

Nouvelle génération d'Optique Adaptative pour l'astronomie Problème Inverse à grands nombres de degrés de liberté.

Clémentine BÉCHET¹, Michel TALLON¹, Éric THIÉBAUT¹, Gérard THOMAS², Éric BLANCO²

¹Centre de Recherche Astrophysique de Lyon (UMR CNRS 5574)
9 avenue Charles André, 69230 Saint-Genis Laval, France

²AMPERE (UMR CNRS 5005), Ecole Centrale de Lyon
36 avenue Guy de Collongue, 69134 Ecully cedex, France

Bechet@obs.univ-lyon1.fr, MTallon@obs.univ-lyon1.fr, Thiebaut@obs.univ-lyon1.fr
Gerard.Thomas@ec-lyon.fr, Eric.Blanco@ec-lyon.fr

Résumé – Les images astrophysiques issues des télescopes au sol sont dégradées par la turbulence de l'atmosphère terrestre. Un système d'Optique Adaptative (OA) doit corriger en temps réel ces perturbations. L'étude d'une nouvelle génération de télescopes, de plus de 30 mètres de diamètre, avec de nouveaux concepts d'OA ayant de 10^4 à 10^5 d.d.l., appelle de nouveaux algorithmes pour la commande du système. L'optimisation d'un critère nous a conduit à un contrôle en boucle fermée par modèle interne. Cette structure confère au système des propriétés de stabilité intéressantes du point de vue de l'OA et permet d'exploiter les *a priori* sur la turbulence en utilisant l'algorithme FrIM. Cette méthode est adaptée aux systèmes à grands nombres de ddl.

Abstract – Astrophysical images coming from ground-based telescopes are spoiled by atmospheric turbulence. Adaptive Optics systems should provide real-time correction of these disturbances. The study of a new generation of Extremely Large Telescopes, and of new AO concepts for them with 10^4 to 10^5 d.o.f., produced new algorithms for the closed-loop AO system control. First, we optimise a criterion and come to a closed-loop law with an Internal Model Control. Thanks to this design, prior information about turbulence statistics can be used to proceed a good regularization, and to benefit from FrIM, a very fast iterative algorithm adapted to high number of d.o.f.

L'observation d'objets de plus en plus faibles dans l'univers requiert des télescopes au sol toujours plus grands. Cependant, la résolution des images fournies par ces grands télescopes est dégradée par la turbulence atmosphérique. Pour allier sensibilité et haute résolution angulaire, les télescopes au sol sont désormais associés à des systèmes d'optique adaptative[1] (OA). L'OA doit corriger en temps réel les effets de la turbulence atmosphérique grâce à une boucle comprenant un analyseur de front d'onde (AFO), un système de commande et un miroir déformable (MD) pour appliquer la correction, comme schématisé sur la Fig. 1.

Les systèmes d'OA actuellement mis en oeuvre en astronomie utilisent des miroirs à quelques centaines d'actionneurs, *i.e.* quelques centaines de degrés de liberté. Ils sont asservis de façon à ce que les mesures de perturbations résiduelles du front d'onde, fournies par l'AFO, soient nulles. Pour les futurs systèmes d'OA, la démarche est désormais différente. D'une part, la portion de front d'onde analysée par l'AFO ne coïncide pas toujours avec la portion de front d'onde destinée à la voie d'observation, ce qui n'implique pas nécessairement l'annulation des mesures résiduelles. De plus, il est souvent nécessaire d'extrapoler le front d'onde reconstruit. Nous traitons ici cette reconstruction par une approche de type problème inverse, avec une régularisation déduite des propriétés statistiques de la turbulence. D'autre part, les OA en astronomie compteront quelques dizaines de milliers d'actionneurs, d'où l'exi-

gence de nouveaux algorithmes de reconstruction rapide, adaptés à ces grands nombres de degrés de liberté.

