ISSN 1561-2430 (Print)
ISSN 2524-2415 (Online)
UDC 535.317.612
https://doi.org/10.29235/1561-2430-2026-62-1-59-72

Received 18.08.2025
Поступила в редакцию 18.08.2025

Piotr K. Piatrou

B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus, Minsk, Republic of Belarus

STRUCTURAL APPROACH TO THE ADAPTIVE OPTICS TOMOGRAPHY PROBLEM: THE VOLUME TURBULENCE RECONSTRUCTION

Abstract. Adaptive Optics as a relatively new trend in the modern optical systems allows correcting the perturbations related to light propagation in the turbulent atmosphere obtaining the diffraction-limited images. Because it requires computer processing of large information flow in real time, a key problem lays in finding the mathematical algorithms with reduced computational complexity. The paper proposes the approach, which takes into account structural properties of both atmospheric turbulence and the most active optical system. The possibility was demonstrated of one-dimensional accounting of sparse and convolution-like (Toeplitz) structure of adaptive correction in one algorithm. The analysis is made of Multi-Conjucate Adaptive Optics and the comparison was conducted of the nature complexity of the algorithm induced by this strucute with the complexity of sparse algorithm. It was demonstrated how the structural approach can be applied to the quite laborious tomographic turbulence volume reconstruction phase with obtaining due to this a significant computational speed-up.

Keywords: Multi-Conjugate Adaptive Optics, tomographic phase reconstruction, structured matrices, real-time signal processing

For citation. Piatrou P. K. Structural approach to the Adaptive Optics tomography problem: the volume turbulence reconstruction. *Vestsi Natsyyanal'nai akademii navuk Belarusi. Seryya fizika-matematychnykh navuk = Proceedings of the National Academy of Sciences of Belarus. Physics and Mathematics series*, 2026, vol. 62, no. 1, pp. 59–72. https://doi.org/10.29235/1561-2430-2026-62-1-59-72

П. К. Петров

Институт физики имени Б. И. Степанова Национальной академии наук Беларуси, Минск, Республика Беларусь

СТРУКТУРНЫЙ ПОДХОД К РЕШЕНИЮ ТОМОГРАФИЧЕСКОЙ ЗАДАЧИ АДАПТИВНОЙ ОПТИКИ: РЕКОНСТРУКЦИЯ ТУРБУЛЕНТНОСТИ В ОБЪЕМЕ

Аннотация. Адаптивная оптика как относительно новое направление развития современных оптических систем позволяет корректировать искажения, связанные с распространением света в турбулентной атмосфере и получать изображения дифракционного качества. Поскольку это требует компьютерной обработки больших потоков информации в реальном времени, ключевой проблемой является снижение вычислительной сложности соответствующих математических алгоритмов. В статье предложен подход, учитывающий структурные свойства как атмосферной турбулентности, так и самой активной оптической системы. Продемонстрирована возможность одновременного учета разреженной и сверточной (теплицевой) структуры адаптивной коррекции в одном алгоритме. Дан анализ многосопряженной адаптивной оптической системы и сравнение естественной сложности алгоритма, индуцированного его структурой, со сложностью разреженного алгоритма. Показано применение структурного подхода к достаточно трудоемкому алгоритму томографической объемной реконструкции турбулентной фазы с получением благодаря этому существенного ускорения вычислений.

Ключевые слова: многосопряженная адаптивная оптика, томографическое фазовое восстановление, структурированные матрицы, обработка сигналов в реальном времени

Для цитирования. Петров, П. К. Структурный подход к решению томографической задачи адаптивной оптики: реконструкция турбулентности в объеме / П. К. Петров // Весці Нацыянальнай акадэміі навук Беларусі. Серыя фізіка-матэматычных навук. – 2026. – Т. 62, № 1. – С. 59–72. https://doi.org/10.29235/1561-2430-2026-62-1-59-72

Introduction. Adaptive Optics (AO) being quite a powerful technology allows overcoming the *dynamic perturbations* caused by the *atmospheric turbulence* and obtaining diffraction quality images suffers from a serious problem often called in the signal processing literature the *curse of dimentionality* [1]: *computational complexity grows too fast with the controlled system scale.* The *system scale* is usually understood as a number of the system's *degrees of freedom* (DoF), i. e. the free parameters under optimization. For Adaptive Optics systems, which represent a classical example of the automatic control systems with *Multiple Inputs and Multiple Outputs* (MIMO), such a set of the DoF is the number of observation and correction channels, which is normally proportional to the entrance aperture area, in other words, to the square of the aperture diameter. The problem is also complicated by the fact that, for tracking the changes in atmosphere on the way of light to the optical system with a characteristic time on the order of 0.01 second, computer signal processing in *hard real time* is required, that is, not only fast but also with stable *timing.* Owing to this, a key problem in the adaptive optics system design is to find mathematical algorithms with reduced computational complexity. Because of quite good linearity the control of MIMO-system can be reduced to a sequence of multiplications of the sensor measurements vector by some *control matrix.* Such an approach, however, looks too straightforward because it completely ignores the *internal structure* of both atmospheric turbulence and the active optical system itself.

On the other hand, a few structural properties of AO systems leading to complexity reduction are already known. In a limiting case, taking AO system structural properties into full consideration will enable to see the *natural complexity* of the problem, which is expected to be much less than the complexity of a straightforward solution. The mathematical problem of finding the natural complexity has not been completely solved at the moment, though even partial solutions taking into account the internal structure already lead to significant complexity reduction. We can mention the following most important findings related to the AO system structure.

– That turbulence active compensation can be efficiently performed by *linear* control follows from the fact that the turbulence optical effects are described to a good approximation [2] by the *stationary Gaussian random process.*

– The most significant complexity reduction from $\mathcal{O}(D^4)$ to $\mathcal{O}(D^2)$, where D is the telescope aperture diameter, was achieved when it had been proven [3] that the control matrix is approximately reducible to a *sparse + low rank* form, but it was overlooked that the control matrix is not only sparse but also *Toeplitz.*

– Convolutional structure of the *turbulence-to-image interaction operator* following from the atmospheric turbulence statistical *shift invariance* was demonstrated in [4, 5], which allows to transfer the AO system analysis into Fourier Domain and use the Fast Fourier Transform resulting in computational complexity $\mathcal{O}(D^2 \log D^2)$.

– The natural discretization of the convolutional turbulence-to-image interaction operator by a finite resolution *Wave Front Sensor* leads to the *Toeplitz* structure of the control matrix [6], which not only results in comlexity $\mathcal{O}(D^2 \log D^2)$ but also allows presenting the control matrix in an extremely compact form.

As the findings mentioned above and taken by one at a time already lead to notable complexity reduction, it appears to be overlooked that the AO system structural properties can be used all together inside one algorithm resulting in some minimal achievable complexity what we will call *natural* within the frame of this work. In particular, it will be shown that taking into account both sparse and Toeplitz structure of the control matrix in one algorithm allows to achieve unexpectedly significant computational speed-up, which is demonstrated on the most important and laborious case of an AO system, the *Multi-Conjugate Adaptive Optics* (MCAO). We however warn that the research on finding the *natural complexity* is a work in progress and the results reported here may be improved in the future.

1. MCAO overview. A classical AO system aims to dynamically correct the electromagnetic wave *distortions* caused by atmospheric turbulence on the light path through the Earth atmosphere to an astronomical telescope. It is done by measuring the distorted wavefront phase in the telescope pupil by some or other type of a *Wave Front Sensor* (WFS), processing the WFS data by a *Real Time Computer* (RTC) to generate commands for the phase-correcting *Deformable Mirror* (DM) in order to *flatten* the wavefront in the telescope pupil. The more advanced concept of Multi-Conjugate Adaptive Optics was suggested in [7] to address the main limitation of the classical AO: very narrow correctable *Field of View* (FoV) in case of correction by a single DM located in (conjugated to) the pupil. It can be shown [8] that correction by multiple DMs distributed in several conjugated positions within the atmospheric turbulence volume above the telescope extends the corrected FoV from several arcseconds to a few arcminutes at the price of lower average image quality. The MCAO system optical design together with its brief principle of operation explanation are presented in Fig. 1.

