



Modified ADMM algorithm for solving proximal bound formulation of multi-delay optimal control problem with bounded control

K.A. Dawodu 

Abstract

This study presents an algorithm for solving optimal control problems with the objective function of the Lagrange-type and multiple delays on both the state and control variables of the constraints, with bounds on the control variable. The full discretization of the objective functional and the multiple delay constraints is carried out by using the Simpson numerical scheme. The discrete recurrence relations generated from the discretization of both the objective functional and constraints are used to develop the matrix operators, which satisfy the basic spectral properties. The primal-dual residuals of the algorithm are derived in order to ascertain the rate of convergence of the algorithm, which performs faster when relaxed with an accelerator variant in the sense of Nesterov. The direct numerical approach for handling the multi-delay control problem is observed to obtain an accurate result at a faster rate of convergence when over-relaxed with an accelerator variant. This research problem is limited to linear constraints and objective functional of the Lagrange-type and can address real-life models with multiple delays as applicable to quadratic optimization of intensity modulated radiation therapy planning. The novelty of this research paper lies in the method of discretization and its adaptation to handle linearly and proximal bound-constrained program formulated from the multiple delay optimal control problems.

AMS subject classifications (2020): 49M29, 93C43.

Keywords: Multi-delay; ADMM; Primal-dual; Proximal operator.

Received 2 March 2021; revised 4 June 2021; accepted 12 July 2021

Kazeem Adebowale Dawodu

Department of Mathematical Sciences, The Federal University of technology Akure,
Ondo-State, Nigeria. e-mail: kadawodu@futa.edu.ng

1 Introduction

Optimal control problems (OCPs) with delays in state and/or control variables play important roles in the modeling of real-life scenarios. Time delay effects in economical, physical, chemical, and biological processes involving optimal control systems cannot be neglected, especially when they involve the transmission of information between different parts of the system. These hereditary effects of constant delays have been deliberated upon by Xue and Duanin [27], Rihan and Anwar [21], Laarabi, Abta, and Hattaf, [14] and Bashier and Patidar [1]. Gollman, Kern and Maurer [9] did extensive work on mixed control-state inequality constraints with single delay on both state and delay variables with the initial and terminal boundary conditions in a general mixed form. Later, Gollmann and Maurer [10] extended their work to multiple time delays in the control and state variables with mixed control-state constraints. Olotu and Dawodu [18] and Olotu, Dawodu, and Yusuf [19] worked on one-dimensional and generalized control problems, respectively, with multiple delay constants and unbounded control variable (vector) using modified alternating direction method of multipliers (ADMM), while their earlier work in [17] on the delay proportional control system was done by the method of Quasi-Newton embedded augmented Lagrangian functional.

The Douglas–Rachford (D-R) splitting method called the ADMM (*Alternating Direction Method of Multipliers*) had been deployed by a plethora of authors in solving OCPs; though very few have deployed it to solving multi-delay OCPs with control bounds. The patronage is based on the fact that ADMM enjoys the strong convergence properties of the method of multipliers and the decomposability property of dual ascent as in the works of Boley [2], Boyd and Vandenberghe [4], Sun, Toh, and Yang [23], and Yao et al [28]; hence its relevance in solving decentralized and block-convex optimization problems like the large-scaled power or distributed systems cannot be overemphasized. Yao et al. [28] discovered that the ADMM could be used as a tool in solving a vehicle routing problem structured as an integer programming problem provided; the quadratic penalty terms used in ADMM can be reduced into simple linear functions. He and Yuan [13] and He, Hong, and Yuan [12] worked on the proximal Jacobian decomposition of the augmented Lagrangian method for the multiple-block separable convex minimization problem. It was observed that it has a strong affinity for ADMM.

Many authors have researched constrained convex quadratic problems with simple bounds. Voglis and Lagaris [24] was among the early authors that developed an algorithm for quadratic programming problems with box constrained using the *active set strategy* with applications mainly to circus tent, random and bi-harmonic problems. However, the scalability, simplicity, and potential of the ADMM in solving large-scaled structured convex optimization problems made it widely acceptable. The ADMM has the ability to solve the primal-dual feasibility updates in parallel and is well amenable to

the Gauss–Seidel accelerator or relaxation in order to improve its efficiency and rate of convergence as in Nesterov [16].

Pasquale and Gerardo [20] also reviewed the main results on global optimality conditions for establishing the global minimization of a quadratic problem with box constraints. However, Carreira-Perpiñán [5] worked on the proximal bound-constrained quadratic problem subject to simple box constraints. The ADMM algorithm was considered using the convex proximal operator by Moreau [15], Rockafellar [22], and Combettes and Pesquet in [6]. The algorithm was found to be particularly efficient in solving a collection of proximal operators that share a same quadratic form with a high rate of convergence and accuracy. Wu and Shang [25] further presented a global optimization method for solving general nonlinear programming problems subject to box constraints using a differential flow on the dual feasible space. The criteria for the global optimality and existence of complete solutions to the original problems were also given.

Xinmin et al. [26] researched the application of the ADMM algorithm to the constrained quadratic optimization of intensity modulated radiation therapy planning. A graph form convex optimization model was formulated as $\min h(y)$ such that $y = Ax$, $0 \leq x \leq u_x$, $l_y \leq y \leq u_y$, where the meaning and conditions for the bounds were well-defined. This approach involves the construction of a *proximal operator* and the formulation of the optimization problem in the graph form of the ADMM algorithm. In the implementation of the algorithm, a clinically acceptable solution was obtained with fewer memory requirements when compared with other optimization solvers.

2 Statement of problem

The generalized form of the multiple delay OCP with bounded (box) control is expressed below with quadratic continuous functional F and linear dynamical function f .

2.1 Optimal control model

Considering the multi-delay OCP with bounded control below:

$$\min J(x, u) = \frac{1}{2} \int_{t_0}^T F(t, x(t), u(t)) dt \quad (1)$$

$$\text{s.t. } \dot{x}(t) \leq Ax + Bu + \sum_{j=1}^d \alpha_j x(t - r_j) + \sum_{l=1}^e \beta_l u(t - q_l) \quad (2)$$

$$\text{s.t. } \dot{x}(t) \leq f(t, x(t), u(t), x^h(t - r_j), u^h(t - q_l)), \quad j(l) = 1, 2, \dots, d(e), \quad (3)$$

$$x(t_0) = x_0, \quad (4)$$

$$x(t) = \phi(t), \quad t_0 - r \leq t \leq t_0, \quad (5)$$

$$u(t) = \psi(t), \quad t_0 - q \leq t \leq t_0, \quad (6)$$

$$\gamma \leq u(t) \leq \sigma, \quad (7)$$

where

$$x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n, \quad \phi = (\phi_1, \phi_2, \dots, \phi_n)^T \in \mathbb{R}^n,$$

$$u = (u_1, u_2, \dots, u_m)^T \in \mathbb{R}^m,$$

$$x^h(t - r_j) = (x(t - r_1), x(t - r_2), \dots, x(t - r_d)) \in \mathbb{R}^{dn},$$

$$u^h(t - q_l) = (u(t - q_1), u(t - q_2), \dots, u(t - q_e)) \in \mathbb{R}^{em},$$

$$\psi = (\psi_1, \psi_2, \dots, \psi_m)^T \in \mathbb{R}^m,$$

$$\gamma = (\gamma_1, \dots, \gamma_m)^T \in \mathbb{R}^m,$$

$$\sigma = (\sigma_1, \dots, \sigma_m)^T \in \mathbb{R}^m,$$

$$r = \max\{r_j\}_{j=1}^d, \text{ and } q = \max\{q_l\}_{l=1}^e.$$

2.2 Assumptions and properties of the model

The underlining assumptions and properties of the functions are stated below:

- (i) The *time interval* $I = [t_0, T] \subseteq \mathbb{R}$ is a subset of real numbers \mathbb{R} ;
- (ii) The *state function* $x(t) : I \rightarrow X$ of the optimal control system is absolutely continuous on I provided the subset X is closed and bounded in \mathbb{R}^n ;
- (iii) The *control function* $u(t) : I \rightarrow U$ of the control system is piece-wise continuous and Lebesgue measurable on I provided the subset U is closed and bounded in \mathbb{R}^m ;
- (iv) The numbers r_1, r_2, \dots, r_d and q_1, q_2, \dots, q_e are the known *monotonically increasing positive delay constants* for the state and control variables, respectively; that is, $r_j \leq r_{j+1}$ for all $r_j \in \mathbb{R}$ for $j = 1, 2, \dots, d$ and $q_l < q_{l+1}$ for all $q_l \in \mathbb{R}$ for $l = 1, 2, \dots, e$;
- (v) The *initial (pre-shaped) functions* $\varphi(t) : [t_0 - r, t_0] \rightarrow \mathbb{R}^n$ and $\psi(t) : [t_0 - q, t_0] \rightarrow \mathbb{R}^m$ for the state and control variables, respectively, are known and piecewise continuous;
- (vi) The numbers $r = \max\{r_j\}_{j=1}^d$ and $q = \max\{q_l\}_{l=1}^e$ represent the values of the *maximum delay constants* for the state and control variables, respectively;

- (vii) The *cost functional* $F : [t_0, T] \times \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \cup \{\infty\}$ is nonlinear, piecewise continuous and measurable;
- (viii) The *constraint functional* $f : [t_0, T] \times \mathbb{R}^{(1+d)n} \times \mathbb{R}^{(1+e)m} \rightarrow \mathbb{R}^n$ is linear and piece-wise continuous;
- (ix) There exist an *admissible pair* $p = (x(\cdot), u(\cdot)) \in \Omega := ([t_0, T], \mathbb{R}^n) \times ([t_0, T], \mathbb{R}^m)$ that satisfies (1)–(7) of the multi-delay OCP, where the functional $J(x(\cdot), u(\cdot))$ is minimized by the the *admissible triple* $\Omega := (x(\cdot), u(\cdot), p(\cdot))$.

3 Methodology

A very simple discretization of (1) and (3) is obtained by means of the Simpsons and Euler numerical schemes, respectively. Recurrence relations generated are used in developing the respective large sparse matrix operators. The augmented Lagrangian functional with relaxation parameter (factor) is then used as an accelerator variant in the formulation of the modified ADMM (M-ADMM) algorithm. The spectral and convergence analyses are carried out to ensure the well-posedness of the algorithm.

3.1 Background and preliminaries

Suppose that the discrete time interval is given as $I_k = [t_k, t_{k+1}]$ by letting $t_k = t_0 + k\delta$ for $k = 0, 1, \dots, N$. Then the discretization operator, say f_x , maps each discrete point t_k in the discrete time interval $I_k \subseteq \mathbb{R}$ into each discrete (grid) point of the concatenated state vector $x_i^{(k)} \in \mathbb{R}^{nN}$ for all $i = 1, 2, \dots, n$, while the operator, say f_u , maps the points into each discrete point of the concatenated control vector $u_j^{(k)} \in \mathbb{R}^{m(N+1)}$ for all $j = 1, 2, \dots, m$. It is then expressed as

$$f_x : [t_0, T] \subseteq \mathbb{R} \longrightarrow x_i^{(k)} \in \bar{x}_i, \quad (8)$$

$$f_u : [t_0, T] \subseteq \mathbb{R} \longrightarrow u_j^{(k)} \in \bar{u}_j, \quad (9)$$

where $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n) \in \mathbb{R}^{nN}$ and $\bar{u} = (\bar{u}_1, \dots, \bar{u}_m) \in \mathbb{R}^{m(N+1)}$ for $\bar{x}_i = (x_i^{(1)}, \dots, x_i^{(N)}) \in \mathbb{R}^N$, and $\bar{u}_j = (u_j^{(0)}, \dots, u_j^{(N)}) \in \mathbb{R}^{N+1}$, respectively. Considering the existence of multiple delay constants $r_j \in \mathbb{R}$ (for $j = 1, 2, \dots, d$) and $q_l \in \mathbb{R}$ (for $l = 1, 2, \dots, e$) on the state and control variables, respectively. Then some basic Theorems are expressed below to aid in the process of discretization.

Theorem 1 (Rationality theorem). Given the real numbers $r_j, r_{j+1} > 0$ for $r_j < r_{j+1}$, then there exists a unique real number $\delta < 1$ such that the ratios of the numbers is a rational number \mathbb{Q} .

Proof. Let the delay constants be arranged in an increasing manner; that is, $r_j \leq r_{j+1}$ and $q_l \leq q_{l+1}$ for every j, l . It then implies that for every r_j and q_l , there exist integers $m_j \in \mathbb{Z}^+$ and $w_l \in \mathbb{Z}^+$ such that

$$\frac{r_j}{m_j} = \frac{\delta q_l}{\delta w_l} \in \mathbb{R}, \quad (10)$$

where δ is known as the step-length or interval length. This clearly shows that the delays, for both the state and control variables, are integer multiples of the step-length. \square

The proof of this transformation technique was suggested by Guinn [11] to derive first-order necessary conditions for unconstrained OCPs with pure state delays. To further elaborate on the assumptions above and the applications, we then introduce Theorem 2 below.

Theorem 2. Given any interval $[a, b]$, there exists a step-length $h \in \mathbb{R}^+$ such that each sub-interval is a constant multiple of the step-length.

