Johann Radon Institute for
Computational and Applied Mathematics
Austrian Academy of Sciences (ÖAW)

RICAM
JOHANN · RADON · INSTITUTE
FOR COMPUTATIONAL AND APPLIED MATHEMATICS

A frequency dependent preconditioned wavelet method for atmospheric tomography

M. Yudytskiy, T. Helin, R. Ramlau

A FREQUENCY DEPENDENT PRECONDITIONED WAVELET METHOD FOR ATMOSPHERIC TOMOGRAPHY

Mykhaylo Yudytskiy[1a], Tapio Helin[2b], and Ronny Ramlau[3c]

[1] Johann Radon Institute for Computational and Applied Mathematics (RICAM), Altenberger-strasse 69, 4040 Linz, Austria
[2] University of Helsinki, Gustaf Hallströmin katu 2b, FI-00014 Helsinki, Finland
[3] Johannes Kepler University Linz, Altenbergerstrasse 69, 4040 Linz, Austria

Abstract. Atmospheric tomography, i.e. the reconstruction of the turbulence in the atmosphere, is a main task for the adaptive optics systems of the next generation telescopes. For extremely large telescopes, such as the European Extremely Large Telescope, this problem becomes overly complex and an efficient algorithm is needed to reduce numerical costs. Recently, a conjugate gradient method based on wavelet parametrization of turbulence layers was introduced [5]. An iterative algorithm can only be numerically efficient when the number of iterations required for a sufficient reconstruction is low. A way to achieve this is to design an efficient preconditioner. In this paper we propose a new frequency–dependent preconditioner for the wavelet method. In the context of a multi conjugate adaptive optics (MCAO) system simulated on the official end-to-end simulation tool OCTOPUS of the European Southern Observatory we demonstrate robustness and speed of the preconditioned algorithm. We show that three iterations are sufficient for a good reconstruction.

1 Introduction

One of the fundamental challenges of an adaptive optics (AO) system is to be able to determine the mirror corrections in a relatively short timeframe from the incoming data. The typical strategy is to use a matrix–vector multiply (MVM) method, in which the inverse of a regularized system matrix is multiplied with the data vector. With the development of next generation AO systems, the dimension of the data increases. The downside of the MVM method is its computational complexity, which increases quadratically with the amount of data and leads to a high computational cost. Similarly, the complexity of system updates increases rapidly.

This especially plays a role for AO systems that utilize several wavefront sensors (WFSs) with the goal of obtaining a good correction within the field of view. These systems, such as the MCAO or laser tomography adaptive optics (LTAO), rely on an accurate reconstruction of the turbulence layers in order to correct for the atmospheric aberrations. For the proposed MCAO system in the European Extremely Large Telescope (E-ELT), the amount of measured data ranges between 10^5 and 10^6, whereas the reconstruction must be performed under 2 ms.

Recently, several iterative methods have been proposed with the goal of reducing the computational complexity. Many of these methods, such as FD-PCG [17, 18] and FrIM [16, 15], are based on solving the minimum variance estimation problem [4] using the conjugate gradient (CG) algorithm (see e.g., [13]). Others utilize a split approach by coupling a fast wavefront

---

[a] mykhaylo.yudytskiy@ricam.oeaw.ac.at
[b] tapio.helin@helsinki.fi
[c] ronny.ramlau@jku.at

reconstructor, such as CuReD [12], with a Kaczmarz iteration [10]. Our paper aims at a contribution to the development of the CG–based wavelet reconstructor, which has been recently proposed in [5].

The computational cost of a CG–based algorithm is related to the sparsity of the operators in the normal equation as well as the convergence speed. Although dependent on the discretization, the system matrix in AO is often very sparse. However, in the minimum variance approach a regularization matrix is typically included in the normal equation. In statistical modeling this corresponds to the inverse prior covariance. Such a matrix is typically full, with the exception of the Fourier basis, and thus sparse approximation schemes are needed. In [5] it was shown that in the wavelet basis the inverse covariance can be approximated by a diagonal matrix.

This work is focused towards the convergence speed of the wavelet method. Due to the time restrictions, a successful iterative method needs to converge within a limited number of iterations. A commonly used technique to accelerate the convergence is to precondition the system. A properly tailored preconditioner can significantly reduce the number of iterations. Each of the CG–based methods mentioned above uses a distinct preconditioner.

