Systems & Control Letters 165 (2022) 105271

Contents lists available at ScienceDirect

Systems & Control Letters

journal homepage: www.elsevier.com/locate/sysconle

Synthesis of accelerated gradient algorithms for optimization and saddle point problems using Lyapunov functions and LMIs

Dennis Gramlich [a,*], Christian Ebenbauer [a], Carsten W. Scherer [b]

[a] Chair of Intelligent Control Systems, RWTH Aachen University, Kopernikusstrasse 16, 52074 Aachen, Germany
[b] Chair of Mathematical Systems Theory, University of Stuttgart, Pfaffenwaldring 5a, 70569 Stuttgart, Germany

ARTICLE INFO

Article history:
Received 16 June 2020
Received in revised form 11 April 2022
Accepted 8 May 2022
Available online 28 May 2022

Keywords:
Accelerated gradient methods
Linear matrix inequalities
Saddle point problems
Convex synthesis

ABSTRACT

This paper considers the problem of designing accelerated gradient-based algorithms for optimization and saddle-point problems. The class of objective functions is defined by a generalized sector condition. This class of functions contains strongly convex functions with Lipschitz gradients but also non-convex functions, which allows not only to address optimization problems but also saddle-point problems. The proposed design procedure relies on a suitable class of Lyapunov functions and, for a fixed convergence rate, is a convex semi-definite program in all but one scalar parameter. The proposed synthesis allows the design of algorithms that reach the performance of state-of-the-art accelerated gradient methods and beyond.

1. Introduction

Gradient-based optimization algorithms are a standard tool in science and engineering. Many of these algorithms take the form of feedback interconnection between a discrete-time linear system and the gradient of the objective function. In case of a convex objective function, the corresponding gradient satisfies a certain sector condition. Hence such a feedback configuration falls in the class of so called Lur'e systems [1], which have been extensively studied in control theory. In recent years, results from Lur'e systems and techniques from robust control theory have been exploited to analyze convergence rates and robustness of known optimization algorithms and to design novel algorithms. Some of those new publications rely on IQCs (integral quadratic constraints) from robust control to generate convergence results. For example, IQCs were used in [2] to find upper bounds on the convergence rates of existing algorithms. This work was later extended to synthesis of algorithms in [3]. These IQC-based approaches gave rise to the development of the Triple Momentum Method [4]. This method has the fastest known upper convergence bound for strongly convex functions with Lipschitz gradients. Other related work that analyzes optimization algorithms from a dynamical systems perspective is for example given in [5,6], where also Lyapunov function techniques and robust control theory are employed, or in [7], where discrete-time algorithms are analyzed based on continuous-time counterparts. In addition, in [8,9] semi-definite programming formulations are proposed to analyze the convergence properties of first order optimization methods. Further related results are discussed in the recent paper [10], where the design of robust algorithms for structured objective functions based on IQCs is considered. Finally, in [11], the authors provide a semi-definite program for the analysis of gradient based optimization algorithms, which can be shown to find the fastest convergence rate that can be certified by quadratic Lyapunov functions.

In this paper, we address *convex design* (convex synthesis) of gradient-based algorithms for optimization and saddle point problems, where the class of objective functions is defined by a generalized sector condition. In particular, the contributions of this paper are as follows. Firstly, we consider classes of functions that are more general than the classes of strongly convex functions usually considered in the literature. In particular, the classes under consideration also contain non-convex functions, which we utilize in our procedure to design algorithms capable of searching for saddle points instead of minima. For example, the ability to search for saddle points allows us to apply the design method to optimization problems with equality constraints. Secondly, based on a rather general class of Lyapunov functions, we derive convex synthesis conditions for algorithm design in the form of linear matrix inequalities. Specifically, we provide a non-conservative convexification in the sense that the analysis matrix inequalities (when algorithm parameters are given) are feasible if and only if the synthesis matrix inequalities (when algorithm parameters are decision variables) are feasible, i.e., our design procedure is not more conservative than the corresponding analysis. This is in contrast to many other results in the literature, where the step from

* Corresponding author.
E-mail addresses: dennis.gramlich@ic.rwth-aachen.de (D. Gramlich), christian.ebenbauer@ic.rwth-aachen.de (C. Ebenbauer), carsten.scherer@mathematik.uni-stuttgart.de (C.W. Scherer).

https://doi.org/10.1016/j.sysconle.2022.105271
0167-6911/

convex analysis to convex synthesis is only possible by imposing additional assumptions (e.g. fixed IQC multipliers or quadratic Lyapunov functions). In the case of strongly convex functions, our design procedure reaches the same convergence rates as the Triple Momentum Method and it allows to incorporate additional structural properties of the objective function to design tailored algorithms with even faster convergence rates, as demonstrated in the paper.

Notation

We denote the spectrum of a matrix by $\sigma(\boldsymbol{A})$ and for the spectral radius we will write $\rho(\boldsymbol{A})$. We will use the notation $\|v\|_{\boldsymbol{A}}^2 = v^\top \boldsymbol{A} v$, which is a semi-norm if the matrix $\boldsymbol{A}$ is positive semi-definite or a norm whenever $\boldsymbol{A}$ is positive definite. However, note that in most places, where we use the notation $\|v\|_{\boldsymbol{A}}$, we do not require $\boldsymbol{A}$ to be positive semi-definite. In large matrix equations, we will sometimes write $\boldsymbol{A}^\top \boldsymbol{B}(\star)$ instead of $\boldsymbol{A}^\top \boldsymbol{B}\boldsymbol{A}$ with $(\star)$ as a placeholder for $\boldsymbol{A}$.

2. Problem statement and preliminary results

2.1. Problem statement

Consider the gradient based algorithm defined by

$$x_{k+1} = \boldsymbol{A}x_k + \boldsymbol{B}\nabla f(\boldsymbol{C}x_k), \tag{1}$$

where $x_k \in \mathbb{R}^n$ and the matrices $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, $\boldsymbol{B} \in \mathbb{R}^{n\times d}$, $\boldsymbol{C} \in \mathbb{R}^{d\times n}$ are the algorithm parameters to be designed. The objective function $f \in C^1(\mathbb{R}^d)$ is assumed to satisfy the following generalized sector condition for all $z_1, z_2 \in \mathbb{R}^d$:

$$\begin{aligned}\frac{1}{2}\|z_1 - z_2\|_{\boldsymbol{M}}^2 &\leq f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2) \\ &\leq \frac{1}{2}\|z_1 - z_2\|_{\boldsymbol{L}}^2,\end{aligned} \tag{2}$$

where $\boldsymbol{M} \preceq \boldsymbol{L} \in \mathbb{R}^{d\times d}$ are given symmetric matrices. In the following, $S(\boldsymbol{M}, \boldsymbol{L})$ denotes the set of all C^1 functions that satisfy (2). Note that $S(m\boldsymbol{I}_d, l\boldsymbol{I}_d)$, $0 < m < l$, is the set of m-strongly convex functions with l-Lipschitz continuous gradients. In addition, any function f that is contained in $S(\boldsymbol{M}, \boldsymbol{L})$ is at least m-weakly convex in the sense that $z \mapsto f(z) + \frac{m}{2} z^\top z$ is convex for $m = \max \operatorname{eig}(-\boldsymbol{M}) \cup \{0\}$. Finally, in the case $f \in C^2(\mathbb{R}^d)$, (2) is equivalent to $\boldsymbol{M} \preceq \nabla^2 f(z) \preceq \boldsymbol{L}$ for all $z \in \mathbb{R}^d$, where $\nabla^2 f$ denotes the Hessian of f.

Requirements. The algorithm design problem of this paper is feasible if and only if $\boldsymbol{M}$ and $\boldsymbol{L}$ are non-singular and have the same eigenvalue signature. This design problem is formally stated as:

Problem 1. For given $n \geq d$, $\boldsymbol{M} \preceq \boldsymbol{L}$, and convergence rate $\rho \in [0, 1[$, we aim to *design* matrices $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times d} \times \mathbb{R}^{d\times n}$ such that there exists $c \in \mathbb{R}_{\geq 0}$ and for any $f \in S(\boldsymbol{M}, \boldsymbol{L})$ some $x_f^* \in \mathbb{R}^n$ such that

$$\nabla f(z_f^*) = 0 \text{ for } z_f^* := \boldsymbol{C}x_f^* \tag{3}$$

holds and for any $x_0 \in \mathbb{R}^n$ the iterates x_k of (1) satisfy

$$\|x_f^* - x_k\| \leq c\rho^k \|x_f^* - x_0\|, \quad \forall k \in \mathbb{N}_0. \tag{4}$$

In our setting, *design* (synthesis) refers to computing the algorithm parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ by solving *convex* optimization problems, i.e., semi-definite programs.

Our goal is solving Problem 1. The following Problem 2 is similar to Problem 1, with the slight modification that all the functions f under consideration have their critical points in $z_f^* = 0$. This is favorable for the application of tools from robust control theory, which are often formulated for fixed-points in zero.

Problem 2. For given $n \geq d$, symmetric matrices $\widetilde{\boldsymbol{L}} \succeq 0$ and $\boldsymbol{M}$, and $\rho \in [0, 1[$, design matrices $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{C}}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times d} \times \mathbb{R}^{d\times n}$ such that the constraint

$$\widetilde{\boldsymbol{C}}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\widetilde{\boldsymbol{B}}\boldsymbol{M} = \boldsymbol{I}_d \tag{5}$$

is satisfied and there exists some $c \in \mathbb{R}_{\geq 0}$ such that for any $f \in S_0(0, \widetilde{\boldsymbol{L}}) := \{f \in S(0, \widetilde{\boldsymbol{L}}) : \nabla f(0) = 0\}$ and $x_0 \in \mathbb{R}^n$ the iterates of (1) satisfy

$$\|x_k\| \leq c\rho^k \|x_0\|, \quad \forall k \in \mathbb{N}_0.$$

Theorem 1. *Let symmetric matrices $\boldsymbol{M} \preceq \boldsymbol{L}$ be given, set $\widetilde{\boldsymbol{L}} := \boldsymbol{L} - \boldsymbol{M}$ and fix $\rho \in [0, 1[$. Then the matrices $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ solve Problem 1 if and only if the matrices $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{C}})$ solve Problem 2, where $\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{B}\boldsymbol{M}\boldsymbol{C}$, $\widetilde{\boldsymbol{B}} = \boldsymbol{B}$, $\widetilde{\boldsymbol{C}} = \boldsymbol{C}$.*

This theorem justifies that we can solve Problem 2 instead of Problem 1. The proof of this theorem and all following ones can be found in the appendix.

2.2. Properties of the class $S(\boldsymbol{M}, \boldsymbol{L})$

This subsection serves the purpose of introducing some important properties of $S(\boldsymbol{M}, \boldsymbol{L})$. The first result gives some equivalent characterizations for when $f \in S(0, \boldsymbol{L})$ holds true. Note, that these conditions can be applied to any class $S(\boldsymbol{M}, \boldsymbol{L})$ by using the fact $f \in S(\boldsymbol{M}, \boldsymbol{L}) \Leftrightarrow (z \mapsto f(z) - \frac{1}{2}z^\top \boldsymbol{M} z) \in S(0, \boldsymbol{L} - \boldsymbol{M})$.

Lemma 2 (*Characterizations for $f \in S(0, \boldsymbol{L})$*). *Let $\boldsymbol{L} \succeq 0$ and $f \in C^1(\mathbb{R}^d)$. All conditions below, holding for all $z_1, z_2 \in \mathbb{R}^n$, are equivalent to $f \in S(0, \boldsymbol{L})$:*

(1) $0 \leq f(z_2) - f(z_1) - (\nabla f(z_1))^\top (z_2 - z_1) \leq \frac{1}{2}\|z_1 - z_2\|_{\boldsymbol{L}}^2$,
(2) $0 \leq (\nabla f(z_1) - \nabla f(z_2))^\top (z_1 - z_2) \leq \|z_1 - z_2\|_{\boldsymbol{L}}^2$,
(3) $\frac{1}{2}\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}^2 \leq f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2)$ *and* $\boldsymbol{\Pi}_{\ker \boldsymbol{L}}(\nabla f(z_1) - \nabla f(z_2)) = 0$,
(4) $\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}^2 \leq (\nabla f(z_1) - \nabla f(z_2))^\top (z_1 - z_2)$ *and* $\boldsymbol{\Pi}_{\ker \boldsymbol{L}}(\nabla f(z_1) - \nabla f(z_2)) = 0$.

Not all possible variations of matrices $\boldsymbol{M} \preceq \boldsymbol{L}$ should be considered for optimization. For example, if there exists a singular matrix $\boldsymbol{Q}$ such that $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$, then the function f defined by $f(z) = \frac{1}{2}z^\top \boldsymbol{Q} z + v^\top z$, where v is not in the range of $\boldsymbol{Q}$, would be an element of $S(\boldsymbol{M}, \boldsymbol{L})$ without any critical point. Therefore this set $S(\boldsymbol{M}, \boldsymbol{L})$ would not make sense as a set of objective functions, since we cannot solve Problem 1 for it. The following Lemma characterizes when such cases can be avoided.

Lemma 3 (*Well-posed Pairs $\boldsymbol{M}, \boldsymbol{L}$*). *Let $\boldsymbol{M}, \boldsymbol{L} \in \mathbb{R}^{d\times d}$ be symmetric matrices with $\boldsymbol{M} \preceq \boldsymbol{L}$. Then the following five statements are equivalent:*

(1) The matrices $\boldsymbol{M}$ and $\boldsymbol{L}$ have the same numbers of positive and negative, and no zero eigenvalues.
(2) Any symmetric matrix $\boldsymbol{Q} \in \mathbb{R}^{d\times d}$ with $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ is non-singular.
(3) $\boldsymbol{L} + \boldsymbol{M}$ is non-singular and the spectral radius of $(\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})$ is smaller than one.
(4) $\boldsymbol{M}$ is non-singular and $\boldsymbol{M}^{-1}\boldsymbol{L}$ has only positive eigenvalues.
(5) $\boldsymbol{M}$ and $\boldsymbol{L}$ are non-singular and congruent, i.e., there exists a non-singular matrix $\boldsymbol{T} \in \mathbb{R}^{d\times d}$ with $\boldsymbol{M} = \boldsymbol{T}^\top \boldsymbol{L}\boldsymbol{T}$.

Remark 4. In Lemma 3, statement (1) serves the purpose of giving the reader a good intuition for the property under consideration. Statement (2) and (3) will be useful in later proofs. Note that in particular (2) prevents the counter-example we constructed in the motivation of this lemma.

Because of the importance of this property we define a new notation for matrices $\boldsymbol{M}, \boldsymbol{L}$ fulfilling one and thus all conditions in Lemma 3.

Definition 5 (*Loewner-congruence Ordering on Symmetric Matrices*)**.** For symmetric matrices $\boldsymbol{M}, \boldsymbol{L} \in \mathbb{R}^{d\times d}$, we introduce the partial ordering

$$\boldsymbol{L} \succeq_c \boldsymbol{M} :\Leftrightarrow \begin{cases} \boldsymbol{L} - \boldsymbol{M} \text{ is positive semi-definite} \\ \boldsymbol{L} \text{ and } \boldsymbol{M} \text{ are congruent.} \end{cases}$$

Under the Loewner-congruence ordering, a critical point exists, is unique, and a simple gradient method converges to the critical point, as stated in the following results.

Proposition 6 (*A simple Gradient Method*)**.** *Let $\boldsymbol{M} \preceq \boldsymbol{L} \in \mathbb{R}^{d\times d}$ be non-singular. Then for any convergence rate $\rho > \rho\left((\boldsymbol{L}+\boldsymbol{M})^{-1}(\boldsymbol{L}-\boldsymbol{M})\right)$ there exists $r \in \mathbb{R}_{>0}$ such that*

$$z \mapsto z - 2(\boldsymbol{M}+\boldsymbol{L})^{-1}\nabla f(z) \tag{6}$$

is a contraction for all $f \in S(\boldsymbol{M}, \boldsymbol{L})$ with contraction constant ρ on the Banach space $(\mathbb{R}^d, \|\cdot\|_{\boldsymbol{P}})$, where $\boldsymbol{P} = (\boldsymbol{L}+\boldsymbol{M})((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})})(\boldsymbol{L}+\boldsymbol{M})$.

Remark 7. For $\boldsymbol{M} \preceq_c \boldsymbol{L}$, the optimizer defined by $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ with $\boldsymbol{A} = \boldsymbol{C} = \boldsymbol{I}_d$ and $\boldsymbol{B} = -2(\boldsymbol{L}+\boldsymbol{M})^{-1}$ realizes the contraction in Proposition 6. As a consequence of the Banach fixed-point theorem, it converges faster than any convergence rate $\rho > \rho_{\mathrm{grad}} := \rho\left((\boldsymbol{L}+\boldsymbol{M})^{-1}(\boldsymbol{L}-\boldsymbol{M})\right)$ and converges monotonically in the norm $\|\cdot\|_{\boldsymbol{P}}$ to the unique critical point. Finally, notice that in the case $\boldsymbol{L}-\boldsymbol{M}$ is singular, the infimal convergence rate may not be attained, since r can go towards infinity as ρ converges towards ρ_{grad}. However, if $\boldsymbol{L}-\boldsymbol{M}$ is non-singular, then r disappears and the constructed gradient method converges at the rate ρ_{grad}.

