Analysis and Synthesis of Switched Optimization Algorithms ⋆

Jared Miller * **Fabian Jakob** ** **Carsten Scherer** * **Andrea Iannelli** **

* *Chair of Mathematical Systems Theory, Department of Mathematics, University of Stuttgart, Stuttgart, Germany (e-mail: {jared.miller, carsten.scherer}@imng.uni-stuttgart.de).*

** *Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany (e-mail: {fabian.jakob, andrea.iannelli}@ist.uni-stuttgart.de).*

Abstract: Deployment of optimization algorithms over communication networks face challenges associated with time delays and corruptions. Fixed time delays can destabilize popular gradient-based algorithms, and this degradation is exacerbated by time-varying delays that may arise from packet drops. This work concentrates on the analysis and synthesis of discrete-time optimization algorithms with certified exponential convergence rates that are robust against switched network dynamics between the optimizer and the gradient oracle. Analysis is accomplished by solving linear matrix inequalities under bisection in the exponential convergence rate, searching over Zames-Falb filter coefficients that can certify convergence. Synthesis is performed by alternating between a search over filter coefficient for a fixed controller, and a search over controllers for a fixed filter. Effectiveness is demonstrated by the synthesis of convergent optimization algorithms over networks with time-varying delays, and networks with unstable channel dynamics.

Keywords: Convex Optimization, Robust Control, Switched Systems

1. INTRODUCTION

Optimization algorithms increasingly operate within networked environments, where information exchange is rarely perfect. Communication links introduce delays, packet drops, and other distortions that alter how optimization algorithms interact with gradients. These effects can have significant consequences – slowing convergence, degrading performance, or even destabilizing the algorithm dynamics.

Algorithms should be designed such that they are applicable to a class of objective functions, such as functions with prescribed strong convexity or smoothness characteristics. In the strongly convex setting, robust control techniques have been used to analyze and synthesize optimization algorithms. Introduced by (Lessard et al., 2016), the essential idea is to treat a first-order algorithm as the interconnection of a linear system with a static gradient nonlinearity, and to frame algorithm convergence as an absolute stability problem of the resultant Lur'e system Input-output relations satisfied by sequences arising from gradients inside the function class are described using Integral Quadratic Constraints (IQCs) (Megretski and Rantzer, 1997), and the analysis and design of robust algorithms can subsequently be cast as a robust control problem (Scherer and Ebenbauer, 2021). Most of this system-theoretic approach has focused on algorithm and communication dynamics that are each time-invariant (Scherer et al., 2023). Even though recently this approach was extended to time-varying objectives and algorithms in (Jakob and Iannelli, 2025), important network-induced phenomena such as *time-varying delays* have not been specifically addressed yet. When such effects are taken into account, the resulting dynamics are more accurately captured by switched systems (Wen'an et al., 2008; Conte et al., 2020). Motivated by this perspective, this work extends the system-theoretic framework to include *switched optimization algorithms*. The design of optimization algorithms in a switched setting has mostly focused on analysis and synthesis of distributed algorithms that are robust against time-varying information delays in the form of stale (time-delayed) gradients (Doostmohammadian et al., 2021). Delay-scheduled algorithms, in analogy to delay-scheduled controllers (Briat et al., 2009), have not yet been addressed.

This work performs analysis of switched optimization algorithms by solving mode-dependent (Apkarian et al., 1995) linear matrix inequalities (LMIs) (Scherer, 2023). The switched optimization synthesis problem is posed as a robustness requirement under a regulation constraint (Francis and Wonham, 1976; Stoorvogel et al., 2000), which is compatible under mode changes (Conte et al., 2021).

⋆ J. Miller and C. Scherer are by supported the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2075 – 390740016. We acknowledge the support by the Stuttgart Center for Simulation Science (SimTech). F. Jakob acknowledges the support of the International Max Planck Research School for Intelligent Systems.

Obedience of the regulation constraint is ensured through introduction of a network-dependent internal model.

Given fixed Zames-Falb filters from the analysis procedure, we perform switched optimization algorithm synthesis by applying controller design techniques for LPV systems (De Caigny et al., 2012). The convergence rates and empirical performance of the resulting algorithms are demonstrated on networks with unstable channel dynamics, and networks with time-varying delays. An extended version of the paper is available at (Miller et al., 2025).

Notation. The set of natural numbers including 0 is $\mathbb{N}$. The set of natural numbers between a and b inclusive is $\{a, \ldots, b\}$. An $\mathbb{R}^n$-valued signal is a map $x : \mathbb{N} \to \mathbb{R}^n$ indexed as $\{x_k\}_{k\in\mathbb{N}}$. The 2-norm of a vector $g \in \mathbb{R}^n$ is $\|g\|_2$. The symbol $e_i \in \mathbb{R}^n$ will denote the standard unit vector. The set of symmetric matrices of dimension n is $\mathbb{S}^n$. Positive (semi)-definitness of a symmetric matrix M will be denoted as $M \succ 0$ ($M \succeq 0$). The identity matrix of dimension n is I_n. The zeros and all-ones matrices of dimension $m \times n$ are $0_{m\times n}$ and $1_{m\times n}$, respectively.

