

Quantification orientée objectif et application au contrôle de puissance

Chao ZHANG¹, Sara BERRI¹, Samson LASAULCE¹, Vineeth VARMA², Sophie TARBOURIECH³

¹Laboratoire des Signaux et Systèmes
3 Rue Joliot-curie, 91190 Gif-sur-Yvette, France

²Centre de Recherche en Automatique de Nancy
Boulevard des Aiguillettes - BP 70239, 54506 Vandœuvre-lès-Nancy, France

³Laboratoire d'Analyse et d'Architecture des Systèmes
7 Avenue du Colonel Roche, 31400 Toulouse, France

Résumé – Dans de nombreux problèmes d'allocation de ressources, où seule une version quantifiée des paramètres pertinents est disponible, des stratégies d'allocation optimales doivent être déterminées, comme par exemple, l'allocation de puissance dans les communications sans fil. Dans ce papier nous trouvons trois principales contributions. Premièrement, nous proposons un nouveau modèle de quantification, basé essentiellement sur le problème classique, que nous avons jugé nécessaire de rappeler également. En effet, la différence majeure réside dans l'objectif, dans le modèle proposé, au lieu de minimiser la distorsion, nous proposons de minimiser l'écart entre le maximum d'une fonction d'utilité générale, atteint lorsque tous les paramètres de la fonction sont connus, et ce qui a été effectivement atteint lorsque seule la version quantifiée de ces paramètres est disponible. Pour déterminer un tel quantificateur, le célèbre algorithme Lloyd-Max est généralisé. De plus, nous exposons comment est-ce que le système proposé peut être appliqué au problème du contrôle de puissance dans les communications sans fil; les résultats numériques montrent clairement son efficacité.

Abstract – In many resource allocation problems, optimal allocation strategies must be determined when only a quantized version of the relevant parameters are available, for instance, power allocation in wireless communications. The contribution of this work is threefold. First, the quantization problem is revisited and a framework which encompasses the classical problem of quantization is proposed. Instead of minimizing the distortion, the goal is to minimize the gap between the maximum of a general payoff function (which would be reached by knowing all parameters of the function) and what is effectively reached when only the quantized version of the parameters is available. Then, to determine such a quantizer, the well-known Lloyd-Max algorithm is generalized. At last, we show how this framework can be applied to the problem of power control in wireless communications; the obtained numerical results clearly show the potential of such a framework.

1 Introduction

La formulation du problème considéré dans ce papier est principalement motivée par les problèmes d'allocation de ressources dans les communications sans fil. Dans ce type de problèmes, il est assez souvent nécessaire de faire face à une situation où il existe une fonction d'utilité $f(x; g)$ à maximiser par rapport à x , dont les paramètres représentés par le vecteur g ne sont pas parfaitement connus. Dans ce papier, nous nous intéressons au cas de présence d'une version quantifiée de ces paramètres. Du point de vue technique, le scénario évoqué dans ce papier n'a pas été abordé dans la littérature, cependant il reste tout de même très employé par des travaux liés aux communications sans fil (consulter, par exemple, [1]). En guise d'exemple, il serait pertinent qu'un émetteur exploite une version quantifiée du canal lorsque il est sur le point d'allouer une puissance, une information lui sera communiquée par le récepteur en utilisant un mécanisme de rétroaction (feedback); dans cet exemple, le récepteur doit quantifier les gains des canaux ou une matrice dans le but de satisfaire certaines contraintes,

comme par exemple, liées à la capacité du canal de rétroaction.

Le problème que nous introduisons dans ce papier consiste en la conception d'un quantificateur produisant un vecteur altéré de paramètres, qui est effectivement disponible et servira à maximiser la fonction f . Plus précisément, l'objectif est de minimiser l'impact qu'exerce le bruit de quantification sur la perte d'optimalité, engendrée par l'utilisation de la version quantifiée de g , lors de la maximisation de f .

Afin de déterminer le meilleur quantificateur correspondant à celui qui minimise la perte d'optimalité engendrée par l'utilisation d'une version altérée \hat{g} au lieu de g , nous proposons de généraliser l'Algorithme Lloyd-Max (ALM) [2]. En effet, le quantificateur basé sur le ALM vise à minimiser la distorsion, c'est-à-dire l'erreur quadratique moyenne entre la source et sa version reconstruite. Cependant, il faut noter que c'est une conception indépendante de l'utilisation de la quantité quantifiée, et il s'avère qu'elle pourrait être améliorée si elle est mesurée en termes d'utilité finale. Il existe des travaux où un critère de performance différent est considéré, comme le cas de [3], où la norme L_p – est utilisée (au lieu de la norme Euclidienne).

