2023 62nd IEEE Conference on Decision and Control (CDC)
December 13-15, 2023. Marina Bay Sands, Singapore

Optimization Algorithm Synthesis based on Integral Quadratic Constraints: A Tutorial

Carsten W. Scherer, Christian Ebenbauer and Tobias Holicki

***Abstract*— We expose in a tutorial fashion the mechanisms which underlie the synthesis of optimization algorithms based on dynamic integral quadratic constraints. We reveal how these tools from robust control allow to design accelerated gradient descent algorithms with optimal guaranteed convergence rates by solving small-sized convex semi-definite programs. It is shown that this extends to the design of extremum controllers, with the goal to regulate the output of a general linear closed-loop system to the minimum of an objective function.**

Numerical experiments illustrate that we can not only recover gradient decent and the triple momentum variant of Nesterov's accelerated first order algorithm, but also automatically synthesize optimal algorithms even if the gradient information is passed through non-trivial dynamics, such as time-delays.

***Index Terms*— Optimization Algorithms, Robust Control, Linear Matrix Inequalities.**

I. Introduction

Accelerated gradient algorithms [1] have a wide range of applications in the current era of machine learning and online optimization-based control. From the perspective of control theory, such algorithms can be viewed as a linear time-invariant discrete-time (LTI) system in feedback with the gradient of the to-be-minimized function as a nonlinearity [2]–[5]. This provides an immediate link to absolute stability theory and offers the possibility to apply advanced tools from robust control for the *automated analysis* of accelerated gradient algorithms [5]. By tuning the algorithm parameters based on these tools, the convergence rate of Nesterov's algorithm [6] has been improved to get the so-called triple momentum algorithm [7].

The *automated synthesis* of optimization algorithms by convex optimization is a much more challenging task. This falls into the area of robust feedback controller design [8], [9]. Recent work [10]–[12] has addressed the synthesis problem from this perspective, but based on heuristic methods without optimality guarantees. An alternative approach to non-convex algorithm design by interpolation techniques can be found in [13], [14].

The purpose of this paper is to develop, in a tutorial fashion, the whole pipeline of analysis techniques that open the avenue for a convex solution to the automated algorithm synthesis problem by solving a moderate-sized convex semi-definite program. Another feature of the presented approach is its flexibility. It offers a convex solution to the so-called extremum control problem, with the goal to regulate the output of a dynamical system to the minimum of some convex cost function. These main results are based on [15], [16]. However, we also present an innovation over [16] which renders synthesis possible for LTI systems without any restrictions on their poles or zeros.

The paper is structured as follows. In Sec. II, we show how the algorithm analysis and synthesis problems translate into one of robustness analysis and synthesis. Sec. III recaps robustness analysis with static integral quadratic constraints (IQCs). Dynamic IQCs are introduced in Sec. IV, while the corresponding robust stability test is given in Sec. V. The design of algorithms is presented in Sec. VI and numerical illustrations are found in Sec. VII. Concluding remarks are given in Sec. VIII. All proofs and some explanatory connections to classical passivity-based stability tests are found in an extended version of the paper on arXiv [33].

Next to standard notations, for matrices A, B we express by $A \leq B$ that $B-A$ is nonnegative entrywise, while $A \prec B$ means that A and B are symmetric and $B - A$ is positive definite. For a tuple of matrices $A = (A_1, \ldots, A_k)$, we use

$$\operatorname{diag}(A) = \begin{pmatrix} A_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & A_k \end{pmatrix} \quad \text{and} \quad \operatorname{col}(A) = \begin{pmatrix} A_1 \\ \vdots \\ A_k \end{pmatrix}$$

if the dimensions are compatible. For the real polynomial $\alpha(\mathrm{z}) = \alpha_0 + \cdots + \alpha_{n-1}\mathrm{z}^{n-1} + \mathrm{z}^n$ of degree n, we denote by $C_\alpha \in \mathbb{R}^{n\times n}$ the standard companion matrix with the last row $(-\alpha_0, \ldots, -\alpha_{n-1})$, and $e_n \in \mathbb{R}^n$ is the last standard unit vector. If $x \in \mathbb{R}^n$ then $\|x\|^2 := x^\top x$ is the Euclidean norm. Finally, l_{2e}^n is the space of all sequences $x : \mathbb{N}_0 \to \mathbb{R}^n$, which are tacitly assumed to be extended as $x_t = 0$ for $t < 0$.

We follow the custom in robust control to express a linear system $x_{t+1} = Ax_t + Bu_t$, $y_t = Cx_t + Du_t$ for $t \in \mathbb{N}_0$ as

$$\begin{pmatrix} x_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_t \\ u_t \end{pmatrix} \quad \text{or} \quad y = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] u.$$

The latter notation is also used to represent the input-output map defined by the system. Moreover, the shorthand notation

$$\left(\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right) \xrightarrow{T} \left(\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right)$$

expresses that T is invertible with $\hat{A} = TAT^{-1}$, $\hat{B} = TB$, $\hat{C} = CT^{-1}$, $\hat{D} = D$. Finally, we abbreviate Kalman's

The first and third author are funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 - 390740016. They acknowledge the support by the Stuttgart Center for Simulation Science (SimTech).

Carsten W. Scherer and Tobias Holicki are with the Department of Mathematics, University of Stuttgart, Germany (email: {carsten.scherer,tobias.holicki}@imng.uni-stuttgart.de). Christian Ebenbauer is the Chair of Intelligent Control Systems, RWTH Aachen University, Germany (email: christian.ebenbauer@ic.rwth-aachen.de).

controllability matrix of the pair $(A, B) \in \mathbb{R}^{n\times(n+m)}$ by $\mathcal{K}(A, B) := (B, AB, \ldots, A^{n-1}B)$.

II. Optimization Algorithms as Feedback Systems

A. The Underlying Function Class

In this paper, we work with the class $\mathcal{S}_{m,L}$ of functions $f : \mathbb{R}^d \to \mathbb{R}$ that are L-smooth and m-strongly convex for $m > 0$ or just convex if $m = 0$. Among the various equivalent ways to express these conditions, the following most intuitive ones do not require any a priori assumptions on differentiability.

Definition 1: Let $L > m \geq 0$ and $q(x) = \frac{1}{2}\|x\|^2$ for $x \in \mathbb{R}^n$. Then $\mathcal{S}_{m,L}$ is the set of all $f : \mathbb{R}^d \to \mathbb{R}$ such that

$$f_m := f - mq \text{ and } f^L := Lq - f \text{ are convex.}$$

Moreover, let $\mathcal{S}^0_{m,L} := \{f \in \mathcal{S}_{m,L} \mid f(0) = 0,\ \nabla f(0) = 0\}$.

The latter makes sense since it can be shown that any $f \in \mathcal{S}_{m,L}$ is differentiable [17]. For the gradients of f, f_m, f^L and any $z \in \mathbb{R}^d$, we record the relation

$$\begin{pmatrix}\nabla f^L(z)\\ \nabla f_m(z)\end{pmatrix} = \begin{pmatrix} LI_d & -I_d\\ -mI_d & I_d\end{pmatrix}\begin{pmatrix} z\\ \nabla f(z)\end{pmatrix}. \quad (1)$$

Among the many known inequalities for $f \in \mathcal{S}_{m,L}$, the one in the following lemma stands out in allowing for a direct construction of integral quadratic constraints. It is also underlying the proof of [5, Lemma 8] and, if evaluated at finitely many points, identical to the central inequality in [13, Theorem 4].

Lemma 2: Let $f \in \mathcal{S}_{m,L}$. Then the function $V(x) := (L - m)f_m(x) - q(\nabla f_m(x))$ satisfies

$$V(u) - V(y) \leq \nabla f_m(u)^\top[\nabla f^L(u) - \nabla f^L(y)] \quad (2)$$

for all $u, y \in \mathbb{R}^d$. If $f \in \mathcal{S}^0_{m,L}$ then V has a global minimum at 0 with value 0, i.e., $0 = V(0) \leq V(x)$ for all $x \in \mathbb{R}^d$.

