Journal of
Applied & Computational Mathematics

Olotu et al., J Appl Computat Math 2019, 8:3

Research Article Open Access

Algorithm for Solving Multi-Delay Optimal Control Problems Using Modified Alternating Direction Method of Multipliers

Olotu O, Dawodu KA* and Yusuf TT

Department of Mathematical Sciences, Federal University of Technology, Akure, Nigeria

Abstract

Purpose: This study presents an algorithm for solving Optimal Control Problems (OCPs) with objective function of the Lagrange-type and multiple delays on both the state and control variables of the constraints.

Design/Methodology/Approach: The full discretization of the objective functional and the multiple delay constraints was carried out using the Simpson numerical scheme. The discrete recurrence relations of both the objective function and constraints were generated and used to develop the matrix operators of the optimal control system. The spectral and convergence analyses of the developed matrix operators and formulated unconstrained optimal control problem were carried out respectively to ascertain the well-posedness of the Modified Alternating Direction Method of Multipliers (M-ADMM). The primal-dual feasibility and their residuals were derived in other to obtain the point of convergence of the algorithm.

Findings: The direct numerical approach for handling the multi-delay control problem was observed to obtain a very accurate result at a super-linear rate of convergence. This makes the algorithm faster when compared to the classical Pontryagin maximum principle.

Research limitations/Implications: The research is limited to linear constraints and objective functional of the Lagrange-type.

Practical implications: This research can address real-life models with multiple delays in epidemiology.

Originality: The novelty of this research paper lies in the method of discretization of the multiple delay constraints and the adaptation of the modified ADMM algorithm in handling linearly constraint multiple delay optimal control problems for better performance in terms of speed and accuracy.

Keywords: Multiple delays; Recurrence relation; Alternating direction method of multipliers; Primal residual; Dual residual; Super-linear convergence; Composite Simpson's rule

Introduction

Optimal Control Systems (OCS) with linear and nonlinear inequality constraints with delays in state and/or control variables play important roles in the modeling of real-life phenomena [1]. Many papers over the years have been devoted to developing the maximum principles for optimal control problems with constant and variable delays because the time delay effects in control systems cannot be neglected especially when it involves the transmission of information between different parts of a system [2]. This hereditary effect which frequently occur in economical, physical, chemical and biological processes have been deliberated upon by Bashier and Patidar [3]. Bertsekas [4] generalized the classical Pontryagin's maximum principle for optimal control problems with delay in the state only. Contemporarily, Boley [5] obtained the maximum principle for the non-autonomous linear-quadratic optimal control problem with multiple delays in the state only [6]. Maximum principle for control systems with a time-dependent delay in the state variable was also derived in Boyd et al. [7]. The maximum principle for linear-quadratic optimal control problems with multiple constant delays in both state and control variables was obtained by Chyung and Lee [8]. Colonius and Hinrichsen [9] provided a unified approach to control problems with delays in the state variable only while Ghadimi et al. [10] established maximum principle for the delay optimal control problem with multiple variable delays by the theory of optimal fields. The direct numerical approach was developed by Gilbert [11] using the Quasi-Newton embedded augmented Lagrangian functional for the proportional control class of Optimal Control Problems. Goldfarb and Schneiberg [12] did an 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 [13]. Later, Gollmann and Maurer [14] extended their work to multiple time delays in the control and state variables with mixed control-state constraints. The Pontryagin-type minimum principle for the class of delayed control problems was derived to ascertain their solutions [15]. It is therefore the aim of this paper to introduce the Douglas-Rachford (D-R) splitting method called the Alternating Direction Method of Multipliers (ADMM) as a direct numerical approach in solving Multiple Delay Optimal Control Problems [16]. ADMM therefore enjoys the strong convergence properties of the method of multipliers and the decomposability property of dual ascent [17]. It has the potential to solve large-scaled structured convex optimization problems by solving the primal and dual feasibility updates in parallel [18]. The method gained wide acceptability due to its simplicity, versatility, scalability and ability to solve constrained optimization problems. A number of authors such

***Corresponding author:** Dawodu KA, Department of Mathematical Sciences, Federal University of Technology, Akure, Nigeria, Tel: 08032541380; E-mail: dawodukazeem@yahoo.com

Received April 24, 2019; **Accepted** May 10, 2019; **Published** May 17, 2019

Citation: Olotu O, Dawodu KA, Yusuf TT (2019) Algorithm for Solving Multi-Delay Optimal Control Problems Using Modified Alternating Direction Method of Multipliers. J Appl Computat Math 8: 445.

as Laarabi et al. [19], Nesterov [20], O'Donoghu et al. [21], Olotu et al. [22], Rihan and Anwar [23] and Soliman [24] further worked on the linear convergence of the ADMM with strictly convex inequality [25]. Others include: Yang and Zhang [26], Zhang et al. [27], who worked on the acceleration of the ADMM, called the Accelerated Alternating Method of Multipliers (A-ADMM), to improve on its efficiency and rate of convergence. Nesterov [20] and Zhang et al. [27] improved on the ADMM by deriving the optimal parameter using the Gauss-Seidel relaxation factor. However, the A-ADMM will be deployed and modified as an optimization tool for solving the discretized continuous multiple control problem.

Statement of Problem

The general form of the multiple delay optimal control problem is:

$$\text{Minimize } J(x,\mathrm{u}) = \frac{1}{2}\int_{t_0}^{T} F(t,x(t),u(t))\,dt \tag{1}$$

$$\textit{Subject to } \dot{x}(t) = f\left(t,x(t),u(t),x^h(t),\mathrm{u}^h(t)\right) \quad t\in[t_0,T] \tag{2}$$

$$x(t)=\varphi(t), \qquad t\in[t_0-r,t_0], \tag{3}$$

$$u(t)=\psi(t), \qquad t\in(t_0-q,t_0), \tag{4}$$

$$x(t_0)=x_0, \tag{5}$$

where, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$, $r=\max\{r_j\}^d_{j=1}$, $q=\max\{q\}^e_{l=1}$ $F:[t_0,T]\times\mathbb{R}^n\times\mathbb{R}^m\to\mathbb{R}$

$f:[t_0, T]\times\mathbb{R}^{(d+1)n}\times\mathbb{R}^{(e+1)m}\to\mathbb{R}^n$, $\boldsymbol{\varphi}(t):[t_0-r;t_0]\to\mathbb{R}^n$, $\boldsymbol{\psi}(t):[t_0-q;t_0]\to\mathbb{R}^m$, $r_j<r_{j+1}$ and

$q_l<q_{l+1}$.

However, with respect to all arguments, the functions F and f are assumed to be at least twice continuously differentiable while the functions $\boldsymbol{\varphi}$ and $\boldsymbol{\psi}$ are only continuous. The pair of functions $(x,u)\in\Omega:=([t_0,T],\mathbb{R}^n)\times([t_0,T],\mathbb{R}^m)$ is an *admissible pair* of the problem above such that the conditions of eqn. (1) to eqn. (5) are satisfied. However, it is imperative to develop the numerical approach to solving the case where the continuous nonlinear functional F and constraint function f are quadratic and linear in nature respectively; hence the optimal control model below:

$$Min\ J(x,u)=\frac{1}{2}\int^{T} t_0\left(x^TPx+u^TQu\right)dt \tag{6}$$

$$s.t\ \dot{x}(t)\le Ax+Bu+\sum_{j=1}^{d}\alpha_j x(t-r_j)+\sum_{l=1}^{e}\beta_l u(t-ql) \quad t_0\le t\le T \tag{7}$$

$$x(t)=\varphi(t), \quad t_0-r\le t\le t_0 \tag{8}$$

$$u(t)=\psi(t), \quad t_0-q\le t\le t_0 \tag{9}$$

$$x(t_0)=x_0 \tag{10}$$

where, $r=\max\limits_{1\le j\le d} r_j, q=\max\limits_{1\le j\le e} ql, x=(x_1,x_2,\dots,x_n)^T\in\mathbb{R}^n, u=(u_1,u_2,\dots,\mathrm{u}_m)^T\in\mathbb{R}^m, P\in\mathbb{R}^{n\times n}, Q\in\mathbb{R}^{m\times m}, A\in\mathbb{R}^{p\times n}, B\in\mathbb{R}^{p\times m}, \alpha_j\in\mathbb{R}^{p\times n}\forall_j, \beta_l\in\mathbb{R}^{p\times m}\forall l$

while, all assumptions are as in the general form of the multiple delay control model above.

Materials and Methods

The concept of "*first discretize fully and then optimize*" approach was adopted in obtaining the optimal solution of the multiple delay OCP. This involves applying an existing numerical scheme to discretize the continuous-time multiple delayed constraint, thereby generating a recurrence relation that was used to obtain the large sparse matrix operators. The Augmented Lagrangian functional was used to convert the discretized constrained problem to an unconstrained form. An accelerator variant (relaxation factor) was introduced in the formulation of the modified ADMM algorithm to accelerate its rate of convergence. The spectral analyses of the matrix operators and the convergence analysis of the modified ADMM were carried out to ensure it is well-posed and implementable on the MATLAB subroutine.

