2-26-2017

Distributed Model Predictive Control via Proximal

Xiaodong Hou
Purdue University, hou39@purdue.edu

Jianghai Hu
Purdue University, jianghai@purdue.edu

Jie Cai
Purdue University, cai40@purdue.edu

James E. Braun
Purdue University - Main Campus, jbraun@purdue.edu

Distributed Model Predictive Control via Proximal Jacobian ADMM for Building Control Applications

Xiaodong Hou, Yingying Xiao, Jie Cai, Jianghai Hu and James E. Braun

Abstract

This paper investigates a distributed model predictive control (DMPC) framework for building control applications. The proposed framework is general in that it can be easily customized to solve the dynamic optimization problem for a broad class of multi-zone buildings with relatively complex HVAC systems. The Proximal Jacobian alternating direction method of multipliers (ADMM), a recent variant of the traditional Gauss-Seidel sequential ADMM is employed and adopted to solve the centralized optimization problem, which ultimately leads to an agent-based parallel updating scheme with guaranteed convergence. A case study on the HVAC energy optimization of a multi-zone building is presented to show the effectiveness of the proposed method.

I. Introduction

The building sector consumes over 40% of primary energy in the United States. Advanced building control strategies, especially the model predictive control (MPC) approach, have shown promise to reduce building's heating, ventilation and air-conditioning (HVAC) energy consumption through the dynamic optimization of the interaction of building thermal dynamics with HVAC systems.

However, it has been commonly recognized in the building control community that the engineering cost for building-specific optimal control designs and the implementation cost for deploying advanced controls such as MPC are extremely high. In addition, centralized MPC for large scale buildings usually suffers from poor scalability and high computational burden as the corresponding HVAC configuration and inter-zonal thermal couplings are very complicated.

Xiaodong Hou, Yingying Xiao and Jianghai Hu are with the School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN 47907, USA {hou39, xiao106, jianghai}@purdue.edu.

Jie Cai and James E. Braun are with the School of Mechenical Engineering, Purdue University, West Lafayette, IN 47907, USA {cai40, jbraun}@purdue.edu. This work was supported by the National Science Foundation under Grant No. 1329875.

Consequently, there has been a growing interest of applying distributed MPC (DMPC) strategies in building control, since they enable the decomposition of the centralized optimization problem into multiple small local problems (structurally or computationally), making it possible for scalable and low computational implementations.

Distributed model predictive control has been studied extensively by the control community. Many of the existing DMPC approaches are built on agent-based negotiation or distributed optimization. For example, an iterative solution approach is introduced in [1] for constrained linear systems coupled through the inputs, and agents negotiate with each other by "making proposals" and evaluating proposals. Dual decomposition and Gauss-Seidel alternating direction method of multipliers (ADMM) are applied to DMPC in [2] and [3]. In addition, the proximal center decomposition method is proposed to solve the DMPC problem [4] for linear systems with coupling in the dynamics; the method in [5] employs a parametric nonconvex decomposition algorithm for DMPC of nonlinear systems with constraint couplings; a Jacobi (parallel) algorithm for cooperative DMPC is provided in [6], considering either dynamic coupling or cost coupling. Interested readers can refer to [7] and [8] for more recent work on DMPC. It should be noted that the aforementioned literature either only considers couplings in the dynamics or only couplings in the cost function, but not both.

On the other hand, there has been much effort of applying DMPC to the building control (especially temperature regulation) problems in recent years. For instance, a DMPC controller based on primal decomposition is given in [9] for a multi-zone building where local control profiles are optimized by the local agents; however, it does not consider the case where there is coupling in the objective function. A simlilar study in [10] utilizes Benders' decomposition to split the centralized problem. A dual decomposition based DMPC is used to solve the temperature regulation problem of a multi-zone building served by a central air handling unit in [11]. However, both of these two studies consider greatly simplified HVAC models.

From the above review, it is evident that there is no general DMPC framework that not only takes dynamic couplings into account, but also handles couplings in the constraints and objective functions. Such a framework is sorely needed because it can be customized to a broad class of realistic building control applications in a straightforward fashion as many building problems involve multiple thermal zones (couplings in thermal dynamics) served by a complex HVAC system with multiple consumers or components (couplings in objective functions and resource constraints).

This paper aims at providing such a DMPC framework for building control applications. To this end, we start from a fairly general control formulation (system dynamics, cost function, constraints). Through a series of transformations, the centralized optimization problem is cast into a standard form. Then, a decomposition technique recently proposed in the literature, Proximal Jacobian ADMM, is utilized to decompose the centralized problem into several smaller scale local subproblems, each of which is solved iteratively by local agents in parallel. The convergence of local solutions to a central optimal is also established under certain conditions.

While the proposed method may not always have superior performance compared to algorithms or implementations designed specifically for a particular case study, our framework is suited for a series of building control formulations and requries relatively less customization. This attribute of the proposed method makes it amenable for a plug-and-play implementation, which could significantly reduce engingeering costs. It is also worth mentioning that there have been some attempts in this direction, which combine agent-based control with DMPC for building energy minimization and demand reduction [12] [13].

The rest of the paper is organized as follows. In Section II, a description of the system structure under study is given. In Section III, the problem formulation and transformations are presented. The Proximal Jacobian ADMM method for solving DMPC problems is introduced in Section IV. A multi-zone building control case study is presented in Section V to show the effectiveness of the proposed method. Finally, some concluding remarks are given in Section VI.

II. System Description

A. Notation

In this paper, the following notations will be used. $\mathbb{N}$: set of non-negative integers. $\mathbb{R}^n$: n-dimensional Euclidean space. $\mathbb{R}^{n\times m}$: set of all $n \times m$ real matrices. $A^\top$: transpose of matrix A. $A \succ 0$: matrix A is symmetric positive definite. $\|\cdot\|$: Euclidean norm of a vector or spectral norm of a matrix. $\|x\|_A = \sqrt{x^\top A x}$: A-norm of vector $x \in \mathbb{R}^n$ for $A \succ 0$. $|\mathcal{P}|$: cardinality of a set $\mathcal{P}$. $\mathrm{diag}\big(A, B, C\big)$: block diagonal matrix with diagonal blocks as matrices A, B and C.

B. System Dynamics

Suppose the building under study consists of L thermal zones. The thermal dynamics of the ith zone are characterized by the following discrete-time affine system,

$$x_i(k+1) = A_{ii}x(k) + \sum_{j\in\mathcal{N}_i} A_{ij}x_j(k) + B_{ii}u_i(k) + \sum_{j\in\mathcal{M}_i} B_{ij}u_j(k) + F_i w_i(k), \quad i = 1,\dots,L, \tag{1}$$

where $k \in \mathbb{N}$, $x_i(k) \in \mathbb{R}^{n_i}$ are the local state variables (zone air temperature, floor, wall temperatures, etc.), $u_i(k) \in \mathbb{R}^{m_i}$ is the local control input (such as supply air temperature, supply air flow rate, or sensible cooling), $w_i(k) \in \mathbb{R}^{p_i}$ is the uncontrollable input (ambient temperature, solar radiation, occupants); $A_{ij} \in \mathbb{R}^{n_i\times n_j}$ and $B_{ij} \in \mathbb{R}^{n_i\times m_i}$ are constant system matrices. We assume that there is one agent dedicated for the local optimization and decision making of each zone, i.e, this agent will determine the control input as well as the corresponding zone temperature profile of the given zone. $\mathcal{N}_i$ and $\mathcal{M}_i$ represent the set of agents that have influence on the state x_i through their local states and controls, respectively.

The overall building thermal dynamics are given by

$$x(k+1) = Ax(k) + Bu(k) + Fw(k),$$

where $x(k)$, $u(k)$ and $w(k)$ are the concatenations of the corresponding local variables. Note that an agent may be coupled through states or control inputs with other agents. A fairly broad class of buildings can be represented by this model as it encompasses the possible inter-zonal thermal couplings (open space, window, corridor) and couplings across zones through control inputs.

