H-infinity Optimal Distributed Control in Discrete Time

Lidström, Carolina; Pates, Richard; Rantzer, Anders

Published in:
2017 IEEE 56th Conference on Decision and Control, CDC 2017

DOI:
10.1109/CDC.2017.8264176

2018

Document Version:
Peer reviewed version (aka post-print)

H-infinity Optimal Distributed Control in Discrete Time

Carolina Lidström, Richard Pates and Anders Rantzer.

***Abstract*— We give closed-form expressions for H-infinity optimal state feedback laws applicable to linear time-invariant discrete time systems with symmetric and Schur state matrix. This class includes networked systems with local dynamics in each node and control action along each edge. Furthermore, the structure of the controllers mimics that of the system, which makes them suitable for distributed control purposes.**

I. INTRODUCTION

We study structured H_∞ control and give a class of linear time-invariant (LTI) discrete time systems for which distributed controllers are optimal. To give a flavour of our results, consider the subsystems

$$x_i(t+1) = a_i x(t) + b \sum_{(i,j)\in\mathcal{E}} u_{ij}(t) + d_i(t).$$

Here $i \in (1, \dots, N)$, $0 < a_i < 1$, $b > 0$, $u_{ij} = -u_{ji}$ and $\mathcal{E}$ is the edge set of a network with N nodes. This system is naturally associated with a graph, such as that in Figure 1. Each subsystem is depicted by a node, and the edges describe the couplings between subsystems through the control signals u_{ij}. We show that the static state feedback law

$$u_{ij}(t) = \frac{b}{a_i - 1} x_i(t) - \frac{b}{a_j - 1} x_j(t)$$

minimizes the H_∞ norm of the closed-loop system from the disturbance d to the state x and control input u when $a_i^2 + 2b^2 k_i < a_i$, where k_i is the degree of node i. This constraint is related to the speed of information propagation through the network as well as its connectivity. Note that each control input u_{ij} is only comprised of the states it directly affects, with a proportional term related to these subsystems. Therefore not only is this control optimal, it is also easy to apply, even when the number of subsystems N is incredibly large.

Control of large-scale and complex systems is most often performed in a distributed manner. This is due to the practical impossibility of having access to information about the overall system when deciding the control actions. However, it is not straightforward to translate the conventional control synthesis methods to synthesis of controllers suitable in a large-scale setting. In fact, the optimal distributed control problem can often be intractable.

The authors are members of the LCCC Linnaeus Center and the ELLIIT Excellence Center at Lund University. This work was supported by the Swedish Research Council through the LCCC Linnaeus Center and by the Swedish Foundation for Strategic Research through the project ICT-Psi.

C. Lidström, R. Pates and A. Rantzer are with the Department of Automatic Control, Lund University, Box 118, SE-221 00 Lund, Sweden. carolina.lidstrom@control.lth.se

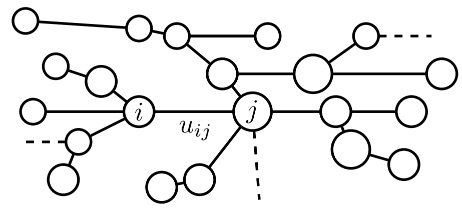

Fig. 1. Part of a network of N nodes. The dashed lines illustrates where the connections are to the rest of the network. Node i depicts subsystem i while edge (i, j) is related to the control input u_{ij} for which $u_{ij} = -u_{ji}$.

Imagine the system described previously to be written compactly as $x(t+1) = Ax + Bu + d$. The optimal control law can then be written as $u = B^T(I - A)^{-1}x$. In fact, this control law is optimal as long as A is symmetric and Schur and $A^2 + BB^T \prec A$. Also, it is on closed-form which is a rarity in the case of H_∞ control. Furthermore, our theory naturally suggests a controller with a structure related to the structure of the considered system, which makes it a candidate for distributed control. This link between the structure of the system and that of the controller is similar to what is described for spatially invariant systems [1], [2]. However, the systems we consider are not restricted to be spatially invariant, though our results are only valid for H_∞ norm performance requirements. Further, synthesis of structured controllers is simplified when the closed-loop system is required to be positive [3], [4]. Although, as we do not include this requirement, it is hard to compare it with our approach. In [5], [6], [7], they instead consider a localized approach to the design of distributed controllers. Of course, this can be conservative. In our case, a local information pattern is optimal whenever the dynamics can be divided into subsystems that only share control inputs.

H_∞ control, let alone distributed H_∞ control, has mainly been treated in the continuous time setting since these questions were first brought up half a century ago. See [8] for the first state space based solution to the centralized non-structured H_∞ control problem. In fact, in order to solve the centralized H_∞ control problem in discrete time an additional criterion on the system needs to be fulfilled [9], [10], [11]. However, it is of great importance to study the problem of distributed H_∞ control in discrete time, as controllers are almost always implemented digitally.

