This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

Prediction-Correction Interior-Point Method for Time-Varying Convex Optimization

Mahyar Fazlyab†, Santiago Paternain, Victor M. Preciado and Alejandro Ribeiro

***Abstract*—In this paper, we develop an interior-point method for solving a class of convex optimization problems with time-varying objective and constraint functions. Using log-barrier penalty functions, we propose a continuous-time dynamical system for tracking the (time-varying) optimal solution with an asymptotically vanishing error. This dynamical system is composed of two terms: (*i*) a *correction* term consisting of a continuous-time version of Newton's method, and (*ii*) a *prediction* term able to track the drift of the optimal solution by taking into account the time-varying nature of the objective and constraint functions. Using appropriately chosen time-varying slack and barrier parameters, we ensure that the solution to this dynamical system globally asymptotically converges to the optimal solution at an exponential rate. We illustrate the applicability of the proposed method in two applications: a sparsity promoting least squares problem and a collision-free robot navigation problem.**

***Index Terms*—Time-Varying Optimization, Dynamic Optimization, Interior-Point Method.**

I. Introduction

The interplay between optimization and control theory is rich and fruitful, resulting in a plethora of efficient computational tools [1]–[8]. Dynamical systems theory provides an array of mathematical tools to analyze the behavior of iterative algorithms proposed to solve standard optimization problems [9], [10]. In this direction, control theory can be used to guarantee the convergence of iterative algorithms to accurate solutions and to analyze the impact of numerical errors and computational delays. Control tools have been extensively exploited in the context of stationary (i.e., time-invariant) optimization problems, in which both the objective function and constraints do not depend on time [1]–[8]. In many practical settings, however, we find optimization problems in which the objective function and/or the constraints depend explicitly on time [11]–[18]. In particular, time-varying optimization problems appear in, for example, the estimation of the path of a stochastic process [11], signal detection with adaptive filters [12], tracking of moving targets [13], and various problems in autonomous systems [14], [15], computer networks [16], learning [17], [18], and online optimization [19].

In time-varying optimization problems, the optimal solution is a function of time; therefore, solving the optimization problem is equivalent to tracking the optimal solution as it varies over time. A natural approach to addressing this problem is to sample the objective and constraint functions at particular times and to solve the corresponding sequence of (time-invariant) constrained optimization problems using iterative algorithms [20], [21]), or their continuous-time counterparts [22], [23]. However, this approach ignores the dynamic aspect of the problem, since each iteration tends to converge towards the optimal point of the sampled time-invariant problem, while the solution of the time-varying case is drifting away over time. Therefore, this approach is likely to induce a steady-state optimality gap (i.e., a tracking error) whose magnitude depends on the time-varying aspects of the problem. This phenomenon has been previously observed in gradient descent algorithms for unconstrained optimization [24], as well as in constrained optimization problems that arise in distributed robotics [25], sequential estimation [11], distributed optimization [26], and neural networks [27]; see [28] for a unified analysis using monotone operator theory.

Alternatively, one can utilize the knowledge of the dynamics of the optimization problem to predict the drift of the optimal solution and incorporate the descent step of an optimization algorithm to correct the prediction. Variations of this idea have been developed in discrete [29] and continuous [30] time. Specifically, the authors in [29] propose a discrete-time prediction-correction scheme for minimizing continuously time-varying unconstrained smooth objective functions. When used in continuous time, the use of a prediction step and a Newton/gradient correction results in asymptotic tracking of the optimal argument of an *unconstrained* optimization problem [30]. Extensions of this method to distributed continuous-time multi-agent systems have also been reported in [31].

In this paper, we consider time-varying smooth convex optimization problems characterized by (*i*) a convex time-varying objective function, and (*ii*) constraints that are expressed as level sets of time-varying convex functions and affine equalities. We propose a continuous-time dynamical system whose state is globally asymptotically driven to the optimal solution at an exponential rate, under certain technical conditions. In particular, we develop a prediction-correction interior-point method that utilizes information about time variations of the optimization problem in order to predict and correct the drift in the optimal solution, resulting in an asymptotically vanishing optimality gap. We emphasize that the main difference between this work and related works in the literature [29], [30] is the inclusion of time-varying functional constraints, which calls for penalty methods (interior-point methods in particular) to handle the constraints.

The paper is structured as follows. In Section II, we formally state the problem under consideration and introduce some regularity assumptions needed in our derivations. We

†Corresponding author: mahyarfa@seas.upenn.edu. This work was supported by the NSF under grants CAREER-ECCS-1651433, IIS-1447470, and by the ONR under grant N00014-12-1-0997. The authors are with the Department of Electrical and Systems Engineering, University of Pennsylvania.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

then consider the particular case of time-varying optimization problems without constraints, as well as affine constraints only (Section III). In both cases, we propose to track the optimal solution using a dynamical system composed of two terms: (*i*) a 'prediction' term that uses information about time variations of the optimization problem, and (*ii*) a 'correction' term based on a continuous-time version of Newton's method. In Section III-C, we propose a dynamical system able to track the solution of time-varying optimization problems with inequality constraints. In this case, we incorporate logarithmic barrier functions with an appropriately chosen time-varying barrier parameter, as well as a time-varying slack variable used to guarantee global convergence. We show that the proposed dynamical system converges asymptotically to the time-varying optimal point for any (and not necessarily feasible) initial condition (Theorem 1), under mild assumptions. These assumptions correspond to standard requirements to prove convergence of interior-point methods. To illustrate our results, we perform two numerical evaluations (Section IV-A and IV-B) and consider two practical applications. The first application is a time-invariant ℓ_1 regularized least squares problem in which we show that the use of a time-varying barrier parameter along with a prediction term speeds up the convergence of conventional interior-point methods (Section IV-C). The second application involves the navigation of a robot in an environment with circular obstacles (Section IV-D). We further consider situations in which the robot is charged with the task of tracking a moving target (Section IV-D2). We make some concluding remarks in Section V.

Notation. Let $\mathbb{R}$, $\mathbb{R}_+$, and $\mathbb{R}_{++}$ be the set of real, nonnegative, and positive numbers. The set $\{1,\dots,n\}$ is denoted by $[n]$. We denote by I_n the n-dimensional identity matrix. We denote by $\mathbb{S}^n$ the space of $n\times n$ symmetric matrices. The gradient of a function $f(x,t)\colon \mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$ with respect to $x\in\mathbb{R}^n$ is denoted by $\nabla_x f(x,t)$. The *partial* derivatives of $\nabla_x f(x,t)$ with respect to x and t are denoted by $\nabla_{xx} f(x,t)$ and $\nabla_{xt} f(x,t)$, respectively.

II. Problem statement

This paper considers a class of convex optimization programs where both the objective and constraint functions are indexed by continuous time. Formally, consider a variable $x\in\mathbb{R}^n$ and let $t\geq 0$ be a continuous time index. We then define a time-varying (TV) objective function $f_0:\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$ taking values $f_0(x,t)$; we also define p time-varying inequality constraint functions $f_i:\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$ taking values $f_i(x,t)$ for $i\in[p]$; and q time-varying affine equality constraint functions $f_i':\mathbb{R}^n\times\mathbb{R}_+\to\mathbb{R}$ taking values $f_i'(x,t)=a_i(t)^\top x-b_i(t)$ for $i\in[q]$. We assume that $f_i(x,t), i\in\{0\}\cup[p]$ and $f_i'(x,t), i\in[q]$ are twice continuously differentiable with respect to x and piecewise continuously differentiable with respect to t, for all $(x,t)\in\mathbb{R}^n\times\mathbb{R}_+$. For any given time $t\geq 0$, the objective function and the constraints define an optimization problem whose optimal argument $x^*(t)$ is defined as:

$$\begin{aligned} x^\star(t) := \ & \operatorname{argmin}\ f_0(x,t), \\ & \text{s.t.} \quad f_i(x,t)\leq 0, \quad i\in[p], \\ & \quad\ A(t)x=b(t), \end{aligned} \tag{1}$$

where $A(t)\colon\mathbb{R}_+\to\mathbb{R}^{q\times n}$ is defined as $A^\top(t)=[a_1(t)\cdots a_q(t)]$, and $b(t)=[b_1(t)\cdots b_q(t)]^\top$. Further, we assume the minimizer $x^\star(t)$ is unique for each t (see Assumption 1). A naive approach to solving (1) is to sample the problem at particular times, say $0\leq t_0<t_1<\cdots$, and solve the corresponding sequence of (time-invariant) optimization problems. In particular, for each $k\in\mathbb{Z}_+$, one could estimate $x^\star(t_k)$ by assuming the objective and constraints are time-invariant on the time interval $[t_k,t_{k+1})$ and performing standard iterative algorithms such as interior-point methods [21], [32]. However, this implementation is likely to induce steady-state tracking errors. In particular, the suboptimality computed at t_{k+1} depends on the number of iterations allowed by our computational capabilities during the time interval $t\in[t_k,t_{k+1})$, as well as how fast the optimal argument has drifted away from $x^\star(t_k)$ during this interval.

Our goal is to develop an alternative approach to solving TV optimization problems with a vanishing tracking error by leveraging information about the temporal variation of the objective and constraints. More precisely, we propose a continuous-time dynamical system $\dot{x}(t)=h(x(t),t)$ whose solution $x(t)$ satisfies $\|x(t)-x^\star(t)\|\to 0$ as $t\to\infty$ for all initial conditions, i.e., global asymptotic convergence. To facilitate the exposition, we first address TV optimization problems without constraints (Section III-A), and then extend the framework to problems with TV constraints (Section III-B and III-C). In order to make the contributions of the paper more precise, we list below the assumptions that we impose on the optimization problem (1).

Assumption 1 (Convexity) The constraint functions $x\mapsto f_i(x,t),\ i\in[p]$ are convex for all $t\geq 0$, and the objective $x\mapsto f_0(x,t)$ is uniformly strongly convex for all $t\geq 0$, i.e., $\nabla_{xx} f_0(x,t)\succeq m_f\mathrm{I}_n$ for some $m_f>0$.

Assumption 2 (Slater's Condition) The interior of the feasible region is nonempty for all $t\geq 0$, i.e., for each $t\geq 0$, there exists a $x^\dagger(t)\in\mathbb{R}^n$ such that $f_i(x^\dagger(t),t)<0$, for all $i\in[p]$, and $a_i(t)^\top x^\dagger(t)=b_i(t)$, for all $i\in[q]$.

Assumption 3 (Full Rank Condition) The number of equality constraints is less than the dimension of the optimization variables, i.e., $q<n$. Moreover, the vectors $a_i(t),\ i\in[q]$ are linearly independent for all $t\geq 0$. This implies that $\operatorname{rank}(A(t))=q$ for all $t\geq 0$.

The uniform strong convexity of the objective function implies that the optimal trajectory $x^\star(t)$ is unique for all $t\geq 0$. By Assumption 2, the optimal solution $x^\star(t)$ in (1) at each $t\geq 0$ can be characterized using the Karush-Kuhn-Tucker (KKT) conditions [20, §5]. Finally, Assumption 3 ensures that the system of equations $a_i(t)^\top x=b_i(t)$ $(i\in[q])$ is consistent and has infinitely many solutions at each $t\geq 0$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

It is worth remarking that we do not make any assumption about asymptotic vanishing time variations in the objective and constraints, namely, the partial derivates of these functions with respect to time are not assumed to converge to zero. However, we will assume that the optimal solution of the TV optimization problem does not vary exponentially fast as a function of time. We will explicitly impose this assumption in Section III-C.

III. Prediction-Correction Methods for Time-Varying Optimization

In this section, we consider the TV optimization problem in (1) without inequality constraints. In particular, we consider two versions of this problem: (*i*) the unconstrained case (Section III-A), and (*ii*) the case with linear equality constraints (Section III-B). We show that, in both cases, it is possible to track the optimal trajectory $x^\star(t)$ with an exponentially vanishing error. The algorithms developed here will be leveraged in Section III-C to derive a prediction-correction interior-point method to track the solution of (1) when inequality constraints are also considered.