Plus vaguement, dans [4], le problème considéré est de minimiser une fonction arbitraire représentant la différence entre le vecteur réel de paramètres et sa version quantifiée mais, encore une fois, ce problème ne correspond pas au cadre de quantification que nous proposons (qui sera référé par QO2 pour quantification/quantificateur orienté objectif). A la connaissance des auteurs, le problème de quantification n'a pas été déjà formulé dans le même sens que dans ce présent papier.

2 Formulation du problème

Nous considérons une fonction $f(x; g) : \mathcal{X} \times \mathcal{G} \rightarrow \mathbb{R}$, appelée fonction d'utilité. Généralement, $x \in \mathcal{X}$ et $g \in \mathcal{G}$ peuvent être des vecteurs avec $\mathcal{X} \subset \mathbb{R}^N$ et $\mathcal{G} \subset \mathbb{R}^K$. Le but est de maximiser la fonction f par rapport à x , et ce en ne disposant que de la version quantifiée de g , notée $\mathcal{Q}(g)$. L'objectif est de déterminer un bon quantificateur QO2, à savoir un quantificateur qui permet de minimiser l'impact de la quantification sur la perte d'optimalité induite par l'utilisation de $\mathcal{Q}(g)$ dans le problème d'optimisation final.

Un quantificateur \mathcal{Q} est donné par une partition de \mathcal{G} en cellules et leur représentant. Notons par M le nombre maximum de cellules. Les cellules sont notées par $\{\mathcal{C}_1, \dots, \mathcal{C}_M\}$ et satisfont $\mathcal{C}_m \subset \mathcal{G}$ tel que $\mathcal{C}_1 \cup \mathcal{C}_2 \cup \dots \cup \mathcal{C}_M = \mathcal{G}$ et $\mathcal{C}_m \cap \mathcal{C}_n = \emptyset$, pour chaque $m \neq n$. La règle de quantification que nous supposons est comme suit : $\mathcal{Q}(g) = r_m$ si $g \in \mathcal{C}_m$, où $m = 1, 2, \dots, M$. L'approche classique consiste à déterminer \mathcal{Q} telle que la distorsion soit minimisée, c'est-à-dire, minimiser

$$D = \int \phi(g) \|\mathcal{Q}(g) - g\|^2 dg, \quad (1)$$

où : ϕ est la fonction de densité de probabilité (f.d.p) de g . L'avantage d'une telle approche est qu'elle permet d'obtenir explicitement le quantificateur (à savoir les représentants et les cellules) et cela conduit à un système indépendant de l'utilité. Cependant, si l'utilité est connue, il est généralement possible d'améliorer davantage les performances lorsqu'elles sont mesurées en termes d'utilité finale. En effet, si on note par F le maximum actuel de f

$$F(g) = \max_x f(x; g), \quad (2)$$

et par \hat{F}

$$\hat{F}(\mathcal{Q}(g)) = \max_x f(x; \mathcal{Q}(g)), \quad (3)$$

le niveau effectivement atteint en ne connaissant que $\mathcal{Q}(g)$, et pas g , est déterminé lorsque F et \hat{F} coïncident. Par conséquent, il serait pertinent de déterminer \mathcal{Q} à travers la relation suivante :

$$\mathcal{Q}^* \in \arg \min_{\mathcal{Q}} \mathbb{E}_g \|F(g) - F(\mathcal{Q}(g))\|^2. \quad (4)$$

La section suivante servira à exhiber des résultats permettant de minimiser la quantité définie par (4). Dans la plupart des cas, comme le cas de ALM, l'optimalité globale n'est pas souvent garantie, nous proposons alors de traiter cette question classique dans nos travaux futurs, et de se focaliser dans ce papier sur ce qui est réellement nouveau.

