

Estimation d'un signal complexe à partir d'un modèle parcimonieux perturbé

Anisia FLORESCU¹, Emilie CHOUZENOUX², Jean-Christophe PESQUET² et Silviu CIOCHINA¹

¹Politehnica University of Bucharest, Telecommunications Dept., Romania

²Université Paris-Est, LIGM, UMR CNRS 8049, 77454 Marne-la-Vallée

anisia.florescu@ugal.ro, silviu@comm.pub.ro

emilie.chouzenoux@univ-mlv.fr, pesquet@univ-mlv.fr

Résumé – Dans ce travail nous abordons le problème de l'estimation, en présence de bruit, d'un signal à valeurs complexes admettant une représentation parcimonieuse dans une famille paramétrique de vecteurs dont les paramètres sont connus de manière imprécise. Nous formulons ce problème d'estimation comme celui de la minimisation d'une fonction de coût non convexe sous contraintes non convexes. Un algorithme de type explicite-implicite (*forward-backward*) permet d'apporter une solution numérique à ce problème d'optimisation. La mise en œuvre de cet algorithme nécessite le calcul d'un opérateur proximal dont nous donnons la forme exacte. Nous montrons que la méthode proposée peut être vue comme une généralisation d'une technique de seuillage dur itératif. En utilisant des résultats récents d'analyse non lisse, nous étudions la convergence de l'algorithme d'optimisation employé. Par ailleurs, nous illustrons le bon comportement numérique de l'approche proposée sur des exemples d'analyse spectrale parcimonieuse.

Abstract – In this paper, we address the problem of estimating a complex-valued signal corrupted with noise. This signal is assumed to have a sparse representation in an uncountable family of vectors the parameters of which are known in an imprecise manner. The problem is formulated as the minimization of a nonconvex function subject to nonconvex constraints. A forward-backward algorithm is employed to solve the corresponding numerical optimization problem. Implementing this algorithm requires the calculus of a proximal operator whose closed-form expression is provided. In addition, this paper shows that the proposed method can be considered as a generalization of an iterative hard thresholding algorithm. The algorithm convergence is also studied using recent nonsmooth analysis results. Simulation results illustrate the good practical performance of the proposed approach when applied to spectral estimation.

1 Introduction

Dans ce travail, nous nous intéressons au problème classique de l'estimation d'un signal $\bar{x} \in \mathbb{C}^Q$ à partir d'observations :

$$y = \bar{x} + w \quad (1)$$

où $w \in \mathbb{C}^Q$ est une réalisation d'un bruit aléatoire. On suppose que ce signal admet une représentation parcimonieuse dans une famille $\mathcal{E} = \{e_\nu \mid \nu \in \mathbb{R}\}$ de vecteurs de \mathbb{C}^Q paramétrés par une variable scalaire $\nu \in \mathbb{R}$. Plus précisément, le signal $\bar{x} \in \mathbb{C}^Q$ peut s'écrire sous la forme

$$\bar{x} = \sum_{n=1}^N \bar{c}_n e_{\bar{\nu}_n} \quad (2)$$

où $\{e_{\bar{\nu}_n} \mid 1 \leq n \leq N\}$ est un sous-ensemble fini de \mathcal{E} , un grand nombre de composantes du vecteur $\bar{c} = (\bar{c}_n)_{1 \leq n \leq N} \in \mathbb{C}^N$ étant supposées nulles [4].

L'originalité de ce travail est de se placer dans le cas où les paramètres $(\bar{\nu}_n)_{1 \leq n \leq N}$ sont connus de manière imprécise, *i.e.* ils sont tels que, pour tout $n \in \{1, \dots, N\}$,

$$\bar{\nu}_n = \theta_n + \bar{\delta}_n \quad (3)$$

où $\theta_n \in \mathbb{R}$ est une valeur fixée et $\bar{\delta}_n \in \mathbb{R}$ est une erreur inconnue qui doit être estimée. Sous les hypothèses que les perturbations $(\bar{\delta}_n)_{1 \leq n \leq N}$ sont petites et que la fonction $\nu \mapsto e_\nu$ est différentiable, un développement de Taylor au premier ordre conduit à

$$(\forall n \in \{1, \dots, N\}) \quad e_{\bar{\nu}_n} \simeq e_{\theta_n} + \bar{\delta}_n e'_{\theta_n} \quad (4)$$

où e'_{θ_n} est le gradient de $\nu \mapsto e_\nu$ en θ_n . Après cette approximation, le modèle (2) prend une forme bilinéaire :

$$\bar{x} = \sum_{n=1}^N (\bar{c}_n e_{\theta_n} + \bar{c}_n \bar{\delta}_n e'_{\theta_n}). \quad (5)$$