Proof. Let the entire length of the interval $l = b - a$ be partitioned into m sub-intervals $[r_{j-1}, r_j]$ for $j = 1, 2, \dots, m$. Then there exists $h \in \mathbb{R}$ with $r_j = hv_j$ for all $v_j \in \mathbb{Z}^+$ such that the length of each sub-interval can be written as $l_j = r_j - r_{j-1}$.

$$\frac{l_j}{h} = \frac{(r_j - r_{j-1})}{h} = n_j + \delta_j \quad n_j \in \mathbb{Z}^+, \delta_j \in [0, 1).$$

Then,

$$\frac{(hv_j - hv_{j-1})}{h} = (v_j - v_{j-1}) = n_j + \delta_j \quad n_j \in \mathbb{Z}^+, \delta_j \in [0, 1). \quad (11)$$

Since $v_j - v_{j-1} \in \mathbb{Z}^+$ then $n_j + \delta_j \in \mathbb{Z}^+$ such that $\delta_j = 0$ or $\delta_j = 1$ for all $j = 1, 2, \dots, d$ (by Theorem 1). Since $\delta_j = 1 \notin [0, 1)$, then $\delta_j = 0$ for all $j = 1, 2, \dots, d$. By slotting $\delta_j = 0$ into (11), we then conclude that the length of each sub-interval $l_j = (r_j - r_{j-1})$ can be expressed as a multiple integer $v_j \in \mathbb{Z}^+$ of h expressed as $r_j = r_{j-1} + hv_j$, where $v_j - v_{j-1} = n_j$ for all $j = 1, 2, \dots, m$. \square

The essence of Theorems 1 and 2 above is to establish the fact that there exists a real number, referred to as *step-length*, $h \in \mathbb{R}$ that divides the entire state and control delay constants with multiples of real positive integers.

Consequent to 1 and 2, the multiple state and control variables are presented as follows:

$$x^{(k-v_j)} = \begin{cases} \phi(t_{k-v_j}), & k - v_j \leq 0; k = 0, 1, 2, \dots, v_j, t \in [t_0 - r, t_0], \text{ (known)} \\ x^{(k-v_j)}, & k - v_j > 0; k = (v_j + 1), \dots, N, t \in [t_0, T], \text{ (unknown)} \\ x(t_0) = x_0 \text{ given} \end{cases} \quad (12)$$

and

$$u^{(k-w_l)} = \begin{cases} \psi(t_{k-w_l}), & k - w_l < 0; k = 0, 1, 2, \dots, (w_l - 1), t \in [t_0 - q, t_0], \text{ (known)} \\ u^{(k-w_l)}, & k - w_l \geq 0 \text{ for } k = w_l, (w_l + 1), \dots, N, t \in [t_0, T], \text{ (unknown)} \end{cases} \quad (13)$$

where $r = r_d = \max\{r_j\}_{j=1}^d$ and $q = q_e = \max\{q_l\}_{l=1}^e$ are the largest delays on the state and control variables, respectively. In other words, the discretized delay state and control variables are represented in the form below:

$$x^{(k-v_j)} = \left(x_1^{(k-v_j)}, x_2^{(k-v_j)}, \dots, x_n^{(k-v_j)} \right)^T \in \mathbb{R}^n, \quad \text{for any } j, \quad (14)$$

$$u^{(k-w_l)} = \left(u_1^{(k-w_l)}, u_2^{(k-w_l)}, \dots, u_m^{(k-w_l)} \right)^T \in \mathbb{R}^m, \quad \text{for any } l, \quad (15)$$

while $x(t)$ and $u(t)$ are estimated within their respective delay intervals by the given delay functions $\phi(t)$ and $\psi(t)$, respectively, such that $\phi(t) \simeq \phi(t_{-k})$ and $\psi(t) \simeq \psi(t_{-k})$, where $t_k = t_0 - k\delta$ for $k = 1, 2, \dots, v_d, \dots, w_e$ if $v_d \leq w_e$. However, the largest state and control delay vectors for each iterate $k = 0, 1, \dots$ are represented below:

$$\hat{x} = \left(x^{(1-v_j)}, x^{(2-v_j)}, \dots, x^{(N-v_j)} \right)^T \in \mathbb{R}^{n(N-v_j)}, \quad \text{for all } j = 1, 2, \dots, d, \quad (16)$$

$$\hat{u} = \left(u^{(-w_l)}, u^{(1-w_l)}, \dots, u^{(N-w_l)} \right)^T \in \mathbb{R}^{m(N+1)}, \quad \text{for all } l = 1, 2, \dots, e. \quad (17)$$

3.2 Discretization of the objective functional

The objective function in (1) is of the form

$$\min J(x, u) = \frac{1}{2} \int_{t_0}^T (x^T P x + u^T Q u) dt, \quad (18)$$

where $x = (x_1, x_2, \dots, x_n)^T \in \mathbb{R}^n$, $u = (u_1, u_2, \dots, u_m)^T \in \mathbb{R}^m$, $P \in \mathbb{R}^{n \times n}$, and $Q \in \mathbb{R}^{m \times m}$. The full discretization of the objective function (1) using the Simpson integration formula gives

$$\min J(x, u) \simeq \frac{1}{2} \sum_{k=1}^{N-1} F(t_k, x^{(k)}, u^{(k)}) \quad (19)$$

and when expanded, it can be expressed as the quadratic function

$$\min_{\bar{x}, \bar{u}} \frac{1}{2} \bar{x}^T \bar{P} \bar{x} + \frac{1}{2} \bar{u}^T \bar{Q} \bar{u} + R, \quad (20)$$

where $R = \frac{\delta}{6}(x^{(0)})^T P x^{(0)} \in \mathbb{R}$ and the concatenated state and control variables are $\bar{x} = (x_1^{(1)}, x_1^{(2)}, \dots, x_n^{(N)})^T \in \mathbb{R}^{nN}$ and $\bar{u} = (u_1^{(0)}, u_1^{(1)}, \dots, u_m^{(N)})^T \in \mathbb{R}^{m(N+1)}$, respectively. However, the block-diagonal coefficient matrices $\bar{P} \in \mathbb{R}^{nN \times nN}$ and $\bar{Q} \in \mathbb{R}^{m(N+1) \times m(N+1)}$ are the block-matrix operators of the objective functional stated below as

$$\bar{P} = \begin{pmatrix} \frac{4\delta}{3}P & 0 & \dots & \dots & 0 \\ 0 & \frac{2\delta}{3}P & \ddots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{4\delta}{3}P & 0 \\ 0 & \dots & \dots & 0 & \frac{\delta}{3}P \end{pmatrix} \quad \text{and} \quad \bar{Q} = \begin{pmatrix} \frac{\delta}{3}Q & 0 & \dots & \dots & \dots & 0 \\ 0 & \frac{4\delta}{3}Q & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \frac{2\delta}{3}Q & \ddots & \dots & \vdots \\ \vdots & \dots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \dots & \dots & \ddots & \frac{4\delta}{3}Q & 0 \\ 0 & \dots & \dots & \dots & 0 & \frac{\delta}{3}Q \end{pmatrix}$$

The detailed derivation is in Appendix as in Olotu, Dawodu, and Yusuf [19]. However, in determining the properties of the derived matrix operators above, the following definitions are then introduced.

Definition 1 (Sylvester criterion). A square ($n \times n$) matrix is positive definite if it is real, symmetric with all the eigenvalues ($\lambda_i > 0$; $i = 1, 2, \dots, n$) or the leading principal minors ($M_i > 0$; $i = 1, 2, \dots, n$) are strictly positive.

Definition 2 (Corollary to Sylvester criterion). If the principal diagonal entries a_{ii} are the only nonzero entries of a square ($n \times n$) matrix such that $a_{ij} = 0$ for $i \neq j$, then the leading principal minors ($M_i > 0$; $i = 1, 2, \dots, n$) are given by $M_i = \prod_{j=1}^i a_{jj}$ for $i = 1, 2, \dots, n$.

Definition 3. Every positive definite matrix is invertible.

Sequel to Definitions 1 and 2, the constructed matrix-operator P is symmetric and positive definite since all the leading principal minors are strictly positive. Consequently, by Definition 3, the matrix P is invertible.

3.3 Discretization of the constraints

The Euler numerical scheme in (21) below is used for the discretization of the multi-delay constraints;

$$x^{(k+1)} = x^{(k)} + \delta f^k, \quad (21)$$

where $f^k = f(t_k, x^{(k)}, u^{(k)}, \hat{x}^{(k)}, \hat{u}^{(k)})$ and expressed in terms of the RHS of (3) as

$$f^k = Ax^{(k)} + Bu^{(k)} + \sum_{j=1}^d \alpha_j x^{(k-v_j)} + \sum_{l=1}^e \beta_l u^{(k-w_l)}, \quad k = 0, 1, 2, \dots, N-1, \tag{22}$$

where $\hat{x}^{(k)} = \{x^{(k-v_1)}, x^{(k-v_2)}, \dots, x^{(k-v_d)}\}$ and $\hat{u}^{(k)} = \{u^{(k-w_1)}, u^{(k-w_2)}, \dots, u^{(k-w_e)}\}$. Adapting (21) and (22) to (3) of the OCP yields the recurrence relation below:

$$\theta x^{(k)} - x^{(k+1)} + \omega u^{(k)} \leq -\delta \sum_{j=1}^d \alpha_j x^{(k-v_j)} - \delta \sum_{l=1}^e \beta_l u^{(k-w_l)}, \tag{23}$$

where $\theta = (I_n + A\delta) \in \mathbb{R}^{n \times n}$, $\omega = B\delta \in \mathbb{R}^{n \times m}$, $I_n \in \mathbb{R}^{n \times n}$ (identity) and $k = 0, 1, \dots, N-1$. Slotting the various values of k into (23) forms the linear inequality below:

$$\bar{A}\bar{x} + \bar{B}\bar{u} \leq \bar{E}\bar{x}^h + \bar{F}\bar{u}^h = C, \tag{24}$$

where \bar{A} is a multi-diagonal matrix with nN rows and nN columns (i.e., $\bar{A} \in \mathbb{R}^{nN \times nN}$); \bar{B} is also a diagonal matrix with nN rows and $m(N+1)$ columns (i.e., $\bar{B} \in \mathbb{R}^{nN \times m(N+1)}$), and \bar{x} is the concatenated vector of the state variable with dimension $nN \times 1$ while \bar{u} is the concatenated row vector of the control variable with dimension $m(N+1) \times 1$. The dimensions of the matrices \bar{E} , \bar{x}^h , \bar{F} and \bar{u}^h are $nN \times n(v_d+1)$, $n(v_d+1) \times 1$, $nN \times mw_e$ and $mw_e \times 1$, respectively. The respective structures of the concatenated vector-matrices are represented below:

$$\bar{A}\bar{x} = \begin{pmatrix} -I_n & 0 & 0 & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & 0 \\ \theta & -I_n & 0 & 0 & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ 0 & \theta & -I_n & 0 & 0 & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \bar{\alpha}_1 & \cdots & 0 & \theta & -I_n & 0 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \bar{\alpha}_d & \cdots & \bar{\alpha}_1 & \cdots & 0 & \theta & -I_n & \ddots & \ddots & \vdots \\ 0 & \bar{\alpha}_d & \cdots & \bar{\alpha}_1 & \cdots & 0 & \theta & -I_n & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \bar{\alpha}_d & \cdots & \bar{\alpha}_1 & \cdots & 0 & \theta & -I_n \end{pmatrix} \begin{pmatrix} x^{(1)} \\ \vdots \\ x^{(v_1+1)} \\ \vdots \\ x^{(v_2-v_1)} \\ \vdots \\ x^{(v_d-v_{d-1})} \\ \vdots \\ x^{(v_d+1)} \\ \vdots \\ x^{(N-1)} \\ x^{(N)} \end{pmatrix},$$

$$\bar{B}\bar{u} = \begin{pmatrix} \omega & 0 & 0 & \dots & \dots & \dots & \dots & 0 \\ 0 & \omega & 0 & 0 & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ \bar{\beta}_1 & \dots & 0 & \omega & 0 & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ \bar{\beta}_e & \dots & \bar{\beta}_1 & \dots & 0 & \omega & 0 & \dots \\ 0 & \bar{\beta}_e & \dots & \bar{\beta}_1 & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ 0 & \dots & 0 & \bar{\beta}_e & \dots & \bar{\beta}_1 & \dots & 0 & \omega & 0 \end{pmatrix} \begin{pmatrix} u^{(0)} \\ \vdots \\ x^{(v_1+1)} \\ \vdots \\ u^{(w_2-w_1)} \\ \vdots \\ u^{(w_e+1)} \\ \vdots \\ u^{(N)} \end{pmatrix},$$

$$\bar{E}\bar{x}^h = \begin{pmatrix} -\theta & \dots & -\bar{\alpha}_1 & \dots & -\bar{\alpha}_s & \dots & -\bar{\alpha}_{d-1} & -\bar{\alpha}_d \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ -\bar{\alpha}_1 & \dots & -\bar{\alpha}_s & \dots & -\bar{\alpha}_{d-1} & -\bar{\alpha}_d & 0 & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ -\bar{\alpha}_s & \dots & -\bar{\alpha}_{d-1} & -\bar{\alpha}_d & \dots & 0 & \dots & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ -\bar{\alpha}_{d-1} & -\bar{\alpha}_d & \dots & 0 & \dots & \dots & \dots & \vdots \\ -\bar{\alpha}_d & \dots & 0 & \dots & \dots & \dots & \dots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \begin{pmatrix} x^{(0)} \\ x^{(-1)} \\ \vdots \\ x^{(v_2-v_3)} \\ \vdots \\ \vdots \\ \vdots \\ x^{(v_d-v_{d-1})} \\ \vdots \\ x^{(1-v_d)} \\ x^{(-v_d)} \end{pmatrix}$$