To support the reader's intuition, we note that (2) for $m = 0$ and $L \to \infty$ boils down to the subgradient inequality for convex functions. Since $f_0(x) = f(x)$, $\nabla f_0(x) = \nabla f(x)$ and $\frac{1}{L}\nabla f^L(x) \to x$ for $L \to \infty$, we infer $\frac{1}{L}V(x) \to f(x)$. After dividing (2) by L, we indeed obtain for $m = 0$ and $L \to \infty$ the inequality $f(u) - f(y) \leq \nabla f(u)^T(u - y)$.

B. Optimization Algorithms and Systems

For $f \in \mathcal{S}_{m,L}$, we recall that the optimization problem

$$\inf_{z\in\mathbb{R}^d} f(z) \quad (3)$$

does admit a unique solution z_* [18]. It is also well-known that the gradient descent algorithm

$$z_{t+1} = z_t - \alpha\nabla f(z_t) \quad (4)$$

for $\alpha = \frac{2}{m+L}$ generates a sequence with $\lim_{t\to\infty} z_t = z_*$. We denote the iteration index by "t" since we want to view (4) as a discrete-time dynamical system for t on the time axis $\mathbb{N}_0$. Even more, (4) can be viewed as the feedback interconnection of the LTI system $x_{t+1} = x_t - \alpha w_t$, $z_t = x_t$ with the static nonlinearity

$$w_t = \nabla f(z_t) \quad (5)$$

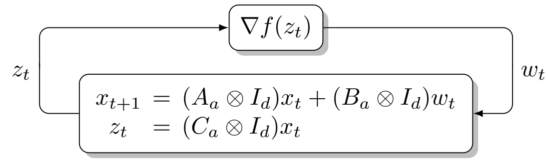

Fig. 1. Feedback representation of optimization algorithms.

for $t \in \mathbb{N}_0$, where $x_t \in \mathbb{R}^d$, $w_t \in \mathbb{R}^d$ and $z_t \in \mathbb{R}^d$ are the state, the input and the output of the linear system. This linear system can actually be expressed as

$$x_{t+1} = (A_a \otimes I_d)x_t + (B_a \otimes I_d)w_t,\ z_t = (C_a \otimes I_d)x_t \quad (6)$$

where $A_a = 1$, $B_a = -\alpha$ and $C_a = 1$. Here $\otimes$ denotes the Kronecker product, which is convenient to compactly describe general algorithms in the sequel. In control, the feedback interconnection (5)-(6) is a so-called Lur'e system. A block diagram of this interconnection is depicted in Fig. 1.

Accelerated versions of gradient descent include a so-called momentum term. A prominent example is Nesterov's algorithm with a description

$$v_{t+2} = v_{t+1} + \beta(v_{t+1} - v_t) - \alpha\nabla f(v_{t+1} + \gamma(v_{t+1} - v_t))$$

for suitable real parameters α, β and $\gamma = \beta$ [1], or the triple momentum version with $\gamma \neq \beta$ [7]. This is nothing but

$$v_{t+2} = v_{t+1} + \beta(v_{t+1} - v_t) - \alpha w_t,\ z_t = v_{t+1} + \gamma(v_{t+1} - v_t)$$

in feedback with (5). Moreover, the latter second order system can be routinely translated into the first-order description (6) with state $x_t = \mathrm{col}(v_{t+1}, v_t)$ and the matrices

$$\begin{pmatrix} A_a & B_a\\ C_a & 0\end{pmatrix} = \left(\begin{array}{cc|c} 1+\beta & -\beta & -\alpha\\ 1 & 0 & 0\\ \hline 1+\gamma & -\gamma & 0\end{array}\right). \quad (7)$$

Hence, also Nesterov's recursion can be expressed as (5)-(6).

These observations provide a strong motivation for investigating the stability properties of the feedback interconnection (5)-(6) for general matrices (A_a, B_a, C_a).

C. Minimal Convergence Requirement

In view of the goal to solve (3), it is a minimal requirement that, for any $f \in \mathcal{S}_{m,L}$ and any initial condition, the signal z_t of the interconnection (5)-(6) should converge to a limit which satisfies the following first order necessary and sufficient condition for optimality:

$$\lim_{t\to\infty} z_t = z_* \text{ with } \nabla f(z_*) = 0. \quad (8)$$

By (5), this implies $w_* := \lim_{t\to\infty} w_t = 0$. If (A_a, C_a) is detectable, we infer $\lim_{t\to\infty} x_t = x_*$ and the limit (x_*, w_*, z_*) satisfies the equilibrium equations

$$x_* = (A_a \otimes I_d)x_*,\quad z_* = (C_a \otimes I_d)x_*,\quad w_* = 0. \quad (9)$$

If we pick $f(z) = \frac{1}{2}m\|z - z_*\|^2$ for $z_* \in \mathbb{R}^d$ with $z_* \neq 0$, the corresponding solution x_* of (9) does not vanish, which in turn shows that 1 is an eigenvalue of A_a. For example in Nesterov's algorithm, this is indeed true since the elements in each row of A_a in (7) sum up to one. As a result, the minimal requirement enforces structural constraints on the algorithm parameters (A_a, B_a, C_a).

In general, we argue in [15, Section 2.2] that (A_a, C_a) can be assumed to be detectable without loss of generality. Then the minimal requirement implies that the algorithm parameters must admit, after a possible state-coordinate change, the structure

$$\begin{pmatrix} A_a & B_a \\ C_a & 0 \end{pmatrix} = \left(\begin{array}{cc|c} A_c & B_c & 0 \\ 0 & 1 & 1 \\ \hline C_c & D_c & 0 \end{array}\right) \text{ with } \det \begin{pmatrix} A_c - I & B_c \\ C_c & D_c \end{pmatrix} \neq 0. \tag{10}$$

The first relation means that the system described with (A_a, B_a, C_a) is the series interconnection of

$$\begin{pmatrix} x^c_{t+1} \\ z_t \end{pmatrix} = \begin{pmatrix} A_c \otimes I_d & B_c \otimes I_d \\ C_c \otimes I_d & D_c \otimes I_d \end{pmatrix} \begin{pmatrix} x^c_t \\ y_t \end{pmatrix} \tag{11}$$

and the discrete-time integrator

$$\begin{pmatrix} x^s_{t+1} \\ y_t \end{pmatrix} = \begin{pmatrix} I_d & I_d \\ I_d & 0 \end{pmatrix} \begin{pmatrix} x^s_t \\ w_t \end{pmatrix} \tag{12}$$

with the transfer matrix $\frac{1}{z-1} I_d$. The second condition in (10) expresses the fact that the pole $z = 1$ of the integrator in the corresponding product of transfer matrices is not canceled.

In other words, the algorithm's parameters must contain a model of the integrator. Although not surprising from the perspective of control, this fact has only been recently clearly emphasized in [11], [15] in the realm of algorithm analysis.

As an illustration, for (7) we note that

$$\left(\begin{array}{cc|c} 1+\beta & -\beta & -\alpha \\ 1 & 0 & 0 \\ \hline 1+\gamma & -\gamma & 0 \end{array}\right) \xrightarrow{\begin{pmatrix} 0 & 1 \\ -\frac{1}{\alpha} & \frac{\beta}{\alpha} \end{pmatrix}} \left(\begin{array}{cc|c} \beta & -\alpha & 0 \\ 0 & 1 & 1 \\ \hline \beta(1+\gamma)-\gamma & -\alpha(1+\gamma) & 0 \end{array}\right).$$

Let us now pinpoint the two essential consequences in case that (A_a, B_a, C_a) does indeed have the structure (10):

1) If (8) is satisfied for (5)-(6) and all $f \in \mathcal{S}^0_{m,L}$, then (8) holds for (5)-(6) and all $f \in \mathcal{S}_{m,L}$.
2) The interconnection (5)-(6) is a controlled uncertain system as familiar in robust control.

