ISSN: 2834-4928

Research Article

Journal of Electrical Electronics Engineering

A Novel Path-Following Method for Time-Varying Optimizations with Optimal Parametric Functions

Mohsen Amidzadeh*

Department of Computer Science, Aalto University, Finland

***Corresponding Author**
Mohsen Amidzadeh, Department of Computer Science, Aalto University, Finland.

Submitted: 2024, May 08; **Accepted:** 2024, Jun 05; **Published:** 2024, Jul 11

Citation: Amidzadeh, M. (2024). A Novel Path-Following Method for Time-Varying Optimizations with Optimal Parametric Functions. *J Electrical Electron Eng, 3*(4), 01-10.

Abstract
We address a broad category of nonlinear constrained optimization problems. We then reformulate it as a time-varying optimization using continuous-time parametric functions and derive a dynamical system to track the respective optimal solution. We then re-parameterize the dynamical system according to a linear combination of parametric functions. By applying the calculus of variations, we optimize these parametric functions to minimize the optimality distance. Consequently, we develop an iterative dynamic algorithm, termed as OP-TVO, to achieve efficient convergence to the solution. We compare the performance of the proposed algorithm with the prediction correction method (PCM) in terms of optimality and computational complexity. The results demonstrate that OP-TVO effectively competes with PCM for the given class of optimization problems, suggesting a promising alternative to PCM. This work also introduces a novel paradigm for solving parametric dynamic systems.

Keywords: Time-Varying Optimization Problem, Functional Optimization Problem, Prediction-Correction Method, Optimality Distance, Dynamical System

1. Introduction

Time-Varying Optimization (TVO) problems involve parametric optimization where the objective function, constraints, or both are expressed using continuously varying functions. This approach is used to determine the optimal trajectory of solutions in continuous-time optimization scenarios. Additionally, for problems where the optimal solution is known for a specific configuration, TVO can be utilized to extra-polate the solution to other settings of interest.

TVO is explored in the domain of parametric programming, where the optimization problem is parameterized using continuous parameters [1-5]. In prediction-correction, methods are developed to address nonlinear constrained TVO problems, providing a mechanism to track a solution trajectory with certain convergence guarantees [3]. A path-following procedure is proposed in to trace the solution path of a parametric nonlinear problem [4]. A quadratic programming is then employed which leads to some convergence properties for their method. In a path-following approach is used to track the solutions of parametric nonlinear constrained programs through a semi-smooth barrier function [5].

TVO can also be viewed as an extension of time-invariant optimization problems for the extrapolation purposes [6-10]. Presents an interior-point method for optimization problems with time-varying objective and constraint functions, formulating a continuous-time dynamical system to track the optimal solution with bounded asymptotic tracking error [8]. In path following, methods are devised in the dual space to track the solutions of time-varying linearly constrained problems [9].

When it goes to the approach for solving TVO problems, prediction-correction schemes emerge as promising tracking algorithms [11-14]. It involves a dynamic-tracking mechanism, called the predictor step to track the solution trajectory over time, accompanied with a Newton based iterative mechanism, called corrector step to adjust the prediction errors. In a discrete-time prediction-correction, approach is proposed to minimize unconstrained time-varying functions, analyzing the asymptotic tracking error to ensure convergence [6]. In prediction-correction methods are introduced to track the optimal solution trajectory in the primal space with a bounded asymptotic error [7]. In prediction-correction, methods are devised in the dual space to track the solutions of time-varying linearly constrained problems [9].

In this paper, we study a category of nonlinear constrained optimization problems where the optimal solution is known for a specific setting and needs to be determined for a target configuration. We utilize the concept of TVO to reformulate the problem using parametric programming. We then investigate a set of parametric functions to optimize the optimality distance. Unlike approaches based on prediction correction methods, our focus is on a functional optimization problem to expedite the convergence rate. Specifically, this paper differs from prediction-correction approaches by designing parametric functions to minimize the optimality distance of the solution, rather than focusing on correcting prediction errors [15,16]. The main contributions of this paper are listed as follows:

• We address a class of nonlinear constrained optimization problems by formulating them as time varying optimization problems using parametric functions. We then use a re-parametrization trick to show that the corresponding dynamical system can be represented based on a linear combination of the parametric functions.
• We globally optimize the parametric functions by a functional optimization procedure, and develop an iterative algorithm with minimum optimality distance. We call the devised algorithm *Optimal Parametric Time-Varying Optimization* (OP-TVO).
• We compare OP-TVO with prediction-correction method from the literature, from the optimality and computational complexity perspectives.

Notations: In this paper, we use lower-case a for scalars, boldface lower-case a for vectors and boldface uppercase **A** for matrices. Further, $\mathbf{A}^{\top}$ is the transpose of **A**, $\|\mathbf{a}\|$ is the Euclidean norm of **a**, $\nabla_a g(\bullet)$ and $\nabla_a^2 g(\bullet)$ are the gradient vector and Hessian matrix of multivariate function $g(\mathbf{a})$ with respect to (w.r.t.) vector **a**, respectively. We show the components of a n-dimensional column vector a using the notation $a = [a_1,...,a_n]^{\top}$. Further, $\{a_n\}_1^N$ collects the components of vector a from $n = 1$ to $n = N$. We use **I, 1** and **0** to denote the identity matrix, all-ones and all-zeros vectors, respectively.

