

Adapting optimal LQG methods to ELT-sized AO systems

Carlos Correia^{a,b,*}, Jean-Marc Conan^a, Caroline Kulcsár^b, Henri-François Raynaud^b and Cyril Petit^a

^aONERA – Office National d’Etudes et de Recherches Aéronautiques,
 29 Av. de la Division Leclerc, 92322 Châtillon, France

^bL2TI – Laboratoire de Traitement et de Transport de l’Information, Univ. Paris 13,
 99 Av. Jean Baptiste Clément, 93430 Villetaneuse, France

*ccorreia@onera.fr

Abstract. The minimum-variance (MV) phase residual Linear-Quadratic-Gaussian (LQG) solution in Adaptive Optics (AO) using fast iterative methods is outlined. Standard LQG computational burden scaling as $O(n^2)$ is shown to be $O(n \log n)$ using spectral methods. Proof-of-concept is made through Monte Carlo simulations.

1 Introduction

The optimal minimum residual phase variance solution for the reconstruction and control in Adaptive Optics (AO) is known to correspond to a discrete-time Linear Quadratic Gaussian (LQG) formulation [1], [2]. It combines a Kalman filter for the estimation of the disturbance phase from noisy measurements with, in the absence of deformable mirror (DM) dynamics, an orthogonal projection onto the DM space. This approach has proven its potential at delivering superior performance in terms of noise rejection, vibration damping and phase reconstruction in classical and wide-field AO [3], [4], [5]. However, LQG AO control, because it involves an explicit real-time reconstruction of the turbulent phase, is more computationally demanding than the standard gain plus integrator control.

With the advent of the Extremely Large Telescopes (ELT) the number n of degrees-of-freedom (DoF) is going to increase considerably. With the computational methods employed so far, the real-time cost of LQG scales as $O(n^2)$ multiplications per sample period, a serious obstacle for ELT-sized systems. Since n grows with the square of the telescope diameter, the ELT’s are at least 2 orders of magnitude more computationally demanding than VLT-like systems.

In this contribution several strategies to render the LQG methods implementable in their new working scenarios in both processing time and memory requirements are outlined. They involve a zonal description of the several parameter spaces together with improved off-line computation of the Kalman filter gain. Real-time operations are shown to be feasible by replacing vector-matrix multiplies by spectral iterative methods that use a sparse approximation to the Riccati equation solution, further bypassing the determination of the Riccati solution itself. The potential of these new approaches is evaluated through Monte Carlo simulations on a down-scaled single-conjugate AO configuration with a Hartmann-Shack (HS) wave-front sensor (WFS).

2 Linear-Quadratic-Gaussian approach in Adaptive Optics

Let’s start with the known and experimentally validated LQG controller, with a compacted control-oriented dynamical state-space representation [6]

$$\begin{cases} \begin{pmatrix} \bar{\phi}_{k+2}^{tur} \\ \bar{\phi}_{k+1}^{tur} \\ u_k \end{pmatrix} = \begin{pmatrix} \mathcal{A}_{tur} & 0 & 0 \\ I & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \bar{\phi}_{k+1}^{tur} \\ \bar{\phi}_k^{tur} \\ u_{k-1} \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ I \end{pmatrix} u_k + \begin{pmatrix} I \\ 0 \\ 0 \end{pmatrix} v_k \\ z_k = D \begin{pmatrix} 0 & I & -N \end{pmatrix} x_k + w_k \end{cases} \quad (1)$$

where $\bar{\phi}_k^{tur}$ is the averaged phase for $t \in [(k-1)T_s, kT_s]$, u_k is the DM control vector and v_k and w_k are zero-mean Gaussian independent white noises of known covariances. The linear operators D