To see 1), we assign to any $f \in \mathcal{S}_{m,L}$ the function f_*, defined with the unique $z_* \in \mathbb{R}^d$ satisfying $\nabla f(z_*) = 0$ as

$$f_*(z) := f(z + z_*) - f(z_*) \quad \text{for} \quad z \in \mathbb{R}^d.$$

Since $f_*(0) = 0$ and $\nabla f_*(0) = 0$, we note that $f_* \in \mathcal{S}^0_{m,L}$. If x_* denotes the unique solution of (9), the system (6) can be equivalently transformed into

$$\begin{aligned} x_{t+1} - x_* &= (A_a \otimes I_d)(x_t - x_*) + (B_a \otimes I_d) w_t, \\ z_t - z_* &= (C_a \otimes I_d)(x_t - x_*). \end{aligned}$$

To be precise, the trajectories (x, w, z) of the interconnection (5)-(6) are in one-to-one correspondence via $\tilde{x} = x - x_*$, $\tilde{w} = w$, $\tilde{z} = z - z_*$ with the trajectories $(\tilde{x}, \tilde{w}, \tilde{z})$ of (6) in feedback with

$$w_t = \nabla f_*(z_t). \tag{13}$$

This proves 1). Even stronger, it shows that exponential stability of the equilibrium $(\widetilde{x}_*, \widetilde{w}_*, \widetilde{z}_*) = (0, 0, 0)$ of the interconnection (6), (13) is equivalent to exponential stability of $(x_*, 0, z_*)$ of the loop (5)-(6).

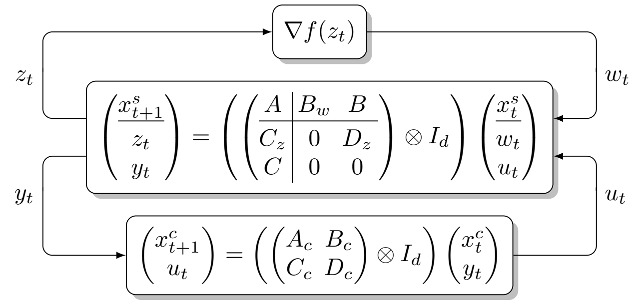

Fig. 2. General algorithm synthesis configuration.

Property 2) is seen by redrawing Fig. 1 as in Fig. 2 with

$$\left(\begin{array}{c|cc} A & B_w & B \\ \hline C_z & 0 & D_z \\ C & 0 & 0 \end{array}\right) = \left(\begin{array}{c|cc} 1 & 1 & 0 \\ \hline 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right), \tag{14}$$

by recalling (11)-(12) and using the auxiliary signal $u_t := z_t$. Then we indeed recognize the uncertainty ∇f in the class $\nabla \mathcal{S}^0_{m,L}$, the to-be-controlled plant defined with (14) and the controller (11). We also emphasize the simplicity of this plant in the realm of algorithms! Still, it is relevant to stress that all our subsequent analysis and synthesis results even apply to general LTI plants as in Fig. 2.

III. Exponential Stability and Passivity

Let us now turn to the development of a test which ensures that the loop (5)-(6) is exponentially stable for all $f \in \mathcal{S}^0_{m,L}$.

One way is based on the exponential signal weighting map

$$T_{\rho^{-1}}(z_0, z_1, z_2, \ldots) = (z_0, \rho^{-1} z_1, \rho^{-2} z_2, \ldots) \tag{15}$$

for some $\rho \in (0, 1]$ [19]. Clearly, $T_{\rho^{-1}}$ is linear and invertible with $T^{-1}_{\rho^{-1}} = T_\rho$. It is then easily checked that the set of trajectories (x, w, z) of (6) are in one-to-one correspondence with trajectories $(\bar{x}, \bar{w}, \bar{z})$ of the system

$$\begin{aligned} \bar{x}_{t+1} &= (\bar{A}_a \otimes I_d)\bar{x}_t + (\bar{B}_a \otimes I_d)\bar{w}_t, \\ \bar{z}_t &= (C_a \otimes I_d)\bar{x}_t \end{aligned} \tag{16}$$

under the signal transformations $\bar{x} = T_{\rho^{-1}} x$, $\bar{w} = T_{\rho^{-1}} w$, and $\bar{z} = T_{\rho^{-1}} z$ and with the ρ-scaled matrices $\begin{pmatrix} \bar{A}_a & \bar{B}_a \end{pmatrix} := \rho^{-1} \begin{pmatrix} A & B \end{pmatrix}$. Similarly, (5) translates into

$$\bar{w}_t = \bar{F}(t, \bar{z}_t) \tag{17}$$

with the static time-varying map $\bar{F}$ associated to ∇f through

$$\bar{F}(t, z) := \rho^{-t} \nabla f(\rho^t z) \quad \text{for} \quad t \in \mathbb{N}_0, \; z \in \mathbb{R}^d. \tag{18}$$

The bars should remind us of the fact that that $(\bar{A}_a, \bar{B}_a)$ and $\bar{F}$ depend on ρ.

As a consequence, if the transformed loop is Lyapunov stable in the sense of $\|\bar{x}_t\| \leq K \|\bar{x}_0\|$ for all $t \in \mathbb{N}_0$, one can conclude that the original loop is exponentially stable with rate ρ in the sense of $\|x_t\| \leq K \rho^{-t} \|x_0\|$ for all $t \in \mathbb{N}_0$. This motivates to develop a robust stability test for the transformed interconnection (16)-(17).

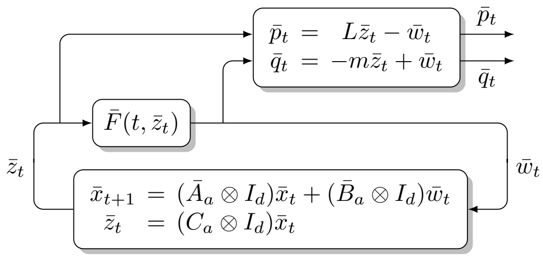

Fig. 3. Block diagram of transformed loop (16)-(17) with signals filtered according to (22).

We start by deriving what is called a valid integral quadratic constraint (IQC) for the nonlinearity by exploiting Lemma 2. If $f \in \mathcal{S}^0_{m,L}$, we conclude from (2) for $y = 0$ that

$$V(u) \leq \nabla f_m(u)^\top \nabla f^L(u) \quad \text{for all} \quad u \in \mathbb{R}^d. \tag{19}$$

Since V is nonnegative and $\rho > 0$, this trivially implies

$$0 \leq \rho^{-t}\nabla f_m(\rho^t z_t)^\top(\rho^{-t}\nabla f^L(\rho^t z_t))$$

for all $z \in l^d_{2e}$ and $t \in \mathbb{N}_0$. Summation leads to the IQC

$$0 \leq \sum_{t=0}^{T-1} \bar{F}_m(t, \bar{z}_t)^\top \bar{F}^L(t, \bar{z}_t) \quad \text{for all} \quad T \in \mathbb{N} \tag{20}$$

and all sequences $\bar{z} \in l^d_{2e}$, where $\bar{F}_m$ and $\bar{F}^L$ are also defined according to (18). Note that the misnomer "IQC" results from a similar concept for continuous-time systems, in which summation is replaced by integration [20]. With (1) we infer

$$\begin{pmatrix} \bar{F}^L(t, \bar{z}_t) \\ \bar{F}_m(t, \bar{z}_t) \end{pmatrix} = \begin{pmatrix} LI_d & -I_d \\ -mI_d & I_d \end{pmatrix} \begin{pmatrix} \bar{z}_t \\ \bar{F}(t, \bar{z}_t) \end{pmatrix}. \tag{21}$$

This motivates to introduce the static filter

$$\begin{pmatrix} \bar{p}_t \\ \bar{q}_t \end{pmatrix} = \begin{pmatrix} LI_d & -I_d \\ -mI_d & I_d \end{pmatrix} \begin{pmatrix} \bar{z}_t \\ \bar{w}_t \end{pmatrix}. \tag{22}$$

If filtering the input-output signals of the nonlinearity (17) accordingly, (21) shows $\bar{p}_t = \bar{F}^L(t, \bar{z}_t)$ and $\bar{q}_t = \bar{F}_m(t, \bar{z}_t)$. Then (20) reads as $\sum_{t=0}^{T-1} \bar{q}_t^\top \bar{p}_t \geq 0$ for all $T \in \mathbb{N}$ and can be interpreted as a passivity property [19] for the outputs of (22) driven by the signals in (17).

In view of Fig. 3 and motivated by the passivity theorem, we expect that stability of (16)-(17) is guaranteed in case that $\sum_{t=0}^{T-1} \bar{q}_t^\top \bar{p}_t < 0$ holds for all $T \in \mathbb{N}$ along the input-output trajectories of the linear system (16) filtered with (22).