Dans un premier temps, nous expliquons la spécificité des signaux mis en jeu dans un système d'OA, notamment leur caractéristiques statistiques. Puis, nous montrons que l'optimisation d'un critère de qualité concernant les images astrophysiques conduit à un reconstructeur à minimum de variance et une loi de commande de type modèle interne. Enfin, nous proposons un algorithme itératif, FrIM, reposant sur un préconditionneur fractal, qui calcule la commande en moins d'opérations qu'aucune autre méthode existante.

1 Caractériser la phase turbulente

Nous considérons ici un système d'OA simple (*cf.* Fig. 1), *i.e.* comprenant un seul AFO et un seul MD, en boucle fermée. Le front d'onde incident sur la pupille du télescope est déformé par la turbulence atmosphérique. D'après le modèle de turbulence de Kolmogorov, l'énergie cinétique turbulente est transmise en cascade des tourbillons à grande échelle vers les plus petites structures. Ceci engendre des fluctuations spatiales et temporelles aléatoires de température dans l'atmosphère, et donc de l'indice de réfraction de l'air. Après propagation au travers de la turbulence atmosphérique, la surface d'onde au niveau de la pupille présente des retards de phase [2]. D'après le mo-

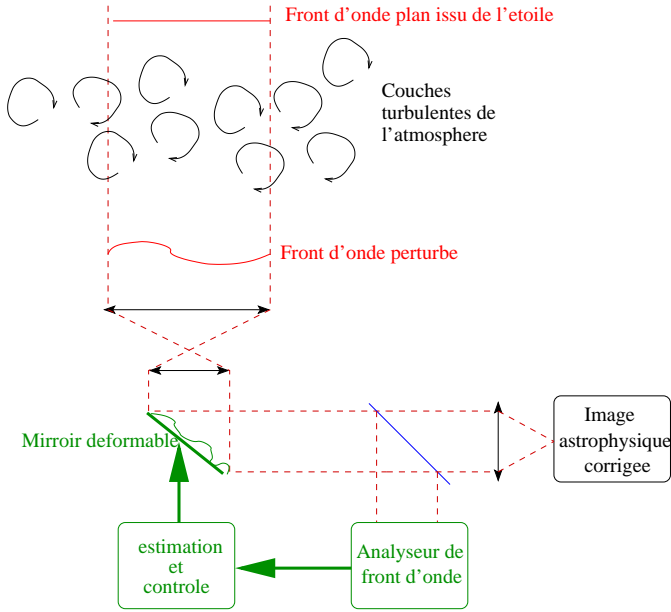


FIG. 1 – Schéma de principe d'une OA simple

dèle de Kolmogorov, ces fluctuations de phase w suivent une statistique gaussienne centrée en chaque point de la pupille et dont la fonction de structure s'écrit :

$$D_w(\boldsymbol{\rho}) \stackrel{\text{def}}{=} \langle [w(\boldsymbol{\rho} + \boldsymbol{\rho}') - w(\boldsymbol{\rho})]^2 \rangle_{\boldsymbol{\rho}} = 6.88 \left(\frac{\|\boldsymbol{\rho}\|}{r_0} \right)^{5/3} \quad (1)$$

où $\langle \cdot \rangle$ représente l'espérance mathématique, $\boldsymbol{\rho}$ et $\boldsymbol{\rho}'$ sont des vecteurs de position dans la pupille du télescope et r_0 est le paramètre de Fried qui caractérise la quantité de turbulence[3]. Par ailleurs, dans le modèle de Kolmogorov, la statistique des fluctuations de la phase w de l'onde a la particularité d'être invariante par changement d'échelle, d'où la nature fractale de la phase turbulente.