and N translate the WFS and DM operation. The observations z_k in (1) relate to the two-steps delayed measurements by $y_k = z_{k-1}$. Let \mathcal{Y}_k be the sequence of all measurements available at $t = kT_s$, namely $\mathcal{Y}_k = \{y_0, \dots, y_k\}$ and define x^{tur} as the part of the state gathering the phase components $x^{tur} \triangleq (\bar{\phi}_{k+1}^{tur,T}, \bar{\phi}_k^{tur,T})^T$ with $\mathcal{A}_d = \begin{pmatrix} \mathcal{A}_{tur} & 0 \\ 0 & 0 \end{pmatrix}$, $\mathcal{B}_d = 0$ and $\mathcal{C}_d = \begin{pmatrix} 0 & I \end{pmatrix}$ concerning only x^{tur} in (1).

With these notations, performed real-time operations at iteration k are

$$\begin{aligned} 1) \quad \widehat{y}_{k|\mathcal{Y}_{k-1}} &= D(0 \ I) \widehat{x}_{k-1|\mathcal{Y}_{k-1}}^{tur} - DNu_{k-2}, \\ 2) \quad \widehat{x}_{k|\mathcal{Y}_k}^{tur} &= \mathcal{A}_d \widehat{x}_{k-1|\mathcal{Y}_{k-1}}^{tur} + \mathcal{A}_d \mathcal{H}_\infty (y_k - \widehat{y}_{k|\mathcal{Y}_{k-1}}), \\ 3) \quad u_k &= P \widehat{\phi}_{k+1|\mathcal{Y}_k}^{tur} = P(I \ 0) \widehat{x}_{k|\mathcal{Y}_k}^{tur}, \end{aligned} \quad (2)$$

with the asymptotic observer gain

$$\mathcal{L}_\infty = \mathcal{A}_d \mathcal{H}_\infty = \mathcal{A}_d \Sigma_\infty \mathcal{C}_d^T (\mathcal{C}_d \Sigma_\infty \mathcal{C}_d^T + \Sigma_w)^{-1}, \quad (3)$$

obtained from the asymptotic solution of the discrete Riccati equation

$$\Sigma_{k+1} = \mathcal{A}_d \Sigma_k \mathcal{A}_d^T + \Gamma^T \Sigma_v \Gamma - \mathcal{A}_d \Sigma_k \mathcal{C}_d^T (\mathcal{C}_d \Sigma_k \mathcal{C}_d^T + \Sigma_w)^{-1} \mathcal{C}_d \Sigma_k \mathcal{A}_d^T. \quad (4)$$

In practice, under the stationary hypothesis, the asymptotic solutions of both the Kalman gain \mathcal{L}_∞ in (3) and the estimation error covariance matrix in (4) are used, which still grants the optimality of the infinite-horizon control decisions. These are obtained off-line.

3 Off-line calculation of Kalman gain \mathcal{L}_∞

The estimation error covariance matrix Σ_∞ is determined by solving a reduced-size algebraic Riccati equation (ARE) that involves only one phase occurrence [3], by doing:

$$\Sigma_\infty = \begin{pmatrix} M & Q^T \\ Q & R \end{pmatrix} = \begin{pmatrix} \mathcal{A}_{tur} R \mathcal{A}_{tur}^T + \Sigma_v & \mathcal{A}_{tur} R^T \\ R \mathcal{A}_{tur}^T & \mathcal{A}_{tur} R \mathcal{A}_{tur}^T + \Sigma_v - \mathcal{A}_{tur}^T R^T D^T (DRD^T + \Sigma_w)^{-1} DR \mathcal{A}_{tur} \end{pmatrix}. \quad (5)$$

Note that the lower-right terms involving R define a Riccati equation on the half-sized system. Upon determination of R , Σ_∞ is found by applying the expressions in Eq. (5) and finally \mathcal{L}_∞ from Eq. (3).

