Identification of Network Dynamics and Disturbance for a Multi-zone Building

Tingting Zeng, Prabir Barooah*
Department of Mechanical and Aerospace Engineering, University of Florida, Gainesville, FL, 32611 USA

***Abstract*—We propose a method that simultaneously identifies a sparse transfer matrix and a disturbance signal for a multi-zone building's temperature dynamics from measurements of inputs and outputs. The proposed method is based on solving a convex optimization problem whose cost function involves an ℓ_1-penalty to promote a sparse solution. The method ensures that the transfer matrix is sparse, so that only dominant interactions among zones are retained in the model. The disturbance, which is mostly occupant-induced, is assumed to be a piecewise-constant signal, which aids identification since the derivative of a piecewise-constant signal is a sparse signal. We test our method on data from a virtual building (a simulation model) and a real building. Results from the virtual building show that the proposed method can accurately identify a sparse network model and a transformed disturbance. Results from the real building data - that does not have a ground truth - show that the method produces sensible results.**

***Index Terms*—system identification, ℓ_1-penalty, sparsity, disturbance, MIMO system**

I. INTRODUCTION

The heating, ventilation, and air conditioning (HVAC) system of a building is responsible for maintaining indoor climate, which directly affects health and comfort of occupants [1], [2]. HVAC systems are also a major consumer of energy [3]. There is a growing and significant interest in improving energy efficiency of HVAC systems through advanced control methods. Occupant comfort and health serves as important constraints to the control objective: comfort plays a crucial role in determining productivity [4], [5]. More recently there is interest in providing occupants of office buildings *personalized comfort* through advanced control. Individuals differ in their perceptions of thermal comfort [6], [7]. Ideally the thermal and air quality measure of each occupant's space should be controlled to meet that particular individual's preference. Products such as comfy$^{\text{TM}}$(`www.comfy.app`) are examples of such efforts. In these applications, energy use may serve as a constraint rather than being the objective.

Such control applications need a low-order model relating the inputs to outputs. For a modern commercial building equipped with variable air volume (VAV) systems, control inputs may include rate of airflow and rate of cooling/heating. Relevant outputs include temperature at a minimum, but can also include humidity, concentration of pollutants etc. In this paper we consider temperature as the output, and denote by "zone" a space whose temperature has an impact on at least one occupant's comfort and health.

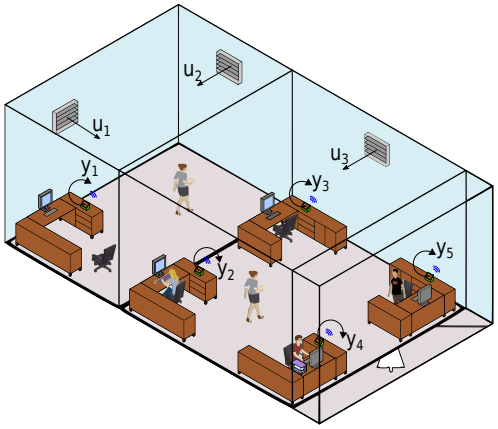

Fig. 1. An example of a building layout with multiple inputs and outputs.

The complexity of the physical processes that determines the input-output behavior makes control-oriented modeling a challenging task. The modeling problem gets exacerbated when one considers an open-plan office, or a building in which there are considerable thermal interactions among the zones, such as the one shown in Figure 1. Due to inter-zone thermal interaction, each actuator (VAV box) may impact several outputs (temperatures at several zones).

This paper is about identifying control-oriented models of multi-zone buildings from measurements of inputs and outputs. It is envisioned that the identified models will be used for predictive control to provide higher energy efficiency and/or personalized comfort, though the control design problem is outside the scope of this paper. A data-driven approach is chosen over a physics-based approach for modeling because of the complexity of the underlying physics.

A key bottleneck in data-driven identification of thermal models of buildings is the presence of large exogenous disturbances (various unknown heat gains). The thermal energy expended by the HVAC system is mostly spent to reject these disturbances; hence these disturbances cannot be assumed to be small. Even without inter-zone interactions, presence of unknown disturbances has been a bottleneck in identifying dynamic models of single-zone building models. Even though there is a long history of identifying building models from data (see, e.g., [8], [9]), only recently has the effect of

* Corresponding author
E-mail: tingtingzeng@ufl.edu (Tingting Zeng), pbarooah@ufl.edu (Barooah Prabir)

This research is partially funded by the National Science Foundation (award # 1463316 and 1646229) and a Renewable Energy and Energy Efficient Technologies (REET) grant from the Florida Department of Consumer Services.

disturbances been addressed in a principled manner; see [10], [11]. These papers are limited to single zones.

The problem we seek to address is that of simultaneously identifying (1) a multiple-input multiple-output (MIMO) transfer matrix - the plant - and (2) the exogenous disturbances, from measurements of inputs and outputs. The transfer matrix can be visualized as a bipartite graph, with edges from the input nodes to output nodes; see Figure 2. We seek to identify a *sparse graph* with a small number of edges. It is hoped that a sparse model will retain information of only the dominant interactions, and thereby help control computations.

Allowing an arbitrary disturbance signal is likely to lead to over-fitting: the disturbance absorbs all model mismatch and thus is not useful for prediction. We assume that exogenous disturbance is *piecewise-constant*, so that its derivative is a sparse signal, i.e., mostly zero. The motivation behind this assumption is as follows. The disturbance is mostly occupant-induced: it is dominated by heat gain from occupants' bodies and from equipments such as computers, lights and printers. And, occupancy in commercial buildings can be modeled as a piecewise-constant signal: people come in at a particular time and leave at the end of the workday, perhaps with a temporarily absence during lunch [12]. Since occupancy varies in an approximately piecewise-constant manner, so should the disturbance induced by occupants.

