{\theta=\bar{\theta}_{k_1}}, \dots, \left. \frac{\partial h(\theta, \mathbf{t})}{\partial \theta} \right|_{\theta=\bar{\theta}_{k_L}} \right]_{N \times L}.$$

Cette modélisation dite structurée est discutée dans [1] et permet de mettre en oeuvre de manière opérationnelle des estimateurs conjoints de \mathbf{s} , du support \mathcal{S} et de l'erreur de grille \mathbf{e} [2]. Différents algorithmes ont été proposés pour prendre en compte l'erreur structurée par l'algorithme de reconstruction parcimonieux tel que le LASSO [3]. L'estimateur proposé ici se place en aval d'une estimation déjà obtenue par un algorithme souffrant de l'erreur de grille.

2 Estimateur proposé

Un estimateur robuste à l'erreur de grille doit

1. Corriger l'estimation de \mathbf{x} en estimant l'erreur de grille \mathbf{e} apparaissant dans le régime des hauts RSBs [4].
2. Ne pas perturber l'estimation de \mathbf{x} dans le régime des bas RSBs où l'erreur de grille est négligeable devant la corruption due au bruit \mathbf{w} .
3. Avoir une complexité calculatoire raisonnable.
4. Etre générique, c'est-à-dire, être en mesure de "robustifier" tout estimateur à contrainte de parcimonie standards.

2.1 Description de l'estimateur OGEC

L'algorithme proposé nommé OGEC pour "Off-Grid Error Correction" satisfait à ces objectifs et est présenté sur la figure 1. L'estimateur OGEC repose sur les étapes suivantes.

2.1.1 Estimation de l'erreur de grille par projection oblique

Supposons qu'un estimateur à contrainte de parcimonie, nous fournisse une estimation du support $\hat{\mathcal{S}}$ et une estimation du vecteur $\hat{\mathbf{x}}$, c'est-à-dire

$$\mathcal{A}(\mathbf{y}) \mapsto \{\hat{\mathbf{x}}, \hat{\mathcal{S}}\}. \quad (1)$$

Le dictionnaire \mathbf{H} et sa "dérivée" $\dot{\mathbf{H}}$ peuvent être exploités grâce à un projecteur oblique [5] des mesures

$$\mathbf{E}_{\dot{\mathbf{H}}_S \mathbf{H}_S} \mathbf{y} = \dot{\mathbf{H}}_S \text{diag}(\hat{\mathbf{x}}) \mathbf{e} + \mathbf{E}_{\dot{\mathbf{H}}_S \mathbf{H}_S} \mathbf{w} \text{ avec}$$

$$\mathbf{E}_{\dot{\mathbf{H}}_S \mathbf{H}_S} = \dot{\mathbf{H}}_S \left(\dot{\mathbf{H}}_S^T \mathbf{P}_{\mathbf{H}_S}^\perp \dot{\mathbf{H}}_S \right)^{-1} \dot{\mathbf{H}}_S^T \mathbf{P}_{\mathbf{H}_S}^\perp \text{ et } \mathbf{P}_{\mathbf{H}_S}^\perp = \mathbf{I} - \mathbf{H}_S \mathbf{H}_S^\dagger$$

On cherche l'erreur permettant de minimiser le bruit projeté, supposant $\mathbf{x} = \hat{\mathbf{x}}$ suivant le critère

$$\hat{\mathbf{e}} = \arg \min_{\mathbf{e} \in \mathcal{A}(\mathbf{y})} \|\mathbf{E}_{\dot{\mathbf{H}}_S \mathbf{H}_S} \mathbf{w}\|^2.$$

La solution de ce problème de minimisation de la norme du bruit projeté est

$$\hat{\mathbf{e}} = \text{diag}(\hat{\mathbf{x}})^{-1} \dot{\mathbf{H}}_S^\dagger \mathbf{E}_{\dot{\mathbf{H}}_S \mathbf{H}_S} \mathbf{y}.$$

2.1.2 Correction du dictionnaire et par ré-estimation

Une fois l'erreur de grille estimée, le dictionnaire corrigé est alors construit selon

$$\tilde{\mathbf{H}}_S = \mathbf{H}_S + \dot{\mathbf{H}}_S \text{diag}(\hat{\mathbf{e}}) \quad (2)$$

Enfin, le vecteur d'intérêt corrigé selon l'erreur de grille, noté $\hat{\hat{\mathbf{x}}}$, est estimé selon la résolution du critère moindres carrés suivant :

$$\hat{\hat{\mathbf{x}}} = \min_{\mathbf{x}} \|\mathbf{y} - \tilde{\mathbf{H}}_S \mathbf{x}\|^2.$$