The work presented covers the non-trivial translation of the continuous time problems studied by the authors in [12] and [13]. Besides the static state feedback law described previously, we give a closed-form expression for a state feed-back controller with integral action which requires minimum control effort and guarantees a specified level of disturbance attenuation. It has similar structure preserving properties to the static state feedback controller as well as track references.

The outline is as follows. In Section II we present general results on the optimal state feedback controllers, and state their closed-forms as well as the system requirements needed for them to be applicable. In Section III, we discuss their relation to distributed control. Section IV displays how the results are related to their continuous time counterparts. The introduction is ended with the notation used.

The set of real numbers is denoted $\mathbb{R}$ and the space of n-by-m real-valued matrices is denoted $\mathbb{R}^{n\times m}$. The identity matrix is written as I. Given a matrix M, the spectral norm of M is denoted $\|M\|$ and the Moore-Penrose pseudo-inverse of M is denoted $M^\dagger$. A square matrix $M \in \mathbb{R}^{n\times n}$ is said to be Hurwitz if all eigenvalues have negative real part. It is said to be Schur if all eigenvalues are strictly inside the unit circle. Furthermore, for a square symmetric matrix M, $M \prec 0$ ($M \preceq 0$) means that M is negative (semi)definite while $M \succ 0$ ($M \succeq 0$) means M is positive (semi)definite. The H_∞ norm of a proper and real rational stable transfer function $T(z)$ is written as $\|T\|_\infty$ and given by $\|T\|_\infty = \sup_{\omega\in\mathbb{R}} \|T(e^{j\omega})\|$.

II. Closed-form H_∞ optimal state feedback in discrete time

The first part of this section treats H_∞ optimal static state feedback when the performance requirement is to minimize the impact from process disturbance on the state and control input. In the second part, we require the controller to track a reference in addition to a disturbance rejection requirement, with minimum control effort. Naturally, the latter controller has integral action. Further, we specify a class of systems for which the optimal controllers can be stated on closed-form. This class includes a broad range of linear networked systems. We will show how these results can be used for distributed control in Section III.

A. Optimal static state feedback

Consider the discrete time LTI system

$$x(t+1) = Ax + Bu + Hd \tag{1}$$

where the state $x \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, disturbance $d \in \mathbb{R}^l$ and the matrices A, B and H are of appropriate dimensions. Furthermore, consider the regulated output

$$\zeta = \begin{bmatrix} x \\ u \end{bmatrix}. \tag{2}$$

The objective is to find a stabilizing static state feedback law $u = Kx$, $K \in \mathbb{R}^{m\times n}$, that minimizes the H_∞ norm of the closed-loop system from d to ζ. We denote the transfer function of the closed-loop system with (1)-(2) and K by $T_{d\to\zeta}[K]$. It is given by

$$T_{d\to\zeta}[K](z) = \begin{bmatrix} I \\ K \end{bmatrix} (zI - A - BK)^{-1} H. \tag{3}$$

The following theorem gives a closed-form optimal control law, with respect to the objective described above, that is applicable to (1) with symmetric and Schur stable matrix A and for which $A^2 + BB^T \prec A$. The latter constraint is related to the sample time used in the time discretization of the equivalent continuous time system. This is discussed in more detail in Section IV.

Theorem 1: Consider (3) with A symmetric and Schur and $A^2 + BB^T \prec A$. Then, $\|T_{d\to\zeta}[K]\|_\infty$ is minimized by $K_{\text{opt}} = B^T(A - I)^{-1}$ and the minimal value of the norm is $\|H^T\left((A - I)^2 + BB^T\right)^{-1} H\|^{\frac{1}{2}}$.

The proof of Theorem 1 is given in the Appendix.

Remark 1: Note that the control law is independent of how the disturbance enters the system, i.e, H in (1).

B. Optimal PI control

Consider the discrete time LTI system

$$\begin{aligned} x(t+1) &= Ax + Bu + Bd \\ q(t+1) &= q + x \end{aligned} \tag{4}$$

where the state $x \in \mathbb{R}^n$, integral of the state x, i.e., $q \in \mathbb{R}^n$, control input $u \in \mathbb{R}^m$, disturbance $d \in \mathbb{R}^m$ and the matrices A and B are of appropriate dimensions. The objective is to find a stabilizing state feedback controller K, which maps $r - x$ to u, where $r \in \mathbb{R}^n$ is a reference signal. Furthermore, it should track r with minimum control effort and also guarantees a certain level of disturbance attenuation. The closed-loop transfer functions from r to u and d to q are denoted $T_{r\to u}[K]$ and $T_{d\to q}[K]$, respectively. The following theorem states a closed-form optimal control law, with respect to the objective described above, that is applicable to (4) with A symmetric and $0 \prec A \prec I$.