The identification problem is posed as an ℓ_1-regularized least squares problem; the ℓ_1-penalty promotes a sparse solution. Recent years have seen advances in both theory and applications of identifying sparse models through ℓ_1-regularization; see [13]–[15] and references therein.

While there are a number of papers on the problem of modeling multi-zone building dynamics (see [16]–[18] and references therein), to the best of our knowledge the only reference that simultaneously identifies a plant model and disturbance for a multi-zone building is [19]. Ref. [19] estimates the plant parameters and an output disturbance that encapsulates the effect of an unknown input disturbance. The method results in a complete interaction model connecting each input-output pair. The method in [19] require solving a non-convex problem. In contrast, the method proposed here involves convex optimization and results in a sparse interaction model. We also show that the proposed method always produces a unique solution. A related reference on multi-zone modeling is [20], which identifies a single-zone equivalent model of a multi-zone building.

A recent body of work on network structure identification is closely related to this paper, e.g., [14], [15], [21], [22]. Ref. [21] is concerned with the blind identification problem in which input measurements are not available, only output measurements are. Ref. [15] identifies the binary structure of the network. Ref. [14] is concerned with experiment design to aid identification of the network structure, and [22] provides an identification method when a subset of nodes can be actuated and measured.

A preliminary version of this work is presented in [23]. Compared to that paper this article makes two contributions. One, we provide results characterizing the optimization problem's solution that furthermore helps in choosing the regularization parameter. Two, we provide evaluation of the method on data from a real building.

The rest of the paper is organized as follows. Section II formulates the problem precisely, Section III describes the proposed algorithm, and Section IV describes evaluation results. The paper concludes with some comments in Section V.

II. Problem formulation

The *output node set*, denoted by $\mathcal{V}_y$, is the set of temperature measurements that are outputs of the model we wish to identify. The output nodes can be visualized as locations with temperature sensors. The number of *output nodes* is denoted by n_y (= $|\mathcal{V}_y|$, $|\cdot|$ denotes cardinality). The *input node set* $\mathcal{V}_u$ contains the set of controllable inputs (cooling or heating rate) and the set of exogenous inputs (outside air temperature and solar irradiance). The controllable inputs can be physically visualized as VAV boxes; there is a one to one mapping between them. In the building shown in Figure 1, there are 5 inputs (3 VAV boxes, outside temperature and solar irradiance) and 5 outputs. Figure 4 shows another example; it has 6 input nodes (4 VAV boxes) and 10 output nodes (8 cubicles and 2 conference rooms).

Time is measured by a discrete counter $k = 0, 1, 2, \ldots$. At time k, let $y_i[k] \in \mathbb{R}$ be the measured air temperature at output node i, $u_j[k]$ be the vector of three known inputs : (1) the rate of heat gain due to the air supply, q_{hvac}(kW), (2) the outside air temperature T_{oa} (°C), (3) the solar irradiance η^{sol}(kW/m^2), at input node j, and $w_i[k]$ be the unknown disturbance (total heat gain due to occupants, plug loads etc.), at output node i. We denote $y[k] = [y_1[k], y_2[k], \ldots, y_{n_y}[k]]^T$, $u[k] = [u_1[k], u_2[k], \ldots, u_{n_u}[k]]^T$, and $w[k] = [w_1[k], w_2[k], \ldots, w_{n_y}[k]]^T$.

Our starting point is a MIMO (multiple-input-multiple output) transfer matrix in Z-transform domain:

$$\begin{aligned} Y(z^{-1}) &= G_u(z^{-1})U(z) + G_w(z^{-1})W(z^{-1}) \\ &= \frac{1}{D(z^{-1})}\left(N_u(z^{-1})U(z^{-1}) + N_w(z^{-1})W(z^{-1})\right) \\ &= \frac{1}{D(z^{-1})}\left(N_u(z^{-1})U(z^{-1}) + \bar{W}(z^{-1})\right), \end{aligned} \tag{1}$$

with $Y, \bar{W} \in \mathbb{C}^{n_y}$, $G_u, N_u \in \mathbb{C}^{N_y \times N_u}$, $D \in \mathbb{C}, G_w$, $N_w \in \mathbb{C}^{N_y \times N_y}$, where the second equality is obtained by making $D(z^{-1})$ the least common multiple of the denominators of $G_u(z^{-1})$ and $G_w(z^{-1})$, and the third equality is obtained by defining $\bar{W}(z^{-1}) := N_w(z^{-1})W(z^{-1})$. Here we assume N_w is diagonal, meaning output y_i is only affected by disturbance w_i. Denote ρ as the order of $D(z^{-1})$, then,

$$D(z^{-1}) = 1 - \sum_{k=1}^{\rho} \theta_k z^{-k},$$

$$N_u(z^{-1}) = \begin{bmatrix} N_u^{(1,1)}(z^{-1}) & \ldots & N_u^{(1,j)}(z^{-1}) & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ N_u^{(i,1)}(z^{-1}) & \ldots & N_u^{(i,j)}(z^{-1}) & \ldots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix},$$

$$\bar{W}(z^{-1}) = \begin{bmatrix} \dots, & N_{w_i}(z^{-1})w_i(z^{-1}), & \dots \end{bmatrix}^T,$$

for some *plant parameters* θ_k's, in which

$$N_u^{(i,j)}(z^{-1}) = \sum_{k=0}^{\rho} \theta_{(n_u(i-1)+j)(\rho+1)+k} z^{-k}.$$

