Linear quadratic control computation for systems with a directed tree structure

Armaghan Zafar * Farhad Farokhi * Michael Cantoni *

* Department of Electrical and Electronic Engineering, The University of Melbourne, Parkville VIC 3010, Australia (e-mail: armaghanz@student.unimelb.edu.au, cantoni,ffarokhi)@unimelb.edu.au).

Abstract: A computational method is proposed for solving a structured quadratic optimal control problem subject to linear discrete-time dynamics arising from a directed tree structured interconnection of heterogeneous sub-systems. The problem is first formulated as a quadratic program with structure along three dimensions of the decision space. A nested algorithm based on block Jacobi iterations is proposed for the linear system of equations obtained from the corresponding first-order optimality conditions. It is shown that the per iteration computational burden scales favorably with increasing problem size in each dimension. The computations at each iteration are amenable to distributed implementation on a network of parallel processors mirroring the tree graph structure of the problem. Numerical experiments, based on a model data for an automated irrigation network, are used to demonstrate aspects of the approach, including the impact of early termination of the inner iterations in agreement with corresponding analysis.

NOTATIONS

Throughout this document, \( \mathbb{R} := (-\infty, +\infty) \) is used to denote the set of real numbers and \( \mathbb{N} := \{1, 2, \ldots\} \) is the set of natural numbers. \( \mathbb{R}^n \) denotes an \( n \)-dimensional real vector and \( \mathbb{R}^{p \times q} \) denotes a real matrix with \( p \) rows and \( q \) columns. \( A \backslash \{b\} \) denotes all elements of the set \( A \) except element \( b \). \( I_q \) is the identity matrix of size \( q \times q \) and \( 0_{p \times q} \) denotes a matrix of zeros of size \( p \times q \). \( A^\prime \) denotes the transpose of a matrix. \( A \succ 0 \) means the symmetric matrix \( A = A^\prime \in \mathbb{R}^{n \times n} \) is positive definite (i.e., there exists \( c > 0 \) such that \( x^\prime A x \geq c x^\prime x \) for all \( x \in \mathbb{R}^n \)) and \( A \succeq 0 \) means \( A \) is a positive semi-definite (i.e., \( x^\prime A x \geq 0 \) for all \( x \in \mathbb{R}^n \)). “\( \text{diag}(\cdot)'' \) represents construction of block diagonal matrix from input arguments and “\( \text{col}(\cdot)'' \) represents concatenation of input arguments as column vector. \( \otimes \) represents the Kronecker product.

1. INTRODUCTION

In this paper, computational aspects of solving structured large-scale finite-horizon linear-quadratic (LQ) optimal control problems are investigated. Specifically problems subject to linear discrete-time dynamics arising from the interconnection of heterogeneous sub-systems by a directed tree graph. The key features are that (i) each node has in degree no greater than 2, (ii) the out degree is 1 for all nodes, except one node with no children, and (iii) the sub-graph comprised of the branch nodes is a directed path, where a branch node means a node of in degree 2 or more. This kind of structure is relevant in various domains, including water irrigation networks (Li et al., 2005), supply chains (Perea-López et al., 2003), and radial power distribution networks (Peng and Low, 2018).

Fig. 1. Directed information flow in a tree graph network

To help clarify, consider a resource distribution network consisting of a primary channel and secondary distribution channels. Suppose that resource flow regulators are located at the source of and along channels immediately downstream of every supply point to a secondary channel or end-user as the case may be. Each of these regulators is controlled to achieve a specified level of the resource in the buffer at the supply points immediately downstream, based on observations of this level and the downstream flow load. Such a feedback control architecture results in automatic demand driven release of the resource from the source; e.g., see (Cantoni et al., 2007) for related work on irrigation networks. Indeed, the information flow corre-

2405-8963 Copyright © 2020 The Authors. This is an open access article under the CC BY-NC-ND license. Peer review under responsibility of International Federation of Automatic Control.
10.1016/j.ifacol.2020.12.069
The set of natural numbers.