We use a (θ) to represent the derivative of $a(\theta)$ w.r.t. θ.

2. Problem Statement

This paper studies a class of nonlinear constrained optimization problems. The problem involves an objective function $f(\cdot) : \mathbb{R}^N \to \mathbb{R}$, vector-valued constraint functions $h_m(\cdot) = [h_{m,1},...,h_{m,N}]^{\top}(\cdot) : \mathbb{R}^N \to \mathbb{R}^N$ and optimization variables $x_m \in \mathbb{R}^N$ for $m \in \{1,...,M\}$. The problem is specifically
described as:

$$P_0 : \quad \min_{\{\boldsymbol{x}_m\}_1^M} \sum_{m=1}^{M} a_m f(\boldsymbol{x}_m) \quad \text{s.t.} \quad \sum_{m=1}^{M} \boldsymbol{h}_m(\boldsymbol{x}_m) = \boldsymbol{u}, \tag{1}$$

Where $u \in \mathbb{R}^N$. The optimal solution of Problem P_0 is known for given non-zero parameters $a_m = p_{0,m}$, with $m \in \{1,..., M\}$. The **aim** is then to find the optimal solution for non-zero target parameters $a_m = p_{\tau,m}$.

Such problems arise in distributed optimizations or multi-agent systems, where individual agents are required to optimize agent-specific rewards contributing to an overall cost function. The aim is thus to obtain agent-specific variables $\{\boldsymbol{x}_m\}_1^M$ that optimize this overall cost function. This class of challenges also arise in constrained problems with an objective established from different cost functions with distinct weights $\{a_m\}_1^M$.

We then intend to express P_0 based on a TVO problem with parametric functions, and devise a path following method with convergence rate being optimized. For this, we follow a functional optimization approach to design the parametric functions. Consequently, we parameterize $\{a_m\}_1^M$ with the parametric functions $\{b_m(\theta)\}_1^M$ and parameter $\theta \in \mathbb{R}^+ \cup \{0\}$, so that

$$\lim_{\theta \to 0} b_m(\theta) = p_{0,m}, \qquad \lim_{\theta \to \tau} b_m(\theta) = p_{\tau,m}, \tag{2}$$

For $m \in \{1,...,M\}$. Note that such parametric functions as presented in (2) are not unique. However, we aim to find those parametric functions that optimize the convergence rate.

We then consider the following TVO problem

$$P_1(\theta) : \quad \boldsymbol{x}^*(\theta) = \operatorname*{argmin}_{\{\boldsymbol{x}_m(\theta)\}_1^M} \sum_{m=1}^{M} b_m(\theta) f\big(\boldsymbol{x}_m(\theta)\big) \quad \text{s.t.} \quad \sum_{m=1}^{M} \boldsymbol{h}_m\big(\boldsymbol{x}_m(\theta)\big) = \boldsymbol{u}, \tag{3}$$

Where $\boldsymbol{x}(\theta) = [\boldsymbol{x}_1^\top, \ldots, \boldsymbol{x}_M^\top]^\top(\theta)$. As declared, we assume that the solution of $P_1(\theta)$ for $\theta = 0$ is given,

And the solution at target $\theta = \tau > 0$ is to be found.

A naive approach to find the solution of $P_1(\theta)$ is to use a Newton-based iterative algorithm. However, this approach may suffer from low convergence rate and its solution optimality depends on a step-length parameter.

Instead, we develop a time-varying approach to exploit the information of optimal solution of $P_1(0)$. We thus formulate the following continuous-time dynamical system

$$\dot{\boldsymbol{x}}_m(\theta) = \boldsymbol{\phi}_m\big(\boldsymbol{x}(\theta), \theta\big): \ \mathbb{R}^{NM} \times \mathbb{R}^+ \cup \{0\} \to \mathbb{R}^N, \tag{4}$$

With optimal trajectory solution denoted by $x^*(\bullet)$. We then devise an iterative approach to predict $x^*(\bullet)$ by $x(\bullet)$ so that the *optimality distance* $\|x(\theta) - x^*(\theta)\|$ is minimized for $\theta \to \tau$. Note that (4) shows a set of ODEs which should be solved with initial condition $x(0)$ to give the desired solution $x(\tau)$. As such, we intend to jointly solve TVO $P_1(\theta)$ and design $\{b_m(\theta)\}_1^M$ for $\theta \in [0,\tau]$ so that the optimality distance is minimized.

We need the following assumptions for TVO $P_1(\theta)$.

Assumption I: The objective function $f(\bullet)$ and constraints $h_m(\bullet)$ are twice continuously differentiable with respect to (w.r.t.) x_m.

Assumption II: The matrices $\{\boldsymbol{\mathcal{J}}_m\}_1^M$ are invertible for $\theta \in [0,\tau]$, where $\boldsymbol{\mathcal{J}}_m \in \mathbb{R}^{N\times N}$ is the transpose of Jacobian matrix of $h_m(\bullet)$ w.r.t. $x_m(\theta)$.

