Systematically structured H_2 optimal control for truss-supported segmented mirrors

Doelman, Reinier; Dominicus, Sander; Bastaits, Renaud; Verhaegen, Michel

DOI
10.1109/TCST.2018.2850306

Publication date
2018

Published in
IEEE Transactions on Control Systems Technology

Systematically Structured H_2 Optimal Control for Truss-Supported Segmented Mirrors

Reinier Doelman, Sander Dominicus, Renaud Bastaits, and Michel Verhaegen, *Member, IEEE*

***Abstract*— A systematic distributed optimal control design procedure is proposed for the rejection of wind load-induced disturbances on a truss-supported segmented mirror. The distributed nature of the controller is achieved by weighing of the interaction matrices between local (per-segment) controllers in a global $\mathcal{H}_2$ optimization. The procedure allows a tradeoff analysis between the controller implementation complexity versus the improved performance the extra communication brings. The procedure is demonstrated on a finite element model of a segmented mirror on a flexible supporting truss to which we apply the combined closed-loop performance and local controller interconnection structure optimization. The resulting set of controllers is compared to a set of baseline controllers including linear–quadratic–Gaussian control, singular value decomposition control, and a distributed controller where local controllers of neighboring segments communicate. The tradeoff analysis for the segmented mirror demonstrates that the communication between the local controllers can be greatly reduced without significantly compromising the rejection of wind-induced wavefront errors.**

***Index Terms*— Active optics, distributed control, optimal control, segmented mirror, structured control.**

I. Introduction

THE primary segmented mirror in large ground-based astronomical telescopes requires active control to compensate for the dynamic disturbance of wind loading on the segments.

As such, the modeling of segmented mirrors, the modeling of dynamical disturbance, design of controllers, and the evaluation of the closed-loop performance are studied during the design phase of telescopes (see [1] for the Keck Telescope, [2]–[4] for the European Extremely Large Telescope (E-ELT), or [5]–[8] for the Thirty Meter Telescope).

The dynamics of the identical segments are not necessarily decoupled. If the segments are mounted on a supporting truss, the backstructure causes interaction between the dynamics of the individual segments. This backstructure interaction is taken into account in the control design [2], [6]. The effect of the backstructure interaction increases with telescope size [9] and depends too on the structural damping in the supporting truss.

From the perspective of the entire closed-loop system, not only the backstructure causes spatial dynamics. A second source of the dynamical interaction is the way the position of the mirror segments is measured, which is through edge sensors that measure the relative displacement between neighboring segments. A third source is spatial correlation in the disturbance model (the force of the wind acting on the segments).

A fourth source is the dynamics of the control loop. For example, References [2] and [10] feature singular value decomposition (SVD)-based modal controllers. In [11], local (per-segment) controllers are designed that are connected to the controllers of neighboring segments. In [12], a centralized $\mathcal{H}_2$ optimal controller is designed and a distributed controller based on spatial invariance assumptions. We observe that, conceptually, the choice of the structure of the controller (centralized, decentralized, and hierarchical) is made first, and subsequently the controller is designed to meet a performance criterion. We propose in this brief to see the choice of the segmented mirror controller structure and the meeting of a performance criterion as something that should and could be done in a single controller design procedure. In this procedure, there is a multicriterion optimization that results in a tradeoff curve of controller complexity versus performance.

Even if current design methods can reach the desired performance, the use of an optimal control theory and controller structure optimization might open up the possibility of less stringent design criteria for other parts of the system, or might improve the end result of the whole optical system.

Several related approaches for static (state or output) feedback that search for a sparse controller structure regardless of a model structure can be found in [13]–[16]. Instead of static output feedback, we focus our attention on dynamic output feedback. We analyze the closed-loop performance of the system in the case that each segment has a local, low-order, dynamic output feedback controller, and a decision has to be taken on if and how these local controllers should be interconnected.

If the structure of the controllers is fixed up front, then the existing nonconvex (gradient descent) optimization methods could be applied [17], [18] to optimize the controller.

We can apply the relevant theory in [13] on appropriate addition of regularization on the decision variables to the objective function to induce sparsity in the resulting controller matrices.

Manuscript received November 23, 2017; revised May 2, 2018; accepted May 26, 2018. Manuscript received in final form June 21, 2018. This work was supported in part by the European Research Council through the Seventh Framework Programme (FP7/2007-2013) under Grant 339681 and in part by the Fonds National de la Recherche Scientifique under Grant FRIA FC76554. Recommended by Associate Editor E. C. Kerrigan. *(Corresponding author: Reinier Doelman.)*

R. Doelman, S. Dominicus, and M. Verhaegen are with the Delft Center for Systems and Control, Delft University of Technology, 2628 CD Delft, The Netherlands (e-mail:r.doelman@tudelft.nl; a.m.dominicus@student.tudelft.nl; m.verhaegen@tudelft.nl).

R. Bastaits is with Ion Beam Applications, Louvain-la-Neuve, Belgium (e-mail: renaud.bastaits@outlook.com).

Digital Object Identifier 10.1109/TCST.2018.2850306

We build on the work in [13], [17], and [18] to derive analytic expressions to compute the gradients of the multicriterion performance measure with respect to the controller matrices.

The result of this approach is that the offline, up front computational cost to design the controller is high, but the online computations are light (see also Section VI). The gradient descent procedure requires the solution of Lyapunov equations, whose computational complexity grows cubically with the number of states of the closed-loop system. The online computations are of a computational complexity that grows linearly with the number of interconnections between the local controllers, and quadratically with the number of states in a local controller. If the computations are done centrally, the optimization of the interconnection structure leads to scalable online computations as well. This is an advantage over a standard optimal linear–quadratic–Gaussian (LQG) controller, whose online computational requirements grow with the number of segments and system states.