To make this precise, we start by emphasizing that the stability test itself is formulated for $d = 1$, while the conclusions are drawn for arbitrary dimensions $d \in \mathbb{N}$. Throughout the paper we slightly abuse the notation and do not indicate the dependence of system signals on d.

For $d = 1$, we note that the input-output signals of (16) filtered with (22) satisfy $\bar{w}_t = mC_a\bar{x}_t + \bar{q}_t$ and $\bar{p}_t = LC_a\bar{x}_t - \bar{w}_t$. With an identical state-trajectory, we hence infer

$$\begin{pmatrix} \bar{x}_{t+1} \\ \bar{p}_t \end{pmatrix} = \underbrace{\begin{pmatrix} \bar{A}_a + \bar{B}_a m C_a & \bar{B}_a \\ (L-m)C_a & -1 \end{pmatrix}}_{\begin{pmatrix} \mathcal{A} & \mathcal{B} \\ \mathcal{C} & \mathcal{D} \end{pmatrix}} \begin{pmatrix} \bar{x}_t \\ \bar{q}_t \end{pmatrix}. \tag{23}$$

As a consequence, also for a general $d \in \mathbb{N}$, the trajectories of (16) filtered with (22) satisfy

$$\begin{pmatrix} \bar{x}_{t+1} \\ \bar{p}_t \end{pmatrix} = \begin{pmatrix} \mathcal{A} \otimes I_d & \mathcal{B} \otimes I_d \\ \mathcal{C} \otimes I_d & \mathcal{D} \otimes I_d \end{pmatrix} \begin{pmatrix} \bar{x}_t \\ \bar{q}_t \end{pmatrix}. \tag{24}$$

This leads to our first analysis result, which involves a passivity property of the system (23) and hence also of (24).

Theorem 3: Suppose there exists some $\mathcal{X} = \mathcal{X}^\top$ with

$$\mathcal{X} \succ 0 \text{ and } \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & 0 \end{pmatrix}^\top \begin{pmatrix} \mathcal{X} & 0 \\ 0 & -\mathcal{X} \end{pmatrix} \begin{pmatrix} \mathcal{A} & \mathcal{B} \\ I & 0 \end{pmatrix} + \\ + \begin{pmatrix} \mathcal{C} & \mathcal{D} \\ 0 & 1 \end{pmatrix}^\top \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \mathcal{C} & \mathcal{D} \\ 0 & 1 \end{pmatrix} \prec 0. \tag{25}$$

Then there exists a constant K such that, for any $f \in \mathcal{S}^0_{m,L}$, all trajectories of the original loop (5)-(6) satisfy

$$\|x_t\| \leq K\rho^t \|x_0\| \quad \text{for all} \quad t \in \mathbb{N}_0. \tag{26}$$

For $\rho = 1$, it is also assured that $\lim_{t\to\infty} x_t = 0$ holds true.

Before addressing the practical application of this robust stability test, we discuss how to substantially improve it by the incorporation of so-called stability multipliers.

IV. Dynamic Integral Quadratic Constraints

It is a classical idea [21], [22] to improve Theorem 3 by imposing a passivity condition after filtering the signal $\bar{p}$ in (22) with a causal and stable time-invariant system. In this context, such a filter is often called a stability multiplier [19].

In fact, passing the signal $\bar{p}_t = \bar{F}^L(t, \bar{z}_t)$ through a delay of time $\nu \in \mathbb{N}$ leads to $\bar{r}_t = \bar{F}^L(t-\nu, \bar{z}_{t-\nu})$ (where we recall our convention that $\bar{z}_{t-\nu} = 0$ and hence $\bar{r}_{t-\nu} = 0$ for $t < \nu$.) The following IQC incorporates this delayed signal $\bar{r}_t$ and is, again, a rather immediate consequence of Lemma 2.

Lemma 4: Let $f \in \mathcal{S}^0_{m,L}$, $\rho \in (0, 1]$ and $\nu \in \mathbb{N}$. Then

$$0 \leq \sum_{t=0}^{T-1} \bar{F}_m(t, \bar{z}_t)^\top (\bar{F}^L(t, \bar{z}_t) - \rho^\nu \bar{F}^L(t-\nu, \bar{z}_{t-\nu})) \tag{27}$$

holds for all $T \in \mathbb{N}$ and all signals $\bar{z} \in l^d_{2e}$.

A conic combination of (20) and (27) for $\nu \in \mathbb{N}$ leads to the IQC with more general filters in the following lemma.

Lemma 5: Let $f \in \mathcal{S}^0_{m,L}$, $\rho \in (0, 1]$ and suppose that $\lambda_0, \lambda_1, \ldots \in \mathbb{R}$ satisfy

$$\lambda_\nu \leq 0 \quad \text{for all} \quad \nu \in \mathbb{N} \quad \text{and} \quad \sum_{\nu=0}^{\infty} \rho^{-\nu}\lambda_\nu > 0. \tag{28}$$

For $\bar{z} \in l^d_{2e}$, let $\bar{p}_t = \bar{F}^L(t, \bar{z}_t)$ be passed through the filter

$$\bar{r}_t = \sum_{\nu=0}^{t} \lambda_\nu \bar{p}_{t-\nu} \quad \text{for} \quad t \in \mathbb{N}_0. \tag{29}$$

With $\bar{q}_t := \bar{F}_m(t, \bar{z}_t)$, the signals $\bar{r}, \bar{q}$ then satisfy the IQC

$$0 \leq \sum_{t=0}^{T-1} \bar{q}_t^\top \bar{r}_t \quad \text{for all} \quad T \in \mathbb{N}. \tag{30}$$

Due to the incorporation of the dynamic filter (29), we call (30) a dynamic IQC. Note that (28) implies $\lambda_0 > 0$.

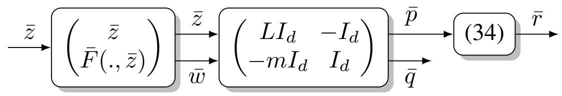

Fig. 4. Signals in dynamic IQC of nonlinearity.

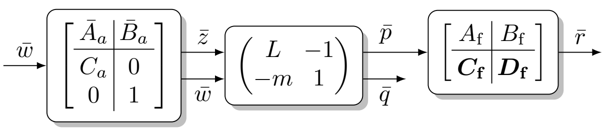

Fig. 5. Signals in filtered linear system.

The infinite impulse response filters (29) are subject to the infinite number of constraints (28). To overcome this trouble for the purpose of computations, we proceed with filters that have a state-space realization $(A_{\rm f}, B_{\rm f}, \boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ with a fixed pole-pair $(A_{\rm f}, B_{\rm f}) \in \mathbb{R}^{l\times(l+1)}$ and free filter coefficients collected in $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}}) \in \mathbb{R}^{1\times l}\times\mathbb{R}$ such that

$$\lambda_0 = \boldsymbol{D_{\rm f}} \quad\text{and}\quad \lambda_{\nu+1} = \boldsymbol{C_{\rm f}} A_{\rm f}^{\nu} B_{\rm f} \text{ for } \nu\in\mathbb{N}_0. \tag{31}$$

The boldface notation reminds us of the fact that $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ is a decision variable in the subsequent stability test. For a suitable choice of $(A_{\rm f}, B_{\rm f})$, we now establish that the infinitely many constraints (28) on the Markov parameters (31) can be expressed by a finite number of linear ones.

Lemma 6: Fix $(A_{\rm f}, B_{\rm f}) = (C_\alpha, e_l)$ with a real polynomial $\alpha(\mathrm{z}) = \alpha_0+\alpha_1\mathrm{z}+\cdots+\alpha_{l-1}\mathrm{z}^{l-1}+\mathrm{z}^l$ of degree $l\in\mathbb{N}$ having all its roots in $\mathbb{D}_\rho := \{\mathrm{z}\in\mathbb{C} \mid |\mathrm{z}|<\rho\}$ and with coefficients satisfying $\alpha_0,\ldots,\alpha_{l-1}\le 0$.