3 Système général de quantification

En présence d'un vecteur g de dimension K ($K > 1$), définir de manière conjointe les cellules et les représentants qui minimisent (4), constitue un problème non-trivial. Ce qui nous a guidé à nous recourir à des systèmes de quantification classiques tel que ALM, afin de déterminer un système qui puisse caractériser les représentants les plus appropriés $\{r_1, \dots, r_M\}$ (localement optimaux) pour un ensemble donné de cellules $\{\mathcal{C}_1, \dots, \mathcal{C}_M\}$, et des cellules localement optimales pour un ensemble donné de représentants. Après que ces deux ensembles soient bien définis, nous serons en mesure de résoudre de manière itérative pour $\{r_1, \dots, r_M\}$ et $\{\mathcal{C}_1, \dots, \mathcal{C}_M\}$, ce qui donnera une solution localement optimale (pas nécessairement globalement optimale).

La détermination des cellules optimales pour un ensemble donné de représentants $\{r_1, \dots, r_M\}$, peut être effectuée de la même manière que l'algorithme classique ALM, et en construisant des partitions similaires à celles de Voronoi :

$$\mathcal{C}_m = \left\{ g \in \mathbb{R}^K : [F(g) - F(r_m)]^2 \leq [F(g) - F(r_n)]^2 \right\}, \quad (5)$$

où : $n \neq m$. Ce choix est motivé par le fait d'être en mesure de chercher un ensemble dont la distorsion est minimale en termes de $F(g)$, au lieu de la distance Euclidienne ou la distorsion de g , comme supposé par les partitions classiques de Voronoi. D'autre part, déterminer un représentant optimal r_m pour une cellule donnée \mathcal{C}_m , pourrait s'avérer ne pas être aussi simple. En effet, lorsque g est un scalaire, F est inversible, et \mathcal{C}_m est définie par un intervalle de type $[t_m, t_{m+1}]$, le meilleur représentant r_m^* est donné par (6)

$$F(r_m^*) = \frac{\int_{t_m}^{t_{m+1}} \phi(g) F(g) dg}{\int_{t_m}^{t_{m+1}} \phi(g) dg}. \quad (6)$$

Les résultats énoncés ci-dessus peuvent être prouvés en appliquant la différentielle à $\mathbb{E}_g (\|F(\hat{g}) - F(g)\|^2)$ par rapport à r_m pour une cellule donnée $\mathcal{C}_m = [t_m, t_{m+1}]$. Comme F est inversible, sa dérivée ne s'annule jamais, et donc (6) pourrait être atteint. Cependant, ceci ne reste pas valable quand g devient un vecteur, en dépit de l'inversibilité de F , car les K éléments de r_m ne peuvent être déterminés en n'utilisant qu'une seule équation (une équation pour K inconnus). Ainsi, lorsque g est un vecteur sous forme $g = (g_1, g_2, \dots, g_K)^T$, certaines hypothèses doivent être ajoutées afin de rendre possible la procédure de détermination.

Hypothèse 3.1 (Hypothèse de décomposition). *La fonction F peut s'écrire comme $F = \sum_{k=1}^K u_k$, avec $u_k : \mathcal{G} \rightarrow \mathbb{R}$ et elle est telle que la fonction vecteur $V(g) := (u_1(g), \dots, u_K(g))^T$ est inversible en \mathcal{C}_m .*

Cette hypothèse est bien adaptée pour maintes applications, où l'utilité totale est définie comme somme de plusieurs composantes, comme par exemple, la somme de plusieurs taux sur

chaque bande. En effet, la fonction V peut ne pas être globalement inversible, cependant, comme nous allons l'illustrer dans la Sec. 4, nos résultats peuvent être appliqués à des partitions \mathcal{C}_m où la condition de décomposition est vérifiée. Pour les systèmes sans fil, cette condition suffisante est souvent satisfaite en raison de la monotonie de la plupart des fonctions d'utilité ; en effet, de nombreuses fonctions d'utilité dans ce type de système sont typiquement croissantes de manière monotone par rapport au Rapport Signal sur Interférence plus Bruit (RSIB).