Theorem 2: Consider (4) with A symmetric and $0 \prec A \prec I$. Define $\gamma = \|\left((I - A)^{-1}B\right)^\dagger\|$ and assume that $\tau > 0$ fulfills

$$\tau\left(\tau I - \gamma B^T(I - A)^{-2}B\right) \succeq B^T(I - A)^{-4}AB.$$

Then, the problem

$$\begin{array}{ll} \text{minimize} & \|T_{r\to u}[K]\|_\infty \\ \text{subject to} & \|T_{d\to q}[K]\|_\infty \le \tau \end{array}$$

over stabilizing K, is solved by

$$\hat{K}_{\text{opt}}(z) = k\left(B^T(I - A)^{-2} + \frac{1}{z-1}B^T(I - A)^{-1}\right),$$

where $k = \gamma/\tau$. The optimal value is γ.

The proof of Theorem 2 is given in the Appendix.

Remark 2: The parameter τ determines the bandwidth of the control loop where a smaller τ corresponds to disturbance rejection over a wider frequency range.

Remark 3: Note that K enters as an ordinary negative feedback as compared to the previous subsection, where the negative feedback sign was incorporated in the control law.

III. DISTRIBUTED IMPLEMENTATION

This section concerns structured control and describes a particular type of systems for which Theorem 1 and 2 result in distributed controllers. The considered systems are comprised of subsystems with local dynamics, that only share control inputs. Furthermore, each control input only affects two subsystems. Depicting the subsystems as nodes and the control inputs as edges between the nodes they affect, the overall system can be illustrated by a network graph.

It is evident from the results given in the previous section that the open-loop system need to be structured itself in order for the optimal control laws to be structured. This is natural from a large-scale system point of view as the dynamics of such systems most definitely would be highly localized. For the networked systems considered in this section, the control signals drive the interaction among the subsystems. In a transportation network this is related to routing of the flow of commodities. Furthermore, the dynamics in each subsystem is diffusive, so they act as buffers. The dynamics could also describe a linear approximation of the behaviour of more involved dynamics around an operating point.

A. Static state feedback case

Consider a network with N nodes or subsystems,

$$x_i(t+1) = a_i x_i + b \sum_{(i,j)\in\mathcal{E}} u_{ij} + d_i \tag{5}$$

where $i \in (1, ..., N)$, $b \in \mathbb{R}$, $b > 0$ and $\mathcal{E}$ is edge set. The overall system can be written on the form (1) with A diagonal and B with columns of one element equal to 1 and one equal to -1, scaled by b, while the remaining elements are zero. See Figure 1 for an illustration of the system. Each node in the system's graph represents a subsystem i while an edge (i, j) illustrates how control signal u_{ij} enters the system. Furthermore, $u_{ij} = -u_{ji}$, i.e., what is drawn from system j is added to system i. This class of systems includes linear models of transportation and buffer networks. The Corollary below follows from Theorem 1 and gives a closed-form expression for an optimal distributed static state feedback law for (5).

Corollary 1: Consider a graph with a set of nodes $\mathcal{V}$ and edges $\mathcal{E}$. Let the dynamics in each node $i \in \mathcal{V}$ be given by (5) with $0 < a_i < 1$, $b > 0$ and $u_{ij} = -u_{ji}$. Furthermore, consider the N subsystems to be written on the form (1)-(2), where $x = \{x_i\}_{i\in\mathcal{V}}$, $u = \{u_{ij}\}_{(i,j)\in\mathcal{E}}$ and $d = \{d_i\}_{i\in\mathcal{V}}$ and assume that $A^2 + BB^T \prec A$. Then, the control law

$$u_{ij} = \frac{b}{a_i - 1} x_i - \frac{b}{a_j - 1} x_j$$

minimizes the H_∞ norm of the transfer function from the disturbance d to the regulated output ζ.

Proof: The overall system is given by $x(t+1) = Ax + Bu + d$, where A is diagonal and $0 \prec A \prec I$ with $A^2 + BB^T \prec A$. The controller structure then follows from Theorem 1. ■

Remark 4: In the distributed case, the constraint $A^2 + BB^T \prec A$ can be approximated by a local constraint. See Section IV for more details.