Theorem 8 (*Existence and Uniqueness of Critical Points*)**.** *Let $\boldsymbol{M} \preceq \boldsymbol{L} \in \mathbb{R}^{d\times d}$ be given symmetric, non-singular matrices. Then the following three statements are equivalent:*

(1) The matrices $\boldsymbol{M}, \boldsymbol{L}$ satisfy $\boldsymbol{M} \preceq_c \boldsymbol{L}$.
(2) For all $f \in S(\boldsymbol{M}, \boldsymbol{L})$ there exists at least one $z_f^* \in \mathbb{R}^d$ with $\nabla f(z_f^*) = 0$.
(3) For all $f \in S(\boldsymbol{M}, \boldsymbol{L})$ there exists at most one $z_f^* \in \mathbb{R}^d$ with $\nabla f(z_f^*) = 0$.

Remark 9. Theorem 8 shows that if we aim to design algorithms that are convergent for the whole class $S(\boldsymbol{M}, \boldsymbol{L})$, we need to require $\boldsymbol{M} \preceq_c \boldsymbol{L}$, because otherwise there would be elements of $S(\boldsymbol{M}, \boldsymbol{L})$ without critical points. Hence the introduced partial ordering plays a key role in our results. Note that it is no coincidence that in Theorem 8 the existence of critical points for all functions in $S(\boldsymbol{M}, \boldsymbol{L})$ and the uniqueness of critical points are two separate, equivalent statements.

3. Main results

In this section, a convex synthesis approach of optimizer parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ for a given convergence rate and for the set of objective functions $S(\boldsymbol{M}, \boldsymbol{L})$ is provided. By Theorem 1, the design for $S(\boldsymbol{M}, \boldsymbol{L})$ reduces to designing algorithms for $S_0(0, \widetilde{\boldsymbol{L}}) = \{f \in S(0, \widetilde{\boldsymbol{L}}) | \nabla f(0) = 0\}$ with $\widetilde{\boldsymbol{L}} = \boldsymbol{L} - \boldsymbol{M}$. Hence, we study Problem 2 instead of Problem 1.

3.1. A class of Lyapunov functions

To design the algorithm parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ with a prescribed convergence rate, we propose the following class of (non-quadratic) Lyapunov function candidates

$$V_f(x) = \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{pmatrix} \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} + f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top \widetilde{\boldsymbol{L}}^\dagger \nabla f(\boldsymbol{C}x) \tag{7}$$

with parameter $0 \prec \boldsymbol{P} = \boldsymbol{P}^\top \in \mathbb{R}^{n+d\times n+d}$. In continuous-time, similar Lyapunov functions have already been applied to Lur'e systems. Those Lyapunov functions share the first term, which is quadratic in the state x and the static non-linearity $\nabla f(z)$. They have been proposed by Yakubovic for $d = 1$ in [12] and are employed e.g. in [13–15].

Our design approach, for a given convergence rate ρ, is based on finding a Lyapunov function ($\boldsymbol{P} \succ 0$) and algorithm parameters by semi-definite programming satisfying the conditions in the next standard Lyapunov theorem.

Theorem 10. *Consider the algorithm (1) for $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}) \in \mathbb{R}^{n\times n} \times \mathbb{R}^{n\times d} \times \mathbb{R}^{d\times n}$ and $f \in C^1(\mathbb{R}^d)$ with $\nabla f(x_f^*) = 0$. Let there exist some function $V : \mathbb{R}^d \to \mathbb{R}$ and $\alpha > 0$, $\beta > 0$ satisfying the quadratic bounds*

$$\alpha\|x - x_f^*\|^2 \le V(x) \le \beta\|x - x_f^*\|^2 \quad \forall x \in \mathbb{R}^n \tag{8}$$

and the ρ-weighted increment inequality

$$V(x^+) - \rho^2 V(x) \le 0 \quad \forall x \in \mathbb{R}^n, \tag{9}$$

where $x^+ = \boldsymbol{A}x + \boldsymbol{B}\nabla f(\boldsymbol{C}x)$. Then (4) holds for some $c \in \mathbb{R}_{\ge 0}$ with $c \le \sqrt{\beta/\alpha}$.

The following two lemmas provide useful bounds for the constants α and β in (8) and the increment (9) of the Lyapunov function candidate V_f in (7).

Lemma 11 (*Upper Bound on the Lyapunov Increment of V_f*)**.** *Assume $f \in S_0(0, \widetilde{\boldsymbol{L}})$. Then, the weighted increment (9) for V_f is upper bounded as follows for arbitrary $\lambda \in [0, \rho^2]$:*

$$V_f(x^+) - \rho^2 V_f(x) \le \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}^\top \left(\begin{array}{cc|cc} -\rho^2\boldsymbol{P}_{11} & -\rho^2\boldsymbol{P}_{12} & 0 & 0 \\ -\rho^2\boldsymbol{P}_{21} & -\rho^2\boldsymbol{P}_{22} & 0 & 0 \\ \hline 0 & 0 & \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ 0 & 0 & \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{array}\right) \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix} + \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}^\top \left(\begin{array}{cc|cc} 0 & 0 & 0 & -\frac{\lambda}{2}\boldsymbol{C}^\top \\ 0 & 0 & 0 & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger \\ \hline 0 & 0 & 0 & \frac{1}{2}\boldsymbol{C}^\top \\ -\frac{\lambda}{2}\boldsymbol{C} & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & \frac{1}{2}\boldsymbol{C} & -\widetilde{\boldsymbol{L}}^\dagger \end{array}\right) \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix},$$

where $w = \nabla f(\boldsymbol{C}x)$, $w^+ = \nabla f(\boldsymbol{C}x^+)$ and $x^+ = \boldsymbol{A}x + \boldsymbol{B}w$.

Lemma 12 (*Quadratic Bounds on V_f*)**.** *Let $f \in S_0(0, \widetilde{\boldsymbol{L}})$ be given. Then the Lyapunov function candidate V_f fulfills the quadratic bounds*

$$\alpha\|x\|^2 \le V_f(x) \le \beta\|x\|^2$$

with the constants $\alpha := \lambda_{\min}(\boldsymbol{P})$ and $\beta := \lambda_{\max}(\boldsymbol{P})(1 + \|\widetilde{\boldsymbol{L}}\|^2\|\boldsymbol{C}\|^2) + \frac{\|\widetilde{\boldsymbol{L}}\|\|\boldsymbol{C}\|^2}{2}$.

3.2. Convex synthesis of algorithms

The following theorem reformulates the condition (9) in Theorem 10 using the established bound in Lemma 11.

Theorem 13 (*Analysis Inequalities*)**.** *Let $\boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times d}$ and $\boldsymbol{C} \in \mathbb{R}^{d\times n}$ be given. Set $\widetilde{\boldsymbol{A}} = \boldsymbol{A} + \boldsymbol{BMC}$. Then Algorithm (1) solves Problem 1 and has convergence rate $\rho \in [0, 1[$, if there exist $\boldsymbol{P} = \boldsymbol{P}^\top \succ 0$, $\lambda \in [0, \rho^2]$ and $r \in \mathbb{R}$ such that $\boldsymbol{I}_d = \boldsymbol{C}(\widetilde{\boldsymbol{A}} - \boldsymbol{I})^{-1}\boldsymbol{BM}$ (see (5)) is satisfied and*

$$\begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \hline \widetilde{\boldsymbol{A}} & \boldsymbol{B} & 0 \\ 0 & 0 & \boldsymbol{I}_d \end{pmatrix}^\top \left(\begin{array}{cc|cc} -\rho^2\boldsymbol{P}_{11} & -\rho^2\boldsymbol{P}_{12} & 0 & 0 \\ -\rho^2\boldsymbol{P}_{21} & -\rho^2\boldsymbol{P}_{22} & 0 & 0 \\ \hline 0 & 0 & \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ 0 & 0 & \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{array}\right)(\star) + \begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \hline \widetilde{\boldsymbol{A}} & \boldsymbol{B} & 0 \\ 0 & 0 & \boldsymbol{I}_d \end{pmatrix}^\top \left(\begin{array}{cc|cc} 0 & 0 & 0 & -\frac{\lambda}{2}\boldsymbol{C}^\top \\ 0 & -r\boldsymbol{\Pi} & 0 & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger \\ \hline 0 & 0 & 0 & \frac{1}{2}\boldsymbol{C}^\top \\ -\frac{\lambda}{2}\boldsymbol{C} & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & \frac{1}{2}\boldsymbol{C} & -\widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} \end{array}\right)(\star) \prec 0 \qquad (10)$$

holds, where $\widetilde{\boldsymbol{L}} = \boldsymbol{L} - \boldsymbol{M}$ and $\boldsymbol{\Pi} = \boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})}$.

Theorem 13 provides sufficient conditions for a given algorithm to achieve a convergence rate ρ. Notice that the conditions in Theorem 13 are affine in the positive definite decision variable $\boldsymbol{P}$ and hence semi-definite programming can be used to verify these conditions. For the synthesis of algorithms, i.e., if in addition to $\boldsymbol{P}$ also $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ are decision variables, the decision variables enter in a non-affine (non-convex) fashion and thus, an efficient synthesis of algorithms with semi-define programming is not possible. Hence, it is of key importance to find *equivalent* conditions in terms of matrix inequalities and equations in which the decision variables enter in an affine fashion. The following theorem shows that this is indeed possible.

Theorem 14 (*Synthesis Inequalities*)**.** *Let $n \geq 3d$. Then there exist matrices $\boldsymbol{A} \in \mathbb{R}^{n\times n}, \boldsymbol{B} \in \mathbb{R}^{n\times d}, \boldsymbol{C} \in \mathbb{R}^{d\times n}$, which render the conditions (5) and (10) in Theorem 13 for a given convergence rate ρ feasible, if and only if there exist $\hat{\boldsymbol{A}} \in \mathbb{R}^{n\times n}, \hat{\boldsymbol{B}} \in \mathbb{R}^{n\times d}, \boldsymbol{C} \in \mathbb{R}^{d\times n}, \boldsymbol{P} = \boldsymbol{P}^\top \in \mathbb{R}^{n+d\times n+d}$, $r \in \mathbb{R}$ and $\lambda \in [0, \rho^2]$ which satisfy the matrix inequality*

$$\left(\begin{array}{cc|c|cc} -\rho^2\boldsymbol{P}_{11} & -\rho^2\boldsymbol{P}_{12} & * & * & * \\ -\rho^2\boldsymbol{P}_{21} & -\rho^2\boldsymbol{P}_{22} - r\boldsymbol{\Pi} & * & * & * \\ \hline \frac{1}{2}\boldsymbol{J}_2\hat{\boldsymbol{A}} - \frac{\lambda}{2}\boldsymbol{C} & \frac{1}{2}\boldsymbol{J}_2\hat{\boldsymbol{B}} + \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & -\widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} & * & * \\ \hline \hat{\boldsymbol{A}} & \hat{\boldsymbol{B}} & \boldsymbol{P}_{12} & -\boldsymbol{P}_{11} & -\boldsymbol{P}_{12} \\ \boldsymbol{J}_3\hat{\boldsymbol{A}} & \boldsymbol{J}_3\hat{\boldsymbol{B}} & \boldsymbol{P}_{22} & -\boldsymbol{P}_{21} & -\boldsymbol{P}_{22} \end{array}\right) \prec 0 \qquad (11)$$

and the affine constraints

$$\hat{\boldsymbol{B}} = (\hat{\boldsymbol{A}} - \boldsymbol{P}_{11})\boldsymbol{J}_1^\top \boldsymbol{M}^{-1}, \qquad \boldsymbol{C}\boldsymbol{J}_1^\top = \boldsymbol{I}_d, \\ \boldsymbol{C} = \boldsymbol{J}_2\boldsymbol{P}_{11}, \qquad \boldsymbol{P}_{21} = \boldsymbol{J}_3\boldsymbol{P}_{11}, \qquad (12)$$

where $\widetilde{\boldsymbol{L}} := \boldsymbol{L} - \boldsymbol{M}$, $\boldsymbol{\Pi} := \boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})}$ and $\boldsymbol{J}_1, \boldsymbol{J}_2, \boldsymbol{J}_3 \in \mathbb{R}^{d\times n}$ are given as

$$\boldsymbol{J}_1 = \begin{pmatrix} \boldsymbol{I}_d & 0 \end{pmatrix}, \quad \boldsymbol{J}_2 = \begin{pmatrix} 0_d & \boldsymbol{I}_d & 0 \end{pmatrix}, \quad \boldsymbol{J}_3 = \begin{pmatrix} 0_d & 0_d & \boldsymbol{I}_d & 0 \end{pmatrix}.$$

With a solution of (11) and (12), the conditions (5) and (10) in Theorem 13 are feasible with the algorithm parameters

$$\boldsymbol{A} := \boldsymbol{P}_{11}^{-1}\hat{\boldsymbol{A}} - \boldsymbol{BMC}, \quad \boldsymbol{B} := \boldsymbol{P}_{11}^{-1}\hat{\boldsymbol{B}} \quad \text{and} \quad \boldsymbol{C}.$$

The conditions for synthesizing algorithms from Theorem 14 are only sufficient, since this is true for analysis based on Theorem 13. The conservatism essentially depends on the choice of the Lyapunov function V_f and the estimates in Lemma 11. By making use of Theorems 13 and 14, we show in the following result that one can always achieve the convergence rate of the gradient method from Proposition 6. Still, the numerical results in Section 4 reveal that it is indeed possible to exceed this rate by designing dedicated algorithms.

Theorem 15 (*Guaranteed Feasibility for the Gradient Method*)**.** *Let $\boldsymbol{M} \preceq_c \boldsymbol{L}$ hold. Then the following statements are true:*

(1) The gradient method defined by $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ with $\boldsymbol{A} = \boldsymbol{I}_d$, $\boldsymbol{B} = -2(\boldsymbol{L} + \boldsymbol{M})^{-1}$, $\boldsymbol{C} = \boldsymbol{I}_d$ fulfills the conditions (10) and (5) of Theorem 13 for any $\rho \in]\rho_{\text{grad}}, 1]$, where $\rho_{\text{grad}} = \rho\left((\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})\right)$.

(2) If $n \geq 3d$, then there exists a solution to (11) and (12) in Theorem 14 for any $\rho \in]\rho_{\text{grad}}, 1]$.

This result is similar to one of [11], in which the authors show that their analysis techniques can certify the fastest convergence rates that can be achieved with quadratic Lyapunov functions. Our approach offers the advantage of using a class of non-quadratic Lyapunov functions.

The rate ρ_{grad} is the optimal convergence rate for the gradient method from Proposition 6. On the other hand, if one minimizes ρ simultaneously over all variables in Theorems 14, 15 implies that one can achieve at least the rate ρ_{grad}, while the examples in Section 4 show that the optimal rate is typically smaller.

4. Examples and numerical results

In the following examples, the LMI formulation from Theorem 14 is used to design optimization algorithms for three problems. In this course, ρ is minimized using a bisection algorithm to find a feasible solution of the matrix inequality (11). Strictly speaking, synthesis based on Theorem 14 is not based on solving an LMI if we consider λ as a decision variable. However, one can optimize over the single scalar parameter λ by an additional line search (see A.14). In case of the subsequent three problems, λ could also just be set equal to ρ^2 since this always turned out to be the optimal value. In fact, the choice $\lambda = \rho^2$ was optimal for all problems with $\boldsymbol{M} \succ 0$ that we have ever tested. Only when $\boldsymbol{M} \not\succ 0$, the optimal value for λ sometimes differed from ρ^2.

4.1. Convergence rates

To demonstrate the performance of our synthesis, we apply it to the class $S(m, l)$, $0 < m \leq l, m, l \in \mathbb{R}$ of strongly convex functions, which is often considered in the literature (for example in [2,10,16]). The algorithm parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ are computed by solving (11) and (12) in Theorem 14 for $\lambda = \rho^2$, where ρ is optimized using a bisection search. Here, setting λ equal to ρ^2 is motivated by the proof of Lemma 11, where $\lambda = \rho^2$ gives the tightest estimate on the increment of the Lyapunov function. The obtained convergence rates are shown in Fig. 1, where they are compared to the convergence rates of the Triple Momentum method from [4] and the theoretical lower bound on the convergence rates obtained by Nesterov. As can be observed, our synthesized algorithm has the same convergence rates as the Triple Momentum method. A result, that is also obtained in [3] using an IQC based approach. In fact Triple Momentum and our synthesized algorithm were basically identical, i.e., they produced the same iterates z_k.

4.2. Structured objective functions

The following (academic) example shall demonstrate the possible benefits of including additional properties of the objective function into algorithm design compared to the design for $S(m, l)$.