2. PRELIMINARIES

2.1 Switched Systems

A switched linear system with N_s modes, n states, n_u inputs, and n_y outputs is described by a tuple $(\Sigma, \mathcal{G})$, where Σ collects together the N_s subsystems, and $\mathcal{G}$ is a directed, unweighted graph with N_s-vertices. Each subsystem Σ_r for $r \in \{1, \ldots, N_s\}$ is assumed to be described by

$$\Sigma_r : \quad \begin{pmatrix} x_{k+1} \\ \hline y_k \end{pmatrix} = \left(\begin{array}{c|c} \mathcal{A}_r & \mathcal{B}_r \\ \hline \mathcal{C}_r & \mathcal{D}_r \end{array}\right) \begin{pmatrix} x_k \\ \hline u_k \end{pmatrix}. \tag{1}$$

The notation Σ_r will refer to the specific representation in (1) at each $r \in 1, \ldots, s$. The vertices $\mathcal{V}$ of the switching graph $\mathcal{G}$ are the modes $r \in \{1, \ldots, N_s\}$. The edges $\mathcal{E}$: $(r, r') \in \mathcal{V} \times \mathcal{V}$ of $\mathcal{G}$ encode possible switching transitions. The set of all infinite-length paths $\{s_k\}_{k\in\mathbb{N}}$ in $\mathcal{G}$ is denoted as Path($\mathcal{G}$), in which Path($\mathcal{G}$) is a subset of all maps $\mathbb{N} \to \{1, \ldots, N_s\}$ with $(s_k, s_{k+1}) \in \mathcal{E} \quad \forall k \in \mathbb{N}$. The switched system $(\Sigma, \mathcal{G})$ can therefore be interpreted as a map $(\Sigma, \mathcal{G}) : (x_0, s, u) \to y$.

A trajectory of (1) is a tuple $(x_k, u_k, y_k, s_k)_{k\in\mathbb{N}}$. This tuple is uniquely specified by an initial condition $x_0 \in \mathbb{R}^n$, an input sequence $u : \mathbb{N} \to \mathbb{R}^{n_u}$, and a switching sequence $s \in$ Path($\mathcal{G}$). The switched system $(\Sigma, \mathcal{G})$ is bounded (Lyapunov Stable) if there exists a $\gamma > 0$ such that when $u = 0$, for all initial points $x_0 \in \mathbb{R}^n$ and switching sequences $s \in$ Path($\mathcal{G}$), the relation $\|x_k\|_2 \leq \gamma\|x_0\|_2$ is obeyed. The switched system is asymptotically stable if when $u = 0$ it holds that $\lim_{k\to\infty} x_k = 0$ for all initial points $x_0 \in \mathbb{R}^n$ and switching sequences $s \in$ Path($\mathcal{G}$). The switched system rejects constant disturbances if it is asymptotically stable and $\lim_{k\to\infty} y_k = 0$ holds for any constant input u (Francis and Wonham, 1976).

Given two *switched* systems $(\Sigma_1, \mathcal{G}) : (x_1, s, (q, z)) \to (p, w)$ and $(\Sigma_2, \mathcal{G}) : (x_2, s, (w, u)) \to (z, y)$ sharing the same transition graph $\mathcal{G}$, we denote the star product $(\Sigma_1 \star \Sigma_2, \mathcal{G}) : ((x_1, x_2), s, (q, u)) \to (p, y)$ as their well-posed feedback interconnection comprised of taking the star product (Zhou and Doyle, 1998, Page 267-269) between the per-mode linear systems as $(\Sigma_1 \star \Sigma_2)_r = \Sigma_{1,r} \star \Sigma_{2,r}$ $\forall r \in 1, \ldots, s$. The common graph $\mathcal{G}$ may be omitted in the presentation of switched systems to simplify notation.

3. SWITCHED OPTIMIZATION ALGORITHMS

Given scalars $0 < m < L < \infty$, we define the class $\mathcal{S}_{m,L}$ as the set of functions $f : \mathbb{R}^c \to \mathbb{R}$ such that $f - \frac{m}{2}\|\cdot\|_2^2$ and $\frac{L}{2}\|\cdot\|_2^2 - f$ are both convex In particular, $\mathcal{S}_{m,L}$ includes the set of twice-continuously-differentiable functions f that have a bounded Hessian as $mI \preceq \nabla^2 f(z) \preceq LI$ for all $z \in \mathbb{R}^c$. The optimization problem under consideration is

$$z^* \in \operatorname*{argmin}_{z\in\mathbb{R}^c} f(z), \tag{2}$$

in which $f \in \mathcal{S}_{m,L}$. Because $m > 0$, the optimum z^* solving (2) is unique. The set $\mathcal{S}^0_{m,L} \subset \mathcal{S}_{m,L}$ is the class of functions $f \in \mathcal{S}_{m,L}$ such that $f(0) = 0$ and $\nabla f(0) = 0$; if $f \in \mathcal{S}^0_{m,L}$, then $z^* = 0$ satisfies (2).

An optimization algorithm is a causal procedure that returns a sequence of iterates $\{z_k\}_{k\in\mathbb{N}}$ given an initial iterate $z_0 \in \mathbb{R}^c$ and a function $f \in \mathcal{S}_{m,L}$. The optimization algorithm is globally *convergent* if for every initial iterate $z_0 \in \mathbb{R}^c$, the returned sequence of iterates $\{z_k\}_{k\in\mathbb{N}}$ satisfies $\lim_{k\to\infty} z_k = z^*$ where z^* satisfies (2).

3.1 Algorithm Model

The algorithmic interconnection of the gradient of a function $f \in \mathcal{S}_{m,L}$ and switched system $(\Sigma, \mathcal{G})$ with $\mathcal{D}_r = 0$ at all modes is

$$\begin{pmatrix} x_{k+1} \\ \hline z_k \end{pmatrix} = \left(\begin{array}{c|c} \mathcal{A}_{s_k} & \mathcal{B}_{s_k} \\ \hline \mathcal{C}_{s_k} & 0 \end{array}\right) \begin{pmatrix} x_k \\ \hline w_k \end{pmatrix}, \quad w_k = \nabla f(z_k). \tag{3}$$

A trajectory of the algorithmic interconnection (3) is a tuple $(x_k, w_k, z_k, s_k)_{k\in\mathbb{N}}$ satisfying (3) for all $k \in \mathbb{N}$.