Then, for any pair $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}}) \in \mathbb{R}^{1\times l}\times\mathbb{R}$, the Markov parameters (31) satisfy the constraints (28) iff

$$\boldsymbol{C_{\rm f}}\mathcal{K}(A_{\rm f}, B_{\rm f}) \le 0 \quad\text{and}\quad \boldsymbol{D_{\rm f}} + \boldsymbol{C_{\rm f}}(\rho I - A_{\rm f})^{-1}B_{\rm f} > 0. \tag{32}$$

Moreover, (32) implies that $\boldsymbol{D_{\rm f}} > 0$ and that all eigenvalues of $A_{\rm f} - B_{\rm f}\boldsymbol{D_{\rm f}}^{-1}\boldsymbol{C_{\rm f}}$ are located in $\mathbb{D}_\rho$.

To summarize, for a fixed polynomial α of degree l as in Lemma 6, we work from now on with the filter matrices

$$A_{\rm f} := C_\alpha,\quad B_{\rm f} := e_l,\quad \boldsymbol{C_{\rm f}}\in\mathbb{R}^{1\times l},\quad \boldsymbol{D_{\rm f}}\in\mathbb{R} \tag{33}$$

such that $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ satisfies the constraints (32). It is then assured that the trajectories of (17) filtered by (22) and

$$\begin{pmatrix}\xi_{t+1}\\ \bar r_t\end{pmatrix} = \begin{pmatrix} A_{\rm f}\otimes I_d & B_{\rm f}\otimes I_d\\ \boldsymbol{C_{\rm f}}\otimes I_d & \boldsymbol{D_{\rm f}}\otimes I_d\end{pmatrix}\begin{pmatrix}\xi_t\\ \bar p_t\end{pmatrix},\quad \xi_0 = 0 \tag{34}$$

(see Fig. 4) satisfy the passivity condition (30) for all $f\in\mathcal{S}^0_{m,L}$. In this way, we have identified a whole nicely parameterized convex family of valid dynamic IQCs for the nonlinearity (17) involving the filters or multipliers (34).

V. Robust Stability Analysis with Dynamic IQCs

In alignment with Sec. III, the robust stability test with dynamic IQCs is now formulated for the correspondingly filtered linear system as depicted in Fig. 5. If recalling (23), we are lead to the system

$$\bar r = \left[\begin{array}{c|c} A_{\rm f} & B_{\rm f}\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}\end{array}\right]\left[\begin{array}{c|c}\bar A_a + \bar B_a m C_a & \bar B_a\\ \hline (L-m)C_a & -1\end{array}\right]\bar q.$$

A realization of this series interconnection is given as

$$\left(\begin{array}{c|c}\mathcal{A} & \mathcal{B}\\ \hline \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}}\end{array}\right) := \left(\begin{array}{cc|c} A_{\rm f} & B_{\rm f}(L-m)C_a & -B_{\rm f}\\ 0 & \rho^{-1}(A_a + B_a m C_a) & \rho^{-1}B_a\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}(L-m)C_a & -\boldsymbol{D_{\rm f}}\end{array}\right). \tag{35}$$

Theorem 7: Fix some $\rho\in(0,1]$ and $(A_{\rm f}, B_{\rm f})$ with the properties in Lemma 6. If there exist an $\mathcal{X} = \mathcal{X}^\top$ and a filter parameter $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ with (32) such that (25) is satisfied, then the same conclusions can be drawn as in Theorem 3.

We use boldface letters to highlight that $(\boldsymbol{\mathcal{C}}\ \boldsymbol{\mathcal{D}})$ depends affinely on the decision variables $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$. Hence (25) constitutes a genuine LMI in the variables $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ and $\mathcal{X}$.

We emphasize that Theorem 7 encompasses Theorem 3 for $l=0$, which means that $A_{\rm f}, B_{\rm f}, \boldsymbol{C_{\rm f}}$ are empty matrices; since the LMI (25) is homogeneous in $(\mathcal{X}, \boldsymbol{\mathcal{D}})$, we can indeed fix $\boldsymbol{\mathcal{D}} = \boldsymbol{D_{\rm f}}$ to the value 1 without loss of generality.

After picking the targeted convergence rate $\rho\in(0,1]$ and either $l=0$ or a characteristic filter polynomial α of degree $l\in\mathbb{N}$ as in Lemma 6, feasibility of the LMIs (32) and (25) (defined with the matrices (35)) thus guarantees exponential stability of (5)-(6) with rate ρ for any $f\in\mathcal{S}^0_{m,L}$ and any dimension $d\in\mathbb{N}$. A more detailed practical recipe for how to apply Theorem 7 can be extracted from Sec. VI.

Note that the dimension of the LMI (25) is independent from $d\in\mathbb{N}$. Moreover, it is remarkable that even only a few filter states ($l = 1,2,3$) can substantially improve the test over $l=0$. For $\alpha(\mathrm{z}) = \mathrm{z}^l$ with $l=1$, Theorem 7 encompasses the algorithm analysis approach in [5].

Related IQC results to ensure robust exponential loop stability can be found in [23], [24], while a more general dissipativity-based robustness analysis framework is exposed in the recent survey article [25] and the references therein.

VI. Convex Algorithm Synthesis

Let us now turn to the design problem for the interconnection in Fig. 2 with the general plant

$$\begin{pmatrix} z\\ y\end{pmatrix} = \left[\begin{array}{c|cc} A & B_w & B\\ \hline C_z & 0 & D_z\\ C & 0 & 0\end{array}\right]\begin{pmatrix} w\\ u\end{pmatrix} \tag{36}$$

where $A\in\mathbb{R}^{n\times n}$. Again, we slightly abuse notation by not indicating the dependence of the signals on d. As for analysis, we first construct the modified system description in order to formulate the key synthesis result. To this end, we transform all signals of (36) according to (15) to get

$$\begin{pmatrix}\bar z\\ \bar y\end{pmatrix} = \left[\begin{array}{c|cc}\rho^{-1}A & \rho^{-1}B_w & \rho^{-1}B\\ \hline C_z & 0 & D_z\\ C & 0 & 0\end{array}\right]\begin{pmatrix}\bar w\\ \bar u\end{pmatrix}.$$

Filtering the signal $\mathrm{col}(\bar z, \bar w)$ as in (22) with $d=1$ leads to

$$\begin{pmatrix}\bar p\\ \bar y\end{pmatrix} = \underbrace{\left[\begin{array}{c|cc}\rho^{-1}(A+B_w m C_z) & \rho^{-1}B_w & \rho^{-1}(B+B_w m D_z)\\ \hline (L-m)C_z & -1 & (L-m)D_z\\ C & 0 & 0\end{array}\right]}_{\left[\begin{array}{c|cc}\tilde A & \tilde B_w & \tilde B\\ \hline \tilde C_z & \tilde D_{zw} & \tilde D_z\\ \tilde C & 0 & 0\end{array}\right]}\begin{pmatrix}\bar q\\ \bar u\end{pmatrix}.$$

This plant filtered with (34) for $d=1$ reads as

$$\begin{pmatrix}\bar r\\ \bar y\end{pmatrix} = \underbrace{\left[\begin{array}{cc|cc} A_{\rm f} & B_{\rm f}\tilde C_z & B_{\rm f}\tilde D_{zw} & B_{\rm f}\tilde D_z\\ 0 & \tilde A & \tilde B_w & \tilde B\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}\tilde C_z & \boldsymbol{D_{\rm f}}\tilde D_{zw} & \boldsymbol{D_{\rm f}}\tilde D_z\\ 0 & C & 0 & 0\end{array}\right]}_{\left[\begin{array}{c|cc}\hat A & \hat B_w & \hat B\\ \hline \boldsymbol{\hat C}_z & \boldsymbol{\hat D}_{zw} & \boldsymbol{\hat D}_z\\ \hat C & 0 & 0\end{array}\right]} \begin{pmatrix}\bar q\\ \bar u\end{pmatrix}. \tag{37}$$

For the system (37), we now pick a controller

$$\bar u = \left[\begin{array}{c|c}\hat A_c & \hat B_c\\ \hline \hat C_c & \hat D_c\end{array}\right]\bar y. \tag{38}$$

Then the resulting interconnection admits the description

$$\bar r = \left[\begin{array}{cc|c}\hat A + \hat B\hat D_c\hat C & \hat B\hat C_c & \hat B_w\\ \hat B_c\hat C & \hat A_c & 0\\ \hline \boldsymbol{\hat C}_z + \boldsymbol{\hat D}_z\hat D_c\hat C & \boldsymbol{\hat D}_z\hat C_c & \boldsymbol{\hat D}_{zw}\end{array}\right]\bar q =: \left[\begin{array}{c|c}\mathcal{A} & \mathcal{B}\\ \hline \boldsymbol{\mathcal{C}} & \boldsymbol{\mathcal{D}}\end{array}\right]\bar q, \tag{39}$$

where all the bold matrices depend affinely on $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$.