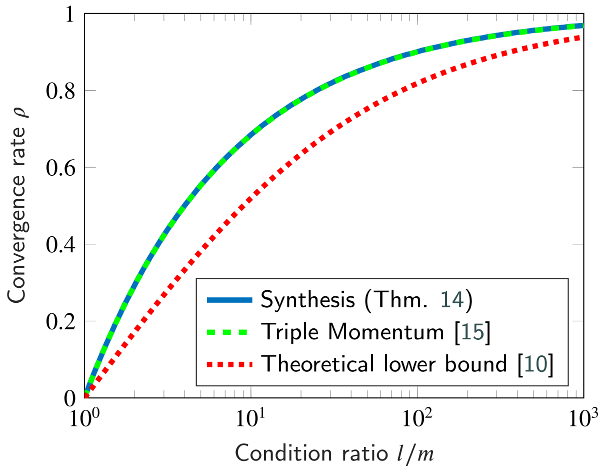

Fig. 1. The convergence rate guarantees achieved by designing algorithms for $S(m, l)$ using Theorem 14 are plotted over the condition number l/m and compared to the rate bound of Triple Momentum $\rho = 1 - \frac{\sqrt{m}}{\sqrt{l}}$ and the theoretical lower bound $\rho = \frac{\sqrt{l}-\sqrt{m}}{\sqrt{l}+\sqrt{m}}$ from [16].

Consider the function class $S(\boldsymbol{M}, \boldsymbol{L})$ with

$$\boldsymbol{M} = \begin{pmatrix} l - m + \frac{m^2}{l} & 0 \\ 0 & m \end{pmatrix}, \quad \boldsymbol{L} = \boldsymbol{S}^\top \begin{pmatrix} l & 0 \\ 0 & 2m - \frac{m^2}{l} \end{pmatrix} \boldsymbol{S},$$
$$\boldsymbol{S} = \begin{pmatrix} \sqrt{1 - \left(\frac{m}{l}\right)^2} & -\frac{m}{l} \\ \frac{m}{l} & \sqrt{1 - \left(\frac{m}{l}\right)^2} \end{pmatrix}.$$

These matrices fulfill $m\boldsymbol{I} \preceq \boldsymbol{M} \preceq_c \boldsymbol{L} \preceq l\boldsymbol{I}$. Moreover, the largest eigenvalue of $\boldsymbol{L}$ is l and the smallest eigenvalue of $\boldsymbol{M}$ is m. Hence, the best "standard method" for the class $S(\boldsymbol{M}, \boldsymbol{L})$ is a method for $S(m, l)$ and has a convergence rate that is not faster than $\frac{\sqrt{l}-\sqrt{m}}{\sqrt{l}+\sqrt{m}}$. The method designed using Theorem 14, on the other hand, has at least the convergence rate $\rho((\boldsymbol{M}+\boldsymbol{L})^{-1}(\boldsymbol{L}-\boldsymbol{M}))$. Fig. 2 illustrates these convergence rates together with the rate of a synthesized algorithm. One can recognize that, in this example, the structured method is superior to any unstructured method.

4.3. Application to constrained optimization

The class $S(\boldsymbol{M}, \boldsymbol{L})$ can contain non-convex functions. If both $\boldsymbol{M}$ and $\boldsymbol{L}$ are indefinite but the condition $\boldsymbol{M} \preceq_c \boldsymbol{L}$ is fulfilled, then $S(\boldsymbol{M}, \boldsymbol{L})$ is a class of functions with unique critical (saddle) points. One particular saddle point problem can be obtained in the context of convex constrained optimization. If one aims to solve the (linearly) constrained optimization problem

$$\begin{aligned} &\text{minimize } g(y), \\ &\text{subject to } y \in \mathbb{R}^d, \ \boldsymbol{A}_{\text{eq}} y = b_{\text{eq}}, \end{aligned} \tag{13}$$

where $g \in S(\boldsymbol{M}, \boldsymbol{L})$, $\boldsymbol{A}_{\text{eq}} \in \mathbb{R}^{d_2 \times d}$ and $0 \prec \boldsymbol{M} \prec \boldsymbol{L}$ holds (such that g is strictly convex), then a solution can be found by solving the saddle point problem

$$\sup_{\lambda \in \mathbb{R}^{d_2}} \inf_{y \in \mathbb{R}^d} g(y) + \lambda^\top (\boldsymbol{A}_{\text{eq}} y - b_{\text{eq}}).$$

Here, the Lagrangian function $L(y, \lambda) = g(y) + \lambda^\top(\boldsymbol{A}_{\text{eq}} y - b_{\text{eq}})$ is an element of $S(\boldsymbol{M}_L, \boldsymbol{L}_L)$, where

$$\boldsymbol{M}_L = \begin{pmatrix} \boldsymbol{M} & \boldsymbol{A}_{\text{eq}}^\top \\ \boldsymbol{A}_{\text{eq}} & 0 \end{pmatrix}, \quad \boldsymbol{L}_L = \begin{pmatrix} \boldsymbol{L} & \boldsymbol{A}_{\text{eq}}^\top \\ \boldsymbol{A}_{\text{eq}} & 0 \end{pmatrix}. \tag{14}$$

If $\boldsymbol{M}_L \preceq_c \boldsymbol{L}_L$ is satisfied, then our design procedure can be applied to design a gradient based (primal–dual) algorithm for $f : z \mapsto L(z)$, where $z = (y, \lambda)$. (This algorithm will then converge to a saddle point of L which solves the constrained optimization problem.) The following lemma shows under rather mild conditions that this is possible.

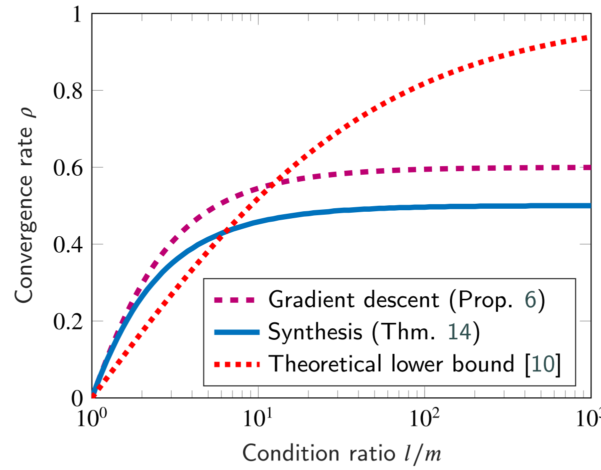

Fig. 2. Convergence rates achieved by the gradient descent algorithm in Proposition 6 and by synthesis with Theorem 14 for $S(\boldsymbol{M}, \boldsymbol{L})$. Note that the theoretical lower bound holds for the class $S(m, l)$ and not for the subset $S(\boldsymbol{M}, \boldsymbol{L}) \subset S(m, l)$, because the subset contains fewer objective functions.

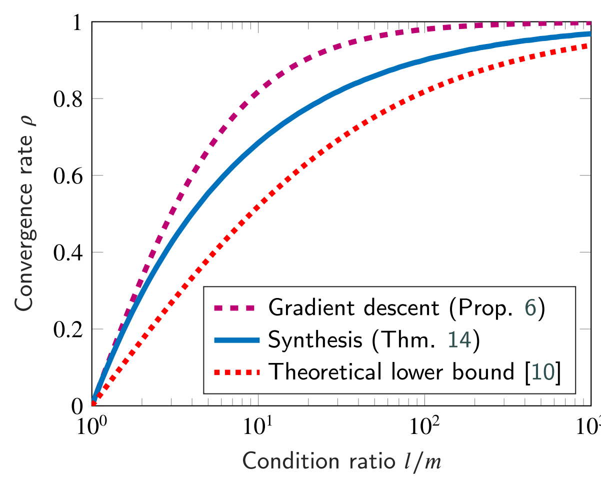

Fig. 3. Convergence rates achieved by the gradient descent algorithm in Proposition 6 and by designing algorithms with Theorem 14 for a constrained optimization problem.

Lemma 16. *Let symmetric matrices $\boldsymbol{M}, \boldsymbol{L} \in \mathbb{R}^{d \times d}$ and $\boldsymbol{A}_{\text{eq}} \in \mathbb{R}^{d_2 \times d}$ be given. Assume that $\boldsymbol{M} \preceq_c \boldsymbol{L}$ holds with $\boldsymbol{M}, \boldsymbol{L}$ being non-singular and that $\boldsymbol{A}_{\text{eq}}$ has full row rank. Consider $\boldsymbol{M}_L, \boldsymbol{L}_L$ defined in (14). Then $\boldsymbol{M}_L \preceq_c \boldsymbol{L}_L$ holds, and $\boldsymbol{M}_L$ and $\boldsymbol{L}_L$ are non-singular.*

As an academic example, consider the constrained optimization problem (13) with $g \in S(m\boldsymbol{I}_2, l\boldsymbol{I}_2)$ and $\boldsymbol{A}_{\text{eq}} = (1 \ \ 1)$. As described above, matrices $\boldsymbol{M}_L \preceq_c \boldsymbol{L}_L$ can be constructed such that the Lagrangian L of this problem is in $S(\boldsymbol{M}_L, \boldsymbol{L}_L)$. This enables algorithms of the form $x_{k+1} = \boldsymbol{A}x_k + \boldsymbol{B}\nabla L(\boldsymbol{C}x_k)$ to be designed. The algorithm parameters $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ can be designed by solving the matrix inequality from Theorem 14. The results are presented in Fig. 3. For the sake of comparison, we added the rates of the descent algorithm from Proposition 6. Interestingly, the convergence rates are exactly equal to the convergence rates for the unconstrained optimization problems. In general, we observed in our experiments that the convergence rates for linearly constrained optimization problems were often faster than those for unconstrained problems, but never slower.

Compared to other works, as in [17], our example is rather simple, involving only equality constraints. However, the purpose of this example is to show the principle applicability of our method to constrained optimization. The applicability to linear inequality constraints is part of future work.

Notice that we have the condition $n \geq 3d$ in Theorem 14, hence the algorithm with one equality constraint has at least dimension 9. However, it is often possible to reduce the dimension of the algorithm as outlined below. For example, we consider the algorithm parameters $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ for $m = 1, l = 15$ designed using Theorem 14. The original matrices had dimension $n = 9$. We observed that the last three modes usually do not contribute much to the dynamics of the algorithm. Hence, it is possible to eliminate them using balanced truncation. We used Theorem 13 to check that the reduced algorithm still converges for $S(m = 1, l = 15)$. The reduced algorithm achieves a convergence rate of at least 0.7422. which is faster than the rate of gradient descent, which is 0.8750 and exactly as fast as the unreduced algorithm. In our example, we obtained the reduced parameters:

$$\boldsymbol{A} = \begin{pmatrix} 1 & 0 & 0 & 0.0135 & -0.0258 & -0.0017 \\ 0 & 1 & 0 & 0.0135 & 0.0258 & -0.0017 \\ 0 & 0 & 1 & -0.6076 & -0.0036 & -0.0363 \\ 0 & 0 & 0 & -0.3097 & -0.0042 & -0.0474 \\ 0 & 0 & 0 & -0.0039 & 0.3909 & -0.0002 \\ 0 & 0 & 0 & 1.1631 & 0.0070 & 0.5255 \end{pmatrix},$$

$$\boldsymbol{B} = \begin{pmatrix} -0.0846 & 0.0707 & -0.1978 \\ 0.0707 & -0.0846 & -0.1978 \\ -0.2758 & -0.2758 & -3.2399 \\ 0.0860 & 0.0940 & -4.7039 \\ 0.6738 & -0.6727 & -0.0264 \\ 0.0896 & 0.0900 & 6.3240 \end{pmatrix},$$

$$\boldsymbol{C} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}.$$

Here, we have chosen a specific representation in which the first d columns of $\boldsymbol{A}$ are the first d unit vectors in $\mathbb{R}^n$ and $\boldsymbol{C}$ takes the form of an identity matrix concatenated with a zero block. The existence of such a representation is guaranteed by (5). From this specific representation, it can be extracted that $\boldsymbol{A}$ will always have d eigenvalues at one. The modes with one eigenvalues play the role of a memory for the current best guess of the algorithm and are therefore necessary.

Remark 17 (*On the Constant c in Problem 1*). So far, we have discussed how to find algorithms of the form (1) which minimize the convergence rate ρ in the convergence estimate (4). However, the constant c_f is also of importance for the performance of algorithms. Our results guarantee that such a constant exists. By combining Theorem 10 and Lemma 12, it is even possible to find the upper bound

$$c \leq \sqrt{\beta/\alpha} = \sqrt{\frac{\lambda_{\max}(\boldsymbol{P})}{\lambda_{\min}(\boldsymbol{P})}(1 + \|\widetilde{\boldsymbol{L}}\|^2\|\boldsymbol{C}\|^2) + \frac{\|\widetilde{\boldsymbol{L}}\|\|\boldsymbol{C}\|^2}{2\lambda_{\min}(\boldsymbol{P})}}. \tag{15}$$

Firstly, notice that this upper bound does not directly depend on the function $f \in S(\boldsymbol{M}, \boldsymbol{L})$ but only on $\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C}$ and $\widetilde{\boldsymbol{L}} = \boldsymbol{L} - \boldsymbol{M}$. Secondly, notice that the upper bound grows with the condition number of the matrix $\boldsymbol{P}$. This is not ideal, since the condition number of $\boldsymbol{P}$ can go to infinity when ρ is minimized toward the feasibility limit. Numerical experiments showed that this can indeed happen. However, the true transient behavior of designed optimization algorithms was usually much better than the bound obtained by (15).

5. Conclusion

We presented a *convex* synthesis procedure for designing gradient-based algorithms based on a general class of Lur'e Lyapunov functions and linear matrix inequalities. The class of objective functions, which was considered, generalizes the class of strongly convex functions and offers the possibility to incorporate additional information into the algorithm design. It should be emphasized that this class of functions also includes non-convex functions — in particular functions which have a saddle point. The usefulness of our novel function class was demonstrated, firstly, by showing that additional information about the objective function can boost the convergence rate of algorithms considerably and, secondly, by showing that it can be used to design algorithms for solving optimization problems with linear equality constraints.

Open future research questions are for example the design of distributed algorithms or the design of optimization algorithms for problems with inequality constraints.

CRediT authorship contribution statement

Dennis Gramlich: Conceptualization, Methodology, Writing – original draft. **Christian Ebenbauer:** Supervision, Writing – review & editing. **Carsten W. Scherer:** Supervision, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The work of C.W. Scherer has been funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 - 390740016. He acknowledges the support by the Stuttgart Center for Simulation Science (SimTech).

Appendix. Proofs

A.1. Projections and pseudo inverses

The pseudo inverse $\boldsymbol{L}^\dagger$ and projection matrix $\boldsymbol{\Pi}_{\ker \boldsymbol{L}}$ and $\boldsymbol{\Pi}_{\mathrm{im}\, \boldsymbol{L}}$ onto the kernel/image of a symmetric matrix $\boldsymbol{L}$ are used at several places in this paper. Hence, some important formulas are summarized below. Let $\boldsymbol{A} = \boldsymbol{U}^\top \boldsymbol{\Sigma} \boldsymbol{V}$ be the singular value decomposition of a matrix $\boldsymbol{A}$, then

$$\boldsymbol{A}^\dagger = \underbrace{\begin{pmatrix} \boldsymbol{V}_1 \\ \boldsymbol{V}_2 \end{pmatrix}^\top}_{\boldsymbol{V}^\top} \underbrace{\left(\begin{array}{c|c} \begin{pmatrix} \sigma_1^{-1} & & \\ & \ddots & \\ & & \sigma_s^{-1} \end{pmatrix} & 0 \\ \hline 0 & 0 \end{array}\right)}_{\boldsymbol{\Sigma}^\dagger} \underbrace{\begin{pmatrix} \boldsymbol{U}_1 \\ \boldsymbol{U}_2 \end{pmatrix}}_{\boldsymbol{U}}$$

and $\boldsymbol{\Pi}_{\ker \boldsymbol{A}} = \boldsymbol{V}_2^\top \boldsymbol{V}_2$, $\boldsymbol{\Pi}_{\mathrm{im}\, \boldsymbol{A}} = \boldsymbol{U}_1 \boldsymbol{U}_1^\top$. We will particularly be interested in the following four identities for the projectors and pseudo inverses of a symmetric positive semidefinite matrix $\boldsymbol{L}$, $r \neq 0$:

$$\boldsymbol{\Pi}_{\mathrm{im}\, \boldsymbol{L}} = \boldsymbol{L}\boldsymbol{L}^\dagger = \boldsymbol{L}^\dagger \boldsymbol{L}, \tag{16}$$

$$\boldsymbol{I}_d = \boldsymbol{\Pi}_{\mathrm{im}\, \boldsymbol{L}} + \boldsymbol{\Pi}_{\ker \boldsymbol{L}}, \tag{17}$$

$$(\boldsymbol{L} + r\boldsymbol{\Pi}_{\ker \boldsymbol{L}})^{-1} = \boldsymbol{L}^{\dagger} + \frac{1}{r}\boldsymbol{\Pi}_{\ker \boldsymbol{L}}, \tag{18}$$

$$(\boldsymbol{L}^{\dagger} + r\boldsymbol{\Pi}_{\ker \boldsymbol{L}})^{-1} = \boldsymbol{L} + \frac{1}{r}\boldsymbol{\Pi}_{\ker \boldsymbol{L}}. \tag{19}$$

These identities follow from the singular value decomposition as shown above.

A.2. Proof of Theorem 1

Notice that the proof of Theorem 1 relies on the properties established in Section 2.2.