L'OA vise à compenser en temps réel les déformations des fronts d'onde. Bien que la phase soit continue dans l'espace et dans le temps, le nombre de degrés de liberté N_{act} , le diamètre D de la pupille ou encore la fréquence de la boucle, définissent certaines limites de performance. Nous considérons ces valeurs comme fixées et nous nous ramonnons ici à un système d'OA linéarisé équivalent, discrétisé à la fois spatialement et temporellement. Les vecteurs \mathbf{w} et \mathbf{w}^c , de \mathbb{R}^N , sont les valeurs échantillonnées des fronts d'onde, respectivement incident et corrigé, sur une grille régulière à 2 dimensions, incluant la pupille du télescope. L'AFO fournit le vecteur des mesures de dérivées spatiales locales, à deux dimensions, \mathbf{d} dans \mathbb{R}^M , et le vecteur des commandes \mathbf{a} est dans $\mathbb{R}^{N_{act}}$. L'équation de mesure du système s'écrit

$$\mathbf{d}_k = \mathbf{S} \cdot (\mathbf{w}_{k-1} - \mathbf{w}_{k-1}^c) + \mathbf{n}_k \quad (2)$$

avec \mathbf{S} le modèle linéaire de l'AFO et \mathbf{n} le vecteur de bruit et d'erreur de modélisation des mesures. Les indices k permettent de représenter les retards caractéristiques du système. ; ici, le temps de pose de l'AFO. Les erreurs de mesure sont supposées Gaussiennes centrées et indépendantes entre elles. D'où une matrice de covariance du bruit \mathbf{C}_n diagonale.

En conséquence, l'objectif est de contrôler un système

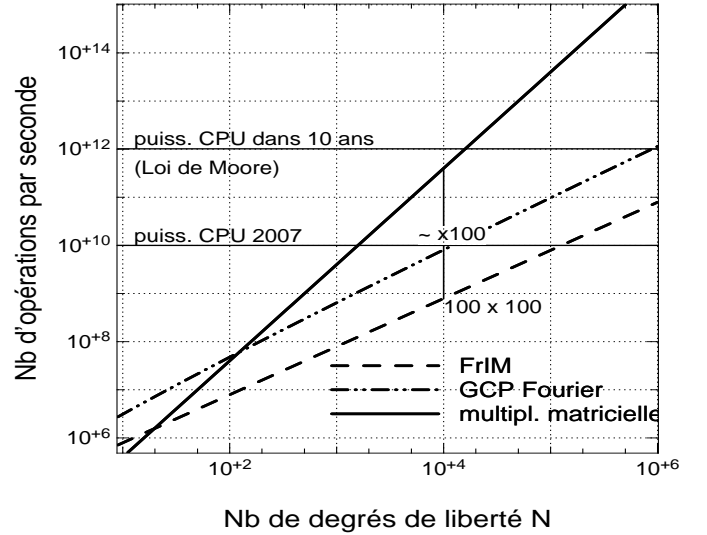


FIG. 2 – Nombres d'opérations par seconde nécessaires pour différents algorithmes de reconstruction, en fonction du nombre de degrés de liberté du système d'OA. Sur les OA actuelles : Multiplication matrice-vecteur. Récemment suggéré : Gradients Conjugués Préconditionnés dans le domaine de Fourier. Mis en oeuvre ici : FrIM, Gradients Conjugués avec Préconditionneur fractal.

multi-entrées multi-sorties à plusieurs milliers de degrés de liberté et à une fréquence de l'ordre de 1 kHz.

2 L'optimisation du critère

La qualité des images corrigées croît lorsque la variance de la phase résiduelle sur la pupille,

$$\langle \epsilon^2 \rangle = \langle \|\mathbf{P} \cdot (\mathbf{w} - \mathbf{w}^c)\|^2 \rangle = \langle \|\mathbf{P} \cdot (\mathbf{w} - \mathbf{M} \cdot \mathbf{a})\|^2 \rangle \quad (3)$$

diminue. \mathbf{M} est le modèle d'influence du miroir. \mathbf{P} est un opérateur linéaire qui ôte à la phase sa composante constante sur la pupille, *le piston*, et annule la phase hors de la pupille. La norme $\|\cdot\|$ représente la norme Euclidienne canonique de \mathbb{R}^N . Partant du critère ci-dessus, nous avons écrit, en toute généralité, la commande \mathbf{a} comme une combinaison linéaire des mesures résiduelles et de la commande appliquée pendant la pose

$$\mathbf{a}_k = \mathbf{R} \cdot \mathbf{d}_{k-1} + \mathbf{Q} \cdot \mathbf{a}_{k-2} \quad (4)$$