3. ODES Associated with TVO P1(θ)

Proposition 1. The solution of Karush–Kuhn–Tucker conditions of problem $P_1(\theta)$, for $\theta \in [0,\tau]$, can
Be found by the pair $(x(\theta),\lambda)$ which follows the dynamical system (4) with:

$$\boldsymbol{\phi}_m\big(\boldsymbol{x}(\theta), \theta\big) = -\Big(b_m(\theta)\nabla_m^2 f + \sum_{n=1}^{N} \lambda_n \nabla_m^2 h_{m,n}\Big)^{-1} \boldsymbol{\mathcal{J}}_m \left(\dot{\boldsymbol{\lambda}} - \frac{\dot{b}_m(\theta)}{b_m(\theta)}\boldsymbol{\lambda}\right), \tag{5}$$

Where

$$\boldsymbol{\lambda} = -b_m(\theta)\boldsymbol{\mathcal{J}}_m^{-1}\nabla_m f, \qquad \text{for} \quad m \in \{1, \ldots, M\}, \tag{6}$$

And $\dot{\lambda}$ is obtained using

$$\sum_{m=1}^{M} \boldsymbol{\mathcal{J}}_m^\top \dot{\boldsymbol{x}}_m(\theta) = \mathbf{0}. \tag{7}$$

Proof. Please refer to Appendix A.
According to (5), the dynamical system has been formulated based on a nonlinear Combination of two parametric functions, $b_m(\theta)$ and $\dot{b}_m(\theta)$. In the following, we use a decomposition trick to express the dynamical system as a linear combination of parametric functions.

A. Reparametrizing based on a Decomposition

We introduce the parametric functions

$$c_m(\theta) := \frac{\dot{b}_m(\theta)}{b_m(\theta)}, \qquad m \in \{1, \ldots, M\}, \tag{8}$$

Which should satisfy $\int_0^\tau c_m(\theta)d\theta = \log\left(\frac{p_{\tau,m}}{p_{0,m}}\right) := \psi_m$ based on (2). We then have:

Theorem 1. *The dynamical system (5) can be re-parameterized based on the following linear combination of* $\{c_m(\theta)\}_1^M$*:*

$$\dot{\boldsymbol{x}}(\theta) = \boldsymbol{\phi}\big(\boldsymbol{x}(\theta), \theta\big) = \boldsymbol{\Gamma}(\theta)\, \boldsymbol{c}(\theta), \tag{9}$$

where $\boldsymbol{\phi}(\cdot,\cdot) = [\boldsymbol{\phi}_1^\top, \ldots, \boldsymbol{\phi}_M^\top]^\top(\cdot,\cdot)$,

$$\mathbf{\Gamma}(\theta) = \begin{pmatrix} \boldsymbol{\gamma}_{11} & \cdots & \boldsymbol{\gamma}_{1M} \\ \vdots & \vdots & \vdots \\ \boldsymbol{\gamma}_{N1} & \cdots & \boldsymbol{\gamma}_{NM} \end{pmatrix} \in \mathbb{R}^{NM \times M},$$

$$\boldsymbol{\gamma}_{nm} = -G_n^{-1}\Big(\Big(\sum_{k=1}^{M} D_k\Big)^{-1} D_m - \delta_{nm}\boldsymbol{I}\Big)\mathbf{1},$$

$$G_n = \operatorname{diag}(\boldsymbol{v})^{-1}\boldsymbol{\mathcal{J}}_n^{-1}\,\Big(\nabla_n^2 f + \sum_{j=1}^{N} v_j \nabla_n^2 h_{n,j}\Big),$$

$$D_m = \boldsymbol{\mathcal{J}}_m^{\top}\, G_m^{-1},$$

for $n \in \{1,\ldots,N\}$ *and* $m \in \{1,\ldots,M\}$ *with*

$$\boldsymbol{v} = -\boldsymbol{\mathcal{J}}_m^{-1}\,\nabla_m f. \tag{10}$$

Proof. We introduce N-by-N matrices $\{G_m\}^M_1$ and exploit the following decomposition:

$$b_m(\theta)\nabla_m^2 f + \sum_{n=1}^{N} \lambda_n \nabla_m^2 h_{m,n} = \boldsymbol{\mathcal{J}}_m \operatorname{diag}(\boldsymbol{\lambda}) G_m. \tag{11}$$

Then, we get:

$$G_m = \operatorname{diag}(\boldsymbol{v})^{-1}\boldsymbol{\mathcal{J}}_m^{-1}\,\Big(\nabla_m^2 f + \sum_{n=1}^{N} v_n \nabla_m^2 h_{m,n}\Big), \tag{12}$$

for which we used $\lambda = b_m(\theta)\, v$ based on (6) and (10). Equation (12) shows that G_m is notably independent of parametric function $b_n(\theta)$. Now, by plugging (11) into (5), we obtain:

$$\dot{\boldsymbol{x}}_n(\theta) = -G_n^{-1}\operatorname{diag}(\boldsymbol{\lambda})^{-1}\big(\dot{\boldsymbol{\lambda}} - c_n(\theta)\boldsymbol{\lambda}\big),$$

and according to (7), we get:

$$\dot{\boldsymbol{\lambda}} = \boldsymbol{\lambda}\,\Big(\sum_{k=1}^{M} D_k\Big)^{-1}\sum_{m=1}^{M} D_m c_m(\theta)\,\mathbf{1},$$

which together yields:

$$\dot{\boldsymbol{x}}_n(\theta) = -G_n^{-1}\Big(\Big(\sum_{k=1}^{M} D_k\Big)^{-1}\sum_{m=1}^{M} D_m c_m(\theta) - c_n(\theta)\boldsymbol{I}\Big)\mathbf{1}.$$

Based on the definition of γ_{nm}, we thus have:

$$\dot{\boldsymbol{x}}_n(\theta) = \sum_{m=1}^{M} \boldsymbol{\gamma}_{nm} c_m(\theta), \quad n \in \{1,\ldots,N\},$$

which proves the statement.