A. Unconstrained Time-Varying Convex optimization

Consider the following unconstrained version of (1):

$$x^\star(t) := \underset{x\in\mathbb{R}^n}{\operatorname{argmin}} \; f_0(x,t). \tag{2}$$

Under sufficient regularity conditions, we could implement a descent method – in particular, Newton's method – on $f_0(x,t)$ that would rapidly converge to $x^\star(t)$. In the limit of infinitesimal steps, the sequence of iterations results in the following continuous-time dynamical system:

$$\dot{x}(t) = -\nabla_{xx}^{-1} f_0(x(t),t)\nabla_x f_0(x(t),t). \tag{3}$$

The trajectory $x(t)$ generated by (3) would approach a neighborhood around $x^\star(t)$, but it does *not* converge exactly to $x^\star(t)$, since the solution itself is changing over time. Observe that—with sufficient regularity—the optimal argument $x^\star(t)$ in (2) satisfies the first-order optimality condition $\nabla_x f_0(x^\star(t),t) = 0,\ \forall t \geq 0$. Since this latter condition is true for all times $t \geq 0$, its derivative must also be null, from where we obtain

$$\begin{aligned} 0 &= \dot{\nabla}_x f_0(x^\star(t),t) \\ &= \nabla_{xx} f_0(x^\star(t),t)\dot{x}^\star(t) + \nabla_{xt} f_0(x^\star(t),t), \end{aligned} \tag{4}$$

where $\dot{\nabla}_x f_0$ denotes the total derivative of $\nabla_x f_0$ with respect to t, while $\nabla_{xt} f_0$ denotes the partial derivative of the gradient $\nabla_x f_0$ with respect to t. Solving (4) for $\dot{x}^\star(t)$ yields the dynamical system

$$\dot{x}^\star(t) = -\nabla_{xx}^{-1} f_0(x^\star(t),t)\nabla_{xt} f_0(x^\star(t),t). \tag{5}$$

If the optimal solution $x^\star(t)$ was known for some $t_0 \geq 0$, the system in (5) could be used to track the evolution of $x^\star(t)$, since (5) guarantees that the optimality condition $\nabla_x f_0(x^\star(t),t) = 0$ is satisfied for all $t \geq t_0$. If we do not have access to $x^\star(t)$ at any point in time, we propose to combine the dynamics in (3) and (5) to build the following dynamical system:

$$\dot{x}(t) = h(x(t),t), \tag{6}$$

where the vector field $h(x,t)\colon \mathbb{R}^n \times \mathbb{R}_+ \to \mathbb{R}^n$ is defined as

$$h(x,t) = -\nabla_{xx}^{-1} f_0(x,t)\Big[P\nabla_x f_0(x,t) + \nabla_{xt} f_0(x,t)\Big]. \tag{7}$$

Here $P \in \mathbb{S}_{++}^n$ is a positive definite matrix satisfying $P \succeq \alpha I_n$ for some $\alpha > 0$. The dynamics in (6) contains two terms: (*i*) a *prediction* term $-\nabla_{xx}^{-1} f_0(x(t),t)\nabla_{xt} f_0(x(t),t)$ that attempts to track the changes in the objective function (see (5)), and (*ii*) a *correction* term $-\nabla_{xx}^{-1} f_0(x(t),t)P\nabla_x f_0(x(t),t)$ that 'pushes' $x(t)$ towards the optimum. Notice that the correction term corresponds to a gradient flow when $P = \nabla_{xx} f_0(x(t),t) \succeq m_f I_n$, and is a Newton flow when $P = \alpha I_n,\ \alpha > 0$. In the next proposition, we show that the dynamical system in (6) globally exponentially converges to the optimal trajectory $x^\star(t)$.

Proposition 1 *Let $x^\star(t)$ be defined as in* (2) *and $x(t)$ be the solution of* (6)*, where the objective function is m_f-strongly convex (Assumption 1). Then, the following inequalities hold:*

$$\|x(t) - x^\star(t)\|_2 \leq C_1 e^{-\alpha t}, \tag{8a}$$

$$0 \leq f_0(x(t),t) - f_0(x^\star(t),t) \leq m_f C_1^2 e^{-2\alpha t}, \tag{8b}$$

where $0 \leq C_1 := \dfrac{1}{m_f}\|\nabla_x f_0(x(0),0)\|_2 < \infty$.

Proof: See Appendix A. ■

Proposition 1 establishes that under the strong convexity assumption, the solution of (6) converges exponentially to a point that satisfies the first-order optimality condition $\nabla_x f_0(x^\star(t),t) = 0$. This exponential convergence is guaranteed in terms of the function value $f_0(x(t),t) - f_0(x^\star(t),t)$, as well as the distance to the optimal solution $\|x(t) - x^\star(t)\|_2$.

For algorithmic implementation, we can discretize the continuous-time dynamics (6) using, for example, Euler's forward method with a constant step size $\tau > 0$ [33, §1]. If the map $x \mapsto h(x,t)$ in (6) is uniformly Lipschitz on a region where the solution is contained, then the discretization error would be of the order $\mathcal{O}(\tau)$ [33, §1]. Alternatively, one could use line search to generate adaptive step sizes that guarantee a strict reduction in the suboptimality at each iteration [34].

Remark 1 (Role of α) Proposition 1 suggests that the continuous-time dynamics (6) can achieve an arbitrarily fast convergence rate by increasing α. However, it should be noted that when we discretize the continuous-time dynamics (6), using Euler's forward method with a constant sampling period $\tau > 0$, the coefficient α affects the maximum allowable step size that can guarantee stability and convergence of the discretized dynamics. In other words, the effective step size in the algorithmic implementation would be $\alpha\tau$ which has an upper bound for stability and convergence; see [34] for more details.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

B. Equality-Constrained Time-Varying Convex Optimization

We consider now a version of (1) in which we incorporate equality constraints:

$$x^\star(t) := \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\ f_0(x,t), \quad \text{s.t. } A(t)x = b(t), \tag{9}$$

where the matrix $A(t)\colon \mathbb{R}_+ \to \mathbb{R}^{q\times n}$ and vector $b(t)\colon \mathbb{R}_+ \to \mathbb{R}^q$ define q TV equality constraints. In order to design a dynamical system to track $x^\star(t)$ in (9), we introduce a Lagrange multiplier $\nu \in \mathbb{R}^q$ and define the Lagrangian associated with the optimization problem in (9) as

$$\mathcal{L}(x,\nu,t) := f_0(x,t) + \nu^\top(A(t)x - b(t)), \tag{10}$$

which is strongly convex in x (Assumption 1) and concave (affine) in ν. From the Lagrangian in (10), we define the TV dual function $\mathcal{G}(\nu,t) := \min_{x\in\mathbb{R}^n}\mathcal{L}(x,\nu,t)$ and the optimal dual argument as $\nu^*(t) := \operatorname{argmax}_{\nu\in\mathbb{R}^q}\ \mathcal{G}(\nu,t)$. By virtue of strong convexity (Assumption 1) and Assumption 3, the optimal primal-dual pair $(x^\star(t),\nu^\star(t))$ is unique at each time $t \geq 0$. Furthermore, we know this optimal pair must satisfy the following KKT conditions:

$$\begin{aligned} 0 &= \nabla_x\mathcal{L}(x^\star(t),\nu^\star(t),t) = \nabla_x f_0(x^\star(t),t) + A(t)^\top\nu^\star(t),\\ 0 &= \nabla_\nu\mathcal{L}(x^\star(t),\nu^\star(t),t) = A(t)x^\star(t) - b(t). \end{aligned} \tag{11}$$

We define the aggregate variable $z := [x^\top,\nu^\top]^\top \in \mathbb{R}^{n+q}$ and the optimal primal-dual solution $z^\star(t) := [x^\star(t)^\top,\nu^\star(t)^\top]^\top$ so as to rewrite (11) in the condensed form $0 = \nabla_z\mathcal{L}(x^\star(t),\nu^\star(t),t)$. Since this latter equation must hold for all $t \geq 0$, we can take the time derivative of both sides, which results in a prediction dynamics of the form $\dot z^\star(t) = -\nabla_{zz}^{-1}\mathcal{L}(z^\star(t),t)\nabla_{zt}\mathcal{L}(z^\star(t),t)$, resembling (5). We combine this prediction term with a Newton-like correction term of the form $-\nabla_{zz}^{-1}\mathcal{L}(z(t),t)P\nabla_z\mathcal{L}(z(t),t)$ to propose the following dynamical system:

$$\dot z(t) = -\nabla_{zz}^{-1}\mathcal{L}(z(t),t)\Big[P\nabla_z\mathcal{L}(z(t),t) + \nabla_{zt}\mathcal{L}(z(t),t)\Big], \tag{12}$$

$P \in \mathbb{S}_{++}^{n+q}$ is a symmetric positive definite matrix satisfying $P \succeq \alpha I_{n+q}$ for some $\alpha > 0$. In the following proposition, we prove that the state of this dynamical system globally exponentially converges to the optimal solution $z^\star(t)$, under appropriate assumptions.

Proposition 2 *Consider the optimization problem in* (9) *satisfying Assumptions 1 and 3. Denote* $z(t) = [x(t)^\top,\nu(t)^\top]^\top$ *as the solution of* (12)*, where* $\mathcal{L}(z,t) = \mathcal{L}(x,\nu,t)$ *is defined in* (10)*. Further, assume that the condition* $\|\nabla_{zz}^{-1}\mathcal{L}(z,t)\|_2 \leq M$ *holds for some* $M > 0$ *and all* $(z,t) \in \mathbb{R}^{n+q}\times\mathbb{R}_+$*. Then, the following inequalities hold:*

$$\|\nabla_z\mathcal{L}(z(t),t)\|_2 \leq \|\nabla_z\mathcal{L}(z(0),0)\|_2 e^{-\alpha t}, \tag{13a}$$

$$\|z(t) - z^\star(t)\|_2 \leq M\|\nabla_z\mathcal{L}(z(0),0)\|_2 e^{-\alpha t}, \tag{13b}$$

where $z(0) \in \mathbb{R}^{n+q}$ *is the initial condition.*

Proof: See Appendix B. ■

Proposition 2 establishes that the solutions of the dynamical system in (12) converge exponentially to the optimal trajectory from arbitrary (not necessarily feasible) initial conditions. For the special case that $P = \operatorname{diag}(P_x,P_\nu)$ where $P_x \in \mathbb{S}_{++}^n$ and $P_\nu \in \mathbb{S}_{++}^q$, the solution of (12) satisfies $\dot\nabla_\nu\mathcal{L}(z(t),t) = -P_\nu\nabla_\nu\mathcal{L}(z(t),t)$ (see (47) in Appendix B). Therefore, since $\nabla_\nu\mathcal{L}(x(t),t) = A(t)x(t) - b(t)$, we obtain:

$$\|A(t)x(t) - b(t)\|_2 \leq \|A(t)x(0) - b(0)\|_2 e^{-\lambda_{\min}(P_\nu)t}. \tag{14}$$

This implies that feasibility is achieved at an exponential rate. Further, if the initial condition $x(0)$ is feasible (i.e., $A(0)x(0) - b(0) = 0$), then the solution $x(t)$ will be feasible all the time, i.e., we have that $A(t)x(t) - b(t) = 0$ for all $t \geq 0$.

In the next subsection, we propose a solution to the most general time-varying optimization problem, where both equality and inequality constraints are considered.