\( N \)

denote the set of real numbers and \( R \) real vector and

\( c > 0 \) throughout this document, supply chains (Perea-López et al., 2003), and radial power distribution networks (Peng and Low, 2018). Various domains, including water irrigation networks (Li et al., 2020), are used to demonstrate aspects of the approach, including mirroring the tree graph structure of the problem. Numerical experiments, based on a model data program with structure along three dimensions of the decision space. A nested algorithm based on block Jacobi iteration is presented. The convergence of these nested Jacobi iterations including error analysis when the inner iterations are terminated early is considered in Section 4. The proposed method is applied to determine the set-point references for an automated irrigation network model. Computation time results for a single processor implementation are presented in Section 5. Finally, some concluding remarks are given in Section 6.

2. PROBLEM FORMULATION

For the sake of simplifying the arguments and notation, the following assumptions are made:

Assumption 1. The number of pools in each secondary channel is equal to \( N_1 = N_2 = \ldots = N_K = N \) represents the common set of secondary channel indexes.

Remark 2. Assumption 1 only aids in the simplification of notations and clarity of presentation. All subsequent developments remain valid if Assumption 1 does not hold. In such a case, the size of the block matrices presented in equation (4) will be different for each index \( j \) and the limit of summation goes to \( N = \max(N_i) \).

Following (Cantoni et al., 2017) and (Zafar et al., 2019), where cascade (i.e., directed path) interconnections are studied within an LQ optimal control context, by defining, \( u_{i,j} := \text{col}(u_{i,j}(0), \ldots, u_{i,j}(T-1)) \in R^{m_j \times T \times T} \) and \( x_{i,j} := \text{col}(x_{i,j}(0), \ldots, x_{i,j}(T)) \in R^{n_j \times (T+1)} \), the problem (2) can be reformulated as the following temporally stacked QP:

\[
\begin{align*}
\min_{x_{i,j}, u_{i,j}} & \quad \frac{1}{2} \sum_{i=1}^{K} \sum_{j=1}^{N_i} \left( \sum_{t=0}^{T-1} [x_{i,j}(t)]' [Q_{i,j} 0 0] [x_{i,j}(t)] + x_{i,j}(T) P_{i,j} x_{i,j}(T) \right) \\
\text{subject to} & \\
& x_{i,j}(0) = \xi_{i,j} \quad \text{for } i \in \mathcal{K}, \quad j \in \mathcal{N}_i, \\
& x_{i,j}(t+1) = A_{i,j} x_{i,j}(t) + B_{i,j} u_{i,j}(t) + E_{i,j} x_{i,j+1}(t) + F_{i,j} x_{i,j+1}(t) \quad \text{for } i \in \mathcal{K}, \quad j \in \mathcal{N}_i, \quad t \in T \setminus \{T\}, \quad \mathcal{E}_{i,j} \in R^{n_{i,j} \times n_{i,j}}, \quad Q_{i,j} \in R^{n_{i,j} \times n_{i,j}} \quad \text{and} \quad P_{i,j} \in R^{n_{i,j} \times n_{i,j}} \quad \text{are symmetric positive definite matrices.} \end{align*}
\]
where

\[
Q_{i,j} = \begin{bmatrix} I_T \otimes Q_{i,j} & 0_{n_{i,j},T \times n_{i,j}^2} \\ 0_{n_{i,j},T \times n_{i,j}^2} & P_{A,i,j} \end{bmatrix} \in \mathbb{R}^{n_{i,j}(T+1) \times n_{i,j}(T+1)},
\]

\[
\bar{R}_{i,j} = I_T \otimes R_{i,j} \in \mathbb{R}^{m_{i,j}(T) \times m_{i,j}(T)},
\]

\[
\bar{H}_{i,j} = \text{col}(I_{n_{i,j},0} n_{i,j} \times \cdots \times 0_{n_{i,j},n_{i,j}}) \in \mathbb{R}^{n_{i,j}(T+1) \times n_{i,j}},
\]

\[
\bar{A}_{i,j} = -I_{n_{i,j}(T+1)} + \begin{bmatrix} 0_{n_{i,j},T \times n_{i,j}} & 0_{n_{i,j},n_{i,j}} \\ I_T \otimes A_{i,j} & 0_{n_{i,j}T \times n_{i,j}} \end{bmatrix},
\]

\[
\bar{B}_{i,j} = \begin{bmatrix} 0_{n_{i,j},m_{i,j}(T)} \\ I_T \otimes B_{i,j} \end{bmatrix},
\]

\[
\bar{F}_{i,j} = \begin{bmatrix} 0_{n_{i,j},T \times n_{i,j}} \\ I_T \otimes F_{i,j} \end{bmatrix},
\]