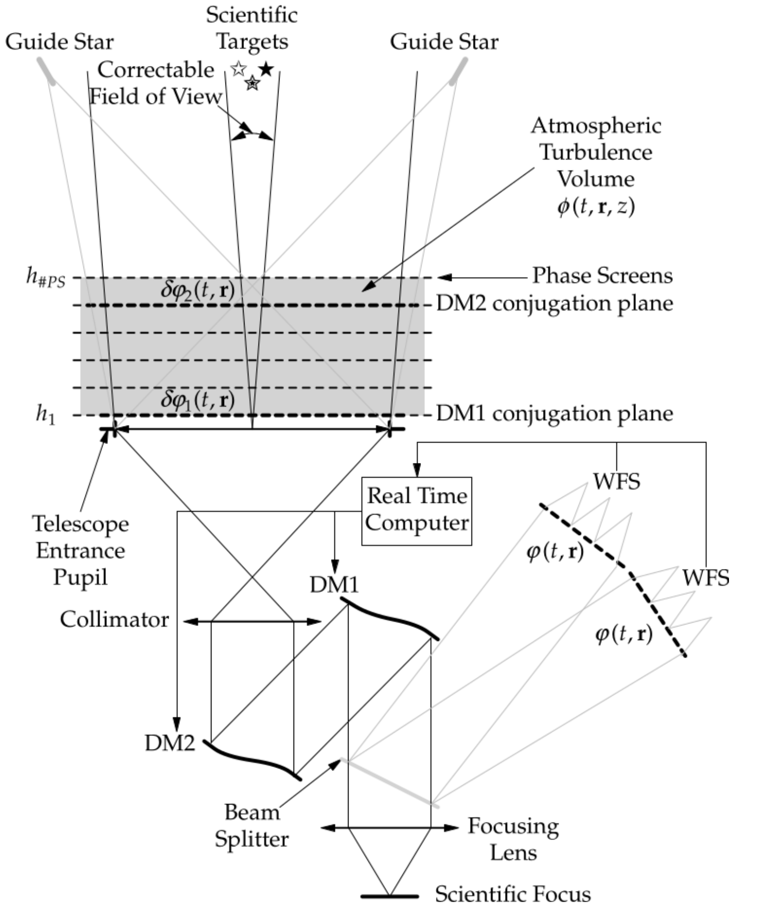


Fig. 1. Optical scheme of a Multi-Conjugate AO system with (at least) two *Deformable Mirrors.* Dynamic 3D *phase distortion* $\phi(t,\mathbf{r},z)$ is probed by passing through the *turbulence volume* the light from multiple light sources, the *Guide Stars* (GS), natural or artificial. The *Wave Font Sensors* measure a set of the 2D *wavefronts* $\varphi_g(t,\mathbf{r})|_{g=1}^{\#Gs}$ accumulated along the paths from multiple Guide Stars to the telescope pupil, which provides information allowing to *tomographically reconstruct* $\phi(t,\mathbf{r},z)$ in the volume of turbulent air above the telescope. Multiple *Deformable Mirrors* optically conjugated over several altitudes within the turbulence volume generate the *phase corrections* $\delta\phi_d(t,\mathbf{r})|_{d=1}^{\#DM}$ in an attempt to compensate the phase distortion not in a plane but in the entire volume, which helps to achieve the turbulence compensation over a broader FoV containing the *Scientific Targets* of interest. Phase distortion reconstruction from the WFS measurements and generation of DM correction commands are performed by the *Real Time Computer.* Light from the Scientific Targets is drawn in black, from Guide Stars in grey. Vector r stands for x,y-coordinates in either phase screen, DM or WFS plane

The most difficult part of the MCAO method is an extremely high computational complexity of the underlying algorithms that need to be run by the RTC. In what follows, mathematical description of the MCAO *tomographic problem* is given.

Assume that an object imaged by an optical system is a point source located at infinity. Then, ideally, the electromagnetic wave or *wavefront* reaching the telescope *entrance pupil* without any distortion induced by some obstacles on the *propagation path,* such as atmospheric turbulence, should be a *plane wave* $E(\mathbf{r}) = A\exp[ik(\mathbf{t}^T\mathbf{r} + p)]$, where A is the constant *amplitude,* $\mathbf{t} \overset{\Delta}{=} (t_x, t_y)$ is the wavefront direction vector or *tilt,* p is the constant phase called *piston.* Note that neither tilt, nor piston induces any image corruption [2]. The plane wave results in the ideal *diffraction-limited* point source image in the telescope *focal plane.* In the case of propagation path distortion, the wavefront entering the telescope aperture is $E(\mathbf{r}) = A(\mathbf{r})\exp[ik\varphi(\mathbf{r})]$, i. e. $A(\mathbf{r}) \neq$ const. (*scintillation*), $\varphi(\mathbf{r}) \neq \mathbf{t}^t\mathbf{r} + p$ (*aberration*). In the case of weak turbulence the scintillation effects are negligible [9], so the main goal of the *adaptive correction* is to *flatten* the wavefront phase, i. e. to minimize $\|\varphi(\mathbf{r})\|$, where $\|\cdot\|$ is some *function norm* defined on the telescope pupil. In the context of the MCAO correction this minimization problem can be written as (we drop the time dependence for clarity)

$$\widehat{\mathcal{R}} = \arg\min_{\mathcal{R}} \left\langle \left\| H_\phi \phi(\mathbf{r}, z) - H_a \mathcal{R}\mathbf{s} \right\|^2 \right\rangle, \tag{1}$$

where $\mathbf{s} = [\mathbf{s}_1 ... \mathbf{s}_{\#GS}]^T$ is the concatenated vector of the WFS measurements from all Guide Stars, $\mathcal{H}_\phi$ the *Propagation Operator* mapping the 3D phase distortion $\phi(\mathbf{r},z)$ to the 2D wavefronts $\boldsymbol{\varphi}(\mathbf{r}) = [\varphi_1(\mathbf{r}) ... \varphi_{\#ST}(\mathbf{r})]^T$ accumulated in the telescope pupil, H_a the propagation operator mapping the DM phase corrections $\delta\boldsymbol{\varphi}(\mathbf{r}) = [\delta\varphi_1(\mathbf{r}) ... \delta\varphi_{\#DM}(\mathbf{r})]^T$ to the pupil, $\mathcal{R}$ is the *Reconstruction Operator* $\mathcal{R}\mathbf{s} = \boldsymbol{\delta\varphi}(\mathbf{r})$. The $\langle \cdot \rangle$ brackets stand for averaging over the turbulence phase statistics as turbulence is a *random process.*

In the *linear approximation,* the minimization problem (1) can be presented in a matrix form. Following [10] the Tomographic Problem can be discretized by approximating all functions involved by certain *basis function sets.* In particular, the turbulence-corrupted phase in the pupil can have a finite-dimensional approximation

$$\varphi(r) \approx \sum_{i=1}^{N} \varphi_i \phi_i(r) \overset{\Delta}{=} \boldsymbol{\phi}^T(\mathbf{r})\boldsymbol{\varphi}, \tag{2}$$

where $\boldsymbol{\phi}(\mathbf{r}) = \phi_i(\mathbf{r})\,|_{i=1}^{N\to\infty}$ is a basis, not necessarily orthogonal, which approximates efficiently enough the turbulence phase in the sense of small *approximation error*

$$\varepsilon^2 = \left\| \varphi(r) - \boldsymbol{\phi}^T(\mathbf{r})\boldsymbol{\varphi} \right\|^2, \tag{3}$$

where the norm $\|f\|^2 = (f, f)$ is defined via a *vector function space scalar product* (a,b), for instance,

$$(a,b) = \int_P a(\mathbf{r}) b(\mathbf{r}) d\mathbf{r} \tag{4}$$

for L^2 space. Note that the very possibility of finite-dimensional approximation (2) follows from the fact that the turbulence phase $\phi(\mathbf{r},z)$ belongs to a *compact subset* of functions in a certain function space, which by itself is an important *structural property.* The approximation weights $\boldsymbol{\varphi} = \varphi_i\,|_{i=1}^{N}$ are found as

$$\boldsymbol{\varphi} = \left(\boldsymbol{\phi}(\mathbf{r}), \boldsymbol{\phi}(\mathbf{r})^T \right)^{-1} \left(\boldsymbol{\phi}(\mathbf{r}), \varphi(\mathbf{r}) \right). \tag{5}$$

This way a continuous function $\varphi(r)$ can be replaced with a vector of its *basis approximation weights* $\boldsymbol{\varphi}$ from (2), (5). Analogously, the 3D *volume phase function* $\phi(\mathbf{r},z)$ can be transformed into a vector form in two steps:

1) $\phi(\mathbf{r},z)$ is replaced with a set of 2D *turbulence layers* (see Fig. 1)

$$\phi(\mathbf{r}, z) \approx \boldsymbol{\phi}(\mathbf{r}) = [\phi(\mathbf{r}, h_1) ... \phi(\mathbf{r}, h_{\#PS})]^T, \quad \phi(\mathbf{r}, h_p) = \int_{h_{p-1}}^{h_p} \phi(\mathbf{r}, z) dz; \tag{6}$$

2) given a basis on each layer, $\boldsymbol{\phi}(\mathbf{r})$ is replaced by its approximation weights $\boldsymbol{\phi} = \left[\boldsymbol{\phi}_1 ... \boldsymbol{\phi}_{\#PS} \right]^T$ such as $\boldsymbol{\phi}(\mathbf{r}) \approx \left[\boldsymbol{\phi}_1^T(\mathbf{r})\phi_1 ... \boldsymbol{\phi}_{\#PS}^T(\mathbf{r})\boldsymbol{\phi}_{\#PS} \right]^T$.