The interactions captured by the transfer matrix model described above can be visualized by a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ with node set $\mathcal{V} = \mathcal{V}_u \cup \mathcal{V}_y$ and edge set $\mathcal{E} \subset \mathcal{V}_u \times \mathcal{V}_y$. An edge $e = (i, j)$ from $j \in \mathcal{V}_u$ to $i \in \mathcal{V}_y$ exists if and only if the $(i, j)^{th}$ entry of the numerator matrix N_u satisfies $N_{ji}(z) \neq 0$ for every z. It is a bipartite graph without cycles: an edge is possible between an input node and an output node (sensor) but not between input pairs or output pairs [24]. In the extreme case when every actuator affects every sensor, the graph will look like what is shown in Figure 2. Note an output node y_i is also affected by disturbance w_i. In practice the graph will be far more sparse: only ambient temperature and solar irradiance can affect many outputs, each of the other inputs can only affect a small number of outputs.

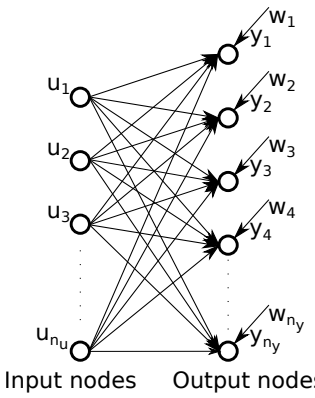

Fig. 2. The most general case of the network structure of a multi-zone HVAC model: a complete directed bipartite graph without cycles, with an edge from every input node to every output node.

Following assumptions are made to reduce the number of free parameters. The first assumption means there is a "building-level disturbance"; the disturbance at a zone is a scaled version of it. The second assumption means the dynamic responses from disturbances to temperatures are the same for every zone.

Assumption 1: Define $w_{bldg}[k] := \sum_{i=1}^{n_y} w_i[k]$. We assume $\exists\ k_w = [k_w^1, \dots, k_w^{n_y}]^T \in \mathbb{R}^{n_y}$, $\sum_{i=1}^{n_y} k_w^i = 1$, such that $w_i = k_w^i w_{bldg}$, $\forall\ i = 1, \dots, n_y$.

Assumption 2:

$\exists\ N_w^{bldg}(z^{-1})$, *s.t.* $N_w^{bldg}(z^{-1}) = N_{w_i}(z^{-1})$, $\forall\ i = 1, \dots, n_y$.

Then $\bar{W}(z^{-1})$ can be rewritten as

$$\bar{W}(z^{-1}) = \begin{bmatrix} \dots, & k_w^i N_w^{bldg}(z^{-1}) w_{bldg}(z^{-1}), & \dots \end{bmatrix}^T.$$

Performing an inverse Z-transformation on (1) and defining

$$\bar{w}_{bldg}(z^{-1}) := Z^{-1}\left(N_w^{bldg}(z^{-1}) w_{bldg}(z^{-1})\right) \tag{2}$$

as the *transformed* disturbance, yields a difference equation, from which we obtain the linear regression form:

$$y[k] = \Phi[k]\theta, \qquad k = \rho + 1, \dots, k_{max} \tag{3}$$

where $\theta = [\theta_{p_d}^T, \theta_{p_n}^T, \bar{w}_{bldg}^T]^T$, in which $\theta_{p_d} = [\theta_1, \dots, \theta_\rho]^T$ are plant parameters that appear in $D(z^{-1})$, $\theta_{p_n} = [\theta_{\rho+1}, \dots, \theta_{n_y n_u(\rho+1)+\rho}]^T$ are plant parameters that appear in $N_u(z^{-1})$, and $\bar{w}_{bldg} = [\bar{w}_{bldg}[\rho+1], \dots, \bar{w}_{bldg}[k_{max}]]^T$ are the transformed disturbance defined in (2), where k_{max} is the number of the time indices for which data are collected. Eq. (3) can be further written as:

$$\Phi := \begin{bmatrix} \vdots \\ \phi^i \\ \vdots \end{bmatrix} \in \mathbb{R}^{n_y(k_{max}-\rho) \times \left(k_{max} + n_y n_u (\rho+1)\right)}, i = 1, \dots, n_y,$$

where each block ϕ^i has the form:

$$\phi^i = \left[\begin{array}{ccc|ccc|c} y_i[\rho+1] & \dots & y_i[1] & \phi^i_{u_1} & \dots & \phi^i_{u_j} \ \dots & k_w^i e_1 \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \\ y_i[k] & \dots & y_i[k-\rho] & \phi^i_{u_1} & \dots & \phi^i_{u_j} \ \dots & k_w^i e_{k-\rho} \\ \vdots & & \vdots & \vdots & & \vdots & \vdots \end{array}\right]$$
$$\in \mathbb{R}^{(k_{max}-\rho) \times \left(k_{max} + n_y n_u (\rho+1)\right)}, \quad j = 1, \dots, n_u,$$

where e_k is the k-th canonical basis vector of $\mathbb{R}^{k_{max}-\rho}$ in which the 1 appears in the k-th place, and

$$\phi^i_{u_j} = \begin{bmatrix} 0_{1\times(\rho+1)(i-1)} & | & u_j[k] & \dots & u_j[k-\rho] & | & 0_{1\times(\rho+1)(n_y-i)} \end{bmatrix}$$
$$\in \mathbb{R}^{1 \times n_y(\rho+1)}.$$

Our goal is to identify from measurements of inputs and outputs, (1) a transfer matrix $G_u(z^{-1})$ that is sparse, meaning its graph has a small number of edges, and (2) the signal $\{\bar{w}_{bldg}[k]\}$, so that the resulting model (transfer matrix and transformed disturbance) predicts the output well.