for \(i \in K, j \in N\). \(\bar{A}_{i,j} \in \mathbb{R}^{n_{i,j}(T+1) \times n_{i,j}(T+1)}\) and \(\bar{B}_{i,j} \in \mathbb{R}^{n_{i,j}(T+1) \times n_{i,j}(T+1)}\) and \(\bar{F}_{i,j} = 0\) for all \((i,j)\) \(i = K, j \neq 1\). \(\bar{E}_{i,j} \in \mathbb{R}^{n_{i,j}(T+1) \times n_{i,j}(T+1)}\) and \(\bar{E}_{i,N} = 0\) (so that effectively \(\bar{e}_{i,N+1} = 0\)). Notice that matrices in the temporally stacked QP (3) are block diagonal except matrices \(\bar{A}_{i,j}\) which are block bi-diagonal.

Now defining \(\hat{n}_i := \sum_{i=1}^{K} n_{i,j}(T+1)\), \(\hat{m}_i := \sum_{i=1}^{K} m_{i,j} T\), \(\hat{v}_i := \sum_{i=1}^{K} n_{i,j}, \hat{u}_i := \text{col}(\bar{u}_{i,1}, \ldots, \bar{u}_{i,j}) \in \mathbb{R}^{\hat{m}_i}\) and \(\hat{x}_i := \text{col}(\bar{x}_{i,1}, \ldots, \bar{x}_{i,j}) \in \mathbb{R}^{\hat{n}_i}\), problem (3) can be reformulated by stacking across the dimension of the primary channel which results in the following sparse and structured quadratic program (QP).

\[
\min \frac{1}{2} \sum_{j \in N} \left( \begin{bmatrix} \hat{x}_j \\ \hat{u}_j \end{bmatrix} \right)^T \begin{bmatrix} \hat{Q}_j & 0 \\ 0 & \hat{R}_j \end{bmatrix} \begin{bmatrix} \hat{x}_j \\ \hat{u}_j \end{bmatrix} + \hat{A}_j \hat{x}_j + \hat{B}_j \hat{u}_j + \hat{E}_j \hat{x}_{j+1} + \hat{H}_j \hat{\xi}_j, \quad j \in N,
\]

(4a)

subject to

\[
0 = \hat{A}_j \hat{x}_j + \hat{B}_j \hat{u}_j + \hat{E}_j \hat{x}_{j+1} + \hat{H}_j \hat{\xi}_j, \quad j \in N,
\]

(4b)

where

\[
\hat{A}_j = \text{diag}(\hat{A}_{1,j}, \ldots, \hat{A}_{K,j}), \quad j \in N \setminus \{1\},
\]

\[
\hat{A}_1 = \begin{bmatrix} \hat{A}_{1,1} & \hat{F}_{1,1} \\ \vdots & \ddots \\ \hat{F}_{K-1,1} & \hat{A}_{K,1} \end{bmatrix} \in \mathbb{R}^{\hat{n}_j \times \hat{n}_j},
\]

\[
\hat{Q}_j = \text{diag}(\hat{Q}_{1,j}, \ldots, \hat{Q}_{K,j}) \in \mathbb{R}^{\hat{n}_j \times \hat{n}_j}, \quad \hat{R}_j = \text{diag}(\hat{R}_{1,j}, \ldots, \hat{R}_{K,j}) \in \mathbb{R}^{\hat{m}_j \times \hat{m}_j}, \quad \hat{B}_j = \text{diag}(\hat{B}_{1,j}, \ldots, \hat{B}_{K,j}) \in \mathbb{R}^{\hat{n}_j \times \hat{m}_j}, \quad \hat{E}_j = \text{diag}(\hat{E}_{1,j}, \ldots, \hat{E}_{K,j}) \in \mathbb{R}^{\hat{n}_j \times \hat{n}_j}, \quad \hat{H}_j = \text{diag}(\hat{H}_{1,j}, \ldots, \hat{H}_{K,j}) \in \mathbb{R}^{\hat{n}_j \times \hat{n}_j},
\]

for \(j \in N\). The variables \(\hat{p}_j = \text{col}(\bar{p}_{1,j}, \ldots, \bar{p}_{K,j}) \in \mathbb{R}^{\hat{n}_j}\) are Lagrange multipliers. Since \(\hat{Q}\) and \(\hat{R}\) are positive definite, the problem is strictly convex and the KKT conditions are necessary and sufficient for optimality (Nocedal and Wright, 2000).