In this brief, we discuss the multicriterion design procedure in Section II. The controller design procedure is demonstrated on an accurate, but numerically challenging the finite element-based virtual model of a segmented mirror on a flexible supporting truss. The model is discussed in Section III, and the necessary adaptations and transformations for optimal control design in Section IV. The complete model includes edge sensors, backstructure interaction, and a spatially correlated wind disturbance model. In Section V, we show the results of the multicriterion tradeoff analysis for this segmented mirror model. The results are discussed in Section VI.

II. Distributed Control Approach

As will be discussed in Section IV, the combined segmented mirror and wind loading model used for control design can be transformed into the following standard discrete-time system description:

$$\begin{pmatrix} x[k+1] \\ z[k] \\ y[k] \end{pmatrix} = \begin{pmatrix} A & B_1 & B \\ C_1 & 0 & 0 \\ C & F & 0 \end{pmatrix} \begin{pmatrix} x[k] \\ w[k] \\ u[k] \end{pmatrix} \tag{1}$$

with $x \in \mathbb{R}^n$, $z \in \mathbb{R}^{r_z}$, $y \in \mathbb{R}^{r_y}$, $w \in \mathbb{R}^{m_w}$, and $u \in \mathbb{R}^{m_u}$. For notational simplicity, time postscripts will be dropped, i.e., x is short for $x[k]$, z, y, w, and u are defined similarly, and x^+ is short for $x[k+1]$. The system matrices have dimensions that can be inferred from the signal dimensions. The vector y contains the edge sensor measurements, and the vector u is the segment position actuator inputs. For the segmented mirror system, the disturbance vector w is the white Gaussian noise that drives the wind model and the sensor noise. The output vector z describes the mirror shape. The global mirror shape is determined by the top position of each hexagonal segment's three position actuators, and the channels of z comprise the deviation of the position of the top of each actuator from the mean of the top positions of all the actuators.

For a system driven by white noise, $w \sim \mathcal{N}(0, I)$, with r_z output channels (indexed z_j), the squared $\mathcal{H}_2$ norm of the system equals the output variance

$$\|T_{w\to z}\|_{\mathcal{H}_2}^2 = \sum_{j=1}^{r_z} E\left[z_j^2\right]$$

where $T_{w\to z}$ is the transfer function from the disturbance vector w to performance channel z. If we use the linear approximation of the wavefront error of segmented mirrors [19], then the rms wavefront error e is related to the $\mathcal{H}_2$ norm by

$$e = \frac{2}{\sqrt{r_z}} \|T_{w\to z}\|_{\mathcal{H}_2}$$

where the factor of 2 comes from the fact that the rms wavefront error is twice the rms surface error. Standard LTI optimal control methods can be used to minimize the effects of wind on the error e.

We concentrate on dynamic output feedback controllers without a direct feedthrough term, i.e., controllers of the form

$$\begin{pmatrix} x_c^+ \\ u \end{pmatrix} = \begin{pmatrix} A_c & B_c \\ C_c & 0 \end{pmatrix} \begin{pmatrix} x_c \\ y \end{pmatrix} \tag{2}$$

where $x_c \in \mathbb{R}^{n_c}$ is the controller states. By closing the loop, we obtain the system

$$\begin{pmatrix} \mathbf{x}^+ \\ z \end{pmatrix} = \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & 0 \end{pmatrix} \begin{pmatrix} \mathbf{x} \\ w \end{pmatrix} \tag{3}$$

where

$$\left(\begin{array}{c|c} \mathcal{A} & \mathcal{B} \\ \hline \mathcal{C} & 0 \end{array}\right) = \left(\begin{array}{cc|c} A & BC_c & B_1 \\ B_cC & A_c & B_cF \\ \hline C_1 & 0 & 0 \end{array}\right)$$

and $\mathbf{x} = (x^T \quad x_c^T)^T$.

With a closed-loop transfer function $T_{cl,w\to z}$, the optimization problem becomes

$$\min_{A_c, B_c, C_c} J(A_c, B_c, C_c) \tag{4}$$

where $J(A_c, B_c, C_c) := \|T_{\mathrm{cl},w\to z}\|_{\mathcal{H}_2}^2$.

With A_c, B_c, and C_c known, the controllability and observability Gramians W_c and W_o are, respectively, determined by solving the Lyapunov equations

$$\begin{aligned} \mathcal{A}W_c\mathcal{A}^T - W_c + \mathcal{B}\mathcal{B}^T &= 0 \\ \mathcal{A}^T W_o\mathcal{A} - W_o + \mathcal{C}^T\mathcal{C} &= 0. \end{aligned} \tag{5}$$

The squared $\mathcal{H}_2$ norm of this system can be computed as follows (see in this context [17], [18])

$$\|T_{\mathrm{cl},w\to z}\|_{\mathcal{H}_2}^2 = (\mathcal{C}W_c\mathcal{C}^T) = (\mathcal{B}^T W_o\mathcal{B}).$$

Equation (5) shows that this criterion is not convex in the controller parameters, even if the controller parametrization is affine. Transformations exist that render the computation of an $\mathcal{H}_2$ controller a convex problem [20] through appropriate substitution of products of decision variables, but such a substitution would hamper imposing a desired sparsity structure on the controller matrices.

The squared $\mathcal{H}_2$ norm of the system is, however, differentiable and analytical expressions which can be derived for the gradients of the norm with respect to the controller matrices in (2). This also allows us to iteratively update the

controller matrices during a gradient descent optimization. The disadvantage is that due to the nonconvex nature of the problem, it cannot be guaranteed that the global optimum will be found. This consequence can be mitigated by trying multiple different starting points for the optimization.