A *fixed orbit* of (3) is a tuple $(\{x^*_r\}_{r=1}^{N_s}, w^*, z^*)$ such that

$$\forall (r, r') \in \mathcal{E} : \begin{pmatrix} x^*_{r'} \\ \hline z^* \end{pmatrix} = \left(\begin{array}{c|c} \mathcal{A}_r & \mathcal{B}_r \\ \hline \mathcal{C}_r & 0 \end{array}\right) \begin{pmatrix} x^*_r \\ \hline w^* \end{pmatrix}, \; w^* = \nabla f(z^*).$$

The interconnection in (3) is a *convergent switched optimization algorithm* if a fixed orbit $(\{x^*_r\}_{r=1}^{N_s}, w^*, z^*)$ exists such that the following properties are obeyed:

1. $w^* = 0$.
2. If $z^* = 0$, then $x^*_r = 0$ for all $r \in \{1, \ldots, N_s\}$.
3. For every trajectory (x_k, w_k, z_k, s_k), and as $k \to \infty$:
$$\|x_k - x^*_{s_k}\|_2 \to 0, \quad \|z_k - z^*\|_2 \to 0, \quad w_k \to 0. \tag{4}$$

3.2 Algorithms over Networks

The setting of switched networked optimization algorithms includes an algorithm (3) in which Σ is formed through the interconnection of a network $(P, \mathcal{G})$ with state x^N and a controller $(K, \mathcal{G})$ with state ξ. For each fixed mode $r \in \{1, \ldots, N_s\}$, the algorithmic interconnection is described by

$$\nabla f : \qquad w_k = \nabla f(z_k), \tag{5a}$$

$$P_r : \quad \begin{pmatrix} x^N_{k+1} \\ \hline z_k \\ y_k \end{pmatrix} = \left(\begin{array}{c|cc} A_r & B_{r,1} & B_{r,2} \\ \hline C_{r,1} & D_{r,11} & D_{r,12} \\ C_{r,2} & D_{r,21} & D_{r,22} \end{array}\right) \begin{pmatrix} x^N_k \\ \hline w_k \\ u_k \end{pmatrix}, \tag{5b}$$

$$K_r : \quad \begin{pmatrix} \xi_{k+1} \\ \hline u_k \end{pmatrix} = \left(\begin{array}{c|c} A^K_r & B^K_r \\ \hline C^K_r & D^K_r \end{array}\right) \begin{pmatrix} \xi_k \\ \hline y_k \end{pmatrix}. \tag{5c}$$

The resulting system is the interconnection $(\Sigma, \mathcal{G}) = (P \star K, \mathcal{G})$. We refer to (Conte et al., 2020) for how time-varying delays can be modeled as a switched system.

4. OPTIMIZATION ALGORITHM ANALYSIS

We begin by providing a sufficient condition based on regulation theory under which switched optimization algorithms can converge to the optimal point z^* (2). We then use LMI methods to upper-bound the exponential convergence rate of a given switched optimization algorithm.

4.1 Regulation of Switched Optimization Algorithms

We first describe a regulation property (Francis and Wonham, 1976; Wen'an et al., 2008) that is sufficient to ensure that trajectories of (3) converge to an optimal point z^*.

Theorem 1. Interconnection (3) converges to an optimal solution z^* for any $s \in \text{Path}(\mathcal{G})$, $f \in \mathcal{S}_{m,L}$ if there exist matrices $(\Theta_r)_{r=1}^{N_s}$ such that

$$\forall (r, r') \in \mathcal{E}: \qquad \mathcal{A}_r \Theta_r - \Theta_{r'} = 0, \tag{6a}$$

$$\forall r \in \{1, \dots, N_s\}: \qquad \mathcal{C}_r \Theta_r = I_c, \tag{6b}$$

and such that that the following switched Luré interconnection is asymptotically stable for all $v_0 \in \mathbb{R}^n$, $f^0 \in \mathcal{S}^0_{m,L}$, $s \in \text{Path}(\mathcal{G})$:

$$v_{k+1} = \mathcal{A}_{s_k} v_k + \mathcal{B}_{s_k} \tilde{w}_k \qquad \tilde{w}_k = \nabla f^0(\tilde{z}_k), \tag{7a}$$

$$\tilde{z}_k = \mathcal{C}_{s_k} v_k. \tag{7b}$$

Proof: See Theorem 4.1 of (Miller et al., 2025). □

4.2 Exponential Convergence Rate Analysis

Theorem 1 offers a condition for asymptotic convergence to z^*. The speed of asymptotic convergence can be judged by its exponential convergence rate. The algorithmic interconnection in (3) is ρ-exponentially convergent for $\rho \in (0, 1)$ if there exists a $\gamma_z > 0$ such that the following inequality is obeyed by any sequence z generated by the algorithm

$$\|z_k - z^*\|_2 \leq \gamma_z \rho^k \|z_0 - z^*\|_2 \qquad \forall k \in \mathbb{N}. \tag{8}$$

The system $(\Sigma, \mathcal{G})$ forms a ρ-exponentially convergent *switched optimization algorithm* with respect to the function $f \in \mathcal{S}_{m,L}$ if (4) and (8) both hold. The system is ρ-exponentially convergent in the state x to the fixed orbit $\{x^*_r\}_{r=1}^{N_s}$ if the following inequality is satisfied at any initial condition x_0 and switching signal $s \in \text{Path}(\mathcal{G})$ of (3):

$$\|x_k - x^*_{s_k}\|_2 \leq \rho^k \|x_0 - x^*_{s_k}\|_2 \qquad \forall k \in \mathbb{N}. \tag{9}$$

The analysis problem over $\mathcal{S}_{m,L}$ is as follows.

Problem 1. Given a function class $\mathcal{S}_{m,L}$, a switched system $(\Sigma, \mathcal{G})$, and a rate $\rho > 0$, certify if $(\Sigma, \mathcal{G})$ is a ρ-convergent switched optimization algorithm for all $f \in \mathcal{S}_{m,L}$.