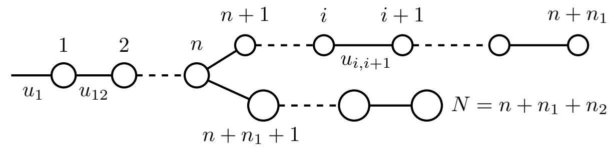

Fig. 2. Buffer network with N buffers. Buffer 1 has inputs u_1 and u_{12}. A general buffer i has inputs $u_{i,k}$, where k is the number of its neighbours. The network has a fork structure where, from the left, the first part has n buffers. The upper branch has n_1 buffers while the lower branch has n_2 buffers.

B. Distributed optimal PI control

Consider a slight variation to the subsystems in (5) with an extra state for each subsystem i as the integral of x_i, denoted q_i,

$$\begin{aligned} x_i(t+1) &= a_i x_i + b_i(u_i + d_i) + \sum_{(i,j)\in\mathcal{E}} (u_{ij} + d_{ij}), \\ q_i(t+1) &= q_i + x_i. \end{aligned} \tag{6}$$

In this system, the disturbances enter in the same way as the control inputs and again $u_{ij} = -u_{ji}$ as well as $d_{ij} = -d_{ji}$. The corollary given next follows from Theorem 2.

Corollary 2: Consider a graph with a set of nodes $\mathcal{V}$ and edges $\mathcal{E}$. Let the dynamics in each node $i \in \mathcal{V}$ be given by (6) with $0 < a_i < 1$, $b_i \neq 0$ for at least one i, $u_{ij} = -u_{ji}$, $d_{ij} = -d_{ji}$ and $q_i(0) = 0$. Denote $e_i = r_i - x_i$, where r_i is the reference signal for subsystem i. Furthermore, consider the overall system written on the form (4). Define $\gamma = \|((I - A)^{-1}B)^+\|/\tau$ and assume that

$$\tau(\tau - \gamma B^T(I - A)^{-2}B) \succeq B^T(I - A)^{-4}AB.$$

Then, the controller

$$\begin{aligned} p_i(t+1) &= p_i(t) + e_i(t), \\ u_{ij}(t) &= k\left(p_i/(1 - a_i) - p_j/(1 - a_j)\right) \\ &\quad + k\left(e_i/(1 - a_i)^2 - e_j/(1 - a_j)^2\right), \\ u_i(t) &= k\left(p_i b_i/(1 - a_i) + e_i b_i/(1 - a_i)^2\right), \end{aligned}$$

with $k = \gamma/\tau$, minimizes the H_∞ norm of the transfer function from r to u while keeping the L_2-gain from d to q bounded by τ.

Proof: The overall system is given by $x(t+1) = Ax + Bu + Bd$, where A is diagonal and $0 \prec A \prec I$. Furthermore, the assumptions in Theorem 2 hold, and the controller structure thus follows from Theorem 2. ■

C. Numerical example

Consider the buffer network depicted in Figure 2, where N is the total number of buffers. The network has a fork structure where the leftmost part, the root, has n buffers while the upper and lower branch has n_1 and n_2 buffers, respectively. The dynamics of the content in the buffers, around some operating point, is

$$x_1(t+1) = a_1 x_1 + b_1(u_1 + d_1) + u_{12} + d_{12},$$
$$x_i(t+1) = a_i x_i + \sum_{(i,j)\in\mathcal{E}} u_{ij} + d_{ij},\ \forall\, i \in 2,\ldots,N.$$

The disturbance d_{ij} enters on edge (i,j). The control inputs and disturbances satisfy $u_{ij} = -u_{ji}$ and $d_{ij} = -d_{ji}$, respectively. Each $0 < a_i < 1$, so given a non-zero initial state the buffers will eventually be empty if no control is used and after any disturbance has abated. The content in the buffers dissipates faster through the lower branch than through the upper.

The system described is part of the class of systems treated in the previous subsections. We will now show the closed-loop behaviour of the system with the distributed static state feedback and distributed PI controller, respectively. As we consider the number of buffers N to be large, a localized control approach is the only practical design for implementation. Our results suggest control inputs u_{ij} that only require local information regardless of the size of the system N.

Figure 3 shows the levels in some of the buffers in the network, over time. At time $t = 20$ a constant disturbance enters in node 1. The disturbance is then processed through the network via the edges. The two upper plots and the lower right plot show the time trajectories of the three first buffers in the head, the upper branch and the lower branch, respectively. The dotted lines show the references, while the dashed and solid lines show the trajectories given the static and PI controller, respectively. The bottom left plot shows the time trajectory of the disturbance. Note that with the distributed PI controller the impact of the disturbance is larger in the lower branch than in the upper branch, even though it has faster dynamics. Also, you can see how the process is evolved through the network, starting at the head and following the branches by comparing the peaks of the time trajectories. It is evident that the PI controller is able to track the reference while the static controller leaves a stationary error under the duration of the disturbance.