A. Optimization Approach

Using the results in [17] and [18], we can derive the following gradients of $J(A_c, B_c, C_c)$ in (4) with respect to the controller matrices:

$$\begin{aligned}
&\frac{\partial J(A_c, B_c, C_c)}{\partial A_c} \\
&\quad = 2\begin{pmatrix}0 & I\end{pmatrix} W_o \mathcal{A} W_c \begin{pmatrix}0\\ I\end{pmatrix} \\
&\frac{\partial J(A_c, B_c, C_c)}{\partial B_c} \\
&\quad = \quad 2\begin{pmatrix}0 & I\end{pmatrix} W_o \mathcal{A} W_c \begin{pmatrix}C^T\\ 0\end{pmatrix} + 2\begin{pmatrix}0 & I\end{pmatrix} W_o \mathcal{B} F^T \\
&\frac{\partial J(A_c, B_c, C_c)}{\partial C_c} \\
&\qquad 2\begin{pmatrix}B^T & 0\end{pmatrix} W_o \mathcal{A} W_c \begin{pmatrix}0\\ I\end{pmatrix} + 2E^T \mathcal{C} W_c \begin{pmatrix}0\\ I\end{pmatrix}. \qquad (6)
\end{aligned}$$

The derivation of these expressions can be found in the Appendix.

The gradients in (6) can be used in a gradient descent scheme to find a locally optimal dynamic output feedback controller

In our implementation, we used the accelerated gradient descent method ADAM [21], because the method selects the step sizes automatically, and the method is efficient in both memory usage and the amount of required additional computations. For a gradient descent procedure with n_v variables, the required computations are of $\mathcal{O}(n_v)$ complexity and the additional required memory of $\mathcal{O}(n_v)$ size. The advantage of a gradient-based optimization is that the structure can be imposed on the controller system matrices. This will be discussed in Section II-B.

B. Discovering a SparseLy Connected Controller in a User-Motivated Global Structure

From a distributed controller design point of view, we can assume that there are N subsystems, with either one or more mirror segments per subsystem, and we would like to assign a local controller to each subsystem. If we assume that each local controller is connected only to the inputs and outputs of its own subsystem and controller states of other local controllers, then we can state that the matrices B_c and C_c have a block diagonal structure after proper renumbering of system inputs and outputs. Denote this as $B_c \in \mathcal{K}_{B_c}$ and $C_c \in \mathcal{K}_{C_c}$.

The matrix A_c can be written in the form

$$A_c = \begin{pmatrix} A_{c,11} & \cdots & A_{c,1N} \\ \vdots & & \vdots \\ A_{c,N1} & \cdots & A_{c,NN} \end{pmatrix} \qquad (7)$$

where $A_{c,ii},\ i = 1, \ldots, N$ constitutes the local controller dynamics of subsystem i and $A_{c,ij},\ i \neq j$ describes together with $A_{c,ji}$ the interaction between local controllers i and j. That is, the local controller update $x_{c,i}^+$ can be described as

$$x_{c,i}^+ = A_{c,ii} x_{c,i} + \sum_{j=1, j\neq i}^{N} A_{c,ij} x_{c,j} + B_{c,i} y_i \qquad (8)$$

where $x_c^T = (x_{c,1}^T \ \cdots \ x_{c,N}^T)$ is the controller state, y_i is local measurements, and $B_{c,i}$ is the appropriate block on the diagonal of B_c. The number of connections between the states of local controllers is computed by

$$\frac{1}{2} \sum_{i,j,\ i\neq j} \operatorname{card}\left(\left\| \begin{pmatrix} A_{c,ij} \\ A_{c,ji}^T \end{pmatrix} \right\|_F \right) \qquad (9)$$

where the function $\operatorname{card}(\cdot)$ denotes the cardinality operator

$$\operatorname{card}(q) = \begin{cases} 1 & q \neq 0, \\ 0 & q = 0. \end{cases}$$

The argument of the cardinality operator in (9) accounts for the fact that the local controller state information can flow both ways if there is an interconnection between the states of controllers i and j.

To find a distributed controller, one could force the matrix A_c to have blocks equal to zero, indicating that there is no possibility for communication between the two local controllers. Two examples are decentralized control and controllers where states of local controllers of neighboring segments are connected.

Based on [16] and [22], we propose to add the block sparsity promoting term

$$H(A_c) := \frac{1}{2} \sum_{i,j,\ i\neq j} \Gamma_{i,j} \left\| \begin{pmatrix} A_{c,ij} \\ A_{c,ji}^T \end{pmatrix} \right\|_F$$

as a weighted convex relaxation of (9), to the objective function (4), in order to tradeoff performance of the closed-loop system with the number of interconnections between the local controllers. The block weights $\Gamma_{i,j}$ influence the optimization's preference for having certain blocks put to 0. Using this weighting term, it is possible to systematically weigh the relative ease of implementation of connections between local controllers against an improved performance of the closed-loop system by varying the parameter γ.

The resulting optimization problem is

$$\begin{aligned}
\min_{A_c, B_c, C_c} \quad & J(A_c, B_c, C_c) + \gamma H(A_c) \\
\text{s.t. } & B_c \in \mathcal{K}_{B_c} \\
& C_c \in \mathcal{K}_{C_c} \qquad (10)
\end{aligned}$$

where γ is a regularization term, whereby we can influence the (general) level of block sparsity in A_c.

The derivative of $H(A_c)$ with respect to the blocks in A_c is

$$\frac{\partial H(A_c)}{\partial A_{c,kl}} = \frac{\Gamma_{k,l} A_{c,kl}}{\left\| \begin{pmatrix} A_{c,kl} \\ A_{c,lk}^T \end{pmatrix} \right\|_F}, \quad l \neq k, \left\| \begin{pmatrix} A_{c,kl} \\ A_{c,lk}^T \end{pmatrix} \right\|_F \neq 0.$$

From the derivatives with respect to the blocks of A_c, the entire gradient of the objective function in (10) with respect to A_c can be constructed.