Remark: The dynamical system (9) depends only on $c(\bullet)$ and not on $b(\bullet)$. Moreover, as $\Gamma(\bullet)$ does not depend on $c(\bullet)$, the dynamical system portrays a linear parametric expression w.r.t. $c(\bullet)$. It enables us to find the condition in which solving (5), or equivalently (9), leads to a unique solution.

In this regard, we make the following assumptions.

Assumption III: The matrices $\{\nabla_n^2(f+\boldsymbol{v}^\top \boldsymbol{h}_n)\}_1^N$, $\sum_{k=1}^M D_k$ and diag(v) are invertible.

Assumption IV: The derivatives of $\{\nabla_m f\}_1^M$, $\{\mathcal{J}_m\}_1^M$ and $\{\nabla_n^2(f+\boldsymbol{v}^\top \boldsymbol{h}_n)\}_1^N$ w.r.t. $\{\boldsymbol{x}_m(\theta)\}_1^M$ are bounded.

Proposition 2. ***If Assumptions I, II, III** and **IV** hold, then the dynamical system (9) has a unique Solution.*

Proof. By utilizing the Picard–Lindelof theorem and leveraging the equivalence between Lipschitz¨ continuity and boundedness of the derivative, the statement follows.

The linear form of (9) also enables us to design $c(\bullet)$ such that the optimality distance is optimized.

Once $c(\bullet)$ is designed, we can get: $b_m(\theta) = p_{0,m} \exp\left(\int_0^\theta c_m(\xi)d\xi\right)$ based on (2) and (8).

4. Optimality Distance

Equation (9) shows a set of ODEs that is intricate to precisely solve due to highly non-linearity w.r.t. θ. However, one approach is to use the Euler method to approximate $x(\bullet)$ by $\hat{x}(\bullet)$:

$$O_d = \|\hat{\boldsymbol{x}}(\tau) - \boldsymbol{x}(\tau)\| \le \frac{\Delta\theta^2}{2}\sum_{j=1}^{L} \|\ddot{\boldsymbol{x}}(\tau - j\Delta\theta)\| + \mathcal{O}(L\Delta\theta^3),$$

where $\Delta\theta$ is the incremental step and $\dot{\hat{x}}(\theta) = \phi(\hat{x}(\theta), \theta)$. For this method, the optimality distance O_d is upper-bounded by:

$$\hat{\boldsymbol{x}}(\theta) = \hat{\boldsymbol{x}}(\theta - \Delta\theta) + \Delta\theta\, \dot{\hat{\boldsymbol{x}}}(\theta - \Delta\theta), \qquad \theta \in (0, \tau], \tag{13}$$

where $\tau = L\Delta\theta$. This shows that optimality distance is limited by order of $\Delta\theta^2$. However, based on (13) another upper-bound can be found as follows:

$$O_d = \left\| \int_0^\tau \big(\dot{\boldsymbol{x}}(\theta) - \hat{\boldsymbol{m}}\big) d\theta \right\| \le \int_0^\tau \|\dot{\boldsymbol{x}}(\theta) - \hat{\boldsymbol{m}}\| d\theta, \tag{14}$$

where

$$\hat{\boldsymbol{m}} = \frac{1}{L}\sum_{j=1}^{L} \dot{\hat{\boldsymbol{x}}}(\tau - j\Delta\theta).$$

Note that $\hat{m}$ does not depend on θ. Consequently, minimizing the upper bound of O_d, i.e., $\int_0^\tau \|\dot{\boldsymbol{x}}(\theta) - \hat{m}\| d_\theta$, leads to the optimality distance being minimized.

5. Optimality Distance Minimization

We consider the following functional optimization problem (FOP) to jointly design the parametric functions $c(\bullet)$ and find the optimum solution $x(\bullet)$:

$$\begin{aligned} J_1: \qquad & \min_{\boldsymbol{x}(\cdot),\boldsymbol{c}(\cdot)} \int_0^\tau \|\dot{\boldsymbol{x}}(\theta) - \hat{\boldsymbol{m}}\|^2 d\theta \\ & \text{s.t.} \quad \dot{\boldsymbol{x}}(\theta) = \boldsymbol{\Gamma}(\theta)\, \boldsymbol{c}(\theta), \\ & \text{s.t.} \quad \int_0^\tau \boldsymbol{c}(\theta) d\theta = \boldsymbol{\psi}. \end{aligned} \tag{15}$$