The DM phase correction in linear approximation is expressed the same way with a difference that the arbitrary basis functions for phase screens are replaced with a finite set of *Actuator Influence Functions* $\mathbf{f}(\mathbf{r}) = f_i(\mathbf{r})\,|_{i=1}^{\#A}$ for DMs that are *fixed by design.* Thus, the DM phase correction is

$$\boldsymbol{\delta\varphi}(\mathbf{r}) = \mathbf{f}^T(\mathbf{r})\mathbf{a}, \tag{7}$$

where $\mathbf{a} = a_i\,|_{i=1}^{\#A}$ is a vector of *actuator commands.* Assuming that $\mathbf{a} = [\mathbf{a}_1 ... \mathbf{a}_{\#DM}]^T$ and $\mathbf{f}(\mathbf{r}) = [\mathbf{f}_1(\mathbf{r}) ... \mathbf{f}_{\#DM}(\mathbf{r})]^T$ are concatenations of commands and influence functions from all DMs, the Reconstruction Operator $\mathcal{R}$ from (1) can be written as

$$\mathcal{R}\mathbf{s} = \boldsymbol{\delta\phi}(\mathbf{r}) = \mathbf{f}^T(\mathbf{r})\mathbf{a} \overset{\Delta}{=} \mathbf{f}^T(\mathbf{r})\mathbb{R}\mathbf{s}, \tag{8}$$

where $\mathbb{R}$ is just a matrix. It is this MCAO *control matrix* mentioned above, which will be called the *Recostructor,* that is to be found from (1).

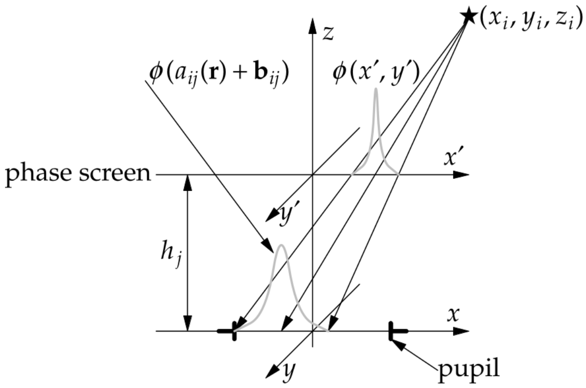


Fig. 2. Geometrical optics propagation from a phase screen (or DM) to the telescope pupil. Function $\phi_j(x',y')$ defined in a coordinate frame (x',y',z') on the jth phase screen is central-projected along the rays emitted from the ith point source with coordinates (x_i,y_i,z_i) to the telescope pupil to become $\phi_j\left(a_{ij}(\mathbf{r})+\mathbf{b}_{ij}\right)$ in the pupil coordinate frame$(\mathbf{r},z)$

The Propagation Operators $\mathcal{H}_{\phi,a}$ from (1) can also be easily presented in a matrix form with a certain reservation: with an assumption of *weak* atmospheric turbulence one can accept the *Rytov Approximation* [9] for the light propagation through the turbulence layers to the telescope pupil, which states that the results of light diffraction by each phase screen $\boldsymbol{\phi}_p(\mathbf{r})\,|_{p=1}^{\#PS}$ can be *summed independently* on the pupil. Thus, the *Propagation Matrices* $\mathbb{H}_{\phi,a}$ equivalent to $H_{\phi,a}$ have the block structure

$$\mathbb{H}_\phi=\begin{bmatrix} w_1 \quad \mathbb{H}_{11} & \dots & w_1 \quad \mathbb{H}_{1\#PS} \\ \vdots & \ddots & \vdots \\ w_{\#ST}\ \mathbb{H}_{\#ST1} & \cdots & w_{\#ST}\ \mathbb{H}_{\#ST\#PS} \end{bmatrix}, \quad \mathbb{H}_a=\begin{bmatrix} w_1 \quad \mathbb{H}_{11} & \dots & w_1 \quad \mathbb{H}_{1\#DM} \\ \vdots & \ddots & \vdots \\ w_{\#ST}\ \mathbb{H}_{\#ST1} & \cdots & w_{\#ST}\ \mathbb{H}_{\#ST\#DM} \end{bmatrix}, \tag{9}$$

where $w_s\,|_{s=1}^{\#ST}$ are the weighting coefficients for each Scientific Target, each block $\mathbb{H}_{sp}$ in the *geometrical optics approximation* (see Fig. 2) can be computed using (5) as

$$\mathbb{H}_{sp}=\left(\mathbf{h}(\mathbf{r}),\mathbf{h}^T(\mathbf{r})\right)^{-1}\left(\mathbf{h}(\mathbf{r}),\boldsymbol{\phi}_p^T(a_{sp}\mathbf{r}+\mathbf{b}_{sp})\right),\quad a_{sp}=1-h_p/z_s,\quad \mathbf{b}_{sp}=h_p(x_s,y_s)/z_s, \tag{10}$$

where $\boldsymbol{\phi}_j(x',y')$ is a basis defined on jth screen, $\mathbf{h}(\mathbf{r})$ is a basis defined on the telescope pupil. Product $\mathbb{H}_{sp}\boldsymbol{\phi}_p$ maps the pth screen's phase function $\phi_p(x',y')$ approximation weights $\phi_p(x',y')\approx\boldsymbol{\phi}_p^T(x',y')\boldsymbol{\phi}_p$ to approximation weights of the same function projected on the pupil: $\phi_p(a_{sp}\mathbf{r}+\mathbf{b}_{sp})\approx\mathbf{h}(\mathbf{r})^T(\mathbb{H}_{sp}\boldsymbol{\phi}_p)$. For $\mathbb{H}_a$ basis $\boldsymbol{\phi}_p(\cdot)$ is replaced in (10) with the Influence Function set $\mathbf{f}_p(\cdot)$ for pth DM, and products $\mathbb{H}_{sp}\boldsymbol{\phi}_p$ are replaced with $\mathbb{H}_{sp}\mathbf{a}_p$, where $\mathbf{a}_p$ is the pth DM actuator command vector. Substitution (8)–(10) into (1) gives

$$\boldsymbol{\sigma}^2=\left\langle\left\|H_\phi\phi(\mathbf{r},z)-H_aR\mathbf{s}\right\|^2\right\rangle\approx\left\langle\left(\mathbb{H}_\phi\boldsymbol{\phi}-\mathbb{H}_a\mathbb{R}\mathbf{s}\right)^T\mathbb{W}\left(\mathbb{H}_\phi\boldsymbol{\phi}-\mathbb{H}_a\mathbb{R}\mathbf{s}\right)\right\rangle, \tag{11}$$

where $\mathbb{W}$ is a block-diagonal matrix with blocks $w_s^2(\mathbf{h}(\mathbf{r}),\mathbf{h}^T(\mathbf{r}))\,|_{s=1}^{\#ST}$. Since (11) presents a *quadratic function,* finding its minimum w.r.t. elements of $\mathbb{R}$ is done by setting its first derivative to zero as

$$\frac{\partial}{\partial\mathbb{R}}\boldsymbol{\sigma}^2=0, \tag{12}$$

which, after applying the standard matrix derivative relations [11], gives for the optimal *Reconstructor Matrix*

$$\hat{\mathbb{R}}=(\mathbb{H}_a^T\mathbb{W}\mathbb{H}_a+\mathbb{N}\mathbb{N}^T+\alpha\mathbb{I})^{-1}\mathbb{H}_a^T\mathbb{W}\mathbb{H}_\phi\cdot\left\langle\boldsymbol{\phi}\mathbf{s}^T\right\rangle\left\langle\mathbf{s}\mathbf{s}^T\right\rangle^{-1}\overset{\Delta}{=}\mathbb{F}\cdot\mathbb{E}, \tag{13}$$

where $\mathbb{N}\mathbb{N}^T+\alpha\mathbb{I}$ is a *Regularization Term* added because, in general, the MCAO Tomography Problem is *ill-posed* [12]. Matrix $\mathbb{N}$ is usually *low-rank*. Eq. (13) formally solves the MCAO Tomographic