Dans le cadre l'acquisition comprimée, il est standard de supposer un à priori Bayésien $p(\mathbf{x})$. L'estimateur optimal est celui de l'EQMM (Erreur Quadratique Moyenne Minimale) faisant intervenir la moyenne à posteriori $\hat{\hat{\mathbf{x}}} = \mathbb{E}[\mathbf{x}|\mathbf{y}]$. Cependant l'implémentation de celui-ci est coûteuse et se trouve en contradiction avec le troisième objectif énuméré au début de cette section. Nous préférons alors exploiter l'estimateur LE-QMM (3) (Linéaire EQMM) [6] définit selon

$$\hat{\hat{\mathbf{x}}} = \tilde{\mathbf{H}}_S^T \left(\tilde{\mathbf{H}}_S \tilde{\mathbf{H}}_S^T + \sigma^2 \mathbf{R}_x^{-1} \right)^{-1} \mathbf{y} \quad (3)$$

où le dictionnaire corrigé est donné par l'expression eq. (2). Contrairement à l'EQMM, le LEQMM fait appel uniquement aux statistiques d'ordre deux, \mathbf{R}_x , du vecteur d'intérêt \mathbf{x} .

2.2 Performances théoriques

L'algorithme OGEC se concentrant principalement sur l'estimation de l'erreur de grille, nous nous intéressons ici aux biais conditionnels et à l'EQM théoriques pour l'erreur de grille \mathbf{e} sachant eq. (1). De plus, selon le premier objectif énoncé au début de cette section, l'erreur de grille limite de manière drastique les performances d'estimation des algorithmes standards dans le régime des hauts RSBs. Dans ce contexte, il est raisonnable de supposer pour notre étude $\hat{\mathcal{S}} = \mathcal{S}$. Cette propriété a été mise en évidence dans [1] et illustrée dans la partie simulation de ce travail.

Le vecteur de biais conditionnellement à $\{\hat{\mathbf{x}}, \mathcal{S}\}$ est

$$\begin{aligned} \mathbf{b}(\hat{\mathbf{e}}|\hat{\mathbf{x}}) &= \mathbb{E}_{\mathbf{y}|\mathbf{x},\hat{\mathbf{x}},\mathbf{e}} [\hat{\mathbf{e}}(\mathbf{y}, \hat{\mathbf{x}}) - \mathbf{e}] \\ &= (\text{diag}(\hat{\mathbf{x}})^{-1} \text{diag}(\mathbf{x}) - \mathbf{I}) \mathbf{e} \end{aligned} \quad (4)$$

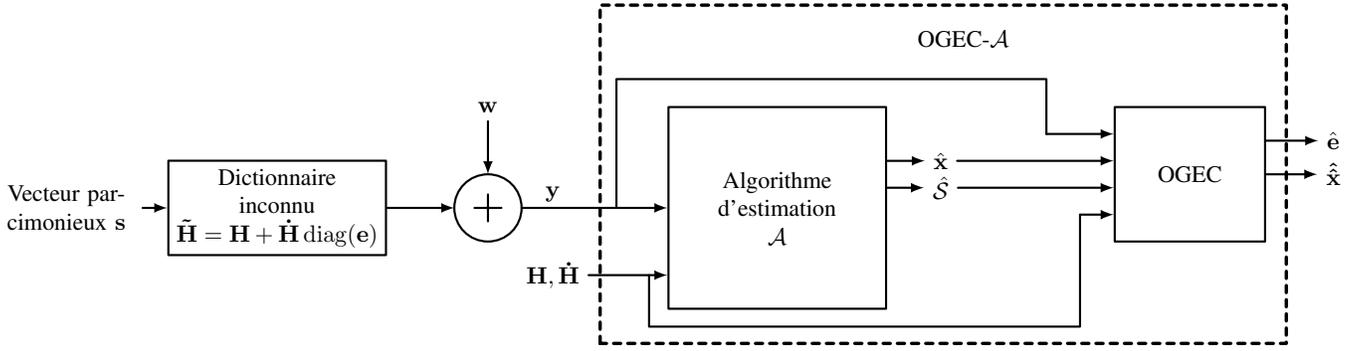


FIGURE 1 – Description de l’algorithme OGEC-A.