Step 1: The equivalence

$$f \in S(\boldsymbol{M}, \boldsymbol{L}) \Leftrightarrow g \in S(0, \boldsymbol{L} - \boldsymbol{M}), \tag{20}$$

where f and g are related by $f(z) = g(z) + \frac{1}{2}z^{\top}\boldsymbol{M}z$, will be used throughout this proof. Subtracting $\frac{1}{2}\|z_1 - z_2\|_{\boldsymbol{M}}^2$ from all terms in (2), we obtain

$$\begin{aligned} 0 &\le g(z_2) - g(z_1) + (\nabla g(z_1))^{\top}(z_1 - z_2) \\ &\le \frac{1}{2}\|z_1 - z_2\|_{\boldsymbol{L}-\boldsymbol{M}}^2, \end{aligned}$$

which is the characterization of $g \in S(0, \boldsymbol{L} - \boldsymbol{M})$.

Step 2: (From Problems 1 to 2, convergence rate) Assume that $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ solves Problem 1. We prove the convergence rate of the algorithm defined by $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ in Problem 2. Let $g_1 \in S_0(0, \widetilde{\boldsymbol{L}}) = S_0(0, \boldsymbol{L} - \boldsymbol{M})$ be an arbitrary function. Then $f_1(z) := g_1(z) + \frac{1}{2}z^{\top}\boldsymbol{M}z$ is an element of $S(\boldsymbol{M}, \boldsymbol{L})$ due to (20). Moreover, $\nabla f_1(0) = 0$. Now consider the iterates of algorithm (1) with the parameters $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ for the objective function g_1:

$$\begin{aligned} x_{k+1} &= \widetilde{\boldsymbol{A}}x_k + \boldsymbol{B}\nabla g_1(\boldsymbol{C}x_k) \\ &= \boldsymbol{A}x_k + \boldsymbol{B}(\nabla g_1(\boldsymbol{C}x_k) + \boldsymbol{M}\boldsymbol{C}x_k) \\ &= \boldsymbol{A}x_k + \boldsymbol{B}\nabla f_1(\boldsymbol{C}x_k). \end{aligned}$$

Since the matrix triple $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ is assumed to solve Problem 1 and $f_1 \in S(\boldsymbol{M}, \boldsymbol{L})$, we know that x_k converges to $x_{f_1}^*$ at rate ρ for any $x_0 \in \mathbb{R}^d$. Notice that $x_{f_1}^*$ must be equal to zero, because zero is a globally attractive fixed-point of the considered iteration (since $\boldsymbol{A}0 + \boldsymbol{B}\nabla f_1(\boldsymbol{C}0) = 0$ by $\nabla f_1(0) = 0$) and such a fixed point must be unique.

Step 3: (From Problems 1 to 2, constraint) It remains to show satisfaction of the constraint (5). For this purpose define $f_2 \in S(\boldsymbol{M}, \boldsymbol{L})$ for some $z_{f_2}^* \in \mathbb{R}^d$ as $f_2(z) := \frac{1}{2}(z - z_{f_2}^*)^{\top}\boldsymbol{M}(z - z_{f_2}^*)$ and note that it satisfies $\nabla f_2(z_{f_2}^*) = 0$. By assumption, Problem 1 is solved, meaning that the iterates of algorithm (1)

$$\begin{aligned} x_{k+1} &= \boldsymbol{A}x_k + \boldsymbol{B}\nabla f_2(\boldsymbol{C}x_k) = \boldsymbol{A}x_k + \boldsymbol{B}\boldsymbol{M}(\boldsymbol{C}x_k - z_{f_2}^*) \\ &= \widetilde{\boldsymbol{A}}x_k - \boldsymbol{B}\boldsymbol{M}z_{f_2}^* \end{aligned}$$

converge to $x_{f_2}^*$ for any x_0 at rate ρ. This implies firstly, that $x_{f_2}^*$ is a solution of the fixed point equation

$$x_{f_2}^* = \widetilde{\boldsymbol{A}}x_{f_2}^* - \boldsymbol{B}\boldsymbol{M}z_{f_2}^* \tag{21}$$

and secondly that the spectral radius of $\widetilde{\boldsymbol{A}}$ must be no larger than ρ. Hence, $\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n$ is non-singular. Then, (21) can be solved for $x_{f_2}^*$ yielding

$$x_{f_2}^* = (\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M}z_{f_2}^* \overset{(3)}{=} \boldsymbol{C}x_{f_2}^*.$$

Since $z_{f_2}^*$ is arbitrary, $\boldsymbol{C}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M} = \boldsymbol{I}_d$ must hold.

Step 4: (From Problems 2 to 1, convergence rate) Now assume, that $(\widetilde{\boldsymbol{A}}, \widetilde{\boldsymbol{B}}, \widetilde{\boldsymbol{C}})$ is a solution of Problem 2. We prove that $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ solves Problem 1. For that, we first consider all functions $f \in S(\boldsymbol{M}, \boldsymbol{L})$ for which there exists a critical point z_f^*. First let $f_3 \in S(\boldsymbol{M}, \boldsymbol{L})$ be given such that there exists $z_{f_3}^*$ with $\nabla f(z_{f_3}^*) = 0$. Then g_3 defined by $g_3(z) = f_3(z + z_{f_3}^*) - \frac{1}{2}z^{\top}\boldsymbol{M}z$ is an element of $S_0(0, \boldsymbol{L} - \boldsymbol{M}) = S_0(0, \widetilde{\boldsymbol{L}})$ due to (20). Hence, the algorithm

$$\tilde{x}_{k+1} = \widetilde{\boldsymbol{A}}\tilde{x}_k + \boldsymbol{B}\nabla g_3(\boldsymbol{C}\tilde{x}_k) \tag{22}$$

converges to zero at rate ρ for any $x_0 \in \mathbb{R}^n$. Now add $x_{f_3}^* := (\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M}z_{f_3}^*$ on both sides of the above equation and consider the new sequence $x_k := \tilde{x}_k + x_{f_3}^*$:

$$\begin{aligned} x_{k+1} &= \tilde{x}_{k+1} + x_{f_3}^* \overset{(22)}{=} \widetilde{\boldsymbol{A}}\tilde{x}_k + \boldsymbol{B}\nabla g_3(\boldsymbol{C}\tilde{x}_k) + x_{f_3}^* \\ &= \underbrace{\widetilde{\boldsymbol{A}}(\tilde{x}_k + x_{f_3}^*)}_{=\boldsymbol{A}x_k + \boldsymbol{B}\boldsymbol{M}\boldsymbol{C}x_k} + \boldsymbol{B}\nabla g_3(\boldsymbol{C}\tilde{x}_k) + \underbrace{x_{f_3}^* - \widetilde{\boldsymbol{A}}x_{f_3}^*}_{\overset{(5)}{=} -\boldsymbol{B}\boldsymbol{M}z_{f_3}^*} \\ &= \boldsymbol{A}x_k + \boldsymbol{B}\boldsymbol{M}(\boldsymbol{C}x_k - z_{f_3}^*) + \boldsymbol{B}\nabla g_3(\boldsymbol{C}\tilde{x}_k) \\ &\overset{(5)}{=} \boldsymbol{A}x_k + \underbrace{\boldsymbol{B}\boldsymbol{M}\boldsymbol{C}\tilde{x}_k + \boldsymbol{B}\nabla g_3(\boldsymbol{C}\tilde{x}_k)}_{=\boldsymbol{B}\nabla f_3(\boldsymbol{C}x_k)} \end{aligned}$$

This is the equation for the iterates x_k of the algorithm defined by $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ and f_3. Since $\tilde{x}_k$ goes to zero at rate ρ, so does x_k go to $x_{f_3}^*$.

Step 5: (From Problems 2 to 1, existence and uniqueness of critical points) Finally, we argue that there cannot be an element of $S(\boldsymbol{M}, \boldsymbol{L})$ with no critical point: If there were an $f \in S(\boldsymbol{M}, \boldsymbol{L})$ with two critical points, then the arguments from Step 4 would prove convergence of (1) to both critical points, which cannot be true. Hence, there exists no such function in $S(\boldsymbol{M}, \boldsymbol{L})$. Consequently, Theorem 8 guarantees that any function in $S(\boldsymbol{M}, \boldsymbol{L})$ has a critical point and thus Step 4 covers all cases for which Problem 2 can be solved. Therefore, $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ solve Problem 1. □

A.3. Proof of Lemma 2

(2) ⇒ (1): The key to prove this statement is that the second term in inequality (1) can be written as the integral

$$\begin{aligned} &\int_0^1 (\nabla f(z_1 + \tau(z_2 - z_1)) - \nabla f(z_1))^{\top}(z_2 - z_1)\,\mathrm{d}\tau \\ &= f(z_2) - f(z_1) - (\nabla f(z_1))^{\top}(z_2 - z_1). \end{aligned} \tag{23}$$

Using (2), the integrand can be upper and lower bounded as

$$\begin{aligned} 0 &\le (\nabla f(z_1 + \tau(z_2 - z_1)) - \nabla f(z_1))^{\top}(z_2 - z_1) \\ &= \frac{1}{\tau}(\nabla f(z_1 + \tau(z_2 - z_1)) - \nabla f(z_1))^{\top}\tau(z_2 - z_1) \\ &\le \frac{1}{\tau}\|\tau(z_1 - z_2)\|_{\boldsymbol{L}}^2 = \tau\|z_1 - z_2\|_{\boldsymbol{L}}^2. \end{aligned} \tag{24}$$

Applying this estimate to the integral expression, we obtain

$$\begin{aligned} &f(z_2) - f(z_1) - (\nabla f(z_1))^{\top}(z_2 - z_1) \\ &\overset{(23)}{=} \int_0^1 (\nabla f(z_1 + \tau(z_2 - z_1)) - \nabla f(z_1))^{\top}(z_2 - z_1)\,\mathrm{d}\tau \\ &\overset{(24)}{\le} \int_0^1 \tau\|z_1 - z_2\|_{\boldsymbol{L}}^2\,\mathrm{d}\tau = \|z_1 - z_2\|_{\boldsymbol{L}}^2 \end{aligned}$$

and

$$\begin{aligned} 0 &\le \int_0^1 \underbrace{(\nabla f(z_1 + \tau(z_2 - z_1)) - \nabla f(z_1))^{\top}(z_2 - z_1)}_{\overset{(24)}{\ge} 0}\,\mathrm{d}\tau \\ &= f(z_2) - f(z_1) - (\nabla f(z_1))^{\top}(z_2 - z_1), \end{aligned}$$

which together imply (1).

(1) ⇒ (3): Let $f \in C^1(\mathbb{R}^d)$ fulfill (1). Define $g(z) = f(z) - (\nabla f(z_1))^{\top}z$. Then $g \in S(0, \boldsymbol{L})$ and $\nabla g(z_1) = 0$ hold. Thus, z_1 is a minimizer of g and we have

$$g(z_1) - g(z_2) \le g(z_2 - \tilde{z}) - g(z_2)$$

for any vector $\tilde{z} \in \mathbb{R}^d$. Here, we can add the term $\nabla g(z_2)^\top \tilde{z}$ on both sides of the inequality to get

$$g(z_1) - g(z_2) + \nabla g(z_2)^\top \tilde{z} \le g(z_2 - \tilde{z}) - g(z_2) + \nabla g(z_2)^\top \tilde{z} \overset{(1)}{\le} \frac{1}{2}\|\tilde{z}\|_{\boldsymbol{L}}^2.$$

By substituting $\boldsymbol{A}\nabla g(z_2)$ for $\tilde{z}$, we get

$$g(z_1) - g(z_2) + \nabla g(z_2)^\top \boldsymbol{A}\nabla g(z_2) \le \frac{1}{2}\|\boldsymbol{A}\nabla g(z_2)\|_{\boldsymbol{L}}^2$$

for any matrix $\boldsymbol{A} \in \mathbb{R}^{d\times d}$ or equivalently

$$\nabla g(z_2)^\top \boldsymbol{A}\nabla g(z_2) - \frac{1}{2}\|\boldsymbol{A}\nabla g(z_2)\|_{\boldsymbol{L}}^2 \le g(z_2) - g(z_1).$$

Now, we substitute $g(z_2) = f(z_2) - (\nabla f(z_1))^\top z_2$ and obtain

$$(\nabla f(z_1) - \nabla f(z_2))^\top \boldsymbol{A}(\nabla f(z_1) - \nabla f(z_2)) - \frac{1}{2}\|\boldsymbol{A}(\nabla f(z_1) - \nabla f(z_2))\|_{\boldsymbol{L}}^2 \le f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2).$$

For $\boldsymbol{A} = \boldsymbol{L}^\dagger$, this is equivalent to

$$\frac{1}{2}\parallel \nabla f(z_1) - \nabla f(z_2) \parallel_{\boldsymbol{L}^\dagger}^2 \le f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2).$$

In the case $\boldsymbol{A} = r\boldsymbol{\Pi}_{\ker \boldsymbol{L}}$, the result is

$$r(\nabla f(z_1) - \nabla f(z_2))^\top \boldsymbol{\Pi}_{\ker \boldsymbol{L}}(\nabla f(z_1) - \nabla f(z_2)) \le f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2),$$

which implies $\boldsymbol{\Pi}_{\ker \boldsymbol{L}}(\nabla f(z_1) - \nabla f(z_2)) = 0$, because r can be chosen arbitrarily large.

(3) ⇒ (4): Adding $\frac{1}{2}\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}^2 \le f(z_2) - f(z_1) + (\nabla f(z_1))^\top (z_1 - z_2)$ and $\frac{1}{2}\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}^2 \le f(z_1) - f(z_2) + (\nabla f(z_2))^\top (z_2 - z_1)$ yields inequality in (4).

(4) ⇒ (2): Let $f \in C^1(\mathbb{R}^d)$ fulfill (4). Then

$$\sqrt{\boldsymbol{L}}\sqrt{\boldsymbol{L}^\dagger}(\nabla f(z_1) - \nabla f(z_2)) = (\nabla f(z_1) - \nabla f(z_2))$$

holds for all $z_1, z_2 \in \mathbb{R}^d$, since $\boldsymbol{\Pi}_{\ker \boldsymbol{L}}(\nabla f(z_1) - \nabla f(z_2)) = 0$ implies, that $\nabla f(z_1) - \nabla f(z_2)$ is in the image of $\boldsymbol{L}$. This observation can be used to derive the bound using the Cauchy–Schwarz-Inequality (CSI)

$$\begin{aligned}\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}^2 &\overset{(4)}{\le} (\nabla f(z_1) - \nabla f(z_2))^\top (z_1 - z_2)\\ &= (\nabla f(z_1) - \nabla f(z_2))^\top \sqrt{\boldsymbol{L}^\dagger}\sqrt{\boldsymbol{L}}(z_1 - z_2)\\ &\overset{\text{CSI}}{\le} \|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}\|z_1 - z_2\|_{\boldsymbol{L}},\end{aligned}$$

which implies $\|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger} \le \|z_1 - z_2\|_{\boldsymbol{L}}$. Now, f fulfills (2), because

$$\begin{aligned}&(\nabla f(z_1) - \nabla f(z_2))^\top (z_1 - z_2)\\ &\overset{\text{CSI}}{\le} \|\nabla f(z_1) - \nabla f(z_2)\|_{\boldsymbol{L}^\dagger}\|z_1 - z_2\|_{\boldsymbol{L}}\\ &\le \|z_1 - z_2\|_{\boldsymbol{L}}^2.\end{aligned}$$

□

A.4. Proof of Lemma 3

(1) ⇒ (2): Let $\boldsymbol{Q}$ with $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ be given and let $(\lambda_i^{(\boldsymbol{M})})_{i=1}^d$, $(\lambda_i^{(\boldsymbol{Q})})_{i=1}^d$, $(\lambda_i^{(\boldsymbol{L})})_{i=1}^d$ be the eigenvalues of those matrices in ascending order. It follows from $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ and the theorem of Courant–Fischer that

$$\lambda_1^{(\boldsymbol{M})} \le \lambda_1^{(\boldsymbol{Q})} \le \lambda_1^{(\boldsymbol{L})}, \dots, \lambda_d^{(\boldsymbol{M})} \le \lambda_d^{(\boldsymbol{Q})} \le \lambda_d^{(\boldsymbol{L})}$$

holds. Since $\lambda_i^{(\boldsymbol{M})}$ and $\lambda_i^{(\boldsymbol{L})}$ always have the same sign and are not equal to zero by assumption, the values $\lambda_i^{(\boldsymbol{Q})}$ cannot be zero for any i. Hence, no eigenvalue of $\boldsymbol{Q}$ can be zero and hence, $\boldsymbol{Q}$ is invertible.