C. General Time-Varying Convex Optimization

In this subsection, we consider another version of (1) where we include the inequality constraints only:

$$x^\star(t) := \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\ f_0(x,t), \quad \text{s.t. } f_i(x,t) \leq 0,\ i \in [p]. \tag{15}$$

In light of the analysis in Subsection III-B, we can always eliminate the equality constraints in (1) by Lagrangian relaxation. Therefore, we ignore equality constraints for now, without losing generality, and will remark on the addition of equality constraints at the end of this subsection. The Lagrangian associated with Problem (15) is

$$\mathcal{L}(x,\lambda,t) := f_0(x,t) + \sum_{i=1}^p \lambda_i f_i(x,t), \tag{16}$$

where $\lambda := [\lambda_1\cdots\lambda_p]^\top \in \mathbb{R}_+^p$ is the vector of Lagrange multipliers. Notice that $\mathcal{L}(x,\lambda)$ is twice differentiable, strongly convex in x, and concave (affine) in λ. The dual function is defined as $\mathcal{G}(\lambda,t) = \min_{x\in\mathbb{R}^n}\mathcal{L}(x,\lambda,t)$, and the corresponding optimal dual solution set is $\Gamma^\star(t) := \operatorname{argmax}_{\lambda\in\mathbb{R}^p}\mathcal{G}(\lambda,t)$. Under Assumptions 1 and 2, the necessary and sufficient KKT conditions [20, §5] for optimality of $x^\star(t)$ in (1) at each $t \geq 0$ read as

$$\begin{aligned} &\nabla_x f_0(x^\star(t),t) + \sum_{i=1}^p \lambda_i^\star\nabla_x f_i(x^\star(t),t) = 0;\\ &\lambda_i^\star f_i(x^\star(t),t) = 0,\ \lambda_i^\star \geq 0,\\ &f_i(x^\star(t),t) \leq 0,\ \text{for all } i \in [p], \end{aligned} \tag{17}$$

where $\lambda^\star = [\lambda_1^\star\cdots\lambda_p^\star]^\top \in \Gamma^\star(t)$. In what follows, we use barrier functions [20, §11] to incorporate the inequality constraints into the objective function. First, we consider the following equivalent representation of (15):

$$x^\star(t) = \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\ f_0(x,t) + \sum_{i=1}^p \mathbb{I}_-(f_i(x,t)), \tag{18}$$

where $\mathbb{I}_-:\mathbb{R}\to\{0,\infty\}$ is defined such that $\mathbb{I}_-(u) = 0$ if $u \leq 0$, and $\mathbb{I}_-(u) = \infty$ if $u > 0$. We now approximate $\mathbb{I}_-(u)$ by a smooth closed convex function of the form $-\frac{1}{c}\log(-u)$, where $c > 0$ is an arbitrary constant called the barrier parameter[1].

[1]Notice that $\lim_{c\to\infty} -\frac{1}{c}\log(-u) = \mathbb{I}_-(u)$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

Therefore, we can approximate the non-smooth problem (18) by the following smooth convex optimization problem:

$$\underset{x\in\mathcal{D}(t)}{\text{minimize}} \quad f_0(x,t) - \frac{1}{c(t)}\sum_{i=1}^{p}\log(-f_i(x,t)), \tag{19}$$

where $c\colon \mathbb{R}_+ \to \mathbb{R}_{++}$ is a time-dependent positive barrier parameter and the domain of the objective function is the open set $\mathcal{D}(t) := \{x \in \mathbb{R}^n\colon f_i(x,t) < 0,\ i \in [p]\}$. Our goal is to design a dynamical system able to track the optimal solution of (19). As we show below, this would require to initialize the dynamical system at a point inside $\mathcal{D}(0)$, i.e., $x(0) \in \mathcal{D}(0)$. To circumvent this restriction, we include a nonnegative time-dependent slack function $s\colon \mathbb{R}_+ \to \mathbb{R}_+$ in the optimization problem (19), and define the *approximate* optimal trajectory as

$$\widehat{x}^\star(t) := \underset{x\in\widehat{\mathcal{D}}(t)}{\operatorname{argmin}}\ f_0(x,t) - \frac{1}{c(t)}\sum_{i=1}^{p}\log(s(t) - f_i(x,t)), \tag{20}$$

where $\widehat{\mathcal{D}}(t) := \{x \in \mathbb{R}^n\colon f_i(x,t) < s(t),\ i \in [p]\}$ is an open set containing $\mathcal{D}(t)$. Notice that for any $x(0) \in \mathbb{R}^n$, we can choose $s(0) > \max_{i\in[p]} f_i(x(0),0)$ so that $x(0) \in \widehat{\mathcal{D}}(0)$, i.e., the initial condition lies in the "enlarged" feasible set $\widehat{\mathcal{D}}(0)$. In the next lemma, we characterize the approximation error in terms of $c(t)$, $s(t)$, and the optimal dual variables in (17).

Lemma 1 *Let $x^\star(t)$ and $\widehat{x}^\star(t)$ be defined as in* (18) *and* (20)*, respectively. Then, under Assumptions 1 and 2, and for any $\lambda^\star \in \Gamma^\star(t)$ and $t \geq 0$, the following inequality holds,*

$$|f_0(\widehat{x}^\star(t),t) - f_0(x^\star(t),t)| \leq \frac{p}{c(t)} + s(t)\Big(\inf_{\lambda^\star\in\Gamma^\star(t)}\|\lambda^\star\|_1\Big). \tag{21}$$

Proof: See Appendix C. ■

The above lemma suggests that if $s(t)$ and $c(t)$ are chosen such that the right-hand side of (21) converges to zero, the approximate solution $\widehat{x}^\star(t)$ converges to the optimal solution $x^\star(t)$ in (18). In what follows, we design a dynamical system whose solution globally asymptotically converges to $\widehat{x}^\star(t)$, defined in (20). Let us define $\Phi\colon \mathbb{R}^n \times \mathbb{R}_{++} \times \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}$ as

$$\Phi(x,c,s,t) := f_0(x,t) - \frac{1}{c}\sum_{i=1}^{p}\log\left(s - f_i(x,t)\right), \tag{22}$$

which is twice differentiable and strongly convex in x. The optimal solution $\widehat{x}^\star(t)$ in (20) satisfies the optimality condition $\nabla_x\Phi(\widehat{x}^\star(t), c(t), s(t), t) = 0$ for all $t \geq 0$. Our goal is to design a dynamical system for $x(t)$ such that the gradient $\nabla_x\Phi$ vanishes exponentially along the trajectory $(x(t), c(t), s(t), t)$. In particular, based on the results in Section III-A, we propose the following dynamical system:

$$\dot{x} = -\nabla_{xx}^{-1}\Phi\Big[\,P\nabla_x\Phi + \nabla_{xs}\Phi\ \dot{s} + \nabla_{xc}\Phi\ \dot{c} + \nabla_{xt}\Phi\,\Big], \tag{23}$$

where $P \in \mathbb{S}^n_{++}$ satisfies $P \succeq \alpha I_n$ for some $\alpha > 0$. The first term on the right-hand side corresponds to a Newton-like term, which is similar to the correction term in (6). The remaining terms in (23) play a role similar to the prediction term in (6), since they account for time variations in $\Phi(x,c,s,t)$ through s, c, and $f_i(x,t),\ i \in \{0\} \cup [p]$.

Notice that it is important for the dynamics in (23) to render a solution $x(t)$ such that the argument of the logarithmic barrier functions in (22) remains positive, i.e., we must have that $x(t) \in \widehat{\mathcal{D}}(t)$ for all $t \geq 0$. The next lemma states that this is indeed the case. Moreover, we show that that the solution to (23) globally exponentially converges to the approximate solution $\widehat{x}^\star(t)$ in (20).

Lemma 2 *Let $\widehat{x}^\star(t)$ be defined in* (20) *and $x(t)$ be the solution of* (23) *for $x(0) \in \widehat{\mathcal{D}}(0)$, $s(0) > \max_i f_i(x(0),0)$, and $c(0) > 0$. Then, under Assumptions 1 and 2, $x(t)$ satisfies $x(t) \in \widehat{\mathcal{D}}(t)$ for all $t \geq 0$. Moreover, the following inequality holds*

$$\|x(t) - \widehat{x}^\star(t)\|_2 \leq C_3 e^{-\alpha t},$$

where $0 \leq C_3 := \dfrac{1}{m_f}\|\nabla_x\Phi(x(0),c(0),s(0),0)\|_2 < \infty$.

Proof: See Appendix D. ■

Next, we need to establish the convergence of the approximate solution $\widehat{x}^\star(t)$ in (20) to the optimal solution $x^\star(t)$ in (18). Intuitively, we need the barrier parameter $c(t)$ to asymptotically diverge to infinity and the slack variable $s(t)$ to asymptotically vanish so that the approximation error vanishes, according to (21). For this to be true, we need to make the following assumption about the optimal dual variables defined in (17).

Assumption 4 For any $\gamma > 0$, the optimal dual variables satisfy $\lim_{t\to\infty}(\sup_{\lambda^\star\in\Gamma^\star(t)}\|\lambda^\star\|_1)\exp(-\gamma t) = 0$, for all $i \in [p]$.

The above assumption excludes the possibility for the optimal dual variables to escape to infinity exponentially fast. Roughly speaking, this condition requires the optimization problem to not have any exponentially growing function of t. By virtue of Assumption 4, the approximation error in (21) vanishes asymptotically if the slack variable $s(t)$ goes to zero exponentially fast and the barrier parameter $c(t)$ diverges to infinity. The next theorem states the main result of this subsection.

Theorem 1 *Consider the optimization problem in* (18) *and the objective function in* (22)*. Let $x(t)$ be the solution of* (23) *with initial conditions $x(0) \in \mathbb{R}^n$, $c(0) > 0$, and $s(0) > \max_{1\leq i\leq p} f_i(x(0),0)$. Let $\lim_{t\to\infty} c(t) = \infty$, and $s(t) = s(0)\exp(-\gamma_s t)$ for some $\gamma_s > 0$. Then, under Assumptions 1, 2, and 4, we have that*

$$\lim_{t\to\infty}\|x(t) - x^\star(t)\|_2 = 0.$$

Proof : By Assumption 4 and Lemma 1, we have that $\lim_{t\to\infty}|f_0(\widehat{x}^\star(t),t) - f_0(x^\star(t),t)| = 0$ when $s(t) = s(0)\exp(-\gamma_s t)$, and $\lim_{t\to\infty} c(t) = \infty$. Since $f_0(x,t)$ is strongly convex (Assumption 1) we obtain $\lim_{t\to\infty}\|\widehat{x}^\star(t) - x^\star(t)\|_2 = 0$. Finally, by Lemma 2 we can conclude $\lim_{t\to\infty}\|x(t) - x^\star(t)\|_2 = 0$ ■

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

According to Theorem 1, the continuous-time dynamics in (23) yields a solution that asymptotically converges to the optimal solution in (18) from arbitrary initial conditions. Some remarks are in order.

Remark 2 (Barrier Parameter) The barrier parameter $c(t)$ is required to be positive, monotonically increasing, asymptotically diverging to infinity, and bounded in finite time. A convenient choice is $c(t) = c(0)\exp(\gamma_c t)$ for $\gamma_c, c(0) > 0$.

Remark 3 (Adding Equality Constraints) As mentioned at the beginning of Section III-C, we ignored equality constraints in our analysis. In order to account for equality constraints, we define the Lagrangian function as

$$\mathcal{L}(x, \nu, t) = \Phi(x, c(t), s(t), t) + \nu^\top (A(t)x - b(t)), \quad (24)$$

where $\Phi(x, c, s, t)$ is defined in (22). Since $f_0(x, t)$ is strongly convex, the map $x \mapsto \Phi(x, c, s, t)$ is also strongly convex. Furthermore, assume that the following condition holds,

$$\left\| \begin{bmatrix} \nabla_{xx}\Phi(x, c, s, t) & A^\top(t) \\ A(t) & 0_{q\times q} \end{bmatrix}^{-1} \right\|_2 \leq N, \text{ for all } t \geq 0,$$

for some $0 < N < \infty$ and all $(z, t) \in \mathbb{R}^{n+q} \times \mathbb{R}_+$. Then, by defining the aggregate vector of decision variables $z = [x^\top \ \nu^\top]^\top \in \mathbb{R}^{n+q}$, and applying the dynamics (12) on the Lagrangian (24), we can use the same arguments as in Proposition 2 and Theorem 1 to conclude that $\lim_{t\to\infty} \|z(t) - z^\star(t)\|_2 = 0$.