To solve Problem 1, we use the following Corollary.

Corollary 1. A sufficient condition for the algorithm in (3) to be ρ-convergent to z^* for all $f \in \mathcal{S}_{m,L}$ as in (8) is if the condition (6) holds, and the following system is bounded:

$$\bar{w}_k = \rho^{-k} \nabla f^0(\rho^k \bar{z}_k),$$

$$\bar{v}_{k+1} = (\rho^{-1}\mathcal{A}_{s_k})\bar{v}_k + (\rho^{-1}\mathcal{B}_{s_k})\bar{w}_k, \tag{10a}$$

$$\bar{z}_k = \mathcal{C}_{s_k} \bar{v}_k. \tag{10b}$$

This proof follows arguments from (Scherer, 2023) about exponential stability. Let $\mathcal{T}_{\rho^{-1}}$ denote the ρ-exponential signal weighting map, transforming a sequence v as

$$\mathcal{T}_{\rho^{-1}}: (v_0, v_1, v_2, \dots) \mapsto (v_0, \rho^{-1} v_1, \rho^{-2} v_2, \dots). \tag{11}$$

Defining the exponentially weighted signals $\bar{v} = \mathcal{T}_{\rho^{-1}} v, \bar{e} = \mathcal{T}_{\rho^{-1}} e, \bar{w} = \mathcal{T}_{\rho^{-1}} w$, we note that boundedness of $\bar{v}$ as $\|\bar{v}_k\|_2 \leq \gamma \|\bar{v}_0\|_2$ over all $k \in \mathbb{N}, x_0 \in \mathbb{R}^n$ for some $\gamma > 0$, implies exponential stability of the unweighted signal v as $\|v_k\|_2 \leq \gamma \rho^{-k} \|v_0\|_2$ for all $k \in \mathbb{N}, v_0 \in \mathbb{R}^n$. □

We also use the bar notation to denote the ρ-exponential weighting of a switched linear system as in (10), such as $\bar{G}$ resulting from G with the following state space descriptions at each $r \in 1, \dots, N_s$:

$$G_r = \left(\begin{array}{c|c} \mathcal{A}_r & \mathcal{B}_r \\ \hline \mathcal{C}_r & \mathcal{D}_r \end{array}\right) \qquad \bar{G}_r = \left(\begin{array}{c|c} \rho^{-1}\mathcal{A}_r & \rho^{-1}\mathcal{B}_r \\ \hline \mathcal{C}_r & \mathcal{D}_r \end{array}\right). \tag{12}$$

Next, we use families of dissipation inequalities satisfied by gradients of functions f in $\mathcal{S}^0_{m,L}$ to construct multiplier relaxations that can certify ρ-convergence.

Lemma 1. (Lemma 5 of (Scherer et al., 2023)). Consider a function $f^0 \in \mathcal{S}^0_{m,L}$, an exponential rate $\rho \in (0, 1]$, and a sequence of scalar coefficients $\{\lambda_\nu\}_{\nu=0}^{\nu_{\max}}$ satisfying

$$\lambda_\nu \leq 0 \quad \forall \nu \geq 1, \qquad \textstyle\sum_{\nu=0}^{\nu_{\max}} \rho^{-\nu} \lambda_\nu > 0. \tag{13}$$

For any sequence z, define the exponentially weighted sequences $\bar{z}, \bar{p}, \bar{q}, \bar{g}$ as

$$\bar{z}_k := \rho^{-k} z_k, \qquad \bar{q}_k := \rho^{-k} \nabla f^0(\rho^k \bar{z}_k) - m\bar{z}_k,$$

$$\bar{p}_k := L\bar{z}_k - \rho^{-k} \nabla f^0(\rho^k \bar{z}_k), \quad \bar{g}_k := \textstyle\sum_{\nu=0}^{k} \lambda_\nu \bar{p}_{k-\nu}. \tag{14}$$

Then the sequences $\bar{q}, \bar{g}$ satisfy the following dissipation inequality for all $T \in \mathbb{N}, T > 0$:

$$\textstyle\sum_{k=0}^{T-1} \bar{q}_k^\top \bar{g}_k \geq 0. \tag{15}$$

The coefficients λ in (13) parameterize a Zames-Falb Finite Impulse Response (FIR) filter $\Psi(\lambda) = \sum_{\nu=0}^{\nu_{\max}} \lambda_\nu \mathbf{z}^{-\nu}$. The choice of $\lambda_0 = 1$ and $\lambda_\nu = 0$ for all $\nu \geq 1$ leads to a passivity property (identity filter). Cases with $\lambda_\nu \neq 0$ for some $\nu \geq 1$ describe dynamic filters.

A state space realization of $\Psi(\lambda)$ may be defined using fixed matrices A_f, B_f, and λ-affine maps $C_f(\lambda), D_f(\lambda)$ as

$$\begin{pmatrix} \psi_{k+1} \\ \bar{g}_k \end{pmatrix} = \left(\begin{array}{c|c} A_f & B_f \\ \hline C_f(\lambda) & D_f(\lambda) \end{array}\right) \begin{pmatrix} \psi_k \\ \bar{p}_k \end{pmatrix}. \tag{16}$$

We conduct a signal transformation in order to render the conditions in Lemma 1 in the form of a star product:

$$\begin{pmatrix} \bar{p}_k \\ \bar{w}_k \end{pmatrix} = \begin{pmatrix} (L-m)I & I \\ I & mI \end{pmatrix} \begin{pmatrix} \bar{q}_k \\ \bar{z}_k \end{pmatrix}. \tag{17}$$