C. Constraints

Two common types of constraints in building problems are considered in this paper: local and shared constraints. Firstly, each thermal zone and its HVAC equipment are subject to local state and control input constraints of the form

$$x_i(k) \in \mathcal{X}_i(k), \quad u_i(k) \in \mathcal{U}_i(k), \quad i = 1,\dots,L, \tag{2}$$

where $\mathcal{X}_i(k)$ and $\mathcal{U}_i(k)$ are assumed to be time-varying compact convex sets. These local constraints usually arise due to thermal comfort requirements and HVAC operation constraints.

For example, zone air temperatures should be maintained inside comfort intervals; sensible cooling injected into the room space from a rooftop unit (RTU) cannot exceed the RTU's capacity.

Secondly, shared constraints often involve different HVAC equipment and their corresponding agents, representing some "shared resources" in the whole system. For instance, several variable air volume (VAV) units may be served by a central air handling unit (AHU), hence the total air flow rate through individual VAVs is smaller than that provided by the AHU. In this study, we consider shared constraints of the following form

$$\sum_{i=1}^{L} h_i(k) q_i\big(x_i(k)\big) \le h_0(k), \quad \sum_{i=1}^{L} g_i(k) r_i\big(u_i(k)\big) \le g_0(k), \tag{3}$$

where $h_i(k)$, $g_i(k) \in \mathbb{R}$, $i = 1, \ldots, L$, are time-dependent constants; $q_i(\cdot)$ and $r_i(\cdot)$ are convex functions of x_i and u_i, respectively. Notice that q_i and r_i will be set to zero if the corresponding zone or equipment is not involved in the shared constraint. The current formulation can be easily generalized to the case with multiple shared constraints.

D. Objective Function

In many building control problems, the objective function to be minimized approximates the total HVAC energy bill in a certain prediction horizon. Extra terms that represent demand charges may also be included in the objective function. It is not uncommon to have couplings in the objective function across different HVAC equipment. A simple example is given by different HVAC components in a single building or a cluster of buildings all sharing the same cooling source, e.g., a central chiller plant.

To accommodate the situation where there are multiple cooling/heating sources, we consider the following infinite horizon global objective function,

$$J_\infty = \sum_{k=0}^{\infty} \sum_{i=1}^{L} f_i(x_i, u_i, x_{j \in \mathcal{P}_i}, u_{j \in \mathcal{Q}_i}). \tag{4}$$

Here, each f_i represents the energy consumption of certain HVAC equipments that share one cooling/heating source, and is assumed to be a closed, convex function in each argument; $\mathcal{P}_i$ and $\mathcal{Q}_i$ (not necessarily the same as $\mathcal{N}_i$ and $\mathcal{M}_i$) represent the set of agents that have influence on the local cost function of agent i through x_j and u_j, respectively.

Remark 1. *The current formulation is very general as it takes into account couplings in states, control inputs, and cost functions among different agents. Many systems in the literature can be treated as special cases of this formulation.*

III. Problem Formulation

A. Centralized Optimization Problem

The MPC scheme replaces the infinite horizon objective function in (4) with a finite horizon one. The prediction horizon is assumed to be N, and the corresponding optimization problem is solved at each time step in a receding horizon fashion. For notational simplicity, the current time index is assumed to be 0 and we will use $x(k)$ and $u(k)$ to represent the predicted state and control input that are k steps after the current step.

The centralized optimization problem at each time instance can then be formulated as

$$\begin{aligned}
\underset{x(k),\,u(k)}{\text{minimize}} \quad & J = \sum_{k=0}^{N-1}\sum_{i=1}^{L} f_i(x_i, u_i, x_{j\in\mathcal{P}_i}, u_{j\in\mathcal{Q}_i}) \\
\text{subject to} \quad & x(k+1) = Ax(k) + Bu(k) + Fw(k), \\
& \sum_{i=1}^{L} h_i(k+1) q_i\big(x_i(k+1)\big) \le h_0(k+1), \\
& \sum_{i=1}^{L} g_i(k) r_i\big(u_i(k)\big) \le g_0(k), \\
& x_i(k+1) \in \mathcal{X}_i(k+1), \quad u_i(k) \in \mathcal{U}_i(k), \\
& i = 1, \ldots, L, \quad k = 0, 1, \ldots, N-1.
\end{aligned}$$

The above optimization problem can be more compactly represented as

$$\begin{aligned}
\underset{\mathbf{x},\,\mathbf{u}}{\text{minimize}} \quad & J = \sum_{i=1}^{L} F_i(\mathbf{x}_i, \mathbf{u}_i, \mathbf{x}_{j\in\mathcal{P}_i}, \mathbf{u}_{j\in\mathcal{Q}_i}) \\
\text{subject to} \quad & \mathbf{x} = \boldsymbol{\Omega} x(0) + \boldsymbol{\Phi}\mathbf{u} + \boldsymbol{\Psi}\mathbf{w}, \\
& \mathbf{H}\widehat{\mathbf{q}}(\mathbf{x}) \le \mathbf{h}_0, \quad \mathbf{G}\widehat{\mathbf{r}}(\mathbf{u}) \le \mathbf{g}_0, \\
& \mathbf{x}_i \in \mathcal{X}_{Ni}, \quad \mathbf{u}_i \in \mathcal{U}_{Ni}, \quad i = 1, \ldots, L,
\end{aligned} \tag{5}$$

where $\mathcal{X}_{Ni} = \prod_{k=1}^{N} \mathcal{X}_i(k)$, $\mathcal{U}_{Ni} = \prod_{k=0}^{N-1} \mathcal{U}_i(k)$, $F_i(\cdot) = \sum_{k=0}^{N-1} f_i(\cdot)$; $\boldsymbol{\Omega}$, $\boldsymbol{\Phi}$, $\boldsymbol{\Psi}$ are block matrices that represent the overall dynamics in the whole prediction horizon, and

$$\mathbf{x}_i = [x_i^\top(1) \cdots x_i^\top(N)]^\top, \quad \mathbf{u}_i = [u_i^\top(0) \cdots u_i^\top(N-1)]^\top,$$

$$\mathbf{x} = [x^\top(1) \cdots x^\top(N)], \quad \mathbf{u} = [u^\top(0) \cdots u^\top(N-1)], \quad \mathbf{w} = [w^\top(0) \cdots w^\top(N-1)]^\top,$$

$$\mathbf{H} = \text{diag}(h^\top(1) \cdots h^\top(N)), \quad \mathbf{G} = \text{diag}(g^\top(0) \cdots g^\top(N-1)),$$

$$h(k) = [h_1(k) \cdots h_L(k)]^\top, \quad g(k) = [g_1(k) \cdots g_L(k)]^\top,$$
$$\widehat{\mathbf{q}}(\mathbf{x}) = \big[\hat{q}^\top\big(x(1)\big) \cdots \hat{q}^\top\big(x(N)\big)\big]^\top, \quad \widehat{\mathbf{r}}(\mathbf{u}) = \big[\hat{r}^\top\big(u(0)\big) \cdots \hat{r}^\top\big(u(N-1)\big)\big]^\top,$$
$$\hat{q}\big(x(k)\big) = [q_1(x_1(k)) \cdots q_L(x_L(k))]^\top, \quad \hat{r}\big(u(k)\big) = [r_1(x_1(k)) \cdots r_L(x_L(k))]^\top.$$