Defining \(\hat{n} = \sum_{j=1}^{N} (\hat{n}_j), \hat{m} = \sum_{j=1}^{N} (\hat{m}_j), \hat{v} = \sum_{j=1}^{N} (\hat{v}_j), \hat{x} = \text{col}(\bar{x}_1, \ldots, \bar{x}_N), \hat{u} = \text{col}(\bar{u}_1, \ldots, \bar{u}_N), \hat{p} = \text{col}(\bar{p}_1, \ldots, \bar{p}_N)\), the KKT conditions can be written as

\[
\begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix} \begin{bmatrix} \hat{x} \\ \hat{u} \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{p} \end{bmatrix},
\]

(5)

where

\[
\hat{A} = \begin{bmatrix} \hat{A}_1 & \hat{A}_2 & \cdots \\ \vdots & \ddots & \vdots \\ \hat{A}_{N-1} & \cdots & \hat{A}_N \end{bmatrix} \in \mathbb{R}^{\hat{n} \times \hat{n}}.
\]

3. BLOCK JACOBI ITERATIONS

Defining \(\hat{\nu} := \text{col}(\hat{x}, \hat{u})\), (5) can be written as

\[
\begin{bmatrix} Z & Y' \end{bmatrix} \begin{bmatrix} \hat{v} \\ \hat{p} \end{bmatrix} = \begin{bmatrix} 0 \\ -\hat{r}_{\hat{p}} \end{bmatrix},
\]

(6)

where

\[
Z = \begin{bmatrix} \hat{Q} & 0 \\ 0 & \hat{R} \end{bmatrix} \quad \text{and} \quad Y = \begin{bmatrix} \hat{A} \end{bmatrix} \hat{B}.
\]

Note that matrix \(Z\) is positive definite and hence, that the inverse is positive definite. Equation (6) can be solved as

\[
(YZ^{-1}Y')\hat{p} = -\hat{r}_p,
\]

(7)

where \(\hat{r}_p = -Y'\hat{p}\). Since \(Z > 0\) and \(Y\) is full row rank, \(YZ^{-1}Y'\) is positive definite. The matrix \(Z\) is block diagonal with block size independent of \(N, K, T\). Therefore computation of \(Z^{-1}\) can be carried out efficiently. However, \(YZ^{-1}Y'\) is a block tri-diagonal matrix and solving (7) is computationally more expensive.

The linear system (7) can be solved by block Jacobi iterations. First, note that \(YZ^{-1}Y'\) has the block tri-diagonal structure

\[
\Delta = YZ^{-1}Y' = \begin{bmatrix} \hat{\Phi}_1 & \hat{\Omega}_1 \\ \hat{\Omega}_1^T & \hat{\Phi}_2 \end{bmatrix} \begin{bmatrix} \hat{\Phi}_1 & \hat{\Omega}_2 \\ \hat{\Omega}_2^T & \hat{\Phi}_N \end{bmatrix},
\]

(8)

where

\[
\hat{\Phi}_j = \hat{A}_j \hat{Q}_j^{-1} \hat{A}_j' + \hat{B}_j \hat{R}_j^{-1} \hat{B}_j' + \hat{E}_j \hat{Q}_j^{-1} \hat{E}_j', \quad j \in \mathbb{N}, \quad \hat{\Omega}_j = \hat{E}_j \hat{Q}_j^{-1} \hat{A}_j', \quad j \in \mathbb{N}.
\]

(9)

for \(j \in \mathbb{N}\) and \(\hat{Q}_{N+1} = 0\). Let

\[
\Psi = \text{diag}(\hat{\Phi}_1, \ldots, \hat{\Phi}_N) \in \mathbb{R}^{\hat{n}_j \times \hat{n}_j}\]
\( \Xi := \Psi - \Delta \). The Jacobi iterations for solving (7) are given by the following:

\[
\Psi^{k+1} = \bar{r}_p + \Xi \tilde{p}^k, \tag{10}
\]

where \( k \) is the iteration index (Hackbusch, 2016). Since \( \Delta > 0 \), the block diagonal matrix \( \Psi > 0 \) (Golub and Van Loan, 2013, Cor 4.2.2). The convergence of these iterations to the solution of (7) is considered in Section 4.