In the remainder of this paper, the doubling algorithm [7] to solve for R has been used. Solution at iteration k is equivalent to solution of fixed-point Riccati iteration 2^k as defined in (4). At iteration k ,

$$\begin{aligned} \alpha_{k+1} &= \alpha_k (I + \beta_k \gamma_k)^{-1} \alpha_k, \\ \beta_{k+1} &= \alpha_k (I + \beta_k \gamma_k)^{-1} \beta_k \alpha_k, \\ \gamma_{k+1} &= \gamma_k + \alpha_k^T \gamma_k (I + \beta_k \gamma_k)^{-1} \alpha_k, \end{aligned} \quad (6)$$

with initial guesses $\alpha_{-1} = \mathcal{A}_{tur}^T$, $\beta_{-1} = D^T \Sigma_w^{-1} D$, $\gamma_{-1} = \Sigma_v$. The solution is found from

$$R = \lim_{k \rightarrow +\infty} \gamma_{k+1}. \quad (7)$$

Figure 1 compares the convergence rate of the doubling vs. the fixed-point Riccati iteration in (4) to the reference non-iterative ARE solution [7]. Since just using a quarter of the total matrix size, with the doubling algorithm overall gains in operations, with a total of 8 iterations, are of the order of 250 (considering a factor 8 by avoiding $O(n^3)$ inversion and $2^8/8 = 32$ less iterations).

4 Iterative LQG solution using a sparse approximation to Σ_∞

Though fast methods exist to determine the off-line Kalman gain, the real-time operations to perform in (2) do involve several vector-matrix multiplies that scale as $O(n^2)$. To relax the computational burden, a zonal description of phase, measurements and DM influence functions has been adopted, like in [8]. Section 4.1 shows that this option tends to sparsify linear operators and thus reduce the burden of (2). Upon this, in sect. 4.2, fast iterative methods are applied provided an approximate Σ_∞ exists that be physically meaningful.

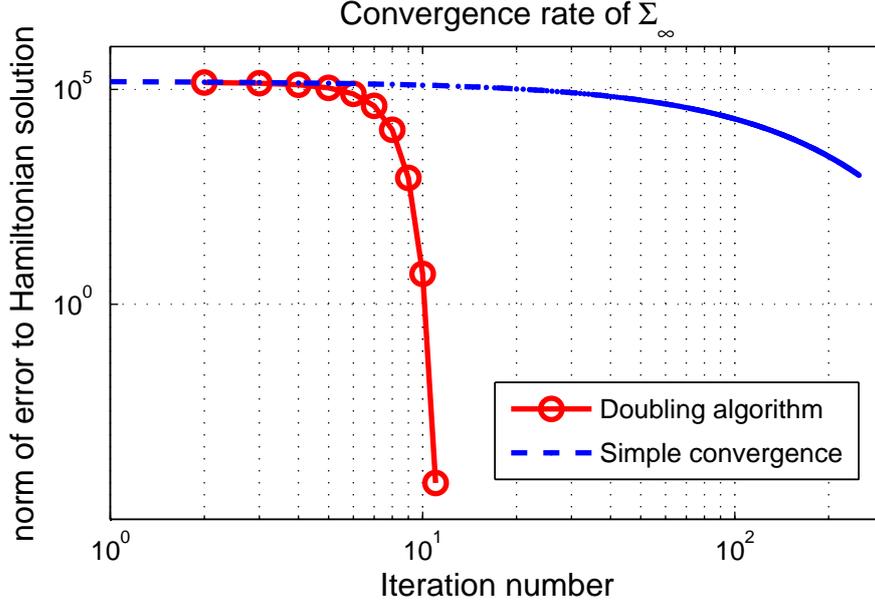


Fig. 1. Convergence rate of the doubling and fixed-point iterative algorithms. The reference solution was computed via the eigendecomposition of the Euler matrix.

4.1 Modelling using sparse techniques

The elementary operations in (2) are detailed separately using a zonal description of phase, measurements and DM influence functions.