Une approche similaire a été proposée dans [7] pour décomposer un signal sous forme de versions traduites d'éléments d'un dictionnaire. Cependant, l'approche ℓ_1 considérée était limitée aux signaux à valeurs réelles. L'étude d'un modèle parcimonieux perturbé a été également effectuée dans [10] en optimisant un critère de type moindres carrés totaux (*total least squares*) pénalisé par un terme ℓ_1 . Notre approche diffère par les points suivants :

- elle est basée sur l'emploi d'une pénalisation ℓ_0 ,
- les erreurs $(\bar{\delta}_n)_{1 \leq n \leq N}$ sont supposées telles que

$$(\forall n \in \{1, \dots, N\}) \quad |\bar{\delta}_n| \leq \Delta_n, \quad (6)$$

les bornes $(\Delta_n)_{1 \leq n \leq N} \in [0, +\infty[^N$ étant fixées au préalable,

- un terme de fidélité quelconque peut être employé pourvu qu'il soit de gradient lipschitzien.

L'article est organisé comme suit : la section 2 présente la démarche d'optimisation proposée et décrit l'algorithme explicite-implicite employé. Ensuite, la section 3 illustre l'applicabilité de cet algorithme à travers un ensemble de résultats en estimation spectrale. Une conclusion est apportée dans la section 4.

2 Méthode proposée

2.1 Formulation variationnelle.

Afin d'identifier le modèle parcimonieux perturbé, nous proposons de résoudre le problème d'optimisation suivant :

$$\begin{aligned} \underset{\substack{c=(c_n)_{1 \leq n \leq N} \in \mathbb{C}^N \\ \delta=(\delta_n)_{1 \leq n \leq N} \in B}}{\text{minimiser}} \quad & \Phi \left(\sum_{n=1}^N (c_n e_{\theta_n} + c_n \delta_n e'_{\theta_n}) - y \right) \\ & + \lambda l_0(c) + \frac{\epsilon}{2} \|c\|^2 \end{aligned} \quad (7)$$

où B est le pavé de \mathbb{R}^N défini par $[-\Delta_1, \Delta_1] \times \dots \times [-\Delta_N, \Delta_N]$, $\Phi: \mathbb{C}^Q \rightarrow \mathbb{R}$ est le terme de fidélité aux données souvent lié à l'anti-log-vraisemblance du bruit entâchant les observations, $\lambda \in]0, +\infty[$ est un paramètre de régularisation servant à imposer la "sparsité" de la représentation et $\epsilon \in [0, +\infty[$. Le dernier terme joue un rôle similaire à celui d'une pénalisation élastique [11].

Dans la suite, on introduit les matrices $E = [e_{\theta_1} \dots e_{\theta_N}] \in \mathbb{C}^{Q \times N}$, $E' = [e'_{\theta_1} \dots e'_{\theta_N}] \in \mathbb{C}^{Q \times N}$, et l'on définit, pour tout $n \in \{1, \dots, N\}$ et $(c_n, d_n) \in \mathbb{C}^2$, la fonction

$$\psi_n(c_n, d_n) = \lambda \chi_{\{0\}}(c_n) + \iota_{S_n}(c_n, d_n) + \frac{\epsilon}{2} |c_n|^2, \quad (8)$$

où ι_{S_n} est la fonction indicatrice du cône fermé

$$S_n = \{(c_n, d_n) \in \mathbb{C}^2 \mid \exists \delta_n \in [-\Delta_n, \Delta_n], d_n = \delta_n c_n\}. \quad (9)$$

On peut alors reformuler le problème (7) comme celui de la minimisation de la fonction

$$(c, d) \mapsto \Phi \left([E \ E'] \begin{bmatrix} c \\ d \end{bmatrix} - y \right) + \sum_{n=1}^N \psi_n(c_n, d_n). \quad (10)$$

2.2 Algorithme proposé.

Si l'on fait l'hypothèse que la fonction Φ est différentiable, on peut apporter une solution numérique au problème (7) en employant l'algorithme 1.