Given a set of filter coefficitents λ and a rate ρ, we form the system $\hat{G}^\lambda$ from a switched system $(\Sigma, \mathcal{G})$ as

$$\hat{\Sigma}^\lambda = \Psi(\lambda) \left[\begin{pmatrix} (L-m)I & I \\ I & mI \end{pmatrix} \star \bar{G} \right] \tag{18}$$

$$\text{as described by} \quad \left(\begin{array}{c|c} \hat{\mathcal{A}}^\lambda_r & \hat{\mathcal{B}}^\lambda_r \\ \hline \hat{\mathcal{C}}^\lambda_r & \hat{\mathcal{D}}^\lambda_r \end{array}\right). \tag{19}$$

Theorem 2. The switched algorithm $(\Sigma, \mathcal{G})$ from (3) is certified as ρ-exponentially convergent for all $f \in \mathcal{S}_{m,L}$ if the condition in (6) is satisfied, and there exists a filter length $\nu_{\max} \in \mathbb{N}$, filter coefficients $\lambda \in \mathbb{R}^{\nu_{\max}}$ satisfying (13), and matrices $M_r \in \mathbb{S}^{n+\nu_{\max}}$, $M_r \succ 0$ for each $r \in 1, \dots, s$ such that for all $\forall (r, r') \in \mathcal{E}$:

$$[\star]^\top \begin{pmatrix} M_r & 0 \\ 0 & -M_{r'} \end{pmatrix} \begin{pmatrix} \hat{\mathcal{A}}_r^\lambda & \hat{\mathcal{B}}_r^\lambda \\ I & 0 \end{pmatrix} + [\star]^\top \begin{pmatrix} 0 & I_c \\ I_c & 0 \end{pmatrix} \begin{pmatrix} \hat{\mathcal{C}}_r^\lambda & \hat{\mathcal{D}}_r^\lambda \\ 0 & I_c \end{pmatrix} \prec 0. \tag{20}$$

Proof: A detailed proof follows by using arguments from Scherer et al. (2023); Scherer and Ebenbauer (2025). The equation in (20) can be interpreted as strict negative-realness of the system in (19). Lemma 1 details that gradients of all functions in $\mathcal{S}^0_{m,L}$ are passive under a sequence of transformations via the dissipation relation (15). The interconnection between a strictly antipassive system and a passive system is bounded, hence, feasibility of (20) ensures that the system in (3) is ρ-exponentially stable for all $\mathcal{S}^0_{m,L}$. □

Exponential convergence rates of Σ can be upper-bounded by performing a bisection on ρ. The accuracy of the bound can be increased by increasing $\nu_{\max}$, but there is no guarantee of convergence to the true rate bound.

5. ALGORITHM SYNTHESIS

We now focus on the synthesis of switching optimization methods. The algorithm synthesis problem with fixed filter $\Psi(\lambda)$ satisfying (13) and convergence bound ρ is as follows.

Problem 2. Given a function class $\mathcal{S}_{m,L}$, a network $(P, \mathcal{G})$ from (5b), and filter coefficients $\lambda \in \mathbb{R}^{\nu_{\max}}$ satisfying (13), synthesize a controller (5c) $(K, \mathcal{G})$ such that $(P \star K, \mathcal{G})$ from (3) forms a ρ-exponentially convergent algorithm when interconnected with any $f \in \mathcal{S}_{m,L}$.

Assumptions. The regulation property from Theorem 1 will be guaranteed through the introduction of an internal model. This internal-model-based regulation technique is only possible if a regulator equation is solvable:

Lemma 2. If $G = P \star K$ is a convergent optimization algorithm for all $f \in \mathcal{S}_{m,L}$, and $\{\Theta_r\}_{r=1}^{N_s}$ are solutions to (6) with respect to the plant representations $\{P_r \star K_r\}_{r=1}^{N_s}$, then there exists a solution $\{\Pi_r, \Gamma_r\}_{r=1}^{N_s}$ of the following *regulator equation*

$$\forall (r, r') \in \mathcal{E}: \quad \begin{pmatrix} A_r & B_{r,2} \\ C_{r,1} & D_{r,12} \end{pmatrix} \begin{pmatrix} \Pi_r \\ \Gamma_r \end{pmatrix} = \begin{pmatrix} \Pi_{r'} \\ I_c \end{pmatrix}. \tag{21}$$

Proof: We can vertically partition each matrix Θ_r as $\Theta_r = [\Pi_r^\top, \Xi_r^\top]^\top$ for every $r \in \{1, \ldots, N_s\}$. Given that the star product $P_r \star K_r$ requires that $(I - D_2^K D_{r,22})$ is invertible at each r, the following matrices

$$\Gamma_r := (I - D_r^K D_{r,22})^{-1} D_r^K (C_{2,r}\Xi_r + C_{2,r}\Pi_r), \tag{22}$$
$$\Phi_r := C_{2,r}\Pi_r + D_{2,r}\Gamma_r, \tag{23}$$

are unique at each $r \in \{1, \ldots, N_s\}$, and solve (21). □

We therefore require the following assumptions to perform synthesis of switched ρ-convergent optimization algorithms that can be certified by Theorem 2.

Assumption 1. Given a network P in (5b), the regulator equation (21) has a solution $\{\Pi_r, \Gamma_r\}_{r=1}^{N_s}$.

Assumption 2. The matrix $(I - D_{r,11}m)$ is invertible for each $r \in \{1, \ldots, N_s\}$.

If Assumption 1 is violated, then there cannot exist a controller K such that $P \star K$ can be certified as convergent by Theorem 1 (Lemma 2). Solvability of 1 is a property only of the network P, and is indepenedent of K. Assumption 2 ensures the interconnection of $D_{r,11}$ and the error system with the known function $f^0(\tilde{z}) = \frac{m}{2}\|\tilde{z}\|_2^2$, $\nabla f^0(\tilde{z}) = m\tilde{z}$ as described by $(mI) \star P_r$ is well-posed for each mode r.