In this paper we propose a new efficient preconditioner for the wavelet method based on heuristic considerations about the temporal evolution of the atmosphere. We modify a Jacobi preconditioner to obtain different preconditioning to the low and high frequency regimes. Apart from reducing the number of iterations, the benefit of using such an approach is an increased robustness and stability of the overall method. Moreover, for the wavelet method our preconditioner can be formulated as a diagonal matrix, which is very efficient regarding the computational cost.

The paper is structured as follows. In Section 2 we summarize the wavelet method. The new preconditioning scheme is proposed and discussed in Section 3. We present numerical convergence results for an E-ELT MCAO configuration in Section 4 using the European Southern Observatory (ESO) end–to–end simulation tool OCTOPUS [7]. Finally, in Section 5 we state our conclusions and future work.

2 The Wavelet Method

Let us give a short overview of the wavelet method for atmospheric tomography [5]. The method is motivated by a wavelet parametrization of the turbulence layers. We use the Daubechies 3 wavelets, which are a well–known orthogonal wavelet basis with compact support. In addition, the Daubechies wavelets have useful decay properties in the frequency domain, which allows a completely diagonal representation of the inverse covariance. For an extensive introduction to wavelet basics we refer to [2].

We assume that the AO system utilizes M guide stars, of which M_{lgs} are laser guide stars (LGS) and the rest are natural guide stars (NGS). Further, we assume a layered atmospheric turbulence model [11], where $\boldsymbol{\phi} = (\phi_1, \ldots, \phi_L)$ is the vector of L turbulence layers. The subproblems for different guide star directions are given by

$$\mathbf{s}_m = \Gamma_m P_m^{\text{LGS}} \boldsymbol{\phi} \quad \text{and} \quad \mathbf{s}_{m'} = \Gamma_{m'} P_{m'}^{\text{NGS}} \boldsymbol{\phi} \tag{1}$$

for $1 \le m \le M_{\text{lgs}}$ and $M_{\text{lgs}} < m' \le M$. Here, P_m^{LGS} and $P_{m'}^{\text{NGS}}$ are geometric propagation operators in directions of the guide stars associated to an LGS with cone effect and an NGS; Γ_m is the Shack–Hartmann operator (for more details, see [5]). Moreover, we write

$$\mathbf{s} = (\mathbf{s}_m)_{m=1}^M = \mathbf{A}\boldsymbol{\phi}, \tag{2}$$

2

where **A** is the concatenation of operators $\Gamma_m P_m^{\text{LGS}}$ and $\Gamma_{m'} P_{m'}^{\text{NGS}}$. Estimating $\boldsymbol{\phi}$ in (2) from a given **s** is called the atmospheric tomography problem.

The atmospheric tomography problem is ill–posed and regularization is necessary to obtain a stable solution from noisy measurements. To this end, the problem can be formulated in Bayesian framework (see e.g., [6]). In this setting, the regularized solution is the maximum a–posteriori (MAP) estimate, which is obtained by solving the minimization problem

$$\min_{\boldsymbol{\phi}} \left(\|\mathbf{C}_\Phi^{-1/2}\boldsymbol{\phi}\|_2^2 + \|\mathbf{C}_\mathcal{E}^{-1/2}(\mathbf{s} - \mathbf{A}\boldsymbol{\phi})\|_2^2 \right). \tag{3}$$

Here, $\mathbf{C}_\Phi$ and $\mathbf{C}_\mathcal{E}$ are the turbulence and measurement noise covariance operators.

The turbulence covariance operator $\mathbf{C}_\Phi = \text{diag}(C_1, \ldots, C_L)$ has a block–diagonal structure, where C_ℓ is the covariance operator of layer ℓ following the Kolmogorov power law. The operator has a very efficient approximation in the wavelet domain. A covariance operator C_ℓ acting on a turbulence layer translates to weights applied to wavelet coefficients in the wavelet domain.