If we stack $\mathbf{x}$ and $\mathbf{u}$ into a single optimization variable $\mathbf{y} = [\mathbf{u}^\top, \mathbf{x}^\top]^\top$, then optimization problem (5) is cast as

$$\begin{aligned} \underset{\mathbf{y}}{\text{minimize}} \quad & J = \sum_{i=1}^{L} F_i(\mathbf{y}_i, \mathbf{y}_{j\in\mathcal{P}_i\cup\mathcal{Q}_i}) \\ \text{subject to} \quad & \widehat{\mathbf{A}}_{\text{eq}}\mathbf{y} = \widehat{\mathbf{b}}_{\text{eq}}, \quad \widehat{\mathbf{A}}_{\text{in}}\widehat{\mathbf{p}}(\mathbf{y}) \le \widehat{\mathbf{b}}_{\text{in}}, \\ & \mathbf{y}_i \in \mathcal{Y}_{Ni}, \quad i = 1, \ldots, L, \end{aligned} \tag{6}$$

where $\mathbf{y}_i = [\mathbf{u}_i^\top, \mathbf{x}_i^\top]^\top$, $\mathcal{Y}_{Ni} = \mathcal{U}_{Ni} \times \mathcal{X}_{Ni}$, and

$$\widehat{\mathbf{A}}_{\text{eq}} = \begin{bmatrix} -\mathbf{\Phi} & \mathbf{I} \end{bmatrix}, \quad \widehat{\mathbf{b}}_{\text{eq}} = \mathbf{\Omega}x(0) + \mathbf{\Psi}\mathbf{w},$$
$$\widehat{\mathbf{A}}_{\text{in}} = \begin{bmatrix} \mathbf{H} & \mathbf{0} \\ \mathbf{0} & \mathbf{G} \end{bmatrix}, \quad \widehat{\mathbf{b}}_{\text{in}} = \begin{bmatrix} \mathbf{h}_0 \\ \mathbf{g}_0 \end{bmatrix}, \quad \widehat{\mathbf{p}}(\mathbf{y}) = \begin{bmatrix} \widehat{\mathbf{q}}(\mathbf{x}) \\ \widehat{\mathbf{r}}(\mathbf{u}) \end{bmatrix},$$

where $\mathbf{I}$ and $\mathbf{0}$ are identity matrix and zero matrix of proper dimensions, respectively. We call $\mathbf{y}_i$ the local private variable of agent i.

B. Consensus Constraint

Because of the couplings in constraints and the objective function, the centralized optimization problem is not readily separable for a distributed solution.

The objective function can be decoupled by introducing local copies of coupling variables for each agent and enforcing all the local copies to have the same value through extra consensus constraints. For example, agent i is coupled with $\mathbf{y}_{j\in\mathcal{P}_i\cup\mathcal{Q}_i}$ in the local cost function F_i. Then, a local copy of $\mathbf{y}_{j\in\mathcal{P}_i\cup\mathcal{Q}_i}$ for agent i can be introduced as $\widehat{\mathbf{y}}_i$. Notice that $\widehat{\mathbf{y}}_i$ is used for the sake of notational simplicity; whereas only copies of $\mathbf{x}_{j\in\mathcal{P}_i}$ and $\mathbf{u}_{j\in\mathcal{Q}_i}$ need to be included in $\widehat{\mathbf{y}}_i$. With these extra variables, the optimization problem becomes

$$\begin{aligned} \underset{\mathbf{y}}{\text{minimize}} \quad & J = \sum_{i=1}^{L} F_i(\mathbf{y}_i, \widehat{\mathbf{y}}_i) \\ \text{subject to} \quad & \widehat{\mathbf{A}}_{\text{eq}}\mathbf{y} = \widehat{\mathbf{b}}_{\text{eq}}, \quad \widehat{\mathbf{A}}_{\text{in}}\widehat{\mathbf{p}}(\mathbf{y}) \le \widehat{\mathbf{b}}_{\text{in}}, \\ & \mathbf{E}_i\mathbf{y} = \widehat{\mathbf{y}}_i, \quad \mathbf{y}_i \in \mathcal{Y}_{Ni}, \quad i = 1, \ldots, L, \end{aligned} \tag{7}$$

where $\mathbf{E}_i$ is a matrix with elements of value 0 or 1. Each row of $\mathbf{E}_i$ has only one non-zero element, which enforces the local copy at agent i to be the same as the original variable at its neighbouring agent $j \in \mathcal{P}_i \cup \mathcal{Q}_i$. Hence, all local copies of the same private variable will have a consistent value.

IV. Distributed Solution via Proximal Jacobian ADMM

By combining the two equality constraints and introducing a slack variable $\mathbf{z}_0$, the optimization problem (7) can be equivalently transformed into

$$\begin{aligned} \underset{\mathbf{z}}{\text{minimize}} \quad & J = \sum_{i=0}^{L} F_i(\mathbf{z}_i) \\ \text{subject to} \quad & \mathbf{A}\mathbf{z} = \mathbf{b}, \quad \mathbf{C}\mathbf{p}(\mathbf{z}) = \mathbf{d}, \\ & \mathbf{z}_i \in \mathcal{Z}_{Ni}, \quad i = 0, \dots, L, \end{aligned} \tag{8}$$

where

$$\mathbf{A} = \begin{bmatrix} \widehat{\mathbf{A}}_{\text{eq}} & \mathbf{0} & \mathbf{0} \\ \mathbf{E} & -\mathbf{I}_{\text{b}} & \mathbf{0} \end{bmatrix}, \quad \mathbf{C} = \begin{bmatrix} \widehat{\mathbf{A}}_{\text{in}} & \mathbf{0} & \mathbf{I} \end{bmatrix},$$

$$\mathbf{p}(\mathbf{z}) = \begin{bmatrix} \widehat{\mathbf{p}}^\top(\mathbf{y}) & \widehat{\mathbf{y}}^\top & \mathbf{z}_0^\top \end{bmatrix}^\top, \quad \mathbf{z} = \begin{bmatrix} \mathbf{y}^\top & \widehat{\mathbf{y}}^\top & \mathbf{z}_0^\top \end{bmatrix}^\top,$$

$$\mathbf{b} = \begin{bmatrix} \widehat{\mathbf{b}}_{\text{eq}}^\top & 0^\top \end{bmatrix}^\top, \quad \mathbf{d} = \widehat{\mathbf{b}}_{\text{in}}, \quad \mathbf{E} = \begin{bmatrix} \mathbf{E}_1^\top & \cdots & \mathbf{E}_L^\top \end{bmatrix}^\top,$$

and $\widehat{\mathbf{y}} = [\widehat{\mathbf{y}}_1^\top, \dots, \widehat{\mathbf{y}}_L^\top]^\top$; $\mathbf{z}_i = [\mathbf{y}_i^\top, \widehat{\mathbf{y}}_i^\top]^\top \in \mathcal{Z}_{Ni} = \{(\mathbf{y}_i, \widehat{\mathbf{y}}_i) | \mathbf{y}_i \in \mathcal{Y}_{Ni}\}$ for $i = 1, \dots, L$; $\mathcal{Z}_{N_0}$ is the positive orthant, and $F_0(\mathbf{z}_0) = 0$. In addition, $\mathbf{I}_{\text{b}}$ is a block identity matrix of proper dimension.

In the optimization problem (8), with the objective function being the sum of local cost functions of individual agents, the only coupling that hinders a direct decomposition into subproblems are the two equality constraints. To overcome this, we introduce the augmented Lagrangian function and dual variables as follows.