Since the number of samples is typically large, Φ is a tall matrix. And, due to the dependency of Φ on (noisy) measurements of inputs and outputs and the presence of canonical basis vectors on its rightmost block, it is straightforward to verify that Φ is full column-rank except in case of degenerate data. That is, Φ is full column-rank almost surely (a.s.). Though a least squares solution to (3) can be obtained, it is not likely to lead to a sparse solution. We therefore seek an alternative to least squares.

III. Proposed Algorithm

We seek to obtain a sparse model capturing only dominant interactions from input nodes to output nodes. Such a model can be promoted by penalizing the ℓ_1-norm of θ_{p_n}, the vector of the coefficients that determines the numerator matrix $N_u(z^{-1})$. Note that the parameters that determine the denominator $D(z^{-1})$ need not to be sparse. As for the disturbance, recall the discussion in Section I that the disturbance can be assumed to be piecewise-constant. Since the transformed disturbance $\bar{w}_{bldg}[k]$ is a linear combination of time-shifted value of the disturbance $w_{bldg}[k]$, the same should hold for $\bar{w}_{bldg}[k]$. The derivative of such a signal with respect to time is sparse, i.e., mostly zero. Hence we seek a *piecewise-constant* $\bar{w}_{bldg}$ by penalizing the ℓ_1-norm of the derivative of $\bar{w}_{bldg}$.

The discrete time derivative is $D_1\bar{w}_{bldg}$, where D_1 is the first difference matrix:

$$D_1 := \begin{bmatrix} -1 & 1 & 0 & \dots & \dots \\ 0 & -1 & 1 & 0 & \dots \\ & & \ddots & \ddots & \\ \dots & \dots & 0 & -1 & 1 \end{bmatrix} \in \mathbb{R}^{(k_{max}-\rho-1)\times(k_{max}-\rho)}.$$

Let $S \in \mathbb{R}^{\left((\rho+1)(n_y n_u-1)+k_{max}\right)\times\left((\rho+1)n_y n_u+k_{max}\right)}$ be defined as

$$S := \left[\begin{array}{c|c|c} 0_{\left((\rho+1)(n_y n_u-1)+k_{max}\right)\times\rho} & I_{(\rho+1)n_y n_u} & 0 \\ & 0 & D_1 \end{array}\right],$$

so that

$$(S\theta)^T = [\theta_{p_n}^T, (D_1\bar{w}_{bldg})^T]. \tag{4}$$

We therefore pose the following optimization problem to estimate the plant transfer matrix and transformed disturbance:

$$\hat{\theta} = \arg\min_{\theta} \frac{1}{2}\|y - \Phi\theta\|_2^2 + \lambda\|S\theta\|_1, \tag{5}$$

where $\lambda \geq 0$ is an user-defined weight. The ℓ_1 penalty is used to encourage a solution $\hat{\theta}$ so that $S\hat{\theta}$ is sparse, which will make the network model have a small number of edges and the disturbance signal piecewise-constant. Since Φ is full column-rank a.s., the solution to (5) is unique a.s. due to strict convexity of the objective function.

The problem (5) is closely connected to the so-called lasso problem [13]:

$$\hat{\chi} = \arg\min_{\chi} \frac{1}{2}\|\boldsymbol{z} - \Psi\chi\|_2^2 + \lambda\|\chi\|_1. \tag{6}$$

Our problem (5) falls into the category of *generalized lasso* problem since the ℓ_1-norm penalty is not on the decision variable θ but on $S\theta$ [25].

A. Regularization Parameter Selection

The solution to (5) depends critically on the value of λ. In this section we propose a method to select λ. This requires first transforming (5) into form (6), then showing that there is a $\lambda_{\max}$ so that for all $\lambda \geq \lambda_{\max}$, the unique solution to (5) satisfies $S\hat{\theta} = 0$, and finally describing a heuristic to select λ by searching in the range $[0, \lambda_{\max}]$.

1) Determining λ_{max}: In order to determine the value of λ_{max}, first we need to transform our problem (5) into form (6) using the following theorem.

Theorem 1 ([26]): For $\Phi \in \mathbb{R}^{n\times p}$, $S \in \mathbb{R}^{m\times p}$, if matrix $H = [\Phi^T, S^T]^T$ has full column rank and S has full row rank, then the generalized lasso problem (5) can be transformed into the lasso problem (6) with transformation:

$$\chi = S\theta, \quad \Psi = Q_1 Q_2^{\dagger},$$
$$\boldsymbol{z} = [I_n - Q_1(I_p - Q_2^{\dagger}Q_2)Q_1^{\dagger}]\boldsymbol{y},$$

in which $Q_1 \in \mathbb{R}^{n\times p}$, $Q_2 \in \mathbb{R}^{m\times p}$ are defined from the unique QR decomposition:

$$H = \begin{bmatrix}\Phi \\ S\end{bmatrix} = QR = \begin{bmatrix}Q_1 \\ Q_2\end{bmatrix} R. \qquad \square$$

Our problem (5) can be transformed into the lasso form (6) as it satisfies conditions in Theorem 1: $H = [\Phi^T, S^T]^T$ has full column rank a.s. since Φ has full column rank a.s., and S has full row rank.