Proposition 3.2 (Représentants optimaux). *Si l'hypothèse 3.1 est satisfaite pour une partition \mathcal{C}_m , avec une fonction décomposée inversible $V(g) = (u_1(g), \dots, u_K(g))^T$, alors le représentant optimal r_m^* qui minimise $\mathbb{E}_g(\|V(\hat{g}) - V(g)\|^2)$ minimise également $\mathbb{E}_g(\|F(\hat{g}) - F(g)\|^2)$, où $\hat{g} = \mathcal{Q}(g)$, et peut être obtenu en résolvant le système de K équations suivant*

$$u_k(r_m^*) = \frac{\int_{\mathcal{C}_m} \phi(g) u_k(g) dg}{\int_{\mathcal{C}_m} \phi(g) dg}. \quad (7)$$

Le cas scalaire (4) peut être obtenu en fixant K à 1. A présent, nous allons exploiter la Proposition 3.2 afin de déterminer un algorithme convenable qui permettra la caractérisation d'un QO2 en connaissant F . Le nouveau quantificateur peut

Entrées : $\phi(g) : \mathbb{R}^K \rightarrow \mathbb{R}_{\geq 0}$, $F = \sum_k u_k$ satisfait

l'hypothèse 3.1, $\{r_1^{(0)}, \dots, r_M^{(0)}\}$

Sorties : $\{r_1^*, \dots, r_M^*\}$, $\{C_1^*, \dots, C_M^*\}$

Initialisation : Fixer l'indice d'itération $q = 0$. Initialiser

les représentants du quantificateur $\{r_1^{(0)}, \dots, r_M^{(0)}\}$. Fixer

$r_m^{(-1)} = 0$ pour chaque $m \in \{1, \dots, M\}$.

while $\sum_{m=1}^M (r_m^{(q)} - r_m^{(q-1)})^2 > \delta$ et $q < Q$ **do**

 Mettre à jour l'indice d'itération : $q \leftarrow q + 1$.

 Pour chaque $m \in \{1, 2, \dots, M\}$, mettre à jour C_m^q à partir de r_m^{q-1} en utilisant (5).

 Pour chaque $m \in \{1, 2, \dots, M\}$, mettre à jour

r_m^q pour chaque partition C_m^q en utilisant (7).

end

$\forall m \in \{1, \dots, M\}$, $r_m^* = r_m^{(q)}$, $t_m^* = t_m^{(q)}$, $t_{M+1}^* = +\infty$

Algorithm 1: Algorithme déterminant un quantificateur via l'utilité

être résumé par l'Algorithme. 1. Il faut noter que l'hypothèse 3.1 peut ne pas être satisfaite pour tout l'ensemble \mathcal{G} , dans ce cas, \mathcal{G} peut être partitionné en plusieurs sous-régions où chaque région peut être quantifiée de manière spécifique. Par exemple, comme on pourra le voir dans la section suivante, lorsque l'objectif est de maximiser l'énergie, les solutions sont telles que seul le meilleur canal est choisi. Ainsi, seul ce canal doit être quantifié, et un nombre minimum de bits peut être attribué aux autres canaux.

4 Application à des fonctions d'utilité typiques au systèmes sans fil

Dans cette section, nous considérons une fonction d'utilité particulière pour le scénario multi-bandes, qui est la fonction d'efficacité énergétique. Les quantificateurs qu'utilise un émetteur pour l'allocation de sa puissance sont fournis par les gains des canaux. Notons par $g_k = |h_k|^2$ le gain du canal au niveau de la bande k , où h_k peut être une réalisation d'une variable aléatoire gaussienne complexe, si on considère un canal avec évanouissement de Rayleigh. La puissance émise sur la bande k est notée par p_k , et elle est supposée limitée : $p_k \geq 0$ et $\sum_{k=1}^K p_k \leq P_{\max}$. Notons par $p = (p_1, \dots, p_K)^T$ et $g = (g_1, \dots, g_K)^T$, les vecteurs colonnes de dimension K formés par les niveaux des puissances et les gains des canaux, respectivement. Dans ce papier, notre choix se porte sur la fonction d'efficacité donnée dans [5], qui est définie comme suit :

$$f(p, g) = \frac{\sum_{k=1}^K e\left(-\frac{c\sigma^2}{g_k p_k}\right)}{\sum_{k=1}^K p_k}, \quad (8)$$

où : σ^2 est le bruit au niveau du récepteur, et $c = 2^r - 1$, r représente l'efficacité spectrale. Afin de déterminer F dans le cas de maximisation de l'efficacité énergétique, nous commençons tout d'abord par déterminer une politique de contrôle de puissance optimale, ce qui constitue l'objet de la prochaine proposition.