(2) ⇒ (3): To show the first statement, consider the case $\boldsymbol{Q} = \frac{1}{2}(\boldsymbol{M}+\boldsymbol{L})$. Then, it holds that $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ and hence, $\boldsymbol{Q} = \frac{1}{2}(\boldsymbol{M}+\boldsymbol{L})$ is invertible. To show the second statement, consider the case $\boldsymbol{Q} = \frac{1}{2}(\boldsymbol{M} + \boldsymbol{L}) + \frac{\alpha}{2}(\boldsymbol{L} - \boldsymbol{M})$. For $\alpha \in [-1, 1]$, it holds that $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ and thus

$$0 \neq \det\left(\frac{1}{2}(\boldsymbol{M} + \boldsymbol{L}) + \frac{\alpha}{2}(\boldsymbol{L} - \boldsymbol{M})\right) \quad \forall \alpha \in [-1, 1].$$

By non-singularity of $(\boldsymbol{M} + \boldsymbol{L})$, the factor $\det \frac{1}{2}(\boldsymbol{M} + \boldsymbol{L})$ can be pulled out of the above expression, which gives

$$0 \neq \det\left(\frac{1}{2}(\boldsymbol{M} + \boldsymbol{L})\right)\det\left(\boldsymbol{I} + \alpha(\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})\right)$$

and consequently

$$0 \neq \det\left(\boldsymbol{I} + \alpha(\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})\right) \quad \forall \alpha \in [-1, 1].$$

This implies, that $(\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})$ cannot have eigenvalues in $\mathbb{R}\backslash] - 1, 1[$. However, since $(\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})$ is similar to the symmetric matrix $\sqrt{\boldsymbol{L} - \boldsymbol{M}}(\boldsymbol{M} + \boldsymbol{L})^{-1}\sqrt{\boldsymbol{L} - \boldsymbol{M}}$, all of its eigenvalues have to be real. (Note that $\sqrt{\boldsymbol{L} - \boldsymbol{M}}$ exists because $\boldsymbol{L} - \boldsymbol{M}$ is positive semi-definite.) Hence, all eigenvalues of $(\boldsymbol{M}+\boldsymbol{L})^{-1}(\boldsymbol{L}-\boldsymbol{M})$ have to be in $] - 1, 1[$ and thus also $\rho((\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})) < 1$ holds.

(3) ⇒ (4): Suppose that $\boldsymbol{M}$ is not invertible, i.e. there exists a vector $z \in \mathbb{R}^d \setminus \{0\}$ with $\boldsymbol{M}z = 0$. Then

$$(\boldsymbol{L} + \boldsymbol{M})z = (\boldsymbol{L} - \boldsymbol{M})z \Rightarrow z = (\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})z$$

implies that z is an eigenvector to the eigenvalue 1 of $(\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})$, which contradicts $\rho((\boldsymbol{M} + \boldsymbol{L})^{-1}(\boldsymbol{L} - \boldsymbol{M})) < 1$. Hence $\boldsymbol{M}$ is non-singular.

Next we show $\sigma(\boldsymbol{M}^{-1}\boldsymbol{L}) \subseteq \mathbb{R}_{>0}$. Consider the identity

$$\begin{aligned}(\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M}) &= \boldsymbol{I} - 2(\boldsymbol{L} + \boldsymbol{M})^{-1}\boldsymbol{M}\\ &= \boldsymbol{I} - 2(\boldsymbol{M}^{-1}\boldsymbol{L} + \boldsymbol{I})^{-1}.\end{aligned}$$

Suppose, that $\boldsymbol{M}^{-1}\boldsymbol{L}$ has an eigenvalue λ with associated eigenvector v. Then $\boldsymbol{M}^{-1}\boldsymbol{L} + \boldsymbol{I}$ has eigenvalue $\lambda + 1$ with eigenvector v and $(\boldsymbol{M}^{-1}\boldsymbol{L} + \boldsymbol{I})^{-1}$ has eigenvalue $\frac{1}{\lambda+1}$ with eigenvector v. Thus

$$\begin{aligned}(\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})v &= (\boldsymbol{I} - 2(\boldsymbol{M}^{-1}\boldsymbol{L} + \boldsymbol{I})^{-1})v\\ &= v - \frac{2}{\lambda + 1}v = \frac{\lambda - 1}{\lambda + 1}v.\end{aligned}$$

Hence, $\frac{\lambda-1}{\lambda+1}$ is an eigenvalue of $(\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M})$ and thus it is in $] - 1, 1[$. This implies $\lambda \in \mathbb{R}_{>0}$. Hence $\sigma(\boldsymbol{M}^{-1}\boldsymbol{L}) \subseteq \mathbb{R}_{>0}$ holds true.

(4) ⇒ (5): Suppose, that $\boldsymbol{L}\boldsymbol{M}^{-1}$ has only positive eigenvalues. Then there exists a symmetric positive definite matrix $\boldsymbol{P} \in \mathbb{R}^{d\times d}$ such that the Lyapunov inequality $\boldsymbol{P}\boldsymbol{L}\boldsymbol{M}^{-1} + \boldsymbol{M}^{-1}\boldsymbol{L}\boldsymbol{P} \succ 0$ is satisfied. A congruence transform with $\boldsymbol{M}$ yields $\boldsymbol{M}\boldsymbol{P}\boldsymbol{L} + \boldsymbol{L}\boldsymbol{P}\boldsymbol{M} \succ 0$. By Lemma 19 we can infer that $\boldsymbol{M}$ and $\boldsymbol{L}$ are congruent.

(5) ⇒ (1): By Sylvester's Law of Inertia, matrices have the same signature, if and only if they are congruent. □

A.5. Proof of Proposition 6

We prove the contraction property of the map $\phi : z \mapsto z - 2(\boldsymbol{M} + \boldsymbol{L})^{-1}\nabla f(z)$, by using the norm $\|z\|_{\boldsymbol{P}}^2 = z^\top \boldsymbol{P} z$, where $\boldsymbol{P} = (\boldsymbol{L} + \boldsymbol{M})((\boldsymbol{L} - \boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})})(\boldsymbol{L} + \boldsymbol{M})$. In a first step, rewrite ϕ as

$$\begin{aligned}\phi(z) &= (\boldsymbol{L}+\boldsymbol{M})^{-1}(\boldsymbol{L}+\boldsymbol{M})z - 2(\boldsymbol{L}+\boldsymbol{M})^{-1}\nabla f(z)\\ &= (\boldsymbol{L}+\boldsymbol{M})^{-1}\left((\boldsymbol{L}-\boldsymbol{M})z - 2(\nabla f(z)-\boldsymbol{M}z)\right)\\ &= (\boldsymbol{L}+\boldsymbol{M})^{-1}\left((\boldsymbol{L}-\boldsymbol{M})z - 2\nabla g(z)\right)\end{aligned}$$

with $g \in S(0, \boldsymbol{L}-\boldsymbol{M})$ defined by $g(z) := f(z) - \frac{1}{2}z^\top \boldsymbol{M}z$. For the following computation, remember that

$$\begin{aligned}\sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger(\boldsymbol{L}-\boldsymbol{M})} &= \sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger}\sqrt{\boldsymbol{L}-\boldsymbol{M}}^2\\ &\overset{(16)}{=} \boldsymbol{\Pi}_{\mathrm{im}\,\boldsymbol{L}-\boldsymbol{M}}\sqrt{\boldsymbol{L}-\boldsymbol{M}} = \sqrt{\boldsymbol{L}-\boldsymbol{M}}\end{aligned} \tag{25}$$

and

$$\begin{aligned}&\sqrt{\boldsymbol{L}-\boldsymbol{M}}\sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger}(\nabla g(z_1)-\nabla g(z_2))\\ &\overset{(16)}{=} \boldsymbol{\Pi}_{\mathrm{im}\,\boldsymbol{L}-\boldsymbol{M}}(\nabla g(z_1)-\nabla g(z_2)) \qquad (26)\\ &\overset{(17)}{=} (\boldsymbol{I}-\boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}})(\nabla g(z_1)-\nabla g(z_2)) \overset{\text{Proposition 6}}{=} \nabla g(z_1)-\nabla g(z_2)\end{aligned}$$

hold. Consider now

$$\begin{aligned}&\|\phi(z_1)-\phi(z_2)\|_{\boldsymbol{P}}^2\\ &= \left\|\sqrt{\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})}\right)}(\boldsymbol{L}+\boldsymbol{M})\left(\phi(z_1)-\phi(z_2)\right)\right\|^2\\ &= \Big\|\sqrt{\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})}\right)}(\boldsymbol{L}+\boldsymbol{M})(\boldsymbol{L}+\boldsymbol{M})^{-1}\\ &\qquad ((\boldsymbol{L}-\boldsymbol{M})(z_1-z_2) - 2(\nabla g(z_1)-\nabla g(z_2)))\Big\|^2\\ &\overset{(\star)}{=} \|\sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger}\left((\boldsymbol{L}-\boldsymbol{M})(z_1-z_2) - 2(\nabla g(z_1)-\nabla g(z_2))\right)\|^2\\ &\overset{(25)}{=} \|\sqrt{\boldsymbol{L}-\boldsymbol{M}}(z_1-z_2) - 2\sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger}(\nabla g(z_1)-\nabla g(z_2))\|^2\\ &\overset{(26)}{=} \underbrace{4\|\nabla g(z_1)-\nabla g(z_2)\|^2_{(\boldsymbol{L}-\boldsymbol{M})^\dagger} - 4(\nabla g(z_1)-\nabla g(z_2))^\top(z_1-z_2)}_{\le 0}\\ &\qquad + \|\sqrt{\boldsymbol{L}-\boldsymbol{M}}(z_1-z_2)\|^2 \le \|z_1-z_2\|^2_{(\boldsymbol{L}-\boldsymbol{M})}.\end{aligned}$$

Concerning ($\star$) notice, that the kernel projector has no contribution, since the products are all zero and the under-braced expression being non-positive follows from Lemma 2. Finally, by Lemma 18 we know that for any $\rho > \rho_{\mathrm{grad}}$ there exists some $r \in \mathbb{R}_{>0}$ such that $\boldsymbol{L}-\boldsymbol{M} \preceq \rho^2\boldsymbol{P}$ holds. Hence, we can overestimate $\|z_1-z_2\|^2_{(\boldsymbol{L}-\boldsymbol{M})}$ by $\rho^2\|z_1-z_2\|^2_{\boldsymbol{P}}$ (by choosing a sufficient value r) resulting in the final estimate

$$\|\phi(z_1)-\phi(z_2)\|^2_{\boldsymbol{P}} \le \|z_1-z_2\|^2_{(\boldsymbol{L}-\boldsymbol{M})} \le \rho^2\|z_1-z_2\|^2_{\boldsymbol{P}}. \qquad \square$$

A.6. Proof of Theorem 8

Note that $S(\boldsymbol{M}, \boldsymbol{L})$ is not empty since $\boldsymbol{M} \preceq \boldsymbol{L}$. It remains to show that the three statements in the theorem are equivalent under the condition $\boldsymbol{M} \preceq \boldsymbol{L}$.

(1) ⇒ (2) and (1) ⇒ (3): Assume $\boldsymbol{M} \preceq_c \boldsymbol{L}$ are non-singular. Let $f \in S(\boldsymbol{M}, \boldsymbol{L})$ be given. Then, by Proposition 6, the mapping $\phi : z \mapsto z - 2(\boldsymbol{M}+\boldsymbol{L})^{-1}\nabla f(z)$ is a contraction on $\mathbb{R}^d$ and $(\boldsymbol{M}+\boldsymbol{L})$ is non-singular. By the Banach fixed point theorem the mapping ϕ has exactly one fixed point z_f^* with $\phi(z_f^*) = z_f^* \Leftrightarrow \nabla f(z_f^*) = 0$. This implies (2) and (3).

¬(1) ⇒ ¬(2) and ¬(1) ⇒ ¬(3): Suppose that $\boldsymbol{M} \preceq_c \boldsymbol{L}$ does not hold, but $\boldsymbol{M} \preceq \boldsymbol{L}$ holds. Then there exists $\boldsymbol{Q} = \boldsymbol{Q}^\top \in \mathbb{R}^{d\times d}$ with $\boldsymbol{M} \preceq \boldsymbol{Q} \preceq \boldsymbol{L}$ and $\det \boldsymbol{Q} = 0$ by Lemma 3. Let $v \in \mathbb{R}^d \setminus \{0\}$ be an element of the kernel of $\boldsymbol{Q}$. Then the function $f_1 \in S(\boldsymbol{M}, \boldsymbol{L})$ defined by $f_1(z) = \frac{1}{2}z^\top\boldsymbol{Q}z + v^\top z$ has no critical point with $\nabla f(z) = 0$, because otherwise

$$v^\top\nabla f_1(z) = v^\top(\boldsymbol{Q}z + v) = v^\top\boldsymbol{Q}z + v^\top v = \|v\|^2$$

would have to be zero. At the same time, the function $f_2 \in S(\boldsymbol{M}, \boldsymbol{L})$ defined by $f_2(z) = \frac{1}{2}z^\top\boldsymbol{Q}z$ has infinitely many critical points with $\nabla f_2(z) = 0$, because any point $z = rv$ with $r \in \mathbb{R}$ is a critical point of f_2 by $\nabla f_2(z) = r\boldsymbol{Q}v = 0$. $\square$

A.7. Proof of Lemma 11

Define the abbreviations $w = \nabla f(\boldsymbol{C}x)$, $w^+ = \nabla f(\boldsymbol{C}x^+)$, $x^+ = \boldsymbol{A}x + \boldsymbol{B}w$ and $\widetilde{\boldsymbol{L}} = \boldsymbol{L}-\boldsymbol{M}$. With that the ρ-weighted increment of the Lyapunov function is

$$\begin{aligned}&V_f(x^+) - \rho^2 V_f(x)\\ &= \begin{pmatrix}x^+\\ w^+\end{pmatrix}^\top\begin{pmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22}\end{pmatrix}\begin{pmatrix}x^+\\ w^+\end{pmatrix} - \rho^2\begin{pmatrix}x\\ w\end{pmatrix}^\top\begin{pmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22}\end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}\\ &\underbrace{+ f(\boldsymbol{C}x^+) - f(0) - \frac{1}{2}\|w^+\|^2_{\widetilde{\boldsymbol{L}}^\dagger} - \rho^2\left(f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\|w\|^2_{\widetilde{\boldsymbol{L}}^\dagger}\right)}_{I}.\end{aligned}$$

To upper bound expression I, we use the estimate

$$\begin{aligned}&\underbrace{-\rho^2}_{\le -\lambda}\underbrace{\left(f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\|w\|^2_{\widetilde{\boldsymbol{L}}^\dagger}\right)}_{\ge 0}\\ &\le -\lambda\left(f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\|w\|^2_{\widetilde{\boldsymbol{L}}^\dagger}\right),\end{aligned}$$

which we can use to obtain

$$\begin{aligned}I &\le (1-\lambda)\underbrace{\left(f(\boldsymbol{C}x^+) - f(0) + \frac{1}{2}\|w^+\|^2_{\widetilde{\boldsymbol{L}}^\dagger}\right)}_{\overset{\text{Lemma 2}}{\le}(w^+)^\top(\boldsymbol{C}x^+-0)}\\ &\quad + \lambda\underbrace{\left(f(\boldsymbol{C}x^+) - f(\boldsymbol{C}x) + \frac{1}{2}\|w^+ - w\|^2_{\widetilde{\boldsymbol{L}}^\dagger}\right)}_{\overset{\text{Lemma 2}}{\le}(w^+)^\top(\boldsymbol{C}x^+-\boldsymbol{C}x)}\\ &\quad \underbrace{-\frac{(2-\lambda)}{2}\|w^+\|^2_{\widetilde{\boldsymbol{L}}^\dagger} - \frac{\lambda}{2}\|w^+ - w\|^2_{\widetilde{\boldsymbol{L}}^\dagger} + \frac{\lambda}{2}\|w\|^2_{\widetilde{\boldsymbol{L}}^\dagger}}_{=(w^+)^\top\widetilde{\boldsymbol{L}}^\dagger(w^+-\lambda w)}\\ &\le (1-\lambda)(w^+)^\top\boldsymbol{C}x^+ + \lambda(w^+)^\top(\boldsymbol{C}x^+ - \boldsymbol{C}x)\\ &\quad - (w^+)^\top\widetilde{\boldsymbol{L}}^\dagger(w^+ - \lambda w)\\ &= (w^+)^\top\left(\boldsymbol{C}x^+ - \lambda\boldsymbol{C}x - \widetilde{\boldsymbol{L}}^\dagger(w^+ - \lambda w)\right).\end{aligned}$$

Now, this estimate for expression I can be used to upper bound $V_f(x^+) - \rho^2 V_f(x)$ as follows:

$$\begin{aligned}V(x^+) - \rho^2 V(x) &\le \begin{pmatrix}x^+\\ w^+\end{pmatrix}^\top\begin{pmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22}\end{pmatrix}\begin{pmatrix}x^+\\ w^+\end{pmatrix}\\ &\quad - \rho^2\begin{pmatrix}x\\ w\end{pmatrix}^\top\begin{pmatrix}\boldsymbol{P}_{11} & \boldsymbol{P}_{12}\\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22}\end{pmatrix}\begin{pmatrix}x\\ w\end{pmatrix}\\ &\quad + (w^+)^\top\left(\boldsymbol{C}x^+ - \lambda\boldsymbol{C}x - \widetilde{\boldsymbol{L}}^\dagger(w^+ - \lambda w)\right),\end{aligned}$$

which corresponds to the inequality in Lemma 11. $\square$

A.8. Proof of Lemma 12

Step 1 (upper bound). We can establish an upper bound on $f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top\widetilde{\boldsymbol{L}}^\dagger\nabla f(\boldsymbol{C}x)$ by the estimates:

$$\begin{aligned}&f(\boldsymbol{C}x) - f(0) \underbrace{-\frac{1}{2}\nabla f(\boldsymbol{C}x)^\top\widetilde{\boldsymbol{L}}^\dagger\nabla f(\boldsymbol{C}x)}_{\le 0}\\ &\overset{(\star)}{\le} \underbrace{f(\boldsymbol{C}x) - f(0) - (\nabla f(0))^\top(\boldsymbol{C}x - 0)}_{\overset{\text{Lemma 2}}{\le}\frac{1}{2}\|\boldsymbol{C}x-0\|^2_{\widetilde{\boldsymbol{L}}}}\\ &\le \frac{\|\boldsymbol{L}-\boldsymbol{M}\|}{2}\|\boldsymbol{C}x - 0\|^2.\end{aligned}$$