By solving FOP J_1, we can achieve the optimal solution $x(\theta)$ which minimizes the optimality distance O_d based on (14). To solve (15), we thus constitute the Hamiltonian $\mathcal{H}$ as:

$$\mathcal{H} = \|\dot{\boldsymbol{x}}(\theta) - \hat{\boldsymbol{m}}\|^2 + \boldsymbol{w}(\theta)^\top \big(\dot{\boldsymbol{x}}(\theta) - \boldsymbol{\Gamma}(\theta)\, \boldsymbol{c}(\theta)\big) + \boldsymbol{\lambda}^\top \Big(\boldsymbol{c}(\theta) - \frac{1}{\tau}\boldsymbol{\psi}\Big),$$

Where $w(\theta)$ is a co-state variables and λ is a Lagrange multiplier. Using calculus of variations, the Functional solution of (15) is obtained by:

$$\begin{cases} \nabla_{\boldsymbol{c}(\theta)}\mathcal{H} = 2\boldsymbol{\Gamma}(\theta)^\top\Big(\boldsymbol{\Gamma}(\theta)\boldsymbol{c}(\theta) - \hat{\boldsymbol{m}} - \tfrac{1}{2}\boldsymbol{w}(\theta)\Big) + \boldsymbol{\lambda} = 0, \\ \nabla_{\boldsymbol{x}(\theta)}\mathcal{H} - \tfrac{d}{d\theta}\nabla_{\dot{\boldsymbol{x}}(\theta)}\mathcal{H} \\ \qquad\qquad = \boldsymbol{w}(\theta)^\top\, \nabla_{\boldsymbol{x}(\theta)}\boldsymbol{\Gamma}(\theta)\, \boldsymbol{c}(\theta) + 2\ddot{\boldsymbol{x}}(\theta) + \dot{\boldsymbol{w}}(\theta) = 0, \\ \dot{\boldsymbol{x}}(\theta) - \boldsymbol{\Gamma}(\theta)\,\boldsymbol{c}(\theta) = 0, \qquad \int_0^\tau \boldsymbol{c}(\theta)d\theta - \boldsymbol{\psi} = 0. \end{cases} \tag{16}$$

This system of conditions are intricate to solve and an estimation of $\hat{m}$ is needed in advance. These motivate us to develop an iterative algorithm to find the solution of J_1.

In this regard, we propose an iterative mechanism as follows: In the beginning, we initialize $c(\bullet)$ such that $\int_0^\tau \boldsymbol{c}(\theta)d\theta = \boldsymbol{\psi}$. Then, we continually follow two consecutive steps till the algorithm converges. These are the *Prediction* and *Parametric-tuning* steps. In the prediction step, we sequentially solve (13) based on (9) and recently updated $c(\bullet)$, in order to predict $x(\theta)$ for $\theta \in (0, \tau]$. In the parametric-tuning step, we minimize the functional objective $\int_0^\tau \|\dot{\boldsymbol{x}}(\theta) - \hat{\boldsymbol{m}}\|^2 d\theta$ w.r.t. $c(\bullet)$ with $\hat{m}$ being obtained based on the solution of prediction step. As declared, we perform these two steps till the convergence. We call this algorithm Optimal Parametric Time-Varying Optimization (OP-TVO).

Specifically, we consider the following FOP, in the second step of OP-TVO, to optimize the parametric functions $c(\bullet)$:

$$\begin{aligned} J_2: \quad & \min_{\boldsymbol{c}(\cdot)} \int_0^\tau \Big\|\boldsymbol{\Gamma}(\theta)\,\boldsymbol{c}(\theta) - \hat{\boldsymbol{m}}\Big\|^2 d\theta + \mu \int_0^\tau \boldsymbol{c}(\theta)^\top \boldsymbol{c}(\theta) d\theta \\ & \text{s.t.} \quad \int_0^\tau \boldsymbol{c}(\theta)d\theta = \boldsymbol{\psi}\ , \end{aligned} \tag{17}$$

Where the term $\int_0^\tau \boldsymbol{c}(\theta)^\top \boldsymbol{c}(\theta)d\theta$ is additionally added to regularize the smoothness of $c(\theta)$ w.r.t. θ, and $0 < \mu \ll 1$ is the regularization coefficient.

Proposition 3. *The globally optimal solution of J_2 is obtained by*:

$$\boldsymbol{c}(\theta) = \boldsymbol{\Pi}^{-1}(\theta)\left(\boldsymbol{\Gamma}(\theta)^\top \hat{\boldsymbol{m}} - \boldsymbol{\lambda}\right), \tag{18}$$

where

$$\begin{aligned} \boldsymbol{\Pi}(\theta) &= \boldsymbol{\Gamma}(\theta)^\top\boldsymbol{\Gamma}(\theta) + \mu\boldsymbol{I}, \\ \boldsymbol{\lambda} &= \left(\int_0^\tau \boldsymbol{\Pi}^{-1}(\theta)d\theta\right)^{-1}\left(\int_0^\tau \boldsymbol{\Pi}^{-1}(\theta)\boldsymbol{\Gamma}(\theta)^\top d\theta\, \hat{\boldsymbol{m}} - \boldsymbol{\psi}\right). \end{aligned}$$

Proof. Considering that the dynamical system (9) has been expressed based on a linear combination of parametric functions $c(\bullet)$, J_2 is a convex FOP. This implies that the globally optimal solution of J_2 can be found by applying the Euler-Lagrange equation on J_2 [17].

$$\begin{cases} \boldsymbol{\Gamma}(\theta)^\top(\boldsymbol{\Gamma}(\theta)\,\boldsymbol{c}(\theta) - \hat{\boldsymbol{m}}) + \mu\boldsymbol{c}(\theta) + \boldsymbol{\lambda} & = 0, \\ \int_0^\tau \boldsymbol{c}(\theta)d\theta - \boldsymbol{\psi} & = 0\ . \end{cases} \tag{19}$$

By solving (19), the statement follows.