When in addition to the sets $\mathcal{K}_{B_c}$ and $\mathcal{K}_{C_c}$, the set $\mathcal{K}_{A_c}$ is similarly defined as the nonzero block pattern of the matrix A_c, (10) can be optimized with $\gamma = 0$ and the additional constraint $A_c \in \mathcal{K}_{A_c}$:

$$\begin{aligned} \min_{A_c, B_c, C_c} \quad & J(A_c, B_c, C_c) \\ \text{s.t. } & A_c \in \mathcal{K}_{A_c} \\ & B_c \in \mathcal{K}_{B_c} \\ & C_c \in \mathcal{K}_{C_c}. \end{aligned} \tag{11}$$

The sets $\mathcal{K}_{A_c}, \mathcal{K}_{B_c}$, and $\mathcal{K}_{C_c}$ can be specified by the user. However, the set $\mathcal{K}_{A_c}$ can also be derived from the solution to (10) by thresholding the Frobenius norm of the blocks of A_c. The subsequent optimization is called "polishing" [13].

III. Segmented Mirror on a Flexible Supporting Truss

A. Mirror Model

We use the mirror model as described in [9] and [10] and Section IV-C. The mirror model has segments with a diameter of 1.8 m and consists of two rings and 18 segments.

The model is created using SAMCEF Finite Element Method (FEM) software and a Craig–Bampton reduction. The FEM model reads

$$\underbrace{\begin{pmatrix} \hat{M}_{11} & \hat{M}_{12} \\ \hat{M}_{21} & I \end{pmatrix}}_{\hat{M}} \begin{pmatrix} \ddot{x}_1 \\ \ddot{\alpha} \end{pmatrix} + \underbrace{\begin{pmatrix} C_{11} & 0 \\ 0 & 0 \end{pmatrix}}_{\hat{C}} \begin{pmatrix} \dot{x}_1 \\ \dot{\alpha} \end{pmatrix} + \underbrace{\begin{pmatrix} \hat{K}_{11} & 0 \\ 0 & \Omega^2 \end{pmatrix}}_{\hat{K}} \begin{pmatrix} x_1 \\ \alpha \end{pmatrix} = \begin{pmatrix} F_1 \\ 0 \end{pmatrix}. \tag{12}$$

The state x_1 contains the bottom and top positions of the actuators, $x_1 = (x_{\text{bottom}}^T \;\; x_{\text{top}}^T)$. α is the vector of modal amplitudes of the fixed boundary modes resulting from the Craig–Bampton reduction.

Forces F_1 are external forces on the truss, either through loading of the mirror segments or reaction forces in the truss support. Modal damping is added to the model as in [10] with a damping ratio of 1%, which is a standard value in segmented mirror research [9], [11], [23].

The matrices $\hat{M}_{11}$, $\hat{M}_{12}$, $\hat{M}_{21}$, Ω, and $\hat{K}_{11}$ are obtained from the Craig–Bampton reduction. The diagonal matrix Ω contains the natural frequencies of the fixed boundary modes.

The three actuators that suspend each segment are mounted on top of the supporting truss. The edge sensors are located on the edges of the mirror segments. Up to six sensors are present per segment, depending on the number of neighboring segments. All sensors measure the relative displacement with respect to the neighboring segment, in the out-of-plane direction. The six sensors pointing toward the middle, where there is no segment, measure the displacement with respect to the supporting truss.

B. Wind Load Disturbance and Control Objective

The control objective is to minimize the effect of wind loading on the shape of the mirror.

Many different disturbances act on the telescope structure, e.g., wind loading, edge sensor noise, and structural vibrations [4], [12], [24], [25]. From these, we incorporate only wind loading and edge sensor noise in the model. Other disturbances can be dealt with by the adaptive optics systems, separate control systems, or are of very little influence on the wavefront.

The wind disturbance considered is the along-wind response of the mirror, modeled with a classic random vibration approach. The turbulent wind force is assumed to follow Davenport's spectrum [26] and act on the model through the force F_1 in (12). The reference mean velocity of the wind is 10 $\mathrm{m\,s^{-1}}$ in the direction perpendicular to the mirror with a constant wind profile over the height of the mirror, and cross correlation of the disturbance on the different segments is assumed to be nonzero and computed according to [9]. The analytically computed cross power spectral density of Davenport's spectrum is approximated by a ninth-order band-limited white Gaussian noise-driven LTI system using the method in [27]. The edge sensors are assumed to have a noise level of $1\mathrm{nm}/\sqrt{\mathrm{Hz}}$.

The Maréchal criterion [28] states that the performance of an optical element is limited by diffraction when the rms wavefront error is lower than $(\lambda_l^2/180)^{1/2} \approx \lambda_l/13.4$, where λ_l is the wavelength of light. Since the smallest λ_l observed by the science instruments in, for example, E-ELT [29] is $\lambda_l = 370$ nm, the objective for the controller design is a closed-loop rms wavefront error below 27.6 nm. Any performance below this value we consider to be sufficient, though lower values indicate that other system requirements could be set more lenient. We do not consider frequency weighting of errors or the possibility of subsequent error compensation through adaptive optics in the performance comparison between different controllers.

IV. Model Adaptations for Optimal Control Engineering

We can write the FEM model into a descriptor (with the subscript d) state-space form

$$\begin{aligned} E_d \dot{x} &= A_d x + B_d u + B_v v \\ y &= C_d x + e \\ E_d &= \begin{pmatrix} \hat{M}_{11} & \hat{M}_{12} & 0 & 0 \\ \hat{M}_{21} & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{pmatrix}, \quad B_d = \begin{pmatrix} S_a k_a \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ A_d &= \begin{pmatrix} -C_{11} & 0 & -\hat{K}_{11} & 0 \\ 0 & 0 & 0 & -\Omega^2 \\ I & 0 & 0 & 0 \\ 0 & I & 0 & 0 \end{pmatrix} \quad B_v = \begin{pmatrix} I \\ 0 \\ 0 \\ 0 \end{pmatrix} \\ C_d &= (0 \;\; 0 \;\; S_y \;\; 0). \end{aligned} \tag{13}$$