Let D_ℓ be a diagonal operator with entries $(D_\ell)_{\lambda\lambda} = c_\rho(\ell)^{-1} 2^{\frac{11}{3}j}$, where j is the scale of the wavelet with global index λ and $c_\rho(\ell)$ describes the measure of the optical turbulence strength at layer ℓ. Then, under certain assumptions the following approximation holds [5],

$$\|\mathbf{C}_\Phi^{-1/2}\boldsymbol{\phi}\|_2^2 = \sum_{\ell=1}^L \|C_\ell^{-1/2}\phi_\ell\|_2^2 \simeq \sum_{\ell=1}^L (D_\ell \mathbf{c}_\ell, \mathbf{c}_\ell)_2 = (\mathbf{D}\mathbf{c}, \mathbf{c})_2, \tag{4}$$

where $\mathbf{D} = \text{diag}(D_1, \ldots, D_L)$ is the block–diagonal operator of diagonal operators D_ℓ and $\mathbf{c} = (\mathbf{c}_1, \ldots \mathbf{c}_L)$ is the vector of wavelet coefficients corresponding to all layers.

The noise covariance operator $\mathbf{C}_\mathcal{E} = \text{diag}\left(\widehat{C}_1, \ldots, \widehat{C}_{M_{\text{lgs}}}, \widehat{C}_{M_{\text{lgs}}+1}, \ldots, \widehat{C}_M\right)$ also has a block–diagonal structure, where the first M_{lgs} blocks correspond to the LGS sensors and the remaining blocks to the NGS sensors. We assume the 2×2 block diagonal representation for the LGS operators $\widehat{C}_m$, $1 \le m \le M_{\text{lgs}}$ corresponding to the elongated measurement noise, described in [1]. The NGS operators $\widehat{C}_{m'}$, $M_{\text{lgs}} < m' \le M$ are modeled by the identity operator scaled with the noise variance.

Laser guide star measurements contain an erroneous tip–tilt component. In the wavelet approach, the tip–tilt component in the LGS measurements is removed by modifying (1) to $(I - T)\mathbf{s}_m = (I - T)\Gamma_m P_m^{LGS}\boldsymbol{\phi}$ for $1 \le m \le M_{\text{LGS}}$. Here, T is the orthogonal projection operator onto the tip–tilt measurements space. For the regularized problem (3), such noise modeling can be interpreted as using an updated inverse measurement noise covariance matrix

$$\mathbf{C}_\mathcal{E}^{-1} = \text{diag}\left((I - T)\widehat{C}_1^{-1}(I - T), \ldots, (I - T)\widehat{C}_{M_{\text{lgs}}}^{-1}(I - T), \widehat{C}_{M_{\text{lgs}}+1}^{-1}, \ldots, \widehat{C}_M^{-1}\right) \tag{5}$$

in combination with the original equations in (1).

The minimization problem (3) formulated in the wavelet domain thus reduces to solving a system of linear equations

$$(\widetilde{\mathbf{A}}^T \mathbf{C}_\mathcal{E}^{-1} \widetilde{\mathbf{A}} + \alpha\mathbf{D})\mathbf{c} = \widetilde{\mathbf{A}}^T \mathbf{C}_\mathcal{E}^{-1}\mathbf{s}. \tag{6}$$

Here, $\widetilde{\mathbf{A}}$ is the discretization of (2) in the Daubechies 3 wavelet basis, $\mathbf{C}_\mathcal{E}^{-1}$ is the inverse noise covariance (5) and $\mathbf{D}$ is the diagonal operator from the approximation (4). As we only have an approximation in (4) instead of equality, a scalar parameter α is introduced for fine–tuning the balance between the fitting and the regularizing terms. The effect of the parameter has already been studied in [5].

3

Equation (6) is formulated for open–loop measurement data. Hence, to apply the algorithm in closed loop, we combine the mirror shapes with the residual measurements in the pseudo–open loop control (POLC) [3]. We use a modified POLC with an integrator gain, see e.g., [9]. Within the POLC, an approximate solution to (6) is found using either the CG algorithm or the preconditioned conjugate gradient (PCG) algorithm (see e.g., [13]). The CG and PCG iterations are warm started, i.e. the solution from previous iteration is used as an initial guess.

Depending on the AO system, the deformable mirror (DM) shapes are chosen with respect to the MAP estimate of the atmosphere using some optimality criteria. For an MCAO system the desired shape of the DMs is obtained by minimizing a least squares functional [4]. This problem, known as the fitting step, is well–posed. Hence, no regularization techniques are needed to solve the corresponding normal equation.