5.1 Regulation

Given a solution $\{\Pi_r, \Gamma_r\}_{r=1}^{N_s}$ to (21), we define $\Phi_r := C_{r,2}\Pi_r + D_{r,22}\Gamma_r$ for each $r \in \{1, \ldots, N_s\}$. We construct controllers K as the interconnection between fixed internal models Q and designed subcontrollers R. An internal model Q inspired by (Stoorvogel et al., 2000) (originally derived for a non-switched system) is described for each $r \in \{1, \ldots, N_s\}$ by

$$Q_r: \quad \begin{pmatrix} \omega_{k+1} \\ \hline u_k \\ \tilde{y}_k \end{pmatrix} = \left(\begin{array}{c|c:cc} I_c & 0 & I & 0 \\ \hline -\Gamma_r & 0 & 0 & I \\ \Phi_r & I & 0 & 0 \end{array}\right) \begin{pmatrix} \omega_k \\ \hline y_k \\ \hdashline \tilde{u}_{1k} \\ \tilde{u}_{2k} \end{pmatrix}. \tag{24}$$

The subcontrollers R are described by

$$R_r: \quad \begin{pmatrix} \tilde{\xi}_{k+1} \\ \hline \tilde{y}_k \end{pmatrix} = \left(\begin{array}{c|c} A_{r,c} & B_{r,c} \\ \hline C_{r,c1} & D_{r,c1} \\ C_{r,c2} & D_{r,c2} \end{array}\right) \begin{pmatrix} \tilde{\xi}_k \\ \hline \tilde{u}_{1,k} \\ \tilde{u}_{2,k} \end{pmatrix}, \tag{25}$$

for which the controller K in (5c) can be expressed as the interconnection $K = Q \star R$. This interconnection may be described for each r as

$$K_r: \left(\begin{array}{c|c} A_r^K & B_r^K \\ \hline C_r^K & D_r^K \end{array}\right) = \left(\begin{array}{c|c:c} I + D_{r,c1}\Phi_r & C_{r,c1} & D_{r,c1} \\ \hline B_{r,c}\Phi_r & A_{r,c} & B_{r,c} \\ -\Gamma_r + D_{r,c2}\Phi_r & C_{r,c2} & D_{r,c2} \end{array}\right). \tag{26}$$

Requiring that $D_{r,11} + D_{r,12}D_{r,c2}D_{r,21} = 0$ at all modes $r \in, 1\ldots, s$ when designing the subcontroller R ensures that $P \star Q \star R$ has no algebraic loops.

Theorem 3. If Assumption 1 is satisfied, then the controller structure (26) ensures that $P \star Q \star R$ satisfies condition (6) for output regulation from Theorem 1.

Proof: See Theorem 5.2 of (Miller et al., 2025). □

5.2 Controller Synthesis

The controller synthesis problem can be posed as finding a sub-controller $(R, \mathcal{G})$ ensuring that $(P \star Q \star R, \mathcal{G})$ is a ρ-convergent optimization algorithm for any $f \in \mathcal{S}_{m,L}$. For a fixed filter λ satisfying (13), the control synthesis problem may be solved by finding per-mode controllers $\{R_r\}_{r=1}^{N_s}$ such that the system $\hat{G}^\lambda$ with individual modes defined for each $r \in \{1, \ldots, N_s\}$ as

$$\hat{G}_r^\lambda := \Psi(\lambda) \left[\begin{pmatrix} (L-m)I & I \\ I & mI \end{pmatrix} \star \bar{P}_r \star \bar{Q}_r \star \bar{R}_r \right], \tag{27}$$

satisfies the LMI in (20) across all arcs $(r, r') \in \mathcal{E}$. The controller synthesis and reconstruction procedure follows established methods for Linear Parameter Varying synthesis, see (De Caigny et al., 2012) for further detail. The alternating search between λ and R for controller synthesis is described in Algorithm 1. At the end of IterMax iterations, Algorithm 1 will return an exponential convergence rate ρ and a subcontroller R such that $(P \star Q \star R, \mathcal{G})$ is ρ-convergent. Certified algorithm convergence is achieved if $\rho < 1$.

Algorithm 1 Synthesis of Switched Optimization Algorithms

Require: A switching graph $\mathcal{G}$, a networked system P (5b), constants (m, L), filter length $\nu_{\max}$

iter $\leftarrow 1$, Initial filter $\Psi(\lambda) \leftarrow 1$
Compute internal models Q from (24)
for iter $\in 1, \ldots,$ IterMax **do**
    Synthesis: $R \leftarrow$ feasible subcontroller with infimal ρ under fixed filter $\Psi(\lambda)$ given $P \star Q$.
    Recover the switched algorithm $(P \star Q \star R)$
    Analysis: $\lambda \leftarrow$ filter coefficients from Theorem 2 for analysis of $(P \star Q \star R)$ with infimal ρ
end for
return rate ρ, filter λ, subcontroller R

6. NUMERICAL EXAMPLES

Numerical experiments are implemented in MATLAB (2024a). All LMIs are solved using LMIlab from the Robust Control Toolbox of MATLAB. The code to generate experiments is publicly available[1]. These examples aim to find the optimal value of the following quadratic plus log-sum-exp function parameterized by terms $\Lambda \in \mathbb{R}^{c\times c}$, $b \in \mathbb{R}^c$, and $\ell \in \mathbb{R}$:

$$f(z; \Lambda, b, \ell) = \frac{1}{2} z^\top \Lambda z + b^\top z + \ell \log\left(\textstyle\sum_{i=1}^d e^{z_i}\right). \tag{28}$$

The optimal value of (28) must be found when the oracle $\nabla f(\cdot; \Lambda, b, \ell)$ is only accessible through a time-varying network. Scalars $0 < m < L < \infty$ are given to define the function class (28). The vector b is randomly drawn in $\mathbb{R}^c$. The matrix Λ is a randomly generated symmetric matrix with eigenvalues between m and L' for a given $L' \in (m, L)$. The scalar ℓ is defined as $\ell := L - L'$. Under these choices of parameters (b, Λ, ℓ), every instance of f from (28) is a member of $\mathcal{S}_{m,L}$.