Measurement estimate $\widehat{y}_{k-1|y_{k-1}} = D(0 I)\widehat{x}_{k|y_{k-1}}^{tur} - DNu_{k-2}$: Using a zonal approach to describe the WF phase, D becomes sparse because the kernel of the HS operation involves discrete differences of phase points around each sub-aperture. Adopting the Fried geometry, only four phase points are needed to compute the HS slopes. Matrix N is also sparse for the DM influence functions are localised (using bi-cubic splines or clipped Gaussians).

Measurement update and prediction $\widehat{x}_{k|y_k}^{tur} = \mathcal{A}_d\widehat{x}_{k-1|y_{k-1}}^{tur} + \mathcal{A}_d\mathcal{H}_\infty(y_k - \widehat{y}_{k|y_{k-1}})$: A sparse \mathcal{A}_{tur} operator is seek that approximates the natural *full* solution from $\mathcal{A}^* = \arg \min_{\mathcal{A}_{tur}} \langle \|\phi_{k+1} - \mathcal{A}_{tur}\phi_k\|^2 \rangle$, yielding $\mathcal{A}^* = \langle \phi_{k+1}\phi_k^T \rangle \langle \phi_k\phi_k^T \rangle^{-1}$. This solution has entries diagonally concentrated. Sparse approximations for both \mathcal{A}_{tur} and \mathcal{H}_∞ are detailed in sect. 4.2 to be used with fast solvers of sparse linear systems.

Projection onto DM influence functions $u_k = P(I 0)\widehat{x}_{k|y_k}^{tur} = (N^T N)^{-1}N^T\widehat{\phi}_{k+1|y_k}^{tur}$: The matrix N being sparse, fast Gaussian elimination methods are applicable to solve for u_k .

4.2 Transforming the LQG solution into a sparse linear system of the type $Ax = b$

All operations but the vector-matrix product involving $\mathcal{A}_d\mathcal{H}_\infty$ are relaxed using sparse modelling techniques: matrix C_d is sparse since both D and N are sparse. Matrix \mathcal{A}_d is sparse altogether.

Consider uncorrelated noise between sub-apertures, $\Sigma_w = \sigma^2 I$. Then \mathcal{H}_∞ in Eq. (3) can be written, in informations form, as

$$\mathcal{A}_d\mathcal{H}_\infty = \mathcal{A}_d(C_d^T C_d + \sigma^2 \Sigma_\infty^{-1})^{-1} C_d^T. \quad (8)$$

Solution x of the multiplication of \mathcal{H}_∞ by a vector can be found from the solution of the linear system $Ax = b$ with $A = (C_d^T C_d + \sigma^2 \Sigma_\infty^{-1})$ and $b = C_d^T x$, x a column vector. Multiplication by \mathcal{A}_d is done right after to complete the solution of the multiplication by (8). Under the assumption that A is sparse, then the computational costs are to decrease according to the sparseness patterns of A and the convergence rate of the iterative algorithms.

Using the matrix inversion lemma and the Shur complement, from (5) one gets

$$\Sigma_\infty^{-1} = \begin{pmatrix} M & Q^T \\ Q & R \end{pmatrix}^{-1} = \begin{pmatrix} S_d^{-1} & -S_d^{-1} Q^T R^{-1} \\ -R^{-1} Q S_d^{-1} & R^{-1} + R^{-1} Q S_d^{-1} Q^T R^{-1} \end{pmatrix}, \quad (9)$$

with

$$S_d = M - Q^T R^{-1} Q = \Sigma_v, \quad (10)$$

where M , Q and $R = R^T$ are defined in (5). This leads to

$$S_d^{-1} = \Sigma_v^{-1}. \quad (11)$$

Simplifying the term

$$-S_d^{-1} Q^T R^{-1} = -\Sigma_v^{-1} \mathcal{A}_{tur} R^T R^{-1} = -\Sigma_v^{-1} \mathcal{A}_{tur}, \quad (12)$$