Problem. The solution gets separated into two independent parts: the *Phase Estimation* presented by matrix $\mathbb{F}$ and *DM Fitting* presented by matrix $\mathbb{F}$. Such a *Separation Principle* is yet another structural property. In what follows we concentrate on the efficient solution of the Estimation problem letting the Fitting problem be a subject for the future work. At this point it should be noted that the Fitting problem, being nearly the same laborious mathematically, is less demanding *computationally.* For the same reason we, for the time being, put aside the problem of *temporal control* involving *optimal time filtering* and *prediction.*

Computing the *Estimation Matrix* $\mathbb{E}$ requires further detalization. In particular, a *linearized model* for the WFS needs to be introduced:

$$\mathbf{s} \approx \mathbb{G}\boldsymbol{\phi} + \mathbf{n}, \tag{14}$$

where $\mathbb{G}$ is a *Sensor Matrix* mapping the 3D turbulence phase to the WFS measurements, $\mathbf{n}$ is the *Measurement Noise.* Substitution of (14) into (13) gives

$$\mathbb{E} = \left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle \mathbb{G}^T \left(\mathbb{G} \left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle \mathbb{G}^T + \left\langle \mathbf{n}\mathbf{n}^T \right\rangle + \mathbb{N}\mathbb{N}^T + \alpha \mathbb{I} \right)^{-} \tag{15}$$

where we used the fact that the turbulence and measurement noise random processes are statistically independent and thus $\left\langle \boldsymbol{\phi}\mathbf{n}^T \right\rangle = 0$ and added the formal *regularization term.* The WFS *noise covariance matrix* $\mathbf{n}\mathbf{n}^T$ is known *by design.*

Computation of the *turbulence covariance matrix* requires a special attention to the *internal structure* of the turbulence phase statistics, as the turbulence theory is well established [2, 13, 14]. For computation of $\left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle$ the following statistical properties are useful.

– The turbulence-induced phase perturbation is approximately a Gaussian random process with zero mean and covariance function

$$C\phi(\rho) = 0.033 C_n^2 (2\pi)^2 \frac{(L_0 / 2\pi)^{5/6} K_{5/6}(2\pi\rho / L_0)\rho^{5/6}}{2^{5/6}\Gamma(11/6)}, \quad \rho = \left| \mathbf{r}_1 - \mathbf{r}_2 \right|, \tag{16}$$

where $C_n^2[m^{-2/3}]$ is the turbulence-induced refractive index fluctuation strength, $L_0[m]$ the turbulence outer scale varying from 5 to 100 m, $K_{5/6}(\cdot)$ the modified Bessel function of second kind and order 5/6, $\Gamma(\cdot)$ is a Gamma-function.

– The turbulence strength profiles $C_n^2(h)$ are measureable with specialized instruments and well documented [2, 14].

– The turbulence correlation length is on the order of the outer scale L_0, which implies that the turbulence layers separated by a few hundred meters are statistically independent.

The Sensor Matrix has a block structure similar to that of the Propagation Matrices $\mathbb{H}$:

$$\mathbb{G} = \begin{bmatrix} \mathbb{G}_{11} & \dots & \mathbb{G}_{1\#PS} \\ \vdots & \ddots & \vdots \\ \mathbb{G}_{\#GS1} & \cdots & \mathbb{G}_{\#GS\#PS} \end{bmatrix}, \quad \mathbb{G}_{sp} = \mathscr{G}_s \boldsymbol{\phi}_p^T (a_{sp}\mathbf{r} + \mathbf{b}_{sp}), \tag{17}$$

where $\mathscr{G}_s = \left[\mathscr{G}_{s1} \dots \mathscr{G}_{s\#S} \right]^T$ is the WFS *measurement operator* that maps the phase in the telescope pupil to an sth WFS measurement vector $\mathbf{s}_s$. The form of this operator depends on the WFS type, for the most popular Shack-Hartmann WFS [15] it is quite simple

$$\begin{bmatrix} s_{ix} \\ s_{iy} \end{bmatrix} = \mathscr{G}_i \phi(\mathbf{r}) = \int_s ds \nabla \phi(\mathbf{r} - \mathbf{r_i}), \quad i = 1, \dots, \#S, \tag{18}$$

where S is the sensor *subaperture* shape, $\mathbf{r_i}$ is the subaperture center position.

The equations presented to the moment give enough information for what we will consider the *straightforward* MCAO Tomographic Problem practical implementation, both Estimation and Fitting parts, with minimal employment of the problem's internal structure, for instance, the fact that matrix

$\left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle$ is block-diagonal. It is assumed that all the involved matrices are *full.* By computing the matrix elements and the subsequent matrix multiplication one can get $\mathbb{E}, \mathbb{F}$ and $\mathbb{R} = \mathbb{F} \cdot \mathbb{E}$. Obviously, complexity of the full matrix application is proportional to the product of its dimensions, i. e. $\mathcal{O}(D^4)$. It is informative to estimate the MCAO algorithm complexity for an example of a large-scale AO system. For this purpose we can take the MCAO system for the Thirty Meter Telescope (TMT) [16]. As our attention will be concentrated on the $\mathbb{E}$-matrix, its dimension for the 30-meter telescope using, for instance, a 6-layer turbulence profile approximation taken from [14] is expected to be about 18000 × 33600, which needs to be applied within 1 ms. This well illustrates the difficulty of the problem. It also should be mentioned that the $\mathbb{E}$-matrix has to be recomputed regularly to adjust to the changing atmospheric conditions, so not only the complexity of $\mathbb{E}\mathbf{s}$-*multiplication* is of interest, but also the complexity of $\mathbb{E}$-*recomputation*, which in case of full matrix is $\mathcal{O}(D^6)$.

2. MCAO internal structure. It is clear that the straightforward full-matrix approach described above has unacceptable complexity and is barelysuitable for use with the large-scale AO systems of the prospective giant telescopes like TMT. The true complexity of the problem is expected however to be *much smaller* thanks to the *amenable internal structure* the straightforward approach completely ignores. In what follows, two the most significant facts will be mentioned, which, if used together, give the dramatic complexity reduction.

Zonal approximation. The computational complexity depends strongly on the choice of the *basis set* for the turbulence phase approximation. First of all, it is important to find a basis *optimal* for the specific turbulence statistics in the sense of the fastest convergence, a problem that still has no good solution. A second, more constructive approach, is to use the basis sets that are less computationally demanding. In particular [17], the basis options get split into two big classes: the *global* ones such as Zernike Polynomials [18], when each basis function $\phi_i(\mathbf{r})$ covers the entire phase screen, and the *zonal* ones, when the basis functions have *limited support* much smaller than the entire phase screen. The last option immediately makes the matrices in (13) *strongly sparse.* The prominent example of zonal basis functions are *B-splines* [19], the extremely computationally simple finite-support functions with an additional critical property: in the case of *Cardinal B-spline* they are also *shift-invariant,* which is the second structural property to be deeply analysed below. It was experimentally discovered [17] that atmospheric phase can be approximated with the same accuracy by a comparable number of the splines and Zernike Polynomials resulting, therefore, in the matrices of approximately the same dimensions but *full* in the latter and *sparse* in the former case.