A. Augmented Lagrangian and Dual Problem

First notice that, the elements in $\mathbf{z}$ can be re-ordered as $\mathbf{z} = [\mathbf{z}_0^\top, \dots, \mathbf{z}_L^\top]^\top$, where $\mathbf{z}_i = [\mathbf{y}_i^\top, \widehat{\mathbf{y}}_i^\top]^\top$, $i = 1, \dots, L$, is the decision variable of agent i (an additional agent will determine $\mathbf{z}_0$). For the optimization problem (8), an augmented Lagrangian function is formulated as,

$$\mathcal{L}_\rho(\mathbf{z}, \lambda, \mu) = \sum_{i=0}^{L} F_i(\mathbf{z}_i) + \lambda^\top(\mathbf{A}\mathbf{z} - \mathbf{b}) + \mu^\top(\mathbf{C}\mathbf{p}(\mathbf{z}) - \mathbf{d}) + \frac{\rho_1}{2}\|\mathbf{A}\mathbf{z} - \mathbf{b}\|^2 + \frac{\rho_2}{2}\|\mathbf{C}\mathbf{p}(\mathbf{z}) - \mathbf{d}\|^2$$

$$= \sum_{i=0}^{L} F_i(\mathbf{z}_i) + \lambda^\top \Big(\sum_{i=0}^{L} \mathbf{A}_i \mathbf{z}_i - \mathbf{b} \Big) + \frac{\rho_1}{2} \Big\| \sum_{i=0}^{L} \mathbf{A}_i \mathbf{z}_i - \mathbf{b} \Big\|^2$$

$$+ \mu^\top \Big(\sum_{i=0}^{L} \mathbf{C}_i \mathbf{p}_i(\mathbf{z}_i) - \mathbf{d} \Big) + \frac{\rho_2}{2} \Big\| \sum_{i=0}^{L} \mathbf{C}_i \mathbf{p}_i(\mathbf{z}_i) - \mathbf{d} \Big\|^2$$

$$= \sum_{i=0}^{L} \mathcal{L}_i(\mathbf{z}_i,\, \lambda,\, \mu) - \lambda^\top \mathbf{b} - \mu^\top \mathbf{d} + \phi(\mathbf{z})$$

where λ, μ are the Lagrange multipliers, or dual variables; $\rho_1 > 0$ and $\rho_2 > 0$ are the penalty parameters; $\mathcal{L}_i(\mathbf{z}_i,\, \lambda,\, \mu) = F_i(\mathbf{z}_i) + \lambda^\top \mathbf{A}_i \mathbf{z}_i + \mu^\top \mathbf{C}_i \mathbf{p}_i(\mathbf{z}_i)$ and $\phi(\mathbf{z}) = \frac{\rho_1}{2} \| \sum_{i=0}^{L} \mathbf{A}_i \mathbf{z}_i - \mathbf{b} \|^2 + \frac{\rho_2}{2} \| \sum_{i=0}^{L} \mathbf{C}_i \mathbf{p}_i(\mathbf{z}_i) - \mathbf{d} \|^2$. Notice that $\mathbf{A}_i$ are the columns of $\mathbf{A}$ that correspond to the elements of $\mathbf{z}_i$; $\mathbf{p}_i(\mathbf{z}_i)$ is the concatenation of the elements in $\mathbf{p}(\mathbf{z})$ with $\mathbf{z}_i$ as argument; similarly, $\mathbf{C}_i$ is obtained by picking out columns of $\mathbf{C}$ that correspond to the elements of $\mathbf{p}_i(\mathbf{z}_i)$.

The dual function is obtained by minimizing the Lagrangian function with respect to the primal variable $\mathbf{z}$, $d(\lambda,\, \mu) = \inf_{\mathbf{z}} \mathcal{L}_\rho(\mathbf{z},\, \lambda,\, \mu)$.

B. Precursor Algorithms

Many primal-dual based decomposition schemes have been proposed to solve (8). Dual decomposition [2] can not be applied to our problem since the objective function is not strongly convex in $\mathbf{z}_0$. The multi-block Gauss-Seidel ADMM [14] is a direct extension of the standard ADMM [15] from the two-block case to the multi-block case. However, it is shown in [16] that the multi-block Gauss-Seidel ADMM cannot guarantee convergenve. A variable splitting method is presented in [17] [18], which transforms the multi-block setting into an equivalent two block setting; however, it requires a large number of auxiliary decision variables and extra constraints.

C. Proximal Jacobian ADMM

From the perspective of implementation, a parallel update scheme across agents is preferred. That is, all the agents perform local optimization simultaneously without having to wait other's updated information as in the serial scheme. One straightforward method is the Jacobian ADMM, where at every iteration each agent will minimize the augmented Lagrangian function with respect to its local decision variable $\mathbf{z}_i$, assuming other parts of $\mathbf{z}$ fixed. However, even in the simplest two block setting, this scheme does not converge in general [19].

In this paper, we adopt the Proximal Jacobian ADMM method from [20], which builds on the Jacobian ADMM with a proximal term $\frac{\varphi_i}{2}\|\mathbf{z}_i - \mathbf{z}_i^v\|^2$ added to regularize each agent's subproblem for some $\varphi_i > 0$. The multi-block Proximal Jacobian ADMM scheme is given in Algorithm 2. At each iteration, local agents solve the local optimization problem (9) in parallel, followed by an update on the dual variables using dual ascent. The dual variables can be thought as a coordinator that facilitates the satisfaction of coupling constraints. Notice that $i+$ and $i-$ represent the indices before and after index i, respectively.

Algorithm 1 Proximal Jacobian ADMM

1: Initialize $(\mathbf{z}^0,\, \lambda^0,\, \mu^0)$, set $v = 0$;

2: **repeat**

3: Update $\mathbf{z}_i$ (in parallel) according to

$$\mathbf{z}_i^{v+1} = \arg\min_{\mathbf{z}_i \in \mathcal{Z}_{Ni}} \big(\mathcal{L}_i(\mathbf{z}_i,\, \lambda^v,\, \mu^v) + \frac{\varphi_i}{2}\|\mathbf{z}_i - \mathbf{z}_i^v\|^2 + \phi(\mathbf{z}_{i-}^v,\, \mathbf{z}_i,\, \mathbf{z}_{i+}^v)\big); \tag{9}$$

4: Update λ and μ according to

$$\lambda^{v+1} = \lambda^v + \rho_1(\mathbf{A}\mathbf{z}^{v+1} - \mathbf{b});$$
$$\mu^{v+1} = \mu^v + \rho_2\big(\mathbf{C}\mathbf{p}(\mathbf{z}^{v+1}) - \mathbf{d}\big);$$

5: $v \leftarrow v + 1$;

6: **until** some stopping criterion is satisfied.

Theorem 1. *(Global Convergence) Suppose the following conditions hold,*

1) *$q_i(\cdot)$ and $r_i(\cdot)$ in* (3) *are linear functions in their respective argument, e.g.,* $\mathbf{C}\mathbf{p}(\mathbf{z}) = \widehat{\mathbf{C}}\mathbf{z}$ *for some constant* $\widehat{\mathbf{C}}$*;*
2) *the parameters in Algorithm 1 satisfy $\rho_1,\, \rho_2 > 0$;*
3) $\varphi_i > (L-1)\left(\rho_1\|\mathbf{A}_i\|^2 + \rho_2\|\widehat{\mathbf{C}}_i\|^2\right), \quad i = 1, \ldots, L.$

Then the sequence $\{\mathcal{S}^v = (\mathbf{z}_0^v, \ldots, \mathbf{z}_L^v, \lambda^v, \mu^v)\}$ generated by Algorithm 1 converges to a fixed point of the mapping defined by (9)*, for $i = 1, \ldots, L$.*

Proof. See Appendix. □

Parameters φ_i not only make such a parallel update structure possible, but also play a crucial role in the convergence. Intuitively, values of φ_i represents how aggressively individual agents should perform local updates: larger values indicate more conservativeness when updating local decision variables. Parameters ρ_1 and ρ_2 reflect the penalty on the violation of coupling constraints: larger values indicate stronger emphasis on satisfaction of the coupling constraints, but potentially could result in more iterations to reach optimal solutions. It often requires some fine tuning of all parameters to obtain a desirable convergence behaviour.

V. CASE STUDY

The proposed method will be applied to a building control case study, which represents a class of multi-zone buildings with local control equipments (RTU, zone-dedicated VAV and opening controllable diffuser). The energy saving potential of coordination between the zones and the corresponding HVAC equipments will be explored. We will also demonstrate the effectiveness of the proposed method in deploying such an agent-based coordination strategy.