Background and preliminaries

The state variable $x(t)=(x_1,x_2,\dots x_n)\in\mathbb{R}^n$ and the control variable $u(t)=(u_1,u_2,\dots,u_m)\in\mathbb{R}^m$ defined over the intervals $[t_0-r,t_0]$ and $[t_0-q,t_0]$ respectively are sets of n and m dimensional Banach spaces of the continuous functions of the state and control vectors respectively. Suppose the time interval is discretized by letting $\delta=\frac{(T-t_0)}{N}\in\mathbb{Z}^+$ such that $N\in\mathbb{Z}^+$ and $t_k=t_0+k\delta$, (for k=0, 1, …N), then the discretization operator f_x maps each discrete point in the time interval $[t_0,T]\subseteq\mathbb{R}$ into each discrete point of the concatenated state vector $x_i^{(k)}(.)\in\mathbb{R}^{nN}$ for all i=1, 2, …, n; while, the operator 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, …, m as expressed below:

$$f_x:[t_0,T]\subseteq\mathbb{R}\to x_i^{(k)}(.)\in\overline{x} \tag{11}$$

$$f_u:[t_0,T]\subseteq\mathbb{R}\to u_j^{(k)}(.)\in\overline{u} \tag{12}$$

where, $\overline{x}=(\overline{x}_1,\overline{x}_2,\dots,\overline{x}_n)\in\mathbb{R}^{nN}$

$\overline{x}_i=\left(\overline{x}_i^{(1)},\overline{x}_i^{(2)},\dots,\overline{x}_i^{(N)}\right)\in\mathbb{R}^N$, $\overline{\mathrm{u}}=(\overline{\mathrm{u}}_1,\overline{\mathrm{u}}_2,\dots,\overline{\mathrm{u}}_m)\in\mathbb{R}^{m(N+1)}$ *and* $\overline{\mathrm{u}}_j=\left(\overline{\mathrm{u}}_j^{(0)},\overline{\mathrm{u}}_j^{(1)},\dots,\overline{\mathrm{u}}_j^{(N)}\right)\in\mathbb{R}^{N+1}$.

Considering the existence of single delay constants r and q on the state and control variables over the interval $[t_0-r,T]$ and $[t_0-q,T]$, respectively, such that $r=\upsilon\delta$ and $q=w\delta$; the discretized delay state and delay control variables are hereby represented in the equations below:

$$f_{x^h}=\begin{cases}x^h(t_{k-v})=\varphi(t_{k-v}):k\le v;k=0,1,2,\dots,\upsilon,t\in[t_0-r,t_0],x(t_0)=x_0\,(\text{known state})\\ x(t_{k-v})=x_{k-v}:k>v;k=(v+1),(v+2),\dots,N,t\in[t_0,T]\,(\text{Unknown state})\end{cases} \tag{13}$$

and

$$f_{u^h}=\begin{cases}u^h(t_{k-w})=\psi(t_{k-w}):k\le w;k=0,1,2,\dots,w,t\in[t_0-q,t_0]\,(\text{known control})\\ u(t_{k-w})=u_{k-w}:k>w;k=(w+1),(w+2),\dots,N,t\in[t_0,T]\ \text{Unknown state.}\end{cases} \tag{14}$$

For the case of multiple delay constants $r_j\in\mathbb{R}$ (for j=1, 2,..., d) and $q_l\in\mathbb{R}$ (for l=1, 2,..., e) on the state and control variables respectively, it is assumed that the numbers are monotonically increasing; that is, ($r_j<r_j+1$) and ($q_l<q_l+1$). Therefore, the multiple state and control variables for eqns. (13) and (14) for j=1,2,...,d and l=1,2,...,e are represented in eqns. (15) and (16):

$$x_{k-n_j}=\begin{cases}x^h_{k-v_j}=\varphi\left(t_{k-v_j}\right):k-v_j\le k=0,1,2,\dots,v_j[t_0-r,t_0]\ (\text{known})\\ x_{k-v_j}:k-v_j>0;k=(v_j+1),\dots,N,t\in[t_0,T]\,(\text{unknown})\\ \qquad x(t_0)=x_0\ \text{given}\end{cases} \tag{15}$$

and

$$u_{k-w_l}=\begin{cases}u^h_{k-w_l}=\psi\left(t_{k-w_l}\right):k-w_l<0;k=0,1,2,\dots,(w_l-1),t\in[t_0-q,t_0],\ (\text{known})\\ u_{k-w_l}\ :k-w_l\ge 0\ \text{for}\ k=w_l,(w_l+1),\dots,N,t\in[t_0,T],\ (\text{unknown})\end{cases} \tag{16}$$

where, $r=r_d=\max\{r_j\}^d_{j=1}$ *and* $q=q_e=\max\{q_l\}^e_{l=1}$ 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:

Citation: Olotu O, Dawodu KA, Yusuf TT (2019) Algorithm for Solving Multi-Delay Optimal Control Problems Using Modified Alternating Direction Method of Multipliers. J Appl Computat Math 8: 445.

$$\hat{x} = \left(x^{(1-v_j)}, x^{(2-v_j)}, \ldots, x^{(N-v_j)}\right) \in \mathbb{R}^{nN} \; for\ all\ j = 1,2,\ldots,d,$$

$$\hat{u} = \left(u^{(-w_l)}, u^{(1-w_l)}, \ldots, u^{(N-w_l)}\right) \in \mathbb{R}^{m(N+1)} \; for\ all\ l = 1,2,\ldots,\mathrm{e},$$

where, for any arbitrary k (for $k = 0,1,2,\ldots,N$), the discretized delay state and control vectors are concatenated giving,

$$\mathrm{x}^{(\mathrm{k}-\mathrm{v_j})} = \left(\mathrm{x}_1^{(\mathrm{k}-\mathrm{v_j})}, \mathrm{x}_2^{(\mathrm{k}-\mathrm{v_j})}, \ldots, \mathrm{x}_\mathrm{n}^{(\mathrm{k}-\mathrm{v_j})}\right) \in \mathbb{R}^\mathrm{n}, \text{ for any j}, \tag{17}$$

$$\mathrm{u}^{(\mathrm{k}-\mathrm{v_j})} = \left(\mathrm{u}_1^{(\mathrm{k}-\mathrm{w_l})}, u_2^{(\mathrm{k}-\mathrm{w_l})}, \ldots, u_\mathrm{n}^{(\mathrm{k}-\mathrm{w_l})}\right) \in \mathbb{R}^\mathrm{m}, \text{ for any l}, \tag{18}$$

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 $\phi(t) \simeq (t_{-k})$ and $\psi(t) \simeq \psi(t_{-k})$ where $t_k = t_0 - kh$ for $k = 1,2,\ldots,v_d (\geq w_e)$.

Discretization of the objective function

In the discretization of the continuous-time objective function, the third order one-third Simpson rule in eqn. (19) below was used:

$$\int_{t_0}^{T} F(x,t))dt \simeq \frac{\delta}{3}\left[f(x_0) + 2\sum_{k=1}^{\frac{N}{2}-1} f(x_{2k}) + 4\sum_{k=1}^{\frac{N}{2}} f(x_{2k-1}) + f(x_N)\right] \tag{19}$$

Then the discretized objective function of the optimal control problem in eqn. (6) using eqn. (19) is expressed below as:

$$\begin{aligned} \min_{x,u} J(x,u) &\simeq \min_x \frac{1}{2}\left(\underbrace{\frac{\delta}{3}x_0^T P x_0 + \frac{2\delta}{3}\sum_{K=1}^{\frac{N}{2}-1} x_{2k}^T P x_{2k} + \frac{4\delta}{3}\sum_{k=1}^{\frac{N}{2}} x_{2k-1}^T P x_{2k-1} + \frac{\delta}{3}x_N^T P x_N}_{\bar{x}^T \bar{P}\bar{x}}\right) \\ &+ \min_u \frac{1}{2}\left(\underbrace{\frac{\delta}{3}u_0^T Q u_0 + \frac{2\delta}{3}\sum_{K=1}^{\frac{N}{2}-1} u_{2k}^T Q x_{2k} + \frac{4\delta}{3}\sum_{k=1}^{\frac{N}{2}} u_{2k-1}^T Q x_{2k-1} + \frac{\delta}{3}u_N^T Q u_N}_{\bar{u}^T \bar{Q}\bar{u}}\right) \\ &\simeq \min_{\bar{x},\bar{u}} \frac{1}{2}\bar{x}^T \bar{P}\bar{x} + \frac{1}{2}\bar{u}^T \bar{Q}\bar{u} + R \end{aligned} \tag{20}$$

where, $R = \frac{\delta}{6}X_0^T P x_0 \in \mathbb{R}$ and the concatenated state and control variables are $\bar{x} = \left(x_1^{(1)}, x_1^{(2)}, \ldots, x_n^{(N)}\right) \in \mathbb{R}^{nN}$ $and\ \bar{u} = \left(u_1^{(0)}, u_1^{(1)}, \ldots, u_m^{(N)}\right) \in \mathbb{R}^{m(N+1)}$ 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:

$$\bar{P} = \begin{bmatrix} \frac{4\delta}{3}P & 0 & \cdots & 0 \\ 0 & \frac{2\delta}{3}P & \ddots & \vdots \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & \frac{\delta}{3}P \end{bmatrix} \; and\ \bar{Q} = \begin{bmatrix} \frac{\delta}{3}Q & 0 & \cdots & \cdots & \cdots & 0 \\ 0 & \frac{2\delta}{3}Q & \ddots & \cdots & \cdots & \vdots \\ \vdots & \ddots & \frac{4\delta}{3}Q & \ddots & \cdots & \vdots \\ \vdots & \cdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & \cdots & \cdots & \ddots & \frac{4\delta}{3}Q & 0 \\ 0 & \cdots & \cdots & \cdots & 0 & \frac{\delta}{3}Q \end{bmatrix} \tag{21}$$

Discretization of the constraints

The $\frac{1}{3}$ Simpson's rule in eqn. (22) was used for the discretization of the multi-delay constraints:

$$x^{(k+2)} = x^{(k)} + \frac{\delta}{3}\left[f^k + 4f^{k+1} + f^{k+2}\right]. \tag{22}$$

$\dot{\mathrm{x}}(\mathrm{t}) = \mathrm{f}(\mathrm{t},\mathrm{x},\mathrm{u})$ implies that $\mathrm{f}(\mathrm{t},\mathrm{x},\mathrm{u}) \approx \mathrm{f}\left(\mathrm{t}^\mathrm{k}, \mathrm{x}^\mathrm{k}, \mathrm{u}^\mathrm{k}, \hat{\mathrm{x}}^\mathrm{k}, \hat{\mathrm{u}}^\mathrm{k}\right) = \mathrm{f}^\mathrm{k}$ where

$$f^k = Ax^{(k)} + Bu^{(k)} + \sum_{j=1}^{d}\alpha_j x^{(k-v_j)} + \sum_{l=1}^{e}\beta_l u^{(k-w_l)}, \tag{23}$$