Here, $x \in \mathbb{R}^{n_d}$ is the system state of the descriptor system, $A_d,\ E_d \in \mathbb{R}^{n_d \times n_d}$, $y \in \mathbb{R}^{n_y}$ is the measurement, e is the

sensor noise, and $u \in \mathbb{R}^{n_u}$ is the input. The matrix $S_a k_a$ in B_d describes the actuator topology and the influence of actuator inputs (displacements) u on the system. v is the output of the LTI wind model. The matrix S_y in C_d describes the sensor topology. The descriptor system in (13) is from a numerical point of view badly conditioned and not immediately suitable for simulation, optimization, and control with standard MATLAB toolboxes. We use the toolbox by Binder *et al*. [30] to transform a system (13) in series with the wind model that produces v, into the staircase canonical form, which transforms the matrix E_d into a diagonal matrix. The resulting state-space mirror model can be simulated accurately but still has a badly conditioned E matrix. We make a minimal realization of the system which removes all uncontrollable and unobservable poles. Since for these mirror models at hand, this removes the generalized eigenvalues at infinity, E_d is no longer badly conditioned and the model can be rewritten into a standard continuous-time state-space model.

Since the implementation of a controller is expected to be in discrete time, the model is converted to a discrete-time model. All dominant dynamics are contained in a bandwidth of 200 Hz, so the model is discretized with a sampling frequency of 1 kHz.

After this series of transformations, the model is in the form of (1) and the techniques outlined in Section II can be applied. A (numerically stable) square-root covariance filter [31] is used to compute a Kalman gain.

V. Numerical Results

The gradient descent procedure described in Section II was applied to a model of the segmented mirror with two rings, for a total number of $N = 18$ segments. This gives $n = 221$ system states, $r_y = 78$ sensors, and $r_z = m_u = 54$ actuators. Several baseline controllers are generated for comparison purposes.

1) A (globally optimal) LQG controller was created.
2) Then, an SVD controller similar to the controller in [9] was implemented to compare the performance of the optimal controllers to a controller based on a modal approach.

A range of differently structured dynamic output feedback controllers were obtained.

3) A dynamic output feedback controller with full matrices A_c, B_c, and C_c and a reduced number of controller states were designed using the gradient descent procedure. Since each segment has six degrees of freedom, we choose each local controller to have three states, for a total $n_c = 3N = 54$ controller states. We refer to this controller as an unstructured reduced order controller.
4) All the structured reduced order ($n_c = 54$) controllers have block diagonal matrices B_c and C_c. The fully interconnected version (all local controllers are connected to all other local controllers) therefore has a full matrix A_c, and block diagonal matrices B_c and C_c.
5) Finally, a structured reduced order ($n_c = 54$) controller is created where local controllers are connected to those local controllers that are associated with neighboring segments, which is reflected in the block structure of A_c. The matrices B_c and C_c are still block diagonal.

TABLE I

Performance of Different Control Strategies of the Two-Ring Mirror Model. The System Matrix Columns Feature an "f" for "Full Matrix," "b" for "Block Diagonal" or "s" for "Block Sparse"

	Controller	n_c	$A_c,\ B_c,\ C_c$	RMS WF error (nm)
1.	LQG	221	f, f, f	6.32
2.	SVD controller	-	-	127.96
3.	Unstruct. red. order	54	f, f, f	6.52
4.	Fully interconnected	54	f, b, b	14.52
5.	Neighbours connected	54	s, b, b	20.26

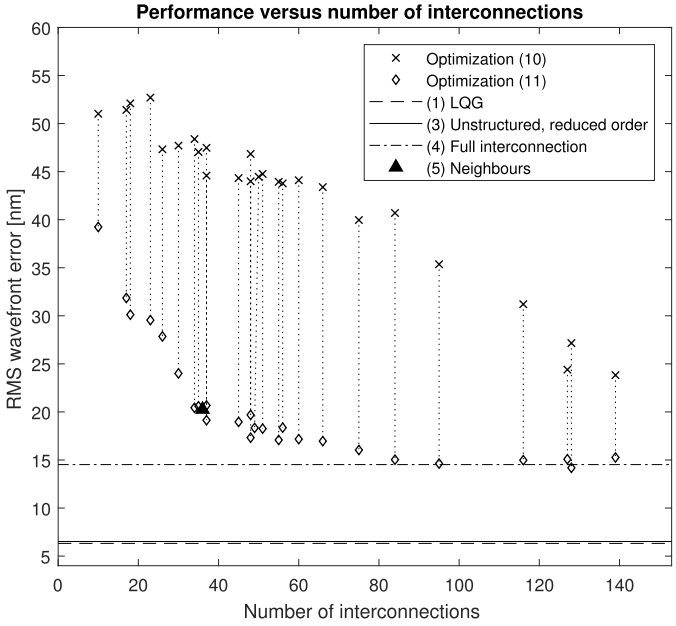

Fig. 1. Tradeoff between number of interconnections and wavefront error. Controllers marked by "x" obtained using (10). Controllers marked by a diamond are solutions to (11). The horizontal axis gives the value of (9) (the number of interconnections between local controllers). The vertical axis is the rms wavefront error as explained in Section III-B. The penalty paid for the use of the L1 norm as stand in for the cardinality operator is indicated by the vertical dotted lines.

Table I records the rms wavefront errors of controllers 1–5.

Fig. 1 displays the performance of the structured reduced order controllers obtained through optimization of (10) for different values of γ, resulting in the points marked "x." The values Γ_{ij} we chose to increase proportional to the square root of the Euclidean distance between the centers of segments i and j. A square root was used to not overly penalize the formation of longer distance connections but reflect a preference for shorter connections. The structure of B_c and C_c are fixed to block diagonal.