and

$$\bar{F}\bar{u}^h = \begin{pmatrix} -\bar{\beta}_1 & \dots & -\bar{\beta}_2 & \dots & -\bar{\beta}_s & \dots & -\bar{\beta}_{e-1} & -\bar{\beta}_e \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \\ -\bar{\beta}_1 & \dots & -\bar{\beta}_s & \dots & -\bar{\beta}_{e-1} & -\bar{\beta}_e & 0 & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ -\bar{\beta}_s & \dots & -\bar{\beta}_{e-1} & -\bar{\beta}_e & \dots & 0 & \dots & \vdots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ -\bar{\beta}_{e-1} & -\bar{\beta}_e & \dots & 0 & \dots & \dots & \dots & \vdots \\ -\bar{\beta}_e & \dots & 0 & \dots & \dots & \dots & \dots & \vdots \\ 0 & \dots & \dots & \dots & \dots & \dots & \dots & \vdots \\ \vdots & \dots & \dots & \dots & \dots & \dots & \dots & 0 \end{pmatrix} \begin{pmatrix} x^{(0)} \\ u^{(-1)} \\ \vdots \\ u^{(w_2-w_3)} \\ \vdots \\ \vdots \\ \vdots \\ u^{(w_e-w_{e-1})} \\ \vdots \\ u^{(1-w_e)} \\ u^{(-w_e)} \end{pmatrix},$$

where the delay coefficients $\bar{\alpha}_k$ and $\bar{\beta}_k$ are defined below as

$$[\bar{\alpha}_k] = \begin{cases} -\delta\alpha_k, & 1 \leq k \leq d, \\ 0 & \text{elsewhere} \end{cases} \quad (25)$$

and

$$[\bar{\beta}_k] = \begin{cases} -\delta\beta_k, & 1 \leq k \leq e, \\ 0 & \text{elsewhere.} \end{cases} \quad (26)$$

However, the entries $[\bar{a}_{ij}] \in \bar{A}$, $[\bar{b}_{ij}] \in \bar{B}$, $[\bar{e}_{ij}] \in \bar{E}$, and $[\bar{f}_{ij}] \in \bar{F}$ of the various matrix structures are described below as follows:

$$[\bar{a}_{ij}] = \begin{cases} -I_n, & i = j \quad 1 \leq i \leq N, \\ \theta, & j = i - 1 \quad 2 \leq i \leq N, \\ \bar{\alpha}_k, & j = i - 1 - v_k \quad 2 + v_k \leq i \leq N, \quad k = 1, 2, \dots, d, \\ 0 & \text{elsewhere,} \end{cases} \quad (27)$$

$$[\bar{b}_{ij}] = \begin{cases} \omega, & i = j \quad 1 \leq i \leq N, \\ \bar{\beta}_k, & j = i - w_k \quad 1 + w_k \leq i \leq N, \quad k = 1, 2, \dots, e, \\ 0 & \text{elsewhere,} \end{cases} \quad (28)$$

$$[\bar{e}_{ij}] = \begin{cases} -\theta, & i = j = 1, \\ -\bar{\alpha}_k, & j = v_k + 1 - i \quad 1 \leq i \leq v_k, \quad k = 1, 2, \dots, d, \\ 0 & \text{elsewhere,} \end{cases} \quad (29)$$

$$[\bar{f}_{ij}] = \begin{cases} -\sigma, & i = j = 1, \\ -\bar{\beta}_k, & j = w_k + 1 - i \quad 1 \leq i \leq w_k, \quad k = 1, 2, \dots, e, \\ 0 & \text{elsewhere.} \end{cases} \quad (30)$$

It is imperative to note that the real symmetric and positive definite properties of the discretized matrices \bar{P} , \bar{Q} , \bar{A} , and \bar{B} affect the well-posedness of the algorithm and the selection of the optimal (relaxation and penalty) parameters; hence the bases for the assumption.

3.4 Splitting method

The D-R splitting technique, also known as ADMM, is a special case of the splitting technique and can be traced back to work done on monotone operators and operator splitting methods into the general Hilbert spaces. Eckstein and Ferris [7] showed in turn that DR splitting is also a special case of the

proximal point algorithm applicable to optimal control systems. Suppose there exist two separable closed convex functions (or proximal operators) $f(x)$ and $g(x)$ that are non-expansive and not necessarily symmetric. Then the iterates of the D-R splitting algorithm for $\min f(x) + g(x)$ starting with any ζ^0 and repeating for $k = 0, 1, \dots$ are stated as

$$\begin{aligned}\zeta^{k+1} &\leftarrow \zeta^k + \text{prox}_g(2x^k - \zeta^k) - x^k, \\ x^{k+1} &\leftarrow \text{prox}_f(\zeta^{k+1}),\end{aligned}$$

where x^k converges to the solution of $0 \in \partial f(x) + \partial g(x)$, if the solution exists. The D-R Algorithm can be re-structured by introducing a new variable $w^k = x^k - \zeta^k$ if when substituted into the D-R splitting Algorithm above, it yields its equivalent iterates expressed below as

$$\begin{aligned}u^{k+1} &\leftarrow \text{prox}_g(x^k + w^k), \\ x^{k+1} &\leftarrow \text{prox}_f(u^{k+1} - w^k), \\ w^{k+1} &\leftarrow w^k + x^{k+1} - u^{k+1},\end{aligned}$$

where $x = \text{prox}_f(u - w)$ satisfies $0 \in \partial f(x) + \partial g(x)$. However, the equivalent D-R Algorithm can be amended by introducing the Nesterov-type accelerated variants or relaxation factor $\alpha \in [0, 2]$ in order to improve the rate of convergence of the iterates. The relaxed iterates is arrived at by replacing u^{k+1} with $\alpha u^{k+1} + (1 - \alpha)x^k$ in the sense of Nesterov [16]. The derived algorithm is stated as

$$\begin{aligned}\text{Initialize at any } \zeta^0, x^0, \text{ fixed } \alpha \text{ and repeat for } k = 0, 1, \dots, \\ u^{k+1} &\leftarrow \text{prox}_g(x^k + w^k), \\ x^{k+1} &\leftarrow \text{prox}_f(\alpha u^{k+1} + (1 - \alpha)x^k - w^k), \\ w^{k+1} &\leftarrow w^k + x^{k+1} - (\alpha u^{k+1} - \alpha x^k),\end{aligned}$$

where it is an over-relaxation for $1 < \alpha < 2$, under-relaxation for $0 < \alpha < 1$ and reduces to its non-relaxed form if $\alpha = 1$. The extension of the splitting method in the computations of the updates of the dual and adjoint variables of a separable convex constrained problem of the form $\min f(x_1) + g(x_2)$ s.t. $h(x_1, x_2) = 0$ requires a sequential minimization of the augmented Lagrangian functional $L_\rho(x_1, x_2, \zeta)$; hence the name alternating direction method of multipliers (ADMM).

The three steps iterates (updates) for the ADMM algorithm are as follows:

$$x^{k+1} \leftarrow \arg \min_{x_1} \left(f(x_1) + \zeta^T h(x_1, x_2^k) + \frac{\rho}{2} \| h(x_1, x_2^k) \|_2^2 \right), \quad (31)$$

$$x_2^{k+1} \leftarrow \arg \min_{x_2} \left(g(x_2) + \zeta^T h(x_1^{k+1}, x_2) + \frac{\rho}{2} \| h(x_1^{k+1}, x_2) \|_2^2 \right), \quad (32)$$

$$\zeta^{k+1} \leftarrow \zeta^k + \rho h(x_1^{k+1}, x_2^{k+1}), \quad (33)$$

where $\rho > 0$ is the penalty parameter and $x_1^{k+1} \equiv \alpha x_1^{k+1} - (1-\alpha)x_2^k$ for any accelerator variant $\alpha \in [0, 2]$ puts the algorithm in a relaxed mode. The ADMM is a more general method for handling all kinds of convex optimization problems which includes the ℓ_1 -regularization of the for $\min \sum_i f_i(w) + \|w\|$, which other traditional methods such as gradient descent, Newton, and so on cannot handle. The ease with which the ADMM can be implemented in parallel for large-scaled data at a faster rate gives it an edge over other algorithms.

3.5 Analysis on bounded control problems

In a general box constrained nonlinear programming problem

$$\min \{f(x) \mid x \in \Omega \subseteq \mathbb{R}^n\}, \tag{34}$$

where $\Omega = \{x \in \mathbb{R}^n \mid l \leq x \leq u\}$ is a feasible space and a closed convex subset of \mathbb{R}^n , $f(x)$ and a strongly convex quadratic function that is twice continuously differentiable in \mathbb{R}^n , l and u are the lower and upper bounds, respectively. Carreira-Perpiñán [5] and Combettes and Pesquet [6] extended the theory of proximal operator to a proximal bound-constrained quadratic program (QP) (34) with the Lagrange formulation presented as

$$L_\rho(x, z, \xi) = f(x) + g(z) + \frac{\rho}{2} \|x - z + \xi\|_2^2 \quad \text{s.t.} \quad x = z, \tag{35}$$

where $z \in \mathbb{R}^n$ is the slack (auxiliary) variable, $\xi = \frac{\lambda}{\rho} \in \mathbb{R}^n$ is the scaled dual (adjoint) vector, $\rho > 0$ is the penalty parameter, and $g(z) = \frac{\mu}{2} \|z - v\|_2^2$ is the convex term (or indicator function for the nonnegative orthant). The proximal M-ADMM formulation is then expressed in the form

$$x^{k+1} \leftarrow \mathbf{prox}_{f,\rho}(z^k - \xi^k) = \arg \min_x \left(f(x) + \frac{\rho}{2} \|x - z^k + \xi^k\|_2^2 \right), \tag{36}$$

$$z^{k+1} \leftarrow \mathbf{prox}_{g,\rho}(x^{k+1} - \xi^k) = \arg \min_z \left(\frac{\mu}{2} \|z - v\|_2^2 + \frac{\rho}{2} \|x^{k+1} - z + \xi^k\|_2^2 \right) \tag{37}$$

$$\text{s.t.} \quad l \leq z \leq u,$$

$$\xi^{k+1} \leftarrow \xi^k + x^{k+1} - z^{k+1}. \tag{38}$$

However, there exist three possible cases of the location of the parabola vertex of the objective functional upon which the optimal solution is determined. See sketches below for illustration. Each of the cases indicated in Figures 1, 2, and 3 represents the locations of the parabola vertex z_i^* within the bounds, below the lower bound, and above the upper bound, respectively, while Figure 4 represents the median of the three points. Therefore, the optimal solution defined on the feasible space is the median of each of the i th component ($i = 1, 2, \dots, N$) of the three vectors; lower bound l , upper bound

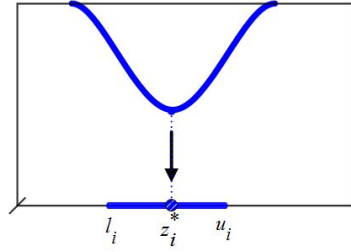


Figure 1: Vertex within bounds: $l_i \leq z_i \leq u_i$

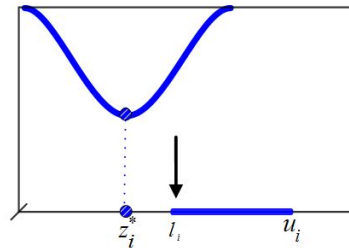


Figure 2: Vertex below lower bound: $z_i \leq l_i$

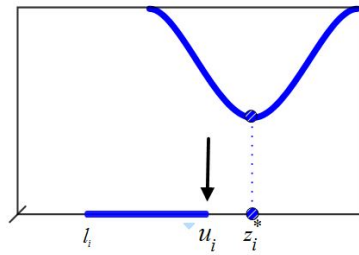


Figure 3: Vertex above upper bound: $u_i \leq z_i$

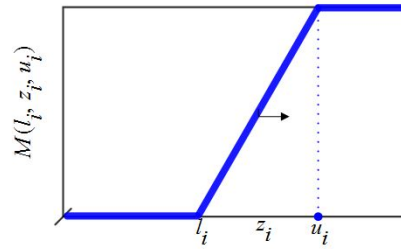


Figure 4: Median points: $M(l_i, z_i, u_i)$

u , and the parabola vertex z_i^* . The choice of the next iterate \bar{z} is determined as follows:

$$\bar{z}_i = \begin{cases} M(z_i^*, l_i, u_i) = \text{Min}[\text{Max}(z_i^*, l_i), u_i] = \text{Min}[l_i, u_i] = l_i, & z_i^* \leq l_i, \\ M(l_i, z_i^*, u_i) = \text{Min}[\text{Max}(l_i, z_i^*), u_i] = \text{Min}[z_i^*, u_i] = z_i^*, & l_i \leq z_i^* \leq u_i, \\ M(l_i, u_i, z_i^*) = \text{Min}[\text{Max}(l_i, u_i), z_i^*] = \text{Min}[u_i, z_i^*] = \bar{u}_i, & z_i^* \geq u_i. \end{cases} \quad (39)$$