Remark 4 (Variable Number of Constraints) In practical applications with a variable number of constraints, a restarting scheme can be used in which the barrier function (22) and the prediction-correction dynamics (23) are updated so as to add or remove constraints. In this case, the slack variable and the barrier parameter must be reset so as to ensure feasibility and asymptotic convergence.

Remark 5 (Second-Order Dynamics) The derivatives of logarithmic barrier functions are singular on the boundary of the feasible set. This may induce numerical instability in the discrete-time implementation of the dynamical system (23). To avoid this issue, we replace the first-order dynamics in (23) by the following second-order dynamics [16]:

$$\begin{aligned} \dot{x} &= -\nabla_{xx}^{-1}\Phi\Big[\, \alpha y + \nabla_{xs}\Phi \ \dot{s} + \nabla_{xc}\Phi \ \dot{c} + \nabla_{xt}\Phi \,\Big], \\ \dot{y} &= -\gamma y + \alpha \nabla_x \Phi, \end{aligned} \quad (25)$$

where $\gamma > 0$ is an arbitrary constant. Intuitively, the gradient function is passed through a first-order low-pass filter (the second ODE in (25)) whose output is then fed into the main dynamics. The resulting dynamics tends to reduce numerical instability induced by discretization. It is shown in [16] that $V = \frac{1}{2}\nabla_x\Phi^\top\nabla_x\Phi + \frac{1}{2}y^\top y$ is a Lyapunov function for (25), establishing that $\lim_{t\to\infty} \|\nabla_x \Phi\| = 0$.

D. Online Implementation

The dynamical system proposed in (23) includes the prediction term $\nabla_{xt}\Phi(x, t)$, whose computation consists in evaluating the terms $\{\nabla_{xt} f_i(x, t)\}_{i=0}^p$ and $\{\frac{\partial}{\partial t} f_i(x, t)\}_{i=1}^p$. In an online setting, we might only have access to limited or noisy information about these terms. More precisely, assume that we have access to an estimate of $\nabla_{xt}\Phi$ denoted by $\widehat{\nabla}_{xt}\Phi$ that satisfies the bound

$$\|\widehat{\nabla}_{xt}\Phi(x, t) - \nabla_{xt}\Phi(x, t)\|_2 \leq \eta, \quad (26)$$

for some known $\eta > 0$ and all $(x, t) \in \widehat{\mathcal{D}}(t) \times \mathbb{R}_+$. In this setting, we consider the following dynamics:

$$\dot{x} = -\nabla_{xx}^{-1}\Phi\Big[\, \alpha\nabla_x\Phi + \nabla_{xs}\Phi \ \dot{s} + \nabla_{xc}\Phi \ \dot{c} + \widehat{\nabla}_{xt}\Phi \,\Big], \quad (27)$$

where we consider a state-dependent α as follows:

$$\alpha = \frac{\alpha_0}{\max(\|\nabla_x\Phi\|_2, \varepsilon)}, \quad (28)$$

where $\alpha_0 > \eta$, and $\varepsilon > 0$ are arbitrary constants. The next theorem states that the solution of (27) converges to an ε-neighborhood of the approximate optimal solution $\widehat{x}^\star(t)$, defined in (20), in finite time and will stay there forever.

Theorem 2 *Denote $x(t)$ as the solution of* (27) *where Φ is defined in* (22)*. Assume $\widehat{\nabla}_{xt}\Phi$ satisfies the bound in* (26)*, and the coefficient α is defined in* (28) *with $\alpha_0 > \eta$ and $\varepsilon > 0$. Then, under Assumptions 1 and 2, the solution $x(t)$ satisfies*

$$\lim_{t\to\infty} \|x(t) - x^\star(t)\|_2 \leq \frac{\eta\varepsilon}{\alpha_0 m_f}.$$

Proof: See Appendix E. ■

E. Prediction-Correction in Time-Invariant Interior-Point Method

As a particular application of our results, we consider the following time-invariant convex optimization problem,

$$x^\star := \arg\min \ f_0(x) \ \text{ s.t. } f_i(x) \leq 0, i \in [p]. \quad (29)$$

Using logarithmic barrier functions to relax the constraints, we define

$$\Phi(x, c) = f_0(x) - \frac{1}{c}\sum_{i=1}^p \log(-f_i(x)), \quad (30)$$

and the corresponding central path

$$x^\star(c) := \arg\min_{x\in\mathcal{D}} \Phi(x, c), \quad (31)$$

where $\mathcal{D} = \{x \in \mathbb{R}^n \colon f_i(x) < 0, i \in [p]\}$ is the *interior* of the feasible set. It follows from Lemma 1 that $\lim_{c\to\infty} \|x^\star(c) - x^\star\| = 0$. In standard implementation of the interior-point method [20, §11], the optimization problem (31) is solved sequentially for a positive growing sequence $(c_k)_{k=1}^\infty$, each starting from the optimal solution of the previous optimization problem. The resulting sequence $(x^\star(c_k))_{k=1}^\infty$ converges to the optimal point $x^\star$ as $c_k \to \infty$. For each fixed c_k, $x^\star(c_k)$ can be found, for

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

instance, using the following Newton dynamics (or its discrete-time equivalent):

$$\dot{x}_k(t) = -\nabla_{xx}^{-1}\Phi(x_k(t), c_k)\nabla_x\Phi(x_k(t), c_k). \tag{32}$$

According to Theorem 1, we have that $\lim_{t\to\infty} \|x_k(t) - x^\star(c_k)\| = 0$. Thus, by choosing the initial points as $x_{k+1}(0) = \lim_{t\to\infty} x_k(t) = x^\star(c_k)$ and let $\lim_{k\to\infty} c_k = \infty$, we can build a continuous path that converges to $x^\star$ in (29). As a less computationally expensive alternative, we propose to consider an increasing time-dependent barrier parameter $c(t)$, in lieu of discontinuous jumps. In this case, the problem in (31) renders a time-varying objective function. The prediction-correction dynamics (23) using the barrier function in (30) yields the ODE

$$\dot{x} = -\nabla_{xx}^{-1}\Phi(x, c)\Big[\alpha\nabla_x\Phi(x, c) + \nabla_{xc}\Phi(x, c)\,\dot{c}\Big],$$

whose solution $x(t)$ satisfies $\lim_{t\to\infty} \|x(t) - x^\star\|_2 = 0$ when $\lim_{t\to\infty} c(t) = \infty$. We will numerically illustrate the performance of this approach in Subsection IV-C.

IV. Numerical Experiments

In this section, we provide four numerical examples to illustrate the time-varying optimization framework herein proposed. In Subsection IV-A, we consider a TV optimization problem with equality constraints. In Subsection IV-B, we solve a TV optimization problem with inequality constraints to illustrate the effectiveness of the prediction-correction interior-point method in solving inequality-constrained problems. In Subsection IV-C, we use the prediction-correction framework discussed in Subsection III-E to solve a time-invariant large-scale ℓ_1-regularized least-squares problem. In Subsection IV-D, we solve a navigation problem to drive a disk-shaped robot towards a potentially moving desired location without colliding with obstacles in the environment.

A. Equality Constraints

Consider the following quadratic optimization problem with equality constraints:

$$\begin{aligned} \text{minimize} \quad & \frac{1}{2}\|x - r(t)\|_2^2 + \mu\exp(1^\top x), \\ \text{s.t.} \quad & \cos(\omega t)x_1 + \sin(\omega t)x_2 + x_3 = 1, \end{aligned} \tag{33}$$

where $\mu = \omega = 0.1$, and $r(t) = [\sin(\omega t)\ \ \cos(\omega t)\ \ 1]^\top$. This problem is an instance of (9) with $n = 3$ primal variables and $q = 1$ equality constraints. We consider the Lagrangian (10) associated with (33). To track the optimal trajectory, we implement the ODE (12) choosing $\alpha = 1$, and using Euler's forward discretization with a constant sampling period $\Delta t > 0$. In Figure 1, we plot $\|\nabla_z\mathcal{L}(z(t), t)\|_2$ versus t for various sampling periods $\Delta t \in \{0.01, 0.1, 0.3\}$. We observe that larger sampling periods increase the steady-state tracking error. In Figure 2, we fix the sampling period to $\Delta t = 0.1$ and plot $\|\nabla_z\mathcal{L}(z(t), t)\|_2$ versus t for $\alpha \in \{0.1, 1, 5, 10\}$. We observe that higher values of α speed up the convergence. However, too large values could lead to poor performance and instability (specifically $\alpha > 12$ in our example)–see Remark 1.

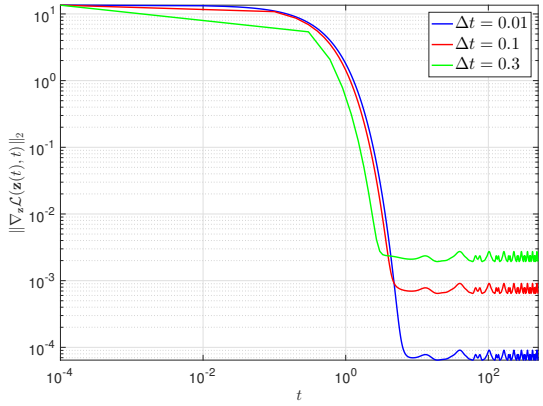

Fig. 1: Plot of $\|\nabla_z\mathcal{L}(z(t), t)\|_2$ where $z(t)$ is the solution to the discretized version of (12) for the numerical example in (33). Different paths correspond to different sampling periods Δt. Less frequent sampling leads to larger steady-state errors.

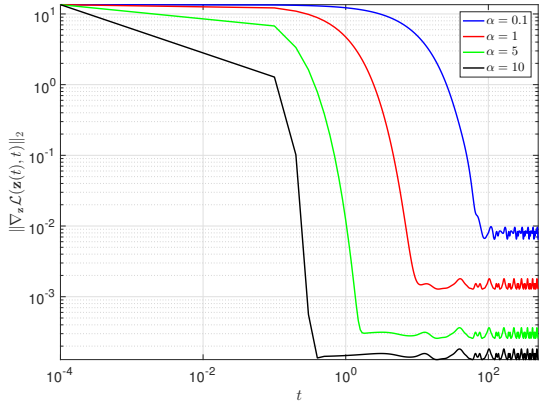

Fig. 2: Plot of $\|\nabla_z\mathcal{L}(z(t), t)\|_2$ where $z(t)$ is the solution to the discretized version of (12) for the numerical example in (33). The sampling time is $\Delta t = 0.1$, and the curves correspond to various values of α. Typically, higher values of α increase the convergence rate, but too large values would eventually destabilize the solution (not shown in the figure).