Proposition 4.1. *Le système d'allocation de puissance optimal qui permet de maximiser l'efficacité énergétique est défini comme suit :*

$$p_k^*(g_k) = \begin{cases} 0 & k \neq \arg \max_i g_i, \\ \min\left(\frac{c\sigma^2}{g_k}, P_{\max}\right) & k = \arg \max_i g_i. \end{cases} \quad (9)$$

Selon la Proposition 4.1, afin de maximiser l'efficacité énergétique, un émetteur ne va transmettre que sur le meilleur canal. Pour mieux estimer l'utilité au niveau de l'émetteur, nous pouvons utiliser la méthode du vecteur de QO2 proposée dans la section précédente. Cependant, ceci engendrerait une complexité de calcul importante si le nombre de bandes est très grand. Comme il n'existe qu'une seule bande active à chaque intervalle de temps, ceci permet de concevoir un système de quantification spécial. En premier lieu, nous divisons la région totale en K sous-régions $\hat{\mathcal{C}}_1, \dots, \hat{\mathcal{C}}_K$, où $\hat{\mathcal{C}}_k = \{g \in \mathbb{R}^K : g_k = \max_i g_i\}$. La région $\hat{\mathcal{C}}_k$ correspond à celle dans laquelle g_k constitue le meilleur canal. Sans perte de généralité, nous considérons le système de quantification pour la région $\hat{\mathcal{C}}_k$. Si la réalisation du canal appartient à $\hat{\mathcal{C}}_k$, seule la bande k sera active. Ceci implique que seule la sélection de la bande k et

de g_k est utile pour l'amélioration de l'efficacité énergétique. Supposons $\mathcal{Q}(g) = (\hat{g}_1, \dots, \hat{g}_K)^T$, la première question peut être résolue aisément en fixant le plus grand nombre de représentants dans $\hat{\mathcal{C}}_k$ à \hat{g}_k , c'est-à-dire $\max_{g \in \hat{\mathcal{C}}_k} \mathcal{Q}(g) = \hat{g}_k$. Il faut noter que la fonction de l'efficacité énergétique dans $\hat{\mathcal{C}}_k$ peut être simplifiée comme suit :

$$F(g \in \hat{\mathcal{C}}_k) = \max_p f(p, g \in \hat{\mathcal{C}}_k) = \frac{e^{\left(-\frac{c\sigma^2}{g_k p_k^*}\right)}}{p_k^*(g_k)}. \quad (10)$$

Par conséquent, la deuxième question, liée à la détermination de la valeur quantifiée optimale de \hat{g}_k , peut être résolue en minimisant $d(\hat{g}_k)$, qui est définie comme suit :

$$d(\hat{g}_k) = \mathbb{E}_{g_k | g_k \geq g_1, \dots, g_k \geq g_K} [F(g_k) - F(\hat{g}_k)]^2. \quad (11)$$

En utilisant l'Algorithme. 1, le quantificateur scalaire avec les niveaux de transitions $\{t_{k,1}, \dots, t_{k,M+1}\}$ et leur représentant $\{l_{k,1}, \dots, l_{k,M}\}$, peut être calculé. Ce qui permettra par la suite de minimiser $d(\hat{g}_k)$ en connaissant g_k , et ainsi obtenir le meilleur canal. Sans perte de généralité, nous supposons que chaque région $\hat{\mathcal{C}}_k$ va être divisée en M différentes cellules de quantification $\{\mathcal{C}_{k,1}, \dots, \mathcal{C}_{k,M}\}$. Le représentant de la région de quantification $\mathcal{C}_{k,m}$ est défini comme un vecteur de dimension K , $r_{k,m} = (r_{k,m}^1, \dots, r_{k,m}^K)^T$. En se basant sur les résultats de la section précédente, le vecteur de la région de quantification $\mathcal{C}_{k,m}$ peut s'exprimer comme suit :

$$\mathcal{C}_{k,m} = \{g \in \mathbb{R}^K : t_{k,m} < g_k \leq t_{k,m+1}\} \cap \hat{\mathcal{C}}_k. \quad (12)$$

Le représentant correspondant $r_{k,m}$ peut être choisi comme :

$$r_{k,m}^k = l_{k,m} \quad (13)$$

$$r_{k,m}^{k'} = \text{const} < r_{k,m}^k (k' \neq k). \quad (14)$$

L'ensemble des régions de quantification $\mathcal{C} = \{\mathcal{C}_{1,1}, \dots, \mathcal{C}_{K,M}\}$ et celui des représentants $r = \{r_{1,1}, r_{1,2}, \dots, r_{K,M}\}$ du vecteur de quantification g , peuvent être déterminés en utilisant (12), (13) et (14). Avec cette nouvelle approche, la complexité de calcul est considérablement réduite.