3.6 M-ADMM algorithm formulation

The original ADMM formulation is modified to accommodate both the state and control variables before imposing the Karush–Khun–Tucker (KKT) optimality conditions. The Gauss–seidel accelerator variant is introduced to improve the convergence properties of the algorithm as illustrated in the convergence analysis below. The re-formulated compact convex quadratic optimization problem is then stated as

$$\text{Min} J(\bar{x}, \bar{u}) = \frac{1}{2} \bar{x}^T P \bar{x} + \frac{1}{2} \bar{u}^T Q \bar{u} + R \quad (40)$$

$$\text{s.t. } G(x, u) = \bar{A} \bar{x} + \bar{B} \bar{u} - C \leq 0, \quad (41)$$

$$\bar{\gamma} \leq \bar{u} \leq \bar{\sigma}, \quad (42)$$

$$\bar{\sigma} - \hat{u} \geq 0, \quad (43)$$

$$\bar{\gamma} + \hat{u} \geq 0, \quad (44)$$

where

$$\begin{aligned} \bar{x} &= (x^{(1)}, x^{(2)}, \dots, x^{(N)}) \in \mathbb{R}^{nN}, \quad \bar{u} = (u^{(0)}, u^{(1)}, \dots, u^{(N)}) \in \mathbb{R}^{m(N+1)}, \\ \bar{A} &\in \mathbb{R}^{nN \times nN}, \quad \bar{C} \in \mathbb{R}^{nN}, \quad \bar{B} \in \mathbb{R}^{nN \times m(N+1)}, \\ \bar{P} &\in \mathbb{R}^{nN \times nN}, \quad \text{and } \bar{Q} \in \mathbb{R}^{m(N+1) \times m(N+1)}, \end{aligned}$$

while $\bar{\gamma} = (1/\sqrt{N} + 1)(\gamma_1^{(1)}, \gamma_1^{(2)}, \dots, \gamma_1^{(N+1)}, \gamma_2^{(1)}, \dots, \gamma_m^{(N+1)}) \in \mathbb{R}^{m(N+1)}$ and $\bar{\sigma} = (1/\sqrt{N} + 1)(\sigma_1^{(1)}, \sigma_1^{(2)}, \dots, \sigma_1^{(N+1)}, \sigma_2^{(1)}, \dots, \sigma_m^{(N+1)}) \in \mathbb{R}^{m(N+1)}$ for all $\gamma_i \in \gamma$ and $\sigma_i \in \sigma$, respectively. In addition, \bar{P} and \bar{Q} are the real, symmetric, and positive-definite matrix operators.

3.7 Proximal-bound formulation

In formulating the *convex-proximal bound* constrained program for the discretized multi-delay OCP in (40)–(42), the Carreira-Perpiñá approach [5] is applied. Suppose that the control bound (42) is re-formulated into two equations, $\bar{\sigma} - \bar{u} \geq 0$ and $\bar{\gamma} + \bar{u} \geq 0$ such that

$$\begin{aligned} 0 &\leq \frac{\mu}{2} \left\| \underbrace{(\sigma - \gamma)}_y - \underbrace{(v_1 + v_2)}_v \right\|_2^2 \\ &\leq \frac{\mu}{2} \left\| \bar{\sigma} - \bar{u} - v_1 \right\|_2^2 + \frac{\mu}{2} \left\| \bar{\gamma} + \bar{u} - \xi_2 \right\|_2^2 + \frac{\mu}{2} \left\| y - v \right\|_2^2, \end{aligned} \quad (45)$$

where $g(y) = \frac{\mu}{2} \left\| y - v \right\|_2^2$ is a control bound function defined for $\bar{u} \geq y$ and $\bar{\gamma} \leq y \leq \bar{\sigma}$; while v is the slack variable replacing the inequalities and μ is the penalty parameter. Therefore, the associated augmented Lagrangian functional for the quadratic formulation is expressed as

$$\begin{aligned} L_{\rho_1, \rho_2}(\bar{x}, \bar{u}, z, v, y, \xi_1, \xi_2, \mu) &= \frac{1}{2} \bar{x}^T P \bar{x} + \frac{1}{2} \bar{u}^T Q \bar{u} + R + \frac{\mu}{2} \left\| y - v \right\|_2^2 \\ &\quad + \frac{\rho_1}{2} \left\| A \bar{x} + B \bar{u} - C + z + \xi_1 \right\|_2^2 \\ &\quad + \frac{\rho_2}{2} \left\| \bar{u} - y + \xi_2 \right\|_2^2 \\ \text{s.t. } & z, v \geq 0 \text{ and } \bar{\gamma} \leq y \leq \bar{\sigma}, \end{aligned} \quad (46)$$

with the M-ADMM iterations stated as

$$\begin{aligned}\bar{x}^{k+1} &\leftarrow \arg \min_{\bar{x}} \left\{ \frac{1}{2} \bar{x}^T P \bar{x} + \frac{\rho_1}{2} \| A \bar{x} + B \bar{u}^k - C + z^k + \xi_1^k \|_2^2 \right\}, \\ \bar{u}^{k+1} &\leftarrow \arg \min_{\bar{u}} \left\{ \frac{1}{2} \bar{u}^T Q \bar{u} + \frac{\rho_1}{2} \| A \bar{x}^{k+1} + B \bar{u} - C + z^k \|_2^2 + \frac{\rho_2}{2} \| \bar{u} - y^k + \xi_2^k \|_2^2 \right\}, \\ z^{k+1} &\leftarrow \arg \min_z \left\{ \frac{\rho_1}{2} \| A \bar{x}^{k+1} + B \bar{u}^{k+1} - C + z + \xi_1^k \|_2^2 \right\} \text{ s.t. } z \geq 0, \\ v^{k+1} &\leftarrow \arg \min_v \left\{ \frac{\mu}{2} \| y - v \|_2^2 \right\} \text{ s.t. } v \geq 0,\end{aligned}$$

$$\begin{aligned}y^{k+1} &\leftarrow \arg \min_y \left\{ \frac{\mu}{2} \| y - v \|_2^2 + \frac{\rho_2}{2} \| \bar{u}^{k+1} - y + \xi_2^k \|_2^2 \right\} \text{ s.t. } \bar{\gamma} \leq y \leq \bar{\sigma}, \\ \xi_1^{k+1} &= \xi_1^k + (A \bar{x}^{k+1} + B \bar{u}^{k+1} - C + z^{k+1}), \\ \xi_2^{k+1} &= \xi_2^k + (\bar{u}^{k+1} - y^{k+1}),\end{aligned}$$

where \bar{x} and \bar{u} are the primal variables, λ is the Lagrange multiplier, $\rho_1, \rho_2 > 0$ are the penalty parameters, $\|\cdot\|_2$ is the euclidean (spectral) norm of a vector (matrix) argument, $\xi_i = \lambda_i/\rho_i$, $i = 1, 2$ are the scaled dual vectors, y is the introduced auxiliary (slack) vector, and $l_+(z)$ is the indicator function for the non-negative orthants defined as $l_+(y) = \frac{\mu}{2} \| y - v \|_2^2$ and ∞ otherwise.

3.8 Derivations of the update formulas

Applying the KKT optimality conditions on the augmented Lagrangian functional (46) for the sequential minimization of the variables requires updating all the critical variables as indicated in the ADMM iterations. Moreover the update formulas of the adjoint variables will be accelerated using the Gauss–Seidel relaxation factor $\alpha \in [0, 2]$ as clearly illustrated in the works of Nesterov [16] and Ghadimi et al. [8]; though this work considers an over-relaxation where $\alpha \in (1, 2]$.

State-update - \bar{x} : Setting $\partial_{\bar{x}} L(\bar{x}, \cdot) = 0$ yields

$$\bar{x}^{k+1} = -\rho_1 (P + \rho_1 A^T A)^{-1} A^T (B \bar{u}^k - C + z^k + \xi_1^k) \quad (\bar{x} - \text{update}), \quad (47)$$

where $(P + \rho_1 A^T A)^{-1}$ is invertible since P is real, symmetric, and positive definite and $A^T A$ is positive semi-definite sequel to Definitions 1, 2, and 3.

Control-update - \bar{u} : Setting $\partial_{\bar{u}} L(\bar{x}^{k+1}, \bar{u}, \cdot) = 0$ and replacing $A \bar{x}^{k+1}$ with $h(\bar{x}^{k+1}, \bar{u}^k, \alpha) = \alpha A \bar{x}^{k+1} - (1 - \alpha)(B \bar{u}^k - C + z^k)$ to relax the algorithm with the relaxation factor yield

$$\bar{u}^{k+1} = -\rho(\bar{Q} + \rho \bar{B}^T \bar{B})^{-1} \bar{B}^T [\alpha (\bar{A} \bar{x}^{k+1} - C + z^k) - (1 - \alpha) \bar{B} \bar{u}^k + v^k]. \quad (48)$$

Auxiliary (Slack)-update - \bar{z} : Setting $\partial_z L(\cdot^{k+1}, z, \cdot^k) = 0$ and relaxing the algorithm by replacing $A\bar{x}^{k+1}$ with $h(\bar{x}^{k+1}, \bar{u}^k, \alpha)$ yields

$$z^* = -\alpha(A\bar{x}^{k+1} + B\bar{u}^{k+1} - C) - (1 - \alpha)[B(\bar{u}^{k+1} - \bar{u}^k) - z^k] - \xi_1^k. \quad (49)$$

Therefore,

$$\begin{aligned} z^{k+1} &= \max \{0, z^*\} \\ &= \max \left\{ 0, -\alpha(A\bar{x}^{k+1} + B\bar{u}^{k+1} - C) - (1 - \alpha)[B(\bar{u}^{k+1} - \bar{u}^k) - z^k] - \xi_1^k \right\}. \end{aligned} \quad (50)$$

Updating the parabola vertex - y : Setting $\partial_y L(\cdot^{k+1}, y, \cdot^k) = 0$ in the derivation of the parabola vertex y in the box constraint yields

$$\partial_y L(\cdot^{k+1}, y, \cdot^k) = \mu(y - v) - \rho_2(\bar{u} - y + \xi_2) = 0 \quad (51)$$

$$y^* = \frac{\mu v^{k+1} + \rho_2(\bar{u}^{k+1} + \xi_2^k)}{(\mu + \rho_2)} \quad \text{s.t.} \quad \bar{\gamma} \leq y^* \leq \bar{\sigma}, \quad (52)$$

where for each component y_i^* of $y^* = (y_1, y_2, \dots, y_N)$, the minimum of the upward parabola is defined in the sub-interval $[\bar{\gamma}_i, \bar{\sigma}_i]$ such that the solution is the median $[\bar{\gamma}_i, y_i^*, \bar{\sigma}_i]$ of $\bar{\gamma}_i, \bar{\sigma}_i$ and each component y_i^* of the parabola vertex derived above is specifically defined as

$$y_i^* = \frac{\mu v_i^{k+1} + \rho_2(\bar{u}_i^{k+1} + \xi_{2i}^k)}{(\mu + \rho_2)} \quad \text{s.t.} \quad \bar{\gamma}_i \leq y_i^* \leq \bar{\sigma}_i \quad \text{and element-wise for all } i. \quad (53)$$

Therefore,

$$\begin{aligned} y^{k+1} &= \text{Min} \left[\bar{\gamma}, \frac{\mu v^{k+1} + \rho_2(\bar{u}^{k+1} + \xi_2^k)}{(\rho_2 + \mu)}, \bar{\sigma} \right], \\ &= \text{Min}[\text{Max}(\bar{\gamma}, y^*), \bar{\sigma}], \end{aligned} \quad (54)$$

with the steps involving element-wise thresholding operations. Considering the extension of the above concept of proximal bound operator on the OCP, let the state variable $\bar{x} \in \mathbb{R}^N$ belong to a set of admissible state functions X (written $\bar{x} \in X \subseteq \mathbb{R}^\infty$) and let the control variable $\bar{u} \in \mathbb{R}^{(N+1)}$ belong to the set of admissible control functions U (written $\bar{u} \in U \subseteq \mathbb{R}^\infty$). We then assume that $v^{(\bar{\gamma})} = (\bar{x}^*, \bar{\gamma})$ and $v^{(\bar{\sigma})} = (\bar{x}^*, \bar{\sigma})$ are the points on the state-control coordinates (\bar{x}, \bar{u}) for which the control is bounded below and above by $\bar{\gamma} \in \mathbb{R}^{(N+1)}$ and $\bar{\sigma} \in \mathbb{R}^{(N+1)}$, respectively. If there exists any point on the i th component of the control vector ($i = 1, 2, \dots, (N+1)$), then $y_i \in [\bar{\gamma}_i, \bar{\sigma}_i]$ is such that $v^{(y_i)} = (\bar{x}_i^*, y_i)$ is the point on the (\bar{x}, \bar{u}) trajectory that prescribes the parabola vertex of the objective function $f(\bar{x}, \bar{u})$ at its minimum. Therefore, the median M of $v^{(\bar{\gamma}_i)}$, $v^{(\bar{\sigma}_i)}$ and the parabola vertex $v^{(y_i)}$ are defined as follows:

$$(\bar{x}_i^*, y_i^*) = \begin{cases} \text{Min}[\text{Max}(v^{(\bar{\gamma}_i)}, v^{(y_i)}), v^{(\bar{\sigma}_i)}] = v^{(y_i)}, & v^{(y_i)} \in [v^{(\bar{\gamma}_i)}, v^{(\bar{\sigma}_i)}], \\ \text{Min}[\text{Max}(v^{(\bar{u}_i)}, v^{(\bar{\gamma}_i)}), v^{(\bar{\sigma}_i)}] = v^{(\bar{\gamma}_i)}, & v^{(y_i)} \leq v^{(\bar{\gamma}_i)}, \\ \text{Min}[\text{Max}(v^{(\bar{\gamma}_i)}, v^{(\bar{\sigma}_i)}), v^{(y_i)}] = v^{(\bar{\sigma}_i)}, & v^{(y_i)} \geq v^{(\bar{\sigma}_i)}, \end{cases} \quad (55)$$

where $\bar{x} = \bar{x}_i^*$ and $\bar{u} = y_i^*$ are the optimal state and control variables, respectively, within the feasible space. Therefore, at $\bar{x} = \bar{x}^{(k+1)}$, the \bar{u} -update is defined as

$$(\bar{x}^{(k+1)}, y^{k+1}) = M [v^{(\bar{\gamma}_i)}, v^{(\bar{\sigma}_i)}, v^{(y_i)}], \text{ element-wise for all } i \in \{1, 2, \dots, (N+1)\}, \quad (56)$$

where $v^{(y_i)} = (\bar{x}_i^*, y_i^*) \equiv (\bar{x}_i^*, y_i^{k+1})$ such that

$$y_i^{k+1} = \frac{\mu v_i^{k+1} + \rho_2(\bar{u}_i^{k+1} + \xi_{2i}^k)}{(\rho_2 + \mu)}, \quad \bar{x}_i^* = \bar{x}_i^{(k+1)}, \quad u_i = y_i, \quad \text{for all } i \in \{1, 2, \dots, (N+1)\}. \quad (57)$$

For further illustration on the analysis of the proximal bound and its effects on the objective functional value, we then consider the figures below.