For the optimization of (10) using gradient descent, different initialization strategies can be applied. First, stabilizing controllers with near-zero values in the system matrices are used. In this way the closed-loop system is stable, and a solution to the Lyapunov equations in (5) can be found. Second, instead of a controller with near-zero matrices, the matrices of baseline controller 4 could have been used. A third option

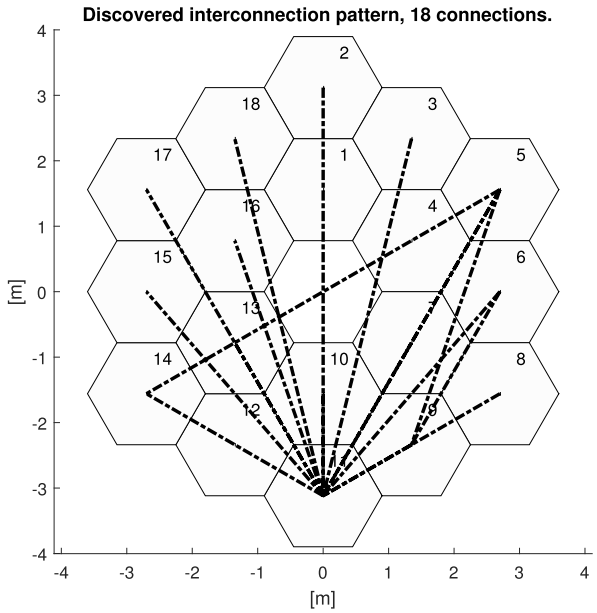

Fig. 2. Example of an interconnection pattern generated by the discovery procedure.

is that the value of γ could have been gradually changed and the optimization started with the controller from the previous optimization. We found that the second option gave the best results.

Once the gradient descent procedure converged, blocks with Frobenius norm below a threshold of 10^{-4} were deemed not to be in the interconnection structure. That is, there is no connection between segments i and j if

$$\left\| \begin{pmatrix} A_{c,kl} \\ A_{c,lk}^T \end{pmatrix} \right\|_F < 10^{-4}. \tag{14}$$

The resulting interconnection structure defines $\mathcal{K}_{A_c}$, and (11) can be used to compute a locally optimal value for the performance of the structured controller. However, given that proper initialization can make a difference for the result in nonconvex optimization, we utilized the availability of the matrices of baseline controller number 4. By setting γ high and Γ_{ij} to zero if segments i and j should be connected according to $\mathcal{K}_{A_c}$, and initializing the gradient descent procedure with the matrices of baseline controller 4, the "fully interconnected" controller transforms into a controller with the proper structure during the gradient descent procedure. The resulting performance is marked in Fig. 1 by a diamond marker. The corresponding "discovery" and "polishing" controllers are connected by a vertical dotted line. Apart from this initialization, a gradient descent method for problem (11) could also have been initialized, like the method for (10), with a controller with near-zero matrices, or the controller found in the "discovery" procedure.

In the same figure, some of the baseline controllers are also indicated by horizontal lines, and baseline controller number 5 is indicated with a star. Even though the "fully interconnected" controller should be plotted by a point in this graph, for comparison with controllers with a sparser interconnection structure, a line is drawn.

In Fig. 2, one of the discovered interconnection patterns (with 18 connections) is plotted on top of an image of the segmented mirror. A dashed line between two segment centers indicates that the local controllers are connected according to the sparsity structure in A_c.

VI. Discussion

Immediately apparent from Table I is the small difference between the global optimum (controller 1 and the LQG controller) and the reduced order unstructured controller (controller 3). When controller 3 is compared with controller number 4, where the difference is the structure in B_c and C_c, we see the wavefront error more than doubles.

The points marked with "x" are the performances of the best controllers found by optimizing (10) for different values of γ and different initial guesses for the controller matrices. A tradeoff can clearly be seen between the number of interconnections and performance. The identified interconnection structures for these controllers were used in a structured controller optimization (11), and the resulting performances are indicated with diamond markers. The difference between the tradeoff curves for the discovery procedure (marked with "x") and the subsequent structured optimization (diamond markers) is relatively large and the resulting curves are not smooth. The large differences in performance clearly show the penalty paid for using the Frobenius norm as a differentiable substitute for the cardinality operator. What is clear from the "polished" curve is that fewer than half of the interconnections between the mirror segment controllers are not necessary for approximately the same performance as baseline controller 4. This observation is not clear from the "discovery" curve, and neither is it from fixing the controller structure heuristically like in the baseline controllers.

The relatively small degradation in performance for controllers with approximately a third of possible interconnections not only justifies a search for optimal interconnection structures in distributed systems but is also relevant for situations where the online computation time of the input signal is critical. Even though the block sparse matrices A_c have an interpretation as being a controller that is distributed (see (8)) one might choose to do the controller computations in a centralized manner. Let n_{lc} be the dimension of the local controller, such that $n_c = Nn_{lc}$, and let p be the maximum number of blocks on a block row of A_c

$$p = \max_i \sum_j^N \text{card}(\|A_{c,ij}\|_F).$$

The multiplication of $A_c x_c$ has a computational complexity of $\mathcal{O}(n_c^2) = \mathcal{O}(N^2\, n_{lc}^2)$ for a dense matrix A_c. For a block sparse matrix A_c, it is possible to exploit the sparsity and parallellize the computation as in (8), and compute the result with the computational complexity $\mathcal{O}(pn_{lc}^2)$. The multiplication $B_c y$ can be computed with computational complexity $\mathcal{O}(n_c r_y) = \mathcal{O}(n_{lc} N r_y)$. However, for block diagonal B_c, this can be efficiently parallelized and computed in $\mathcal{O}(n_{lc} r_{ly})$, where r_{ly} is the maximum number of measurements per segment. A similar argument can be given for the computation of $C_c x_c$. Through an efficient use of the sparsity in the controller matrices in a centralized implementation of the controller,

the online computation time is not related to the number of subsystems, but to the prechosen number of states n_{lc}, and through p to the introduced level of sparsity. We see that for the analyzed segmented mirror, there can be a strong improvement in online computation time with only a small loss in performance. By optimizing the interconnection structure, the online computation time can be traded off against system performance.