B. Inequality Constraints

Consider the following TV quadratic optimization problem:

$$\begin{aligned} x^\star(t) := \ & \operatorname{argmin} \ \frac{1}{2}(x_1 + \sin(t))^2 + \frac{3}{2}(x_2 + \cos(t))^2, \\ & \text{s.t.} \quad x_2 - x_1 - \cos(t) \le 0. \end{aligned} \tag{34}$$

In the following simulation, we show how to track $x^\star(t)$ using the continuous-time dynamics in (23). In order to illustrate the usage of the time-dependent slack variable $s(t)$, we choose the initial condition $x(0) = (-2, 0)^\top$, which is infeasible at $t = 0$. As discussed in Subsection III-C, we include the slack variable to enlarge the feasible set. In this example, the augmented objective function in (22) takes the form:

$$\begin{aligned} \Phi(x, s, c, t) = & \frac{1}{2}\Big(x_1 + \sin(t)\Big)^2 + \frac{3}{2}\Big(x_2 + \cos(t)\Big)^2 \\ & - \frac{1}{c}\log\Big(s + \cos(t) + x_1 - x_2\Big). \end{aligned}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

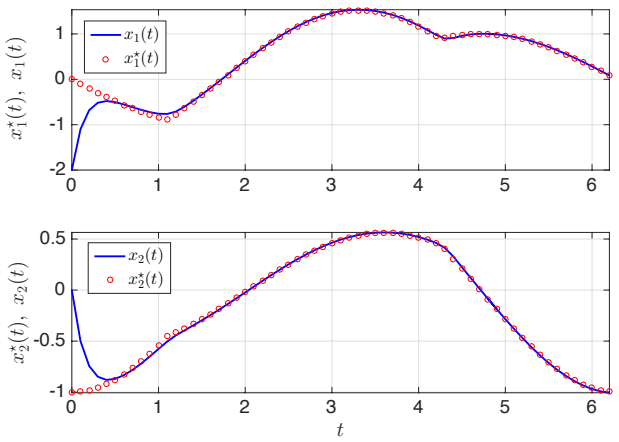

Fig. 3: Plot of the coordinates of the optimal trajectory $x^\star(t)$, defined in (34), and the tracking trajectory $x(t)$.

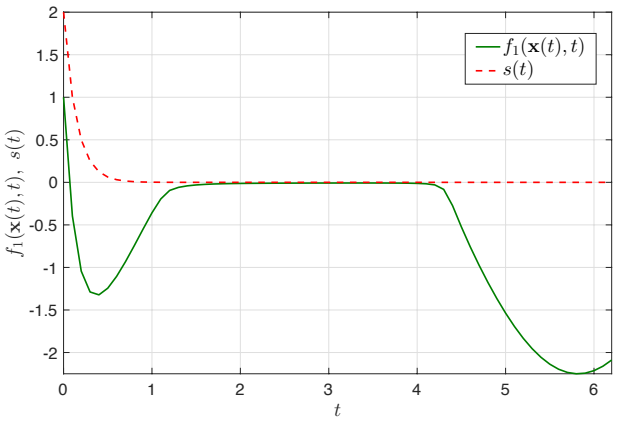

Fig. 4: Plot of the constraint function $f_1(x(t),t) = x_2(t) - x_1(t) - \sin(t)$ and the slack variable $s(t)$ against t, for the problem (34).

In our simulations, we consider the following time-dependent barrier parameter and slack variable: $c(t) = 10e^t$ and $s(t) = 2e^{-5t}$. The initial value of the slack variable is $s(0) = 2$; hence, $x(0)$ is initially feasible with respect to the enlarged feasible set. Using these particular values, all the conditions of Theorem 1 are satisfied. We numerically solve the ODE in (23) for the time interval $t \in [0, 2\pi]$ using Euler's forward method with step size $\tau = 0.1$. In Fig. 3, we plot the trajectory of the resulting solution $x(t) = (x_1(t), x_2(t))^\top$ along with the optimal solution $x^\star(t) = (x_1^\star(t), x_2^\star(t))$ defined in (34). In Fig. 4 we plot the time evolution of the constraint function $f_1(x,t) := x_2(t) - x_1(t) - \cos(t)$, as well as the slack variable $s(t)$. Notice that at $t = 0$ the state $x(t)$ violates the constraint $f_1(x,t) \leq 0$. However, $x(t)$ converges to the feasible set exponentially fast as the slack variable $s(t)$ vanishes exponentially.

C. ℓ_1-Regularized Least Squares

In this subsection, we illustrate how to leverage prediction-correction in time-invariant problems (described in Subsection III-E) to solve the following (time-invariant) ℓ_1-regularized least-squares problem:

$$x^\star(\lambda) = \underset{x\in\mathbb{R}^n}{\operatorname{argmin}}\ \|Ax - b\|_2^2 + \lambda\|x\|_1, \tag{35}$$

where $A \in \mathbb{R}^{m\times n}$ and $b \in \mathbb{R}^m$ are given, and $\lambda > 0$ is a constant regularizer which is commonly used to prevent over-fitting whenever $m < n$. Since the objective function in (35) is not differentiable, we analyze the following (differentiable) equivalent convex program [35]:

$$\begin{aligned}(x^\star(\lambda), u^\star(\lambda)) = \underset{x,u\in\mathbb{R}^n}{\operatorname{argmin}} \quad & \|Ax - b\|_2^2 + \lambda\sum_{i=1}^n u_i, \\ \text{s.t.} \quad & -u_i \leq x_i \leq u_i,\ i \in [n]. \end{aligned} \tag{36}$$

In the following numerical experiment, we generate a sparse vector $x^\star \in \mathbb{R}^{2048}$ with 10 entries of value ± 1, and all other entries equal to zero. The entries of the measurement matrix $A \in \mathbb{R}^{256\times 1024}$ are independently generated according to the standard normal density. The measured vector $b \in \mathbb{R}^{256}$ is generated by $b = Ax^\star + v$ where v is a contaminating noise drawn from the Gaussian distribution $\mathcal{N}(0_{256}, 0.01\, I_{256})$. The regularizer parameter is chosen to be $\lambda = 2$. For these numerical values, we solve (36) using three methods: (*i*) the Standard Newton Interior-Point Method (SNIPM) [20, Chapter 11] where the central points $x^\star(c)$ are computed using Newton's method with the sequence $c_k = 10 \times 5^k,\ k \geq 0$; (*ii*) the Prediction-Correction Newton Interior-Point Method (PCNIPM), described in Subsection III-E, where the barrier parameter is equal to $c(t) = 10e^t$; and (*iii*) the Truncated Newton Interior-Point Method (TNIPM), described in [35], where a preconditioned conjugate gradient method was proposed to compute the Newton step, and the barrier parameter is updated at each iteration. For all these three methods, we use a backtracking line search to adaptively select the step size. To assess the progress of the algorithms, we use the following quantity (as proposed in [35]),

$$\frac{\eta}{g(\nu)} := \frac{\|Ax - b\|_2^2 + \lambda\|x\|_1 - g(\nu)}{g(\nu)}. \tag{37}$$

Here, $g(\nu)$ is the dual function of the constrained problem akin to (35):

$$\min\ z^\top z + \lambda\|x\|_1\ \text{ s.t. } z = Ax - b, \tag{38}$$

and $\nu \in \mathbb{R}^m$ is the dual vector associated with the constraint $z = Ax - b$. The quantity in (37) is an upper bound of the relative duality gap $[p^\star - g(\nu)]/g(\nu)$, where $p^\star = \|Ax^\star(\lambda) - b\|_2^2 + \lambda\|x^\star(\lambda)\|_1$ is the primal optimal value (see [35] for more details). Fig. 5 illustrates the evolution of $\eta/g(\nu)$ against the iteration number for these three algorithms. For the stopping criterion, we choose $\eta/g(\nu) \leq 10^{-4}$. In our simulations, the SNIPM takes 42 iterations, while the PCNIPM proposed in Subsection III-E takes 17 iterations. Notice that the performance of the prediction-correction method is comparable to TNIPM, since in the latter method the barrier parameter is also updated at each iteration, but the prediction term is not included.

D. Robot Navigation

In this subsection, we solve the navigation problem of driving a disk-shaped robot of radius $r > 0$ to a given configuration x_d without colliding with obstacles in the environment. More precisely, let us consider a closed and convex

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

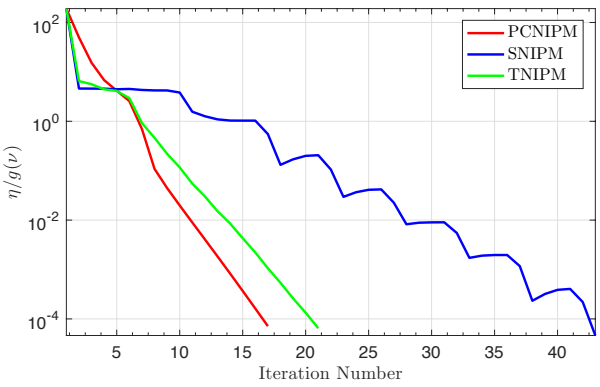

Fig. 5: Plot of the upper bound $\eta/g(\nu)$ on the relative suboptimality against the number of iterations for the three algorithms described in Subsection IV-C.

workspace $\mathcal{W} \subset \mathbb{R}^n$ of possible configurations that the robot can take. Assume that the workspace is populated with m non-intersecting spherical obstacles, where the center and radius of the i-th obstacle are denoted by $x_i \in \mathcal{W}$ and $r_i > 0$, respectively. We define the *free space*, denoted by $\mathcal{F}$, as the set of configurations in the workspace in which the robot does not collide with any of the obstacles. More formally,

$$\mathcal{F} = \left\{x \in \mathcal{W} \colon \overline{B}(x, r) \subseteq \mathcal{W} \setminus \cup_{i=1}^{m} B(x_i, r_i)\right\}, \tag{39}$$

where $B(x, r)$ is the n-dimensional open ball centered at x with radius r, and $\overline{B}(x, r)$ represents its closure.

Let us denote the center of mass of the robot by x_c. Given a final desired configuration $x_d \in \mathcal{F}$, the navigation problem under consideration consists of finding a trajectory of x_c such that $x_c(t) \in \mathcal{F}$ for all $t \geq 0$, and $\lim_{t\to\infty} x_c(t) = x_d$. In [36], the authors proposed a solution to this problem using the idea of *projected goal*, as described below. This idea consists of continuously computing the projection of the destination x_d onto a neighborhood around the center of mass of the robot in which there are no obstacles. Denote this projection by $\bar{x}_d$ —yet to be properly defined— then, the control law $\dot{x}_c = -(x_c - \bar{x}_d)$ ensures convergence of the center of mass of the robot to the desired configuration while avoiding the obstacles [36]. As we describe below, this technique can be interpreted as the solution of a TV convex optimization problem. To formulate this problem, we first need to provide some definitions.

We define first the notion of *power distance* between a point x and a disk $B(x_i, r_i)$ as $\mathcal{P}(x, B(x_i, r_i)) = \|x - x_i\|_2^2 - r_i^2$, [37]. We define the so-called *local workspace* around x_c as

$$\mathcal{LW}(x_c) = \{x \in \mathcal{W} \colon \mathcal{P}(x, B(x_c, r)) \leq \mathcal{P}(x, B(x_i, r_i)),\ \forall i\},$$

i.e., the set of points in $\mathcal{W}$ that are closer (in power distance) to the robot than to any of the obstacles. The local workspace defines a polytope whose boundaries are hyperplanes, such as the polygon marked with a thick light blue line in Fig. 6 (see [36, Eq. (6)] for an explicit expression of these hyperplanes). Furthermore, the *collision-free local workspace* around x_c is defined as [36]:

$$\mathcal{LF}(x_c) = \left\{x \in \mathcal{W} \colon a_i(x_c)^\top x - b_i(x_c) \leq 0,\ i = 1 \dots m\right\},$$

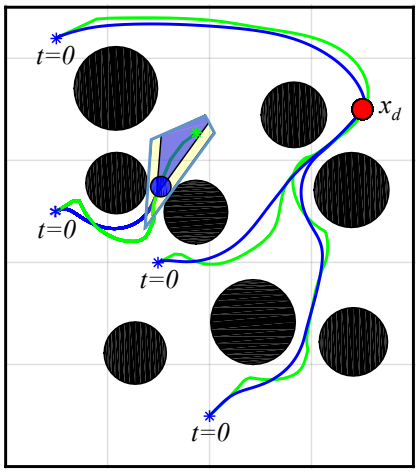

Fig. 6: The red circle represents the desired configuration x_d. The green and blue lines represent, respectively, the trajectories of the estimates of the projected goal $\widehat{x}(t)$ and the trajectories of the robot $x_c(t)$ for 4 different initial conditions.

where,

$$a_i(x_c) = x_i - x_c, \quad \theta_i(x_c) = \frac{1}{2} - \frac{r_i^2 - r^2}{2\|x_i - x_c\|^2}, \tag{40}$$
$$b_i(x_c) = (x_i - x_c)^\top \left(\theta_i x_i + (1 - \theta_i)x_c + r\frac{x_c - x_i}{\|x_c - x_i\|}\right).$$

Assuming that the robot follows the integrator dynamics $\dot{x}_c = u(x_c)$, the controller proposed in [36] is given by

$$\dot{x}_c = -K(x_c - x^\star), \tag{41}$$

where $K > 0$ is the gain of the controller and $x^\star$ is the orthogonal projection of the desired configuration x_d onto the collision-free local workspace $\mathcal{LF}(x_c)$. Under the assumption that the distance between the center of any two obstacles i and j is larger than $r_i + r_j + 2r$, it can be shown that the controller law in (41) solves the navigation problem ([36, Theorem 1]). In what follows, we cast the navigation problem as a TV convex optimization program that can be solved using the tools developed in this paper.