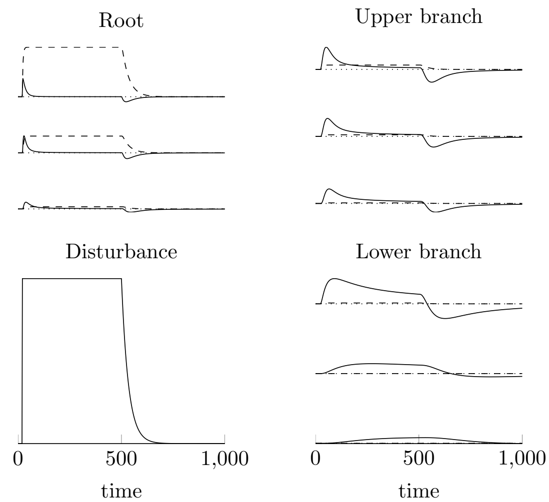

Fig. 3. Numerical example with buffer network of N buffers. The two upper plots and the lower right plot show the time trajectories of the three first buffers in the root, upper branch and lower branch, respectively. For instance the upper left plot show the time trajectories of node 1 in the top, node 2 in the middle and node 3 at the bottom. The plots for the branches are constructed in the same manner. The dotted lines show the references, while the dashed and solid lines show the trajectories given the static and PI controllers, respectively. The bottom left plot shows the disturbance that enters in node 1. Note that in the PI control case, the effect of the disturbance is larger in the lower branch than in the upper branch, even though it has faster dynamics. Also, it is evident that the PI controller is able to track the reference while the static controller leaves a stationary error.

IV. Discussion

A. Comparison to continuous time results

We will now compare Theorem 1 with its continuous time counterpart stated by the authors in [12]. For clarity, we include that result next.

Theorem 3 ([12]): Consider

$$G_{d\to\zeta}[K](s) = \begin{bmatrix} I \\ K \end{bmatrix} (sI - A_c - BK)^{-1} H,$$

with A_c symmetric and Hurwitz. Then $\|G_{d\to z}[K]\|_\infty$ is minimized by $K_{\text{opt}} = B^T A_c^{-1}$ and the minimal value of the norm is given by $\|H^T(A_c^2 + BB^T)^{-1}H\|^{\frac{1}{2}}$.

The statements in Theorem 1 and 3 are very similar, however, with one main difference. That is, the extra requirement on the matrices of the system's state space representation that is required in the discrete time case, i.e., $A^2 + BB^T \prec A$. If we consider the discretization of the open-loop continuous time system with time period h, we get

$$x(t+1) = e^{A_c h} x(t) + \int_0^h e^{A_c\tau} Bu(\tau) d\tau.$$

For small h and given the assumption that $u(t)$ is constant during each time period, we can use the approximation

$$x(t+1) \approx (I + A_c h)x(t) + hBu(t).$$

The constraint then becomes

$$(I + A_c h)^2 + h^2 BB^T \prec I + A_c h,$$

which is equivalent to $h < \|A_c^{\frac{1}{2}}(A_c^2 + BB^T)^{-1}A_c^{\frac{1}{2}}\|$. Thus, for small enough h, it is always fulfilled. Similarly to the discussion above, one can compare Theorem 2 to its continuous time equivalent stated in [13].

The constraint $A^2 + BB^T \prec A$ reveals that $A \succ 0$ in the discrete time case. Thus, the class of discrete time systems that can be considered for both Theorem 1 and Theorem 2 are non-oscillatory. This also maps to the class of continuous time systems considered in Theorem 3, as symmetric matrices do not have imaginary eigenvalues.

The optimal controller given by Theorem 3 is clearly related to the controller resulting from bisecting over the continuous time algebraic Riccati equation (CARE). That is, if we denote the solution to the CARE by P, the controller is given by $K = -B^T P$ [14]. In the discrete time setting, they are not as clearly related. The controller given by bisection over the discrete time ARE (DARE) is $K = -(I + B^T PB)^{-1} B^T PA$ where P is the solution to the DARE [14]. The expression for the DARE controller is more involved than the expression we give for the controller proposed in Theorem 1.