In Fig. 2, one of the discovered sparsity patterns is displayed. One thing to notice is the absence of (rotational) symmetry and difference with a controller where neighboring segments are connected. In a sense such a symmetry was expected, since the segment configuration is rotationally symmetric, all segments are the same, and a flat mean wind velocity profile perpendicular to the mirror plane was assumed for the disturbance. We do see that for this sparsity level, optimization (10) resulted in a master–slave type of controller, where the controller of segment 11 connects to nearly all other controllers, and there are only a few connections among the other local controllers. Similar master–slave type of controllers can be observed for larger number of interconnections, with a few more local controllers with a "master" role.

Another interesting property of the controllers is that the computation of the control action does not feature the numerical issues of the original model.

Finally, it is important to note that many of the structured controllers in Fig. 1 have a remaining wavefront error that is lower than the maximum allowed wavefront error of 27.6 nm, meaning that the tradeoff analysis can and should play a role in the system design.

VII. Conclusion

We demonstrated how the performance—the residual wavefront error of the closed-loop system—of a structured controller, with an interpretation as a distributed controller, can be systematically traded off with the complexity of the distributed controller—interpreted as the number of interconnections between the local controllers. The method was applied to a challenging virtual model of a segmented mirror on a flexible support truss and a wind disturbance model with spatial correlation. For this system, a range of structured controllers were computed. The results show that compared to the performance of fully interconnected local controllers—controller 4 in Table I—the amount of interconnections can be greatly reduced without a significant loss of performance. Furthermore, for small amounts of interconnections and the particular weighting of interconnections we chose, controllers with a centralized aspect seem to be preferred over controllers where interconnections are distributed in a spatial sense.

For many different amounts of interconnections, the designed controllers have residual wavefront errors below the maximum allowed error and a performance improvement can be found by optimizing the interconnection structure of local controllers with respect to heuristic interconnection structures.

VIII. Future Work

As future work, we recommend that robustness against modeling uncertainty is investigated. That is, do the resulting interconnection patterns change if model uncertainty is taken into account? Or if the mirror design parameters change? Furthermore, is the centralized character of the optimal interconnection pattern, observed in Fig. 2, also present in the interconnection pattern of mirrors with more segments? The gap in performance between the controllers with (nearly) full A_c matrices and structured B_c and C_c matrices on the one hand and the unstructured reduced order controller on the other hand would motivate the further inclusion of block sparsity promoting terms in B_c and C_c in the objective function of optimization in the discovery procedure. It also leads to the question whether this inclusion affects the identified interconnection structures to a significant degree.

Appendix
Gradients of the $\mathcal{H}_2$ Norm With Respect to the Controller Matrices for the Discrete-Time Case

Using [17] and [18], we can derive that for the discrete-time case, the gradients of the squared $\mathcal{H}_2$ norm with respect to the closed-loop system matrices are

$$\frac{\partial(\mathcal{C}W_c\mathcal{C}^T)}{\partial\mathcal{A}} = 2W_o\mathcal{A}W_c$$
$$\frac{\partial(\mathcal{C}W_c\mathcal{C}^T)}{\partial\mathcal{B}} = 2W_o\mathcal{B}$$
$$\frac{\partial(\mathcal{B}^TW_o\mathcal{B})}{\partial\mathcal{C}} = 2\mathcal{C}W_c.$$

Using matrix calculus as described in [32], we have

$$\frac{\partial \operatorname{vec}(\mathcal{B})}{\partial \operatorname{vec}(B_c)} = \frac{\partial \operatorname{vec}\left(\begin{pmatrix}0\\ I\end{pmatrix}B_cF\right)}{\partial \operatorname{vec}(B_c)} = F \otimes \begin{pmatrix}0 & I\end{pmatrix}$$

and similarly

$$\frac{\partial \operatorname{vec}(\mathcal{A})}{\partial \operatorname{vec}(B_c)} = \begin{pmatrix}C & 0\end{pmatrix} \otimes \begin{pmatrix}0 & I\end{pmatrix}.$$

Using the Generalized Chain Rule (see [32, Th. 5.3]), we arrive at

$$\frac{\partial(\mathcal{C}W_c\mathcal{C}^T)}{\partial B_c} = 2\begin{pmatrix}0 & I\end{pmatrix}W_o\mathcal{A}W_c\begin{pmatrix}C^T\\ 0\end{pmatrix} + 2\begin{pmatrix}0 & I\end{pmatrix}W_o\mathcal{B}F^T.$$

The results for the matrices C_c and A_c in (6) can be derived along the same lines.

References

[1] J.-N. Aubrun, K. Lorell, T. Mast, and J. Nelson, "Dynamic analysis of the activley controlled segmented mirror of the W. M. Keck ten-meter telescope," *IEEE Control Syst. Mag.*, vol. 7, no. 6, pp. 3–10, Dec. 1987.

[2] B. Sedghi, M. Müller, M. Dimmler, B. Bauvir, T. Erm, H. Bonnet, and M. Cayrel, "Dynamical aspects in control of E-ELT segmented primary mirror (M1)," *Proc. SPIE*, vol. 7733, p. 77332E, Aug. 2010.

[3] B. Sedghi, M. Müller, H. Bonnet, M. Dimmler, and B. Bauvir, "Field stabilization (tip/tilt control) of E-ELT," *Proc. SPIE*, vol. 7733, p. 773340, Aug. 2010.

[4] M. Dimmler *et al.*, "E-ELT primary mirror control system," *Proc. SPIE*, vol. 7012, p. 70121O, Jul. 2008.

[5] D. G. MacMynowski and T. Andersen, "Wind buffeting of large telescopes," *Appl. Opt.*, vol. 49, no. 4, pp. 625–636, 2010.