Algorithm 1 shows the pseudo-code of OP-TVO. For each iteration (iter), the prediction and parametric tuning steps are executed to jointly predict the optimal solution $x^*(\tau)$ and design the parametric functions $c(\bullet)$. We also need a metric for the convergence to stop the algorithm. For this, we track the value of $\hat{O}_d := \sum_{j=1}^{L}\Big\|\boldsymbol{\phi}\Big(\hat{\boldsymbol{x}}(\tau - j\Delta\theta), \tau - j\Delta\theta\Big) - \hat{\boldsymbol{m}}\Big\|^2$ as an estimation of the optimality distance. We thus consider that the algorithm has converged if the value of $\hat{O}_d$ lies below a threshold O_{th}.

6. Numerical Results and Discussion

To evaluate the devised algorithm OP-TVO, we compare it with a Prediction-Correction Method (PCM) [6], as well as with a *Benchmark* solution obtained by an extremely small incremental step $\Delta\theta = 10^{-6}$ regardless of its computational complexity. Note that, we consider this *Benchmark* as the optimal solution, by which we can compute the optimality distance O_d. These algorithms have been implemented using Matlab R2022a on a 8 × 1.70 GHz Intel Core i5-10310U Processor, equipped with 16 GB of memory and 12 Mbytes of data cache.

Algorithm 1: Optimal Parametric Time-Varying Optimization (OP-TVO)

Input: Optimal solution $\boldsymbol{x}(0)$.

Outputs: Optimal solution $\boldsymbol{x}(\tau)$ and parametric functions $\boldsymbol{c}(\cdot)$.

Initialization:

Initialize $\boldsymbol{c}(\cdot)$ so that $\int_0^\tau \boldsymbol{c}(\theta)d\theta = \boldsymbol{\psi}$.

Set $\text{flag} = 1$ and $\text{iter} = 0$.

while flag **do**

iter = iter+1.

Prediction step:

Predict $\boldsymbol{x}(\theta)$ for $\theta \in (0, \tau]$ using (13) and (9) and based on updated $\boldsymbol{c}(\cdot)$.

Parametric-tuning step:

Update $\boldsymbol{c}(\cdot)$ using (18) and based on predicted $\boldsymbol{x}(\theta)$ with $\theta \in (0, \tau]$.

if *Convergence* **then**

$\text{flag} = 0$.

end

end

We then consider a constrained optimization problem E_1 and change the constraints to add non-linearity to the problem [18.19].

$$E_1 : \quad \min_{\{\boldsymbol{x}_m\}_1^M} \sum_{m=1}^{M} a_m \operatorname{erfc}\left(\gamma_0 \frac{x_{m,1}}{\sqrt{2^{\frac{0.1}{x_{m,2}}} - 1}}\right)$$
$$\text{s.t.} \quad \sum_{m=1}^{M} \log(1 + x_{m,1}) = L,$$
$$\text{s.t.} \quad \sum_{m=1}^{M} x_{m,2}^2 = 1,$$

where the optimization variables are $\boldsymbol{x}_m = [x_{m,1}, x_{m,2}]^\top$ for $m \in \{1, \ldots, M\}$, $M = 100$, $a_m = m^{-\tau} / \sum_{m=1}^{M} m^{-\tau}$, $\tau = 3$ and $\gamma_0 = 40$. For E_1, it can be verified that the optimal solution for $a_m = \frac{1}{M}$ is obtained as:

$$x_{m,1} = \exp(L/M - 1), \quad x_{m,2} = \sqrt{1/M}, \quad m \in \{1, \ldots, M\}.$$

Therefore, we constitute a TVO problem exactly as E_1 but with parametric functions $\{b_m(\theta)\}_1^M$ replacing