If (38) is a controller for which the controlled system (39) satisfies the hypotheses of Theorem 7, it is not difficult to verify that the original plant (36) controlled with

$$u = \left[\begin{array}{c|c}\rho\hat A_c & \rho\hat B_c\\ \hline \hat C_c & \hat D_c\end{array}\right] y \tag{40}$$

(see Figure 2) satisfies (26) for any x_0 and any $f\in\mathcal{S}^0_{m,L}$.

Let us now recap a slight variant of a seminal result obtained in [26], [27], a convex solution for the design problem if $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$ satisfying (32) is *held fixed*. To this end, we pick so-called annihilator matrices $\hat U$ and $\hat V$ with

$$\hat U = \mathrm{diag}(\hat C_\perp\ 1)\quad\text{and}\quad \hat V^\top = \begin{pmatrix}\hat B^\top & \boldsymbol{\hat D_z}^\top\end{pmatrix}_\perp \tag{41}$$

where $M_\perp$ means that the columns of this matrix form a basis of the kernel of the matrix M. Then there exist a controller (38) for (37) such that the closed loop system (39) renders the analysis LMIs in Theorem 7 feasible iff there exist symmetric matrices X and $\hat Y$ which satisfy

$$[\bullet]^\top\left(\begin{array}{cc|cc} X & 0 & 0 & 0\\ 0 & -X & 0 & 0\\ \hline 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{array}\right)\left[\begin{pmatrix}\hat A & \hat B_1\\ I & 0\\ \boldsymbol{\hat C}_z & \boldsymbol{\hat D}_{zw}\\ 0 & 1\end{pmatrix}\hat U\right]\prec 0, \tag{42}$$

$$\left[\hat V\left(\begin{array}{cc|cc}-I & \hat A & 0 & \hat B_1\\ 0 & \boldsymbol{\hat C}_z & -1 & \boldsymbol{\hat D}_{zw}\end{array}\right)\right]\left(\begin{array}{cc|cc}\hat Y & 0 & 0 & 0\\ 0 & -\hat Y & 0 & 0\\ \hline 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{array}\right)[\bullet]^\top \succ 0, \tag{43}$$

$$\begin{pmatrix}\hat Y & I\\ I & X\end{pmatrix}\succ 0. \tag{44}$$

For reasons of space, we use the bullet notation to indicate that one should substitute (on the left/right) the respective matrix in square brackets (on the right/left) to render the inequalities symmetric. For fixed $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$, these constraints are affine in X and $\hat Y$. However, this nice structural property is destroyed for (43) if viewing $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$ as an additional decision variable.

To overcome this trouble, we note that the subsystem $\bar u\to\bar r$ of (37) is actually given by

$$\bar r = \left[\begin{array}{c|c}A_{\rm f} & B_{\rm f}\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}\end{array}\right]\left[\begin{array}{c|c}\tilde A & \tilde B\\ \hline \tilde C_z & \tilde D_z\end{array}\right]\bar u = \left[\begin{array}{c|c}\tilde A & \tilde B\\ \hline \tilde C_z & \tilde D_z\end{array}\right]\left[\begin{array}{c|c}A_{\rm f} & B_{\rm f}\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}\end{array}\right]\bar u, \tag{45}$$

a series interconnection of two commuting SISO systems. As the key to convexification, we exploit the fact that this commutation property is reflected by a state-coordinate change for the corresponding natural realizations as

$$\left[\begin{array}{cc|c}A_{\rm f} & B_{\rm f}\tilde C_z & B_{\rm f}\tilde D_z\\ 0 & \tilde A & \tilde B\\ \hline \boldsymbol{C_{\rm f}} & \boldsymbol{D_{\rm f}}\tilde C_z & \boldsymbol{D_{\rm f}}\tilde D_z\end{array}\right]\xrightarrow{\boldsymbol{T}}\left[\begin{array}{cc|c}A_{\rm f} & 0 & B_{\rm f}\\ \tilde B\boldsymbol{C_{\rm f}} & \tilde A & \tilde B\boldsymbol{D_{\rm f}}\\ \hline \tilde D_z\boldsymbol{C_{\rm f}} & \tilde C_z & \tilde D_z\boldsymbol{D_{\rm f}}\end{array}\right]$$

with the specifically structured transformation matrix

$$\boldsymbol{T} = \begin{pmatrix}L^{-1} & -L^{-1}K\\ \boldsymbol{N}L^{-1} & \boldsymbol{M}-\boldsymbol{N}L^{-1}K\end{pmatrix} = \begin{pmatrix}T_{\rm f}\\ \boldsymbol{\tilde T}\end{pmatrix}. \tag{46}$$

An analogous commutation property is used for convexification in [15] based on the Youla-Parametrization and in the state-space approach of [16]. The latter is confined to $\alpha(\mathrm{z}) = \mathrm{z}^l$ for the characteristic polynomial of $A_{\rm f}$, which induces some limitation on the plant (37). None are required in the next result due to the novel flexibility of choosing α.

Theorem 8: Pick α as in Lemma 6 such that the eigenvalues of $A_{\rm f}$ are different from the eigenvalues of $\tilde A$ and from the zeros of $\tilde C_z(\mathrm{z}I-\tilde A)^{-1}\tilde B + \tilde D_z$. With the solutions $K, L, \boldsymbol{M}$ and $\boldsymbol{N}$ of the linear equations

1) $A_{\rm f}K - K\tilde A + B_{\rm f}\tilde C_z = 0$,
2) $L\,\mathcal{K}(A_{\rm f},B_{\rm f}) = \mathcal{K}(A_{\rm f}, B_{\rm f}\tilde D_z - K\tilde B)$,
3) $\boldsymbol{M}\alpha(\tilde A) = \boldsymbol{D_{\rm f}}\alpha(\tilde A) + (\boldsymbol{C_{\rm f}}\otimes I_l)\mathrm{col}(I,\tilde A,\ldots,\tilde A^{l-1})$,
4) $\tilde A\boldsymbol{N} - \boldsymbol{N}A_{\rm f} + \tilde B\boldsymbol{C_{\rm f}} = 0$,

define $\boldsymbol{\tilde T} = \begin{pmatrix}\boldsymbol{N}L^{-1} & \boldsymbol{M}-\boldsymbol{N}L^{-1}K\end{pmatrix}$. Moreover, let

$$\hat U = \mathrm{diag}(I_l,\tilde C_\perp,1)\quad\text{and}\quad V^\top = \begin{pmatrix}\tilde B^\top & \tilde D_z^\top\end{pmatrix}_\perp. \tag{47}$$

Then the following statements are equivalent.

(a) There exists a controller (38) for the plant (37) such that the controlled interconnection (39) satisfies the hypotheses in Theorem 7.
(b) There exist $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$ with (32) and symmetric matrices X, $\tilde Y$ satisfying the LMIs (42) and

$$\left[V\left(\begin{array}{cc|cc}-I & \tilde A & 0 & \boldsymbol{\tilde T}\hat B_1\\ 0 & \tilde C_z & -1 & \boldsymbol{\hat D}_{zw}\end{array}\right)\right]\left(\begin{array}{cc|cc}\tilde Y & 0 & 0 & 0\\ 0 & -\tilde Y & 0 & 0\\ \hline 0 & 0 & 0 & 1\\ 0 & 0 & 1 & 0\end{array}\right)[\bullet]^\top\succ 0, \tag{48}$$

$$\begin{pmatrix}\tilde Y & \boldsymbol{\tilde T}\\ \boldsymbol{\tilde T}^\top & X\end{pmatrix}\succ 0. \tag{49}$$

By its very definition, $\boldsymbol{\tilde T}$ depends *affinely* on $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$ and, thus, the constraints (32), (42) and (48)-(49) constitute *affine constraints on all decision variable* $(\boldsymbol{C_{\rm f}},\boldsymbol{D_{\rm f}})$, X and $\tilde Y$. Hence, their feasibility can be verified by standard SDP-solvers. We emphasize that the complexity of these synthesis LMIs is determined by the dimensions of A and $A_{\rm f}$ only.