1) Interior-Point Method to Estimate the Projected Goal: We now show that the prediction-correction interior-point method developed in Section III can be used to efficiently build an estimator $\widehat{x}$ of the projection of x_d onto the set $\mathcal{LF}(x_c)$, which we denote by $x^\star$. First, observe that $x^\star$ can be defined as the solution of the following convex optimization problem,

$$x^\star := \underset{x\in\mathbb{R}^n}{\operatorname{argmin}} \frac{1}{2}\|x - x_d\|^2$$
$$\text{s.t.} \quad a_i(x_c)^\top x - b_i(x_c) \leq 0, \quad i = 1 \dots m. \tag{42}$$

Observe that since $a_i(x_c)$ and $b_i(x_c)$ depend on the position of the center of mass of the robot, the above optimization problem has an implicit dependence on time through x_c. We estimate the projected goal $x^\star$ as the solution to the ODE in (23) with initial condition $x(0) = x_c(0)$, i.e., the initial position of the robot, and the following objective function:

$$\Phi(x, x_c, t) = \frac{1}{2}\|x - x_d\|^2 - \frac{1}{c(t)}\sum_{i=1}^{m} \log(b_i(x_c) - a_i(x_c)^\top x).$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

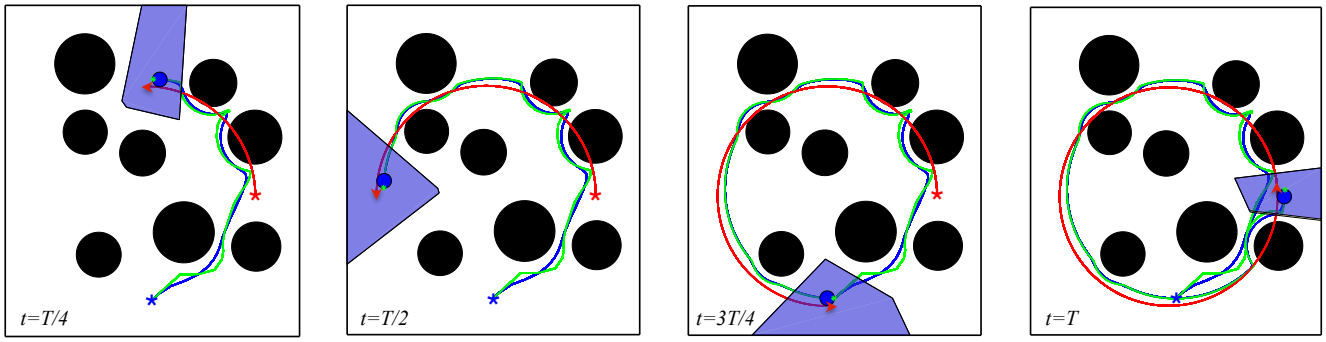

Fig. 7: Trajectory of the estimate of the projected goal $\widehat{x}$ (the green line, starting at the blue star) and trajectory of the robot (the blue line, starting at the blue star) tracking a moving target following a circular trajectory (red line, starting at the red star). The subplots correspond to time instances $t = T/4, T/2, 3T/4$, and T. The local workspace for each time instance is represented by a solid purple polygon containing the disk-shaped robot (the blue circle).

In Appendix F, we derive explicit expressions for all the terms involved in this ODE.

Next, we consider the control law (41), with the difference that we use an estimate of the projected goal instead of the projected goal itself, i.e., we consider the closed loop dynamics

$$\dot{x}_c = -K(x_c - \widehat{x}), \tag{43}$$

where the estimator $\widehat{x}(t)$ is the solution to the ODE in (23) with initial condition $\widehat{x}(0) = x_c(0)$. An important feature of $\widehat{x}(t)$ is that it is feasible at all times, i.e., $\widehat{x}(t) \in \mathcal{LF}(x_c)$. This follows from Lemma 2 and the fact that the initial condition is assumed to be feasible, i.e., the robot is initially located in the free space. Moreover, the estimator $\widehat{x}(t)$ converges exponentially to the projection of x_d onto the collision-free local workspace, denoted by $x^\star$.

To evaluate the performance of the proposed controller and optimizer, we consider a square workspace $\mathcal{W} = [-20, 20]^2$ containing 8 circular obstacles (black circles in Fig. 6). In Fig. 6, we also depict the trajectories followed by a disc-shaped robot of radius equal to one (blue circle) for four different initial conditions. The green and blue lines represent, respectively, the trajectories of the estimates $\widehat{x}(t)$ of the projected goal onto the collision-free local workspace, and the trajectories of the center of mass of the robot $x_c(t)$ for 4 different initial conditions. The blue circle represents a particular configuration of the robot, where the local workspace $\mathcal{LW}(x)$ (resp., the collision-free local workspace $\mathcal{LF}(x_c)$) is the polygon enclosed within light blue lines (resp., the polygon filled in solid purple). For these particular realizations, we have set $\alpha = 5$ in (23), and $K = 0.01$ in (43). Finally, the barrier parameter in (22) is chosen to be $c(t) = e^{0.001t}$. In Fig. 6, we observe how the robot succeeds in converging to the desired destination. Collision avoidance is ensured due to the fact that the estimate of the projected goal $\widehat{x}$ remains always in the collision-free local workspace $\mathcal{LF}(x_c)$.

2) Moving Targets: In our final experiment, we consider a similar navigation problem, but in this case the disk-shaped robot must track a moving target, i.e., $x_d(t) : \mathbb{R}_+ \to \mathcal{W}$. In this scenario, no theoretical guarantees are provided by the technique proposed in [36]; however, we demonstrate how our approach can be used to track a moving target.

In our experiment, we consider a moving target $x_d(t)$ following a circumference of radius 15, centered at the origin, and moving periodically with a period $T = 2 \times 10^3 s$. Observe that the target trajectory (red line in Fig. 7) is allowed to intersect the circular obstacles (e.g., obstacles are on the ground, and the target is flying). To track this target, we use the controller (43) where $\widehat{x}$ is the solution to the ODE in (23) with the following barrier function:

$$\Phi(x, x_c, t) = \frac{1}{2}\|x - x_d(t)\|^2 - \frac{1}{c(t)} \sum_{i=1}^{m} \log(b_i(x_c) - a_i(x_c)^\top x),$$

where $a_i(x_c) = x_i - x_c$ and $b_i(x_c)$ are given by (40). The parameter selection for our simulation is $K = 0.05$, $\alpha = 30$, and $c(t) = 100e^{\alpha t}$ with $\alpha = 0.001$. In Fig. 7, we depict the trajectory followed by the disk-shaped robot (blue circle) over time. As we can observe, the robot succeeds in tracking the moving target while avoiding the circular obstacles.

V. Conclusions

In this paper, we have developed a prediction-correction scheme for solving convex optimization problems with time-varying objective and constraint functions. Using log-barrier penalty functions, we have proposed a continuous-time dynamical system for tracking the solution of the time-varying problem. This dynamical system consists of a correction term, which is a continuous-time implementation of Newton's method, as well as a prediction term that takes into account the time-varying nature of the objective and constraint functions. Under reasonable assumptions, our method globally asymptotically converges to the time-varying optimal solution of the original problem with a vanishing tracking error. We have illustrated the applicability of the proposed method in two practical applications: a sparsity promoting least squares problem and a collision-free robot navigation problem.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

APPENDIX

A. Proof of Proposition 1

By strong convexity of $f_0(x,t)$ (Assumption 1), the Hessian inverse $\nabla_{xx}^{-1} f_0(x(t),t)$ is defined and uniformly bounded for all $t \geq 0$. The time derivative of the gradient at $(x(t),t)$ can be written as

$$\dot{\nabla}_x f_0(x(t),t) = \nabla_{xx} f_0(x(t),t)\dot{x}(t) + \nabla_{xt} f_0(x(t),t).$$

Substituting $\dot{x}(t)$ from (6), we obtain the ODE $\dot{\nabla}_x f_0(x(t),t) = -P\nabla_x f_0(x(t),t)$ whose solution is

$$\nabla_x f_0(x(t),t) = e^{-Pt}\nabla_x f_0(x(0),0),$$

where $x(0) \in \mathbb{R}^n$ is the initial point. This implies that

$$\begin{aligned}\|\nabla_x f_0(x(t),t)\|_2 &\leq \|e^{-Pt}\|_2 \|\nabla_x f_0(x(0),0)\|_2, \\ &\leq e^{-\alpha t}\|\nabla_x f_0(x(0),0)\|_2. \end{aligned} \tag{44}$$

The second inequality follows from the fact that $P \succeq \alpha \mathrm{I}_n$. Next, we use the mean-value theorem to expand $\nabla_x f_0(x(t),t)$ with respect to its first argument around $x^\star(t)$ satisfying $\nabla_x f_0(x^\star(t),t) = 0$ as follows,

$$\nabla_x f_0(x(t),t) = \nabla_{xx} f_0(\eta(t),t)(x(t) - x^\star(t)),$$

where $\eta(t)$ is a convex combination of $x(t)$ and $x^\star(t)$. It follows from uniform strong convexity of $f_0(x,t)$ (Assumption 1) that $\|\nabla_{xx}^{-1} f_0(x(t),t)\|_2 \leq m_f^{-1}$. By the Cauchy-Shwartz inequality, we can write

$$\begin{aligned}\|x(t) - x^\star(t)\|_2 &= \|\nabla_{xx}^{-1} f_0(\eta(t),t)\nabla_x f_0(x(t),t)\|_2 \\ &\leq \frac{1}{m_f}\|\nabla_x f_0(x(t),t)\|_2. \end{aligned} \tag{45}$$

On the other hand, convexity of $f_0(x,t)$ implies that for each $t \geq 0$,

$$0 \leq f_0(x(t),t) - f_0(x^\star(t),t) \leq \nabla_x f_0(x(t),t)^\top (x(t) - x^\star(t)).$$

By applying the Cauchy-Swhartz inequality on the right-hand side of the above inequality and invoking (45), we obtain

$$0 \leq f_0(x(t),t) - f_0(x^\star(t),t) \leq \frac{1}{m_f}\|\nabla_x f_0(x(t),t)\|_2^2. \tag{46}$$

By substituting (44) in the right-hand side of (45) and (46), the proof becomes complete.