Based on this observation the MCAO Tomographic Problem was completely reformulated in the *sparse form* [3] using the simplest, 1st-order *bi-linear, shift-invariant* spline approximation

$$\phi(\mathbf{r}) \approx \sum_{ij} \phi_{ij} Bi(\mathbf{r} - \mathbf{r}_{ij}), \quad \mathbf{r} = (x, y), \quad \mathbf{r}_{ij} = d(i, j),$$

$$Bi(r) = \begin{cases} (1 - |x/d|)(1 - |y/d|), |x|, |y| \le d, \\ 0, \ \text{otherwize}, \end{cases} \tag{19}$$

where, $\phi_{ij} = \phi(\mathbf{r}_{ij})$, which corresponds to a simple *bi-linear interpolation.* With the latter property Eq. (17) for $\left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle$ takes the especially simple form

$$\left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle_{ij,i',j'} = C_\phi(|\mathbf{r}_{ij} - \mathbf{r}_{i'j'}). \tag{20}$$

In Eq. (15) for the *Estimation Matrix* in the case of *bi-linear interpolation* the Sensor Matrix becomes sparse, the regularization term $\mathbb{N}\mathbb{N}^T + \alpha\mathbb{I}$ is low-rank and only $\left\langle \boldsymbol{\phi}\boldsymbol{\phi}^T \right\rangle$ is block-diagonal due to statistical independence of turbulence layers but with full blocks. This difficulty gets eliminated by application to (15) *Sherman-Morrison Fofmula (Inversion Lemma)* [20] to get (we skip the regularization term for clarity)

$$\mathbb{E} = \left\langle \boldsymbol{\phi\phi}^T \right\rangle \mathbb{G}^T \left(\mathbb{G} \left\langle \boldsymbol{\phi\phi}^T \right\rangle \mathbb{G}^T + \left\langle \mathbf{nn}^T \right\rangle \right)^{-1} = \left(\mathbb{G} \left\langle \mathbf{nn}^T \right\rangle^{-1} \mathbb{G} + \left\langle \boldsymbol{\phi\phi} \right\rangle^{-1} \right)^{-1} \mathbb{G}^T \left\langle \mathbf{nn}^T \right\rangle^{-1}, \qquad (21)$$

where each block $\left\langle \boldsymbol{\phi\phi} \right\rangle_p^{-1} \approx cW_p^{-1}\mathbb{LL}$, p=1…#PS, $\mathbb{L}$ is the *Laplacian Matrix* computed via finite differences on the $\mathbf{r}_{ij}$-grid, W_p is the layer's relative weight, c is the normalization constant. Since $\mathbb{L}$ is strongly sparse, the entire Estimation Matrix becomes presented as a combination of sparse and low-rank matrices. To stay within the sparse structure, the inversion of $\left(\mathbb{G}^T \left\langle \mathbf{nn}^T \right\rangle^{-1} \mathbb{G} + \left\langle \boldsymbol{\phi\phi} \right\rangle^{-1} \right)$ is rejected and solving iteratively the equivalent linear equation system

$$\left(\mathbb{G} \left\langle \mathbf{nn}^T \right\rangle^{-1} \mathbb{G}^T + \left\langle \boldsymbol{\phi\phi}^T \right\rangle^{-1} \right) \mathbf{x} = \mathbf{v}, \qquad (22)$$

by the *Preconditioned Conjugate Gradient Iteration* [17] done instead, which, unlike the inversion, preserves the sparsity. It was shown in [3] that the sparse approach results in about 40× speed-up in comparison to the *straightforward* full-matrix approach on a TMT-scale MCAO system even taking into account that the iterative linear equation system solution is required for each new $\mathbb{E}\mathbf{s}$-product computation.

Shift-invariance and Toeplitz structure. The sparse reformulation [3] of the MCAO Tomographic Problem had a significant impact and is currently considered a baseline solution for the practical AO systems. It is quite surprizing that another obvious and well-known structural property, the *shift-invariance,* manifesting itself:

1) in the *turbulence covariance function* (16);
2) in the *spline basis* (in case of a uniform grid);
3) in the *subaperture pattern* of the standard Shack-Hartmann WFS,

was not seriously considered for the purpose of further complexity reduction. It was explicitly used only recently in, for instance [6, 21] where it was noticed that the shift-invariance results in the *convolutional type* Propagation Operators with the *Toeplitz* matrix representation [22]. That shift-invariance mathematically results in the Toeplitz matrix structure is the most obvious on an example of the $\left\langle \boldsymbol{\phi\phi}^T \right\rangle$-matrix, which, according to (20), is constructed from the shift-invariant turbulence covariance function $C_\phi(\rho)$. To see the Toeplitz structure, it is enough to consider just the matrix of distances between the grid points $\rho_{ii',jj'} = |\mathbf{r}_{ij} - \mathbf{r}'_{i'j'}|$. In the case of two 1D grids we can write directly from Fig. 3:

$$\begin{bmatrix} \rho_{11} & \rho_{21} & \rho_{31} \\ \rho_{12} & \rho_{22} & \rho_{32} \\ \rho_{13} & \rho_{23} & \rho_{33} \\ \rho_{14} & \rho_{24} & \rho_{34} \end{bmatrix} = \begin{bmatrix} \rho_{11} & \rho_{12} & \rho_{13} \\ \rho_{12} & \rho_{11} & \rho_{12} \\ \rho_{13} & \rho_{12} & \rho_{11} \\ \rho_{14} & \rho_{13} & \rho_{12} \end{bmatrix}, \qquad \begin{array}{c} \rho_{ii'} = \rho_{i'j} \,(\text{symmetric}), \\ \rho_{ii'} = \rho_{kk'}, if \quad i - i' = k - k' (\text{Toeplitz}). \end{array} \qquad (23)$$

The case of two 2D grids is analogous with the difference that $\rho_{ij,i',j'}$ has additional block structure with equal blocks repeating the Toeplitz pattern (*Block-Toeplitz with Toeplitz Block*s or BTTB).

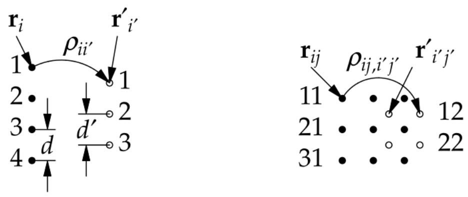


Fig. 3. Left: 1D grid point distances leading to the Toeplitz matrix. Right: 2D grid point distances leading to the BTTB matrix. Important limitations: 1) grid spacings should be *equal*: $d_x = d'_x$, $d_y = d'_y$; 2) the grids should be *parallel* (no grid-to-grid rotation!)

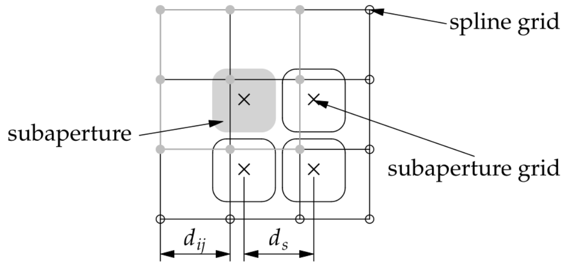


Fig. 4. The Matched Projection of the phase screen bi-linear spline basis grid to the Shack-Hartmann subaperture grid. The locations of splines affecting one subaperture marked in grey are also shown in grey. The grey contour shows the single bi-linear spline support border. The d_s is the subaperture grid spacing, d_{ij} is the jth phase screen spacing center-projected to the telescope pupil along the rays emitted from the ith Guide Star position as on Fig. 2

$$\begin{bmatrix} \rho_{11,11}\rho_{11,21} & \rho_{11,12}\rho_{11,22} \\ \rho_{21,11}\rho_{21,21} & \rho_{21,12}\rho_{21,22} \\ \rho_{31,11}\rho_{31,21} & \rho_{31,12}\rho_{31,22} \\ & \\ \rho_{12,11}\rho_{12,21} & \rho_{12,12}\rho_{12,22} \\ \rho_{22,11}\rho_{22,21} & \rho_{22,12}\rho_{22,22} \\ \rho_{32,11}\rho_{32,21} & \rho_{32,12}\rho_{32,22} \\ & \\ \rho_{13,11}\rho_{13,21} & \rho_{13,12}\rho_{13,22} \\ \rho_{23,11}\rho_{23,21} & \rho_{23,12}\rho_{23,22} \\ \rho_{33,11}\rho_{33,21} & \rho_{33,12}\rho_{33,22} \\ & \end{bmatrix} = \begin{bmatrix} \rho_{11,11}\rho_{11,21} & \rho_{11,12}\rho_{11,22} \\ \rho_{11,21}\rho_{11,11} & \rho_{11,22}\rho_{11,12} \\ \rho_{31,11}\rho_{11,21} & \rho_{31,12}\rho_{11,22} \\ & \\ \rho_{12,11}\rho_{12,21} & \rho_{11,11}\rho_{11,21} \\ \rho_{12,21}\rho_{12,11} & \rho_{11,21}\rho_{11,11} \\ \rho_{32,11}\rho_{12,21} & \rho_{31,11}\rho_{11,21} \\ & \\ \rho_{13,11}\rho_{13,21} & \rho_{12,11}\rho_{12,21} \\ \rho_{13,21}\rho_{13,11} & \rho_{12,21}\rho_{12,11} \\ \rho_{33,11}\rho_{13,21} & \rho_{32,11}\rho_{12,21} \\ & \end{bmatrix}, \quad \begin{matrix} \rho_{ij,i',j'} = \rho_{i'j,ij'} \text{(symmetric)}, \\ \rho_{ii',jj'} = \rho_{kk',ll'}, if \\ i-j=k-l, i'-j'=k'-l' \\ \text{(Block-Toeplitz with} \\ \text{Toeplitz Blocks or BTTB).} \end{matrix} \tag{24}$$