B. Local condition for $A^2 + BB^T \prec A$

Consider the systems described in Section III. They are of the form (1) with A diagonal and B sparse. In fact, given the network description of these systems, the matrix B relates the nodes and edges. B/b is generally called the node-link incidence matrix of the network graph. From this, it is possible to approximate the constraint $A^2 + BB^T \prec A$ by a local constraint. Denote the i:th diagonal element of A by a_i. Furthermore, write $BB^T = b^2 L$. The inequality can then be written as

$$A^2 - A + b^2 L \prec 0.$$

Further, define the diagonal matrix D by $D_{ii} := L_{ii}$. Then the inequality becomes

$$D^{\frac{1}{2}} \left(D^{-1}(A^2 - A) + b^2 D^{-\frac{1}{2}} L D^{-\frac{1}{2}} \right) D^{\frac{1}{2}} \prec 0,$$
$$D^{-1}(A^2 - A) + b^2 L_{\text{sym}} \prec 0,$$

where L_{sym} is the symmetric normalized Laplacian of the network's graph. It is well-known that $\lambda_{\max}(L_{\text{sym}}) \leq 2$. Therefore, satisfying the local condition

$$a_i^2 - a_i + 2b^2 D_{ii} < 0.$$

is sufficient to guarantee that $A^2 + BB^T \prec A$. The entry D_{ii} is the degree of node i, i.e., the number of nodes it is directly connected to, and often denoted k_i. Thus, the constraint $A^2 + BB^T \prec A$ is related to the speed of information propagation through the network and its connectivity, via a_i, b, the bound on the maximum eigenvalue of the symmetric normalized Laplacian and the node degree. Note that a similar analysis can be made for the inequality constraint on τ in Corollary 2, where it is necessary that $\tau I - \gamma B^T (I - A)^{-2} B \succ 0$.

V. CONCLUSIONS AND FUTURE WORKS

We define a class of systems and performance objectives for which the optimal H_∞ controller is structured. It includes networked systems with local dynamics in each node and control action along each edge. Some assumptions on this class of systems are only sufficient. It is left as future work to characterize the both sufficient and necessary systems properties. Furthermore, combining the discrete time results presented with their continuous time counterparts could bring some further intuition into the case of sampled data control for networked systems.

APPENDIX

To prove Theorem 1, we need the following lemma.

Lemma 1: Assume $A \in \mathbb{R}^{n \times n}$ symmetric and Schur, and $B \in \mathbb{R}^{n \times m}$. Then, the following statements are equivalent

(i) $A^2 + BB^T \prec A$,

(ii) $(A - I)\left((A - I)^2 + BB^T\right)^{-1}(A - I) + A - I \succ 0$.

Proof: Note that $A - I \prec 0$ as A is symmetric and Schur. Then,

$$\text{(ii)} \iff \left((A - I)^2 + BB^T\right)^{-1} + (A - I)^{-1} \succ 0$$
$$\iff -A + I \succ (A - I)^2 + BB^T \iff \text{(i)},$$

where in the first step we have multiplied (ii) with $(A - I)^{-1}$ from both left and right. ■

Next, we give the proof of Theorem 1.

Proof of Theorem 1: The proof is divided into two parts. The first part considers a lower bound on the minimal norm-value. In the second part, we show stabilizability of K_{opt} and that the lower bound is achieved for K_{opt}.

The minimal norm-value can be lower bounded as follows

$$\inf_K \|T_{d\to\zeta}[K]\|_\infty = \inf_K \sup_\omega \|T_{d\to\zeta}[K](e^{j\omega})\|$$
$$\geq \inf_K \|T_{d\to\zeta}[K](1)\|. \quad (7)$$

The latter minimization problem is equivalent to the following least-squares problem

$$\begin{aligned} &\text{minimize } \|\zeta\|_2 \\ &\text{subject to } \begin{bmatrix} I - A & -B \end{bmatrix} \zeta = Hd, \ \|d\|_2 \leq 1, \end{aligned}$$

which has the optimal solution

$$\zeta_* = \begin{bmatrix} I - A^T \\ -B^T \end{bmatrix} \left((A - I)(A - I)^T + BB^T\right)^{-1} Hd_*.$$

Thus, given (7) and symmetry of A we have that

$$\inf_K \|T_{d\to\zeta}[K]\|_\infty \geq \|H^T((A - I)^2 + BB^T)^{-1} H\|^{\frac{1}{2}}.$$

We will now prove that $K_{\text{opt}} = B^T(A - I)^{-1}$ is optimal by showing that it is stabilizing and achieves the lower bound given above. For K_{opt} to be stabilizing, $A_{cl} := A + BB^T(A - I)^{-1}$ has to be Schur. It is equivalent to existence of a matrix $P \succ 0$ such that $A_{cl} P A_{cl}^T - P \prec 0$. One such P is $P = (A - I)^2$, which is valid as $A \prec I$. Note that, given the assumptions on A and B, $A_{cl} P A_{cl}^T - P \prec 0$ with $P = (A - I)^2$ is equivalent to $(A - I)^2 + BB^T \succ 0$, which is true as $A \prec I$.