Un deuxième retard majeur du système est donc implicitement pris en compte ici ; le temps de lecture du CCD et de calcul de la commande d'après l'Eq. (4). Pour simplifier les notations, on considère que ce deuxième retard est égal au premier, le temps de pose. $\langle \epsilon^2 \rangle$ est un critère quadratique vis-à-vis de \mathbf{R} et de \mathbf{Q} . L'Eq. (4) tient compte du fait que le miroir a influencé les mesures, grâce à l'introduction d'une matrice \mathbf{Q} . Par ailleurs, l'équation de mesure (2) implique

$$\mathbf{a}_k = \mathbf{R} \cdot (\mathbf{S} \cdot \mathbf{w}_{k-2} + \mathbf{n}_{k-1}) + (\mathbf{Q} - \mathbf{R} \cdot \mathbf{S} \cdot \mathbf{M}) \cdot \mathbf{a}_{k-2} \quad (5)$$

En choisissant $\mathbf{Q} = \mathbf{R} \cdot \mathbf{S} \cdot \mathbf{M}$, le second terme dans l'Eq. (5) disparaît. On obtient finalement une commande en pseudo boucle ouverte,

$$\mathbf{a}_k = \mathbf{R} \cdot \widetilde{\mathbf{d}}_{k-1} = \mathbf{R} \cdot (\mathbf{d}_{k-1} + \mathbf{S} \cdot \mathbf{M} \cdot \mathbf{a}_{k-2}) \quad (6)$$

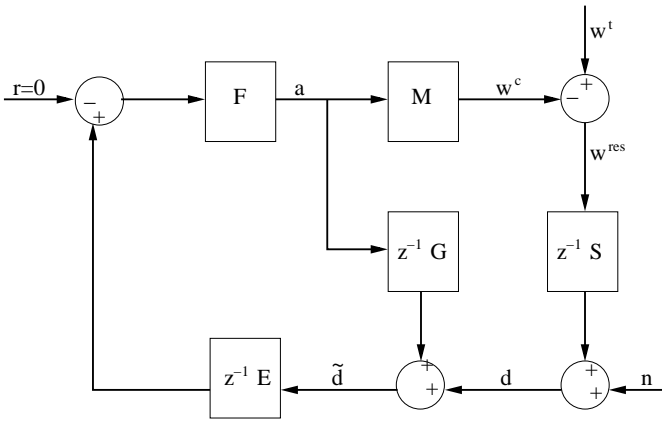


FIG. 3 – Diagramme du système d’OA en boucle fermée

où $\widetilde{\mathbf{d}}_{k-1}$ fournit une estimation de mesures en boucle ouverte, *i.e.* sans la correction du miroir. Ensuite, l’optimisation du critère $\langle \epsilon^2 \rangle$ par rapport à \mathbf{R} fournit la commande optimisée

$$\mathbf{a}_k = \mathbf{F} \cdot \mathbf{E} \cdot \widetilde{\mathbf{d}}_{k-1} \quad (7)$$

avec

$$\mathbf{F} = \mathbf{M}^\dagger \cdot \mathbf{P} \quad (8)$$

$$\mathbf{E} = \mathbf{C}_{w,2} \cdot \mathbf{C}_{w,0}^{-1} \cdot (\mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{S} + \mathbf{C}_{w,0}^{-1})^{-1} \cdot \mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \quad (9)$$

où \mathbf{M}^\dagger représente l’inverse généralisée de la matrice du MD, $\mathbf{C}_{w,0} = \langle \mathbf{w}_k \mathbf{w}_k^\top \rangle$ et $\mathbf{C}_{w,2} = \langle \mathbf{w}_k \mathbf{w}_{k-2}^\top \rangle$ sont respectivement des matrices de covariances spatiale et spatio-temporelle de la phase turbulente.