and

$$-R^{-1} Q S_d^{-1} = -R^{-1} R^T \mathcal{A}_{tur}^T \Sigma_v^{-1} = -\mathcal{A}_{tur}^T \Sigma_v^{-1}, \quad (13)$$

and finally

$$R^{-1} + R^{-1} Q S_d^{-1} Q^T R^{-1} = R^{-1} + \mathcal{A}_{tur}^T \Sigma_v^{-1} \mathcal{A}_{tur}, \quad (14)$$

one hence ends up with a matrix inverse of the form

$$\Sigma_\infty^{-1} = \begin{pmatrix} M & Q^T \\ Q & R \end{pmatrix}^{-1} = \begin{pmatrix} \Sigma_v^{-1} & -\Sigma_v^{-1} \mathcal{A}_{tur} \\ -\mathcal{A}_{tur}^T \Sigma_v^{-1} & R^{-1} + \mathcal{A}_{tur}^T \Sigma_v^{-1} \mathcal{A}_{tur} \end{pmatrix}, \quad (15)$$

A sparse approximation to Σ_∞^{-1} is thus needed to comply with the beneficial use iterating a sparse linear system to get the solution of the product of a vector x by \mathcal{H}_∞ .

4.3 Sparse approximation to Σ_∞^{-1}

From Eq. (15), Σ_∞^{-1} has a sparse representation if both \mathcal{A}_{tur} , Σ_v^{-1} and R^{-1} do as well. Consider a simple auto-regressive (AR) model of the first order

$$\bar{\phi}_{k+1} = \mathcal{A}_{tur} \bar{\phi}_k + v_k, \quad (16)$$

to model the temporal evolution of the atmospheric phase. Then

$$\Sigma_v = \Sigma_\phi - \mathcal{A}_{tur} \Sigma_\phi \mathcal{A}_{tur}^T. \quad (17)$$

With a zonal description of the atmospheric phase, coefficients of \mathcal{A}_{tur} are all diagonally equal. Considering simple first order AR model then $\mathcal{A}_{tur} = aI$, with a found to match the normalised one-step auto-correlation function [6], a rather sparse representation. In that case it comes

$$\Sigma_v = \Sigma_\phi(1 - a^2) \quad (18)$$

which is a *full* matrix. However, its inverse involves Σ_ϕ^{-1} for which a sparse approximation proposed by Ellerbroek [9] exists $\Sigma_\phi^{-1} \approx \gamma \nabla^4$. Hence Σ_v^{-1} can have a sparse representation.

The third element missing is a sparse approximation to R^{-1} the inverse of the estimation error covariance matrix. Under good signal-to-noise conditions, the spectrum of the error propagated though the loop is $\propto f^{-2}$ [10], with f the modulus of the spatial frequency. This means that R is well approximated by a residual spectrum of known power-law. Furthermore, in direct space, $R^{-1} \approx \alpha \nabla^2$. This operator has at most 5 elements per DoF.

In conclusion, $A = \mathcal{A}_d (C_d^T C_d + \sigma^2 \Sigma_\infty^{-1})$ has now a sparse approximation, the validity of which will be evaluated using numerical simulations in sect. 5.

4.4 Conjugate Gradients (CG) and Preconditioned CG

A specially adapted iterative algorithm to the AO problem is the Conjugate-Gradients (CG) [11], with preconditioning in the Fourier space (FD-PCG) [12]. This algorithm involves vector dot products and multiplications by A and the preconditioner. The multiplication of vector x by \mathcal{H}_∞ is now determined from the solution x of the linear system $Ax = b$, followed by multiplication by \mathcal{A}_d . The CG algorithm involves products of the form

$$Ax = (C_d^T C_d + \sigma^2 \Sigma_\infty^{-1})x = \begin{pmatrix} 0 & 0 \\ 0 & \mathbf{D}^T \mathbf{D} \end{pmatrix} x + \sigma^2 \begin{pmatrix} \Sigma_v^{-1} & -\Sigma_v^{-1} \mathcal{A}_{tur} \\ -\mathcal{A}_{tur}^T \Sigma_v^{-1} & R^{-1} + \mathcal{A}_{tur}^T \Sigma_v^{-1} \mathcal{A}_{tur} \end{pmatrix} x. \quad (19)$$