It is less obvious, but the Sensor Matrix $\mathbb{G}$ from (17), (18) is also BTTB at least for the Shack-Hartmann WFS. It can be seen from Fig. 4 that the pattern of 9 bi-linear splines projected on the subaperture grid in the telescope pupil *repeats itself* while being shifted from one subaperture to another, which is a consequence of the *shift*-invariance inherent in both basis splines and WFS subapertures.

The same two conditions need to be met as in the case of the covariance matrix:

1) the spline and subaperture grids should be *parallel*;

2) the spline and subaperture grid spacings should be *equal*:

$$d_s = d_{ij} = d_j / a_{ij}, \quad a_{ij} = 1 - h_j / z_i, \tag{25}$$

where h_j, z_i are the jth phase screen and ith Guide Star altitudes, d_j is the spline grid spacing right on the screen, which *has to be chosen* as $d_j = a_{ij} d_s$. Because d_j should not depend on i, one more requirement is necessary:

3) all Guide Star *altitudes should be the same,* i. e. z_i = const., which is not very strict and can be technically met.

Sparse Toeplitz matrices. As a result, all matrices making the Estimation Matrix in (21) can be made either block-diagonal with (2 × 2)-blocks like the $\left\langle \mathbf{nn}^T \right\rangle^{-1}$-matrix, low-rank like the regularization term $\mathbb{NN}^T$, or Toeplitz/BTTB like $\mathbb{G}$ and $\left\langle \boldsymbol{\phi\phi}^T \right\rangle^{-1} \approx c\mathbb{LL}$ (the Laplacian Matrix is also BTTB [3]). At the same time, because of the *Zonal Approximation,* these matrices are also sparse. This combination of sparsity and Toeplitz structure seems to have been overlooked.

An important property of the Toeplitz/BTTB-matrices is their *compact representation:* to define an $n \times m$ Toeplitz matrix $\mathbb{T}$, because of many equal elements, it is enough to store the values of non-equal elements only, for instance, the ones along the first row and the first column (see e.g. Eq. (23)), which are convenient to keep as a single vector

$$t(\mathbb{T}) = [\mathbb{T}_{n1} \ldots \mathbb{T}_{11} \mathbb{T}_{12} \ldots \mathbb{T}_{1m-1} \mathbb{T}_{1m}]^T. \tag{26}$$

Analogously, the compact representation of the BTTB-matrix with $N \times M$ blocks of $n \times m$ size is a 2D array, columns of which are ordered as

$$k(\mathbb{T}) = [\mathbf{t}_{Ni} \mathbf{t}_{N-11} \ldots \mathbf{t}_{11} \mathbf{t}_{12} \ldots \mathbf{t}_{1M-1} \mathbf{t}_{1M}]^T, \tag{27}$$

where each column is the representation (26) of the corresponding block.

Moreover, as the product $\mathbb{T}\mathbf{v}$ of the Toeplitz/BTTB-matrix by a vector is a *discrete convolution*, the key property of the compact representation is that it is exactly the *convolution kernel*.

In the case of a sparse Toeplitz/BTTB-matrix, the compact representation (or the convolution kernel) becomes also sparse, i. e. even more compact. For instance, the Laplacian compact representation is [3]

$$\begin{matrix} & \vdots & \vdots & \\ \ldots 0 & 1 & 0 \ldots \\ & 1 & -4 & 1 \\ \ldots 0 & 1 & 0 \ldots \\ & \vdots & \vdots & \end{matrix} \tag{28}$$

where "..." means zero padding to the necessary matrix size. Analogously, compact representation of the Sensor $\mathbb{G}$-matrix block, as it is seen from Fig. 4, has only 3 × 3 non-zero values. These non-zero parts are the actual amount of data defining the $\mathbb{L}$- and $\mathbb{G}$-matrices regardless the actual matrix size. Thus, in a great contrast with the full matrices requiring a long computation of all their elements, in a *compact Toeplitz* case it is necessary to calculate only a hand-full number of values. Returning to the TMT MCAO System matrix sizes one can make a comparison. Dimensions of one block of the $\mathbb{G}$-matrix are:

2800 × 3000 *in the full representation vs*;

≈2800 × 3 *non-zeros in just sparse representation* [3] *vs*;

3 × 3 *non-zeros in the sparse BTTB compact representation,*

and there is approximately the same picture for the $\mathbb{L}$-matrix. Such a dramatic complexity reduction allows to think that the described above structural observations approach is very close to the manifested *natural complexity* of the MCAO Tomography Problem. This, in particular, completely solves the $\mathbb{E}$-*recomputation* problem mentioned in Introduction: it takes virtually no time to recompute all matrices that make the $\mathbb{E}$-matrix, even in comparison with the $\mathbb{E}\mathbf{s}$-*multiplication* operation, which in turn requires some additional attention.

Another important property of the Toeplitz/BTTB-matrices is that the matrix-vector product being a convolution can be performed in the Fourier Space. This fact was explicitly used in [5] to reduce the complexity of the Tomography Problem to $\mathcal{O}(D^2 \log D^2)$. This approach, however, does not employ the existing matrix sparsity. Moreover, in the Fourier Space full and sparse Toeplitz matrices are indistinguishable w.r.t. computational complexity.

If the matrices are both Toeplitz/BTTB and strongly sparse, as it is the case for the ones in Eq. (21), the direct matrix-vector multiplication in the Coordinate Space appears to be more efficient, with complexity $NZ \cdot \mathcal{O}(D^2)$, where NZ is the (quite small) number of non-zero elements in the matrix compact representation. Consider a BTTB-matrix $\mathbb{T}$ with the *sparse compact representation* (non-zeros only) $\mathbf{k}(\mathbb{T})$ given as a 2D $(-d:u, -l:r)$-array. Since this array is a convolution kernel, the matrix-vector product $\mathbb{T}\mathbf{v}$ can be interpreted as a direct *2D convolution* of the $\mathbf{k}(\mathbb{T})$-kernel with a 2D $(1-d:n+u, 1-l:M+r)$-array presenting vector $\mathbf{v}$ to get a 2D $(1:n, 1:N)$-array presenting vector $\mathbb{T}\mathbf{v}$:

$$(\mathbb{T}\mathbf{v})(i, I) = \sum_{\beta=-1}^{r} \sum_{\alpha=-d}^{u} \mathbf{k}(\alpha, \beta)\mathbf{v}(di + i + \alpha, dI + I + \beta), \quad i = 1 \ldots n, \quad I = 1 \ldots N. \tag{29}$$

Note the extended limits of the array $\mathbf{v}$. This is what can be called the *Minimal Embedding* of $\mathbb{T}\mathbf{v}$ into $\mathbf{v}$: the array $\mathbf{v}$ gets broadened to such a minimal extent that the index $(i + \alpha, I + \beta)$ never falls out the array borders. The actual size of the **v**-array cannot be smaller than the minimal embedding but can be larger. For instance, for computing $\mathbb{G}_{sp}\phi_p$, it is necessary to additionally compute the shift of the subaperture grid projection (10) onto the *p*th Phase Screen along the *s*th Guide Star:

$$(dI, di)_{sp} = \left[\frac{\mathbf{b}_{sp}}{a_{sp} d_s}\right], \tag{30}$$

where [·] stands for rounding to the nearest integer. There exists another sort of *embedding*: because the telescope aperture is never rectangular, both left and right parts of (29) need to be *masked* by exclusion the *iI*-indexes of non-active phase/subaperture grid points.