Adapting eqns. (22) and (23) to eqn. (7) of the optimal control problem then yielded the recurrence relation below:

$$\begin{aligned} &(\sigma + I_x)x^{(k)} + 4\sigma x^{(k+1)} + (\sigma - I_x)x^{(k+2)} + \gamma u^{(k)} + 4\gamma u^{(k+2)} = \\ &-\frac{\delta}{3}\sum_{j=1}^{d}\alpha_j\left(x^{(k-j)} + x^{(k+1-j)} + x^{(k+2-j)}\right) - \frac{\delta}{3}\sum_{l=1}^{e}\beta_l\left(u^{(k-l)} + u^{(k+1-l)} + u^{(k+2-l)}\right) \end{aligned} \tag{24}$$

where,

$$\sigma = \frac{A\delta}{3} \in \mathbb{R}^{p\times n}, \gamma = \frac{B\delta}{3} \in \mathbb{R}^{p\times m}, I_x \in \mathbb{R}^{p\times n} \; and\ k = 0,1,\ldots,N-2$$

Setting k=0 in eqn. (24) yielded:

$$\begin{aligned} &(\sigma + I_x)x^{(0)} + 4\sigma x^{(1)} + (\sigma - I_x)x^{(2)} + \gamma u^{(0)} + 4\gamma u^{(1)} + \gamma u^{(2)} = \\ &-\frac{\delta}{3}\sum_{j=1}^{d}\alpha_j\left(x^{(-j)} + x^{(1-j)} + x^{(2-j)}\right) - \frac{\delta}{3}\sum_{l=1}^{e}\beta_l\left(u^{(-l)} + u^{(1-l)} + u^{(2-l)}\right) \end{aligned}$$

Expanding and collecting like-terms of x and u yielded

$$\begin{aligned} &(4\sigma + \frac{\delta}{3}\alpha_1)x^{(1)} + (\sigma - I_x)x^{(2)} + [\gamma + \frac{\delta}{3}(\beta_1 + \beta_2)]u^{(0)} + (4\gamma + \frac{\delta}{3}\beta_1)u^{(1)} + \gamma u^{(2)} = -[\sigma + I_x + \\ &\frac{\delta}{3}(\alpha_1 + \alpha_2)]x^{(0)} - \frac{\delta}{3}(\alpha_1 + \alpha_2 + \alpha_3)x^{(-1)} - \frac{\delta}{3}(\alpha_2 + \alpha_3 + \alpha_4)x^{(-2)} - \cdots - \frac{\delta}{3}(\alpha_{d-2} + \alpha_{d-1} + \\ &\alpha_d)x^{(2-d)} - \frac{\delta}{3}(\alpha_{d-1} + \alpha_d)x^{(1-d)} - \frac{\delta}{3}\alpha_d x^{(-d)} - \frac{\delta}{3}(\beta_1 + \beta_2 + \beta_3)u^{(-1)} - \frac{\delta}{3}(\beta_2 + \beta_3 + \beta_4)u^{(-2)} - \\ &\cdots - \frac{\delta}{3}(\beta_{e-2} + \beta_{e-1} + \beta_e)u^{(2-e)} - \frac{\delta}{3}(\beta_{e-1} + \beta_e)u^{(1-e)} - \frac{\delta}{3}\beta_e u^{(-e)}. \end{aligned} \tag{25}$$

Setting k=1 in eqn. (24) yielded:

$$\begin{aligned} &[\sigma + I_x + \frac{\delta}{3}(\alpha_1 + \alpha_2)]x^{(1)} + (4\sigma + \frac{\delta}{3}\alpha_1)x^{(2)} + (\sigma - I_x)x^{(3)} + \frac{\delta}{3}(\beta_1 + \beta_2 + \beta_3)u^{(0)} + [\gamma + \\ &\frac{\delta}{3}(\beta_1 + \beta_2)]u^{(1)} + (4\gamma + \frac{\delta}{3}\beta_1)u^{(2)} + \gamma u_3 = -\frac{\delta}{3}(\alpha_1 + \alpha_2 + \alpha_3)x^{(0)} - \frac{\delta}{3}(\alpha_2 + \alpha_3 + \alpha_4)x^{(-1)} - \cdots \\ &-\frac{\delta}{3}(\alpha_{d-2} + \alpha_{d-1} + \alpha_d)x^{(3-d)} - \frac{\delta}{3}(\alpha_{d-1} + \alpha_d)x^{(2-d)} - \frac{\delta}{3}\alpha_d x^{(1-d)} - \frac{h}{3}(\beta_2 + \beta_3 + \beta_4)u^{(-1)} \\ &-\frac{\delta}{3}(\beta_3 + \beta_4 + \beta)u^{(-2)} - \cdots - \frac{\delta}{3}(\beta_{e-2} + \beta_{e-1} + \beta_e)u^{(3-e)} - \frac{\delta}{3}(\beta_{e-1} + \beta_e)u^{(2-e)} - \frac{\delta}{3}\beta_e u^{(1-e)}. \end{aligned} \tag{26}$$

Setting k=d-4=e-4 in eqn. (24) yielded:

$$\begin{aligned} &(\sigma + I_x)x^{(d-4)} + 4\sigma x^{(d-3)} + (\sigma - I_x)x^{(d-2)} + \gamma u^{(d-4)} + 4\gamma u^{(d-3)} + \gamma u^{(d-3)} = \\ &\frac{\delta}{3}[\sum_{j=1}^{d}\alpha_j(x^{(d-4-j)} + x^{(d-3-j)} + x^{(d-2-j)}) + \sum_{l=1}^{e(\leq d)}\beta_l(u^{(e-4-l)} + u^{(e-3-l)} + u^{(e-2-l)})] \end{aligned}$$

Expanding and collecting the like-terms yielded:

$$\begin{aligned} &\frac{\delta}{3}(\alpha_{d-5} + \alpha_{d-4} + \alpha_{d-3})x^{(1)} + \frac{\delta}{3}(\alpha_{d-6} + \alpha_{d-5} + \alpha_{d-4})x^{(2)} + \cdots + \frac{\delta}{3}(\alpha_{d-4-s} + \alpha_{d-3-s} + \\ &\alpha_{d-2-s})x^{(s)} + \cdots + \frac{\delta}{3}(\alpha_1 + \alpha_2 + \alpha_3)x^{(d-5)} + [\frac{\delta}{3}(\alpha_1 + \alpha_2) + \sigma + I_x]x^{(d-4)} + [\frac{\delta}{3}\alpha_1 + 4\sigma]x^{(d-3)} + \\ &(\sigma - I_x)x^{(d-2)} + \frac{\delta}{3}(\beta_{e-4} + \beta_{e-3} + \beta_{e-2})u^{(0)} + \frac{\delta}{3}(\beta_{e-5} + \beta_{e-4} + \beta_{e-3})u^{(1)} + \frac{\delta}{3}(\beta_{e-6} + \beta_{e-5} + \\ &\beta_{e-4})u^{(2)} + \cdots + \frac{\delta}{3}(\beta_{e-4-s} + \beta_{e-3-s} + \beta_{e-2-s})u^{(s)} + \cdots + \frac{\delta}{3}(\beta_1 + \beta_2 + \beta_3)u^{(e-5)} + [\frac{\delta}{3}(\beta_1 + \beta_2) + \\ &\gamma]u^{(e-4)} + [\frac{\delta}{3}\beta_1 + 4\gamma]u^{(e-3)} + \gamma u^{(e-2)} = -\frac{\delta}{3}(\alpha_{d-4} + \alpha_{d-3} + \alpha_{d-2})x^{(0)} - \frac{\delta}{3}(\alpha_{d-3} + \alpha_{d-2} + \\ &\alpha_{d-1})x_{-1} - \frac{\delta}{3}(\alpha_{d-2} + \alpha_{d-1} + \alpha_d)x^{(-2)} - \frac{\delta}{3}(\alpha_{d-1} + \alpha_d)x^{(-3)} - \frac{\delta}{3}\alpha_d x^{(-4)} - \frac{\delta}{3}(\beta_{e-3} + \beta_{e-2} + \\ &\beta_{e-1})u^{(-1)} - \frac{\delta}{3}(\beta_{e-2} + \beta_{e-1} + \beta_e)u^{(-2)} - \frac{\delta}{3}(\beta_{e-1} + \beta_e)u^{(-3)} - \frac{\delta}{3}\beta_e u^{(-4)} \end{aligned} \tag{27}$$

Setting k=d=e in eqn. (24) yielded:

$$\begin{aligned} &\frac{\delta}{3}(\alpha_{d-1} + \alpha_d)x^{(1)} + \frac{\delta}{3}(\alpha_{d-2} + \alpha_{d-1} + \alpha_d)x^{(2)} + \frac{\delta}{3}(\alpha_{d-3} + \alpha_{d-2} + \alpha_{d-1})x^{(3)} + \frac{\delta}{3}(\alpha_{d-4} + \alpha_{d-3} + \\ &\alpha_{d-2})x^{(4)} + \cdots + \frac{\delta}{3}(\alpha_{d-s} + \alpha_{d+1-s} + \alpha_{d+2-s})x^{(s)} + \cdots + \frac{\delta}{3}(\alpha_1 + \alpha_2 + \alpha_3)x^{(d-1)} + [\frac{\delta}{3}(\alpha_1 + \alpha_2) + \\ &\sigma + I_x]x^{(d)} + [\frac{\delta}{3}\alpha_1 + 4\sigma]x_{d+1} + (\sigma - I_x)x^{(d+2)} + \frac{\delta}{3}\beta_e u^{(0)} + \frac{\delta}{3}(\beta_{e-1} + \beta_e)u^{(1)} + \frac{\delta}{3}(\beta_{e-2} + \beta_{e-1} + \\ &\beta_e)u^{(2)} + \frac{\delta}{3}(\beta_{e-3} + \beta_{e-2} + \beta_{e-1})u^{(3)} + \frac{\delta}{3}(\beta_{e-4} + \beta_{e-3} + \beta_{e-2})u^{(4)} + \frac{\delta}{3}(\beta_{e-5} + \beta_{e-4} + \beta_{e-3})u^{(5)} \\ &+ \cdots + \frac{\delta}{3}(\beta_{e-s} + \beta_{e+1-s} + \beta_{e+2-s})u^{(s)} + \cdots + \frac{\delta}{3}(\beta_1 + \beta_2 + \beta_3)u^{(e-1)} + [\frac{\delta}{3}(\beta_1 + \beta_2) + \gamma]u^{(e)} \\ &+ (\frac{\delta}{3}\beta_1 + 4\gamma)u^{(e+1)} + \gamma u^{(e+2)} = \frac{\delta}{3}\alpha_d x^{(0)}. \end{aligned} \tag{28}$$