Using the approximations $\Sigma_\phi \propto \nabla^4$, and $R^{-1} \approx \alpha \nabla^2$ in Eq. (15), it comes

$$\begin{pmatrix} \Sigma_v^{-1} & -\Sigma_v^{-1} \mathcal{A}_{tur} \\ -\mathcal{A}_{tur}^T \Sigma_v^{-1} & R^{-1} + \mathcal{A}_{tur}^T \Sigma_v^{-1} \mathcal{A}_{tur} \end{pmatrix} = \begin{pmatrix} \frac{1}{1-a^2} \gamma \nabla^4 & -\gamma \frac{a}{1-a^2} \nabla^4 \\ -\frac{a}{1-a^2} \gamma \nabla^4 & \alpha \nabla^2 + \frac{a^2}{1-a^2} \gamma \nabla^4 \end{pmatrix}. \quad (20)$$

Preconditioning an iterative method consists of finding a sparse matrix M^{-1} such that $M^{-1}Ax = M^{-1}b$ has a better eigenvalue clustering, on which convergence is highly dependent [11]. Of course, if M^{-1} is the exact inverse of A then the solution is found in one iteration. Though, M^{-1} would be in general full, thus not granting any computational gains. It turns out that R and R^{-1} being approximated by a residual spatial spectrum, the system can be straightforwardly preconditioned in the Fourier domain, where a diagonal representation is achievable. This being the case, multiplication by the preconditioner in Fourier space is fast since only n multiplications are performed. The fundamentals of Fourier preconditioning in AO are detailed in [12].

5 Proof-of-concept through Monte Carlo numerical simulations

To validate the proposals detailed so far, a Monte Carlo simulation was run with high-fidelity phase, WFS and DM models on a medium-size AO system with 20 sub-apertures across a 8m telescope. Zernike-modal and zonal basis results are compared to understand pros and cons of each basis, modelled with 400 degrees of freedom to represent turbulent phase. The results obtained with a zonal basis are further compared to the results of using a sparse approximation to Σ_ϕ^{-1} in direct and Fourier domains as explained in sect. 4.2. Results are summarised in Table. 1.

Table 1. Long-exposure Strehl-ratio for the modal and zonal configurations. Simulation conditions: SCAO on a 8m telescope with 20 sub-apertures across. $V/D=1.25$, $L_0=25\text{m}$, seeing = 0.65", RON = $2e^-$, $\Sigma_{\infty,modal}^{-1}$ with 400 modes estimated. 500Hz frame-rate. Zonal resolution of 0.5m^{-1} . FD-PCG uses 8 iterations per sample interval.

Type of Σ_∞ used	Long-exposure Strehl-ratio (%)
Modal with $\Sigma_{\infty,modal}^{-1}$	80.56
Zonal with $\Sigma_{\infty,zonal}^{-1}$	81.55
Zonal with sparse approximate to $\Sigma_{\infty,zonal}^{-1}$	81.53
Zonal with FD-PCG iterative method	81.07

Attained performance is as expected very alike in both the modal and zonal cases. The former suffers from a high conditioning number in the projection step, which justifies the 1% decrease compared to the zonal cases. Furthermore the use of the sparse approximation to Σ_ϕ^{-1} does not entail performance loss, even in the iterative case. The warm-restarted FDPCG-LQG uses 8 iterations per interval, a still relatively high amount that is being currently improved.

6 Summary and outlook

This contribution acts in two fronts: decreasing computational requirements to implement the MV LQG optimal solution both off-line and in real-time.