3.1 Structure of Matrices

At each Jacobi iteration, the linear system of equations in (10) needs to be solved. The computational complexity of each Jacobi iteration depends on the structure of the matrix \( \Psi \). It can be seen from (9) that \( \Psi > 0 \) is block diagonal. Each diagonal block \( \Phi_j \) is also very sparse and structured. For \( j \in \mathcal{N} \setminus \{1\} \), each \( A_j \) is block diagonal, whereby both \( \Phi_j \) and \( \Omega_j \) are block diagonal. However, \( \Phi_1 \) is block tri-diagonal because \( A_1 \) is block bi-diagonal. In summary,

\[
\hat{\Phi}_j = \begin{bmatrix}
\hat{\Phi}_{1,j} & \cdots & \hat{\Phi}_{n,j}
\end{bmatrix} \in \mathbb{R}^{n_j \times n_j} 	ext{ for } j \in \mathcal{N} \setminus \{1\},
\tag{11a}
\]

\[
\hat{\Phi}_1 = \begin{bmatrix}
\hat{\Phi}_{1,1} & \phi_{1,1} & \phi_{1,2} & \cdots & \phi_{1,n_j} \\
n_{1,1} & \cdots & \cdots & \cdots & \phi_{1,n_j} \\
\phi_{1,1} & \phi_{1,2} & \cdots & \cdots & \phi_{1,n_j} \\
\phi_{1,2} & \phi_{1,3} & \cdots & \cdots & \phi_{1,n_j} \\
\phi_{1,n_j-1} & \phi_{1,n_j} & \cdots & \cdots & \phi_{1,n_j}
\end{bmatrix} \in \mathbb{R}^{n_1 \times n_1},
\tag{11b}
\]

where

\[
\hat{\Phi}_{i,j} = \bar{A}_{i,j} \bar{Q}_{i,j}^{-1} A'_{i,j} + \bar{B}_{i,j} \bar{R}^{-1}_{i,j} \bar{B}'_{i,j} + \bar{E}_{i,j} \bar{Q}_{i,j}^{-1} \bar{E}'_{i,j} \\
\in \mathbb{R}^{n_{i,j} \times n_{i,j}} \text{ for } i \in \mathcal{K}, j \in \mathcal{N} \setminus \{1\},
\]

\[
\hat{\Phi}_{1,1} = \bar{A}_{1,1} \bar{Q}_{1,1}^{-1} A'_{1,1} + \bar{B}_{1,1} \bar{R}^{-1}_{1,1} \bar{B}'_{1,1} + \bar{E}_{1,1} \bar{Q}_{1,1}^{-1} \bar{E}'_{1,1} \\
+ F_{1,1} \bar{Q}_{1,1}^{-1} \bar{F}'_{1,1} + \mathbb{I} \in \mathbb{R}^{n_{1,1} \times n_{1,1}} \text{ for } i \in \mathcal{K}, j \in \mathcal{N} \setminus \{1\},
\]

\[
\phi_{1,1} = F_{1,1} \bar{Q}_{1,1}^{-1} \bar{F}'_{1,1} + \mathbb{I} \in \mathbb{R}^{n_{1,1} \times n_{1,1}} \text{ for } i \in \mathcal{K}, j \in \mathcal{N} \setminus \{1\},
\]

and \( \bar{Q}_{1,1} = \bar{Q}_{1,1} + \bar{Q}_{1,1}^{-1} \bar{F}'_{1,1} \). A remark on the structure of \( \hat{\Phi}_j \) is that it scales linearly with \( T \). Moreover, \( \hat{\Phi}_{i,j} \) is block tri-diagonal with block-size that is independent of \( K \), \( N \) and \( T \). As such, in solving (10), for \( i \in \mathcal{K}, j \in \mathcal{N} \setminus \{1\} \), the computation of each block component \( \tilde{p}_j^k \) of \( \tilde{p}_j^{k+1} \) decomposes into a small block tri-diagonal equation \( \hat{\Phi}_{i,j} \tilde{p}_j^k = \bar{r}_j^k \) to solve. This can be achieved by backward and forward recursions (Meurant, 1992), with aggregate complexity that scales linearly in \( T \), the number of blocks, and cubically with respect to \( n_{i,j} \), the size of each block. On the other hand, for \( j = 1 \), a block tri-diagonal system of equations with block-size that depends linearly on \( T \) needs to be solved. With a view to retaining linear in \( T \) complexity, it is proposed to (approximately) solve this system of equations by inner Jacobi iterations.