Suppose that (\bar{x}^*, y^*) is the vertex parabola for which $f(\bar{x}, \bar{u})$ is at minimum. Then

$$F(\bar{x}^*, y^*) \leq F(\bar{x}^*, \bar{u}), \text{ s.t. } \bar{x}^* \in X, \text{ for all } \bar{u} \in U \text{ and } y^* \in [\bar{\gamma}, \bar{\sigma}].$$

Assume that there exists a bound $[\bar{\gamma}, \bar{\sigma}] \subseteq U$ on the control variable, simply called bounded (box) constraint, for which the objective functional value $F(\bar{x}, \bar{u})$ is at minimum; then in the process of the algorithm, the three expected cases are stated as

$$F(\bar{x}^*, y^*) = \begin{cases} F(\bar{x}^*, y^*), & \text{for all } \bar{u} \in [\bar{\gamma}, \bar{\sigma}] \subseteq U & \text{Case I} \\ F(\bar{x}^*, \bar{\gamma}), & \text{for all } \bar{u} \leq \bar{\gamma} \subseteq U & \text{Case II} \\ F(\bar{x}^*, \bar{\sigma}), & \text{for all } \bar{u} \geq \bar{\sigma} \subseteq U & \text{Case III} \end{cases} \quad (58)$$

Case I: $\bar{\gamma} \leq \bar{u} \leq \bar{\sigma}$

In Figure 5, the vertex point (\bar{x}^*, \bar{u}^v) that falls within the control bound is same as the optimal state-control trajectories (\bar{x}^*, \bar{u}^*) such that $F(\bar{x}, \bar{u}^*) = F(\bar{x}, \bar{u}^v)$. The functional Ω within the region of feasibility is then defined as

$$\Omega^* = \{F(\bar{x}^*, \bar{u}^v) \mid \bar{x}^* \in X \text{ and } \bar{u} \in [\bar{\gamma}, \bar{\sigma}]\}. \quad (59)$$

Case II: $\bar{u} \leq \bar{\gamma}$

In Figure 6, the vertex point (\bar{x}^*, \bar{u}^v) that falls below the lower control bound is same as the optimal state-control trajectories $(\bar{x}^*, \bar{\gamma})$ such that $F(\bar{x}, \bar{u}^*) = F(\bar{x}, \bar{\gamma})$. The functional Ω within the region of feasibility is then defined as

$$\Omega^* = \{F(\bar{x}^*, \bar{\gamma}) \mid \bar{x}^* \in X \text{ and } \bar{u} \leq \bar{\gamma} \text{ for all } \bar{u} \in U\} \quad (60)$$

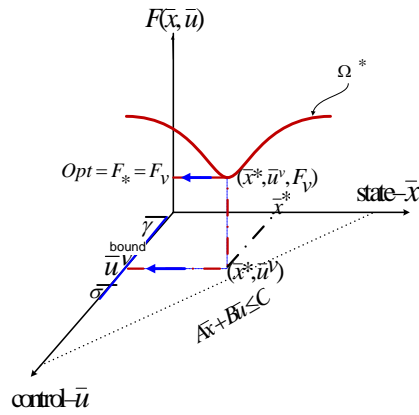


Figure 5: Control within bounds: $\bar{u}^v \leq \bar{\sigma}$

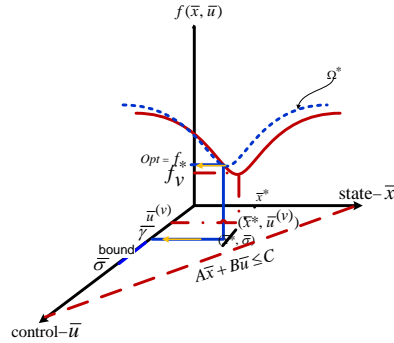


Figure 6: control below lower bound: $\bar{u}^v \leq \bar{\gamma}$

Case III: $\bar{u} \geq \bar{\sigma}$

In Figure 7, the vertex point (\bar{x}^*, \bar{u}^v) falls above the upper control bound with optimal state-control trajectories given as $(\bar{x}^*, \bar{\sigma})$ such that $F(\bar{x}, \bar{u}^*) = F(\bar{x}, \bar{\sigma})$. The functional Ω within the region of feasibility is then defined as

$$\Omega^* = \{F(\bar{x}^*, \bar{\gamma}) \mid \bar{x}^* \in X \text{ and } \bar{u} \geq \bar{\sigma} \text{ for all } \bar{u} \in U\}. \tag{61}$$

Updating the scaled-dual variables - ξ_1, ξ_2 : It yields

$$\xi_1^{k+1} = \xi_1^k + (A\bar{x}^{k+1} + B\bar{u}^{k+1} - C + z^{k+1}), \tag{62}$$

and

$$\xi_2^{k+1} = \xi_2^k + \bar{u}^{k+1} - y^{k+1}. \tag{63}$$

3.9 Primal-dual convergence

The convergence of the M-ADMM for the bounded control in terms of its primal-dual feasibility requires deriving the primal-dual residuals and their behaviors for large iterations.

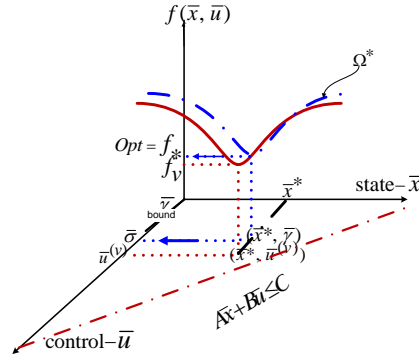


Figure 7: Control above upper bound:
 $\bar{\sigma} \leq \bar{u}^v$

Theorem 3. Given the bounded control problem (40)–(42) with the multipliers for the linear and bounded control constraints given as λ_1 and λ_2 , respectively, then there exists a dual residual $s^{k+1} = \rho_1 B^T [z^{k+1} - z^k] - \rho_2 [u^{k+1} - y^k]$ that converges to zero for given penalty parameters ρ_1 and ρ_2 for the linear and control bounds, respectively.

Proof. Given the objective function $p^k = f(x^k) + g(u^k)$, linear inequality constraint $Ax + Bu \leq C$ and $y = \sigma - \gamma$, the associated Lagrangian with slack z is stated as

$$\begin{aligned}
 L(\bar{x}, \bar{u}, z, \lambda_1, \lambda_2 : \rho_1, \rho_2) &= f(\bar{x}) + g(\bar{u}) + \lambda_1^T (A\bar{x} + B\bar{u} - C + z) \\
 &\quad + \frac{\rho_1}{2} \|A\bar{x} + B\bar{u} - C + z\|_2^2 + \frac{\mu}{2} \|y - v\|_2^2 \\
 &\quad + \lambda_2^T (\bar{u} - y) + \frac{\rho_2}{2} \|\bar{u} - y\|_2^2
 \end{aligned}$$

Applying the optimality conditions KKT to the Lagrangian function above yields

$$\begin{aligned}
& \partial g(\bar{u}) + B^T [\lambda_1^k + \rho_1 (A\bar{x}^{k+1} + B\bar{u} - C + z^k)] + \lambda_2^k + \rho_2 (\bar{u} - y^k) = 0, \\
& \partial g(\bar{u}^{k+1}) + B^T [\lambda_1^k + \rho_1 \underbrace{(A\bar{x}^{k+1} + B\bar{u}^{k+1} - C + z^{k+1})}_{r_1^{k+1}}] + \lambda_2^k \\
& + \rho_2 \underbrace{(\bar{u}^{k+1} - y^{k+1})}_{r_2^{k+1}} - \rho_1 B^T z^{k+1} + \rho_1 B^T z^k + \rho_2 y^{k+1} - \rho_2 y^k = 0, \\
& \partial g(\bar{u}^{k+1}) + B^T [\underbrace{\lambda_1^k + \rho_1 r_1^{k+1}}_{\lambda_1^{k+1}}] + \underbrace{\lambda_2^k + \rho_2 r_2^{k+1}}_{\lambda_2^{k+1}} \\
& = \underbrace{\rho_1 B^T [z^{k+1} - z^k] - \rho_2 [y^{k+1} - y^k]}_{s^{k+1}},
\end{aligned}$$

where the **primal residuals** are expressed as

$$r_1^{k+1} = A\bar{x}^{k+1} + B\bar{u}^{k+1} - C + z^{k+1}, \quad (64)$$

$$r_2^{k+1} = \bar{u}^{k+1} - y^{k+1}. \quad (65)$$

If $x^{k+1} = x^*$ minimizes the convex function, then

$$\partial g(\bar{u}^*) + B^T \lambda_1^* + \lambda_2^* = 0 \quad (66)$$

and setting $u^{k+1} = y^{k+1}$ completes the proof by expressing the **dual residual** as

$$s^{k+1} = \rho_1 B^T [z^{k+1} - z^k] - \rho_2 [u^{k+1} - y^k]. \quad (67)$$

□

Let the state-control variables be concatenated as $\hat{w} = (\bar{x}, \bar{u})$ with dimension $nN + m(N + 1) \times 1$, $\hat{A} = [\hat{A}, \hat{B}]$ with dimension $nN \times (nN + mN + m)$, and $\hat{P} = \begin{bmatrix} [c|c] \hat{P} & \hat{0} \\ \hat{0}^T & \hat{Q} \end{bmatrix}$ with dimension $(nN + mN + m) \times (nN + mN + m)$.

This yields a re-formulated augmented problem of the form

$$\text{Min} \frac{1}{2} \hat{w}^T \hat{P} \hat{w} + R \quad \text{s.t.} \quad \hat{A} \hat{w} \leq C, \quad (68)$$

which results into Theorem 4, as an extension of Ghadimi et al. [8].

Theorem 4. Given a convex quadratic programming (QP) problem $\min \frac{1}{2} \hat{w}^T \hat{P} \hat{w} + q^T \hat{w}$ such that $\hat{A} \hat{w} \leq b$, where $\hat{P} \in \mathbb{R}^{n \times n}$, $\hat{w} \in \mathbb{R}^n$, $q \in \mathbb{R}^n$, $\hat{A} \in \mathbb{R}^{m \times n}$, and $b \in \mathbb{R}^m$, then the optimal step-size for the QP is

$$\bar{\rho}_{CAL} = \left[\sqrt{\lambda_{\min}(\hat{A} \hat{P}^{-1} \hat{A}^T) \lambda_{\max}(\hat{A} \hat{P}^{-1} \hat{A}^T)} \right]^{-1} \quad (69)$$

and the convergence factor is

$$\xi_R^* = \frac{\lambda_{\max}(\hat{A}\hat{P}^{-1}\hat{A}^T) - \sqrt{\lambda_{\min}(\hat{A}\hat{P}^{-1}\hat{A}^T)\lambda_{\max}(\hat{A}\hat{P}^{-1}\hat{A}^T)}}{\lambda_{\max}(\hat{A}\hat{P}^{-1}\hat{A}^T) + \sqrt{\lambda_{\min}(\hat{A}\hat{P}^{-1}\hat{A}^T)\lambda_{\max}(\hat{A}\hat{P}^{-1}\hat{A}^T)}}. \quad (70)$$

For $\alpha \in (1, 2]$, $0 < \xi_R < 1$, \hat{P} is symmetric and positive definite and $\lambda_{\min}(\hat{A}\hat{P}^{-1}\hat{A}^T)$ and $\lambda_{\max}(\hat{A}\hat{P}^{-1}\hat{A}^T)$ are the minimum and maximum eigenvalues of the matrix $\hat{A}\hat{P}^{-1}\hat{A}^T$, respectively.