Setting k=d+t(1≤t<N−2−d) in eqn. (24) yielded:

$$\frac{\delta}{3}\alpha_d x^{(t)}+\frac{\delta}{3}(\alpha_{d-1}+\alpha_d)x^{(1+t)}+\frac{\delta}{3}(\alpha_{d-2}+\alpha_{d-1}+\alpha_d)x^{(2+t)}+\frac{\delta}{3}(\alpha_{d-3}+\alpha_{d-2}+\alpha_{d-1})x^{(3+t)}$$
$$+\cdots+\frac{\delta}{3}(\alpha_{d-s}+\alpha_{d+1-s}+\alpha_{d+2-s})x^{(s+t)}+\cdots+\frac{\delta}{3}(\alpha_1+\alpha_2+\alpha_3)x^{(d+t-1)}+[\frac{\delta}{3}(\alpha_1+\alpha_2)+$$
$$\sigma+I_x]x^{(d+t)}[\frac{\delta}{3}\alpha_1+4\sigma]x^{(d+t+1)}+(\sigma-I_x)x^{(d+t+2)}+\frac{\delta}{3}\beta_e u^{(t)}+\frac{\delta}{3}(\beta_{e-1}+\beta_e)u^{(t+1)}+\frac{\delta}{3}(\beta_{e-2} \quad (29)$$
$$+\beta_{e-1}+\beta_e)u^{(t+2)}+\frac{\delta}{3}(\beta_{e-3}+\beta_{e-2}+\beta_{e-1})u^{(t+3)}+\cdots+\frac{\delta}{3}(\beta_{e-s}+\beta_{e+1-s}+\beta_{e+2-s})u^{(s+t)}+$$
$$\cdots+\frac{\delta}{3}(\beta_1+\beta_2+\beta_3)u^{(e+t-1)}+[\frac{\delta}{3}(\beta_1+\beta_2)+\gamma]u^{(e+t)}+(\frac{\delta}{3}\beta_1+4\gamma)u^{(e+t+1)}+\gamma u_{(e+t+2)}=0$$

Setting k=N-2 in eqn. (24) yielded:

$$\frac{\delta}{3}\alpha_d x^{(N-2-d)}+\frac{\delta}{3}(\alpha_{d-1}+\alpha_d)x^{(N-1-d)}+\frac{\delta}{3}(\alpha_{d-2}+\alpha_{d-1}+\alpha_d)x^{(N-d)}+\frac{\delta}{3}(\alpha_{d-3}+\alpha_{d-2}+$$
$$\alpha_{d-1})x^{(N+1-d)}+\cdots+\frac{\delta}{3}(\alpha_{d-s}+\alpha_{d+1-s}+\alpha_{d+2-s})x^{(N+s-2-d)}+\cdots+\frac{\delta}{3}(\alpha_1+\alpha_2+\alpha_3)x^{(N-3)}$$
$$+[\frac{\delta}{3}(\alpha_1+\alpha_2)+\sigma+I_x]x^{(N-2)}+[\frac{\delta}{3}\alpha_1+4\sigma]x^{(N-1)}+(\sigma-I_x)x^{(N)}+\frac{\delta}{3}\beta_e u^{(N-2-e)}+\frac{\delta}{3}(\beta_{e-1} \quad (30)$$
$$+\beta_e)u^{(N-1-e)}+\frac{\delta}{3}(\beta_{e-2}+\beta_{e-1}+\beta_e)u^{(N-e)}+\frac{\delta}{3}(\beta_{e-3}+\beta_{e-2}+\beta_{e-1})u^{(N+1-e)}+\cdots+\frac{\delta}{3}(\beta_{e-s}$$
$$+\beta_{e+1-s}+\beta_{e+2-s})u^{(N+s-2-e)}+\cdots+\frac{\delta}{3}(\beta_1+\beta_2+\beta_3)u^{(N-3)}+[\frac{\delta}{3}(\beta_1+\beta_2)+\gamma]u^{(N-2)}$$
$$+(\frac{\delta}{3}\beta_1+4\gamma)u^{(N-1)}+\gamma u^{(N)}=0$$

The combination of all the equations for various values of k forms the linear inequality below:

$$\overline{A}\overline{x}+\overline{B}\overline{u}\le\overline{E}\overline{x}^h+\overline{F}\overline{u}^h=C \quad (31)$$

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

$$\overline{A}\overline{x}=\begin{bmatrix}(4\sigma+\frac{\delta}{3}\alpha_1) & (\sigma-I_x) & 0 & \cdots & \cdots & \cdots & \cdots & 0\\ (\sigma+I_x+\overline{\alpha}_0) & \ddots & \ddots & 0 & \cdots & \ddots & \cdots & \vdots\\ \overline{\alpha}_1 & \ddots & \ddots & \ddots & 0 & \cdots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \vdots\\ \overline{\alpha}_d & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \ddots & 0\\ 0 & \cdots & \overline{\alpha}_d & \cdots & \overline{\alpha}_1 & (\sigma+I_x+\overline{\alpha}_0) & (4\sigma+\frac{\delta}{3}\alpha_1) & (\sigma-I_x)\end{bmatrix}\begin{bmatrix}x^{(1)}\\ x^{(2)}\\ x^{(3)}\\ \vdots\\ x^{(d)}\\ \vdots\\ x^{(N-1)}\\ x^{(N)}\end{bmatrix},$$

$$\overline{B}\overline{u}=\begin{bmatrix}(\gamma+\frac{\delta}{3}\overline{\beta}_0) & (4\gamma+\frac{\delta}{3}\beta_1) & \gamma & 0 & \cdots & \cdots & \cdots & \cdots & 0\\ \overline{\beta}_1 & \ddots & \ddots & \ddots & 0 & 0 & \cdots & \ddots & \vdots\\ \overline{\beta}_2 & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \ddots & \vdots\\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & 0 & 0 & \vdots\\ \overline{\beta}_e & \cdots & \overline{\beta}_2 & \overline{\beta}_1 & \ddots & \ddots & \ddots & \ddots & \vdots\\ \vdots & \ddots & \cdots & \ddots & \ddots & (\gamma+\frac{\delta}{3}\overline{\beta}_0) & (4\gamma+\frac{\delta}{3}\beta_1) & \gamma & 0\\ 0 & \cdots & \overline{\beta}_e & \cdots & \overline{\beta}_2 & \overline{\beta}_1 & (\gamma+\frac{\delta}{3}\overline{\beta}_0) & (4\gamma+\frac{\delta}{3}\beta_1) & \gamma\end{bmatrix}\begin{bmatrix}u^{(0)}\\ u^{(1)}\\ \vdots\\ u^{(e)}\\ \vdots\\ .\\ u^{(N-1)}\\ u^{(N)}\end{bmatrix},$$

$$\overline{E}\overline{x}^h=\begin{bmatrix}-(\sigma+I_x+\overline{\alpha}_0) & -\overline{\alpha}_1 & \cdots & -\overline{\alpha}_s & \cdots & -\overline{\alpha}_{d-1} & -\overline{\alpha}_d\\ -\overline{\alpha}_1 & \cdots & -\overline{\alpha}_s & \cdots & -\overline{\alpha}_{d-1} & -\overline{\alpha}_d & 0\\ \vdots & -\overline{\alpha}_s & \cdots & -\overline{\alpha}_{d-1} & -\overline{\alpha}_d & 0 & \vdots\\ -\overline{\alpha}_s & \cdots & -\overline{\alpha}_{d-1} & -\overline{\alpha}_d & 0 & \vdots & \vdots\\ \vdots & -\overline{\alpha}_{d-1} & -\overline{\alpha}_d & 0 & \vdots & \vdots & \vdots\\ -\overline{\alpha}_{d-1} & -\overline{\alpha}_d & 0 & \vdots & \vdots & \vdots & \vdots\\ -\overline{\alpha}_d & 0 & \vdots & \vdots & \vdots & \vdots & 0\\ 0 & \vdots & \vdots & \vdots & \vdots & 0 & \vdots\\ \vdots & \vdots & \vdots & \vdots & 0 & \vdots & \vdots\\ 0 & \cdots & \cdots & 0 & \cdots & \cdots & 0\end{bmatrix}\begin{bmatrix}x^{(0)}\\ x^{(-1)}\\ x^{(-2)}\\ \vdots\\ \vdots\\ \vdots\\ \vdots\\ x^{(2-d)}\\ x^{(1-d)}\\ x^{(-d)}\end{bmatrix}$$

and

$$\overline{F}\overline{u}^h=\begin{bmatrix}-\overline{\beta}_1 & -\overline{\beta}_2 & \cdots & -\overline{\beta}_{e-1} & -\overline{\beta}_e\\ -\overline{\beta}_2 & \cdots & -\overline{\beta}_{e-1} & -\overline{\beta}_e & 0\\ \vdots & -\overline{\beta}_{e-1} & -\overline{\beta}_e & 0 & \vdots\\ -\overline{\beta}_{e-1} & -\overline{\beta}_e & 0 & \cdots & \vdots\\ -\overline{\beta}_e & 0 & \cdots & \cdots & \vdots\\ 0 & \cdots & \cdots & \cdots & \vdots\\ \vdots & \cdots & \cdots & \cdots & \vdots\\ 0 & \cdots & \cdots & \cdots & 0\end{bmatrix}\begin{bmatrix}u^{(0)}\\ x^{(-1)}\\ u^{(-2)}\\ \vdots\\ \vdots\\ \vdots\\ u^{(1-d)}\\ u^{(-d)}\end{bmatrix},$$

where the delay coefficients $\overline{\alpha}_k$ and $\overline{\beta}_k$, for k=0,1,2,…,d are defined below as:

$$\overline{\alpha}_k=\begin{cases}\frac{\delta}{3}(\alpha_{k+1}+\alpha_{k+2}) & k=0\\ \frac{\delta}{3}(\alpha_k+\alpha_{k+1}+\alpha_{k+2}) & k=1,2,...,d-2\\ \frac{\delta}{3}(\alpha_k+\alpha_{k+1}) & k=d-1\\ \frac{\delta}{3}\alpha_k & k=d\end{cases} \quad (32)$$

and

$$\overline{\beta}_k=\begin{cases}\frac{\delta}{3}(\beta_{k+1}+\beta_{k+2}) & k=0\\ \frac{\delta}{3}(\beta_k+\beta_{k+1}+\beta_{k+2}) & k=1,2,...,e-2\\ \frac{\delta}{3}(\beta_k+\beta_{k+1}) & k=e-1\\ \frac{\delta}{3}\beta_k & k=e\end{cases} \quad (33)$$