3 Preconditioning

The convergence of the CG algorithm is controlled by two factors: the eigenvalue distribution of the operator and the decomposition of the residual vector in terms of the eigenvectors, see e.g., [14]. In this work favorably structured residual vectors are obtained by using warm restart. The standard approach to improving the eigenvalue distribution of the operator is a technique called preconditioning. Below we discuss a modified Jacobi preconditioning scheme used in the algorithm and refer to [13] for a general introduction to the topic.

Consider a linear system of equations $\mathbf{Mc} = \mathbf{b}$. The left–preconditioned problem is $\mathbf{J}^{-1}\mathbf{Mc} = \mathbf{J}^{-1}\mathbf{b}$, where $\mathbf{J}$ is an operator that approximates $\mathbf{M}$ but is easy to invert. The Jacobi preconditioner, defined by the diagonal matrix $\mathbf{J} = \text{diag}(\mathbf{M})$ is one of the basic choices for a preconditioner. Although not necessarily spectrally equivalent to $\mathbf{M}$, the Jacobi preconditioner has the effect of centering all the Gerschgorin discs, in which the eigenvalues are contained, around the value 1. Being a diagonal matrix, $\mathbf{J}$ is easy to invert.

The diagonal Jacobi preconditioner to equation (6) has the form $\mathbf{J} = \text{diag}(\widetilde{\mathbf{A}}^T\mathbf{C}_\varepsilon^{-1}\widetilde{\mathbf{A}}) + \alpha\mathbf{D}$. We propose a frequency–dependent preconditioner which is obtained by modifying $\mathbf{J}$ according to

$$\mathbf{J}_\tau = \text{diag}(\widetilde{\mathbf{A}}^T\mathbf{C}_\varepsilon^{-1}\widetilde{\mathbf{A}}) + \alpha\max(\mathbf{D}, \tau\mathbf{I}), \tag{7}$$

where τ is a non–negative scalar parameter. Above, the maximum is taken component–wise.

Low and high wavelet scales correspond to low and high frequency regimes of the layers, respectively. Since the values of $\mathbf{D}$ increase fast with respect to the scales, by increasing the parameter τ we are able to affect the level of damping for the lower frequencies in the preconditioner. In practice, the standard Jacobi preconditioner will damp the high scales too much with respect to the low scales. The preconditioner above introduces a way to balance the standard Jacobi preconditioner with the non–preconditioned CG in a scale–dependent way. Note that for $\tau = 0$ one arrives at the standard Jacobi preconditioner, whereas for a very large choice of $\tau >> 1$, the identity–term dominates (7) and the preconditioner is approximately $\tau\mathbf{I}$. A scaled identity preconditioner has the same effect as the identity preconditioner, hence, the algorithm reduces to the standard CG in this case. We demonstrate this property numerically in Section 4.2.

For convenience we control $\tau = \alpha_J\min(\text{diag}\mathbf{D})$ by scaling the smallest diagonal element of $\mathbf{D}$ with a thresholding parameter $\alpha_J \geq 0$. We plot diagonals of the block corresponding to the highest altitude layer L of $\mathbf{J}_\tau$ for several choices of α_J in Figure 1.
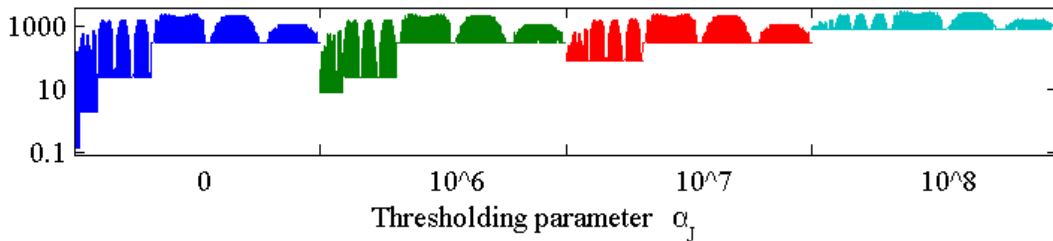
4

Fig. 1. The effect of thresholding to the preconditioner is illustrated at layer *L*. The values on the diagonal of $\mathbf{J}_\tau$ are plotted for four thresholding values α_J. The values are plotted from left to right starting from the coarsest scale. Each color corresponds to different value of α_J. Regular Jacobi is obtained with $\alpha_J = 0$.