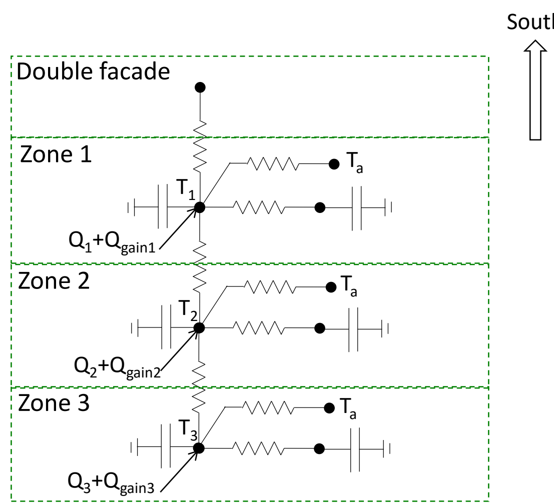

Fig. 1. Purdue Living Lab office room multi-zone thermal network structure

The case study building is the Purdue Living Lab 3 at the Center for High Performance Buildings, West Lafayette, IN, USA. The Purdue Living Lab 3 is a large open space office, whose thermal structure is given in Fig. 1. The room space is divided into three zones according to their relative distance to the south facing double facade. There is a significant load imbalance across the three zones due to different occupancy schedules, solar radiation, and couplings to the ambient through the double facade.

Cooling/heating into the room is provided by a central AHU, which receives chilled water from an air cooled chiller. It is assumed that the supply air at constant temperature goes into the room space through three overhead controllable diffusers, one for each zone. These diffusers allow continuously adjustable air flow rates. This feature provides the energy-saving potential by coordination between zones and diffusers, utilizing couplings between zones (direct air exchange) and building's thermal storage (concrete, furnitures, etc.).

A. Model Description

1) Envelope model: A multi-zone thermal network linear state-space model in the form of (1) was obtained. Model parameters were estimated from building construction information. For each zone i, $\mathcal{N}_i$ represents the adjacent zones, and $\mathcal{M}_i$ is empty. The controllable input $Q_i = u_i^+ - u_i^-$ is the total sensible cooling/heating provided to zone i. $u_i^- \geq 0$ is the cooling energy provided by AHU through diffuser i into zone i. We assume that the supply air temperature is fixed and u_i^- is proportional to the diffuser opening or the air flow rate. $u_i^+ \geq 0$ is the VAV's gas reheat towards zone i. Local state variable x_i includes the average zone air temperature T_i, lumped wall and ceiling temperatures. The exogenous input w_i includes solar radiation, outdoor air temperature, double facade temperature and internal gains $Q_{\text{gain},i}$ (occupancy, computers, etc.). Data of the exogenous inputs are collected during May, 2015.

2) Objective function: The objective function (4) characterizes the total HVAC energy bill in a prediction horizon,

$$\sum_{k=1}^{N} \left(P_e(k) \cdot Pow \left(\sum_{i=1}^{3} u_i^-(k) \right) + P_g \cdot Gas \left(\sum_{i=1}^{3} u_i^+(k) \right) \right)$$

where Pow is the power consumption function, fitted by a fourth order convex polynomial for every ambient temperature, and correlates the power consumption rate to the total sensible cooling. Gas denotes the gas used for the VAV reheat, which is assumed to be a linear function of u_i^+. P_g is the constant gas price, and $P_e(k)$ is the Time of Use (TOU) electricity price, which will be specified later. Notice that Gas is separable with respect to the local decision variables whereas Pow is not; thus this objective function is a special case of (4).

3) Constraints: Local constraints (2) of agent i (zone i) consist of a diffuser constraint $u_i^- \in [u_{i,\text{min}}^-, u_{i,\text{max}}^-]$, which models the minimum and maximum openings of the diffuser; and a thermal comfort constraint $T_i \in [T_{i,\text{min}}, T_{i,\text{max}}]$.

Because the three overhead diffusers are served by the same AHU and VAV, their total air flow is bounded by the supply air flow from the AHU, or equivalently, the total sensible coolings going into the room is bounded by the AHU capacity: $C_{\text{min}} \leq u_1^- + u_2^- + u_3^- \leq C_{\text{max}}$. Another shared constraint captures the thermal couplings between zones. Since we have a linear thermal dynamics model, this constraint can be cast in the standard form of Section III.

Remark 2. *Notice that all the upper and lower bounds of the constraints could be time-varying, determined by either the ambient temperature or the occupancy preferences.*

B. Controllers

The performance of three different controllers will be compared: 1) baseline controller; 2) centralized MPC; 3) distributed MPC using Proximal Jacobian ADMM.

1) Baseline controller: A simple feedback-type controller is designed to maintain the zone temperatures at the upper/lower bounds. Each diffuser will be adjusted locally by the corresponding agent without coordinating with each other. Zone temperatures are allowed to float freely between upper and lower bounds. Whenever the temperature in zone i is about to go above the pre-specified upper bound, diffuser i opens more to maintain it at its temperature upper bound. Clearly, this greedy control strategy is not able to utilize inter-zonal coordination.

2) Centralized MPC: This controller solves the centralized optimization problem (5) in a receding horizon fashion, and only the first control input is applied at each step.

3) Distributed MPC: Three agents solve problem (8) cooperatively using Algorithm 1, and only the first control input of each agent is applied. The following stopping criterion is used: $200 \leq v \leq 600$, and $\max(\|\mathbf{z}^v - \mathbf{z}^{v-1}\|, \|\mathbf{A}\mathbf{z}^v - \mathbf{b}\|, \|\widehat{\mathbf{C}}\mathbf{z}^v - \mathbf{d}\|) \leq 0.001$, where v is the iteration number.

C. Simulation Setup

TOU electricity prices are used in the simulation: \$0.1/kwh for on-peak hours (10am-17pm) and \$0.03/kwh for off-peak hours (17pm-10am). The gas price is assumed to be \$0.03/kwh. The sampling time step is $0.5h$ and the prediction horizon $N = 12h$ is used for both the centalized MPC and the distributed MPC. After a three-day warm-up period, the simulation is run for 7 days. It is assumed that 9am-17pm are occupied hours. The local thermal comfort interval for each zone is assumed to be [21.5°C, 23.5°C] during occupied hours and [20.5°C, 24.5°C] during unoccupied hours. Other parameters are set to be: $u^+_{i,\min} = 0$, $u^+_{i,\max} = $ 2kw, $u^-_{i,\min} = 0$, $u^-_{i,\max} = $ 4kw, $C_{\min} = 0$, $C_{\min} = $ 6kw. Optimization problems are solved numerically by CVX [21].

D. Simulation Results

The simulation results of the baseline controller, centralized MPC and distributed MPC are given in Fig. 2, Fig. 3 and Fig. 4, respectively. We only plot the first two days' zone temperature and control profiles due to space limitation. The total energy bills with the three controllers are summarized in TABLE 1 with the energy saving percentages of both MPCs compared to the baseline controller marked in parentheses.

TABLE I

Control Strategy	Energy Consumption ($)
Baseline Conroller	91.65
Centralized MPC	82.35 (10.15%↘)
Distributed MPC	83.64 (8.74%↘)

Several observations can be made based on the simulation results. Zone 1 has the highest load among all three zones during most of the days simulated, since physically it is closest to the double facade, and thus most heavily influenced by the ambient. With the baseline controller, each diffuser only focuses on its own local zone temperature in a short prediction window; thus there is no coordination between zones and the thermal storage of the building is not utilized. However, with both the centralized MPC and the distributed MPC, local agents are able to take advantage of the building's thermal storage as well as the TOU electricity pricing: agents 2 and 3 pre-cool zones 2 and 3 before 10am when the electricity price is low, in order to store extra cooling energy in the building thermal mass and release it into the room space during the peak hours. In addition, we can observe some coordination between agent 2 and agent 3 as they shift their pre-cooling peaks to different periods so that the total power level is relatively flat. The reason that agent 1 does not pre-cool zone 1 even though the three zones have similar thermal mass is that zone 1 is coupled to the double facade. If agent 1 also pre-cooled zone 1, much of the pre-cooling energy would be lost due to the coupling with double facade, resulting in a lower energy storage efficiency.