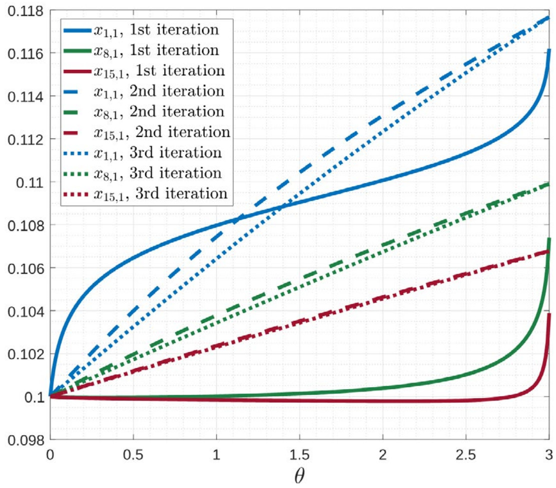

Figure 1: Solution Trajectories for $x_{m,1}$ with $m \in \{1,8,15\}$

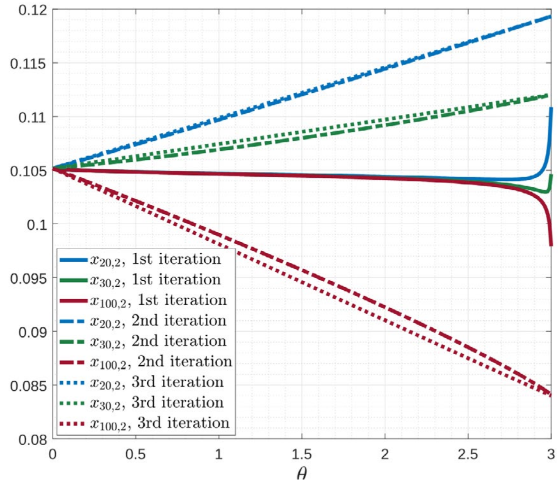

Figure 2: Solution Trajectories for $x_{m,2}$ with $m \in \{20,30,100\}$

$\{a_m\}_1^M$ such that:

$$b_m(0) = \frac{1}{M}, \quad b_m(\tau) = a_m, \quad m \in \{1, \dots, M\}.$$

We further use the re-parametrizing functions $c(\bullet)$ with $c_m(\theta) := \frac{\dot{b}_m(\theta)}{b_m(\theta)}$. We then apply Algorithm 1 with hyper-parameters $\mu = 10^{-7}$, $O_{th} = 10^{-5}$ and $\Delta\theta = 10^{-2}$ to find the optimal solution.

Figures 1 and 2 illustrate the solution trajectories $\{\boldsymbol{x}_m(\theta)\}_1^M$, obtained by Algorithm 1, as a function of θ for different iterations. For the first iteration, the trajectory solutions portray a highly nonlinear behavior. However, as the number of iteration increases, this non-linearity reduces. When the algorithm converges (3rd iteration), the linear curvature of the solution trajectories indicate that a precise solution with minimal optimality distance has been obtained.

We also apply a PCM on Problem E_1 to obtain the solution. To have a fair comparison, we adjust the hyper-parameter $\Delta\theta$ for PCM such that the corresponding optimal value is almost equal to that of Algorithm 1. As such, we need to set $\Delta\theta = 10^{-4}$.

Table I compares the performance results of PCM, Benchmark and OP-TVO. The second column presents the optimal values achieved by these approaches. The third column quantifies the extent of constraints violation, represented as the summation of violations of all constraints. The fourth column indicates the elapsed time in seconds for the computations. And the fifth and sixth columns show the optimality distance O_d and its estimation $\hat{O}_d$, respectively.

Despite PCM achieving a slightly lower objective function value than Benchmark, it is noteworthy that Benchmark excels significantly in terms of constraint satisfaction compared to PCM. Therefore, the Benchmark solution is considered the reference in this comparison. Based on the values of $\hat{O}_d$,

Approach	Optimal Value	Constraint Violations	Elapsed Time [s]	O_d	$\hat{O}_d$
Benchmark	5.6089×10^{-7}	1.312×10^{-6}	2066	0.0	N/A
PCM	5.6086×10^{-7}	1.739×10^{-4}	271	0.00135	N/A
OP-TVO, iter=1	3.2443×10^{-5}	0.419	3	0.8291	N/A
iter=2	5.6088×10^{-7}	5.129×10^{-5}	52	5.132×10^{-4}	0.228
iter=3	5.6088×10^{-7}	5.051×10^{-5}	105	5.130×10^{-4}	2.68×10^{-6}
iter=4	5.6088×10^{-7}	5.050×10^{-5}	160	5.130×10^{-4}	2.14×10^{-6}

Table 1: Performance Result of OP-TVO and PCM

OP-TVO chieves convergence after three iterations. OP-TVO with iter = 3 outperforms PCM from the computational complexity as it converges within 10^5 seconds while PCM converges after 271 seconds. Furthermore, OP-TVO with iter = 3 exhibits a superior optimality distance O_d compared to PCM. Not to mention that OP-TVO better satisfies the constraints than PCM. These results indicate that OP-TVO provides a more accurate solution than PCM, demonstrating its capability to achieve optimal solutions with lower computational complexity and higher performance precision.

7. Conclusion

In this paper, we reformulated a class of nonlinear constrained optimization problems using a time varying optimization approach incorporating parametric functions. By applying a re-parametrization technique, we transformed the problem into a dynamical system expressed linearly in terms of these parametric functions. Our objective was to minimize the optimality distance traced by this dynamical system, which led us to formulate a functional minimization problem. To achieve this, we introduced an iterative algorithm named OP-TVO, designed specifically to determine the trajectory of solutions with an optimal optimality distance. Our experimental results demonstrate that OP-TVO surpasses the Prediction-Correction Method (PCM) in terms of both optimality distance and convergence rate. These findings highlight OP-TVO as a promising alternative to PCM for addressing distributed time-varying optimization problems. By building upon the results gained from this study, future work can explore optimization problems involving time-varying constraints.