Dropping the explicit dependence on the outer iteration index \( k \), the first block equation of (10) is given by
Remark 5. For $j \in \mathcal{N}\setminus\{1\}$, $K$ block tri-diagonal systems of equations are each solved only once per outer iteration $k$. However, for $j = 1$ these $K$ block tri-diagonal systems of equations are solved up to $S_{\text{max}}$ times. Therefore, the total computational time required per outer iteration is upper bounded by that of agent 1. With a fixed bound on the number of inner iterations, the per iteration computational complexity of every outer block Jacobi iteration scales linearly with $N$, $K$ and $T$.

4. NESTED JACOBI ITERATIONS

Let $\Gamma := \Psi^{-1}\Xi$ denote the iteration matrix of the outer Jacobi iterations (10). Let $\bar{p}^*$ denote the exact solution of (7). This is a fixed-point of the iterations, for which $\bar{p}^k \to \bar{p}^*$ as $k \to \infty$ from any initial value if and only if

$$\text{rad}(\Gamma) < 1,$$

(15)

where $\text{rad}(\cdot)$ denotes spectral radius. See (Hackbusch, 2016, Thm 2.16), for example. Since the positive definite matrix $\Delta = YZ^{-1}Y' = \Psi - \Xi$ is block tri-diagonal, it follows that (15) holds, as noted in (Zafar et al., 2019, Theorem 6). Similarly, the inner iterates $\hat{p}_1$ converge, as $s \to \infty$, to the solution $\hat{p}_1^{*}$ of (12) if and only if

$$\text{rad}(\hat{\Upsilon}^{-1}_1\hat{\Lambda}_1) < 1.$$

Again, this holds because $\Phi_1 = \hat{\Upsilon}_1 - \hat{\Lambda}_1 > 0$ is also block tri-diagonal.

When the number of inner Jacobi iterations is limited to $S_{\text{max}} > 0$, the update at outer iteration $k$ corresponds to

$$\bar{p}^{k+1} = \Psi^{-1}(r_{\bar{p}} + \Xi \bar{p}^k) + w^k$$

(16)

where

$$w^k = \text{diag}(\Theta_1, 0, \ldots, 0)(r_{\bar{p}} + \Xi \bar{p}^k)$$

and $\Theta_1 = (\sum_{i=0}^{S_{\text{max}}} (\hat{\Upsilon}_1^{-1}\hat{\Lambda}_1)^i\hat{\Upsilon}_1 - \hat{\Lambda}_1)^{-1}$. In particular,

$$\bar{p}^{k+1} = (\Psi^{-1} + \text{diag}(\Theta_1, 0, \ldots, 0))(r_{\bar{p}} + \Xi \bar{p}^k).$$

This converges to a fixed point if and only if $\text{rad}(\Gamma + \text{diag}(\Theta_1, 0, \ldots, 0)\Xi) < 1$, which holds provided $\Theta_1$ is sufficiently small because eigenvalues vary continuously (Horn and Johnson, 2013, Thm D1) and $\text{rad}(\Gamma) < 1$. It is possible to make $\Theta_1$ arbitrarily small since $\text{rad}(\hat{\Upsilon}_1^{-1}\hat{\Lambda}_1) < 1$, whereby $\sum_{i=0}^{S_{\text{max}}} (\hat{\Upsilon}_1^{-1}\hat{\Lambda}_1)^i\hat{\Upsilon}_1 - \hat{\Lambda}_1 \to \hat{\Phi}_1^{-1}$ as $S_{\text{max}} \to \infty$. The proximity of the perturbed fixed-point to the solution $\bar{p}^*$ of (10) can also be made arbitrarily close by increasing $S_{\text{max}}$.

In (16), if $\|w^k\|_{\infty} \leq \epsilon_w$ for all outer iterations $k$, then the $\ell_{\infty}$-to-$\ell_{\infty}$ gain from the disturbance $w^k$ to the error defined relative to the unperturbed trajectory of the iterates $\bar{p}^k$ is relevant to nested algorithm performance. This gain is finite since the corresponding dynamics is bounded-input bounded-output (BIBO) stable because $\text{rad}(\Psi^{-1}\Xi) < 1$ (Bacciotti, 2019). As such, smaller $\epsilon_w$ implies a smaller perturbation from the exact solution.