4 Numerical Results

For the simulations, we use the proposed MCAO configuration for the E-ELT. The telescope diameter is 42 m, of which roughly 28 percent are obstructed. There are six LGSs positioned in a circle with a diameter of 2 arcmins with 84×84 Shack–Hartmann wavefront sensors (SH-WFS). The LGSs suffer from tip–tilt uncertainty and spot elongation. Moreover, there are three NGSs positioned in a circle with a diameter of 8/3 arcmins with low resolution SH-WFS (one with 2×2 and two with 1×1 subapertures) for tip–tilt correction. LGS flux varies between 50 and 500 photons per subaperture and frame; NGS flux is simulated with 500 photons. LGS and NGS sensors suffer from 3 and 5 electrons per pixel readout noise respectively. In the proposed MCAO configuration three DMs are conjugated to altitudes 0 km, 4 km and 12.7 km.

We simulate one second of the nine layer ESO median seeing atmosphere on OCTOPUS [7], i.e. 500 steps in closed loop. A two–step delay is observed. As the quality evaluation criteria, we use the long exposure (LE) Strehl ratio in K band (for a wavelength of 2200 nm) averaged over a grid of 25 probe star directions within the field of view (2 arcmin square).

Our reconstruction method is described in Section 2 and we use the frequency–dependent preconditioner proposed in Section 3 when referring to PCG. We refer to the method as CG when no preconditioner is used. The standard parameter settings in all test cases are: reconstruction of nine layers, regularization parameter $\alpha = 4$ in equation (6), gain of 0.4 and four CG iterations to solve the least squares fitting equation. We compare our method to the MVM (see e.g., [8]), in which the MAP estimate is defined on three layers of the atmosphere. MVM reference results are provided by the ESO.

In the following we investigate the convergence property of the preconditioned algorithm and the influence of the thresholding parameter α_J. Moreover, we demonstrate stability of the preconditioned algorithm with respect to measurement noise.

4.1 Convergence of the frequency–dependent preconditioned wavelet method

We investigate the number of iterations needed to obtain a sufficient reconstruction quality in Figure 2. The simulation is performed for 100 photons per subaperture and frame of the LGS and we use $\alpha_J = 10^7$ for the PCG. Here we compare the field average LE Strehl of the CG and the PCG reconstructors configured for a different number of iterations. Additionally, the MVM result is plotted.

It can be observed that the PCG with 3 iterations attains a sufficient quality similar to CG with 10 iterations. PCG with more iterations delivers only a small improvement, and the best value is achieved with 6 iterations. Quality of CG rises gradually with the number of iterations.
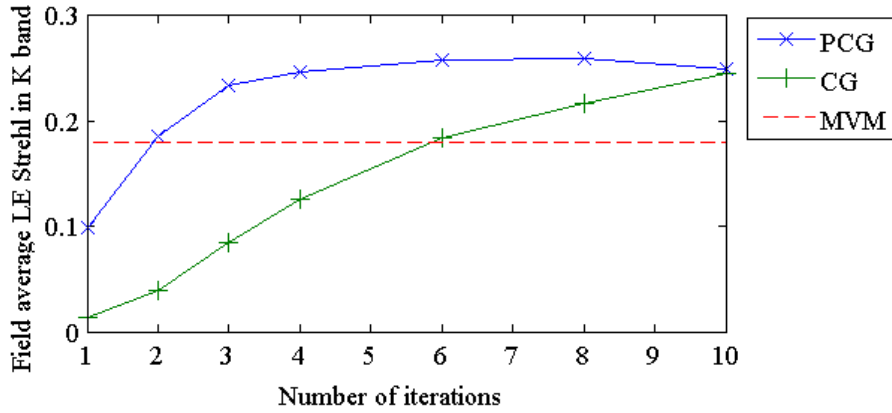
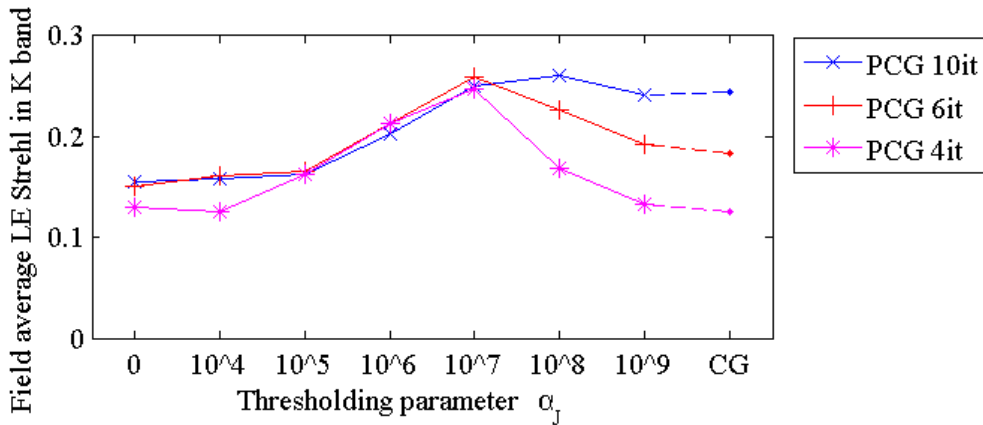**Fig. 2.** Performance of CG and PCG with different number of iterations.