pour une erreur de grille i.i.d. centrée et de matrice de covariance $\sigma_e^2 \mathbf{I}$. On pourra se référer à [1] pour le détail des calculs. Notons que si l’estimation de $\hat{\mathbf{x}}$ fournie par l’algorithme $\mathcal{A}(\mathbf{y})$ est parfaite, l’estimateur OGEC est conditionnellement non biaisé. L’EQM Bayésienne est donnée selon

$$\text{BEQM}_{\text{OGEC}}(\hat{\mathbf{e}}|\hat{\mathbf{x}}) = \sigma_e^2 \sum_{l=1}^L \frac{\text{EQM}_{\mathcal{A}}(\hat{x}_l)}{\hat{x}_l^2} + \sigma^2 \text{Tr} \left[\text{diag}(\hat{\mathbf{x}})^{-2} \left(\hat{\mathbf{H}}_S^T \mathbf{P}_{\hat{\mathbf{H}}_S}^\perp \hat{\mathbf{H}}_S \right)^{-1} \right] \quad (5)$$

où $\text{EQM}_{\mathcal{A}}(\hat{x}_l) = \mathbb{E} [(x_l - \hat{x}_l)^2]$ est l’EQM de l’estimateur \mathcal{A} associée au vecteur $\hat{\mathbf{x}}$ entachée de l’erreur de grille.

Soit $\mathbf{P}_{\hat{\mathbf{H}}_S}^\perp = \mathbf{U}\mathbf{U}^T$ où $\mathbf{U}^T \mathbf{U} = \mathbf{I}$. Le vecteur de mesures projeté admet pour distribution $\mathbf{U}^T \mathbf{y} | \mathbf{e}, \mathbf{x} \sim \mathcal{N}(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ où $\boldsymbol{\mu} = \mathbf{U}^T \hat{\mathbf{H}}_S \text{diag}(\mathbf{x}) \mathbf{e}$ et $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$. Considérons la borne de Cramér-Rao moyennée pour un estimateur à biais donné en eq. (4) opérant sur $\mathbf{U}^T \mathbf{y} | \mathbf{e}, \mathbf{x}$. Celle-ci admet pour expression [8] :

$$\mathcal{C} = \text{Tr} \left[\left(\mathbf{I} + \frac{\partial \mathbf{b}(\hat{\mathbf{e}}|\hat{\mathbf{x}})}{\partial \mathbf{e}} \right) \mathbf{F}(\mathbf{e} | \mathbf{U}^T \mathbf{y})^{-1} \left(\mathbf{I} + \frac{\partial \mathbf{b}(\hat{\mathbf{e}}|\hat{\mathbf{x}})}{\partial \mathbf{e}} \right)^T \right] + \mathbb{E}_{\mathbf{e}} (\|\mathbf{b}(\hat{\mathbf{e}}|\hat{\mathbf{x}})\|^2) \quad (6)$$

où $\mathbf{F}(\mathbf{e} | \mathbf{U}^T \mathbf{y}) = \left(\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{e}} \right)^T \boldsymbol{\Sigma}^{-1} \left(\frac{\partial \boldsymbol{\mu}}{\partial \mathbf{e}} \right)$ est la matrice d’information de Fisher pour le vecteur d’intérêt \mathbf{e} . A la lumière des eq. (5) et eq. (6) (voir [1] pour les calculs), on peut conclure que $\text{BEQM}_{\text{OGEC}} = \mathcal{C}$ et donc que l’estimateur OGEC est statistiquement efficace selon la borne \mathcal{C} .

3 Echantillonnage comprimé de signaux à bande non-limitée

3.1 Description du modèle

Afin d’illustrer les performances de l’algorithme OGEC, nous considérons le problème d’importance [9, 10] de l’échantillonnage comprimé d’un signal à bande non limitée défini selon $x(t) = \sum_{l \in \mathcal{S}} x_l \delta(t - \theta_l)$ où θ_l et x_l sont les délais et amplitudes inconnus de la l -ème impulsion. Les échantillons sont

obtenus suite à un pré-filtrage par un noyau $\phi(t)$ et une compression selon la matrice $\boldsymbol{\Psi}$:

$$\mathbf{c} = \mathbf{H}_S \mathbf{x} + \mathbf{w} \quad \text{avec} \quad \mathbf{H}_S = \boldsymbol{\Psi} \boldsymbol{\Phi}_S$$

où $[\boldsymbol{\Phi}]_{kk'} = \phi\left(\frac{\theta_{k'} - (k-1)}{T_S}\right)$ est une $K \times K$ matrice orthogonale où T_S est la période d’échantillonnage. Le noyau $\phi(t)$ de type sincardinal, permet d’assurer la contrainte d’orthogonalité de la matrice $\boldsymbol{\Phi}$.