In the next section, numerical experiments are performed to illustrate the effect of early termination of inner iterations on the convergence of outer iterations to a preset tolerance.

5. NUMERICAL RESULTS

In this section, numerical experiments are performed by applying the Algorithm 1 to solve an optimal control problem involving model data for an automated irrigation network under so-called distant-downstream control (Cantoni et al., 2007). The irrigation network consists of $K$ number of controlled pools in the primary channel and $N$ number of controller pools in each secondary channel originating from the pools of primary channel. After temporal discretization, the dynamics of each controlled pool can be described by a sub-system in the form of (1). For each sub-system, the control input $u_{i,j}(t)$ is the adjustment of the water-level reference, while the state variable $x_{i,j}(t)$ is the deviation of state trajectory from equilibrium. For each sub-system $i \in \mathcal{K}$, $j \in \mathcal{N}$, which takes the form (1) with $n_{i,j} = 4$, $m_{i,j} = 1$, the corresponding cost is set such that $Q_{i,j} = I$ and $R_{i,j} = 1$.

The implementation involves a single thread on one processor. That is, the possibility of exploiting parallelism in Algorithm 1 and 2 is not considered. The experiments are performed by setting $K = N$ and $T = K + N$ and varying the value of $N$ from 10 to 100, whereby the largest problem considered has 10.04 million primal variables and 8.04 million dual variables.

Fig. 2 shows the processor time for single thread implementations of Algorithm 1 with stopping condition set to $\epsilon = 10^{-3}$ and Algorithm 2 with stopping condition set to $\epsilon_{in} = 10^{-9}$ and varying $S_{\text{max}}$. It shows that the increase in computational time has a quartic slope. Due to different number of inner iterations, the total computational time is different with $S_{\text{max}} = 1$ being fastest for each of the above implementations.

Fig. 3 shows the total number of iterations performed by Algorithm 1 until convergence. The total number of iterations grow modestly with system size due to increase in $\text{rad}(\Gamma)$. Also, the total number of outer iterations are same for different $S_{\text{max}}$. This is by virtue of the stopping criteria $\|\bar{p}^{k+1} - \bar{p}^k\|_{\infty} < \epsilon$, $\forall j \in \mathcal{N}$. It shows that each component $\bar{p}_1^k$ of $\bar{p}$ converges to a stationary point $\bar{p}_1^*$. Accuracy of the solution is shown in Fig. 4.

Fig. 4 shows the error $\|\hat{x} - \bar{x} + \hat{u} - \bar{u} + \hat{\xi}\|_{\infty}$ for the outcome $\hat{x}$ and $\bar{u}$ of nested Jacobi iterations. For $S_{\text{max}} = 1$, 3, this error is large but its modest for $S_{\text{max}} = 5$, 7. For $S_{\text{max}} = 9$, the error is of order $10^{-5}$. However, for fixed number of inner iterations, the error increases with increase in system complexity.
Remark 5. The computational overhead scales linearly with the number of exchanges. Therefore, the overall per-iteration communication complexity of every outer block Jacobi iteration scales linearly with the number of inner iterations, the per-iteration computational time has a quartic slope. Due to different computations at each iteration can be distributed across $N \times K$ parallel processing agents in a network mirroring the tree structure of the graph. A preliminary convergence analysis of the proposed algorithm is provided. It is observed numerically that there is a modest growth in spectral radius of the iteration matrix $\Gamma$ which results in increase of the number of iterations required to converge in the size $N = K$, $T = N + K$ of problems. Future work includes a detailed theoretical analysis for convergence of nested Jacobi iterations subject to early termination of inner iterations. It also includes the consideration of inequality constraints on the state and control input variables and extending the results for a more general class of tree graphs.

7. ACKNOWLEDGMENTS

This work was supported by the Australian Research Council (LP160100666). Aspects of it were done while F. Farokhi held a joint appointment with CSIRO’s Data61, Docklands, VIC 3008, Australia.

REFERENCES


Minerva Access is the Institutional Repository of The University of Melbourne

Author/s:
Zafar, A; Farokhi, F; Cantoni, M

Title:
Linear quadratic control computation for systems with a directed tree structure

Date:
2020-01-01

Citation:

Persistent Link:
http://hdl.handle.net/11343/272712