To show that K_{opt} achieves the lower bound, rewrite

$$T_{d\to\zeta}[K_{\text{opt}}](e^{j\omega})^* T_{d\to\zeta}[K_{\text{opt}}](e^{j\omega}) = H^T G^{-1}(j\omega) H$$

where

$$\begin{aligned} G(j\omega) &= (e^{j\omega} - I)(A - I)M^{-1}(A - I)(e^{-j\omega} - I) \\ &\quad - (e^{j\omega} + e^{-j\omega} - 2I)(A - I) + M \\ &= (2 - 2\cos(\omega)) \underbrace{\left((A - I)M^{-1}(A - I) + A - I\right)}_{=:N} + M, \end{aligned}$$

and $M := (A - I)^2 + BB^T$. It follows from Lemma 1 that $N \succ 0$ as $A^2 + BB^T \prec A$ by assumption. Therefore, it holds that $G(j\omega) \succeq M$ and moreover that $G(j\omega)^{-1} \preceq M^{-1}$ for $\omega \in [0, 2\pi)$. Hence,

$$\begin{aligned} T_{d\to\zeta}[K_{\text{opt}}](e^{j\omega})^* T_{d\to\zeta}[K_{\text{opt}}](e^{j\omega}) &= H^T G^{-1}(j\omega) H \\ \preceq H^T M^{-1} H = H^T((A - I)^2 + BB^T)^{-1} H &= \\ T_{d\to\zeta}[K_{\text{opt}}](1)^* T_{d\to\zeta}[K_{\text{opt}}](1) \end{aligned}$$

from which it follows that

$$\|T_{d\to\zeta}[K_{\text{opt}}]\|_\infty = \|H^T((A-I)^2+BB^T)^{-1}H\|^{\frac{1}{2}}$$

and the proof is complete. ■

To prove Theorem 2 we need the following lemma.

Lemma 2 ([13]): Let $A \in \mathbb{C}^{n\times m}$. Then,

$$\min_{X\in\mathbb{C}^{q\times n}} \|X\| \text{ s.t. } AXA = A$$

has the minimal value $\|A^\dagger\|$, attained by $\hat{X} = A^\dagger$.

Proof of Theorem 2: Define $P(z) := (zI-A)^{-1}B$. Then, $T_{r\to u}[K] = (I+KP)^{-1}K$. Now, define

$$M := I - kB^T(I-A)^{-2}B$$

and note that the given assumption on τ together with $A \succ 0$ yields $M \succ 0$. Firstly, we will show that $\hat{K}_{\text{opt}}$ is stabilizing. Factorize B as $B = GF^T$, where G and F have full column rank. Then,

$$\begin{aligned} T_{r\to u}[\hat{K}_{\text{opt}}](z) &= k\left((z-1)I + kB^T(I-A)^{-2}B\right)^{-1}\cdot \\ &\quad B^T(I-A)^{-2}(zI-A) \\ &= kH\left((z-1)I + kG^T(I-A)^{-2}GF^TF\right)^{-1}\cdot \\ &\qquad G^T(I-A)^{-2}(zI-A), \end{aligned}$$

so the poles of $T_{r\to u}[\hat{K}_{\text{opt}}]$ are the eigenvalues of

$$I - kG^T(I-A)^{-2}GF^TF,$$

i.e., the eigenvalues of M that are not equal to 1. Clearly, as $0 \prec M \preceq I$, $T_{r\to u}[\hat{K}_{\text{opt}}]$ is stable. Thus, $\hat{K}_{\text{opt}}$ is stabilizing.

Now, we will show that $\|T_{d\to q}[\hat{K}_{\text{opt}}]\|_\infty \le \tau$. From the definition of k we have that $I \preceq k^2\tau^2B^T(I-A)^{-2}B$ which is equivalent to

$$\begin{aligned} &B^T(2(1-\cos(\omega))A + (I-A)^2)^{-1}B \\ &\quad\preceq \tau^2\left[2(1-\cos(\omega))M + k^2(B^T(I-A)^{-2}B)^2\right] \end{aligned}$$

and further to

$$\left\|\tfrac{1}{e^{j\omega}-1}P(e^{j\omega})\left(I+\hat{K}_{\text{opt}}(e^{j\omega})P(e^{j\omega})\right)^{-1}\right\| \le \tau.$$

The latter inequality is precisely $\|T_{d\to q}[\hat{K}_{\text{opt}}](e^{j\omega})\| \le \tau$ and thus $\hat{K}_{\text{opt}}$ fulfills the constraint.