3.10 The Error dynamics of the relaxed Algorithm

Let us consider the augmented convex OCP in (68) with ℓ_2 -regularized estimation, with the regularization parameters $\delta > 0$, as stated below:

$$\text{Min} \frac{1}{2} \hat{w}^T \hat{P} \hat{w} + R + \frac{\delta}{2} \|z\|_2^2 \quad \text{s.t.} \quad \hat{A} \hat{w} + z - C = 0. \quad (71)$$

The resulting relaxed ADMM iterations take the form

$$\hat{w}^{k+1} = - \left(\hat{P} + \rho \hat{A}^T \hat{A} \right)^{-1} \hat{A}^T (\lambda^k + \rho(z^k - C)) \quad (72)$$

$$z^{k+1} = \frac{- \left[\lambda^k + \rho \left(\alpha (\hat{A} \hat{w}^{k+1} - C) - (1 - \alpha) z^k \right) \right]}{\delta + \rho} \quad (73)$$

$$\lambda^{k+1} = \lambda^k + \rho \left(\alpha (\hat{A} \hat{w}^{k+1} + z^{k+1} - C) + (1 - \alpha) (z^{k+1} - z^k) \right). \quad (74)$$

Considering the convergence of the relaxed iterations, (72)–(74), it is imperative to note that the relaxed iterations return the standard iterations of the convex control problem by setting $\alpha = 1$ and $\delta = 0$ with the augmented variable $\hat{w}^{k+1} = [\hat{x}^{k+1} \hat{u}^{k+1}]^T$. In characterizing the convergence of the relaxed iterations, we then analyze the *error dynamics* of the ADMM to know how the errors associated with \hat{w}^k or z^k vanish. Making λ^k subject of formula in (73) and inserting into (74) yield $\lambda^{k+1} = -\delta z^{k+1}$. Inserting the w -update into the z -update and using the fact that $\lambda^k = -\delta z^k$, we arrive at

$$z^{k+1} = \frac{1}{\delta + \rho} \underbrace{\left[\delta I + \rho \left((1 - \alpha) I + \alpha (\rho - \delta) \hat{A} (\hat{P} + \rho \hat{A}^T \hat{A})^{-1} \hat{A}^T \right) \right]}_E z^k - \frac{\alpha \rho}{\delta + \rho} \left[\rho \hat{A} (\hat{P} + \rho \hat{A}^T \hat{A})^{-1} \hat{A}^T - I \right] C. \quad (75)$$

For $z = z^*$ a fixed-point of (75), that is,

$$z^* = Ez^* - \frac{\alpha\rho}{\delta + \rho} \left[\rho\hat{A}(\hat{P} + \rho\hat{A}^T\hat{A})^{-1}\hat{A}^T - I \right] C,$$

then the dual error $e^{k+1} = z^{k+1} - z^*$ evolves as $e^{k+1} = Ee^k$ and $z^{k+1} - z^* = E(z^k - z^*)$ such that the z - update of the relaxed iterations above converge if and only if the spectral radius of the error matrix E in the above linear iterations is less than one. The eigenvalues of E can then be written as

$$\xi_{\hat{R}}(\alpha, \rho, \lambda_i(\hat{P})) = 1 - \frac{\alpha\rho(\lambda_i(\hat{P}) + \delta)}{(\rho + \lambda_i(\hat{P}))(\rho + \delta)}, \quad (76)$$

for ρ , δ and $\lambda_i(\hat{P}) \in \mathbb{R}_+$ (i.e., positive) provided $\alpha \in \left(1, \frac{2(\rho + \lambda_i(\hat{P}))(\rho + \delta)}{\rho(\lambda_i(\hat{P}) + \delta)}\right)$. To ascertain the upper bounds ρ and α that guarantee an improvement with strictly smaller convergence factor compared to that of the classical ADMM iterations, we take the derivatives $\xi_{\hat{R}}^*$ with respect to the parameters ρ and α given by $(\rho^*, \alpha^*) = \underset{\rho, \alpha}{\operatorname{argmin}} \max_i |\xi_{\hat{R}}(\alpha, \rho, \lambda_i(\hat{P}))|$ while that for the lower bound of $\xi_{\hat{R}}$ is $\xi_{\hat{R}}^* = \max_i |\xi_{\hat{R}}(\rho^*, \alpha^*, \lambda_i(\hat{P}))|$. This yields the jointly optimal step-size and convergence factor expressed as $(\rho^*, \xi_{\hat{R}}^*)$, where the detail proof of the convergence analysis can be seen in [8]. We then conclude that the over-relaxed iterations are guaranteed to converge faster for all $\alpha \in (1, 2]$. The developed M-ADMM algorithm is stated below; see Appendix for flowchart.

3.11 Algorithm: M-ADMM for proximal bound-constrained program

Input parameters and operators ρ , α , A , B , C , ϵ^{prim} , ϵ^{dual}
Initialize \bar{x}^0 , \bar{u}^0 , z^0 , $\xi_1^0 = \lambda_1^0/\rho_1$, $\xi_2^0 = \lambda_2^0/\rho_2$
Set $P \succeq 0$, $Q \succeq 0$ (symmetric and positive definite)
For $k = 0, 1, 2, \dots$, **do**
| Compute \bar{x}^{k+1} using (47)
| Compute \bar{u}^{k+1} using (48)
| Compute \bar{z}^{k+1} using (50)
| Compute \bar{y}^{k+1} using (54) and set $\bar{u}_1^{k+1} = y_i^{k+1}$, element-wise
for all i
| Update ξ_1^{k+1} and ξ_2^{k+1} using (62) and (63), respectively.
| Compute $\|r^{k+1}\|_2^2$ and $\|s^{k+1}\|_2^2$ using (64) and (67), respectively.
| **If** $\|r^{k+1}\|_2^2 \leq \epsilon^{\text{prim}}$ and $\|s^{k+1}\|_2^2 \leq \epsilon^{\text{dual}}$
| | **stop** and return (output)
| | **else**
| | **repeat** process
| **end (If)**

end (for)

Return $\bar{x}^{k+1}, \bar{u}^{k+1}, z^{k+1}, J^*, \xi_1^{k+1}, \xi_2^{k+1}$

3.12 Computational Complexity

Iteration complexity: Each step of the M-ADMM updates (iterations) is $O(N)$ with large computational efforts required. However, catching the Cholesky factorization of the matrices $(P + p_1 A^T A)$ and $(Q + \rho_1 B^T B + \rho_2 I)$ in the form RR^T (with R the upper triangular) will make the computation more effective, provided P and Q are real symmetric and positive definite. If P or Q is dense, then the computation of the Cholesky factor is a one-time cost of $O(\frac{1}{3}N^3)$ while the linear system is $O(N^2)$ by solving the two triangular systems. However, if P or Q is large and (sufficiently) sparse, then computing the Cholesky factor and solving the linear system are both possibly $O(N)$. The solution can be obtained in few iterations of the M-ADMM algorithm when the CGM or L-BFGS iterative solver is deployed with a warm-start to carry out faster convergence of the updates. In practice, the warm-start in which the values of $z = v$ and $\xi_1 = \xi_2 = 0$ are usually adopted.

Termination criteria: The earlier derived stopping criteria, in (64)–(67), can be deployed with a tolerance range of 10^{-k} (for $k = 3, \dots, 6$). Moreover, the speed at which the M-ADMM converges depends significantly on the quadratic penalty parameters by Boyd et al. in [3]. However, the primal $\epsilon^{Prim} > 0$ and dual $\epsilon^{Dual} > 0$ tolerances are selected as the termination (stopping) criteria for the convergence of the M-ADMM, and their choices depend on the relative ϵ^{rel} and absolute ϵ^{abs} criteria on account that the ℓ_2 norms are in \mathbb{R}^n and \mathbb{R}^m , respectively. The selected primal and dual tolerances are so small such that the algorithm converges (terminates) whenever $\|r^{k+1}\|_2 \leq \epsilon^{Prim}$ and $\|s^{k+1}\|_2 \leq \epsilon^{Dual}$. Usually in literature, the values, $\epsilon^{rel} = 10^{-3}$ and $\epsilon^{abs} = 10^{-3}$, are used as reasonable stopping criteria for the ADMM algorithms and can be reduced to improve the accuracy. Stated below are basic computations in literature as in [3].

$$\epsilon^{Prim} = \sqrt{n}\epsilon^{abs} + \epsilon^{rel} \cdot \max\{\|Ax^{k+1}\|_2, \|Bu^{k+1}\|_2, \|z^{k+1}\|_2, \|C\|_2\} \quad (77)$$

$$\epsilon^{Dual} = \sqrt{m}\epsilon^{abs} + \epsilon^{rel} \cdot \|\rho u\|_2 \quad (78)$$

4 Numerical experiments

In this section, the algorithm was implemented with numerical examples on a *Core i3* CPU. The quasi-Newton solver is interfaced with the M-ADMM in obtaining the optimizer. The convergence factor of ADMM for varying

tolerances and penalty parameters of the linear constraint is later illustrated. These examples demonstrate that the M-ADMM method converges for an over-relaxed factor $\alpha^* \in [1.8, 2.0]$ for faster convergence.

Example 1. Consider the one-dimensional convex multi-delay bounded OCP:

$$\begin{aligned}
 \text{Minimize } J(x(t), u(t)) &= \frac{1}{2} \int_0^1 [2x^2(t) + u^2(t)] dt \\
 \text{subject to } \dot{x}(t) &\leq -2x(t) + 3u(t) + 2x(t - 0.1) + x(t - 0.2) \\
 &\quad -x(t - 0.5) + u(t - 0.1) - 2u(t - 0.3) \\
 x(0) &= 1, \quad t \in [0, 1], \\
 x(t) &= 2t^2 + 1, \quad t \in [-0.5, 0], \\
 u(t) &= 3t + 2, \quad t \in [-0.3, 0], \\
 0 &\leq u(t) \leq 2.
 \end{aligned}$$

The discretized positive definite matrices $\bar{P}(\succ 0)$ and $\bar{Q}(\succ 0)$ for the objective functional, using $\delta = 0.1$ for purpose of illustration, are, respectively, given as

$$\bar{P} = \begin{pmatrix} 0.0889 & 0 & \dots & \dots & 0 \\ 0 & 0.0444 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0.0889 & 0 \\ 0 & \dots & \dots & 0 & 0.0222 \end{pmatrix} \text{ and } \bar{Q} = \begin{pmatrix} 0.0444 & 0 & \dots & \dots & 0 \\ 0 & 0.0222 & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0.0222 & 0 \\ 0 & \dots & \dots & 0 & 0.0111 \end{pmatrix}.$$

While the constraint coefficients are as given in subsection 3.3 using the optimal relaxation parameter $\alpha^* = 2.0$. Table 1 demonstrates the various effects of changing relaxation parameters (factors) on the convergence and results of the M-ADMM algorithm measured by the iterative cycles. The optimal relaxation factor is observed to be $\alpha^* = 1.90$.

Figures 8 and 9 show the effects of the varying penalty parameters ρ_2 on the rate of convergence of the M-ADMM algorithm for increasing relaxation factors α . Clearly, the lines overlap as the algorithm progresses. The optimum values were also obtained at $\alpha^* = 1.90$, as shown in the simulation, indicating an increase in the rate of convergence. Table 2 shows the results when the algorithm was ran for various values of the step-length δ and tolerances for fixed values of ρ_1, ρ_2, ρ_3 , and μ .

Figure 10 illustrates primal and dual convergences of the M-ADMM algorithm using the optimal over-relaxation factor $\alpha = 2.00$ while Figure 11 demonstrates the responses of the state-control variables and the objective values for the chosen tolerance (Tol.= 10^{-3}). The objective values $J^* = 0.3824$

Tolerances	10^{-3}	10^{-4}	10^{-5}	10^{-6}
No. of iterations	(k)	(k)	(k)	(k)
$\alpha = 1.00$	66	128	193	259
$\alpha = 1.20$	55	107	160	216
$\alpha = 1.40$	47	98	138	189
$\alpha = 1.60$	42	80	122	170
$\alpha = 1.80$	37	71	109	155
$\alpha^* = 2.00$	34	69	130	204

Table 1: Effects of varying relaxation factors on fixed penalty parameters ($\mu = 0.3$, $\rho_1 = 0.1$, $\rho_2 = 0.2$, $\rho_3 = 0.3$: $\delta = 0.1$)

Step-lengths	$\delta = 0.10$		$\delta = 0.05$		$\delta = 0.01$	
	10^{-3}	10^{-4}	10^{-3}	10^{-4}	10^{-3}	10^{-4}
$\ x\ $	2.3788	2.3919	2.8830	2.9223	3.2970	3.6359
$\ u\ $	0.8937	0.8933	0.9271	0.9263	0.9812	0.9796
$\ J^*\ $	0.3827	0.3900	0.3968	0.3982	0.3623	0.3730
ξ_R^*	0.5827	0.5827	0.0528	0.0528	0.0471	0.0471
$\bar{\rho}_{CAL}$	0.0557	0.0557	0.7379	0.7379	0.9383	0.9383
k	11	16	14	25	34	69
Time (secs)	0.0504	0.05623	0.0886	0.0305	0.3778	0.2057

Table 2: Summary of results at optimum relaxation parameter ($\mu = 0.3$, $\rho_1 = 0.1$, $\rho_2 = 0.2$, $\rho_3 = 0.3$, $\alpha^* = 2.0$)

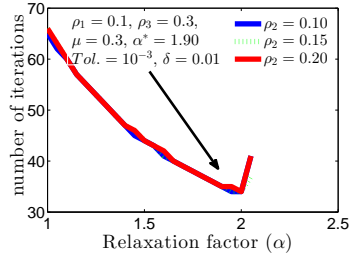


Figure 8: Effects of varying α, ρ on conv. as $\|r^k\|_2^2 \leq \epsilon^{dual}$ and $\|s^k\|_2^2 \leq \epsilon^{primal}$

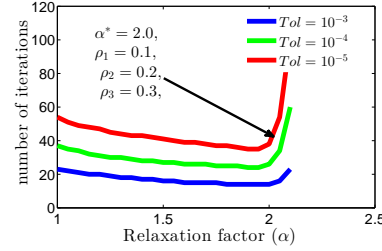


Figure 9: Effects of varying $\alpha, Tol.$ on conv. as $\|r^k\|_2^2 \leq \epsilon^{dual}$ and $\|s^k\|_2^2 \leq \epsilon^{primal}$

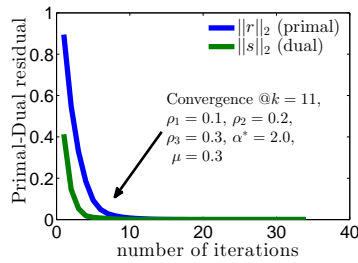


Figure 10: Primal-Dual Convergence: $\|r^k\|_2 \rightarrow r^*, \|d^k\|_2^2 \rightarrow 0$

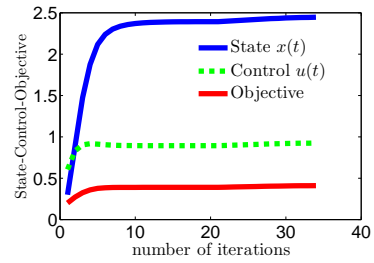


Figure 11: Optimal trajectories: State, control, and objective

were arrived at with the respective optimal state and control values 2.3788 and $0.8937 \in [0, 2]$, respectively.