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

$$[\overline{a}_{ij}]=\begin{cases}(\sigma+I_x+\overline{\alpha}_0) & i=j\quad 1\le i\le N-1\\ (\sigma-I_x) & j=i+1\quad 1\le i\le N-1\\ (\sigma+I_x+\overline{\alpha}_0) & j=i-1\quad 2\le i\le N-1\\ \overline{\alpha}_k & j=i-1-k\quad 2+k\le i\le N-1;\quad k=1,2,...,d\\ 0 & \text{elsewhere}\end{cases} \quad (34)$$

$$[\overline{b}_{ij}]=\begin{cases}(\gamma+\frac{\delta}{3}\overline{\beta}_0) & i=j\quad 1\le i\le N-1\\ (4\gamma+\frac{\delta}{3}\beta_1) & j=i+1\quad 1\le i\le N-1\\ \overline{\beta}_k & j=i-k\quad 1+k\le i\le N-1;\quad k=1,2,...,e\\ 0 & \text{elsewhere}\end{cases} \quad (35)$$

$$[\overline{e}_{ij}]=\begin{cases}-(\sigma+I_x+\overline{\alpha}_0) & i=j=1\quad 1\le i\le N-1;\\ -\overline{\alpha}_k & j=k+2-i\quad 1\le i\le k+1;\quad k=1,2,...,d\\ 0 & \text{elsewhere}\end{cases} \quad (36)$$

$$[\overline{f}_{ij}]=\begin{cases}-\overline{\beta}_k & j=i+k-1\quad 1\le i\le k;\quad k=1,2,...,e\\ 0 & \text{elsewhere}\end{cases} \quad (37)$$

Analysis of matrix operators

The well-posessness and optimal parameter selections are dependent on the nature and properties of the discretized matrices $\overline{P},\overline{Q},\overline{A}$ and $\overline{B}$. It is then imperative that $\overline{P}$ and $\overline{Q}$ be real symmetric and positive definite to avoid ill-conditioning during the convergence of the algorithm.

Theorem 3.1 (Sylvester's criterion): A real, symmetric matrix is positive definite if and only if all the principal minors are positive definite [11].

Corollary 3.1: Given a matrix $P = \text{diag}[p_1, p_2, \cdots p_n]$ with p_i's the principal diagonal entries and zeros elsewhere. Then the matrix *P* is positive definite if and only if all the p_i's are strictly positive (i.e., no zero principal diagonal entry).

Proof: Let $M_j, j = 1,2,\cdots,n$ be the principal minors of the real, symmetric matrix *P*. Then,

$$M_1 = |p_1| = p_1 > 0$$

$$M_2 = \begin{vmatrix} p_1 & 0 \\ 0 & p_2 \end{vmatrix} = p_1 p_2 > 0$$

$$\vdots$$

$$\underset{1\le j\le n}{M_n} = \begin{vmatrix} p_1 & 0 & \cdots & 0 \\ 0 & \ddots & \ddots & \vdots \\ \vdots & \ddots & p_{n-1} & 0 \\ 0 & \cdots & 0 & p_n \end{vmatrix} = p_1 p_2 \cdots p_n = \prod_{j=1}^{n} p_j > 0 \in \mathbb{R}$$

Since all the M_j's are positive, then *P* is strictly positive definite. Considering the matrices $\bar{P}$ and $\bar{Q}$, all the principal diagonal entries *p*,2*p* and 4*p* are strictly positive and as such they are positive definite.

The modified ADMM formulation

The original ADMM formulation was modified to accommodate both the state and control variables before imposing the Karush-Khun-Tucker (KKT) optimality conditions. It was later accelerated using the Gauss-Seidel accelerator variant to speed up the rate of convergence of the M-ADMM algorithm. The re-formulated compact convex quadratic optimization problem is then stated thus:

$$\min_{\bar{x},\bar{u}} \frac{1}{2}\bar{x}^T \bar{P}\bar{x} + \frac{1}{2}\bar{u}^T \bar{Q}\bar{u} + R \quad s.t \quad \bar{A}\bar{x} + \bar{B} \le C \tag{38}$$

Where,

$\bar{x} \in \mathbb{R}^{nN}, \bar{u} \in \mathbb{R}^{m(N+1)}, \bar{P} \in \mathbb{R}^{nN\times nN}, \bar{Q} \in \mathbb{R}^{m(N+1)\times m(N+1)}, \bar{A} \in \mathbb{R}^{p(N-1)\times nN}, \bar{B} \in \mathbb{R}^{p(N-1)\times m(N+1)}$

and $C \in \mathbb{R}^{p(N-1)}$ and $\bar{P}$ and $\bar{Q}$ are real, symmetric and positive-definite. The associated augmented Lagrangian of eqn. (38) is then given as:

$$\begin{aligned}\min_{\bar{x},\bar{u}} L_\rho(\bar{x},\bar{u},z,\lambda) &= \min_{\bar{x},\bar{u}}\{\frac{1}{2}\bar{x}^T\bar{P}\bar{x} + \frac{1}{2}\bar{u}^T\bar{Q}\bar{u} + R + l_+(z) + \lambda^T(\bar{A}\bar{x} + \bar{B}\bar{u} - C + z) \\ &+ \frac{\rho}{2}\| \bar{A}\bar{x} + \bar{B}\bar{u} - C + z\|_2^2\} \; s.t. \; z \ge 0,\end{aligned} \tag{39}$$

where λ is the Lagrange multiplier, ρ>0 is the penalty parameter, $||.||_2$ is the euclidean (spectral) norm of a vector (matrix) argument, *z* is the introduced slack vector and $l_+(z)$ is the indicator function for the non-negative orthants defined as $l_+(z) = 0$ for *z*≥*0* and $l_+(z) = +\infty$ otherwise.

The scaled augmented Lagrangian is of the form:

$$\begin{aligned}\min_{\bar{x},\bar{u}} L_\rho(\bar{x},\bar{u},z,v) &= \min_{\bar{x},\bar{u}}\{\frac{1}{2}\bar{x}^T\bar{P}\bar{x} + \frac{1}{2}\bar{u}^T\bar{Q}\bar{u} + R + l_+(z) + \\ &+ \frac{\rho}{2}\| \bar{A}\bar{x} + \bar{B}\bar{u} - C + z + v\|_2^2\} \; s.t. \; z \ge 0,\end{aligned} \tag{40}$$

where *v*=*λ*/*ρ* is the scaled dual variable.

The following Karush-Kuhn-Tucker (KKT) optimality conditions in eqn. (4) below were imposed on the Augmented Lagrangian for the derivation of the ADMM algorithm.

$$KKT \Leftrightarrow \begin{cases} \nabla_{\bar{x}} L_\rho\left(\bar{x},\bar{u}^k,\lambda^k\right) = \dfrac{\partial L_p\left(\bar{x},\bar{u}^k,\lambda^k\right)}{\partial \bar{x}} = 0 \\ \nabla_{\bar{u}} L_\rho\left(\bar{x}^{k+1},\bar{u},\lambda^k\right) = \dfrac{\partial L_p\left(\bar{x}^{k+1},\bar{u},\lambda^k\right)}{\partial_{\bar{u}}} = 0 \\ \nabla_{\lambda} L_\rho\left(\bar{x}^{k+1},\bar{u}^{k+1},\lambda\right) = \dfrac{\partial L_p\left(\bar{x}^{k+1},\bar{u}^{k+1},\lambda\right)}{\partial \lambda} = \dfrac{\lambda^{k+1} - \lambda^k}{\rho} \end{cases} \tag{41}$$

In the derivation of the ADMM algorithm, the optimality conditions were then applied to eqn. (40) above through the sequential minimization of $\bar{x}$ and $\bar{u}$. The update of the state variable $\bar{x}$ is of the form:

$$\nabla_{\bar{x}} L_\rho(\bar{x},\bar{u},z,v) = \bar{P}\bar{x} + \rho\bar{A}^T\bar{A}\bar{x} + \rho\bar{A}^T[\bar{B}\bar{u} - C + z + v] \tag{42}$$

Then at optimum, $\nabla_{\bar{x}} L_\rho(\bar{x},\bar{u}^k,z^k,v^k) = 0$ such that $\bar{x} = \bar{x}^{k+1}$ yielded:

$$\bar{x}^{k+1} = -\rho\left(\bar{P} + \rho\bar{A}^T\bar{A}\right)^{-1}\bar{A}^T\left(\bar{B}\bar{u}^k - C + z^k + v^k\right) \quad \bar{x}-\text{update} \tag{43}$$

Likewise, the update of the control variable $\bar{u}$ is of the form:

$$\nabla_{\bar{u}} L_\rho(\bar{x},\bar{u},z,v) = R\bar{u} + \rho\bar{B}^T\bar{B}\bar{u} + \rho B^T[\bar{A}\bar{x} - C + z + v] \tag{44}$$