3.2 Illustrations numériques

Pour obtenir les résultats de l’algorithme OGEC, nous utilisons trois algorithmes différents d’estimation parcimonieuse : l’OMP [11], le CoSAMP [12] deux algorithmes gloutons et le SPGL1 [13] un algorithme d’optimisation résolvant le problème du basis pursuit [14]. Nous considérons le cas où le nombre d’inconnues L est fixé. L’algorithme SPGL1 n’exploitant par la connaissance a priori de l’ordre du modèle, L , nous gardons les L plus grandes amplitudes estimées par l’algorithme. Nous fixons $N = 100$, $L = 10$ et $K = 300$ et faisons $1e5$ réalisations de Monte-Carlo pour estimer les espérances mathématiques. La contrainte principale dans le choix des paramètres est de garantir l’estimation correcte du support à haut RSB. Nous prenons pour cela une variance de l’erreur $\sigma_e^2 = 10^{-5}$. Les courbes sont tracées en fonction du rapport interférence sur bruit, que nous définissons par : $\text{ISB} = \sigma_e^2 / \sigma^2$.

La figure 2 présente le résultat de l’algorithme pour l’estimation des amplitudes. Le résultat des différents algorithmes sans correction OGEC est indiqué par une croix, tandis que la BEQM sur l’amplitude après correction LEQMM est indiquée par des ronds. On observe que quand le bruit diminue, les estimateurs atteignent assez rapidement un palier sur l’amplitude (pour un ISB de -10dB) tandis que la correction effectuée permet d’améliorer nettement les performances.

Nous étudions à présent le problème de l’estimation de l’erreur de grille, dont les résultats sont visualisés sur la figure 3. Sont tracés à la fois le résultat de l’estimation sans aucune correction d’erreur marqué par une croix, l’erreur moyenne sur l’estimation des retards une fois la correction OGEC effectuée marquée par un rond et enfin la borne BEQM (5) pour chaque

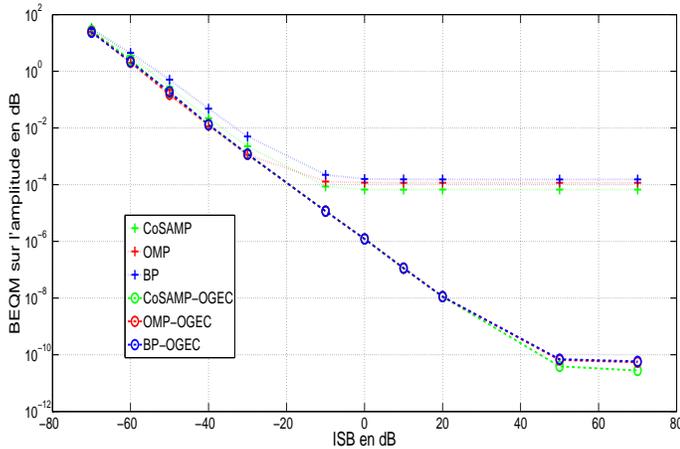


FIGURE 2 – EQM pour les amplitudes Vs. l'ISB

algorithme en pointillé et marquée par une étoile. Pour obtenir son expression, on moyenne l'erreur sur 100 réalisations de \mathcal{A} . Tout comme pour les amplitudes, l'algorithme OGEC permet d'améliorer les performances d'estimation sur les retards. Sans correction, l'erreur obtenue correspond à l'erreur de discrétisation de la grille à haut RSB qui produit un palier constant quel que soit l'algorithme dès un ISB de -10dB [4]. Quand le bruit est assez faible pour que l'erreur atteigne le palier de discrétisation, l'utilisation de l'algorithme OGEC améliore les performances et permet d'estimer l'erreur de grille. Les résultats pratiques de l'algorithme suivent la BEQM théorique, dans la région de bruit d'intérêt. Pour des bruits plus importants, le calcul de la BEQM est trop optimiste puisqu'il part du principe que les supports sont correctement estimés ce qui n'est pas le cas dans cette région de bruit.

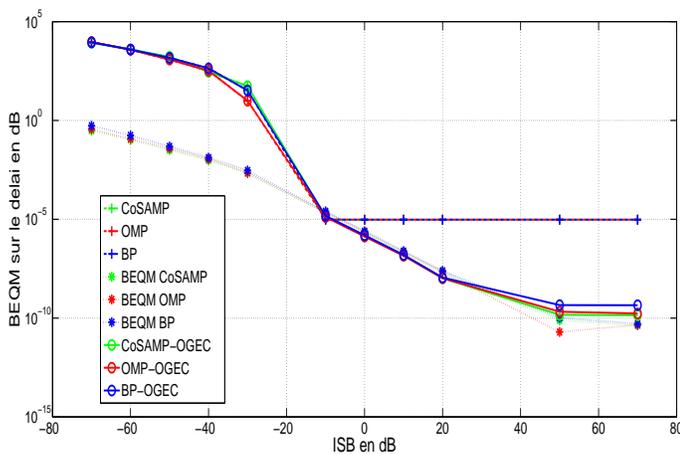


FIGURE 3 – EQM pour les retards Vs. l'ISB

4 Conclusion

Dans ce travail, nous proposons un nouvel algorithme permettant de lutter contre le phénomène d'erreur de grille inhé-