Example 2. Considering the generalized multi-delay OCP with bounded control:

$$\text{Min}J(x, u) = \int_0^{0.5} [2x_1^2 + x_1x_2 + x_2^2 + u_1^2 + u_1u_2 + u_2^2] dt \quad (79)$$

$$\begin{aligned} \text{s.t. } \dot{x}_1(t) &\leq 2x_1(t) + x_2(t) + u_1(t) + 3u_2(t) + x_1(t - 0.1) + x_2(t - 0.1) \\ &\quad + 2x_1(t - 0.2) + x_2(t - 0.2) - x_1(t - 0.3) + 2u_1(t - 0.1) \\ &\quad + u_2(t - 0.1) + u_1(t - 0.3) + 2u_2(t - 0.2), \\ \dot{x}_2(t) &\leq x_1(t) - u_1(t) + 2u_2(t) - x_1(t - 0.1) + x_1(t - 0.2) \\ &\quad + 2x_2(t - 0.2) - x_1(t - 0.3) + u_2(t - 0.1) + u_1(t - 0.2) \\ &\quad + 3u_2(t - 0.2), \quad t \in [0, 0.5] \end{aligned} \quad (80)$$

$$x(0) = (1, 1), \quad (81)$$

$$x(t) = (2t + 1, t^2 + 1), \quad -0.3 \leq t \leq 0, \quad (82)$$

$$u(t) = (2, 2 + t), \quad -0.2 \leq t \leq 0, \quad (83)$$

$$(1, 1) \leq u(t) \leq (2, 3), \quad (84)$$

where $x(t) \in \mathbb{R}^2$ and $u(t) \in \mathbb{R}^2$.

The problem is re-formulated into the format in (1)–(7) to make it amenable to the M-ADMM algorithm. The coefficient matrices are defined below.

$$P = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}, Q = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \gamma = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \sigma = \begin{bmatrix} 2 \\ 3 \end{bmatrix},$$

$$\alpha_1 = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix}, \alpha_2 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \alpha_3 = \begin{bmatrix} -1 & 0 \\ -1 & 0 \end{bmatrix}, \beta_1 = \begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}, \text{ and } \beta_2 = \begin{bmatrix} 1 & 2 \\ 1 & 3 \end{bmatrix}.$$

The coefficient matrices, P and Q , of the quadratic functional are symmetric, invertible (non-singular), and positive definite (i.e., $P \succ 0$ and $Q \succ 0$) since their respective eigenvalues ($\lambda_1 = 1.585$, $\lambda_2 = 4.4142$) and ($\lambda_1 = 1$, $\lambda_2 = 3$) are all positive. This ensures that the operators are well-posed for the algorithm. The choice of the step-length $\delta = 0.1$ was made with $N = 5$ such that the coefficient matrices P and Q of the state and control variables for the continuous-time problem yield the block matrices $\bar{P} \in \mathbb{R}^{10 \times 10}$ and $\bar{Q} \in \mathbb{R}^{12 \times 12}$, respectively. The structures of their block matrix operators, are as defined in (3.2). However, the recurrence relation from the discretization of the constraint, as derived in (23), yields

$$\theta x^{(k)} - x^{(k+1)} + \omega u^{(k)} = -\delta \sum_{j=1}^3 \alpha_j x^{(k-v_j)} - \delta \sum_{l=1}^2 \beta_l u^{(k-w_l)}, \quad (85)$$

where $v = r_3/\delta = 3$, $w = q_2/\delta = 2$ and the various associated matrices are defined below:

$$\theta = I_n + A\delta = \begin{bmatrix} 1.2000 & 0.1000 \\ 0.1000 & 1.0000 \end{bmatrix}, \omega = B\delta = \begin{bmatrix} 0.1000 & 0.3000 \\ -0.1000 & 0.2000 \end{bmatrix},$$

with the formulated block-matrices \bar{A} , \bar{B} , and C are as defined in subsection 3.3. Table 3 shows the results in the first phase of the algorithm. The M-ADMM was ran for various values of the relaxation parameter α evenly spaced between 1.0 and 2.0 with a step-size of 0.05 for a fixed value of the penalty parameters ($\rho_1 = 0.1$, $\rho_2 = 0.2$, $\rho_3 = 0.3$, $\alpha^* = 2.0$) and tolerance ($Tol = 10^{-3}$) to ascertain the optimum relaxation factor (α^*) using the number of iterations as the measure of convergence.

Table 4 also demonstrates the effects of some of the selected relaxation parameters (factors) on the optimum values of the optimization problem as the tolerances decrease geometrically for fixed values of the penalty parameters. Figures 12 and 13 show the rate of convergence of M-ADMM Algorithm using the primal and dual residuals for increasing values of the iteration. Clearly, both lines converge to zero as the algorithm progresses for the varying values of α and ρ_2 as shown in Figure 14, where the average elapsed time per/cycle falls within a second when run on an Intel computer (*Inspiron 15 Core i3*). The optimum values were obtained at $\alpha^* = 2.0$, as shown in the number of

Step-lengths Tolerances	$\delta = 0.10$		$\delta = 0.05$		$\delta = 0.01$	
	10^{-3}	10^{-4}	10^{-3}	10^{-4}	10^{-3}	10^{-4}
$\ x\ $	14.5367	14.1757	0.138099	0.135782	0.138099	0.135782
$\ u\ $	2.7139	2.7203	0.15	144	0.22	164
$\ J^*\ $	18.3905	18.6315	0.16	140	0.22	164
ξ_R^*	0.6910	0.6910	0.16	140	0.22	164
$\bar{\rho}_{CAL}$	0.1612	0.1612	0.16	140	0.22	164
k	117	138	0.16	140	0.22	164
Time (secs)	0.2678	0.3852	0.16	140	0.22	164

Table 3: Summary of results at optimum relaxation parameter ($\rho_1 = 0.1, \rho_2 = 0.2, \rho_3 = 0.3, \alpha^* = 2.0$)

Tolerances	10^{-3}	10^{-4}	10^{-5}	10^{-6}
No. of iterations	(k)	(k)	(k)	(k)
$\alpha = 1.00$	239	467	703	940
$\alpha = 1.20$	199	389	585	786
$\alpha = 1.40$	175	345	523	700
$\alpha = 1.60$	155	307	464	622
$\alpha = 1.80$	140	277	419	562
$\alpha^* = 2.00$	127	253	383	514

Table 4: Effects of varying relaxation factors on fixed penalty parameters ($\rho_1 = 0.1, \rho_2 = 0.2, \rho_3 = 0.3 : \delta = 0.1$)

iterations per cycle, indicating a decrease in the rate of convergence. The M-ADMM ran for various values of the step-length δ geometrically spaced between 0.010 and 0.10 with a constant factor of 0.5 using the tolerances of 10^{-3} and 10^{-4} . Figure 14 further illustrates the faster rate of convergence

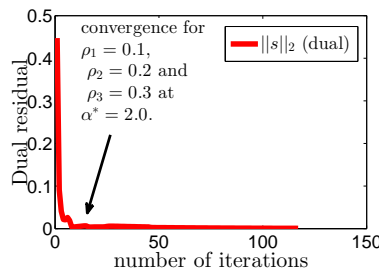


Figure 12: Dual convergence: $\|d^k\|_2 \rightarrow 0$, μ and $\alpha^* = 2.0$

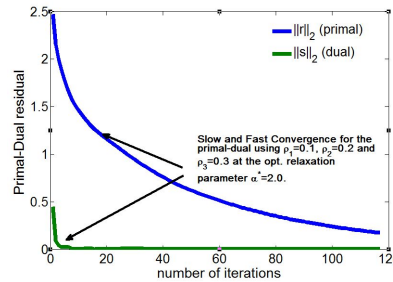


Figure 13: Primal-dual convergence: $\|r^k\|_2 \rightarrow r^*, \|d^k\|_2 \rightarrow 0$

of the M-ADMM with the optimal over-relaxation factor $\alpha = 2.00$ compared to the non-relaxed ADMM for which $\alpha^* = 1$. However, Figure 15 demon-

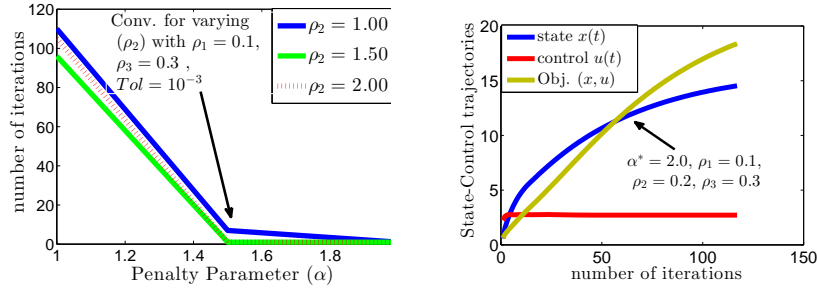


Figure 14: Convergence for varying α, ρ_2 : Figure 15: Optimal trajectories: State, control, and objective ($\rho_1 = 0.1, \rho_3 = 0.3, \alpha^* = 2.0$)

states the responses of the state-control variables and the objective values to increasing iterations at chosen tolerance ($Tol.=10^{-3}$), relaxation parameters ($\rho_1 = 0.1, \rho_2 = 0.2, \rho_3 = 0.3$) and optimal penalty parameter $\alpha^* = 2.0$. In addition, the objective values are observed to increase with increasing values of the state variables until the optimal objective value $J^* = 18.3905$ was arrived at with the respective optimum state and control trajectories stated as

$$x^* = \begin{bmatrix} 4.1416 \\ 13.9343 \end{bmatrix} \text{ and } u^* = \begin{bmatrix} 1.0000 \\ 2.5230 \end{bmatrix} \text{ while } \begin{bmatrix} \|x^*\| \\ \|u^*\| \end{bmatrix} = \begin{bmatrix} 14.5367 \\ 2.7139 \end{bmatrix}$$

and $\|(1, 2)^T\| \leq \|u^*\| \leq \|(2, 3)^T\|$. Applying the formulas in (69) and (70) yields the results for the calculated penalty parameter and convergence factor as $\bar{\rho}^* = 0.1612$ and $xi_{\bar{R}}^* = 0.6910$, respectively, where $\bar{P} \in \mathbb{R}^{(4N+2) \times (4N+2)}$ and $\bar{A} = [A, B] \in \mathbb{R}^{2N \times (4N+2)}$.

5 Conclusion

The ADMM for strictly convex QP with a particular focus on the QPs arising from discretized multi-delay OCP with bounded control was studied. A convergence proof and a method of analysis of the problem were proposed that allows the selection of the optimal relaxation parameter for the acceleration of the algorithm for varying penalty parameters of the augmented Lagrangian functional; while other parameters remain fixed. A major drawback of this approach is that, the penalty parameters for both the linear and bound constraints cannot be derived simultaneously. In the future, we will extend the computational framework to include the effects of both varying parameters in the computational complexity of the algorithm. This research was limited to OCP with linear constraints of non-fractional order. However, in subsequent work, the scope of the work will be extended to convex constrained OCPs

with coupled and/or nonlinear differential constraints with both integer and (or) fractional order.

Acknowledgments

I am grateful to Prof. Montaz Ali of *The University of The Witwatersrand, SA* for introducing me to the ADMM and for his helpful comments in improving the presentation of this article.