Note, that in $(\star)$ the term $(\nabla f(0))^\top(\boldsymbol{C}x - 0)$ can be added because $\nabla f(0) = 0$. This allows the following bound on V_f:

$$\begin{aligned} V_f(x) &= \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{pmatrix} \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} \\ &\quad + f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top \widetilde{\boldsymbol{L}}^\dagger \nabla f(\boldsymbol{C}x) \\ &\leq \lambda_{\max}(\boldsymbol{P}) \left\| \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} \right\|^2 + \frac{\|\widetilde{\boldsymbol{L}}\|}{2}\|\boldsymbol{C}x\|^2 \\ &= \lambda_{\max}(\boldsymbol{P})(\|x\|^2 + \underbrace{\|\nabla f(\boldsymbol{C}x)\|^2}_{\leq \|\widetilde{\boldsymbol{L}}\|^2\|\boldsymbol{C}x\|^2}) + \frac{\|\widetilde{\boldsymbol{L}}\|}{2}\|\boldsymbol{C}x\|^2 \\ &\leq \left(\lambda_{\max}(\boldsymbol{P})(1 + \|\widetilde{\boldsymbol{L}}\|^2\|\boldsymbol{C}\|^2) + \frac{\|\widetilde{\boldsymbol{L}}\|\|\boldsymbol{C}\|^2}{2}\right)\|x\|^2 \\ &= \beta\|x\|^2. \end{aligned}$$

Step 2 (lower bound). We can establish a lower bound on $f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top \widetilde{\boldsymbol{L}}^\dagger \nabla f(\boldsymbol{C}x)$ by the estimate

$$\begin{aligned} 0 &\leq f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\|\nabla f(\boldsymbol{C}x)\|^2_{\widetilde{\boldsymbol{L}}^\dagger} - (\nabla f(0))^\top \boldsymbol{C}x \\ &= f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top \widetilde{\boldsymbol{L}}^\dagger \nabla f(\boldsymbol{C}x), \end{aligned}$$

where the inequality sign follows from Lemma 2 and the equality sign follows from the fact $\nabla f(0) = 0$. This allows now the following lower bound on V_f:

$$\begin{aligned} V_f(x) &= \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{pmatrix} \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} \\ &\quad \underbrace{+ f(\boldsymbol{C}x) - f(0) - \frac{1}{2}\nabla f(\boldsymbol{C}x)^\top \widetilde{\boldsymbol{L}}^\dagger \nabla f(\boldsymbol{C}x)}_{\geq 0} \\ &\geq \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix}^\top \begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{pmatrix} \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} \\ &\geq \lambda_{\min}(\boldsymbol{P}) \left\| \begin{pmatrix} x \\ \nabla f(\boldsymbol{C}x) \end{pmatrix} \right\|^2 \\ &\geq \lambda_{\min}(\boldsymbol{P})\|x\|^2 = \alpha\|x\|^2. \end{aligned}$$

□

A.9. Proof of Theorem 13

First recall that Theorem 1 shows that an algorithm with parameters $(\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{C})$ has convergence rate ρ for $S(\boldsymbol{M}, \boldsymbol{L})$ if an algorithm with parameters $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$, which satisfy the constraint (5), has convergence rate ρ for $S_0(0, \widetilde{\boldsymbol{L}})$. Hence, in the following we show convergence of $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ for $S_0(0, \widetilde{\boldsymbol{L}})$. By Theorem 10, an algorithm defined by $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ is asymptotically stable and has convergence rate ρ, if there exists a Lyapunov function $V_f : \mathbb{R}^n \to \mathbb{R}$, such that

$$\begin{aligned} &\alpha\|x - x_f^*\|^2 \leq V_f(x) \leq \beta\|x - x_f^*\|^2, \\ &V_f(x^+) - \rho^2 V_f(x) \leq 0 \end{aligned}$$

holds for all $x \in \mathbb{R}^n$ and $f \in S_0(0, \widetilde{\boldsymbol{L}})$ with $\beta \geq \alpha > 0$. The considered class of Lyapunov function candidates fulfills these requirements by Lemmas 12 and 11 if

$$\begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}^\top \left(\begin{array}{cc|cc} -\rho^2\boldsymbol{P}_{11} & -\rho^2\boldsymbol{P}_{12} & 0 & 0 \\ -\rho^2\boldsymbol{P}_{21} & -\rho^2\boldsymbol{P}_{22} & 0 & 0 \\ \hline 0 & 0 & \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ 0 & 0 & \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{array}\right) \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}$$

$$+ \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}^\top \left(\begin{array}{cc|cc} 0 & 0 & 0 & -\frac{\lambda}{2}\boldsymbol{C}^\top \\ 0 & 0 & 0 & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger \\ \hline 0 & 0 & 0 & \frac{1}{2}\boldsymbol{C}^\top \\ -\frac{\lambda}{2}\boldsymbol{C} & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & \frac{1}{2}\boldsymbol{C} & -\widetilde{\boldsymbol{L}}^\dagger \end{array}\right) \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}$$

is smaller than zero for all $x \in \mathbb{R}^n$, $w = \nabla f(\boldsymbol{C}x)$, $w^+ = \nabla f(\boldsymbol{C}x^+)$ and $x^+ = \boldsymbol{A}x + \boldsymbol{B}w$. At this point we can even improve the estimate by the observation that due to Lemma 2

$$\begin{aligned} &0 = \boldsymbol{\Pi}_{\ker\widetilde{\boldsymbol{L}}}(\nabla f(\boldsymbol{C}x) - \nabla f(z_f^*)) = \boldsymbol{\Pi}_{\ker\widetilde{\boldsymbol{L}}} w, \\ &0 = \boldsymbol{\Pi}_{\ker\widetilde{\boldsymbol{L}}}(\nabla f(\boldsymbol{C}x^+) - \nabla f(z_f^*)) = \boldsymbol{\Pi}_{\ker\widetilde{\boldsymbol{L}}} w^+ \end{aligned}$$

hold true. This implies, that the term

$$\begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}^\top \left(\begin{array}{cc|cc} 0 & 0 & 0 & 0 \\ 0 & -r\boldsymbol{\Pi}_{\ker\widetilde{\boldsymbol{L}}} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -r\boldsymbol{\Pi}_{\widetilde{\boldsymbol{L}}} \end{array}\right) \begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix}$$

is zero for all $r \in \mathbb{R}$ and can hence be added to the estimate. Since the quantities x, w, x^+, w^+ are given by

$$\begin{pmatrix} x \\ w \\ \hline x^+ \\ w^+ \end{pmatrix} = \begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \hline \widetilde{\boldsymbol{A}} & \boldsymbol{B} & 0 \\ 0 & 0 & \boldsymbol{I}_d \end{pmatrix} \begin{pmatrix} x \\ w \\ w^+ \end{pmatrix}$$

negativity of $V_f(x^+) - \rho^2 V_f(x)$ follows now from (10). Hence, as a consequence of Theorem 10, the algorithm defined by $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ has convergence rate ρ for $S_0(0, \boldsymbol{L} - \boldsymbol{M})$. □

A.10. Proof of Theorem 14

We need to show that the matrix inequality (11) in the transformed variables $\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \boldsymbol{C}, \boldsymbol{P}$ is equivalent to (10). The proof of this theorem works in two steps. The first step is to apply the Schur complement to (10). The second (key) step is to define a linearizing change of variables.

Step 1. First, define $\boldsymbol{Z}$ as follows

$$\begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \hline \widetilde{\boldsymbol{A}} & \boldsymbol{B} & 0 \\ 0 & 0 & \boldsymbol{I}_d \end{pmatrix}^\top \left(\begin{array}{cc|cc} 0 & 0 & 0 & -\frac{\lambda}{2}\boldsymbol{C}^\top \\ 0 & -r\boldsymbol{\Pi} & 0 & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger \\ \hline 0 & 0 & 0 & \frac{1}{2}\boldsymbol{C}^\top \\ -\frac{\lambda}{2}\boldsymbol{C} & \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & \frac{1}{2}\boldsymbol{C} & -\widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} \end{array}\right)(\star)$$

$$= \begin{pmatrix} 0 & 0 & \frac{1}{2}\widetilde{\boldsymbol{A}}^\top\boldsymbol{C}^\top - \frac{\lambda}{2}\boldsymbol{C}^\top \\ 0 & -r\boldsymbol{\Pi} & \frac{1}{2}\boldsymbol{B}^\top\boldsymbol{C}^\top + \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger \\ \frac{1}{2}\boldsymbol{C}\widetilde{\boldsymbol{A}} - \frac{\lambda}{2}\boldsymbol{C} & \frac{1}{2}\boldsymbol{C}\boldsymbol{B} + \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & -\widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} \end{pmatrix} =: \boldsymbol{Z}.$$

With $\boldsymbol{Z}$, (10) becomes

$$\begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \hline \widetilde{\boldsymbol{A}} & \boldsymbol{B} & 0 \\ 0 & 0 & \boldsymbol{I}_d \end{pmatrix}^\top \left(\begin{array}{cc|cc} -\rho^2\boldsymbol{P}_{11} & -\rho^2\boldsymbol{P}_{12} & 0 & 0 \\ -\rho^2\boldsymbol{P}_{21} & -\rho^2\boldsymbol{P}_{22} & 0 & 0 \\ \hline 0 & 0 & \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ 0 & 0 & \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{array}\right)(\star) + \boldsymbol{Z}$$

$$= \begin{pmatrix} \boldsymbol{I}_n & 0 & 0 \\ 0 & \boldsymbol{I}_d & 0 \\ \boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} & \boldsymbol{P}_{11}\boldsymbol{B} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21}\widetilde{\boldsymbol{A}} & \boldsymbol{P}_{21}\boldsymbol{B} & \boldsymbol{P}_{22} \end{pmatrix}^\top \begin{pmatrix} -\rho^2\boldsymbol{P} & 0 \\ 0 & \boldsymbol{P}^{-1} \end{pmatrix}(\star) + \boldsymbol{Z} \prec 0.$$

The matrix $\boldsymbol{P}$ is positive definite by assumption of Theorem 13 and as a consequence of the matrix inequality from Theorem 14. Hence, this algebraic manipulation allows to apply the Schur complement, which states that the above inequality is equivalent to

$$\begin{pmatrix} -\rho^2 \boldsymbol{P}_{11} & -\rho^2 \boldsymbol{P}_{12} & * & * & * \\ -\rho^2 \boldsymbol{P}_{21} & -\rho^2 \boldsymbol{P}_{22} - r\boldsymbol{\Pi} & * & * & * \\ \frac{1}{2}\boldsymbol{C}\widetilde{\boldsymbol{A}} - \frac{\lambda}{2}\boldsymbol{C} & \frac{1}{2}\boldsymbol{C}\boldsymbol{B} + \frac{\lambda}{2}\widetilde{\boldsymbol{L}}^\dagger & -\widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} & * & * \\ \boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} & \boldsymbol{P}_{11}\boldsymbol{B} & \boldsymbol{P}_{12} & -\boldsymbol{P}_{11} & -\boldsymbol{P}_{12} \\ \boldsymbol{P}_{21}\widetilde{\boldsymbol{A}} & \boldsymbol{P}_{21}\boldsymbol{B} & \boldsymbol{P}_{22} & -\boldsymbol{P}_{21} & -\boldsymbol{P}_{22} \end{pmatrix}$$

being negative definite.

Step 2. If we have a solution $(\hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \dots)$ of (11) and constraint (12), then we can just substitute $\widetilde{\boldsymbol{A}} = \boldsymbol{P}_{11}^{-1}\hat{\boldsymbol{A}}, \boldsymbol{B} = \boldsymbol{P}_{11}^{-1}\hat{\boldsymbol{B}}$ into (11) and we see that we obtain the above inequality and hence a solution of (10). This solution also satisfies constraint (5) since

$$\begin{aligned} \boldsymbol{C}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M} &= \boldsymbol{C}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{P}_{11}^{-1}\hat{\boldsymbol{B}}\boldsymbol{M} \\ &= \boldsymbol{C}(\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \boldsymbol{P}_{11})^{-1}\hat{\boldsymbol{B}}\boldsymbol{M} \\ &= \boldsymbol{C}(\hat{\boldsymbol{A}} - \boldsymbol{P}_{11})^{-1}\hat{\boldsymbol{B}}\boldsymbol{M} \\ &= \boldsymbol{C}\boldsymbol{J}_1^T = \boldsymbol{I}_d. \end{aligned}$$

On the other hand, if we are given a solution of (10), (5) with $\boldsymbol{P} \succ 0$ and we want to construct a solution of (11) by substituting $\hat{\boldsymbol{B}} = \boldsymbol{P}_{11}\boldsymbol{B}$, $\hat{\boldsymbol{A}} = \boldsymbol{P}_{11}\widetilde{\boldsymbol{A}}$ and by expressing all the nonlinear expressions $\boldsymbol{C}\widetilde{\boldsymbol{A}}, \boldsymbol{P}_{21}\widetilde{\boldsymbol{A}}, \boldsymbol{P}_{11}\widetilde{\boldsymbol{A}}, \boldsymbol{C}\boldsymbol{B}, \boldsymbol{P}_{21}\boldsymbol{B}, \boldsymbol{P}_{11}\boldsymbol{B}$ in terms of $\hat{\boldsymbol{A}}$ and $\hat{\boldsymbol{B}}$, we cannot guarantee that (12) holds. However, in the following we show that there exists a state transformation of the algorithm such that this can be indeed guaranteed. Hence, any solution of (10), (5) is a solution of (11), (12) by an appropriate coordinate transformation.

If there exists a (non-singular) transformation matrix $\boldsymbol{T}$ such that the transformed variables $\widetilde{\boldsymbol{A}}' = \boldsymbol{T}^{-1}\widetilde{\boldsymbol{A}}\boldsymbol{T}, \boldsymbol{B}' = \boldsymbol{T}^{-1}\boldsymbol{B}, \boldsymbol{C}' = \boldsymbol{C}\boldsymbol{T}$, $\boldsymbol{P}'_{11} = \boldsymbol{T}^\top\boldsymbol{P}_{11}\boldsymbol{T}, \boldsymbol{P}'_{12} = \boldsymbol{T}^\top\boldsymbol{P}_{12}, \boldsymbol{P}'_{21} = \boldsymbol{P}_{21}\boldsymbol{T}, \boldsymbol{P}'_{22} = \boldsymbol{P}_{22}$ fulfill

$$(\widetilde{\boldsymbol{A}}' - \boldsymbol{I}_n)\boldsymbol{J}_1^\top = \boldsymbol{B}'\boldsymbol{M}, \quad \boldsymbol{J}_2\boldsymbol{P}'_{11} = \boldsymbol{C}', \quad \boldsymbol{J}_3\boldsymbol{P}'_{11} = \boldsymbol{P}'_{21},$$

then we have

$$\begin{pmatrix} \boldsymbol{C}'\widetilde{\boldsymbol{A}}' \\ \boldsymbol{P}'_{21}\widetilde{\boldsymbol{A}}' \\ \boldsymbol{P}'_{11}\widetilde{\boldsymbol{A}}' \end{pmatrix} = \begin{pmatrix} \boldsymbol{J}_2 \\ \boldsymbol{J}_3 \\ \boldsymbol{I}_n \end{pmatrix} \hat{\boldsymbol{A}}', \qquad \begin{pmatrix} \boldsymbol{C}'\boldsymbol{B}' \\ \boldsymbol{P}'_{21}\boldsymbol{B}' \\ \boldsymbol{P}'_{11}\boldsymbol{B}' \end{pmatrix} = \begin{pmatrix} \boldsymbol{J}_2 \\ \boldsymbol{J}_3 \\ \boldsymbol{I}_n \end{pmatrix} \hat{\boldsymbol{B}}'$$

and the transformed variables still form a solution of inequality (10). The arguments from Step 1 show that in this case $\widetilde{\boldsymbol{A}}', \boldsymbol{B}', \boldsymbol{C}'$ and $\boldsymbol{P}'$ form also a solution of (11) and by substituting $\widehat{\boldsymbol{A}}'$ and $\widehat{\boldsymbol{B}}'$ for the nonlinear terms it becomes clear that there exists a solution to (11), (12) from Theorem 14.