The $\mathbb{E}\mathbf{s}$ operation also requires the *transpose matrix-vector* multplication. For a BTTB matrix this has a form of the *reverse convolution,* which cannot be written as a single formula but can be expressed as a recursion

$$\begin{aligned} &T^T\mathbf{v} = 0,\\ &(T^T\mathbf{v})(di+1+\alpha, dI+1+\beta) = \mathbf{k}(\alpha,\beta)\upsilon(1,1), \quad \alpha = -d...u, \quad \beta = -l...r,\\ &(T^T\mathbf{v})(di+i+\alpha, dI+I+\beta) = \mathbf{k}(\alpha,\beta)\upsilon(i,I) + (T^T\mathbf{v})(di+i-1+\alpha, dI+I-1+\beta),\\ &\alpha = -d...u, \quad \beta = -l...r, \quad i = 2...n, \quad I = 2...N, \end{aligned} \tag{31}$$

which looks clumsy in comparison to (29) but is easy to implement as a computer code.

The complete set of actions needed to compute $\mathbb{E}\mathbf{s}$ contains:

1. Eqs. (29) and (31) for multiplication by $\mathbb{G}$- and $\mathbb{L}$-matrices.

2. Multiplications by the block-diagonal $\left\langle \mathbf{nn}^T \right\rangle^{-1}$ and the full but low-rank $\mathbb{N}\mathbb{N}^T$.

3. Iterative solving Eq. (22), which reduces mostly to matrix-vector multiplications.

With the low-rank, sparse or sparse + BTTB structure of the involved matrices all the listed operations become extremely fast. A deep difference between the full and even just sparse $\mathbb{E}$-matrix and its *natural complexity* representation and *natural* operations is that neither *explicit* matrix allocation in memory nor *explicit* matrix operations are involved but replaced by specialized algorithms, which allows to talk about the *matrix-free* approach, a term that was firstly introduced, at least in the AO context, in [23].

3. Large scale AO simulations. In this section we present a demonstration of the complexity reduction effect comparing real practical computer implementation of the *Sparse Toeplitz MCAO Estimator* (21) with its classical implementations. Attention is focused on a single-step, open-loop, noiseless calculation, that is, on testing the

1) correctness of the computer code and

2) computational speedup,

ignoring estimation of an AO system performance as such. A 30-meter telescope AO system with the following parameters was taken as a *large-scale* testing case:

– a simple round annular 30-meter pupil;

– 6 *Sodium Layer Laser Guide Stars* located at finite altitude ≈90 km above ground, with central launch at 30″ off-axis;

– 5 *Scientific Natural Stars* at infinite distance, 1 on-axis and 4 at 20″ off-axis as it is shown in Fig. 5.

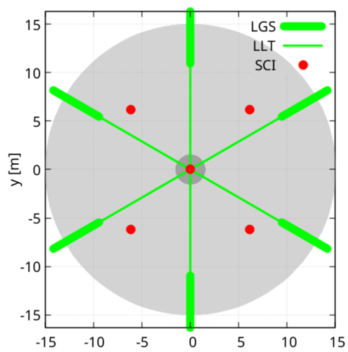


Fig. 5. 30-meter telescope pupil and the star constellation for the tested AO system. *LGS* is the Laser Guide Star, *SCI* is the Scientific Direction, *LLT* is the Laser Launch Telescope

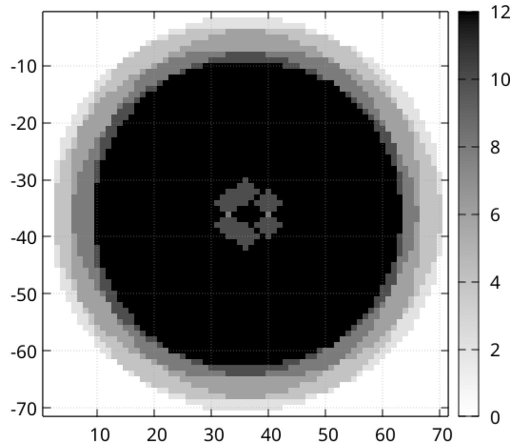
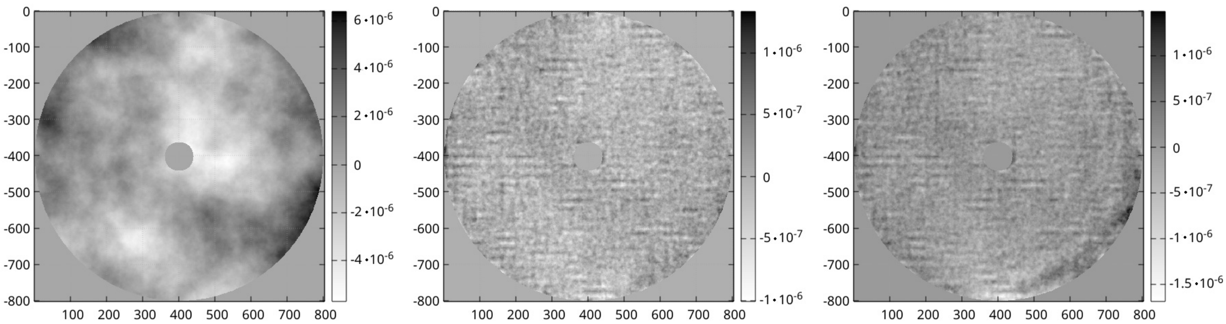


Fig. 6. The active point mask on the highest estimation phase screen obtained by projection of all sensor subaperture grids along the Guide Stars



RMS ≈2000 nm *RMS* ≈200 nm *RMS* ≈240 nm

Fig. 7. The *Piston,Tip,Tilt*-removed turbulent phase in the entrance pupil before and after AO correction. Left: phase on-axis before correction, the total $C_n^2 = 7\cdot 10^{-13}\,\mathrm{m}^{-2/3}$. Center: phase on-axis after correction. Right: phase 20″ off-axis after correction. The *Root Mean Square* error (RMS) is given in units of the *Opical Path Difference* (OPD)

– The atmospheric turbulence is presented as a Cerro Tololo 6-layer model [14], with random-generated high resolution phase screens.

– The AO system is presented as 6 LGS-illuminated Shack-Hartmann WFSs with square subapertures, the subaperture spacing is 0.5 m. Number of active subapertures per sensor is ≈2800.

– The low-resolution estimation phase screens $\boldsymbol{\phi}_p$ coincide with the model ones (the simplest model), have the same spacing as the sensor subapertures up to the *Guide Star Projection* (25), which is exactly the requirement to obtain the *Toeplitz Structure*. The number of active points on each estimation phase screen is slightly larger than the number of sensor subapertures to ensure the 1) *Minimal Embedding* (see Eqs. (29), (31)), 2) projection center shifts (30). Active points corresponding to the sensor subaperture grid onto the highest estimation phase screen are shown in Fig. 6. The number of active grid points on average is ≈3600. These numbers are important for computational complexity comparison.

The computation results on the desrcibed AO system that use the *Sparse Toeplitz* approach by iteratively solving the linear system (22) eploying the BTTB representations of the corresponding matrices and fast matrix-vector multiplications via Eqs. (29), (31) are presented in Fig. 7.

The most simple non-predictioned *Conjugate Gradient* iteration without additional regularization was used, which led to the large number of 100 iterations for acceptable convergence (Fig. 8), a separate difficult problem attempted to be solved elsewhere [12, 17] but not a key issue for this work.

Figs. 7, 8 demonstrate correctness of the *MCAO Estimation* algorithm in its *Sparse Toeplitz* implementation. As to its efficieny, it can be compared with two other key implementations of the same algorithm: the *Full Matrix* and *Sparse Matrix* ones. We need to take into account that in case of full matrix iterative solving of Eq. (22) is not necessary. Instead, Eq. (15) is computed as a single $\mathbb{E}$-matrix.