Then at optimum, $\nabla_{\bar{u}} L_\rho(\bar{x}^{k+1},\bar{u},z^k,v^k) = 0$ such that $\bar{u} = \bar{u}^{k+1}$ yielded:

$$\bar{u}^{k+1} = -\rho\left(\bar{Q} + \rho\bar{B}^T\bar{B}\right)^{-1}\bar{B}^T\left(\bar{A}\bar{x}^{k+1} - C + z^k + v^k\right) \quad \bar{u}-\text{update} \tag{45}$$

The update of the slack variable is of the form:

$$\nabla_z L_\rho(\bar{x},\bar{u},z,v) = \rho(\bar{A}\bar{x} + \bar{B}\bar{u} - C + z + v) \tag{46}$$

Then at optimum, $\nabla_z L_\rho(\bar{x}^{k+1},\bar{u}^{k+1},z,v^k) = 0$ such that $z = z^{k+1}$ yielded:

$$z^{k+1} = \max\{0, -(\bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C + v^k)\} \quad z-\text{update} \tag{47}$$

Introducing the over-relaxation factor $\alpha \in [1.5,1.8]$ into the augmented lagrangian, in the sense of Nesterov [20], by replacing $\bar{A}\bar{x}^{k+1}$ with $h^{k+1} = \alpha^k\bar{A}\bar{x}^{k+1} - (1-\alpha^k)(\bar{B}\bar{u}^k - C + z^k)$ in eqn. (45) yielded:

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

while eqn. (47) yielded:

$$z^{k+1} = max\{0, -h^{k+1} - Bu^{k+1} + C - v^k\}. \tag{49}$$

Upon expansion and re-arrangement of eqn. (49) yielded the over-relaxed $z-$ update

$$\begin{aligned} z^{k+1} &= max\{0, -\alpha(\bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C) - (1-\alpha)\bar{B}(\bar{u}^{k+1} - \bar{u}^k) + (1-\alpha)z^k - v^k\} \\ &= max\{0, -\alpha(\bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C) - (1-\alpha)[\bar{B}(\bar{u}^{k+1} - \bar{u}^k) - z^k] - v^k\} \end{aligned} \tag{50}$$

Updating the dual variable *v* requires that

$$v^{k+1} = v^k + \bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C + z^{k+1} \tag{51}$$

When h^{k+1} was substituted in place of $\bar{A}\bar{x}^{k+1}$ in eqn. (51) yielded:

$$v^{k+1} = v^k + h^{k+1} + \bar{B}\bar{u}^{k+1} - C + z^{k+1} \tag{52}$$

Upon expansion and re-arrangement of eqn. (52) yielded:

$$v^{k+1} = v^k + \alpha(Ax^{k+1} + Bu^{k+1} + z^{k+1} - C) + (1-\alpha)B(u^{k+1} - u^k) \\ +(1-\alpha)(z^{k+1} - z^k) \tag{53}$$

The over-relaxed M-ADMM computes the new iterates as follows:

$$\begin{cases} \bar{x}^{k+1} = -\rho\left(P + \rho\bar{A}^T\bar{A}\right)\bar{A}^T\left(\bar{B}\bar{u}^k - C + z^k + v^k\right) \ \bar{x} - \text{update} \\ \bar{u}^{k+1} = -\rho\left(\bar{Q} + \rho\bar{B}^T\bar{B}\right)^{-1}\bar{B}^T\left[\alpha\left(\bar{A}\bar{x}^{k+1} - C + z^k\right) - (1-\alpha)\bar{B}\bar{u}^k + v^k\right], \\ z^{k+1} = \max\left\{0, -\alpha\left(\bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C\right) - (1-\alpha)\left[\bar{B}\left(\bar{u}^{k+1} - \bar{u}^k\right) - z^k\right] - v^k\right\} \\ v^{k+1} = v^k + \alpha\left(Ax^{k+1} + Bu^{k+1} + z^{k+1} - C\right) + (1-\alpha)B\left(u^{k+1} - u^k\right) \end{cases} \tag{54}$$

Convergence analysis of the modified ADMM

In the development of the M-ADMM algorithm, the augmented Lagrangian functional(ALF) will be deployed because the penalty term associated with the ALF in order to improve the convergence of the M-ADMM algorithm. In literature, the convergence of ADMM algorithm to a solution, in general convex optimization problems, is guaranteed provided the solution exists. Hence the need to carry out spectral analyses regarding the spectrum, symmetry, consistency and positive-definiteness of the matrix operators to ascertain that they were well-posed for the algorithm. However, the limit of the proposed M-ADMM iterates which satisfied the set of first-order optimality conditions produced a certificate of either primal or dual feasibility or both as illustrated below.

In the derivation of the convergence residues, the objective function of the multi-delay optimal control problem was assumed to be p^k at the kth iteration(cycle), which converges to the optimal value p^*($p^k \to p^*$) for large values of k ($k \to \infty$). Suppose the objective function is a closed, proper, convex and sub-differentiable function of f and g expressed on the form $p^k = f^k + g^k$, then the residuals generated at each iteration, known as the dual residual, converges to zero. In the same light, the primal residual $r^k = \bar{A}\bar{x}^k + \bar{B}\bar{u}^k - C$ of the constraint at the k*th* iteration approaches zero as the algorithm approaches optimality. The derivation of the residuals is presented in theorem 5.1 below:

Theorem 5.1: *Let* $p^k = f(\bar{x}^k) + g(\bar{u}^k)$ *be the k-th iterate value of the closed, proper, convex and sub-differentiable objective functions f and g such that it converges to its optimal objective value p*. Given the constraint* $\bar{A}\bar{x} + \bar{B}\bar{u} \le C$ *and multiplier λ, then there exists a dual residual* $d^{k+1} = \rho\bar{A}^T[\bar{B}(\bar{u}^{k+1} - \bar{u}^k) + (v^{k+1} - v^k)]$ *that converges to zero for a given penalty parameter ρ.*

Proof: Given the objective function $p^k = f(\bar{x}^k) + g(\bar{u}^k)$ and linear inequality constraint $\bar{A}\bar{x} + \bar{B}\bar{u} \le C$, the associated Lagrangian with slack z is stated thus:

$$L_\rho(\bar{x},\bar{u},z,\lambda) = f(\bar{x}) + g(\bar{u}) + \lambda^T(\bar{A}\bar{x} + \bar{B}\bar{u} - C + z) + \frac{\rho}{2}\| \bar{A}\bar{x} + \bar{B}\bar{u} - C + z \|_2^2, \tag{55}$$

Applying the optimality conditions (KKT) to obtain

$$\partial f(\bar{x}) + \bar{A}^T\lambda^k + \rho(\bar{A}^T\bar{A}\bar{x} + \bar{A}^T\bar{B}\bar{u} - \bar{A}^TC + 2\bar{A}^Tz) = 0$$
$$\partial f(\bar{x}^{k+1}) + \bar{A}^T\lambda^k + \rho\bar{A}^T(\bar{A}x^{k+1} + \bar{B}\bar{u}^k - C + z^k) = 0$$
$$\partial f(\bar{x}^{k+1}) + \bar{A}^T\lambda^k + \rho\bar{A}^T\underbrace{(\bar{A}\bar{x}^{k+1} + \bar{B}\bar{u}^{k+1} - C + z^{k+1})}_{r^{k+1}} -$$
$$\rho\bar{A}^T\bar{B}\bar{u}^{k+1} - \rho\bar{A}^Tz^{k+1} + \rho\bar{A}^T\bar{B}\bar{u}^k + \rho\bar{A}^Tz^k = 0,$$

where the primal residual is given below as:

$$r^{k+1} = (\bar{A}\bar{x}^{k+1} + B\bar{u}^{k+1} - C + z^{k+1}). \tag{56}$$

Therefore,

$$\partial f(\bar{x}^{k+1}) + \bar{A}^T\lambda^k + \rho\bar{A}^Tr^{k+1} - \rho\bar{A}^T\bar{B}(\bar{u}^{k+1} - \bar{u}^k) - \rho\bar{A}^T(z^{k+1} - z^k) = 0$$
$$\rho\bar{A}^T[\bar{B}(\bar{u}^{k+1} - \bar{u}^k) + (z^{k+1} - z^k)] = \partial f(\bar{x}^{k+1}) + \bar{A}^T\underbrace{[\lambda^k + \rho r^{k+1}]}_{\lambda^{k+1}}$$

Since at the ADMM, the update $\partial f(\bar{x}^{k+1}) + \bar{A}^T\lambda^{k+1} \to 0$, then its dual residual

$$d^{k+1} = \rho\bar{A}^T[\bar{B}(\bar{u}^{k+1} - \bar{u}^k) + (z^{k+1} - z^k)] \to 0 \tag{57}$$

which completes the proof. Q.E.D.

The convergence of the primal-dual feasibility to zero in eqn. (57) is a clear indication that the algorithm is superlinearly convergent.

Theorem 5.2: *Given a convex quadratic programming (QP) problem* $\min\frac{1}{2}y^TPy + q^Ty$ *such that* $Ay \le b$*, where* $P \in \mathbb{R}^{n\times n}, y \in \mathbb{R}^n, A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$*, then the optimal step-size for the QP is*

$$\rho^* = \left[\sqrt{\lambda_{\min}\left(AP^{-1}A^T\right)\lambda_{\max}\left(AP^{-1}A^T\right)}\right]^{-1} \tag{58}$$

and the convergence factor

$$\xi_R^* = \frac{\lambda_{\max}\left(AP^{-1}A^T\right) - \sqrt{\lambda_{\min}\left(AP^{-1}A^T\right)\lambda_{\max}\left(AP^{-1}A^T\right)}}{\lambda_{\max}\left(AP^{-1}A^T\right) + \sqrt{\lambda_{\min}\left(AP^{-1}A^T\right)\lambda_{\max}\left(AP^{-1}A^T\right)}} \tag{59}$$

for $\alpha \in (0,2], \lambda_{\min}(AP^{-1}A^T)$ and $\lambda_{\max}(AP^{-1}A^T)$ are the minimum and maximum eigenvalues of the matrix $(AP^{-1}A^T)$ respectively, Ghadimi et al. [10].