Such a transformation $\boldsymbol{T}$ must now fulfill the constraints

$$\boldsymbol{J}_2\underbrace{\boldsymbol{T}^\top\boldsymbol{P}_{11}\boldsymbol{T}}_{=\boldsymbol{P}'_{11}} = \underbrace{\boldsymbol{C}\boldsymbol{T}}_{=\boldsymbol{C}'}, \qquad \boldsymbol{J}_3\boldsymbol{T}^\top\boldsymbol{P}_{11}\boldsymbol{T} = \underbrace{\boldsymbol{P}_{21}\boldsymbol{T}}_{=\boldsymbol{P}'_{21}},$$

$$\underbrace{\boldsymbol{T}^{-1}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)\boldsymbol{T}}_{=\widetilde{\boldsymbol{A}}' - \boldsymbol{I}_n}\boldsymbol{J}_1^\top = \underbrace{\boldsymbol{T}^{-1}\boldsymbol{B}}_{=\boldsymbol{B}'}\boldsymbol{M}.$$

Rearranging and canceling terms above gives

$$\boldsymbol{T}\boldsymbol{J}_1^\top = (\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M}, \quad \boldsymbol{J}_2\boldsymbol{T}^\top = \boldsymbol{C}\boldsymbol{P}_{11}^{-1}, \quad \boldsymbol{J}_3\boldsymbol{T}^\top = \boldsymbol{P}_{21}\boldsymbol{P}_{11}^{-1}.$$

For the choice $\boldsymbol{J}_1 = (\boldsymbol{I}_d\ 0_d\ 0_d\ 0), \boldsymbol{J}_2 = (0_d\ \boldsymbol{I}_d\ 0_d\ 0), \boldsymbol{J}_3 = (0_d\ 0_d\ \boldsymbol{I}_d\ 0)$, these equations have the solution

$$\boldsymbol{T} = \left((\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M} \quad \boldsymbol{P}_{11}^{-T}\boldsymbol{C}^T \quad \boldsymbol{P}_{11}^{-T}\boldsymbol{P}_{21}^\top \quad \boldsymbol{T}_4\right),$$

provided, that $n \geq 3d$. It remains to show that the transformation is non-singular. Notice that $(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M}$ and $\boldsymbol{C}$ must have full rank because of $\boldsymbol{C}(\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M} = \boldsymbol{I}_d$. Moreover $\boldsymbol{P}_{11}, \boldsymbol{P}_{21}$ can be slightly perturbed (without violating the strict definiteness of $\boldsymbol{P}$ and the matrix inequality (10)), such that $\left((\widetilde{\boldsymbol{A}} - \boldsymbol{I}_n)^{-1}\boldsymbol{B}\boldsymbol{M} \ \ \boldsymbol{P}_{11}^{-T}\boldsymbol{C}^T \ \ \boldsymbol{P}_{11}^{-T}\boldsymbol{P}_{21}^\top\right)$ has full rank too. Finally, $\boldsymbol{T}_4 \in \mathbb{R}^{n \times n-3d}$ can be chosen such that $\boldsymbol{T}$ is non-singular. Hence, all constraints of Theorem 14 are satisfied by construction of $\boldsymbol{T}$, where $\boldsymbol{C}\boldsymbol{J}_1 = \boldsymbol{I}_d$ is implied by (5). Consequently, it is possible to construct solutions related to Theorem 14 from solutions related to Theorem 13 and vice versa. □

A.11. Proof of Theorem 15

Again, we introduce the abbreviations $\widetilde{\boldsymbol{L}} = \boldsymbol{L} - \boldsymbol{M}$ and $\boldsymbol{\Pi} = \boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}}$. In this proof, it will be necessary to find explicit solutions for the matrix inequality (10) from Theorem 13. Therefore, it is purposeful to multiply out the matrix products in this inequality for $\lambda = 0$, resulting in:

$$\begin{pmatrix} \widetilde{\boldsymbol{A}}^\top\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \rho^2\boldsymbol{P}_{11} & \widetilde{\boldsymbol{A}}^\top\boldsymbol{P}_{11}\boldsymbol{B} - \rho^2\boldsymbol{P}_{12} & \widetilde{\boldsymbol{A}}^\top(\boldsymbol{P}_{12} + \frac{1}{2}\boldsymbol{C}^\top) \\ \boldsymbol{B}\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \rho^2\boldsymbol{P}_{21} & \boldsymbol{B}^\top\boldsymbol{P}_{11}\boldsymbol{B} - \rho^2\boldsymbol{P}_{22} - r\boldsymbol{\Pi} & \boldsymbol{B}^\top(\boldsymbol{P}_{12} + \frac{1}{2}\boldsymbol{C}^\top) \\ (\boldsymbol{P}_{21} + \frac{1}{2}\boldsymbol{C})\widetilde{\boldsymbol{A}} & (\boldsymbol{P}_{21} + \frac{1}{2}\boldsymbol{C})\boldsymbol{B} & \boldsymbol{P}_{22} - \widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi} \end{pmatrix}.$$

(1) : This step will be quite lengthy. We will show, that the matrices $(\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C})$ given by

$$\begin{aligned} \widetilde{\boldsymbol{A}} &= \boldsymbol{A} + \boldsymbol{B}\boldsymbol{M}\boldsymbol{C} = \boldsymbol{I}_d - 2(\boldsymbol{L} + \boldsymbol{M})^{-1}\boldsymbol{M} \\ &= (\boldsymbol{L} + \boldsymbol{M})^{-1}(\boldsymbol{L} - \boldsymbol{M}) \\ \boldsymbol{B} &= -2(\boldsymbol{L} + \boldsymbol{M})^{-1} \\ \boldsymbol{C} &= \boldsymbol{I}_d \end{aligned}$$

fulfill all the convergence rate conditions of Theorem 13 for an arbitrary given $\rho \in]\rho_{\text{grad}}, 1[$. Here, the matrix $\widetilde{\boldsymbol{A}}$ fulfills the Lyapunov inequality $\widetilde{\boldsymbol{A}}^\top\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{A}} - \rho^2\widetilde{\boldsymbol{P}} \prec 0$ for $\widetilde{\boldsymbol{P}} := (\boldsymbol{L}+\boldsymbol{M})\left((\boldsymbol{L} - \boldsymbol{M})^\dagger + r\boldsymbol{\Pi}\right)(\boldsymbol{L} + \boldsymbol{M})$ and large enough $r \in \mathbb{R}_{>0}$ by Lemma 18, since $\widetilde{\boldsymbol{A}}^\top\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{A}} = \boldsymbol{L} - \boldsymbol{M}$. To show, that the convergence conditions from Theorem 13 are met we choose $\lambda = 0$ and the following value for $\boldsymbol{P}$:

$$\begin{pmatrix} \boldsymbol{P}_{11} & \boldsymbol{P}_{12} \\ \boldsymbol{P}_{21} & \boldsymbol{P}_{22} \end{pmatrix} = \begin{pmatrix} \frac{\rho^2}{4}\left(\widetilde{\boldsymbol{P}} - \varepsilon(\boldsymbol{L} + \boldsymbol{M})^2\right) & -\frac{1}{2}\boldsymbol{I}_d \\ -\frac{1}{2}\boldsymbol{I}_d & \widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi} - \frac{\varepsilon}{2}\boldsymbol{I}_d \end{pmatrix},$$

where $\varepsilon > 0$, and $r \in \mathbb{R}$ is the same as above. There are three things to show:

(1) The constraint (5) of Theorem 13 is satisfied for $\widetilde{\boldsymbol{A}}, \boldsymbol{B}, \boldsymbol{C}$.
(2) For large enough r and small enough ε, $\boldsymbol{P}$ solves the matrix inequality (10) of Theorem 13.
(3) For large enough r and small enough ε, $\boldsymbol{P}$ is positive definite.

Verifying (1) can be done by a simple calculation of formulas in the constraint.

We will now show (2). Note that $(\boldsymbol{P}_{21} + \frac{1}{2}\boldsymbol{C}) = \frac{1}{2}(\boldsymbol{I}_d - \boldsymbol{I}_d) = 0$ holds, which is why (10) from Theorem 13 simplifies to

$$\begin{pmatrix} \widetilde{\boldsymbol{A}}^\top\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \rho^2\boldsymbol{P}_{11} & \widetilde{\boldsymbol{A}}^\top\boldsymbol{P}_{11}\boldsymbol{B} - \rho^2\boldsymbol{P}_{12} & 0 \\ \boldsymbol{B}\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \rho^2\boldsymbol{P}_{21} & \boldsymbol{B}^\top\boldsymbol{P}_{11}\boldsymbol{B} - \rho^2\boldsymbol{P}_{22} - r\boldsymbol{\Pi} & 0 \\ 0 & 0 & -\frac{\varepsilon}{2}\boldsymbol{I}_d \end{pmatrix} \prec 0.$$

Here it is left to show, that the left upper 2×2 block can be made negative definite by choosing r big and ε small. This is done by dividing the matrix inequality by $\frac{\rho^2}{4}$ and calculating the entries of the left upper blocks:

The first block is

$$\begin{aligned} &\frac{4}{\rho^2}\left(\widetilde{\boldsymbol{A}}^\top\boldsymbol{P}_{11}\widetilde{\boldsymbol{A}} - \rho^2\boldsymbol{P}_{11}\right) \\ &= \widetilde{\boldsymbol{A}}^\top\left(\widetilde{\boldsymbol{P}} - \varepsilon(\boldsymbol{L} + \boldsymbol{M})^2\right)\widetilde{\boldsymbol{A}} - \rho^2\left(\widetilde{\boldsymbol{P}} - \varepsilon(\boldsymbol{L} + \boldsymbol{M})^2\right) \\ &= \widetilde{\boldsymbol{A}}^\top\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{A}} - \rho^2\widetilde{\boldsymbol{P}} - \varepsilon\left((\boldsymbol{L} - \boldsymbol{M})^2 - \rho^2(\boldsymbol{L} + \boldsymbol{M})^2\right). \end{aligned}$$

The second block is

$$\begin{aligned}&\frac{4}{\rho^2}\left(\widetilde{\boldsymbol{A}}^\top \boldsymbol{P}_{11}\boldsymbol{B} - \rho^2 \boldsymbol{P}_{12}\right)\\ &= 2\boldsymbol{I}_d + \widetilde{\boldsymbol{A}}^\top\left(\widetilde{\boldsymbol{P}} - \varepsilon(\boldsymbol{L}+\boldsymbol{M})^2\right)\boldsymbol{B}\\ &= 2\boldsymbol{I}_d - 2(\boldsymbol{L}-\boldsymbol{M})\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}\right) - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M})\\ &= 2\boldsymbol{I}_d - 2\underbrace{(\boldsymbol{L}-\boldsymbol{M})(\boldsymbol{L}-\boldsymbol{M})^\dagger}_{\overset{(16)}{=}\boldsymbol{\Pi}_{\mathrm{im}(\boldsymbol{L}-\boldsymbol{M})}} - 2r\underbrace{(\boldsymbol{L}-\boldsymbol{M})\boldsymbol{\Pi}}_{=0} - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M})\\ &= 2\boldsymbol{I}_d - 2\boldsymbol{\Pi}_{\mathrm{im}(\boldsymbol{L}-\boldsymbol{M})} - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M})\\ &\overset{(17)}{=} 2\boldsymbol{\Pi}_{\ker(\boldsymbol{L}-\boldsymbol{M})} - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M}).\end{aligned}$$

The third block is:

$$\begin{aligned}&\frac{4}{\rho^2}\left(\boldsymbol{B}^\top \boldsymbol{P}_{11}\boldsymbol{B} - \rho^2 \boldsymbol{P}_{22} - r\boldsymbol{\Pi}\right)\\ &= \boldsymbol{B}^\top\widetilde{\boldsymbol{P}}\boldsymbol{B} - 4\varepsilon\boldsymbol{I}_d - 4\left(\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi} - \frac{\varepsilon}{2}\boldsymbol{I}_d\right) - \frac{4}{\rho^2}r\boldsymbol{\Pi}\\ &= 4(\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi}) - 4\varepsilon\boldsymbol{I}_d - 4\left(\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi} - \frac{\varepsilon}{2}\boldsymbol{I}_d\right) - \frac{4}{\rho^2}r\boldsymbol{\Pi}\\ &= 4\left(\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi} - \widetilde{\boldsymbol{L}}^\dagger - r\boldsymbol{\Pi}\right) - 2\varepsilon\boldsymbol{I}_d - \frac{4}{\rho^2}r\boldsymbol{\Pi}.\end{aligned}$$

By the calculation of these blocks the upper 2×2 block is

$$\begin{pmatrix}\widetilde{\boldsymbol{A}}^\top\widetilde{\boldsymbol{P}}\widetilde{\boldsymbol{A}} - \rho^2\widetilde{\boldsymbol{P}} - \varepsilon\left((\boldsymbol{L}-\boldsymbol{M})^2 - \rho^2(\boldsymbol{L}+\boldsymbol{M})^2\right) & 2\boldsymbol{\Pi} - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M})\\ 2\boldsymbol{\Pi} - 2\varepsilon(\boldsymbol{L}-\boldsymbol{M}) & -2\varepsilon\boldsymbol{I}_d - \frac{4}{\rho^2}r\boldsymbol{\Pi}\end{pmatrix},$$

which is negative definite for $\varepsilon > 0$ small enough and r big enough.

Now it is left to show (3), namely that $\boldsymbol{P}$ is positive definite for small enough ε and large enough r. Therefore, we can show that $\boldsymbol{P}$ is positive definite for $\varepsilon = 0$. Then it will also be positive definite for the small perturbation with $\varepsilon > 0$. By the Schur complement, the matrix $\boldsymbol{P}$ for $\varepsilon = 0$ is positive definite if and only if:

$$\begin{aligned}&0 \prec \widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi}\\ &0 \prec \frac{\rho^2}{4}\widetilde{\boldsymbol{P}} - \left(-\frac{1}{2}\boldsymbol{I}_d\right)\underbrace{(\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi})^{-1}}_{\overset{(19)}{=}(\boldsymbol{L}-\boldsymbol{M})+\frac{1}{r}\boldsymbol{\Pi}}\left(-\frac{1}{2}\boldsymbol{I}_d\right)\\ &\quad = \frac{\rho^2}{4}\widetilde{\boldsymbol{P}} - \frac{1}{4}(\boldsymbol{L}-\boldsymbol{M}) - \frac{1}{4r}\boldsymbol{\Pi}.\end{aligned}$$

Since $\rho > \rho_{\mathrm{grad}}$, the matrix $\rho^2\widetilde{\boldsymbol{P}} - (\boldsymbol{L}-\boldsymbol{M})$ is positive definite by Lemma 18 and the matrix $\widetilde{\boldsymbol{L}}^\dagger + r\boldsymbol{\Pi}$ is positive definite by construction. Thus, $\frac{\rho^2}{4}\widetilde{\boldsymbol{P}} - \frac{1}{4}(\boldsymbol{L}-\boldsymbol{M}) - \frac{1}{4r}\boldsymbol{\Pi}$ is positive definite for large values of r. Hence, we only have to make ε small enough and r big enough, such that $\boldsymbol{P}$ becomes positive definite.