B. Proof of Proposition 2

The Hessian of the Lagrangian in (10) with respect to $z = [x^\top \nu^\top]^\top$ is given by

$$\nabla_{zz}\mathcal{L}(z,t) = \begin{bmatrix} \nabla_{xx} f_0(x,t) & A(t)^\top \\ A(t) & 0_{q\times q} \end{bmatrix}.$$

The strong convexity of $f_0(x,t)$ and the full-rank condition $\mathrm{rank}(A(t)) = q < n$ is sufficient for $\nabla_{zz}\mathcal{L}(z,t)$ to be invertible [20, §10]. Therefore, the Hessian inverse $\nabla_{zz}^{-1}\mathcal{L}(z,t)$ in (12) exists. The time evolution of $\nabla_z \mathcal{L}(z(t),t)$ can be written as

$$\dot{\nabla}_z \mathcal{L}(z(t),t) = \nabla_{zz}\mathcal{L}(z(t),t)\dot{z}(t) + \nabla_{zt}\mathcal{L}(z(t),t).$$

Substituting $\dot{z}(t)$ from (12), we obtain the ODE

$$\dot{\nabla}_z \mathcal{L}(z(t),t) = -P\nabla_z \mathcal{L}(z(t),t), \tag{47}$$

which yields the solution $\nabla_z \mathcal{L}(z(t),t) = e^{-Pt}\nabla_z \mathcal{L}(z(0),0)$ for the initial condition $z(0) \in \mathbb{R}^{n+q}$. Since $P \succeq \alpha \mathrm{I}_n$, we obtain

$$\|\nabla_z \mathcal{L}(z(t),t)\|_2 \leq \|\nabla_z \mathcal{L}(z(0),0)\|_2 e^{-\alpha t}.$$

Next, we apply the mean-value theorem to expand $\nabla_z \mathcal{L}(z(t),t)$ with respect to its first argument around the optimal solution $z^\star(t)$ satisfying $\nabla_z \mathcal{L}(z^\star(t),t) = 0$ (see (11)) as follows,

$$\nabla_z \mathcal{L}(z(t),t) = \nabla_{zz}\mathcal{L}(\eta(t),t)(z(t) - z^\star(t)),$$

where $\eta(t)$ is a convex combination of $z(t)$ and $z^\star(t)$. By assumption, we have that $\|\nabla_{zz}^{-1}\mathcal{L}(z,t)\|_2 \leq M$ for $(z,t) \in \mathbb{R}^{n+q} \times \mathbb{R}_+$. Therefore, we obtain from the last identity that

$$\begin{aligned}\|z(t) - z^\star(t)\|_2 &= \|\nabla_{zz}^{-1}\mathcal{L}(\eta(t),t)\nabla_z \mathcal{L}(z(t),t)\|_2 \\ &\leq \|\nabla_{zz}^{-1}\mathcal{L}(\eta(t),t)\|_2 \|\nabla_z \mathcal{L}(z(t),t)\|_2 \\ &\leq M\|\nabla_z \mathcal{L}(z(0),0)\|_2 e^{-\alpha t}. \end{aligned}$$

The proof is complete.

C. Proof of Lemma 1

Define $\widetilde{x}^\star(t)$ as

$$\widetilde{x}^\star(t) := \operatorname*{argmin}_{x\in\mathbb{R}^n} f_0(x,t) + \sum_{i=1}^p \mathbb{I}_-(f_i(x,t) - s(t)), \tag{48}$$

which is a perturbed version of the original optimization problem (18) after including the slack variable $s(t)$ in the constraints. By perturbation and sensitivity analysis [20, §5], we can establish the following inequality when $s(t) \geq 0$,

$$0 \leq f_0(x^\star(t),t) - f_0(\widetilde{x}^\star(t),t) \leq \sum_{i=1}^p \lambda_i^\star s(t), \tag{49}$$

where the inequality holds for all $\lambda^\star \in \Gamma^\star(t)$. The left inequality is based on the fact that the feasible set is enlarged when $s(t) \geq 0$, and hence, the optimal value does not increase. The right inequality follows directly from a sensitivity analysis of the original problem [20, §5]. On the other hand, by replacing the indicator function $\mathbb{I}_-(u)$ in (48) by the logarithmic barrier function $-1/c\log(-u)$ (see (20)), we obtain the bound [20, §11]

$$f_0(\widehat{x}^\star(t),t) - f_0(\widetilde{x}^\star(t),t) \leq \frac{p}{c(t)}. \tag{50}$$

It follows from (49), (50), and the triangle inequality that

$$\begin{aligned}|f_0(\widehat{x}^\star(t),t) - f_0(x^\star(t),t)| &\leq \frac{p}{c(t)} + \sum_{i=1}^p \lambda_i^\star s(t) \\ &\leq \frac{p}{c(t)} + s(t)(\inf_{\lambda\in\Gamma^\star(t)} \|\lambda^\star\|_1). \end{aligned}$$

The proof is complete.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

D. Proof of Lemma 2

Since $f_0(x,t)$ is strongly convex, and $c(t)$ is strictly positive, it follows that $\nabla_{xx}\Phi$ is m_f-strongly convex for $x \in \widehat{\mathcal{D}}(t)$ and, therefore, $\nabla_{xx}^{-1}\Phi$ exists and is uniformly bounded. The time derivative of $\nabla_x\Phi$ can be written as

$$\dot{\nabla}_x\Phi = \nabla_{xx}\Phi\ \dot{x} + \nabla_{xc}\Phi\ \dot{c} + \nabla_{xs}\Phi\ \dot{s} + \nabla_{xt}\Phi. \tag{51}$$

Substituting $\dot{x}$ from (23) into the last identity results in the closed-loop dynamics $\dot{\nabla}_x\Phi = -\alpha\nabla_x\Phi$, which in turn implies that

$$\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2 = e^{-\alpha t}\|\nabla_x\Phi(x(0),c(0),s(0),0)\|_2. \tag{52}$$

In particular, $\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2$ is bounded for all $t \geq 0$. Notice that

$$\nabla_x\Phi(x,c,s,t) := \nabla_x f_0(x,t) + \frac{1}{c}\sum_{i=1}^{p}\frac{\nabla_x f_i(x,t)}{s - f_i(x,t)}, \tag{53}$$

implying that $\|\nabla_x\Phi\|_2$ is unbounded (singular) at the boundary of $\widehat{\mathcal{D}}(t)$. Therefore, since $\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2$ is bounded for all $t \geq 0$, it must hold that $x(t) \in \widehat{\mathcal{D}}(t)$ for all $t \geq 0$. Finally, it follows from m_f-strong convexity of Φ that

$$\|x(t) - \widehat{x}^\star(t)\|_2 \leq \frac{1}{m_f}\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2.$$

Combining the last inequality with (52) yields the desired inequality. The proof is complete.

E. Proof of Theorem 2

We first define the following energy functional,

$$V(t) := \frac{1}{2}\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2^2,$$

which is zero along the approximate optimal trajectory $\widehat{x}^\star(t) \in \widehat{\mathcal{D}}(t)$. For any solution $x(t)$ to the ODE (27), the time derivative of $V(t)$ is

$$\begin{aligned}\dot{V}(t) &= \nabla_x\Phi^\top\dot{\nabla}_x\Phi \\ &= \nabla_x\Phi^\top(\nabla_{xx}\Phi\,\dot{x} + \nabla_{xs}\Phi\dot{s} + \nabla_{xc}\Phi\dot{c} + \nabla_{xt}\Phi) \\ &= \nabla_x\Phi^\top(-\alpha\nabla_x\Phi + \nabla_{xt}\Phi - \widehat{\nabla}_{xt}\Phi).\end{aligned}$$

When $\|\nabla_x\Phi\|_2 \geq \varepsilon$, we have that $\alpha = \dfrac{\alpha_0}{\|\nabla_x\Phi\|_2}$ by (28), and therefore, $\dot{V}(t)$ is given by

$$\begin{aligned}\dot{V}(t) &= \nabla_x\Phi^\top(-\alpha_0\frac{\nabla_x\Phi}{\|\nabla_x\Phi\|_2} - \widehat{\nabla}_{xt}\Phi + \nabla_{xt}\Phi) \\ &= -\alpha_0\|\nabla_x\Phi\|_2 + \nabla_x\Phi^\top(-\widehat{\nabla}_{xt}\Phi + \nabla_{xt}\Phi).\end{aligned}$$

Using the assumption $\|\widehat{\nabla}_{xt}\Phi - \nabla_{xt}\Phi\|_2 \leq \eta$, we obtain the inequality

$$\dot{V}(t) \leq (\eta - \alpha_0)\|\nabla_x\Phi\|_2 = (\eta - \alpha_0)\sqrt{2V(t)}.$$

By the Comparison lemma [38], we can write $V(t) \leq W(t)$, where $W(t)$ is the solution of the initial value problem $\dot{W}(t) = (\eta - \alpha_0)\sqrt{2W(t)}$, $W(0) = V(0)$. From the last ODE, we obtain the solution $W(t) = \frac{1}{2}(\sqrt{2V(0)} - (\alpha_0 - \eta)t)^2$. Hence, $V(t)$ satisfies the bound

$$2V(t) \leq (\sqrt{2V(0)} - (\alpha_0 - \eta)t)^2,$$

or, equivalently,

$$\|\nabla_x\Phi\|_2 \leq \|\nabla_x\Phi_0\|_2 - (\alpha_0 - \eta)t.$$

The above inequality implies that $\|\nabla_x\Phi\|_2$ becomes equal to ε in finite time. When $\|\nabla_x\Phi\|_2 \leq \varepsilon$, we have that $\alpha = \alpha_0/\varepsilon$ by (28), and the time derivative of $V(t)$ is given by

$$\begin{aligned}\dot{V}(t) &= \nabla_x\Phi^\top(-\frac{\alpha_0}{\varepsilon}\nabla_x\Phi - \widehat{\nabla}_{xt}\Phi + \nabla_{xt}\Phi) \\ &= -\frac{\alpha_0}{\varepsilon}\|\nabla_x\Phi\|_2^2 + \nabla_x\Phi(-\widehat{\nabla}_{xt}\Phi + \nabla_{xt}\Phi) \\ &\leq \|\nabla_x\Phi\|_2(-\frac{\alpha_0}{\varepsilon}\|\nabla_x\Phi\|_2 + \eta).\end{aligned}$$

which is negative when $\|\nabla_x\Phi\|_2 > \eta\varepsilon/\alpha_0$. As a result, we have that $\lim_{t\to\infty}\|\nabla_x\Phi\|_2 \leq \eta\varepsilon/\alpha_0$. Finally, it follows by m_f-convexity of Φ that

$$\|x(t) - \widehat{x}^\star(t)\|_2 \leq \frac{1}{m_f}\|\nabla_x\Phi(x(t),c(t),s(t),t)\|_2.$$

Therefore,

$$\lim_{t\to\infty}\|x(t) - \widehat{x}^\star(t)\|_2 \leq \frac{\varepsilon\eta}{\alpha_0 m_f}.$$

The proof is complete.