[6] D. G. MacMynowski, P. M. Thompson, J. C. Shelton, L. C. Roberts, Jr., M. M. Colavita, and M. J. Sirota, "Control system modeling for the Thirty Meter Telescope primary mirror," *Proc. SPIE*, vol. 8336, p. 83360R, Nov. 2011.

[7] D. G. MacMartin and K. Vogiatzis, "Unsteady wind loads for TMT: replacing parametric models with CFD," *Proc. SPIE*, vol. 9150, p. 91500Q, Aug. 2014.

[8] D. G. MacMynowski, M. M. Colavita, W. Skidmore, and K. Vogiatzis, "Primary mirror dynamic disturbance models for TMT: Vibration and wind," *Proc. SPIE*, vol. 7738, p. 77380E, Jul. 2010.

[9] R. Bastaits, "Extremely large segmented mirrors: Dynamics, control and scale effects," Ph.D. dissertation, Univ. Libre Bruxelles, Brussels, Belgium, 2010.

[10] R. Bastaits, G. Rodrigues, B. Mokrani, and A. Preumont, "Active optics of large segmented mirrors: Dynamics and control," *J. Guid., Control, Dyn.*, vol. 32, no. 6, pp. 1795–1803, 2009.

[11] A. Sarlette and R. J. Sepulchre, "Control limitations from distributed sensing: Theory and extremely large telescope application," *Automatica*, vol. 50, no. 2, pp. 421–430, 2014.

[12] S. Jiang, P. G. Voulgaris, L. E. Holloway, and L. A. Thompson, "H_2 control of large segmented telescopes," *J. Vib. Control*, vol. 15, no. 6, pp. 923–949, 2009.

[13] F. Lin, "Structure identification and optimal design of large-scale networks of dynamical systems," Ph.D. dissertation, Dept. Elect. Comput. Eng., Univ. Minnesota, Minneapolis, MN, USA, 2012.

[14] F. Lin, M. Fardad, and M. R. Jovanovic, "Augmented Lagrangian approach to design of structured optimal state feedback gains," *IEEE Trans. Autom. Control*, vol. 56, no. 12, pp. 2923–2929, Dec. 2011.

[15] F. Lin, M. Fardad, and M. R. Jovanović, "Design of optimal sparse feedback gains via the alternating direction method of multipliers," *IEEE Trans. Autom. Control*, vol. 58, no. 9, pp. 2426–2431, Sep. 2013.

[16] M. R. Jovanović and N. K. Dhingra, "Controller architectures: Tradeoffs between performance and structure," *Eur. J. Control*, vol. 30, pp. 76–91, Jul. 2016.

[17] M. Mercadal, "H_2, fixed architecture, control design for large scale systems," Ph.D. dissertation, Massachusetts Inst. Technol., Cambridge, MA, USA, 1990.

[18] D. Petersson, "A nonlinear optimization approach to H_2-optimal modeling and control,"Ph.D. dissertation, Linköping Univ., Linköping, Sweden, 2013.

[19] G. Chanan, D. G. MacMartin, J. Nelson, and T. Mast, "Control and alignment of segmented-mirror telescopes: Matrices, modes, and error propagation," *Appl. Opt.*, vol. 43, no. 6, pp. 1223–1232, 2004.

[20] C. Scherer, P. Gahinet, and M. Chilali, "Multiobjective output-feedback control via LMI optimization," *IEEE Trans. Autom. Control*, vol. 42, no. 7, pp. 896–911, Jul. 1997.

[21] D. P. Kingma and J. Ba. (2014). "Adam: A method for stochastic optimization." [Online]. Available: https://arxiv.org/abs/1412.6980

[22] M. Yuan and Y. Lin, "Model selection and estimation in regression with grouped variables," *J. Roy. Statist. Soc., B (Statist. Methodol.)*, vol. 68, no. 1, pp. 49–67, 2006.

[23] D. G. MacMynowski, C. Blaurock, and G. Z. Angeli, "Initial control results for the thirty meter telescope," in *Proc. AIAA Guid., Navigat. Control Conf.*, 2005.

[24] B. Ulutas, D. Kerley, J. Dunn, A. Suleman, and E. J. Park, "Distributed H_∞ control of dynamically coupled segmented telescope mirrors: Design and simulation," *Mechatronics*, vol. 22, no. 1, pp. 121–135, 2012.

[25] B. Sedghi, M. Müller, and M. Dimmler, "Analyzing the impact of vibrations on E-ELT primary segmented mirror," *Proc. SPIE*, vol. 9911, p. 991111, Aug. 2016.

[26] A. G. Davenport, "The spectrum of horizontal gustiness near the ground in high winds," *Quart. J. Roy. Meteorol. Soc.*, vol. 87, no. 372, pp. 194–211, 1961.

[27] K. Hinnen, M. Verhaegen, and N. Doelman, "Robust spectral factor approximation of discrete-time frequency domain power spectras," *Automatica*, vol. 41, no. 10, pp. 1791–1798, 2005.

[28] A. Maréchal, "Étude des effets combinés de la diffraction et des aberrations géométriques sur l'image d'un point lumineux," *Rev. Opt.*, vol. 26, pp. 257–277, 1947.

[29] *The E-ELT Project Office, E-ELT Construction Proposal*, ESO, München, Germany, 2011.

[30] A. Binder, V. Mehrmann, A. Miedlar, and P. Schulze, "A MATLAB toolbox for the regularization of descriptor systems arising from generalized realization procedures," Inst. Math., Technische Univ. Berlin, Berlin, Germany, Tech. Rep., 2015.

[31] M. Verhaegen and V. Verdult, *Filtering and System Identification: A Least Squares Approach*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[32] D. A. Turkington, *Generalized Vectorization, Cross-Products, and Matrix Calculus*. Cambridge, U.K.: Cambridge Univ. Press, 2013.