Soit \mathcal{H} un espace de Hilbert de dimension finie. Nous rappelons que l'opérateur proximal d'une fonction $\varphi: \mathcal{H} \rightarrow]-\infty, +\infty]$ propre, semi-continue inférieurement telle que $\inf \varphi > -\infty$, est défini par

$$(\forall u \in \mathcal{H}) \quad \text{prox}_{\varphi}(u) \in \arg \min_{v \in \mathcal{H}} \frac{1}{2} \|u - v\|^2 + \varphi(v). \quad (11)$$

Algorithme 1 (algorithme explicite-implicite)

$$(c^{(0)}, d^{(0)}) \in (\mathbb{C}^N)^2$$

$$(\underline{\gamma}, \bar{\gamma}) \in]0, +\infty[^2$$

Pour $k = 0, 1, \dots$

$$\left[\begin{array}{l} \gamma^{(k)} \in]\underline{\gamma}, \bar{\gamma}[\\ D^{(k)} = \nabla \Phi \left([E \ E'] \begin{bmatrix} c^{(k)} \\ d^{(k)} \end{bmatrix} - y \right) \\ \tilde{c}_n^{(k)})_{1 \leq n \leq N} = c^{(k)} - \gamma^{(k)} E^H D^{(k)} \\ \tilde{d}_n^{(k)})_{1 \leq n \leq N} = d^{(k)} - \gamma^{(k)} (E')^H D^{(k)} \\ (c_n^{(k+1)}, d_n^{(k+1)})_{1 \leq n \leq N} = (\text{prox}_{\gamma^{(k)} \psi_n}(\tilde{c}_n^{(k)}, \tilde{d}_n^{(k)}))_{1 \leq n \leq N} \end{array} \right.$$

La mise en œuvre de l'algorithme 1 nécessite le calcul des opérateurs proximaux [6] des fonctions $(\psi_n)_{1 \leq n \leq N}$. Ces fonctions sont non convexes, mais on peut montrer que leurs opérateurs proximaux prennent la forme relativement simple suivante :

$$\begin{aligned} & (\forall n \in \{1, \dots, N\}) \quad (\forall (c_n, d_n) \in \mathbb{C}^2) \quad (\forall \gamma \in]0, +\infty[) \\ & \text{prox}_{\gamma \psi_n}(c_n, d_n) = \begin{cases} (0, 0) & \text{si } |c_n|^2 + |d_n|^2 < \frac{\widehat{\delta}_n c_n - d_n + 2\gamma \lambda (1 + \gamma \epsilon + \widehat{\delta}_n^2)}{1 + \widehat{\delta}_n^2} \\ \frac{c_n + \widehat{\delta}_n d_n}{1 + \gamma \epsilon + \widehat{\delta}_n^2} (1, \widehat{\delta}_n) & \text{sinon,} \end{cases} \end{aligned}$$

où, pour $\text{Re}(c_n d_n^*) \neq 0$,

$$\widehat{\delta}_n = \min \left\{ \frac{\eta_n + (1 + \gamma \epsilon) |d_n|^2 - |c_n|^2}{2 |\text{Re}(c_n d_n^*)|}, \Delta_n \right\} \text{sign}(\text{Re}(c_n d_n^*))$$

$$\text{et } \eta_n = \sqrt{((1 + \gamma \epsilon) |d_n|^2 - |c_n|^2)^2 + 4(1 + \gamma \epsilon) (\text{Re}(c_n d_n^*))^2}.$$

Il est intéressant de remarquer que, dans le cas particulier où Φ est un critère de moindres carrés, $\Delta_n \equiv 0$ et $\epsilon = 0$, l'algorithme se réduit à un seuillage dur itératif [2].