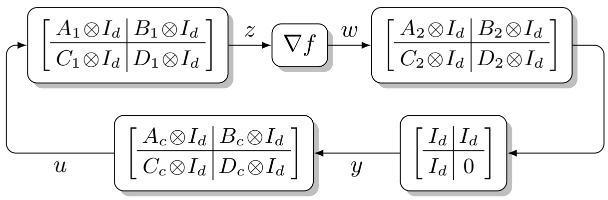

Fig. 6. Optimization over communication channels.

Note that the set of all eigenvalues and zeros of $\tilde{A}$ and $\tilde{C}_z(zI - \tilde{A})^{-1}\tilde{B} + \tilde{D}_z$, respectively, are given by $\rho^{-1}\Lambda$ and $\rho^{-1}\Xi$ with two finite and ρ-independent sets $\Lambda, \Xi \subset \mathbb{C}$. For a particular choice of $(A_{\rm f}, B_{\rm f})$, let us now summarize a concrete procedure for the synthesis of controllers as follows:

1) Fix $0 < m < L$. If $0 \notin \Lambda \cup \Xi$ set $z_0 = 0$. Otherwise choose $z_0 > 0$ close to zero such that $|\lambda| > z_0$ holds for all $\lambda \in (\Lambda \cup \Xi) \setminus \{0\}$.
2) Pick $l \in \mathbb{N}$ and $A_{\rm f} := C_\alpha$, $B_{\rm f} := e_l$ for $\alpha(z) := z^l - z_0^l$.
3) For any $\rho \in (z_0, 1]$, set up the system of LMIs (32), (42), (48)-(49) in the variables $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$, X and $\tilde{Y}$.
4) By bisection, determine the best possible (infimal) rate $\rho_* \in [z_0, 1]$ such that the resulting LMIs are feasible.
5) For some $\rho \in (\rho_*, 1]$ close to the optimal value ρ_*, set up the plant (37) with $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ as obtained from a feasible solution of these LMIs.
6) Design a controller (38) such that the LMI (25) for the closed-loop system (39) is feasible in $\mathcal{X}$.
7) Define $(A_c, B_c, C_c, D_c) := (\rho\hat{A}_c, \rho\hat{B}_c, \hat{C}_c, \hat{D}_c)$.

Since the assumptions on α in Theorem 8 are satisfied, it is possible to set up the LMIs in Step 3); indeed, all the eigenvalues of $A_{\rm f}$ have absolute value z_0 and, hence, none of them is contained in $\rho^{-1}\Lambda \cup \rho^{-1}\Xi$. As for Theorem 7, the case $l = 0$ is covered with empty matrices $A_{\rm f}$, $B_{\rm f}$, $\boldsymbol{C_{\rm f}}$ and $\boldsymbol{D_{\rm f}} = 1$, which boils down to choosing $\tilde{\boldsymbol{T}} = I$ to set up the LMIs in Step 3). Moreover, Theorem 8 guarantees that a controller as in Step 6) does indeed exist (possibly after a slight perturbation of $(\boldsymbol{C_{\rm f}}, \boldsymbol{D_{\rm f}})$ as seen in the proof in [33].)

A numerically stable and constructive procedure to design a controller as in Step 6) and based on the classical synthesis conditions (42)-(44) is found, e.g., in [28].

With the controller in Step 7), for any $f \in \mathcal{S}_{m,L}^0$ and any dimension $d \in \mathbb{N}$, it is guaranteed that all trajectories of the interconnection in Fig. 2 decay exponentially with rate ρ.

VII. Numerical Illustrations

We illustrate our results by an extremum control problem. The purpose is minimize any $f \in \mathcal{S}_{m,L}$ over communication channels. Concretely, the algorithm needs to transmit the actual iterate u_t via a channel modeled by an LTI system with transfer function $G_1(z) = C_1(zI - A_1)^{-1}B_1 + D_1$ to generate z_t. This is fed into the gradient to return $w_t = \nabla f(z_t)$. In turn, this signal is communicated back to the algorithm via a channel with transfer function $G_2(z) = C_2(zI - A_2)^{-1}B_2 + D_2$. To enforce integral action, we are led to the configuration in Fig. 6 with a to-be-designed controller (A_c, B_c, C_c, D_c). To avoid cancelation of the integrator's pole in the loop, we assume that $\det(A_2 - I) \neq 0$ and $G_2(1) \neq 0$.

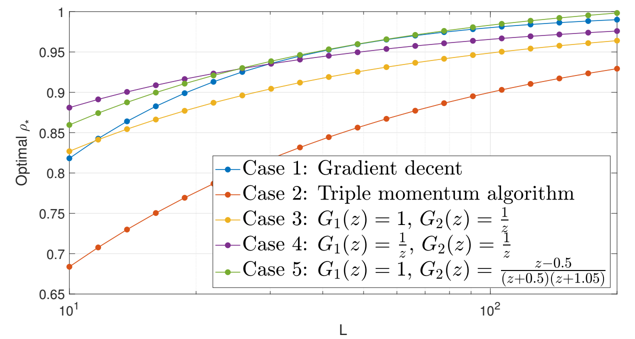

Fig. 7. Optimal convergence rates plotted over L for $m = 1$, $z_0 = 10^{-2}$ and $l = 2$ and for various communication dynamics.

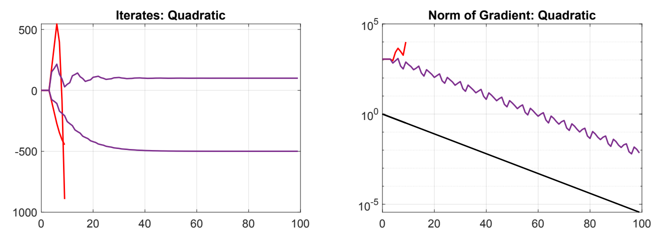

Fig. 8. Iterations for Case 4 with $m = 1$, $L = 10$ and a quadratic function for optimal algorithm (purple) and gradient decent (red). The black line indicates the optimal rate for synthesis.

Since the configuration in Fig. 6 can be subsumed to the one in Fig. 2, we can follow the procedure in Sec. VI to compute optimal rates and close-to-optimal algorithms for this optimization problem under communication constraints.

If choosing $G_1(z) = G_2(z) = 1$, Fig. 6 is identical to Fig. 2 for standard optimization. With $z_0 = 0$ in our synthesis procedure, we recover both the optimal convergence rates and the algorithm parameters for gradient decent ($l = 0$) and $\rho_{\rm tm} := 1 - \sqrt{\frac{m}{L}}$ for the triple momentum algorithm ($l = 1$) [1], [6], [7]. Remarkably, Theorem 8 permits to show that $\rho_{\rm tm}$ is indeed the best possible rate that is achievable among all algorithms and any $l \in \mathbb{N}_0$ [15, Corollary 4.8].

In the subsequent numerical experiments, we pick $m = 1$, $z_0 = 10^{-2}$, $l = 2$ and compute the optimal rates ρ_* for the communication filters in Fig. 7. The results are plotted over the so-called condition number L of the class $\mathcal{S}_{1,L}$.

If compared to the tripel momentum algorithm (Case 2), the rates increase if the gradients are processed with a one-step delay (Case 3), but they are still mostly better than for gradient descent (Case 1). For lager values of L, this is even true for delays in both channels (Case 4). Our approach allows for unstable dynamics in the optimization loop (Case 5), which affects the convergence rates adversely.

Figs. 8 and 9 depict the iterates with the optimal algorithm in Case 4, both for the quadratic and non-quadratic functions $f(x) = h(x-b)$ with $h(x) = Lx_1^2 + mx_2^2$, $h(x) = \frac{1}{2}(x_1^2 + _2^2) + 9\log(\exp(-x_1) + \exp(\frac{1}{3}x_1 + x_2) + \exp(\frac{1}{3}x_1 - x_2))$ [29] and $b = (100, -500)$ for $m = 1$ and $L = 10$, respectively. The optimal rates (depicted by the black line) are matched in the quadratic case and give an upper bound for the non-quadratic function. Gradient decent fails to converge in both cases.