Proposition 1: If $\lambda \geq \lambda_{\max} := \|\Psi^T\boldsymbol{z}\|_{\infty}$, then the unique solution $\hat{\theta}$ to (5) produces a network with no edges ($\hat{\theta}_{p_n} = 0$) and a constant disturbance ($\hat{\bar{w}}_{bldg} = c\mathbf{1}$ for some c). $\square$

Proof of Proposition 1: From optimality condition for (6), we know

$$-\Psi^T(\boldsymbol{z} - \Psi\chi) + \lambda\nu = 0, \tag{7}$$

where ν is the sub-gradient of χ,

$$\nu_i \in \begin{cases} \{+1\}, & \chi_i > 0 \\ \{-1\}, & \chi_i < 0, \\ [-1, 1], & \chi_i = 0 \end{cases} \quad i = 1, ..., k_{\max} - 2. \tag{8}$$

Substituting a λ that is larger than $\|\Psi^T\boldsymbol{z}\|_{\infty}$ and $\chi = 0$ into (7) and solving for ν, we get $\nu = \frac{\Psi^T\boldsymbol{z}}{\lambda}$, which implies (8) also holds: $|\nu_i| \leq 1\ \forall i$ since $\lambda \geq \|\Psi^T\boldsymbol{z}\|_{\infty}$. Therefore $\chi = 0$ is a solution to (6). Notice $\Psi \in \mathbb{R}^{n\times m}$, with $n > m$, has full column rank due to

$$\begin{aligned} rank(\Psi) &= rank(Q_1 Q_2^{\dagger}) \\ &\geq rank(Q_1) + rank(Q_2^{\dagger}) - p = m. \end{aligned}$$

Therefore, our transformed lasso problem (6) - with $\chi = S\theta$ - is strictly convex and thus has a unique solution. From the analysis above, we know this unique solution will produce $\hat{\chi} = 0 = S\hat{\theta}$. The conclusion now follows from (4). ■

B. Heuristic for selecting λ

From the discussion above it follows that when $\lambda \geq \lambda_{max}$, the unique solution satisfies $S\hat{\theta} = 0$. That is, the estimated transfer matrix is 0 and the estimated transformed disturbance is a constant signal. As λ decreases, a non-zero transfer matrix will result. For $\lambda = 0$ we will get the least squares solution to $y = \Phi\theta$. This solution may overfit the data, resulting in a noisy transformed disturbance and a highly connected network model. Therefore, we expect that for a λ at some value between 0 and $\lambda_{\max}$ we will obtain a sparse $\hat{\theta}_{p_n}$ and a non-constant $\hat{\bar{w}}_{bldg}$ that will fit the data well without overfitting it, and are useful for prediction. The remaining problem is to pick a λ within the range $[0,\ \lambda_{\max}]$.

Two common heuristics for choosing λ for the lasso problem (6) are cross-validation [13] and L-curve-based curvature methods [27]. Cross validation divides datasets into multiple folders and requires that parameters to be estimated are the same in each folder. The parameters in our problem contain transformed disturbances, which may differ from one day to the next. The L-curve, which is a log-log plot of the norm of the *solution norm* $\|\chi\|_1$ vs. the corresponding *residual norm* $\|z - \Psi\chi\|_2$, can graphically display the trade-off between the size of a regularized solution and how well it fits the data. An optimal regularization parameter that minimize the trade-off lies at the corner of such L-curve. For the L-curve method, a solution path that changes monotonically with respect to λ is essential, i.e., $(\Psi^T\Psi)^{-1}$ in (6) needs to be diagonally dominant [26]. However, since Ψ depends on data, one can

not show this is satisfied in our case. Thus, none of these methods are suitable.

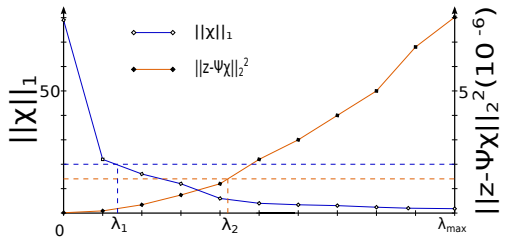

Fig. 3. Illustration of regularization parameter selection

1) Heuristic for selecting λ: The proposed heuristic to choose λ is inspired by the L-curve method. First, we plot both the solution norm and residual norm individually against λ by repeatedly solving Problem (5) for $\lambda \in [0, \lambda_{max}]$. An illustration of these two plots is shown in Figure 3. Second, identify a value λ_1 so that the solution norm is smaller than a user-defined threshold for any $\lambda > \lambda_1$, and then identify λ_2 so that the residual norm is smaller than another user-defined threshold for any $\lambda < \lambda_2$. If $\lambda_2 > \lambda_1$, choose λ to be λ_1. If not, pick another threshold and continue until this condition is met.

IV. Evaluation

The proposed method is evaluated in two ways. In the first case, data generated by a dynamic model of a multi-zone building - a set of coupled ODEs - is used to test the method. The simulation model used to generate data is referred as a "virtual building" in the sequel. In the second case, the proposed algorithm is applied to measurements collected from a real building. In both cases, the building is limited to one with 6 inputs (4 zone cooling powers, ambient temperature, and solar irradiance). A 6th order model is the lowest order that can have different time-constants for each input. Hence we test the method with $\rho = 6$ for both cases.

All numerical results presented in this paper are obtained by using the `cvx` package for solving convex problems in MATLAB© [28].