2.3 Théorème de convergence.

Le problème d'optimisation considéré étant non convexe, la convergence de l'algorithme explicite-implicite n'est pas évidente. En utilisant des résultats récents d'analyse non lisse pour les fonctions à valeurs réelles [1], on peut cependant démontrer la propriété suivante :

Proposition 2.1. *Si Φ est une fonction semi-algébrique de gradient L -lipschitzien avec $L \in]0, +\infty[$, si les bornes $\underline{\gamma}$ et $\bar{\gamma}$ sur le pas de l'algorithme sont choisies telles que*

$$0 < \underline{\gamma} \leq \bar{\gamma} < L^{-1} (\|E\|^2 + \|E'\|^2)^{-1}, \quad (12)$$

et si $\epsilon > 0$, alors toute suite $(c^{(k)}, d^{(k)})_{k \in \mathbb{N}}$ générée par l'algorithme 1 converge vers un point critique de la fonction (10).

Rappelons qu'une fonction de \mathbb{C}^Q vers \mathbb{R} est semi-algébrique si son graphe est une union d'un nombre fini d'ensembles définis par un nombre fini d'inégalités polynomiales (en les parties

réelles et imaginaires de ses variables). Par ailleurs, $\|E\|$ (resp. $\|E'\|$) désigne la norme spectrale de la matrice E (resp. E').

Par ailleurs, nous avons pu établir que, si Φ est convexe, tout point critique de la fonction (10) est, sous une condition peu restrictive, minimiseur local de cette fonction.

3 Application à l'analyse spectrale

Afin d'illustrer le bon comportement de la méthode proposée, nous considérons un problème d'estimation spectrale parcimonieuse [3]. Nous proposons ici une alternative aux méthodes de raffinement de grilles de recherche [8, 9]. Nous considérons un signal analogique à valeurs complexes exprimé comme suit :

$$(\forall t \in \mathbb{R}) \quad s(t) = \sum_{n=1}^N \bar{c}_n \exp(i\bar{\nu}_n t) \quad (13)$$

où $(\bar{\nu}_n)_{1 \leq n \leq N} \in [0, 2\pi]^N$ sont des fréquences angulaires distinctes et $(\bar{c}_n)_{1 \leq n \leq N}$ sont des amplitudes complexes inconnues supposées parcimonieuses. Ce signal est perturbé par un bruit additif v . Le signal observé z s'écrit ainsi

$$(\forall t \in \mathbb{R}) \quad z(t) = s(t) + v(t). \quad (14)$$

Nous disposons de Q observations de z échantillonnées irrégulièrement à des instants $(\tau_q)_{1 \leq q \leq Q}$ choisis de manière aléatoire sur $[0, Q]$. Cela conduit au modèle (1)-(2) en posant $y = (z(\tau_q))_{1 \leq q \leq Q}$ et $w = (v(\tau_q))_{1 \leq q \leq Q}$. Le vecteur w est supposé gaussien complexe circulaire, centré et de matrice de covariance $\sigma^2 I_Q$.

Le dictionnaire est constitué de cisoides définies par

$$(\forall \nu \in \mathbb{R}) \quad e_\nu = (\exp(i\nu\tau_q))_{1 \leq q \leq Q} \quad (15)$$

et nous avons donc

$$(\forall \nu \in \mathbb{R}) \quad e'_\nu = (i\tau_q \exp(i\nu\tau_q))_{1 \leq q \leq Q}. \quad (16)$$

Une approche classique dans ce contexte est d'échantillonner uniformément le domaine fréquentiel :

$$(\forall n \in \{1, \dots, N\}) \quad \theta_n = 2\pi \frac{n-1}{N}. \quad (17)$$

En prenant, pour tout $n \in \{1, \dots, N\}$, $\Delta_n = \pi/N$, la méthode proposée permet de prendre en compte des pulsations qui ne sont pas multiples de $2\pi/N$.

Nous illustrons les performances de notre méthode dans le cas d'un signal complexe s correspondant à la somme de 4 cisoides. Le dictionnaire employé est composé de $N = 500$ cisoides de pulsations régulièrement échantillonnées sur $[0, 2\pi]$. Notons que les pulsations des composantes parcimonieuses de s n'appartiennent pas à la grille de recherche. Nous présentons dans le tableau 1 l'erreur quadratique normalisée du signal estimé avec l'algorithme 1, lorsque Φ est le carré de la norme euclidienne. Les valeurs présentées sont des moyennes obtenues par une étude de Monte Carlo menée sur 100 réalisations et pour différentes valeurs du rapport signal sur bruit (RSB).

TABLE 1 – Valeurs de l'erreur quadratique calculées pour 100 réalisations (4 cisoides).