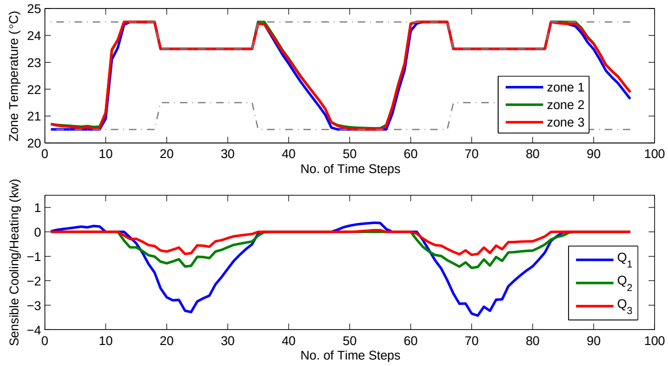

Fig. 2. Baseline controller

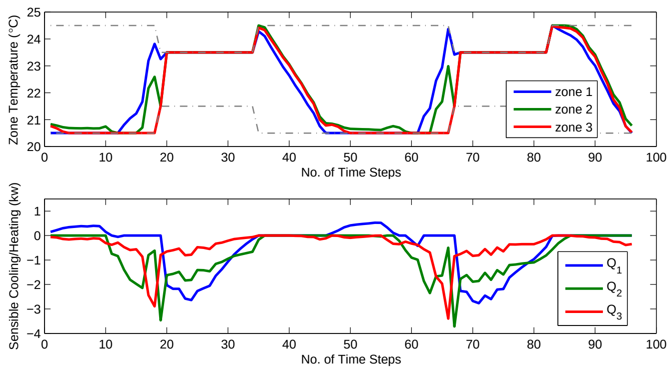

Fig. 3. Centralized MPC

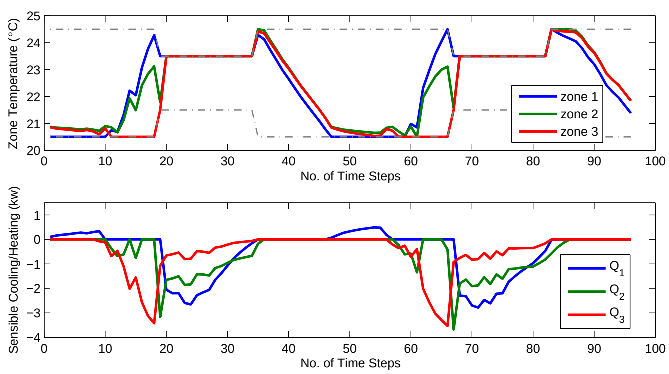

Fig. 4. Distributed MPC

E. Computation Time

With the stopping criteria defined in Section V.B, DMPC is able to terminate in less than 400 iterations most of the times, resulting a computation time of 3 minutes (assuming a perfectly synchronized parallel computation setting), which is much smaller than the $0.5h$ decision/sampling time. To further demonstrate the superiority of the proposed algorithm to the standard Gauss-Seidel ADMM, where different agents perform updates in serial instead of in parallel, we compare their convergence speed. Specifically, we plot the maximum absolute violation of the shared equality constraints, as well as the objective function value at each iteration in Fig. 5.

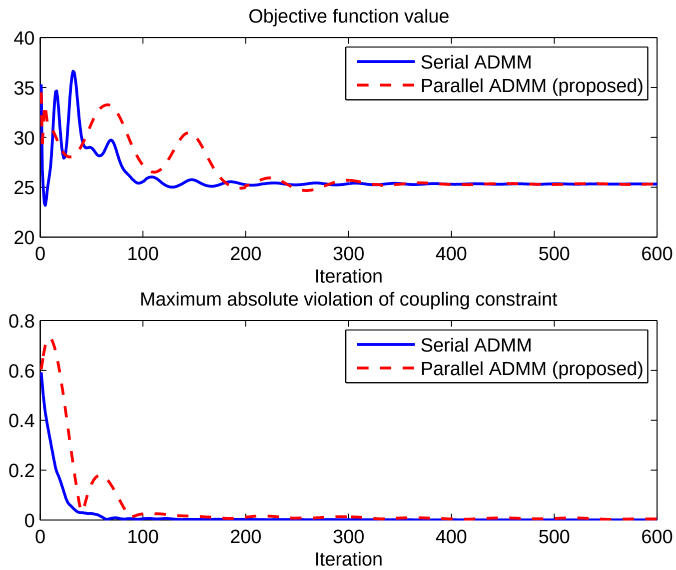

Fig. 5. Convergence comparison under same ρ_1 and ρ_2

From the plots, it is noted that both algorithms achieve almost feasible solutions after approximately 100 iterations. As for the objective function values, the proposed algorithm takes about 400 iterations to converge to the optimal value, compared to 300 iterations with the serial multi-block Gauss-Seidel ADMM. However, if implemented in a truly parallel scheme and ignoring communication overheads, the proposed algorithm will take much less time to converge. And as the number of agent increases, this advantage will become increasingly more significant.

Finally, for this particular case study, the proposed distributed solution does not outperform the centralized MPC in terms of convergence/computation time due to the relatively small scale of the problem and the very strong couplings (in the thermal dynamics and cost functions). However, many real life building control problems have significantly larger size, which makes the centralized problem very difficult or even impossible to solve. In comparison, with the distributed solution, a central coordinator only needs to update the dual variable without solving any optimization problem, making the DMPC scheme much more scalable to problem size.

VI. Conclusion

This paper presents a DMPC framework via the Proximal Jacobian ADMM method for building control applications. The proposed framework can be applied to solve a class of problems involving multi-zone buildings served by multiple HVAC equipment. The distributed solution method using Proximal Jacobian ADMM leads to an agent-based parallel updating scheme with guaranteed convergence. The case study for the HVAC energy minimization of a multi-zone building not only demonstrates the effectiveness of the proposed DMPC solution, but also shows the benefit of coordination between zones and their local control equipment. Future directions include incorporating demand charges into the objective function and investigating more realistic case studies.

VII. Appendix

Proof of Theorem 1. This proof is based on the work in [20], with some modifications.

Assumption 1. $\mathcal{S}^* = (\mathbf{z}_0^*, \ldots, \mathbf{z}_L^*, \lambda^*, \mu^*)$ *is a solution to problem* (8)*, thus satisfies its KKT condition.*

Proof. First write out the first order optimality condition of the unconstrained strongly convex optimization problem (9),

$$\begin{aligned} \mathbf{s}_{i1} = &- \mathbf{A}_i^\top \Big(\lambda^v + \rho_1\big(\mathbf{A}_i\mathbf{z}_i^{v+1} + \sum_{j\neq i}\mathbf{A}_j\mathbf{z}_j^v - \mathbf{b}\big)\Big) \\ &- \widehat{\mathbf{C}}_i^\top \Big(\mu^v + \rho_2\big(\widehat{\mathbf{C}}_i\mathbf{z}_i^{v+1} + \sum_{j\neq i}\widehat{\mathbf{C}}_j\mathbf{z}_j^v - \mathbf{d}\big)\Big) + \varphi_i(\mathbf{z}_i^v - \mathbf{z}_i^{v+1}) \in \partial\widehat{F}_i(\mathbf{z}_i^{v+1}), \end{aligned} \tag{10}$$

where $\widehat{F}_i = F_i + I_{\mathcal{Z}_{Ni}}$, and $I_{\mathcal{Z}_{Ni}}(\mathbf{z}_i)$ is a convex indicator function that takes the value of 0 if $\mathbf{z}_i$ is inside $\mathcal{Z}_{Ni}$, $+\infty$ othewise. By Assumption 2, $(\mathbf{z}_0^*, \ldots, \mathbf{z}_L^*, \lambda^*, \mu^*)$ satisfies the KKT condition of problem (8),

$$\begin{aligned} &\mathbf{s}_{i2} = -\mathbf{A}_i^\top\lambda^* - \widehat{\mathbf{C}}_i^\top\mu^* \in \partial\widehat{F}_i(\mathbf{z}_i^*), \quad i = 0, 1, \ldots, L, \\ &\mathbf{A}\mathbf{z}^* = \mathbf{b}, \quad \widehat{\mathbf{C}}\mathbf{z}^* = \mathbf{d}. \end{aligned} \tag{11}$$