6.1 Inhomogenous Networked System

We first consider minimization of f (28) over a network governed by $N_s = 4$ switching modes. The switching logic is displayed in Figure 1.

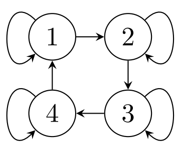

Figure 1. Ring-structured switching graph

The modes $\{P_r\}_{r=1}^4$ of the network P are described by

$$P_1 : \left(\begin{array}{cc|cc} 0 & 0 & 1 & 0 \\ 0 & 0.2 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{array}\right) \otimes I_c, \quad P_2 : \left(\begin{array}{cc|cc} 0.2 & 0 & 0.25 & 0 \\ 0 & 0.9 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -0.4 & 0 & -0.5 & 3 \end{array}\right) \otimes I_c,$$

$$P_3 : \left(\begin{array}{cc|cc} -0.3 & 0 & 0.5 & 0 \\ 0 & -0.5 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ -0.3 & 0 & 0.5 & 1 \end{array}\right) \otimes I_c, \quad P_4 : \left(\begin{array}{cc|cc} 1.2 & 0 & 1 & 0 \\ 0 & -0.2 & 0 & 1 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 2 \end{array}\right) \otimes I_c \tag{29}$$

The dynamics in (29) are chosen such that unstable channel dynamics are present in the subsystems (in P_4), and that (21) admits a solution with non-identical Π and Γ values between modes, with the specific values of

$$\begin{pmatrix} \Gamma_r \\ \Phi_r \end{pmatrix}_{r=1}^4 = \begin{pmatrix} -0.8 & -0.1 & -1.5 & -1.2 \\ 0 & -0.3 & -1.5 & -2.4 \end{pmatrix} \otimes I_c, \tag{30}$$

and $\Pi_r = [0, 1]^\top \otimes I_c$ for all $r \in 1..4$. The solutions (Π, Γ, Φ) in (30) are unique. Thus, the method from Zattoni et al. (2013) requiring mode-independent solutions of the regulator equations cannot be used to certify regulation.

We perform a parameter sweep over increasing L/m. The reference algorithms are gradient descent and triple momentum (Van Scoy et al., 2017) with $\rho_{\text{gd}} = \frac{2}{m+L}$ and $\rho_{\text{TM}} = 1 - \sqrt{\frac{m}{L}}$, which are tuned to optimally converge in the setting of no network dynamics. Figure 2 plots an upper-bound on the convergence rate ρ as a function of L/m developed by solving the analysis program in Theorem 2 with filters of length $\nu_{\max} = 3$. The compared curves in Figure 2 are gradient descent, triple momentum, our proposed synthesis with identity filters in which M_r may be different between modes $r \in \{1, \ldots, N_s\}$, and synthesis under an identity filter with the additional conservatism-introducing constraint that the matrix M_r in the LMI (20) is the same among all modes $r \in \{1, \ldots, N_s\}$. Setting M_r to the same value among all modes r ignores the ring structure of the switching graph switching. All comparisons are performed with respect to $m = 1$, in which L is swept in the range $(1, 5]$. The synthesized path-dependent controller has $\rho < 1$ when $L < 3.360$, and the common storage controller is stable only when $L < 1.660$. Gradient descent and triple momentum are both empirically unstable even when $L = 1.01, L' = 1.005$.

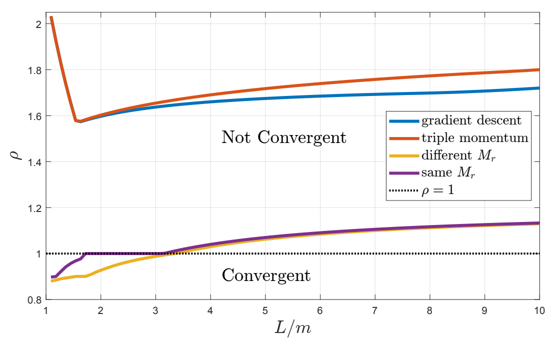

Figure 2. Upper bounds on ρ via Analysis ($\nu_{\max} = 3$)

6.2 Time-Varying Delay: Increasing Delay

Our next example involves synthesis of algorithms with increasing delay before the oracle ∇f. We consider two network structures: Unrestricted Delay and Packet Drops. Both networks involve a maximal delay $H_{\max}$ ranging from $H_{\max} = 0$ to $H_{\max} = 5$. The Unrestricted Delay setting involves a delay of $h_k \in \{0, \ldots, H_{\max}\}$, in which h_k and h_{k+1} are independent. The Packet Drop setting involves the delay increasing by 1 until the maximum delay of $H_{\max}$, or jumps to delay 0.

Figure 3 plots the computed convergence rate ρ by synthesis under identity multipliers λ for $m = 1$ as L

[1] https://github.com/jarmill/time_var_opt

is swept in the range $(1, 10]$. In each subplot, the color indicates the maximal delay $H_{\max}$ between 0 and 5. The top subplot shows the computed bounds on ρ under unrestricted delay (ρ_{all}). The middle subplot depicts bounds on ρ under packet drop logic (ρ_{drop}). The rates ρ rise as the maximal delay increases. The bottom subplot draws the difference between ρ_{all} and ρ_{drop}. The computed rates for $H_{\max} \in \{0, 1\}$ are identical between the unrestricted delay and packet drop cases, when $H_{\max} \geq 2$ the packet-drop-synthesized controller has a smaller ρ than a controller developed for the unrestricted delay case.