(2) : From (1) it is clear that we have a special solution to the conditions in Theorem 13 for $n = d$ and a convergence rate $\rho \in [\rho_{\mathrm{grad}}, 1[$. Let $\widetilde{\boldsymbol{A}}^{(d)}, \boldsymbol{B}^{(d)}, \boldsymbol{C}^{(d)}, \boldsymbol{P}^{(d)}$ be this special solution. This solution can be extended to a solution for arbitrary dimension $n \geq d$ by setting

$$\begin{aligned}&\widetilde{\boldsymbol{A}} = \begin{pmatrix}\widetilde{\boldsymbol{A}}^{(d)} & 0_{d\times n-d}\\ 0_{n-d\times d} & 0_{n-d\times n-d}\end{pmatrix}, \boldsymbol{B} = \begin{pmatrix}\boldsymbol{B}^{(d)}\\ 0_{n-d\times d}\end{pmatrix},\\ &\boldsymbol{C} = \begin{pmatrix}\boldsymbol{C}^{(d)} & 0_{d\times n-d\cdot}\end{pmatrix}, \boldsymbol{P}_{22} = \boldsymbol{P}_{22}^{(d)},\\ &\boldsymbol{P}_{11} = \begin{pmatrix}\boldsymbol{P}_{11}^{(d)} & 0_{d\times n-d}\\ 0_{n-d\times d} & \boldsymbol{I}_{n-d}\end{pmatrix}\boldsymbol{P}_{12} = \begin{pmatrix}\boldsymbol{P}_{12}^{(d)}\\ 0_{n-d\times d}\end{pmatrix}.\end{aligned}$$

Showing that these values satisfy the constraints and the LMI of Theorem 13 is straight forward. Therefore, we get a solution to the conditions in Theorem 13 for any $n \geq d$ for the convergence rate $\rho \in [\rho_{\mathrm{grad}}, 1[$. As stated in Theorem 14, the constraints and matrix inequality of this theorem are equivalent to the conditions of Theorem 13 in the case $n \geq 3d$. Hence, there exists also exists a solution to the conditions in Theorem 14 if $n \geq 3d$ for the considered convergence rate ρ. □

A.12. Proof of Lemma 16

We have to check whether inequality (2) holds for the Lagrangian function $L \in C^1$. Let therefore arbitrary values $z_1, z_2 \in \mathbb{R}^d$ and $\lambda_1, \lambda_2 \in \mathbb{R}^{d_2}$ be given. The lower bound in inequality (2) follows from

$$\begin{aligned}&\frac{1}{2}\begin{pmatrix}z_1 - z_2\\ \lambda_1 - \lambda_2\end{pmatrix}^\top\begin{pmatrix}\boldsymbol{M} & \boldsymbol{A}_{\mathrm{eq}}^\top\\ \boldsymbol{A}_{\mathrm{eq}} & 0\end{pmatrix}\begin{pmatrix}z_1 - z_2\\ \lambda_1 - \lambda_2\end{pmatrix}\\ &= \frac{1}{2}(z_1 - z_2)^\top\boldsymbol{M}(z_1 - z_2) + (\lambda_1 - \lambda_2)^\top\boldsymbol{A}_{\mathrm{eq}}(z_1 - z_2)\\ &\leq f(z_2) - f(z_1) + (\nabla f(z_1))^\top(z_1 - z_2)\\ &\quad + (\lambda_1 - \lambda_2)^\top\boldsymbol{A}_{\mathrm{eq}}(z_1 - z_2)\\ &= \underbrace{f(z_2) + \lambda_2^\top(\boldsymbol{A}_{\mathrm{eq}}z_2 - b_{\mathrm{eq}})}_{L(z_2,\lambda_2)} - \underbrace{(f(z_1) + \lambda_1^\top(\boldsymbol{A}_{\mathrm{eq}}z_1 - b_{\mathrm{eq}}))}_{L(z_1,\lambda_1)}\\ &+ \underbrace{(\nabla f(z_1) + \boldsymbol{A}_{\mathrm{eq}}^\top\lambda_1)^\top(z_1 - z_2) + (\lambda_1 - \lambda_2)^\top(\boldsymbol{A}_{\mathrm{eq}}z_1 - b_{\mathrm{eq}})}_{(\nabla L(z_1,\lambda_1))^\top\begin{pmatrix}z_1 - z_2\\ \lambda_1 - \lambda_2\end{pmatrix}}.\end{aligned}$$

The upper bound can be shown analogously. □

A.13. Auxiliary results

Lemma 18. *Let $\boldsymbol{M} \preceq_c \boldsymbol{L}$ be non-singular, symmetric matrices. Then for any $\rho > \rho\left((\boldsymbol{M}+\boldsymbol{L})^{-1}(\boldsymbol{L}-\boldsymbol{M})\right)$ there exists an $r_0 \in \mathbb{R}_{>0}$ such that for all real numbers $r \geq r_0$*

$$\boldsymbol{L} - \boldsymbol{M} \prec \rho^2(\boldsymbol{L}+\boldsymbol{M})\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}}\right)(\boldsymbol{L}+\boldsymbol{M}).$$

Proof. Let $\rho > \rho\left((\boldsymbol{M}+\boldsymbol{L})^{-1}(\boldsymbol{L}-\boldsymbol{M})\right)$ be given. Define $\boldsymbol{\Pi} := \boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}}$ and

$$\begin{aligned}\tilde{\rho} &:= \rho\left((\boldsymbol{L}+\boldsymbol{M})^{-1}\left(\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}\right)\right)\\ &\overset{(\star)}{=} \rho\left(\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}(\boldsymbol{L}+\boldsymbol{M})^{-1}\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}\right)\\ &= \left\|\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}(\boldsymbol{L}+\boldsymbol{M})^{-1}\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}\right\|.\end{aligned}$$

Here, the equality $(\star)$ holds by a similarity transform with $\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}$. This definition of $\tilde{\rho}$ implies the matrix inequality

$$\tilde{\rho}^2\boldsymbol{I}_d \succeq \left(\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}(\boldsymbol{L}+\boldsymbol{M})^{-1}\sqrt{\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}}\right)^2.$$

A congruence transform with $\left(\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}\right)^{-\frac{1}{2}}(\boldsymbol{L}+\boldsymbol{M})$ yields

$$\tilde{\rho}^2(\boldsymbol{L}+\boldsymbol{M})\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}\right)(\boldsymbol{L}+\boldsymbol{M}) \succeq \boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi},$$

since $\left(\boldsymbol{L}-\boldsymbol{M}+\frac{1}{r}\boldsymbol{\Pi}\right)^{-\frac{1}{2}} \overset{(19)}{=} \sqrt{(\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}}$. By the expression of $\tilde{\rho}$ through the spectral norm and the continuity of the norm, $\tilde{\rho}$ converges to $\rho\left((\boldsymbol{M}+\boldsymbol{L})^{-1}(\boldsymbol{L}-\boldsymbol{M})\right)$ for $r \to \infty$. Hence, we can choose r large enough, such that $\tilde{\rho}$ is small than ρ and thus,

$$\begin{aligned}\boldsymbol{L}-\boldsymbol{M} &\preceq \tilde{\rho}^2(\boldsymbol{L}+\boldsymbol{M})\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}}\right)(\boldsymbol{L}+\boldsymbol{M})\\ &\prec \rho^2(\boldsymbol{L}+\boldsymbol{M})\left((\boldsymbol{L}-\boldsymbol{M})^\dagger + r\boldsymbol{\Pi}_{\ker \boldsymbol{L}-\boldsymbol{M}}\right)(\boldsymbol{L}+\boldsymbol{M}).\end{aligned}$$

Since increasing r corresponds to adding a positive definite term to the right hand side of this inequality, the inequality remains valid for larger values of r. □

Lemma 19 (*Congruence Lemma*)**.** *Let $\boldsymbol{M}, \boldsymbol{L} \in \mathbb{R}^{d\times d}$ be two symmetric matrices such that there exists a positive definite matrix $\boldsymbol{P} = \boldsymbol{P}^\top \in \mathbb{R}^{d\times d}$ with $\boldsymbol{MPL} + \boldsymbol{LPM} \succ 0$. Then $\boldsymbol{M}$ and $\boldsymbol{L}$ are congruent, i.e. there exists a non-singular matrix $\boldsymbol{T}$ such that $\boldsymbol{T}^\top \boldsymbol{MT} = \boldsymbol{L}$.*

Proof. By $\boldsymbol{P}$ being positive definite, there exists a symmetric positive definite matrix $\sqrt{\boldsymbol{P}} \in \mathbb{R}^{d\times d}$ with $\sqrt{\boldsymbol{P}}^2 = \boldsymbol{P}$. A congruence transform with $\sqrt{\boldsymbol{P}}$ yields

$$\sqrt{\boldsymbol{P}}\boldsymbol{M}\sqrt{\boldsymbol{P}}\sqrt{\boldsymbol{P}}\boldsymbol{L}\sqrt{\boldsymbol{P}} + \sqrt{\boldsymbol{P}}\boldsymbol{L}\sqrt{\boldsymbol{P}}\sqrt{\boldsymbol{P}}\boldsymbol{M}\sqrt{\boldsymbol{P}} \succ 0. \tag{27}$$

The matrices $\widetilde{\boldsymbol{M}} := \sqrt{\boldsymbol{P}}\boldsymbol{M}\sqrt{\boldsymbol{P}}$ and $\widetilde{\boldsymbol{L}} := \sqrt{\boldsymbol{P}}\boldsymbol{L}\sqrt{\boldsymbol{P}}$ are congruent to $\boldsymbol{M}$ and $\boldsymbol{L}$. Hence, it is sufficient to show that the matrices $\widetilde{\boldsymbol{M}}$ and $\widetilde{\boldsymbol{L}}$ are congruent.

Therefore, let $\boldsymbol{T}$ be an orthogonal matrix, such that

$$\boldsymbol{T}^\top \widetilde{\boldsymbol{M}}\boldsymbol{T} = \begin{pmatrix} \boldsymbol{D}_1 & 0 \\ 0 & \boldsymbol{D}_2 \end{pmatrix},$$

where $\boldsymbol{D}_1$ is the diagonal matrix of all positive eigenvalues of $\widetilde{\boldsymbol{M}}$ and $\boldsymbol{D}_2$ is the matrix of all negative eigenvalues of $\widetilde{\boldsymbol{M}}$. Now, a congruence transform with $\boldsymbol{T}$ can be applied to (27):

$$\begin{aligned} 0 &\prec \boldsymbol{T}^\top \widetilde{\boldsymbol{L}}\widetilde{\boldsymbol{M}}\boldsymbol{T} + \boldsymbol{T}^\top \widetilde{\boldsymbol{M}}\widetilde{\boldsymbol{L}}\boldsymbol{T} \\ &= \underbrace{\boldsymbol{T}^\top \widetilde{\boldsymbol{L}}\boldsymbol{T}}_{:=\boldsymbol{E}^\top}\boldsymbol{T}^\top \widetilde{\boldsymbol{M}}\boldsymbol{T} + \boldsymbol{T}^\top \widetilde{\boldsymbol{M}}\boldsymbol{T}\underbrace{\boldsymbol{T}^\top \widetilde{\boldsymbol{L}}\boldsymbol{T}}_{:=\boldsymbol{E}} \\ &= \begin{pmatrix} \boldsymbol{E}_{11} & \boldsymbol{E}_{12} \\ \boldsymbol{E}_{21} & \boldsymbol{E}_{22} \end{pmatrix}\begin{pmatrix} \boldsymbol{D}_1 & 0 \\ 0 & \boldsymbol{D}_2 \end{pmatrix} + \begin{pmatrix} \boldsymbol{D}_1 & 0 \\ 0 & \boldsymbol{D}_2 \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_{11} & \boldsymbol{E}_{12} \\ \boldsymbol{E}_{21} & \boldsymbol{E}_{22} \end{pmatrix}. \end{aligned}$$

From this inequality, one can read off

$$\boldsymbol{E}_{11}\boldsymbol{D}_1 + \boldsymbol{D}_1\boldsymbol{E}_{11} \succ 0, \qquad \boldsymbol{E}_{22}\boldsymbol{D}_2 + \boldsymbol{D}_2\boldsymbol{E}_{22} \succ 0$$

from the diagonal blocks. Hence, by the Lyapunov inequality, $\boldsymbol{E}_{11} \succ 0$ and $\boldsymbol{E}_{22} \prec 0$. Now, $\boldsymbol{E}$ is positive definite on the subspace corresponding to $\boldsymbol{E}_{11}$ and negative definite on the subspace corresponding to $\boldsymbol{E}_{22}$. Consequently, $\boldsymbol{E}$ has exactly $\dim \boldsymbol{E}_{11} = \dim \boldsymbol{D}_1$ positive and exactly $\dim \boldsymbol{E}_{22} = \dim \boldsymbol{D}_2$ negative eigenvalues according to Sylvester's law of inertia. Thus the matrices $\boldsymbol{M}$ and $\boldsymbol{L}$, which are congruent to $\boldsymbol{D}$ and $\boldsymbol{E}$, are congruent to each other. □

A.14. Efficiently solving (11)

Let $\boldsymbol{F}(\boldsymbol{P}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}, r, \lambda, \rho^2) \prec 0$ denote (11). To find the fastest algorithm which can be synthesized using Theorem 14, we solve the optimization problem

$$\begin{aligned} &\text{minimize: } \rho^2 \\ &\text{subject to } \boldsymbol{F}(\boldsymbol{P}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}, r, \lambda, \rho^2) \prec 0, \\ &\qquad (12) \text{ and } \lambda \in [0, \rho^2]. \end{aligned} \tag{28}$$

This problem is not convex, since $\boldsymbol{F}(\boldsymbol{P}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}, r, \lambda, \rho^2) \prec 0$ may be convex in $\boldsymbol{P}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}$ and r, but not simultaneously in ρ^2 and λ. Therefore, we suggest the following procedure:

Line search for λ: Firstly, observe that a value for ρ^2 is feasible, whenever there exists a $\lambda \in [0, \rho^2]$ such that

$$J(\rho^2, \lambda) = \inf_{\substack{\boldsymbol{F}(\boldsymbol{P}, \hat{\boldsymbol{A}}, \hat{\boldsymbol{B}}, \hat{\boldsymbol{C}}, r, \lambda, \rho^2) \prec \gamma \boldsymbol{I}, \\ (12)}} \gamma < 0 \tag{29}$$

holds, where $J(\rho^2, \lambda)$ can be computed by solving a semi-definite program. Hence, we suggest the following algorithm:

Algorithm 1 Algorithm Synthesis

$\rho_{\text{upper}} \leftarrow 1$, $\rho_{\text{lower}} \leftarrow 0$, $k \leftarrow 0$
while $k \le k_{\max}$ **do**
  $\rho \leftarrow \frac{\rho_{\text{upper}} + \rho_{\text{lower}}}{2}$
  Compute $\gamma = \inf_{\lambda \in [0,\rho^2]} J(\rho^2, \lambda)$ by line search.
  if $\gamma < 0$ **then**
    $\rho_{\text{upper}} \leftarrow \rho$
  else
    $\rho_{\text{lower}} \leftarrow \rho$
  end if
  $k \leftarrow k + 1$
end while

This algorithm consists of an inner loop that minimizes $J(\rho^2, \lambda)$ over λ using a line search and thereby checks whether the current value of ρ^2 is feasible, and an outer loop that optimizes over ρ^2 using a bisection search. Here, the use of a bisection search for ρ^2 is justified, since investigating (11) yields that the feasible values for ρ^2 are going to be an interval (if there are any). As line search algorithm, we employed grid-search and golden sectioning. We propose the use of golden sectioning, since the function $\lambda \mapsto J(\rho^2, \lambda)$ appeared to be convex in all experiments that we made and therefore golden sectioning should be faster. Nevertheless, we always used a grid-search to backup the golden sectioning.

References

[1] Anatoliy I. Lur'e, Vladimir N. Postnikov, On the theory of stability of control systems, Appl. Math. Mech. 8 (3) (1944) 246–248.

[2] Laurent Lessard, Benjamin Recht, Andrew Packard, Analysis and design of optimization algorithms via integral quadratic constraints, SIAM J. Optim. 26 (1) (2016) 57–95.

[3] Laurent Lessard, Peter Seiler, Direct synthesis of iterative algorithms with bounds on achievable worst-case convergence rate, 2019, arXiv preprint arXiv:1904.09046.

[4] Bryan Van Scoy, Randy A. Freeman, Kevin M. Lynch, The fastest known globally convergent first-order method for minimizing strongly convex functions, IEEE Control Syst. Lett. 2 (1) (2017) 49–54.

[5] Mahyar Fazlyab, Alejandro Ribeiro, Manfred Morari, Victor M. Preciado, Analysis of optimization algorithms via integral quadratic constraints: Nonstrongly convex problems, SIAM J. Optim. 28 (3) (2018) 2654–2689.

[6] Simon Michalowsky, Christian Ebenbauer, The multidimensional n-th order heavy ball method and its application to extremum seeking, in: Proc. of the 53rd IEEE Conference on Decision and Control (CDC), Los Angeles, USA, 2014, pp. 2660–2666.

[7] Ashia C. Wilson, Benjamin Recht, Michael I. Jordan, A Lyapunov analysis of momentum methods in optimization, 2016, arXiv preprint arXiv:1611.02635.

[8] Yoel Drori, Marc Teboulle, Performance of first-order methods for smooth convex minimization: a novel approach, Math. Program. 145 (1–2) (2014) 451–482.

[9] Adrien B. Taylor, Julien M. Hendrickx, François Glineur, Smooth strongly convex interpolation and exact worst-case performance of first-order methods, Math. Program. 161 (1–2) (2017) 307–345.

[10] Simon Michalowsky, Carsten Scherer, Christian Ebenbauer, Robust and structure exploiting optimization algorithms: An integral quadratic constraint approach, Int. J. Control (2020) (ArXiv Preprint arXiv:1905.00279).

[11] Adrien Taylor, Bryan Van Scoy, Laurent Lessard, Lyapunov functions for first-order methods: Tight automated convergence guarantees, 2018, arXiv preprint arXiv:1803.06073.

[12] Vladimir A. Yakubovich, The method of matrix inequalities in the stability theory of nonlinear control systems. II. Absolute stability in a class of nonlinearities with a condition on the derivative, Autom. Remote Control 26 (1965) 577–592.

[13] R. Josselson, G.V.S. Raju, Absolute stability of control systems with many sector and slope-restricted non-linearities, Internat. J. Control 19 (3) (1974) 609–614.

[14] PooGyeon Park, Stability criteria of sector-and slope-restricted lur'e systems, IEEE Trans. Automat. Control 47 (2) (2002) 308–313.
[15] Jak Suykens, Joos Vandewalle, Bart De Moor, An absolute stability criterion for the Lur'e problem with sector and slope restricted nonlinearities, IEEE Trans. Circuits Syst. I 45 (9) (1998) 1007–1009.
[16] Yurii Nesterov, Introductory lectures on convex programming volume i: Basic course, Lect. Notes 3 (4) (1998) 5.
[17] Ion Necoara, Yu Nesterov, Francois Glineur, Linear convergence of first order methods for non-strongly convex optimization, Math. Program. 175 (1–2) (2019) 69–107.