Observing (10) and (11), from the monotonicity property of subdifferential of convex functions, one has

$$(\mathbf{s}_{i1} - \mathbf{s}_{i2})^\top(\mathbf{z}_i^{v+1} - \mathbf{z}_i^*) \geq 0, \tag{12}$$

Notice $\lambda^{v+1} = \lambda^v + \rho_1(\mathbf{A}\mathbf{z}^{v+1} - \mathbf{b})$, and $\mu^{v+1} = \mu^v + \rho_2(\widehat{\mathbf{C}}\mathbf{z}^{v+1} - \mathbf{d})$. Rewrite $\mathbf{s}_{i1}$ in terms of λ^{v+1} and μ^{v+1}, then plug $\mathbf{s}_{i1}$ and $\mathbf{s}_{i2}$ in to (12), one obtains

$$-\Big(\lambda^{v+1} - \lambda^* + \rho_1\sum_{j\neq i}\mathbf{A}_j\big(\mathbf{z}_j^v - \mathbf{z}_j^{v+1}\big)\Big)^\top\mathbf{A}_i(\mathbf{z}_i^{v+1} - \mathbf{z}_i^*)$$

$$-\Big(\mu^{v+1}-\mu^*+\rho_2\sum_{j\neq i}\widehat{\mathbf{C}}_j\big(\mathbf{z}_j^v-\mathbf{z}_j^{v+1}\big)\Big)^\top\widehat{\mathbf{C}}_i(\mathbf{z}_i^{v+1}-\mathbf{z}_i^*)+\varphi_i(\mathbf{z}_i^v-\mathbf{z}_i^{v+1})^\top(\mathbf{z}_i^{v+1}-\mathbf{z}_i^*)\geq 0.$$

The above inequality holds true for all $i=0,1,\ldots,L$. Sum up the inequalities for all i and use the facts that $\sum_{j\neq i}\mathbf{A}_j\big(\mathbf{z}_j^v-\mathbf{z}_j^{v+1}\big)=\mathbf{A}\big(\mathbf{z}^v-\mathbf{z}^{v+1}\big)-\mathbf{A}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)$ and $\sum_{j\neq i}\widehat{\mathbf{C}}_j\big(\mathbf{z}_j^v-\mathbf{z}_j^{v+1}\big)=\widehat{\mathbf{C}}\big(\mathbf{z}^v-\mathbf{z}^{v+1}\big)-\widehat{\mathbf{C}}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)$, one obtains

$$\begin{aligned}&-\big(\lambda^{v+1}-\lambda^*\big)^\top\mathbf{A}(\mathbf{z}^{v+1}-\mathbf{z}^*)-\big(\mu^{v+1}-\mu^*\big)^\top\widehat{\mathbf{C}}(\mathbf{z}^{v+1}-\mathbf{z}^*)\\&+\sum_{i=0}^{L}(\mathbf{z}_i^v-\mathbf{z}_i^{v+1})^\top(\varphi_i\mathbf{I}+\rho_1\mathbf{A}_i^\top\mathbf{A}_i+\rho_2\widehat{\mathbf{C}}_i^\top\widehat{\mathbf{C}}_i)(\mathbf{z}_i^{v+1}-\mathbf{z}_i^*)\\&-\big(\mathbf{z}^v-\mathbf{z}^{v+1}\big)^\top\big(\rho_1\mathbf{A}^\top\mathbf{A}+\rho_2\widehat{\mathbf{C}}^\top\widehat{\mathbf{C}}\big)(\mathbf{z}^{v+1}-\mathbf{z}^*)\geq 0.\end{aligned}\tag{13}$$

Notice that $\mathbf{A}(\mathbf{z}^{v+1}-\mathbf{z}^*)=\mathbf{A}\mathbf{z}^{v+1}-\mathbf{b}=\frac{1}{\rho_1}(\lambda^{v+1}-\lambda^v)$, and $\widehat{\mathbf{C}}(\mathbf{z}^{v+1}-\mathbf{z}^*)=\widehat{\mathbf{C}}\mathbf{z}^{v+1}-\mathbf{d}=\frac{1}{\rho_2}(\mu^{v+1}-\mu^v)$. Substitute them into (13) and re-arrange terms yields

$$\begin{aligned}&\frac{1}{\rho_1}(\lambda^v-\lambda^{v+1})^\top(\lambda^{v+1}-\lambda^*)+\frac{1}{\rho_2}(\mu^v-\mu^{v+1})^\top(\mu^{v+1}-\mu^*)\\&\qquad+\sum_{i=0}^{L}(\mathbf{z}_i^v-\mathbf{z}_i^{v+1})^\top(\varphi_i\mathbf{I}+\rho_1\mathbf{A}_i^\top\mathbf{A}_i+\rho_2\widehat{\mathbf{C}}_i^\top\widehat{\mathbf{C}}_i)(\mathbf{z}_i^{v+1}-\mathbf{z}_i^*)\\&\geq-\big(\mathbf{z}^v-\mathbf{z}^{v+1}\big)^\top\big(\mathbf{A}^\top(\lambda^v-\lambda^{v+1})+\widehat{\mathbf{C}}^\top(\mu^v-\mu^{v+1})\big)\\&\geq-\frac{1}{2}\sum_{i=0}^{L}\Big(\frac{\rho_1}{\kappa_i}\|\mathbf{A}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)\|^2+\frac{\kappa_i}{\rho_1}\|\lambda^v-\lambda^{v+1}\|^2\Big)\\&\qquad-\frac{1}{2}\sum_{i=0}^{L}\Big(\frac{\rho_2}{\kappa_i}\|\widehat{\mathbf{C}}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)\|^2+\frac{\kappa_i}{\rho_2}\|\mu^v-\mu^{v+1}\|^2\Big)\\&=\frac{-\sum_{i=0}^{L}\kappa_i}{2\rho_1}\|\lambda^v-\lambda^{v+1}\|^2+\frac{-\sum_{i=0}^{L}\kappa_i}{2\rho_2}\|\mu^v-\mu^{v+1}\|^2\\&\qquad-\sum_{i=0}^{L}\big(\frac{\rho_1}{2\kappa_i}\|\mathbf{A}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)\|^2\big)-\sum_{i=0}^{L}\big(\frac{\rho_2}{2\kappa_i}\|\widehat{\mathbf{C}}_i\big(\mathbf{z}_i^v-\mathbf{z}_i^{v+1}\big)\|^2\big)\end{aligned}\tag{14}$$

The last inequality in (14) is due to the traingle inequality.