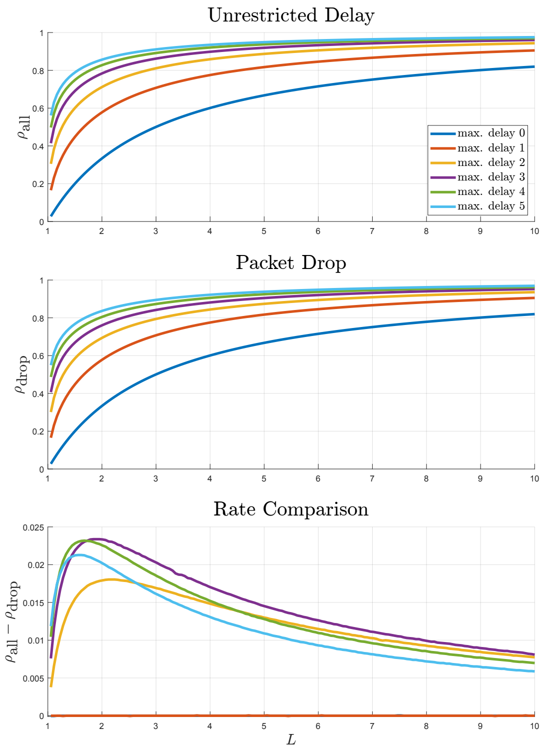

Figure 3. Exponential rates under increasing delay and parameter L

7. CONCLUSION

This paper presented a method to analyze and synthesize switched optimization algorithms. Convergence to the optimal solution may be certified if the per-mode systems solve a regulator equation and achieve stability (Theorem 1). Analysis occurs through convex searches over Zames-Falb filter coefficients λ. Synthesis of switched algorithms given a network P and fixed filters $\Psi(\lambda)$ is accomplished by using regulator theory and output feedback techniques. The analysis and synthesis tasks are jointly convex, and an alternating approach can be used to develop certified switching algorithms with exponential convergence guarantees. Future work involves the establishing of necessary conditions for convergence of switched optimization algorithms, and finding delay-scheduled controllers in the sampled data case, and to synthesize controllers with certified convergence rates in the unknown delay cases.

REFERENCES

Apkarian, P., Gahinet, P., and Becker, G. (1995). Self-scheduled H-infinity control of linear parameter-varying systems: a design example. *Automatica*, 31(9), 1251–1261.

Briat, C., Sename, O., and Lafay, J.F. (2009). Delay-scheduled state-feedback design for time-delay systems with time-varying delays–a LPV approach. *Systems & Control Letters*, 58(9), 664–671.

Conte, G., Perdon, A., Zattoni, E., and Animobono, D. (2021). Disturbance decoupling and model matching problems for discrete-time systems with time-varying delays. *Nonlinear Analysis: Hybrid Systems*, 41, 101043.

Conte, G., Perdon, A.M., and Zattoni, E. (2020). Modeling discrete time systems with variable delays as switching systems without delays. In *2020 European Control Conference (ECC)*, 1591–1594. IEEE.

De Caigny, J., Camino, J.F., Oliveira, R.C., Peres, P.L., and Swevers, J. (2012). Gain-scheduled dynamic output feedback control for discrete-time LPV systems. *Int. J. Robust Nonlinear Control.*, 22(5), 535–558.

Doostmohammadian, M., Khan, U.A., Pirani, M., and Charalambous, T. (2021). Consensus-based distributed estimation in the presence of heterogeneous, time-invariant delays. *IEEE Control Syst. Lett.*, 6, 1598–1603.

Francis, B.A. and Wonham, W.M. (1976). The internal model principle of control theory. *Automatica*, 12(5), 457–465.

Jakob, F. and Iannelli, A. (2025). A Linear Parameter-Varying Framework for the Analysis of Time-Varying Optimization Algorithms. *arXiv:2501.07461*.

Lessard, L., Recht, B., and Packard, A. (2016). Analysis and design of optimization algorithms via integral quadratic constraints. *SIAM J. Optim.*, 26(1), 57–95.

Megretski, A. and Rantzer, A. (1997). System analysis via Integral Quadratic Constraints. *IEEE transactions on automatic control*, 42(6), 819–830.

Miller, J., Jakob, F., Scherer, C., and Iannelli, A. (2025). Analysis and Synthesis of Switched Optimization Algorithms. *arxiv:2510.21490*.

Scherer, C. and Ebenbauer, C. (2021). Convex synthesis of accelerated gradient algorithms. *SIAM J. Control Optim.*, 59(6), 4615–4645.

Scherer, C.W. (2023). Robust Exponential Stability and Invariance Guarantees with General Dynamic O'Shea-Zames-Falb Multipliers. *IFAC-PapersOnLine*, 56(2), 5799–5804.

Scherer, C.W. and Ebenbauer, C. (2025). A Tutorial on Convex Design of Optimization Algorithms by Integral Quadratic Constraints. *Annual Review of Control, Robotics, and Autonomous Systems*, 9, 215–242.

Scherer, C.W., Ebenbauer, C., and Holicki, T. (2023). Optimization algorithm synthesis based on integral quadratic constraints: A tutorial. In *Proc. IEEE Conf. Decis. Control*, 2995–3002.

Stoorvogel, A.A., Saberi, A., and Sannuti, P. (2000). Performance with regulation constraints. *Automatica*, 36(10), 1443–1456.

Van Scoy, B., Freeman, R.A., and Lynch, K.M. (2017). The fastest known globally convergent first-order method for minimizing strongly convex functions. *IEEE Control Systems Letters*, 2(1), 49–54.

Wen'an, Z., Li, Y., and Hongbo, S. (2008). A switched system approach to networked control systems with time-varying delays. In *2008 27th Chinese Control Conference*, 424–427.

Zattoni, E., Perdon, A.M., and Conte, G. (2013). The output regulation problem with stability for linear switching systems: A geometric approach. *Automatica*, 49(10), 2953–2962.

Zhou, K. and Doyle, J.C. (1998). *Essentials of Robust Control*, volume 104. Prentice hall Upper Saddle River, NJ.