RSB (dB)	ℓ_1	ℓ_0	Méthode proposée
14.82	0.452	0.224	0.043
19.82	0.445	0.206	0.024
24.82	0.443	0.198	0.013
29.72	0.442	0.201	0.007

A titre de comparaison, nous donnons également les valeurs d'erreur correspondant à l'utilisation de coûts ℓ_1 et ℓ_0 . Notre approche conduit à des gains substantiels par rapport à l'utilisation d'une stratégie basée sur la simple minimisation d'un coût ℓ_1 ainsi que sur l'emploi direct d'un algorithme de seuillage dur itératif. Notons que la minimisation ℓ_1 correspond plus précisément à la résolution du problème suivant :

$$\underset{c \in S'}{\text{minimiser}} \quad \ell_1(c) \quad (18)$$

où $S' = \{c \in \mathbb{C}^N \mid \|Ec - y\|^2 \leq Q\sigma^2\}$. Les résultats de cette minimisation servent d'initialisation à notre algorithme.

Un algorithme primal-dual proximal [5] est mis en œuvre pour résoudre ce problème de manière efficace. Par ailleurs, le paramètre de régularisation λ est choisi de manière automatique, par une approche dichotomique.

Les résultats présentés sur les figures 1(a) et (b) (resp. 1(c) et (d)) permettent d'évaluer la qualité des estimations typiquement obtenues lors de l'identification de 4 (resp. 6) cisoides.

4 Conclusion

Dans cet article, nous avons proposé une approche variationnelle pour l'estimation de signaux parcimonieux. L'originalité du modèle considéré réside dans le fait que les éléments du dictionnaire sont paramétrés par une variable réelle variant continuellement.

Un algorithme itératif proximal dont nous avons étudié les propriétés de convergence locale nous permet de résoudre le problème non convexe associé. Il convient de noter qu'en l'absence de perturbations, l'algorithme proposé se réduit à un algorithme de seuillage dur itératif.

Les résultats de simulation obtenus dans un contexte d'estimation spectrale, lorsque les observations sont irrégulièrement échantillonnées de façon aléatoire, nous ont permis d'évaluer la qualité des estimations produites.

Références

- [1] H. Attouch, J. Bolte, and B. F. Svaiter. Convergence of descent methods for semi-algebraic and tame problems : proximal algorithms, forward backward splitting, and regularized Gauss-Seidel methods. *Math. Prog.*, 137(1-2) : 1–39, Fév. 2013.