A. Case 1: Evaluation with Virtual Building Data

1) Virtual building description: The floor plan of the virtual building is shown in Figure 4. A RC (resistor-capacitor) network model is used as the virtual building as shown in Figure 5. RC network model is a common modeling paradigm for building thermal dynamics [29]. The parameters of the model were chosen in the following manner. Reference [30] provides an estimate of the RC network model parameters of a single zone by calibration with data collected from a zone in a building at the University of Florida campus (Pugh Hall). This single zone model has two C's and two R's. We use these parameters for each of the eight small zones in the virtual building. For the two larger zones, the C's are chosen as three times these values and R's as one thirds. Other parameters, such as effective area for solar irradiance, are chosen to be the same as those in [30]. The virtual building is a set of coupled ODEs with 15 states (10 measurable

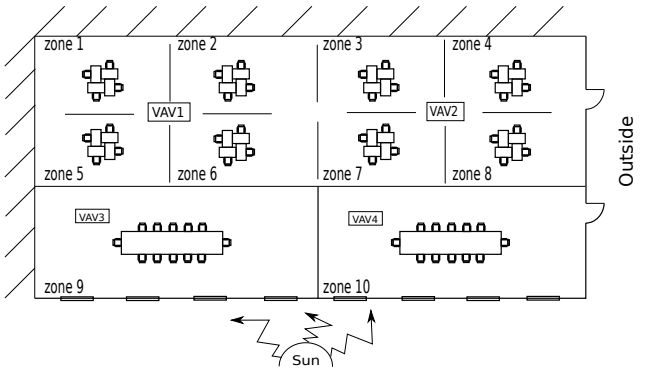

Fig. 4. Floor plan of the virtual building used to generate data: VAV boxes 1 and 2 directly deliver air to each of the four nearby zones while VAV boxes 3 and 4 only deliver to zones 9 and 10, respectively.

temperatures, one for each output node, and 5 unmeasurable temperatures at 5 inter-connection points of output nodes), 6 inputs (4 controllable heat gains from HVAC system, ambient temperature, and solar irradiance) and 10 exogenous disturbances, one for each zone. The input signals are chosen

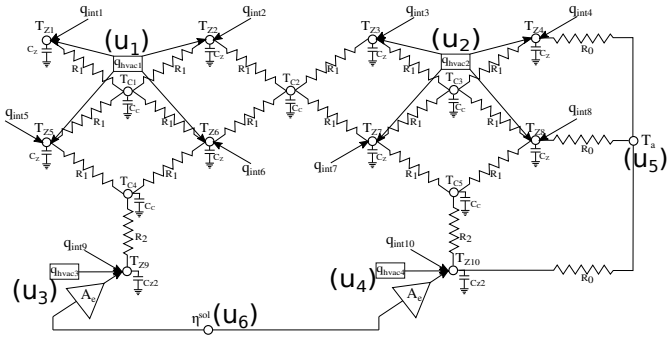

Fig. 5. A schematic of the "RC" network model that constitutes the virtual building. q_{int}'s are the "internal heat gain", i.e., the disturbances. A_e is the effective area.

as follows: the input q_{hvac} is chosen by scaling the measured cooling power from the same zone in Pugh Hall that was used for the single zone model in [30]. This signal is shown in Figure 6 on the top. Ambient temperature is taken from weatherunderground.com, and solar irradiance data is taken from NSRDB: https://nsrdb.nrel.gov/, both for Gainesville, FL. These are shown in Figure 6 (bottom). Disturbances are chosen somewhat arbitrarily, by scaling CO_2 data from Pugh Hall, and are shown in Figure 7. It is important to note that disturbances do not satisfy Assumption 1.

2) Virtual building results:

a) Zone temperature prediction: The transfer matrix and transformed disturbance identified using data from one week (training dataset) are used to predict zone temperatures for the next week (validation dataset). Table I shows the RMS values of the prediction errors for each zone. The maximum RMS value is 2.09°C, which occurs in zone 1, and the minimum RMS value is 0.599°C from zone 8. Figure 8 shows the measured and predicted temperature for these two zones.

b) Plant/network identification: Comparison of identified transfer matrix with the ground truth can be done

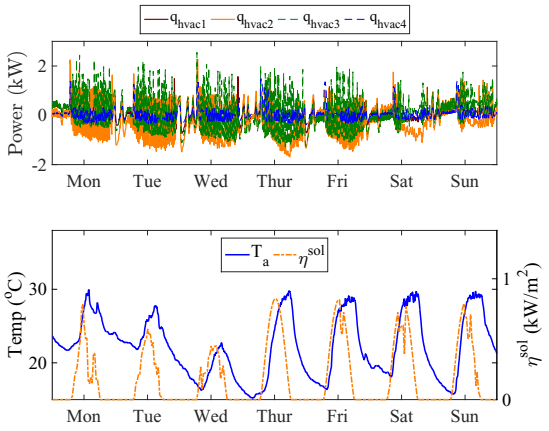

Fig. 6. Training data for algorithm evaluation: q_{hvac}'s are for virtual building evaluation; ambient temperature and solar irradiance are for both virtual building and real building evaluations.

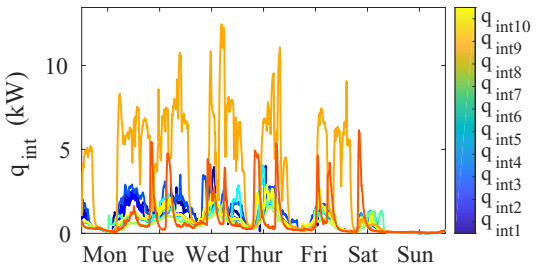

Fig. 7. Virtual building: disturbance data for algorithm evaluation. Notice $w_i \neq k_w^i w_{bldg}$.

in a straightforward manner by comparing the frequency responses. But this method is cumbersome due to the large number of input-output pairs (a total of 60). In addition, it does not help visualizing the sparsity of the identified model $\hat{G}_u$ or the ground truth G_u. Another issue is that the entries of the transfer matrix of neither the virtual building nor the identified model will be strictly zero. There are no zero entries in the matrix G_u due to the interaction among states of the virtual building, which makes every entry of G_u non-zero. As for the identified model $\hat{G}_u$ that is constructed from $\hat{\theta}$, finite-precision numerical calculations cause the entries of $\hat{\theta}$ to be non-zero (albeit extremely small) even when they are

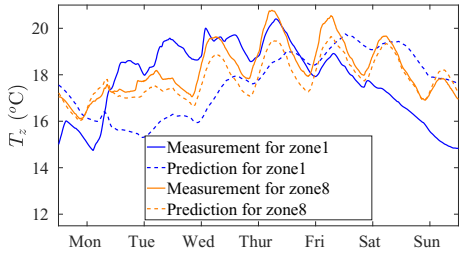

Fig. 8. Evaluation on virtual building: temperature measurements and predictions for zones 4 and 6 (validation dataset).