The result of optimal parameter selection (stepsize) of the convex optimization problem stated above by Ghadimi et al. [10] can be extended to a convex optimal control problem of the form,

$$Min\frac{1}{2}\bar{x}^T\bar{P}\bar{x} + \bar{u}^T\bar{Q}\bar{u} + R \quad s.t \quad \bar{A}\bar{x} + \bar{B}\bar{u} \le C \tag{60}$$

where $\bar{x} \in \mathbb{R}^{nN\times 1}$, $\bar{u} \in \mathbb{R}^{m(N+1)\times m(N+1)}$,

$\bar{P} \in \mathbb{R}^{nN\times nN}, \bar{Q} \in \mathbb{R}^{m(N+1)\times m(N+1)}, \bar{A} \in \mathbb{R}^{p(N-1)\times nN}, R \in \mathbb{R}, \bar{B} \in \mathbb{R}^{p(N-1)\times m(N+1)}$

and $C \in \mathbb{R}^{p(N-1)\times 1}$. Therefore, eqn. (60) above can be re-formulated into a convex optimal control problem of the form:

$$Min\frac{1}{2}\hat{w}^T\hat{P}\hat{w} + R \quad s.t \quad \hat{A}\hat{w} \le C \tag{61}$$

where,

$\hat{w} = (\bar{x},\bar{u}) \in \mathbb{R}^{p(N-1)\times(nN+mN+m)}$, $\hat{A} = \left[\bar{A},\bar{B}\right] \in \mathbb{R}^{p(N-1)\times(nN+mN+m)}$,

$$\hat{P} = \begin{bmatrix} [c\,|\,c]\bar{P} & \hat{0} \\ \hat{0}^T & \bar{Q} \end{bmatrix} \in \mathbb{R}^{(nN+mN+m)\times(nN+mN+m)}$$ and

$\hat{0} \in \mathbb{R}^{nN\times m(N+1)}$. Then, the optimal stepsize for the concatenation of the DOCP, by theorem (5.2) above, is:

$$\rho^* = \left[\sqrt{\lambda_{\min}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)\lambda_{\max}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)}\right]^{-1} \tag{62}$$

and the convergence factor is then,

$$\xi_{\hat{R}}^* = \frac{\lambda_{\max}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right) - \sqrt{\lambda_{\min}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)\lambda_{\max}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)}}{\lambda_{\max}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right) + \sqrt{\lambda_{\min}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)\lambda_{\max}\left(\hat{A}\hat{P}^{-1}\hat{A}^T\right)}} \tag{63}$$

for $\alpha \in (1,2], 0 < \xi^*_{\hat{R}} < 1$ and $\hat{P}$ is symmetric and positive definite.

The primal $\varepsilon^{Prim} > 0$ and dual $\varepsilon^{Dual} > 0$ residuals, selected as termination (stopping) criterion for the convergence of the ADMM, were so small such that $\| r^{k+1} \|_2 \le \varepsilon^{Prim}$ and $\| d^{k+1} \|_2 \le \varepsilon^{Dual}$. However, the choices of our tolerances depend on the relative and absolute criteria on account that the ℓ_2 norms are in $\mathbb{R}^n$ and $\mathbb{R}^p$, respectively. Usually in practice or literature, the following values, $\varepsilon^{rel} = 10^{-3}$ and $\varepsilon^{abs} = 10^{-4}$, are used as reasonable stopping criteria for the ADMM algorithms. Stated below are basic computations in literature as in Boyd et al. [7].

$$\varepsilon^{Prim} = \sqrt{p}\varepsilon^{abs} + \varepsilon^{rel}.max\{\| Ax^{k+1} \|_2, \| Bu^{k+1} \|_2, \| z^{k+1} \|_2, \| C \|_2\} \quad (64)$$

$$\varepsilon^{Dual} = \sqrt{n}\varepsilon^{abs} + \varepsilon^{rel}.\| \rho u |_2 \quad (65)$$

The formulated algorithm:

Algorithm 3.3: Modified-ADMM For multiple DOCP,

Step 0 : **Input** $A, B, C, P, Q, R, \rho, \alpha, Tol., \varepsilon^{Prim}, \varepsilon^{Dual}$

Step 1 : **Initialize** $x^{0,0}, u^{0,0}, \lambda^{0,0}, V^{0,0}$

Step 2 : **Set** $k = 0$

Step 3 : **Compute**

$x^{k+1,i} = -\rho(P + \rho A^T A)^{-1} A^T (Bu^{k,i} - C + z^{k,i} + v^{k,i})$ eqn. (43)

Step 4 : **Stop** if $\| x^{k,i+1} - x^{k,i} \| \le Tol.$

Step 5 : **Compute**

$u^{k,i+1} = -\rho(Q + \rho B^T B)^{-1} B^T[\alpha(Ax^{k+1,i} - C + z^{k,i}) - (1-\alpha)Bu^{k,i} + v^{k,i}]$ eqn. (48)

Step 6 : **Stop** if $\| u^{k,i+1} - u^{k,i} \| \le Tol.$

Step 7 : **Output** x^{k+1}, u^{k+1} and go to step 9 otherwise

Step 8 : **Repeat** steps 3 & 5 for $i = 1, 2, \cdots$ **until** steps 4 & 6 are satisfied

Step 9 : **Compute**

$z^{k+1} = Max\{0, -\alpha(Ax^{k+1} + Bu^{k+1} - C) + (1-\alpha)[B(u^{k+1} - u^k) - z^k] - v^k\}$ eqn. (50)

$$v^{k+1} = v^k + \alpha(Ax^{k+1} + Bu^{k+1} - C + z^{k+1}) + (1-\alpha)B(u^{k+1} - u^k) + (1-\alpha)(z^{k+1} - z^k) \quad \text{eqn. (53)}$$

Step 10 : **Stop** if $\| r^{k+1} \|_2 = \sqrt{(Ax^{k+1} + Bu^{k+1} - C + z^k)} \le \varepsilon^{Prim}$ and eqn. (56)

$\cdot \; \| d^{k+1} \|_2 = \sqrt{\rho A^T[B(u^{k+1} - u^k) + (z^{k+1} - z^k)]} \le \varepsilon^{Dual}$ else go to step 12 eqn. (57)

Step 11 : **Output** x^{k+1}, u^{k+1} and **Compute**

ρ^* (optimal steplength) eqn. (58)

$\xi^*_{\bar{R}}$ (Convergence profile) eqn. (59)

Step 12 : **Set** $k = k + 1$ and go to Step 2.

Results and Discussion

The multi-delay optimal control problem below was considered for the implementation of the algorithm:

$$MinJ(x,u) = \frac{1}{2}\int_0^{0.5} [2x_1^2 + x_1x_2 + x_2^2 + u_1^2 + u_1u_2 + u_2^2]dt$$

$$\begin{aligned} s.t \quad \dot{x}_1(t) &= 2x_1(t) + x_2(t) + u_1(t) + 3u_2(t) + x_1(t-0.1) + x_2(t-0.1) + 2x_1(t-0.2) + \\ &\quad x_2(t-0.2) - x_1(t-0.3) + u_1(t-0.1) + u_2(t-0.1) + u_1(t-0.3) + 2u_2(t-0.2), \\ \dot{x}_2(t) &= x_1(t) - u_1(t) + 2u_2(t) - x_1(t-0.1) + x_1(t-0.2) + 2x_2(t-0.2) - x_1(t-0.3) + \\ &\quad u_2(t-0.1) + u_1(t-0.2) + 3u_2(t-0.2), \end{aligned}$$

$$x(0) = (0,0); \qquad 0 \le t \le 0.5,$$

$$x(t) = (2t+1, t^2+1); \qquad -0.3 \le t \le 0,$$

$$u(t) = (2, 2+t); \qquad -0.2 \le t \le 0,$$

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

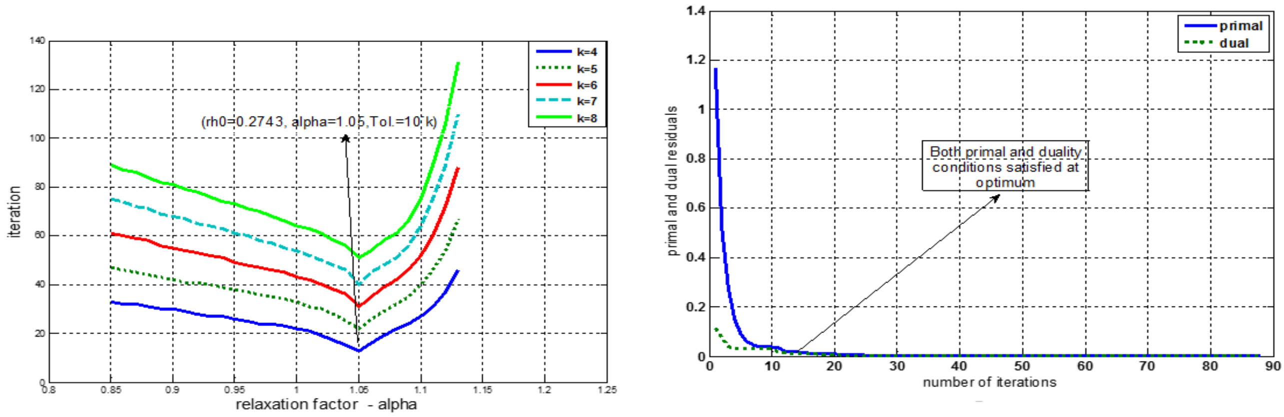
In adopting the M-ADMM algorithm, the coefficient matrices are defined below as:

$$P = \begin{bmatrix} 4 & 1 \\ 1 & 2 \end{bmatrix}, Q = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, A = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}, B = \begin{bmatrix} 1 & 3 \\ -1 & 2 \end{bmatrix}, \alpha_1 = \begin{bmatrix} 1 & 1 \\ -1 & 0 \end{bmatrix},$$