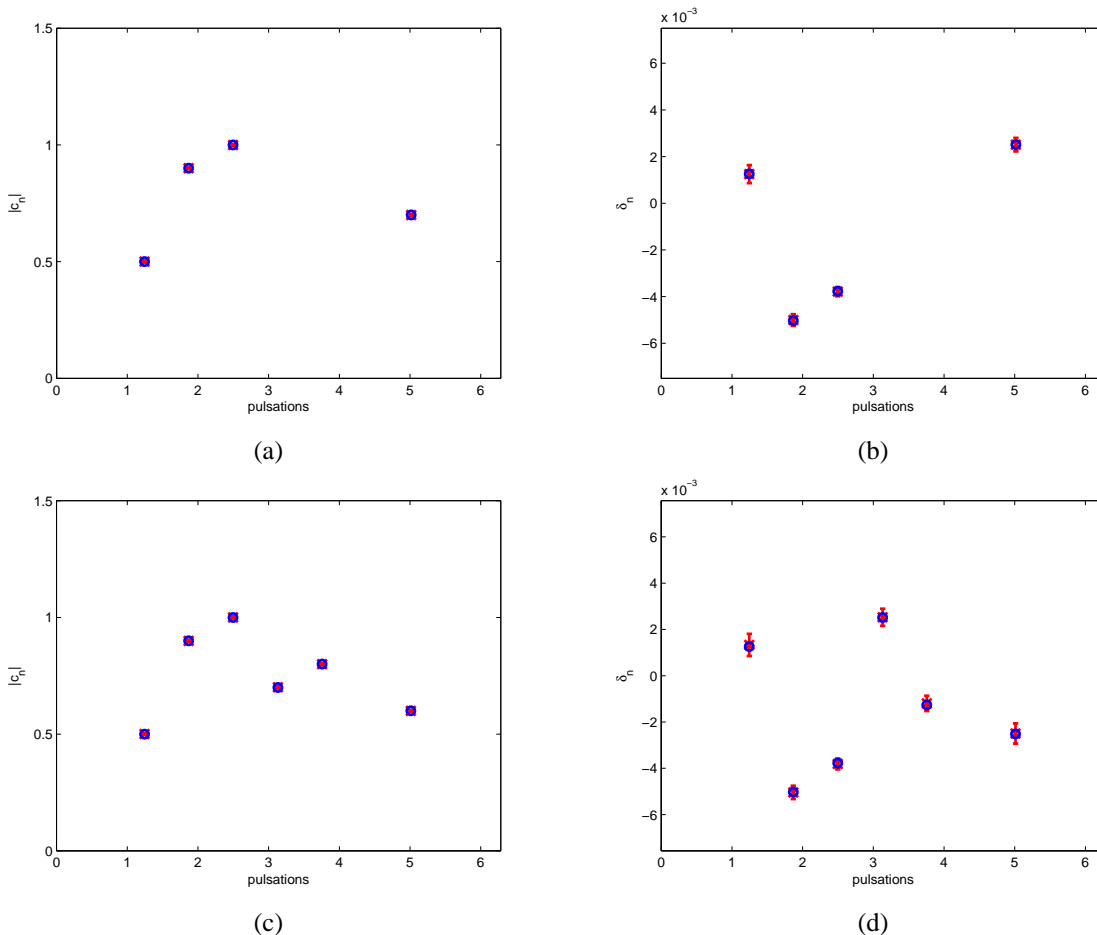


FIGURE 1 – Estimation de 4 cisoïdes pour un RSB = 19.82 dB (a)-(b) et de 6 cisoïdes pour un RSB = 23.27 dB (c)-(d) : amplitudes $(|c_n|)_{1 \leq n \leq N}$ (a)-(c), perturbations $(\delta_n)_{1 \leq n \leq N}$ en fonction des pulsations $(\theta_n)_{1 \leq n \leq N}$ sur la grille (b)-(d). Les valeurs exactes sont représentées par des cercles bleus et les estimées avec leur intervalle de confiance en rouge.

- [2] T. Blumensath and M. E. Davies. Iterative thresholding for sparse approximations. *J. Fourier Anal. Appl.*, 14(5-6) :629–654, 2008.
- [3] S. Bourguignon, H. Carfantan, and J. Idier. A sparsity-based method for the estimation of spectral lines from irregularly sampled data. *IEEE J. Sel. Topics Signal Process.*, 1(4) :575–585, Déc. 2007.
- [4] E. J. Candès and M. B. Wakin. An introduction to compressive sampling. *IEEE Signal Process. Mag.*, 25(2) :21–30, Mar. 2008.
- [5] A. Chambolle and T. Pock. A first-order primal-dual algorithm for convex problems with applications to imaging. *J. Math. Imag. Vis.*, 40(1) :120–145, 2011.
- [6] P. L. Combettes and J.-C. Pesquet. Proximal splitting methods in signal processing. In H. H. Bauschke, R. S. Burachik, P. L. Combettes, V. Elser, D. R. Luke, and H. Wolkowicz, editors, *Fixed-Point Algorithms for Inverse Problems in Science and Engineering*, pages 185–212. Springer-Verlag, New York, 2011.
- [7] C. Ekanadham, D. Tranchina, and E. P. Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *IEEE Trans. Signal Process.*, 59(10) :4735–4744, Oct. 2011.
- [8] S. Sahnoun, E. H. Djerroune, C. Soussen, and D. Brie. Sparse multidimensional modal analysis using a multigrid dictionary refinement. *EURASIP J. Adv. Sig. Proc.*, 60, Mar. 2012. 10 pp.
- [9] P. Stoica and P. Babu. Sparse estimation of spectral lines : Grid selection problems and their solutions. *IEEE Trans. Signal Process.*, 60(2) :962–967, Fév. 2012.
- [10] H. Zhu, G. Leus, and G. B. Giannakis. Sparsity-cognizant total least-squares for perturbed compressive sampling. *IEEE Trans. Signal Process.*, 59(5) :2002–2016, Mai 2011.
- [11] H. Zou and T. Hastie. Regularization and variable selection via the elastic net. *J. R. Statist. Soc. A*, 67(2) :301–320, 2005.