rent à tout estimateur basé sur la contribution d'un dictionnaire redondant. Les erreurs de grille pénalisent gravement les performances des estimateurs à contrainte de parcimonie dans le régime des hauts RSBs. Traditionnellement les approches proposées sont développées de manière *ad-hoc* à une classe d'estimateur donnée. Dans ce travail, nous prenons le parti de proposer une méthodologie générique de post-traitement venant s'ajouter telle une sur-couche permettant de "robustifier" tout type d'estimateur par dictionnaire pour un cout calculatoire réduit. Après avoir présenté l'estimateur OGEC, nous analysons son biais et son EQM de manière théorique. Enfin, les performances de l'OGEC sont illustrées dans le contexte de l'échantillonnage comprimé de signaux à bande non-limitée.

Références

- [1] S. Bernhardt, *Performances et méthodes pour l'échantillonnage comprimé : Robustesse à la méconnaissance du dictionnaire et optimisation du noyau d'échantillonnage*, Thèse, Université Paris-Saclay, Dec. 2016.
- [2] Z. Tan, P. Yang, and A. Nehorai, "Joint sparse recovery method for compressed sensing with structured dictionary mismatches," *Signal Processing, IEEE Transactions on*, vol. 62, no. 19, pp. 4997–5008, 10 2014.
- [3] Chaitanya Ekanadham, Daniel Tranchina, and Eero P Simoncelli, "Recovery of sparse translation-invariant signals with continuous basis pursuit," *IEEE transactions on signal processing*, vol. 59, no. 10, pp. 4735–4744, 2011.
- [4] S. Bernhardt, R. Boyer, B. Zhang, S. Marcos, and P. Larzabal, "Performance analysis for sparse based biased estimator : Application to line spectra analysis," in *2014 IEEE 8th Sensor Array and Multichannel Signal Processing Workshop (SAM)*, 6 2014, pp. 365–368, Invited paper.
- [5] R. T Behrens and L. L Scharf, "Signal processing applications of oblique projection operators," *Signal Processing, IEEE Transactions on*, vol. 42, no. 6, pp. 1413–1424, 1994.
- [6] S. M. Kay, *Fundamentals of Statistical Signal Processing : Detection Theory*, vol. II, Prentice-Hall, Upper Saddle River, New Jersey, 07458, 1998.
- [7] R. Boyer, R. Couillet, B-H Fleury, and P. Larzabal, "Large-system estimation performance in noisy compressed sensing with random support of known cardinality ?a bayesian analysis," *IEEE Transactions on Signal Processing*, vol. 64, no. 21, pp. 5525–5535.
- [8] R. Boyer, "Oblique projection for source estimation in a competitive environment : algorithm and statistical analysis," *Signal Processing*, vol. 89, no. 12, pp. 2547–2554, 2009.
- [9] M. Vetterli, P. Marziliano, and T. Blu, "Sampling signals with finite rate of innovation," *Signal Processing, IEEE Transactions on*, vol. 50, no. 6, pp. 1417–1428, 2002.
- [10] S. Bernhardt, R. Boyer, S. Marcos, Y. Eldar, and P. Larzabal, "Nouveaux noyaux pour l'échantillonnage de signaux impulsifs à bande non-limitée," in *25ème édition du colloque GRETSI*, 2015.
- [11] Y. C. Pati, R. Rezaifar, Y. C. Pati, and P. S. Krishnaprasad, "Orthogonal matching pursuit : Recursive function approximation with applications to wavelet decomposition," pp. 40–44, 1993.
- [12] D. Needell and R. Vershynin, "Signal recovery from incomplete and inaccurate measurements via regularized orthogonal matching pursuit," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 310–316, 2010.
- [13] E. van den Berg and M. P. Friedlander, "Probing the pareto frontier for basis pursuit solutions," *SIAM Journal on Scientific Computing*, vol. 31, no. 2, pp. 890–912, 2008.
- [14] S. Chen, D. Donoho, and M. Saunders, "Atomic decomposition by basis pursuit," *SIAM Journal on Scientific Computing*, vol. 20, no. 1, pp. 33–61, 1998.