Fig. 3. Sensitivity of PCG with respect to thresholding parameter α_J. Standard CG results with the same number of iterations are displayed at the far right side of each graph.

4.2 Influence of the thresholding parameter

The thresholding parameter α_J has a very strong influence on the quality of the frequency–dependent preconditioner. This result can be observed in Figure 3, where we plot the field average LE Strehl for PCG with four, six and ten iterations and a varying parameter α_J.

The performance of the standard Jacobi preconditioner ($\alpha_J = 0$) is rather unstable. Quality begins to improve when a larger thresholding parameter is used, with an optimal value at 10^7 for PCG with four and six iterations, and at 10^8 for PCG with ten iterations. The structure of the preconditioner used to obtain these results can be observed in Figure 1, where we plot the values of $\mathbf{J}_\tau$ on the highest layer L for $\alpha_J = 0, 10^6, 10^7$ and 10^8. Notice that for $\alpha_J = 10^7$ only the finest scale is unaffected. By increasing the parameter α_J beyond its optimal value, we notice that the performance of the algorithm tends towards the non–preconditioned CG result, which is the right–most value in the plot in Figure 3. As has already been observed in Figure 2, higher number of CG iterations leads to a higher result in quality.

4.3 Performance with variable flux

In this example we consider the performance of our methods with respect to varying noise level in the LGS measurements. In Figure 4 we plot the field average LE Strehl obtained for the
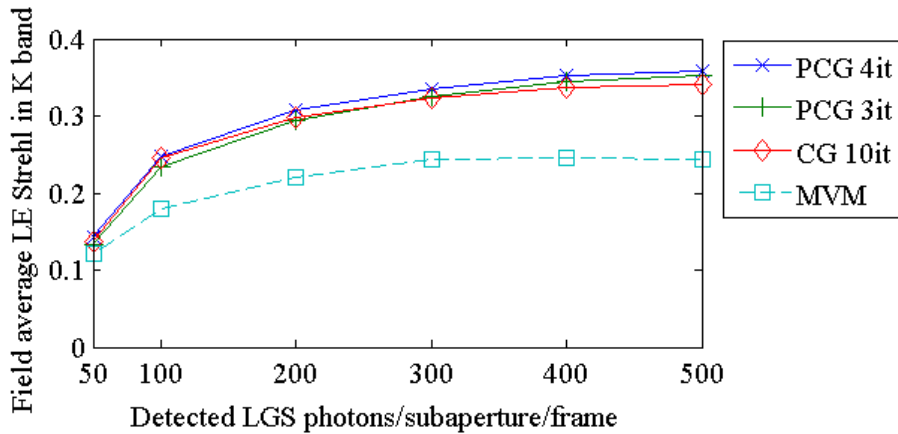
6

Fig. 4. Field average LE Strehl vs. detected number of photons per subaperture and frame of LGS sensors.

following methods: CG with ten iterations, PCG with three and four iterations and the MVM. The LGS flux varies between 50 and 500 photons. We use the thresholding parameter $\alpha_J = 10^7$, which was found to be optimal in Section 4.2, for all test cases.

We observe stable performance of PCG with respect to the flux, both with three and four iterations. At higher flux levels the frequency–dependent preconditioned method outperforms CG, as it seems the CG needs more iterations. Recall that the MVM reconstructs only three layers; thus it is expected to perform worse.