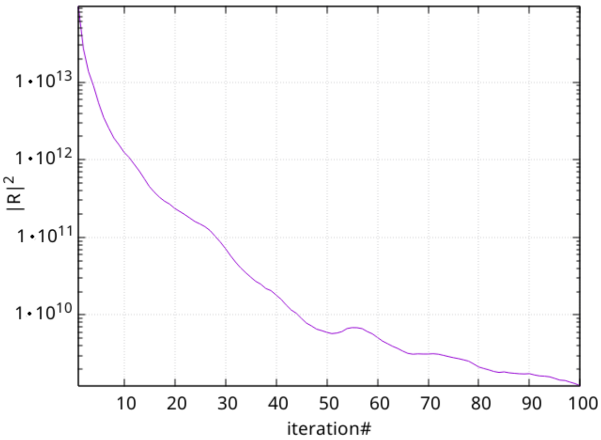


Fig. 8. *Conjugate Gradients* residual error evolution over iterations

Its size is (6 × 3600) × (12 × 2800) = 21600 × 33600. The time for direct multiplication of this matrix by a vector should be compared with the time of iterative solution of Eq. (22). In the just sparse matrix case the iterative solution is used, thus it is enough to compare the time needed for one iteration. Given below are the factual computation timing results obtained on the 64-bit Intel Core i3 2500 MHz processor on a single core without parallelization:

100 iterations of Sparse Toeplitz Algorithm	100 iterations of Sparse Algorithm	Direct matrix-vector multiplication
0.2 s	0.6 s	0.3 s

For the direct multiplication a standard BLAS library subroutine SGEMV() was used. For the sparse matrix-vector multiplication because each block of the involved matrices can be represented as a sparse band matrix the BLAS SGBMV() was used.

The comparison results are quite interesting. Obvious is that just Sparse approach is slower than the Sparse Toeplitz one. It is quite surprizing that the difference with the full-matrix approach is not very profound. The explanation is in the fact that the convergence of the iterative algorithm, which is driven entirely by the involved matrix *condition number,* is too slow. If for convergence just ONE iteration was enough, the Sparse Toeplitz algorithm would be more than two orders of magnitude faster, but slow convergence of the iteration process eliminates this advantage. Apparently, the future work needs to be concentrated on finding the efficient *preconditioner* for the Conjugate Gradients and, possibly, some optimized AO system configuration resulting in lower condition number. It should also be mentioned that the $\mathbb{E}$-matrix *recomputation* in the case of full matrix takes about an hour, whereas recomputation of BTTB-kernels for the matrices involved in (22) by numerical evaluating the integrals in Eq. (18) takes less than 0.1 second, thus solving the recomputation problem completely.

References

1. Trunk G. V. A Problem of Dimensionality: A Simple Example. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 1979, vol. PAMI-1, no. 3, pp. 306-307. https://doi.org/10.1109/TPAMI.1979.4766926

2. Roggemann M. C., Welsh B. M. *Imaging Through Turbulence*. Boca Raton, CRC Press, 1996. 320 p.

3. Ellerbroek B. Efficient computation of minimum-variance wave-front reconstructors with sparse matrix techniques. *Journal of the Optical Society of America A*, 2002, vol. 19, pp. 1803–1816. https://doi.org/10.1364/JOSAA.19.001803

4. Poyneer L. A., Gavel D. T., Brase J. M. Fast Wavefront Reconstruction in large adaptive optics systems with use of the Fourier transform. *Journal of the Optical Society of America A*, 2002, vol. 19, pp. 2100–2111. https://doi.org/10.1364/JOSAA.19.002100

5. Massioni P., Kulcsar C., Raynaud H.-F., Conan G.-M. Fast computation of an optimal controller for large-scale adaptive optics. *Journal of the Optical Society of America A*, 2011, vol. 28, pp. 2298–2309. https://doi.org/10.1364/JOSAA.28.002298

6. Ono X., Correia C., Conan R., Blanco L., Benoit N., Fusco T. Fast iterative tomographic wavefront estimator with recursive Toeplitz reconstructor structure for large-scale systems. *Journal of the Optical Society of America A*, 2018, vol. 35, pp. 1330–1345. https://doi.org/10.1364/JOSAA.35.001330

7. Beckers J. M. Increasing the size of isoplanatic path with multi-conjugate adaptive optics. *Very Large Telescopes and their Instrumentation, ESO Conference and Workshop Proceedings, Proceedings of a ESO Conference on Very Large Telescopes and their Instrumentation, held in Garching, March 21–24, 1988.* Garching, European Southern Observatory (ESO), 1988, pp. 693.

8. Tokovinin A., Le Luarn M., Sarazin M. Isoplanatism in a multi-conjugate adaptive optics. *Journal of the Optical Society of America A*, 2000, vol. 17, pp. 1819–1827. https://doi.org/10.1364/JOSAA.17.001819

9. Roddier F. The effects of atmospheric turbulence in optical astronomy. *Progress in Optics*, 1981, vol. 19, pp. 281–376. https://doi.org/10.1016/S0079-6638(08)70204-X

10. Gavel D. T., Wiberg D. M. Towards strehl-optimizing a aptive optics controllers. *SPIE Proceedings*, 2003, vol. 4839, pp. 890–901. https://doi.org/10.1117/12.459684

11. *Matrix calculus*. Available at: https://en.wikipedia.org/wiki/Matrix_calculus (accessed at 23.07.2025).

12. Ellerbroek B. L., Vogel C. R. Inverse Problems in Astronomical Adaptive Optics. *Inverse Problems*, 2009, vol. 25, no. 6, pp. 063001. https://doi.org/063001 10.1088/0266-5611/25/6/063001

13. Tatarski V. I. *Wave Propagation in a Turbulent Medium*. Courier Dover Publications, 2016. 288 p.

14. Hardy J. W. *Adaptive Optics for Astronomical Telescopes*. London, Oxford University Press, 1998. 448 p. https://doi.org/10.1093/oso/9780195090192.001.0001

15. Vernin J., Agabi A., Avila R., Azouit M., Conan R., Martin F., Masciadri, Sanchez I., Ziad A. 1998 *Gemini site testing campaign: Cerro Pachon and Cerro Tololo.* Gemini Doc. RTP-AO-G0094. Gemini Observatory, Hilo, 2000.

16. *Shack-Hartmann wavefront sensor.* Available at: https://en.wikipedia.org/wiki/Shack-Hartmann_wavefront_sensor (accessed at 23.07.2025).

17. Crane J., Herriot G., Andersen D., Atwood J., Byrnes P., Densmore A., Dunn J. [et al.]. NFIRAOS adaptive optics for the Thirty Meter Telescope. *Adaptive Optics Systems VI*, 2018, vol. 10703, pp. 107033V. https://doi.org/10.1117/12.2314341

18. Gilles L., Ellerbroek B. L., Vogel C. R. Preconditioned conjugate gradient wave-front reconstructors for multiconjugate adaptive optics. *Applied Optics*, 2003, vol. 42, no. 26, pp. 5233–5250. https://doi.org/10.1364/ao.42.005233

19. Noll R. J. Zernike polynomials and atmospheric turbulence. *Journal of the Optical Society of America*, 1976, vol. 66, no. 3, pp. 207–211. https://doi.org/10.1364/josa.66.000207

20. Golub G., VanLoan C. *Matrix Computations*. Baltimore John Hopkins University Press, 1989. 780 p.

21. Cranney J., Piatrou P., De Dona J., Rigaut F., Korkiakoski V. Evaluation of statistical turbulence models for use in the Distributed Kalman Filter. *Proceedings of the Adaptive Optics for Extremely Large Telescopes 5*, 2017, vol. 5. https://doi.org/10.26698/ao4elt5.0087

22. *Toeplitz matrix*. Available at: https://en.wikipedia.org/wiki/Toeplitz_matrix (accessed at 23.07.2025).

23. Ramlau R., Stadtler B. An augmented wavelet reconstructor for atmospheric tomography. *Arxiv* [Preprint], 2020. Available at: arXiv:2011.06842v1; https://doi.org/10.48550/arXiv.2011.06842

Information about the author

Piotr K. Piatrou – Ph. D. (Physics and Mathematics), Leading Researcher, B. I. Stepanov Institute of Physics of the National Academy of Sciences of Belarus (68-2, Nezavisimosti Ave., 220072, Minsk, Republic of Belarus). E-mail: p.piatrou@dragon.bas-net.by. https://orcid.org/0000-0001-9215-3570

Информация об авторе

Петров Петр Казимирович – кандидат физико-математических наук, ведущий научный сотрудник, Институт физики имени Б. И. Степанова Национальной академии наук Беларуси, Минск, Республика Беларусь (пр. Независимости, 68-2, 220072, Минск, Республика Беларусь). E-mail: p.piatrou@dragon.bas-net.by. https://orcid.org/0000-0001-9215-3570