F. Expressions for the Numerical Examples

We now derive explicit expressions for the terms in the ODE (23) for the example in Subsection IV-D. The gradient of the augmented objective function with respect to x takes the form

$$\nabla_x\Phi(x,x_c,t) = x - x_d + \frac{1}{c(t)}\sum_{i=1}^{m}\frac{a_i(x_c)}{b_i(x_c) - a_i(x_c)^\top x},$$

and its Hessian reads as

$$\nabla_{xx}\phi(x,x_c,t)(x,x_c) = I_n + \frac{1}{c(t)}\sum_{i=1}^{m}\frac{a_i(x_c)a_i(x_c)^\top}{(b_i(x_c) - a_i(x_c)^\top x)^2}.$$

Furthermore, the time derivative of the gradient of the barrier function can be written as

$$\begin{aligned}\nabla_{xt}\phi(x,x_c,t) = &-\frac{\dot{c}(t)}{c(t)^2}\sum_{i=1}^{m}\frac{a_i(x_c)}{b_i(x_c) - a_i(x_c)^\top x} \\ &+\frac{1}{c(t)}\sum_{i=1}^{m}\frac{\dot{a}_i(x_c)}{b_i(x_c) - a_i(x_c)^\top x} \\ &-\frac{1}{c(t)}\sum_{i=1}^{m}a_i(x_c)\frac{\dot{b}_i(x_c) - \dot{a}_i(x_c)^\top x}{(b_i(x_c) - a_i(x_c)^\top x)^2}.\end{aligned}$$

The expressions for $\dot{a}_i(x_c)$ and $\dot{b}_i(x_c)$ are derived below. The expression of $a_i(x_c)$ for every $i = [m]$ is given by $a_i(x_c) = (x_i - x_c)$. Thus, its time derivative is given by

$$\dot{a}_i(x_c) = -\dot{x}_c = K(x_c - \hat{x}),$$

where the last equality comes from replacing the time derivative of x_c by the control law (43). We derive next the expression for the time derivative of $b_i(x_c)$ defined in (40). To do

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

so, we compute the time derivative of $\theta_i(x_c)$. Differentiating $\theta_i(x_c)$, defined in (40), yields

$$\dot{\theta}_i(x_c) = -\frac{(x_c - x_i)^\top \dot{x}_c}{\|x_c - x_i\|^4} = K\frac{(x_c - x_i)^\top (x_c - \hat{x})}{\|x_c - x_i\|^4}. \quad (54)$$

Differentiating $b_i(x)$ in (40) yields

$$\begin{aligned}\dot{b}_i(x_c) = &- \dot{x}_c^\top (\theta_i x_i + (1-\theta_i)x_c) \\ &+ \dot{\theta}_i \|x_i - x_c\|^2 (r - \theta_i)(x_i - x_c)^\top \dot{x}_c,\end{aligned}$$

where in the above equation, $\dot{\theta}_i$ and $\dot{x}_c$ are respectively given by (54) and (43).

REFERENCES

[1] U. Helmke and J. B. Moore, *Optimization and dynamical systems*. Springer Science & Business Media, 2012.

[2] D. Feijer and F. Paganini, "Stability of primal–dual gradient dynamics and applications to network optimization," *Automatica*, vol. 46, no. 12, pp. 1974–1981, 2010.

[3] J. Wang and N. Elia, "Control approach to distributed optimization," in *Communication, Control, and Computing (Allerton), 2010 48th Annual Allerton Conference on*, pp. 557–561, IEEE, 2010.

[4] J. Wang and N. Elia, "A control perspective for centralized and distributed convex optimization," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*, pp. 3800–3805, IEEE, 2011.

[5] B. Gharesifard and J. Cortés, "Distributed continuous-time convex optimization on weight-balanced digraphs," *IEEE Transactions on Automatic Control*, vol. 59, no. 3, pp. 781–786, 2014.

[6] S. S. Kia, J. Cortés, and S. Martínez, "Distributed convex optimization via continuous-time coordination algorithms with discrete-time communication," *Automatica*, vol. 55, pp. 254–264, 2015.

[7] A. Cherukuri, E. Mallada, and J. Cortés, "Asymptotic convergence of constrained primal–dual dynamics," *Systems & Control Letters*, vol. 87, pp. 10–15, 2016.

[8] M. Fazlyab, A. Koppel, V. Preciado, and A. Ribeiro, "A variational approach to dual methods for constrained convex optimization," in *IEEE American Control Conference, Seattle, WA*, 2017.

[9] C. Botsaris, "A class of methods for unconstrained minimization based on stable numerical integration techniques," *Journal of mathematical analysis and applications*, vol. 63, no. 3, pp. 729–749, 1978.

[10] A. Brown and M. C. Bartholomew-Biggs, "Some effective methods for unconstrained optimization based on the solution of systems of ordinary differential equations," *Journal of Optimization Theory and Applications*, vol. 62, no. 2, pp. 211–224, 1989.

[11] F. Y. Jakubiec and A. Ribeiro, "D-map: Distributed maximum a posteriori probability estimation of dynamic systems," *Signal Processing, IEEE Transactions on*, vol. 61, no. 2, pp. 450–466, 2013.

[12] R. L. Cavalcante and S. Stanczak, "A distributed subgradient method for dynamic convex optimization problems under noisy information exchange," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 7, no. 2, pp. 243–256, 2013.

[13] K. Zhou, S. Roumeliotis, *et al.*, "Multirobot active target tracking with combinations of relative observations," *Robotics, IEEE Transactions on*, vol. 27, no. 4, pp. 678–695, 2011.

[14] S.-Y. Tu and A. H. Sayed, "Mobile adaptive networks," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 5, no. 4, pp. 649–664, 2011.

[15] S. Lee, Y. Diaz-Mercado, and M. Egerstedt, "Multirobot control using time-varying density functions," *Robotics, IEEE Transactions on*, vol. 31, pp. 489–493, April 2015.

[16] W. Su, *Traffic Engineering and Time-varying Convex Optimization*. PhD thesis, The Pennsylvania State University, 2009.

[17] H. Myung and J.-H. Kim, "Time-varying two-phase optimization and its application to neural-network learning," *IEEE Transactions on Neural Networks*, vol. 8, pp. 1293–1300, Nov 1997.

[18] Y. Zhao and W. Lu, "Training neural networks with time-varying optimization," in *Neural Networks, 1993. IJCNN '93-Nagoya. Proceedings of 1993 International Joint Conference on*, vol. 2, pp. 1693–1696 vol.2, Oct 1993.

[19] A. Jadbabaie, A. Rakhlin, S. Shahrampour, and K. Sridharan, "Online optimization: Competing with dynamic comparators," in *Artificial Intelligence and Statistics (AISTATS), 2015*, 2015.

[20] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.

[21] A. V. Fiacco and G. P. McCormick, *Nonlinear programming: sequential unconstrained minimization techniques*, vol. 4. Siam, 1990.

[22] J. Bolte and M. Teboulle, "Barrier operators and associated gradient-like dynamical systems for constrained minimization problems," *SIAM journal on control and optimization*, vol. 42, no. 4, pp. 1266–1292, 2003.

[23] F. Alvarez, J. Bolte, and O. Brahic, "Hessian riemannian gradient flows in convex programming," *SIAM journal on control and optimization*, vol. 43, no. 2, pp. 477–501, 2004.

[24] A. Y. Popkov, "Gradient Methods for Nonstationary Unconstrained Optimization Problems," *Automation and Remote Control*, vol. 66, pp. 883–891, June 2005.

[25] M. M. Zavlanos, A. Ribeiro, and G. J. Pappas, "Network integrity in mobile robotic networks," *Automatic Control, IEEE Transactions on*, vol. 58, no. 1, pp. 3–18, 2013.

[26] Q. Ling and A. Ribeiro, "Decentralized dynamic optimization through the alternating direction method of multipliers," in *Signal Processing Advances in Wireless Communications (SPAWC), 2013 IEEE 14th Workshop on*, pp. 170–174, IEEE, 2013.

[27] Y. Zhang, Y. Yang, and G. Ruan, "Performance analysis of gradient neural network exploited for online time-varying quadratic minimization and equality-constrained quadratic programming," *Neurocomputing*, vol. 74, no. 10, pp. 1710–1719, 2011.

[28] A. Simonetto, "Time-varying convex optimization via time-varying averaged operators," *arXiv preprint arXiv:1704.07338*, 2017.

[29] A. Simonetto, A. Mokhtari, A. Koppel, G. Leus, and A. Ribeiro, "A class of prediction-correction methods for time-varying convex optimization," *IEEE Transactions on Signal Processing*, vol. 64, pp. 4576–4591, Sept 2016.

[30] M. Baumann, C. Lageman, and U. Helmke, "Newton-type algorithms for time-varying pose estimation," in *Intelligent Sensors, Sensor Networks and Information Processing Conference, 2004. Proceedings of the 2004*, pp. 155–160, IEEE, 2004.

[31] S. Rahili and W. Ren, "Distributed continuous-time convex optimization with time-varying cost functions," *IEEE Transactions on Automatic Control*, vol. 62, no. 4, pp. 1590–1605, 2017.

[32] F. A. Potra and S. J. Wright, "Interior-point methods," *Journal of Computational and Applied Mathematics*, vol. 124, no. 1, pp. 281–302, 2000.

[33] A. Iserles, *A first course in the numerical analysis of differential equations*. No. 44, Cambridge University Press, 2009.

[34] M. Fazlyab, C. Nowzari, G. J. Pappas, A. Ribeiro, and V. M. Preciado, "Self-triggered time-varying convex optimization," in *Decision and Control (CDC), 2016 IEEE 55th Conference on*, pp. 3090–3097, IEEE, 2016.

[35] S.-J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, "An interior-point method for large-scale l 1-regularized least squares," *Selected Topics in Signal Processing, IEEE Journal of*, vol. 1, no. 4, pp. 606–617, 2007.

[36] O. Arslan and D. E. Koditschek, "Exact robot navigation using power diagrams," in *Robotics and Automation, 2016 IEEE International Conference on*, 2016.

[37] F. Aurenhammer, "Power diagrams: properties, algorithms and applications," *SIAM Journal on Computing*, vol. 16, no. 1, pp. 78–96, 1987.

[38] H. K. Khalil and J. Grizzle, *Nonlinear systems*, vol. 3. Prentice hall New Jersey, 1996.

Mahyar Fazlyab received his B.Sc. and M.Sc. degrees in Mechanical Engineering from Sharif University of Technology, Tehran, Iran, in 2010 and 2013, respectively. He has been a Ph.D. student with the Department of Electrical and Systems Engineering at the University of Pennsylvania since September 2013. His research interests include the analysis, optimization, and control of (networked) dynamical systems.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAC.2017.2760256, IEEE Transactions on Automatic Control

Santiago Paternain received the B.Sc. degree in Electrical Engineering from Universidad de la República Oriental del Uruguay, Montevideo, Uruguay in 2012. Since August 2013, he has been working toward the Ph.D. degree in the Department of Electrical and Systems Engineering, University of Pennsylvania. His research interests include optimization and control of dynamical systems.

Victor M. Preciado received his Ph.D. degree in Electrical Engineering and Computer Science from the Massachusetts Institute of Technology in 2008. He is currently the Raj and Neera Singh Assistant Professor of Electrical and Systems Engineering at the University of Pennsylvania. He is a member of the Networked and Social Systems Engineering (NETS) program and the Warren Center for Network and Data Sciences. His research interests include network science, dynamic systems, control theory, and convex optimization with applications in socio-technical systems, technological infrastructure, and biological networks. Dr. Preciado received the NSF CAREER Award in 2017 and best student paper (as advisor) at the 2016 Stochastic Network Conference.

Alejandro Ribeiro received the B.Sc. degree in electrical engineering from the Universidad de la Republica Oriental del Uruguay, Montevideo, in 1998 and the M.Sc. and Ph.D. degrees in electrical engineering from the Department of Electrical and Computer Engineering, the University of Minnesota, Minneapolis, MN, USA, in 2005 and 2007, respectively. From 1998 to 2003, he was a member of the technical staff at Bellsouth Montevideo. After his M.Sc. and Ph.D. studies, in 2008, he joined the University of Pennsylvania, Philadelphia, PA. USA, where he is currently an Assistant Professor at the Department of Electrical and Systems Engineering. His research interests are in the applications of statistical signal processing to the study of networks and networked phenomena. His current research focuses on wireless networks, network optimization, learning in networks, networked control, robot teams, and structured representations of networked data structures. Dr. Ribeiro received the 2012 S. Reid Warren, Jr. Award presented by the University of Pennsylvanias undergraduate student body for outstanding teaching and the NSF CAREER Award in 2010. He is also a Fulbright scholar and the recipient of student paper awards at the 2013 American Control Conference (as adviser), as well as the 2005 and 2006 International Conferences on Acoustics, Speech and Signal Processing.