A partir des pseudo-mesures $\widetilde{\mathbf{d}}_{k-1}$ dans l’Eq. (7), on estime la phase turbulente $\widetilde{\mathbf{w}}_{k-2}$ grâce à $\mathbf{R}^\dagger = (\mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{S} + \mathbf{C}_{w,0}^{-1})^{-1} \cdot \mathbf{S}^\top \cdot \mathbf{C}_n^{-1}$. Cet opérateur linéaire, \mathbf{R}^\dagger , est plus connu sous le nom de reconstituteur à minimum de variance. Étant donné le double retard avec lequel les commandes seront envoyées au miroir, l’Eq. (9) contient un modèle d’évolution temporelle, défini à partir de l’opérateur linéaire $\mathbf{C}_{w,2} \cdot \mathbf{C}_{w,0}^{-1}$, permettant d’estimer la phase turbulente à venir. Enfin, \mathbf{F} illustre la projection de la correction de phase sur l’espace des actionneurs.

A partir de ces équations, il est possible de représenter le système d’OA discretisé par le diagramme de la figure (3). \mathbf{G} est le modèle interne du produit $\mathbf{S} \cdot \mathbf{M}$, utilisé pour décrire la relation entre la commande et les pentes induites. Cet opérateur \mathbf{G} est appelé la matrice d’interactions.

3 Algorithme rapide, FrIM, pour la reconstruction

Pour un système à $N = 10^4$ degrés de liberté et environ 2 fois plus de mesures M , d’après les Eq. (7), (8) et (9), si la matrice $\mathbf{F} \cdot \mathbf{E}$ est précalculée, son application à un vecteur de pseudo-mesures à la fréquence de 1kHz impliquerait $\simeq 4 \times N^2 \times 10^3 \approx 4.10^{11}$ opérations par seconde. On voit sur la Fig. 2 que, même si cette multiplication matricielle est la méthode implémentée sur les systèmes d’OA existants, elle n’est plus applicable pour de grands

systèmes. Étant données les capacités des processeurs actuellement disponibles dans le commerce, et leur supposée évolution, il est nécessaire de développer des méthodes plus rapides. Pour cela, il a été suggéré[5, 6, 4] de résoudre itérativement, dans un premier temps, le système

$$\mathbf{A} \cdot \widehat{\mathbf{w}}_{k-1} = \mathbf{b} \quad (10)$$

avec

$$\mathbf{A} = \mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{S} + \mathbf{C}_{w,0}^{-1} \quad \text{et} \quad \mathbf{b} = \mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \widetilde{\mathbf{d}}_k \quad (11)$$

par un algorithme de gradients conjugués préconditionnés. Puis, il faudra appliquer la prédiction et la projection. La structure creuse des matrices \mathbf{S} , \mathbf{M}^\dagger , \mathbf{C}_n^{-1} , \mathbf{P} permet de les appliquer, elles et leurs transposées, à un vecteur en un nombre d’opérations de l’ordre de N . En revanche, $\mathbf{C}_{w,0}$ et $\mathbf{C}_{w,2}$, tout comme leurs inverses, ne sont pas creuses, d’où la difficulté de prendre en compte des *a priori* sur la turbulence. Une bonne approximation consiste d’abord à simplifier l’expression de la commande dans l’Eq. (9) en n’introduisant pas d’*a priori* temporels, ce qui revient à écrire $\mathbf{C}_{w,0} = \mathbf{C}_{w,2}$. Ceci se justifie par le fait que la turbulence change effectivement peu d’une boucle à l’autre. Il reste encore à appliquer $\mathbf{C}_{w,0}^{-1}$ à un vecteur pour le calcul des résidus de (10).