Finally, we will show that $\hat{K}_{\text{opt}}$ is in fact optimal. Again, consider the constraint with τ. In general $PT_{r\to u}[K]P = P - P(I+KP)^{-1}$ so the constraint demands

$$P(1)T_{r\to u}[K](1)P(1) = P(1).$$

Now, consider the minimization of $\|T_{r\to u}[K](1)\|$ subject to the equality above. Then, by Lemma 2, $T_{r\to u}[K](1) = P(1)^\dagger$ with optimal value $\|P(1)^\dagger\|$. Now, as $T_{r\to u}[\hat{K}_{\text{opt}}](1) = P(1)^\dagger$, it follows that $\hat{K}_{\text{opt}}$ is a feasible solution to the static problem at $\omega = 0$. Furthermore, to show that $\hat{K}_{\text{opt}}$ is optimal for the non static problem we need to show that $\|T_{r\to u}[\hat{K}_{\text{opt}}]\|_\infty$ is achieved at $\omega = 0$. Consider $T_{r\to u}[\hat{K}_{\text{opt}}]T_{r\to u}[\hat{K}_{\text{opt}}]^* \preceq \gamma^2 I$ which is equivalent to

$$\begin{aligned} &k^2B^T(I-A)^{-2}(2(1-\cos(\omega))A + (I-A)^2)(I-A)^{-2}B \\ &\quad\preceq \gamma^2\left[2(1-\cos(\omega))M + k^2(B^T(I-A)^{-2}B)^2\right]. \end{aligned}$$

This inequality holds trivially for $\omega = 0$. It holds for all other other $\omega \in \mathbb{R}$ provided that

$$k^2B^T(I-A)^{-4}AB \preceq \gamma^2\left[I - kB^T(I-A)^{-2}B\right]$$

which is equivalent to the assumption on τ. Thus, $\|T_{r\to u}[\hat{K}_{\text{opt}}]\|_\infty$ takes the minimal value γ and the proof is complete. ■

References

[1] B. Bamieh, F. Paganini, and M. A. Dahleh, "Distributed control of spatially invariant systems," *IEEE Transactions on automatic control*, vol. 47, no. 7, pp. 1091–1107, 2002.

[2] B. Bamieh and P. G. Voulgaris, "A convex characterization of distributed control problems in spatially invariant systems with communication constraints," *Systems & Control Letters*, vol. 54, no. 6, pp. 575–583, 2005.

[3] T. Tanaka and C. Langbort, "The bounded real lemma for internally positive systems and H-infinity structured static state feedback," *IEEE transactions on Automatic Control*, vol. 56, no. 9, pp. 2218–2223, 2011.

[4] A. Rantzer, "Scalable control of positive systems," *European Journal of Control*, vol. 24, pp. 72–80, 2015.

[5] J.-C. Delvenne and C. Langbort, "The price of distributed design in optimal control," in *Decision and Control, 2006 45th IEEE Conference on*. IEEE, 2006, pp. 3640–3645.

[6] Y.-S. Wang, N. Matni, S. You, and J. C. Doyle, "Localized distributed state feedback control with communication delays," in *American Control Conference (ACC), 2014*. IEEE, 2014, pp. 5748–5755.

[7] R. Pates and G. Vinnicombe, "Scalable design of heterogeneous networks," *IEEE Transactions on Automatic Control*, 2016.

[8] J. C. Doyle, K. Glover, P. P. Khargonekar, and B. A. Francis, "State-space solutions to standard H2 and H-infinity control problems," *IEEE Transactions on Automatic control*, vol. 34, no. 8, pp. 831–847, 1989.

[9] D. Limebeer, M. Green, and D. Walker, "Discrete-time H-infinity control," in *Decision and Control, 1989., Proceedings of the 28th IEEE Conference on*. IEEE, 1989, pp. 392–396.

[10] I. Yaesh and U. Shaked, "Minimum H-infinity norm regulation of linear discrete-time systems and its relation to linear quadratic discrete games," in *Decision and Control, 1989., Proceedings of the 28th IEEE Conference on*. IEEE, 1989, pp. 942–947.

[11] D.-W. Gu, M. Tsai, S. O'young, and I. Postlethwaite, "State-space formulae for discrete-time H-infinity optimization," *International Journal of Control*, vol. 49, no. 5, pp. 1683–1723, 1989.

[12] C. Lidström and A. Rantzer, "Optimal H-infinity state feedback for systems with symmetric and hurwitz state matrix," in *American Control Conference (ACC), 2016*. IEEE, 2016, pp. 3366–3371.

[13] A. Rantzer, C. Lidström, and R. Pates, "Structure preserving H-infinity optimal PI control," *arXiv preprint arXiv:1612.07824. To appear in the proceedings of IFAC World Congress*, 2017.

[14] A. A. Stoorvogel, *The H-infinity Control Problem: A State Space Approach*. Prentice Hall, New Jersey, 1992.