5 Analyse numérique des performances

Dans cette section, nous allons présenter des résultats de simulation afin de mettre en avant les performances du QO2 proposé dans un scénario à un seul utilisateur et multi-bandes. Concernant la comparaison avec l'algorithme LM classique, nous examinons la perte d'optimalité engendrée par la quantification définie par la relation (15).

$$\Delta F(\%) = \mathbb{E}_g \left[\left| \frac{F(g) - \hat{F}(\mathcal{Q}(g))}{F(g)} \right| \right] \times 100 \quad (15)$$

nommée perte d'optimalité relative, où l'espérance est calculée par le biais de simulations de Monte-Carlo pour des valeurs de gain du canal g . Pour chaque k , le gain du canal g_k au niveau de la bande k est supposé suivre une distribution exponentielle,

à savoir, sa f.d.p s'écrit $\phi(g_k) = \exp(-g_k)$; ceci correspond à l'hypothèse standard de l'évanouissement de Rayleigh. Le scénario considéré pour toutes les simulations est : la puissance d'émission $P_{\max} = 20$ dBm et la puissance du bruit reçu $\sigma^2 = 0$ dBm, ce qui donne ainsi RSB= 20 dB.

Dans la Fig. 1, la fonction d'utilité f correspond à la fonction d'utilité liée à l'efficacité énergétique définie par (8). La perte d'optimalité relative est tracée en fonction du nombre de bits de quantification, et ce lorsque le nombre de bandes est $K = 16$. Le QO2 proposé est obtenu en utilisant (12), (13) et (14). Comme on pouvait s'y attendre, lorsque le nombre de bits augmente, la perte d'optimalité relative des deux quantificateurs diminue. Le QO2 proposé dans la Sec.4 surpasse le quantificateur vectoriel LM classique. Il est également important de noter que la perte d'optimalité relative du QO2 proposé est très proche de 0 lorsque nous avons plus de 7 bits de quantification, en revanche la perte d'optimalité relative du quantificateur LM demeure significative même avec 9 bits de quantification.

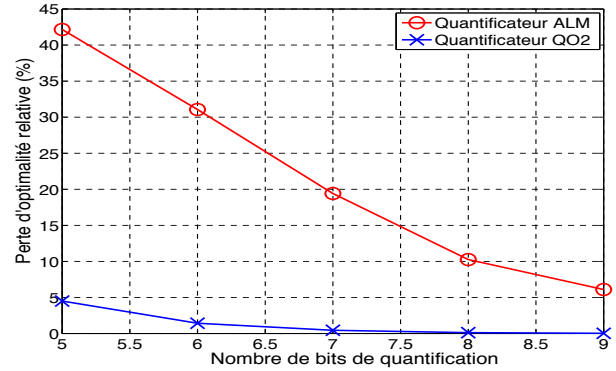


FIGURE 1 – Comparaison des performances en termes d'utilité finale entre le quantificateur classique (qui vise à minimiser la distorsion) et le QO2 proposé.

Références

- [1] G. Caire, N. Jindal, M. Kobayashi et N. Ravindran. *Multiuser MIMO achievable rates with downlink training and channel state feedback*. IEEE Transactions on Information Theory, 56(6), 2845-2866, 2010.
- [2] S. Lloyd. *Least squares quantization in PCM*. IEEE transactions on information theory, 28(2), 129-137, 1982.
- [3] L. Wang, N. Piovto et D. Schonfeld. *Boosting quantization for L p norm distortion measure*. 2012 IEEE Statistical Signal Processing Workshop (SSP), IEEE, 2012.
- [4] J. G. Proakis. *Digital Communications*. 4th ed. New York : McGrawHill, Inc., 2001.
- [5] E. V. Belmaga et S. Lasaulce. *Energy-efficient precoding for multiple-antenna terminals*. IEEE. Transactions on Signal Processing, 59(1), 329-340, 2011.