References

1. Bashier, E.B.M. and Patidar, K.C. *Optimal control of an epidemiological model with multiple time delays*, Appl. Math. Comput. 292 (2017), 47–56.
2. Boley, D. *Local linear convergence of the alternating direction method of multipliers on quadratic or linear programs*, SIAM J. Optim. 23(4) (2013), 2183–2207.
3. Boyd, S., Parikh, N., Chu, E. and Eckstein, J. *Distributed optimization and statistical learning via the alternating direction method of multipliers*, Found. Trends Mach. Learn. 3(1) (2011), 11–22.
4. Boyd, S. and Vandenberghe, L. *Convex Optimization*, Cambridge University Press, NY, 2004.
5. Carreira-Perpiñán, M.A. *An ADMM algorithm for solving a proximal bound-constrained quadratic program*, arXiv:1412.8493[math.OC], 2014.
6. Combettes, P. and Pesquet, J. *Proximal splitting methods in signal processing*, Fixed-point algorithms for inverse problems in science and engineering, 185–212, Springer Optim. Appl., 49, Springer, New York, 2011.
7. Eckstein, J. and Ferris, M. *Operator-splitting methods for monotone affine variational inequalities, with a parallel application to optimal control*, INFORMS J. Comput. 10(2) (1998), 218–235.
8. Ghadimi, E., Teixeira, A., Shames, I. and Johansson, M. *Optimal Parameter Selection for the Alternating Direction Method of Multipliers (ADMM): Quadratic Problems*, Linnaeus center, Elect. Eng., Royal Inst. of Technology, Sidney, 2014.
9. Gollman, L., Kern, D. and Maurer, H. *Optimal control problems with delays in state and control variables subject to mixed control-state constraints*, Optimal Control Appl. Methods 30(4) (2009), 341–365.

10. Gollman, L. and Maurer, H. *Theory and Applications of Optimal Control Problems with Multiple Delays*, J. Ind. Manag. Optim. 10 (2014), 413–441.
11. Guinn, T. *Reduction of delayed optimal control problems to nondelayed problems*, J. Optim. Theory Appl. 18 (1976), 371–377.
12. He, B., Hong, X. and Yuan, X. *On the proximal Jacobian decomposition of ALM for multiple-block separable convex minimization problems and its relationship to ADMM*, J. Sci. Comput. 66 (2016), 1204–1217.
13. He, B. and Yuan, X. *Block-wise alternating direction method of multipliers for multiple-block convex programming and beyond*, SSMAI J. Comput. Math. 691 (2015), 145–174.
14. Laarabi, H., Abta, A. and Hattaf, K. *Optimal control of a delayed SIRS epidemic model with vaccination and treatment*, Acta Biotheor. 63 (2015), 87–97.
15. Moreau, J.J. *Fonctions convexes duales et points proximaux dans un espace hilbertien*, (French) C. R. Acad. Sci. Paris 255 (1962), 2897–2899.
16. Nesterov, Y.E. *A method for solving the convex programming problem with convergence rate $O(1/k^2)$* , Dokl. Akad. Nauk SSSR, 269 (1983), 543–547.
17. Olotu, O. and Dawodu, K.A. *On the discretized algorithm for optimal proportional control constrained by delay differential equation*, Journal of Mathematical Theory and Modelling by IISTE, 3(8) (2013), 157–169.
18. Olotu, O. and Dawodu, K.A. *Numerical solution to one-dimensional multi-delay system using the modified alternating direction method of multipliers*, Journal of the Nigerian Association of Mathematical Physics, Nigeria. 50 (2019), 107–116.
19. Olotu, O., Dawodu, K.A. and Yusuf, T.T. *Algorithm for solving multi-delay optimal control problems using modified alternating direction method of multipliers*, J. Appl. Computat. Math. 8 (3) (2019), 1000445.
20. Pasquale, L.D. and Gerardo, T. *Quadratic programming with bound constraints*, Encyclopedia of Optimization, USA, 2009.
21. Rihan, F.A. and Anwar, M.N. *Qualitative analysis of delayed SIR epidemic model with a saturated incidence rate*, Int. J. Differ. Equ. 2012, Art. ID 408637, 13 pp.
22. Rockafellar, R.T. *Monotone operators and the proximal point algorithm*, SIAM J. Control and Optim., 14 (5), (1976), 877–898.

23. Sun, D., Toh, K.C. and Yang, L. *A convergent 3-block semi-proximal alternating direction method of multipliers for conic programming with 4-type of constraints*, SIAM J. Optim., 25(2) (2015) 882–915.
24. Voglis, C. and Lagaris, I.E. *Boxcqp: An algorithm for bound constrained convex quadratic problems* 1st IC-SCCE, Athens, 2004.
25. Wu, D. and Shang, Y. *Complete solutions to general box-constrained global optimization problems*, J. Appl. Math. 2011, Art. ID 478608, 17 pp.
26. Xinmin, L., Andrew, H.B., Zachary, G. and Rodney, D. *Constrained quadratic optimization for radiation treatment planning*, American Control Conference (ACC), Boston, USA, 2016.
27. Xue, Y. and Duan, X. *Dynamic analysis of an SIR epidemic model with nonlinear incidence rate and double delays*, Intl. J. Inf. Syst. Sci. Computer Inf. 7 (2011), 92–102.
28. Yao, Y., Zhu, X., Dong, H., Wu, S., Wu, H., Carol Tong, L. and Zhou, X. *ADMM-based problem decomposition scheme for vehicle routing problem with time windows*, Transp. Res. B: Meth. 129, (2019), 156–174.

Appendix

A. Derivation of \bar{P} and \bar{Q}

By Simpson's rule, if for each $k = 1, 2, \dots, \frac{N}{2} - 1$, $f(t_{2k})$ appears in the term corresponding to the interval $[t_{2k-2}, t_{2k}]$ and $f(t_{2k-1})$ appears in the term corresponding to the interval $[t_{2k}, t_{2k+2}]$ for $k = 1, 2, \dots, \frac{N}{2}$, then the approximation of the integral in (18), given that $F(t, x(t), u(t)) = x^T P x + u^T Q u$, can be expressed as

$$I := \frac{1}{2} \int_{t_0}^T F(t, x(t), u(t)) dt \simeq \frac{1}{2} \sum_{k=1}^{N-1} F(t_k, x^{(k)}, u^{(k)}) \quad (86)$$

$$\simeq \frac{1}{2} \left[\frac{\delta}{3} \left(f^{(0)} + 2 \sum_{k=1}^{\frac{N}{2}-1} f^{(2k)} + 4 \sum_{k=1}^{\frac{N}{2}} f^{(2k-1)} + f^{(N)} \right) \right],$$

where

$$f^{(0)} = (x^{(0)})^T P x^{(0)} + (u^{(0)})^T Q u^{(0)}, \quad (87)$$

$$f^{(2k)} = (x^{(2k)})^T P x^{(2k)} + (u^{(2k)})^T Q u^{(2k)}, \quad (88)$$

$$f^{(2k-1)} = (x^{(2k-1)})^T P x^{(2k-1)} + (u^{(2k-1)})^T Q u^{(2k-1)}, \text{ and} \quad (89)$$

$$f^{(N)} = (x^{(N)})^T P x^{(N)} + (u^{(N)})^T Q u^{(N)}. \quad (90)$$

Substituting equations (87)–(90) into (86) yields

$$I \approx \frac{\delta}{6} (x^{(0)})^T P x^{(0)} + \frac{1}{2} \left[\frac{2\delta}{3} \sum_{k=1}^{\frac{N}{2}-1} (x^{(2k)})^T P x^{(2k)} + \frac{4\delta}{3} \sum_{k=1}^{\frac{N}{2}} (x^{(2k-1)})^T P x^{(2k-1)} \right. \\ \left. + \frac{\delta}{3} (x^{(N)})^T P x^{(N)} \right] + \frac{1}{2} \left[\frac{\delta}{3} (u^{(0)})^T P u^{(0)} + \frac{2\delta}{3} \sum_{k=1}^{\frac{N}{2}-1} (u^{(2k)})^T P u^{(2k)} \right. \\ \left. + \frac{4\delta}{3} \sum_{k=1}^{\frac{N}{2}} (u^{(2k-1)})^T P u^{(2k-1)} + \frac{\delta}{3} (u^{(N)})^T P u^{(N)} \right]. \quad (91)$$

Expanding (91) for various values of k gives

$$\begin{aligned}
 I \approx & \frac{\delta}{6} (x^{(0)})^T P x^{(0)} + (x_1^{(1)}, x_1^{(2)}, \dots, x_n^{(N)}) \begin{pmatrix} \frac{4\delta}{3}P & 0 & \dots & \dots & 0 \\ 0 & \frac{2\delta}{3}P & \ddots & \dots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \frac{4\delta}{3}P & 0 \\ 0 & \dots & \dots & 0 & \frac{\delta}{3}P \end{pmatrix} \begin{pmatrix} x_1^{(1)} \\ x_1^{(2)} \\ \vdots \\ \vdots \\ x_n^{(N)} \end{pmatrix} \\
 & + (u_1^{(0)}, u_1^{(1)}, \dots, u_m^{(N-1)}, u_m^{(N)}) \begin{pmatrix} \frac{\delta}{3}Q & 0 & \dots & \dots & \dots & 0 \\ 0 & \frac{4\delta}{3}Q & \ddots & \dots & \dots & \vdots \\ \vdots & \ddots & \frac{2\delta}{3}Q & \ddots & \dots & \vdots \\ \vdots & \dots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \dots & \dots & \ddots & \frac{4\delta}{3}Q & 0 \\ 0 & \dots & \dots & \dots & 0 & \frac{\delta}{3}Q \end{pmatrix} \begin{pmatrix} u_1^{(0)} \\ u_1^{(1)} \\ \vdots \\ \vdots \\ u_m^{(N-1)} \\ u_m^{(N)} \end{pmatrix}
 \end{aligned}$$

expressed as

$$I \approx \min_{\bar{x}, \bar{u}} \frac{1}{2} \bar{x}^T \bar{P} \bar{x} + \frac{1}{2} \bar{u}^T \bar{Q} \bar{u} + R, \tag{92}$$

where any elements $\bar{p}_{ij} \in \bar{P}$ and $\bar{q}_{ij} \in \bar{Q}$ are described as

$$[\bar{p}_{ij}] = \begin{cases} \frac{4\delta}{3}P, & i = j(\text{odd}) & 1 \leq i \leq N - 1, \\ \frac{2\delta}{3}P, & i = j(\text{even}) & 2 \leq i \leq N, \\ \frac{2\delta}{3}P, & i = j, i = N, \\ 0 & \text{elsewhere} \end{cases}$$

and

$$[\bar{q}_{ij}] = \begin{cases} \frac{\delta}{3}Q, & i = j, i = 1, N, \\ \frac{4\delta}{3}Q, & i = j(\text{even}) & 2 \leq i \leq N, \\ \frac{2\delta}{3}Q, & i = j(\text{odd}) & 3 \leq i \leq N, \\ 0 & \text{elsewhere.} \end{cases}$$

B. Flowchart for the M-ADMM Algorithm

The diagram below (Figure 16) is the flowchart demonstrating the procedure of the M-ADMM algorithm in handling the discretized optimal control (or constrained optimization) problem with bounded (boxed) control. The left-side of the flowchart is the M-ADMM subroutine, which is the outer loop of Algorithm. However, the right-side of the flowchart is the optimization solver, which in this case is the "quasi-Newton solver" of the low-memory BFGS type.

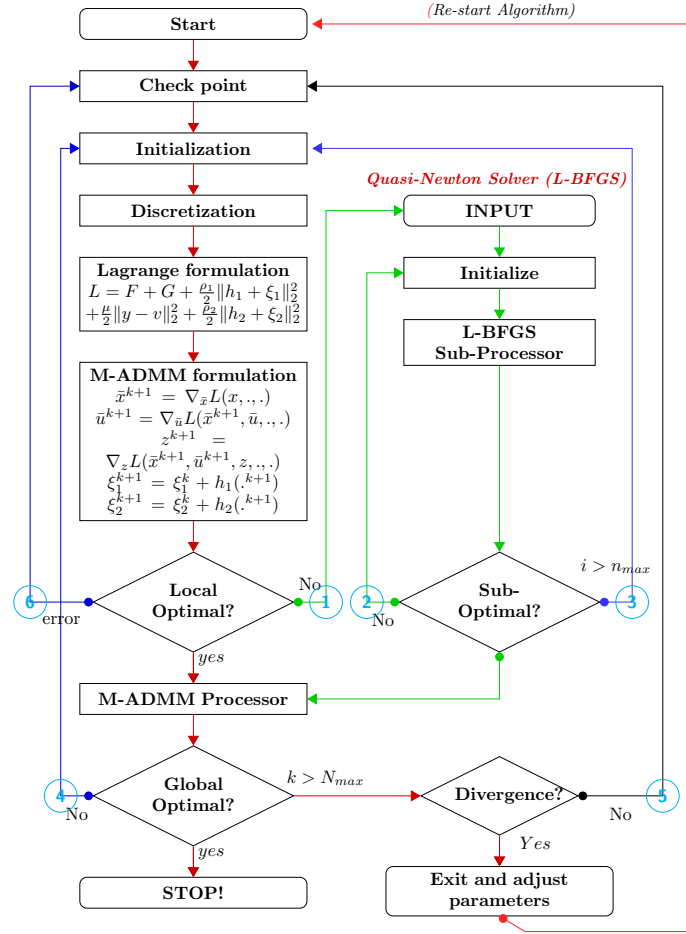


Figure 16: Flowchart for M-ADMM with bounded control

How to cite this article

Dawodu, K. Modified ADMM algorithm for solving proximal bound formulation of multi-delay optimal control problem with bounded control. *Iranian Journal of Numerical Analysis and Optimization*, 2022; 12(1): 1-36. doi: 10.22067/ijnao.2021.69185.1020.