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

The coefficient matrices, $\bar{P}$ and $\bar{Q}$, 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. However, the algorithm runs for various numbers of iterations plotted against 29 values of the relaxation parameters α evenly spaced between 0.85 and 1.13. At each iteration (cycle), the penalty parameter ρ was kept fixed at its optimum $\rho^* = 0.2743$, based on the step-size heuristically computed from eqn. (58). As shown in Figure 1A, 5 values of the dual and primal tolerances 10^{-k} for k=4,5,..8 geometrically spaced between 10^{-4} and 10^{-8}. Each line on the graph corresponds to a fixed value of the dual and primal tolerance (10^{-k}), where the relaxation factor runs over the entire chosen feasible values α∈[0,2]. Clearly, each line converges at the point $\alpha^* = 1.05$, which demonstrates the *optimum over-relaxation factor* for reducing values of α, though few values of *k* were chosen to keep the plot manageable and the cycle of the algorithm not exceeding 200 iterations Figure 1B illustrates the convergence of the M-ADMM iterations for the Quadratic program at the optimum step-size ρ^* and over-relaxation factor $\alpha \in [0,2]$. It demonstrates the responses of the dual and primal residuals for various iteration counters at the optimum relaxation factor $\alpha^* = 1.05$ and at the optimum penalty parameter of $\rho^* = 0.2743$. When comparing the convergence of the dual and primal residuals at start, the dual residual r_k is far lower than that the primal residual d^k which gives it high propensity to satisfy the feasibility conditions than that of the primal. However, the rate of convergence of the primal residual towards zero for large iteration ($\| d^k \|_2^2 \to 0$ as $k \to \infty$) is faster than that of the dual as demonstrated in Figure 1B.

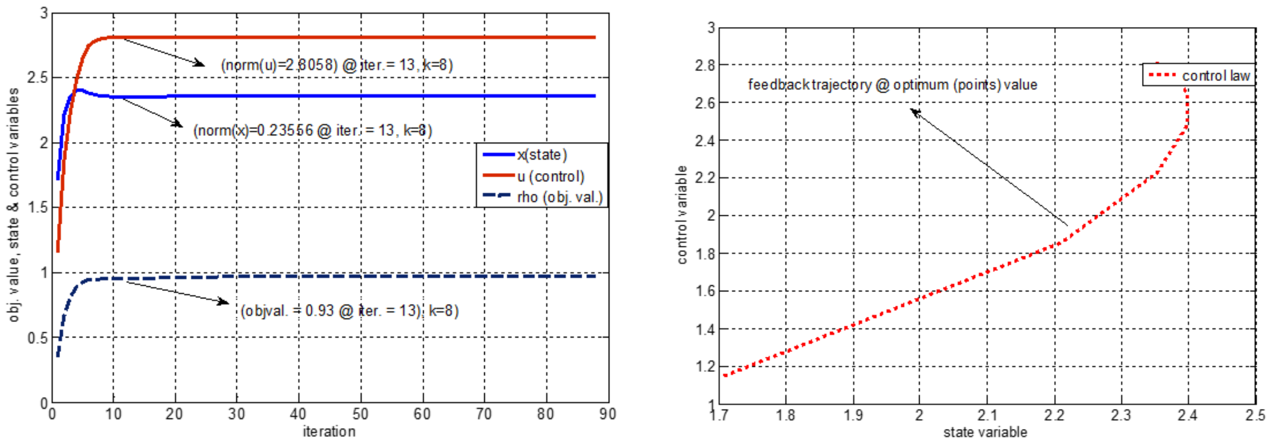
Figure 2A demonstrates the responses of the state-control trajectories and the objective values to increasing iterations at chosen tolerance, fixed relaxation factor and optimal penalty parameter ($10^{-k}, \alpha^*, \rho^*) = (10^{-8}, 1.05, 0.2743)$. However, it was observed that the objective values increase with increasing values of the state and control variables until the optimum objective value $p^* = 0.93$ was arrived at with the optimum trajectories $w^* = (x^*, u^*) = (2.8058, 0.2356)$. Moreover, Figure 2B is the feedback law which demonstrates the degree of relationship between the state *x*(*t*) and control *u*(*t*) trajectories within the specified time domain [0,0.5]. The control (feedback) law is a *time-variant* measure unlike the proportional control system where the feedback law is *time-invariant* (real constant).



(a) Response of Algorithm to Relaxation factor

(b) Effects of Primal-Dual on convergence $||r^k||_2 \longrightarrow 0$ and $||d^k||_2^2 \longrightarrow 0$ as $k \longrightarrow \infty$

Figure 1: The operators for algorithm.



(a) Optimal state-control trajectories

(b) Feedback law

Figure 2: The optimum parameters.

Conclusion and Future Work

The solution to optimal control problem with multiple constant delays using the modified alternating direction method have been established. The convergence of the algorithm was discovered to be at a superlinear rate and provides an explicit expressions for the optimum convergence parameter. We also considered the over-relaxation of the M-ADMM, for various values of the penalty parameters, for which the algorithm is guaranteed to convergence at its optimum. Numerical example was given to establish the effectiveness and general performance of the algorithm in terms of its lower iteration cycle and time in obtaining the optimal solution to the multi-delay control problem. In the future, the algorithm can be extended to other classes of control problems with objective functional of the Bolza-type constrained by bounded or mixed inequality constraints.

References

1. Angell TS, Kirsch A (1990) On the necessary conditions for optimal control of retarded systems. Appl Math and Optimization 22:117-145.
2. Bakke VL (1981) Optimal Fields for Problems with Delays. Opt Theo and Appl 33: 567-593.
3. Bashier EBM, Patider KC (2017) Optimal Control of an Epidemiological Model with Multiple Time Delay. Journal of Applied Mathematics and Computation 292 :47-56.
4. Bertsekas DP (1982) Constrained Optimization and Lagrange Multiplier methods (1stedn), Academic press, Inc., Belmont, USA.
5. Boley D (2013) Local Linear Convergence of the Alternating Direction Method of Multipliers on Quadratic or Linear Programs. SIAM Journal on Optimization 23: 2183-2207.
6. Boyd S, Vandenberghe L (2004) Convex Optimization, New York, USA, Cambridge University Press.
7. Boyd S, Parikh N,Chu E, Peleato B, Eckstein J (2011) Distributed optimization and statistical learning via the alternating direction method of multipliers. Foundations and Trends in Machine Learning 3: 11-22.
8. Chyung DH, Lee EB (1966) Linear Optimal Systems with Time Delays. SIAM J Control 4: 548-575.
9. Colonius F, Hinrichsen D (1978) Optimal Control of Functional Differential Systems. SIAM Journal on Control and Optimization 16: 861-879.

10. Ghadimi E, Teixeira A, Shames I, Johansson M (2014) Optimal Parameter Selection for the Alternating Direction Method of Multipliers (ADMM): Quadratic Problems. IEEE Transactions on Automatic Control 60(3): 644-658.

11. Gilbert GT (1991) Positive Definite Matrices and Sylvester's Criterion. The American Mathematical Monthly 98: 44-46.

12. Goldfarb D, Schneiberg K (2012) Fast Alternating Linearization Methods for Minimizing the Sum of Two Convex Functions. Mathematical Programming 141: 349-382.

13. Gollman LD, Kern D, Maurer H (2008) Optimal Control Problems with Delays in State and Control Variables Subject to Mixed Control-State Constraints. Optimal Control Applications and Methods 30: 341-365.

14. Gollman LD, Maurer H (2014) Theory and Applications of Optimal Control Problems with Multiple Delays. Journal of Industrial and Management Optimization 10: 413-441.

15. Harastivilli GL (1961)The Principle in the Theory of Procces Involving Delay. Dokl Akad, Nauk SSSR 136, pp: 39-42.

16. Kadkhodaie M, Christakopoulou K, Sanjabi M, Banerjee A (2015) Accelerated alternating direction method of multipliers. In: Proceedings of the 21th ACM SIGKDD international conference on knowledge discovery and data mining, ACM, pp: 497-506.

17. Kharatishvili GL (1961) The maximum principle in the theory of optimal processes involving delay. Doklady Akademii Nauk USSR 136: 39-42.

18. Kharastishvilli GL (1967) A Maximum Principle in External Problems with Delays. In: Balakrishnan AV, Neustadt LW (eds.), Mathematical Theory of Control; Academic Press, New York, USA.

19. Laarabi H, Abta A, Hahaf K (2015) Optimal Control of a Delayed SIRS Epidemic Model with Vaccination and Treatment. Acta Biotheor 63: 87-97.

20. Nesterov YE (1983) A method for Solving the Convex Programming Problem with Convergence Rate $O(1/k^2)$. Dokl Akad Nauk SSSR 269: 543-547.

21. O'Donoghue B, Stathopoulos G, Boyd S (2013) A Splitting Method for Optimal Control. IEEE Trans. Control Systems Technology 21: 2432-2442.

22. Olotu O, Dawodu KA, Adebowale K (2013) On the Discretized Algorithm for Optimal Proportional Control Constrained by Delay Differential Equation. Journal of Mathematical Theory and Modelling by IISTE, USA 3, pp: 157-169.

23. Rihan FA, Anwar MN (2012) Qualitative Analysis of Delayed SIR Epidemic Model with a Saturated Incidence rate. Intl J Diff Equation 408637:1-13.

24. Soliman MA (1972) Optimal Feedback Control for Linear-Quadratic Systems having Time Delays. Int J Control 15: 609-627.

25. Xue Y, Duan X (2011) Dynamic Analysis of an SIR Epidemic Model with Nonlinear Incidence rate and Double Delays. Intl J Inf Syst Sci Computer Inf 7: 92-102.

26. Yang J, Zhang Y (2011) An Alternating Direction Algorithms for l_1 Problems in Compressive Sensing. SIAM J Sci Comput 33: 250-278.

27. Zhang X, Berger M, Osher M (2011) A Unified Primal-Dual Algorithm Framework Based on Bregman Iteration. Journal of Scientific Computing 1: 20-46.