References

1. Guddat, J., Vazquez, F. G., & Jongen, H. T. (1990). Parametric optimization: singularities, path following and jumps. Stuttgart: BG Teubner.
2. Dinh, Q. T., Savorgnan, C., & Diehl, M. (2012). Adjoint-based predictor-corrector sequential convex programming for parametric nonlinear optimization. *SIAM Journal on Optimization, 22*(4), 1258-1284.
3. Kungurtsev, V., & Diehl, M. (2014). Sequential quadratic programming methods for parametric nonlinear optimization. *Computational Optimization and Applications, 59*, 475-509.
4. Kungurtsev, V., & Jaschke, J. (2017). A predictor-corrector path-following algorithm for dual-degenerate parametric optimization problems. *SIAM Journal on Optimization, 27*(1), 538-564.
5. Liao-McPherson, D., Nicotra, M. M., & Kolmanovsky, I. V. (2018). A semismooth predictor corrector method for real-time constrained parametric optimization with applications in model predictive control. In *2018 IEEE Conference on Decision and Control (CDC)* (pp. 3600-3607). IEEE.
6. Simonetto, A., Mokhtari, A., Koppel, A., Leus, G., & Ribeiro, A. (2016). A class of prediction-correction methods for time-varying convex optimization. *IEEE Transactions on Signal Processing, 64*(17), 4576-4591.
7. Simonetto, A., & Dall'Anese, E. (2017). Prediction-correction algorithms for time-varying constrained optimization. *IEEE Transactions on Signal Processing, 65*(20), 5481-5494.
8. Fazlyab, M., Paternain, S., Preciado, V. M., & Ribeiro, A. (2018). Prediction-correction interior-point method for time-varying convex optimization. *IEEE Transactions on Automatic Control, 63*(7), 1973-1986.
9. Simonetto, A. (2019). Dual prediction–correction methods for linearly constrained time-varying convex programs. *IEEE Transactions on Automatic Control, 64*(8), 3355-3361.
10. Simonetto, A., Dall'Anese, E., Paternain, S., Leus, G., & Giannakis, G. B. (2020). Time-varying convex optimization: Time-structured algorithms and applications. Proceedings of the IEEE, 108(11), 2032-2048.
11. Guo, D., Lin, X., Su, Z., Sun, S., & Huang, Z. (2018). Design and analysis of two discrete-time ZD algorithms for time-varying nonlinear minimization. *Numerical Algorithms, 77*, 23-36.
12. Paternain, S., Morari, M., & Ribeiro, A. (2019). Real-time model predictive control based on prediction-correction algorithms. In *2019 IEEE 58th Conference on Decision and Control (CDC)* (pp. 5285-5291). IEEE.
13. Bastianello, N., Simonetto, A., & Carli, R. (2019). Prediction-correction splittings for nonsmooth time-varying optimization. In *2019 18th European Control Conference (ECC)* (pp. 1963-1968). IEEE.
14. Bernstein, A., Dall'Anese, E., & Simonetto, A. (2019). Online primal-dual methods with measurement feedback for time-varying convex optimization. *IEEE Transactions on Signal Processing, 67*(8), 1978-1991.
15. Amidzadeh, M. (2023). Time-Varying Optimization with Optimal Parametric Functions. In *2023 62nd IEEE Conference on Decision and Control (CDC)* (pp. 2300-2305). IEEE.
16. Iserles, A. (2008). A first course in the numerical analysis of differential equations (No. 44). Cambridge university press.
17. Seierstad, A., & Sydsaeter, K. (1977). Sufficient conditions in optimal control theory. *International Economic Review*, 367-391.
18. Amidzadeh, M., Al-Tous, H., Tirkkonen, O., & Caire, G. (2020). Cellular network caching based on multipoint multicast transmissions. In *GLOBECOM 2020-2020 IEEE Global Communications Conference* (pp. 1-6). IEEE.
19. Amidzadeh, M., Al-Tous, H., Caire, G., & Tirkkonen, O. (2022). Cellular traffic offloading with optimized compound single-point unicast and cache-based multipoint multicast. In *2022 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 2268-2273). IEEE.

Appendix

A. Proof of Proposition 1.

We constitute the Lagrangian function, and obtain the Karush–Kuhn–Tucker conditions which read:

$$\begin{cases} b_m(\theta)\nabla_m f + \boldsymbol{\mathcal{J}}_m \boldsymbol{\lambda} = \mathbf{0}, \\ \sum_{m=1}^{M} \boldsymbol{h_m}(\boldsymbol{x}_m) - \boldsymbol{u} = \mathbf{0}, \end{cases} \tag{20}$$

where λ is the Lagrange multiplier. By considering Assumption 1 and taking derivative of the recent conditions w.r.t. θ, we get:

$$\begin{cases} \dot{b}_m(\theta)\nabla_m f + \Big(b_m(\theta)\nabla_m^2 f + \sum_{j=1}^{N} \lambda_j \nabla_m^2 h_{m,j}\Big)\dot{\boldsymbol{x}}_n(\theta) \\ \qquad + \boldsymbol{\mathcal{J}}_m \dot{\boldsymbol{\lambda}} = \mathbf{0}, \\ \sum_{m=1}^{M} \boldsymbol{\mathcal{J}}_m^\top \dot{\boldsymbol{x}}_m = \mathbf{0}, \end{cases} \tag{21}$$

By combining (20) and (21) and considering **Assumption 2**, the statement follows.