Define $P_{\mathbf{z}}=\mathrm{diag}\big(\varphi_0\mathbf{I}+\rho_1\mathbf{A}_0^\top\mathbf{A}_0+\rho_2\widehat{\mathbf{C}}_0^\top\widehat{\mathbf{C}}_0,\ldots,\varphi_L\mathbf{I}+\rho_1\mathbf{A}_L^\top\mathbf{A}_L+\rho_2\widehat{\mathbf{C}}_L^\top\widehat{\mathbf{C}}_L\big)\succ 0$, and $P=\mathrm{diag}\big(P_{\mathbf{z}},\frac{1}{\rho_1},\frac{1}{\rho_2}\big)\succ 0$. Then we calculate the difference of two consecutive elements in the sequence $\mathcal{P}^v=\|\mathcal{S}^v-\mathcal{S}^*\|_P^2$,

$$\|\mathcal{S}^v-\mathcal{S}^*\|_P^2-\|\mathcal{S}^{v+1}-\mathcal{S}^*\|_P^2=\|\mathcal{S}^v-\mathcal{S}^{v+1}\|_P^2+2\big(\mathcal{S}^v-\mathcal{S}^{v+1}\big)^\top P\big(\mathcal{S}^{v+1}-\mathcal{S}^*\big).$$

Notice that $\big(\mathcal{S}^v - \mathcal{S}^{v+1}\big)^\top P\big(\mathcal{S}^{v+1} - \mathcal{S}^*\big)$ is the first term (most left hand side) of inequality (14), after simple manipulations one obtains

$$\|\mathcal{S}^v - \mathcal{S}^*\|_P^2 - \|\mathcal{S}^{v+1} - \mathcal{S}^*\|_P^2 \geq \sum_{i=0}^{L}(\mathbf{z}_i^v - \mathbf{z}_i^{v+1})^\top$$
$$\left(\varphi_i\mathbf{I} + \big(1 - \frac{1}{\kappa_i}\big)\big(\rho_1\mathbf{A}_i^\top\mathbf{A}_i + \rho_2\widehat{\mathbf{C}}_i^\top\widehat{\mathbf{C}}_i\big)\right)(\mathbf{z}_i^v - \mathbf{z}_i^{v+1})$$
$$+ (1 - \sum_{i=0}^{L}\kappa_i)\left(\frac{1}{\rho_1}\|\lambda^v - \lambda^{v+1}\|^2 + \frac{1}{\rho_2}\|\mu^v - \mu^{v+1}\|^2\right),$$

If we assume $\kappa_i < \frac{1}{L}$, then $1-\sum_{i=0}^{L}\kappa_i > 0$. And the assumption $\varphi_i > (L-1)(\rho_1\|\mathbf{A}_i\|^2+\rho_2\|\widehat{\mathbf{C}}_i\|^2)$ implies $\varphi_i\mathbf{I} + \big(1 - \frac{1}{\kappa_i}\big)\big(\rho_1\mathbf{A}_i^\top\mathbf{A}_i + \rho_2\widehat{\mathbf{C}}_i^\top\widehat{\mathbf{C}}_i\big) \succ \mathbf{0}$. Therefore,

$$\|\mathcal{S}^v - \mathcal{S}^*\|_P^2 - \|\mathcal{S}^{v+1} - \mathcal{S}^*\|_P^2 \geq \beta\|\mathcal{S}^v - \mathcal{S}^{v+1}\|_P^2. \tag{15}$$

for some $\beta > 0$. Sum up inequality (15) for all v yields

$$\sum_{v=0}^{\infty}\|\mathcal{S}^v - \mathcal{S}^{v+1}\|_P^2 \leq \frac{1}{\beta}\big(\|\mathcal{S}^0 - \mathcal{S}^*\|_P^2 - \|\mathcal{S}^\infty - \mathcal{S}^*\|_P^2\big) \leq \frac{1}{\beta}\|\mathcal{S}^0 - \mathcal{S}^*\|_P^2,$$

which means that the sequence $\mathcal{Q}^v = \sum\|\mathcal{S}^v - \mathcal{S}^{v+1}\|_P^2$ is monotonically increasing and upper bounded. Therefore, $\mathcal{Q}^v$ converges $\lim_{v\to\infty}\|\mathcal{S}^v - \mathcal{S}^{v+1}\|_P^2 = 0$.

In addition, (15) tells us $\mathcal{P}^v = \|\mathcal{S}^v - \mathcal{S}^*\|_P^2$ is a monotonically decreasing sequence, and this sequence is bounded below by 0. Thus, $\|\mathcal{S}^v - \mathcal{S}^*\|_P^2$ also converges (not necessarily to 0). Then, $\mathcal{S}^v$ converges to a fixed point of the mapping defined by (9). □

References

[1] J. Maestre, D. M. De La Pena, E. Camacho, and T. Alamo, "Distributed model predictive control based on agent negotiation," *Journal of Process Control*, vol. 21, no. 5, pp. 685–697, 2011.

[2] Y. Wakasa, M. Arakawa, K. Tanaka, and T. Akashi, "Decentralized model predictive control via dual decomposition," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 2008, pp. 381–386.

[3] F. Farokhi, I. Shames, and K. H. Johansson, *Distributed MPC via dual decomposition and alternative direction method of multipliers*. Springer, 2014.

[4] I. Necoara, D. Doan, and J. A. Suykens, "Application of the proximal center decomposition method to distributed model predictive control," in *Decision and Control, 2008. CDC 2008. 47th IEEE Conference on*. IEEE, 2008, pp. 2900–2905.

[5] J.-H. Hours and C. N. Jones, "A parametric nonconvex decomposition algorithm for real-time and distributed nmpc," *IEEE Transactions on Automatic Control*, vol. 61, no. 2, pp. 287–302, 2016.

[6] D. Groß and O. Stursberg, "On the convergence rate of a jacobi algorithm for cooperative distributed mpc," in *52nd IEEE Conference on Decision and Control*. IEEE, 2013, pp. 1508–1513.

[7] R. Scattolini, "Architectures for distributed and hierarchical model predictive control–a review," *Journal of Process Control*, vol. 19, no. 5, pp. 723–731, 2009.

[8] P. D. Christofides, R. Scattolini, D. M. de la Peña, and J. Liu, "Distributed model predictive control: A tutorial review and future research directions," *Computers & Chemical Engineering*, vol. 51, pp. 21–41, 2013.

[9] M. Y. Lamoudi, M. Alamir, and P. Béguery, "Distributed constrained model predictive control based on bundle method for building energy management," in *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE, 2011, pp. 8118–8124.

[10] P.-D. Moroşan, R. Bourdais, D. Dumur, and J. Buisson, "Distributed model predictive control based on benders' decomposition applied to multisource multizone building temperature regulation," in *49th IEEE Conference on Decision and Control (CDC)*. IEEE, 2010, pp. 3914–3919.

[11] Y. Ma, G. Anderson, and F. Borrelli, "A distributed predictive control approach to building temperature regulation," in *Proceedings of the 2011 American Control Conference*. IEEE, 2011, pp. 2089–2094.

[12] J. Cai, D. Kim, R. Jaramillo, J. E. Braun, and J. Hu, "A general multi-agent control approach for building energy system optimization," *Energy and Buildings*, vol. 127, pp. 337–351, 2016.

[13] J. Cai, J. E. Braun, D. Kim, and J. Hu, "General approaches for determining the savings potential of optimal control for cooling in commercial buildings having both energy and demand charges," *Science and Technology for the Built Environment*, pp. 1–18, 2016.

[14] M. Hong and Z.-Q. Luo, "On the linear convergence of the alternating direction method of multipliers," *arXiv:1208.3922*, 2012.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.

[16] C. Chen, B. He, Y. Ye, and X. Yuan, "The direct extension of admm for multi-block convex minimization problems is not necessarily convergent," *Mathematical Programming*, vol. 155, no. 1-2, pp. 57–79, 2016.

[17] D. P. Bertsekas and J. N. Tsitsiklis, *Parallel and distributed computation: numerical methods*. Prentice hall Englewood Cliffs, NJ, 1989, vol. 23.

[18] X. Wang, M. Hong, S. Ma, and Z.-Q. Luo, "Solving multiple-block separable convex minimization problems using two-block alternating direction method of multipliers," *arXiv preprint arXiv:1308.5294*, 2013.

[19] B. He, L. Hou, and X. Yuan, "On full jacobian decomposition of the augmented lagrangian method for separable convex programming," *SIAM Journal on Optimization*, vol. 25, no. 4, pp. 2274–2312, 2015.

[20] W. Deng, M.-J. Lai, Z. Peng, and W. Yin, "Parallel multi-block admm with o (1/k) convergence," *arXiv preprint arXiv:1312.3040*, 2013.

[21] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 2.1," http://cvxr.com/cvx, Mar. 2014.