Yang et al.[6] ont proposé d’approximer $\mathbf{C}_{w,0}^{-1}$ dans le domaine de Fourier par une matrice diagonale. Dans ce cas, le coût d’application des Gradients Conjugués préconditionnés dans le domaine de Fourier est représenté sur la Fig. 2. Il est proportionnel à $N \log(N)$. *Thiébaud & Tallon*[4] ont, eux, utilisé un opérateur fractal, inversible, \mathbf{K} , décrit par un algorithme de points-milieux et tel que $\mathbf{K} \cdot \mathbf{K}^\top \approx \mathbf{C}_{w,0}$. La nature fractale de la turbulence permet à cette approximation de respecter la fonction de structure de Kolmogorov (*cf.* Eq. (1)) pour la phase et de définir une bijection $\mathbf{w} = \mathbf{K} \cdot \mathbf{u}$. L’application de l’opérateur fractal, de son inverse ou de sa transposée, à un vecteur nécessite $\simeq 6 \times N$ opérations. $\mathbf{u} = \mathbf{K}^{-1} \cdot \mathbf{w}$ est alors un vecteur de coefficients aléatoires gaussiens centrés de variance unité. Grâce à ce changement de variables comme préconditionnement à droite et à gauche des gradients conjugués, la méthode fractale, dénommée FrIM[4], permet de résoudre itérativement le système équivalent à (10)

$$\mathbf{A}' \cdot \widehat{\mathbf{u}}_{k-1} = \mathbf{b}' \quad (12)$$

avec

$$\mathbf{A}' = \mathbf{K}^\top \cdot \mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \mathbf{S} \cdot \mathbf{K} + \mathbf{I}_d \quad \text{et} \quad \mathbf{b}' = \mathbf{K}^\top \cdot \mathbf{S}^\top \cdot \mathbf{C}_n^{-1} \cdot \widetilde{\mathbf{d}}_k \quad (13)$$

Initialement utilisée pour la reconstruction de surfaces d’ondes en boucle ouverte, FrIM est utilisée ici pour l’estimation en boucle fermée. Le changement de variables accélère la convergence des gradients conjugués, et nos simulations montrent que l’ajout d’un préconditionneur diagonal à gauche[4] permet même d’obtenir la solution du système (12) en 2 itérations, pour un rapport signal à bruit moyen.

En conclusion, nous avons vu que l’optimisation du critère nous conduit à une reconstruction de front d’onde turbulent, à variance d’erreur minimale. L’utilisation de la méthode FrIM pour ce calcul fournit le vecteur de commande en quelques $80 \times N$ opérations. D’après la Fig. 2, le

gain en nombre d'opérations par seconde sur la multiplication matricielle pour $N = 10^4$ est d'un facteur 100. De fait, la méthode FrIM permettrait donc dès aujourd'hui, avec les processeurs existants, de commander un système d'OA d'une telle taille.

Références

- [1] F. Roddier. Adaptive Optics in Astronomy. *Cambridge University Press.*, 1999.
- [2] F. Roddier. The effects of atmospheric turbulence in optical astronomy. *Progress in optics*, Vol. 19 : 281–376, Amsterdam, North-Holland Publishing Co., 1981.
- [3] D.L. Fried. Statistics of a geometric representation of wavefront distortion. *J. Opt. Soc. Am. A*, 55 :1435, 1965.
- [4] É. Thiébaud and M. Tallon. Ultra fast maximum a posteriori wave front reconstruction for extremely large telescopes. *submitted to J. Opt. Soc. Am. A*, 2007.
- [5] L. Gilles, C.R. Vogel and B.L. Ellerbroek. Multigrid preconditioned conjugate-gradient method for large-scale wave-front reconstruction. *J. Opt. Soc. Am. A*, 19 :1817–1822, 2002.
- [6] Q. Yang, C.R. Vogel and B.L. Ellerbroek Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography. *Applied Optics*, 45 :5281–5293, 2006.