TABLE I
EVALUATION ON VIRTUAL BUILDING: RMS VALUES OF THE ZONE TEMPERATURE PREDICTION ERROR (°C) (VALIDATION DATASET).

zone 1	zone 2	zone 3	zone 4	zone 5
2.09	1.31	1.48	0.934	1.67
zone 6	zone 7	zone 8	zone 9	zone 10
1.33	0.795	0.599	1.00	1.84

theoretically guaranteed to be 0 (cf. Proposition 1).

In view of the above, a transfer matrix is visualized through the total energy of the impulse response of $G_{p\ell}(z^{-1})$, which is a measure of gain between input-output pair u_ℓ and y_p. Denoting by $g_{p\ell}[k]$ the impulse response, we define

$$E_{G_{p\ell}} := \sum_0^\infty |g_{p\ell}[k]|^2 = \frac{1}{2\pi}\int_0^{2\pi} |G_{pl}(e^{j\Omega})|^2 d\Omega$$

where the equality follows from Perseval's relation [31] and Ω is the angular frequency in radians per sample. When the integral is computed from the identified transfer matrix elements, it is denoted as $E_{\hat{G}_{p\ell}}$. The interaction strengths in the various input-output pairs can be visualized by examining the values of the matrix $E_G := [E_{G_{p\ell}}]$ (or its counterpart $E_{\hat{G}}$).

Figure 9 shows the heat maps of E_G and $E_{\hat{G}}$. Markers with darker color represent higher energy. The ground-truth on input-output interactions are consistent with what we expect from the building layout shown Figure 4. For instance, VAV box 1 supplies air to zones 1,2,5,6, so u_1 should have a strong impact on y_1, y_2, y_5, y_6, while its effect on the other outputs are expected to be weak. Similar trends are also expected for the other inputs from VAV boxes. Zones 4 and 8 are perimeter zones while the rest are interior zones, hence T_a should affect zones 4 and 8 much more than other zones. Zones 9,10 should be strongly impacted by η^{sol} since they have windows directly exposed to the sun. These trends are observed in Figure 9 (ground truth).

Comparing the estimates with the ground truth we see that the proposed algorithm identifies dominant network interactions quite accurately. The estimated transfer functions from exogenous inputs u_5 (T_a) and u_6 (η^{sol}) seem to be less accurate than from the controllable inputs. We believe it is because of lack of adequate excitation in the exogenous inputs: both ambient temperature and solar irradiance is dominated by a single periodic component with a 24-hour period. Among these two, the estimate of transfer functions from input T_a is less accurate than that from input η^{sol} because signal T_a is more correlated to q_{hvac} and disturbance than η^{sol}; see Figure 6.

c) Disturbance: There is no ground truth $\bar{w}_{bldg}[k]$ that can be compared to the estimated $\hat{\bar{w}}_{bldg}$ since Assumptions 1 and 2 are not satisfied in the virtual building. We therefore compare the estimated transformed disturbance, $\hat{\bar{w}}_{bldg}$, with the "total building disturbance" defined as $w_{bldg}[k] := \sum_{i=1}^{i=n_y} w_i[k]$. Figure 10 shows a comparison between w_{bldg} and $\hat{\bar{w}}_{bldg}$, which shows that the signals are highly correlated. In fact, the covariance between these two time series is 0.821. This is quite high since $\hat{\bar{w}}_{bldg}$ is an estimate of a linear transformation of the disturbance, not of the disturbance itself.

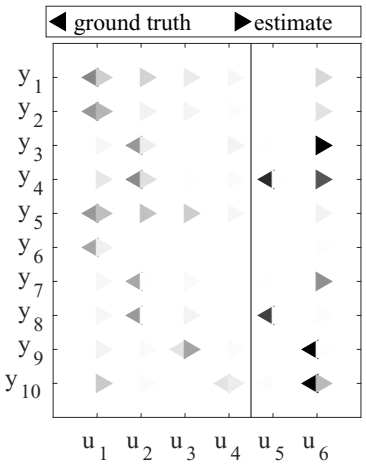

Fig. 9. Evaluation on virtual building: comparison of the heat map between the E_G and $E_{\hat{G}}$: markers shown in darker color represent higher energy, and empty slots correspond to 0. Left from the solid lines are controllable inputs (q_{hvac}) and right from the solid line are exogenous inputs u_5 (T_a) and u_6 (η^{sol}).

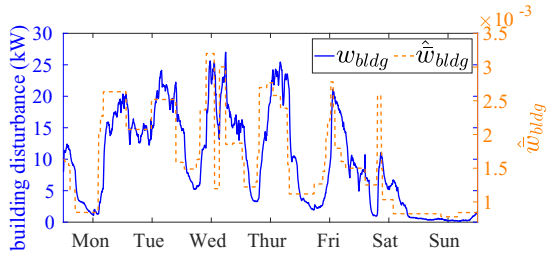

Fig. 10. Evaluation on virtual building: comparison between total building disturbance and estimated transformed disturbance.