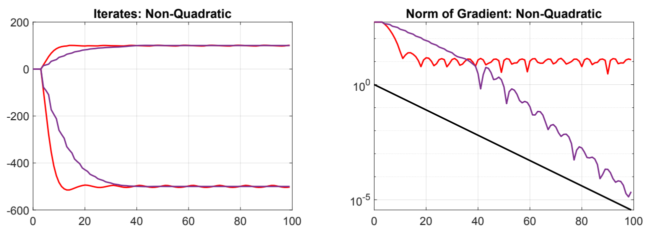

Fig. 9. Plots corresponding to Fig. 8 for a non-quadratic function.

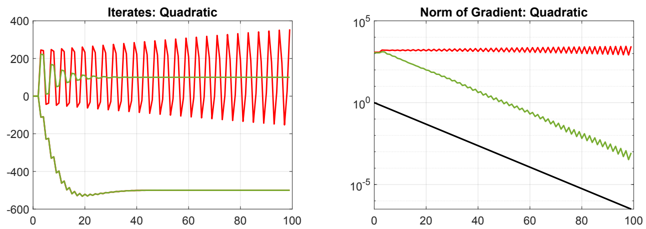

Fig. 10. Iterations for Case 5 with optimal algorithm designed for $m = 1$ and $L = 10$ and employed for a quadratic function with $m = 1$, $L = 10$ (green) as well as for $m = 1$, $L = 11$ (red), respectively. The black line indicates the optimal rate for synthesis.

Finally, Fig. 10 reveals that optimal algorithms are working well for the classes of functions they are designed for. As expected from robust control, however, they can be sensitive to deviations from the assumptions, as they lead to instability for the class $\mathcal{S}_{1,11}$ that is only slightly larger than $\mathcal{S}_{1,10}$. All presented results can be reproduced with the software at [32].

VIII. Conclusions

In this paper we have presented the full pipeline to design optimal optimization algorithms or extremum controllers based on causal dynamic stability multipliers. A novel parametrization of these filters overcomes technical assumptions as required in previous work. Future work is devoted to incorporating inexact gradient information and to handling anti-causal multipliers as well as performance objectives in synthesis.

References

[1] Y. Nesterov, *Lectures on Convex Optimization*, ser. Springer Optimization and Its Applications. Springer International Publishing, 2018, vol. 137.

[2] B. Polyak, *Introduction to Optimization*. Optimization Software, Inc., New York, 1987.

[3] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *Proceedings of the IEEE Conference on Decision and Control and European Control Conference, Orlando, FL, 2011*.

[4] H.-B. Dürr and C. Ebenbauer, "On a class of smooth optimization algorithms with applications in control," IFAC Proc. vol. 45, no. 17. Elsevier BV, 2012, pp. 291–298.

[5] L. Lessard, B. Recht, and A. Packard, "Analysis and Design of Optimization Algorithms via Integral Quadratic Constraints," *SIAM Journal on Optimization*, vol. 26, no. 1, pp. 57–95, 2016.

[6] Y. Nesterov, "A method for unconstrained convex minimization problem with the rate ofconvergence $O(\frac{1}{k^2})$." *Doklady AN SSSR*, vol. 269, pp. 543–547, 1983, (In Russian; translated as SovietMath. Docl.).

[7] B. V. Scoy, R. A. Freeman, and K. M. Lynch, "The fastest known globally convergent first-order method for minimizing strongly convex functions," *IEEE Control Systems Letters*, vol. 2, no. 1, pp. 49–54, 2018.

[8] K. Zhou, J. Doyle, K. Glover, *Robust and Optimal Control*, . Prentice Hall, 1996.

[9] C. Scherer, S. Weiland, "Linear matrix inequalities in control," in *The Control Systems Handbook, Second Edition: Control System Advanced Methods*, CRC Press, Chapter 24, pp. 1-30, 2011.

[10] L. Lessard and P. Seiler, "Direct synthesis of iterative algorithms with bounds on achievable worst-case convergence rate," in *2020 American Control Conference*, 2020, pp. 119–125.

[11] S. Michalowsky, C. Scherer, and C. Ebenbauer, "Robust and structure exploiting optimisation algorithms: an integral quadratic constraint approach," *International Journal of Control*, vol. 94, pp. 1–24, 2021.

[12] D. Gramlich, C. Ebenbauer, and C. W. Scherer, "Synthesis of accelerated gradient algorithms for optimization and saddle point problems using Lyapunov functions and LMIs," *Systems & Control Letters*, vol. 165, 2022.

[13] A. B. Taylor, J. M. Hendrickx, and F. Glineur, "Smooth strongly convex interpolation and exact worst-case performance of first-order methods," *Mathematical Programming*, vol. 161, no. 1-2, pp. 307–345, 2016.

[14] A. B. Taylor and Y. Drori, "An optimal gradient method for smooth strongly convex minimization," *Mathematical Programming*, vol. 199, no. 1-2, pp. 557–594, 2022.

[15] C. Scherer and C. Ebenbauer, "Convex synthesis of accelerated gradient algorithms," *SIAM Journal on Control and Optimization*, vol. 59, no. 6, pp. 4615–4645, 2021.

[16] T. Holicki and C. W. Scherer, "Algorithm design and extremum control: Convex synthesis due to plant multiplier commutation," in *60th IEEE Conference on Decision and Control*, 2021, pp. 3249–3252.

[17] C. W. Scherer, "Robust exponential stability and invariance guarantees with general dynamic O'Shea-Zames-Falb multipliers," in *Proc. IFAC World Congress, to appear*, 2023.

[18] A. Beck, *First-Order Methods in Optimization*. Society for Industrial and Applied Mathematics, 2017.

[19] C. Desoer and M. Vidyasagar, *Feedback Systems: Input-Output Approach*. London: Academic Press, 1975.

[20] A. Megretski and A. Rantzer, "System analysis via Integral Quadratic Constraints," *IEEE T. Automat. Contr.*, vol. 42, pp. 819–830, 1997.

[21] J. Willems and R. Brockett, "Some new rearrangement inequalities having application in stability analysis," *IEEE T. Automat. Contr.*, vol. 13, no. 5, pp. 539–549, 1968.

[22] G. Zames and P. L. Falb, "Stability conditions for systems with monotone and slope-restricted nonlinearities," *SIAM Journal of Control*, vol. 6, pp. 89–109, 1968.

[23] R. Boczar, L. Lessard, and B. Recht, "Exponential convergence bounds using integral quadratic constraints," in *54th IEEE Conference on Decision and Control (CDC)*, 2015, pp. 7516–7521.

[24] B. Hu and P. Seiler, "Exponential decay rate conditions for uncertain linear systems using integral quadratic constraints," *IEEE Transactions on Automatic Control*, vol. 61, no. 11, pp. 3631–3637, 2016.

[25] C. W. Scherer, "Dissipativity and integral quadratic constraints: Tailored computational robustness tests for complex interconnections," *IEEE Control Systems Magazine*, vol. 42, no. 3, pp. 115–139, 2022.

[26] P. Gahinet and P. Apkarian, "A linear matrix inequality approach to H_∞ control," *Internat. J. Robust Nonlinear Control*, vol. 4, pp. 421–448, 1994.

[27] T. Iwasaki and R. Skelton, "All controllers for the general $\mathcal{H}_\infty$ control problem: LMI existence conditions and state space formulas," *Automatica*, vol. 30, pp. 1307–1317, 1994.

[28] P. Gahinet, "A new parametrization of H_∞ suboptimal controllers," pp. 1031–1051, 1994.

[29] L. Lessard, "The analysis of optimization algorithms: A dissipativity approach," *IEEE Control Systems Magazine*, vol. 42, no. 3, pp. 58–72, 2022.

[30] K. J. Aström and R. M. Murray, *Feedback Systems: An Introduction for Scientists and Engineers*. Princeton Universety Press, 2009.

[31] E. de Souza and S. Bhattacharyya, "Controllability, observability and the solution of $AX - XB = C$," *Linear Algebra and its Applications*, vol. 39, pp. 167–188, 1981.

[32] A Matlab implementation of the results of this paper can be found at `https://zenodo.org/badge/latestdoi/691960972`

[33] C.W. Scherer, C. Ebenbauer and T. Holicki, "Optimization Algorithm Synthesis based on Integral Quadratic Constraints: A Tutorial," `https://doi.org/10.48550/arXiv.2306.00565`, 2023