A migration from modal to zonal description of phase, measurements and DM influence functions is proposed, giving rise to rather sparse linear operators.

The use of the doubling algorithm in a reduced-size system allows for more than two orders of magnitude gains when compared to the standard fixed-point iteration to solve for the estimation Riccati equation solution. Memory requirements (not quantified) are due to decrease, since matrix inversion is of half-sized matrices.

In real-time, a novel approach using iterative methods to solve for the product of a vector by the Kalman gain have been proposed. It requires a sparse approximation to the estimation error covariance matrix. In doing so, the ARE is no longer computed. It has been adopted the approximation by a spectral operator $\propto f^{-2}$, which suggests that multiplication by both direct and inverse linear operators can be accomplished in Fourier space using fast iterative methods. Using the FD-PCG algorithm, the LQG computational burden is therefore decreased from $O(n^2)$ to $O(n \log(n))$ of the Fourier transforms.

Generalisation to wide-field AO systems is ongoing, by taking advantage of the fact that Fourier modes are almost independent between turbulent layers. Similar ideas could be applied to closed-loop MCAO at a later step.

Faster convergence is a topic of the authors' current research as well as quantifying the suitability of the approximation used for phase prediction. A fair comparison with the pseudo-open-loop [13] and the FrIM+IMC [14] is being currently pursued. It will allow to infer which of these strategies is wiser: start from a computationally friendly MV method and adapt it to AO closed-loop operation or start from the optimal MV control solution and make it computationally friendly.

References

1. B. Le Roux, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, L. M. Mugnier, and T. Fusco, "Optimal control law for classical and multiconjugate adaptive optics," *JOSA A* **21**, 1261–1276 (2004).
2. C. Kulcsár, H.-F. Raynaud, C. Petit, J.-M. Conan, and P. V. de Lesegno, "Optimal control, observers and integrators in adaptive optics," *Opt. Express* **14**, 7464–7476 (2006).
3. C. Petit, "Etude de la commande optimale en OA et OAMC, validation numérique et expérimentale," Ph.D. thesis, E.D. Galilée, Univ. Paris XIII (2006).
4. C. Petit, J.-M. Conan, C. Kulcsár, and H.-F. Raynaud, "Linear quadratic gaussian control for AO and MCAO: experimental and numerical analysis," *JOSA A* **26**, 1307–1325 (2009).
5. A. Costille, C. Petit, J.-M. Conan, C. Kulcsár, H.-F. Raynaud, and T. Fusco, "Wide field AO laboratory demonstration with closed loop tomographic control," submitted to *JOSA A* (2009).
6. C. Correia, H.-F. Raynaud, C. Kulcsár, and J.-M. Conan, "On the optimal reconstruction and control in adaptive optics with mirror dynamics," submitted to *JOSA A* (2009).
7. B. D. O. Anderson and J. B. Moore, *Optimal Filtering* (Dover Publications Inc., 1995).
8. P. Piatrou and M. C. Roggemann, "Performance study of kalman filter controller for multiconjugate adaptive optics," *Appl. Opt.* **46**, 1446–1455 (2007).
9. B. L. Ellerbroek, "Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques," *JOSA A* **19**, 1803–1816 (2002).
10. F. Roddier, *Adaptive Optics in Astronomy* (Cambridge University Press, New York, 1999).
11. Y. SAAD, *Iterative Methods for sparse linear systems* (SIAM, Boston, MA, 1996), 2nd ed.
12. Q. Yang, C. Vogel, and B. Ellerbroek, "Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography," *JOSA A* (2005).
13. L. Gilles, "Closed-loop stability and performance analysis of least-squares and minimum-variance control algorithms for multiconjugate adaptive optics," *Appl. Optics* **44**, 993–1002 (2005).
14. C. Béchet, M. Tallon, and É. Thiébaud, "Closed-loop ao performance with FrIM," in "Adaptive Optics Topical Meetings" (Optical Society of America, 2007), p. JTUA4.