5 Conclusions and Outlook

In this paper we introduced a new frequency–dependent preconditioner for the wavelet method. Using numerical simulations we found an optimal parameter for this preconditioning scheme. Furthermore, in numerical simulations we demonstrated that the wavelet method with this preconditioner needs three PCG iterations, independent of the flux.

As future work, a more sophisticated scheme for controlling the scales of the preconditioner will be investigated. For instance, a more accurate choice for the cut–off parameter τ in (7) might improve the quality and give more insight into the way the frequency–dependent preconditioning works.

Acknowledgements

References

1. C. Béchet, M. Le Louarn, R. Clare, M. Tallon, I. Tallon-Bosc, and É. Thiébaut. Closed-loop ground layer adaptive optics simulations with elongated spots: impact of modeling noise correlations. In *AO4ELT Proceedings*, page 03004, 2010.

2. I. Daubechies. *Ten Lectures on Wavelets*, volume 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*. Soc. for Industrial and Applied Mathematics, Philadelphia, Pa., 1992.

3. B. Ellerbroek and C. R. Vogel. Simulations of closed-loop wavefront reconstruction for multiconjugate adaptive optics on giant telescopes. *Proc. SPIE*, 5169:206–217, 2003.

4. T. Fusco, J. M. Conan, G. Rousset, L. M. Mugnier, and V. Michau. Optimal wave-front reconstruction strategies for multiconjugate adaptive optics. *J. Opt. Soc. Am.*, 18(10):2527, 2001.

5. T. Helin and M. Yudytskiy. Wavelet methods in multi-conjugate adaptive optics. *Inverse Problems*, 29(8):085003, 2013.

6. J. Kaipio and E. Somersalo. *Statistical and Computational Inverse Problems*, volume 160 of *Applied Mathematical Sciences*. Springer Science+Business Media, Inc, 2005.

7. M. Le Louarn, C. Vérinaud, V. Korkiakoski, N. Hubin, and E. Marchetti. Adaptive optics simulations for the European Extremely Large Telescope. In *Proc. SPIE 6272, Advances in Adaptive Optics II*, page 627234, 2006.

8. E. Marchetti, N. Hubin, E. Fedrigo, et al. MAD the ESO multi-conjugate adaptive optics demonstrator. In *Proc. SPIE 4839, Adaptive Optical System Technologies II*, pages 317–328, 2003.

9. P. Piatrou and L. Gilles. Robustness study of the pseudo open-loop controller for multiconjugate adaptive optics. *Appl. Opt.*, 44(6):1003–1010, Feb 2005.

10. R. Ramlau and M. Rosensteiner. An efficient solution to the atmospheric turbulence tomography problem using Kaczmarz iteration. *Inverse Problems*, 28(9):095004, 2012.

11. M. Roggeman and B. Welsh. *Imaging Through Turbulence*. Cambridge University Press, Cambridge, 1996.

12. M. Rosensteiner. Wavefront reconstruction for extremely large telescopes via CuRe with domain decomposition. *J. Opt. Soc. Am. A*, 29(11):2328–2336, Nov 2012.

13. Y. Saad. *Iterative Methods for Sparse Linear Systems*. SIAM, second edition, April 2003.

14. Zdeněk Strakoš and Petr Tichý. On error estimation in the conjugate gradient method and why it works in finite precision computations. *Electron. Trans. Numer. Anal.*, 13:56–80 (electronic), 2002.

15. M. Tallon, I. Tallon-Bosc, C. Béchet, F. Momey, M. Fradin, and É. Thiébaut. Fractal iterative method for fast atmospheric tomography on extremely large telescopes. In *Proc. SPIE 7736, Adaptive Optics Systems II*, pages 77360X–77360X–10, 2010.

16. E. Thiébaut and M. Tallon. Fast minimum variance wavefront reconstruction for extremely large telescopes. *J. Opt. Soc. Am.*, 27(5):1046, 2010.

17. C. R. Vogel and Q. Yang. Fast optimal wavefront reconstruction for multi-conjugate adaptive optics using the Fourier domain preconditioned conjugate gradient algorithm. *Optics Express*, 14(17):7487–7498, 2006.

18. Q. Yang, C. R. Vogel, and B. Ellerbroek. Fourier domain preconditioned conjugate gradient algorithm for atmospheric tomography. *Appl. Opt.*, 45(21):5281, 2006.