B. Case 2: Evaluation with Real Building Data

1) Building description: We now apply the algorithm from data collected from Pugh Hall, a building in the University of Florida campus that houses offices, classrooms and conference rooms. Evaluation of the method with data from a real building is challenging since there is no ground truth to compare with. However, if the zones are widely separated in space, it is reasonable to believe that the control input (cooling power) to one zone should not have any effect on the output (zone temperature) at the other zones. That is, for any two distinct zones i and j, we expect $G_{ij}(z) = 0$. Checking for such a sparse structure of the estimated transfer matrix provides a sanity check.

We collect input-output data from 4 zones that are widely separated, and in fact are located on two different floors of the building. They are shown in Figure 11. Other input components, ambient temperature and solar irradiance, are the same as in virtual building case described previously.

2) Real building results:

a) Transfer matrix: The heat map of $E_{\hat{G}}$ is shown in Figure 12. We see from the figure that the graph $E_{\hat{G}}$ is approximately diagonal for $i, j = 1, \ldots, 4$. This indicates the algorithm correctly identifies the lack of thermal interaction among these four zones. Note that $u_6(\eta^{\text{sol}})$ appears to have a

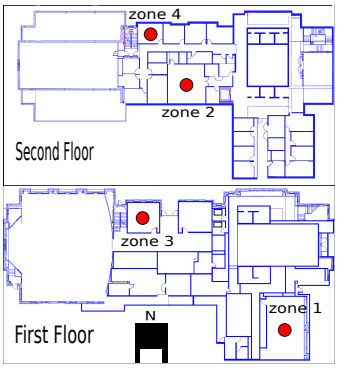

Fig. 11. Floor plan of Pugh Hall. The four zones that are chosen to form the "multi-zone building", on which the method is tested, are shown as red circles.

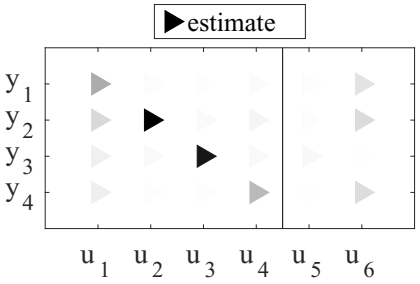

Fig. 12. Evaluation on real building: comparison of the heat map between the $E_{G_{ij}}$ and $E_{\hat{G}_{ij}}$. Left from the solid lines are controllable inputs (q_{hvac}) and right from the solid line are exogenous inputs u_5 (T_a) and u_6 (η^{sol}).

strong effect on the temperature of these zones except for zone 3. This is reasonable because all four zones have windows, but zone 3 is the only zone that has a window facing north which is shadowed by trees. The estimates for input T_a is less accurate, likely because this signal is correlated to q_{hvac}.

b) Disturbance: Since the disturbances are unknown, we compare the estimated transformed disturbance $\hat{w}_{bldg}$ to measurements of CO_2. If the unmeasured disturbance is indeed mostly due to occupants, and since CO_2 concentration is affected strongly by occupancy, we expect the disturbance in a zone to be highly correlated to CO_2 concentration in that zone. For the purpose of comparison with the building-level disturbance, we define the "building-level CO_2 concentration" as the sum of CO_2 level measurements from the four zones of the building: $\overline{CO}_2 := \sum_{i=1}^{4} CO_2^i$. The correlation between the true building-level disturbance (which is the sum of the zone-level disturbances) and building-level CO_2 will be arguably lower than those for each zone.

Figure 13 shows the estimated transformed disturbance $\hat{w}_{bldg}$, along with the building-level CO_2 concentration $\overline{CO}_2$. The covariance between the two is $cov(\overline{CO}_2, \hat{w}_{bldg}) = 0.45$. Although this is not large enough to claim victory, it nonetheless provides a positive sanity check.

c) Zone temperature fit: Figure 14 shows the temperature in zone 2 and the fit from the identified transfer matrix and disturbance. Zone 2 has the worst fit: the RMS values of the prediction errors for zones 1 through 4 are 0.305, 0.616, 0.445, 0.372 °F.

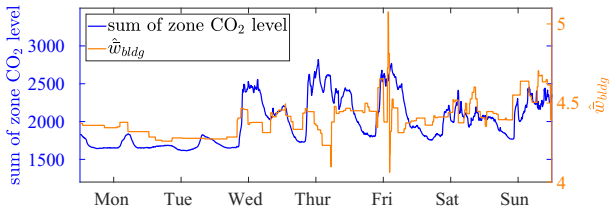

Fig. 13. Evaluation on real building: comparison between building's CO_2 level and estimated transformed disturbance.

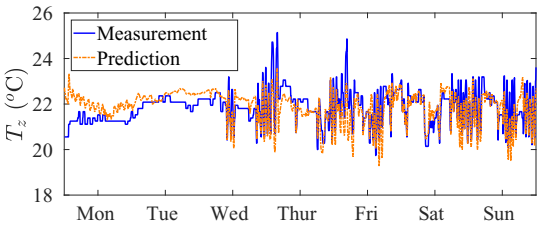

Fig. 14. Evaluation on real building: temperature measurement and prediction for zone 2.

V. Conclusion

We proposed an algorithm to simultaneously identify a MIMO transfer matrix of a multi-zone building, and a transformed version of the building-wide exogenous disturbance, by solving a convex optimization problem. Application of the method to data from a simulation model and data collected from a real building shows promising results.

A limitation of the proposed method is that it only identifies a transformed disturbance, which is related to the disturbance by a non-invertible linear transformation. Additionally, it is assumed that the disturbance at every zone is related to the total building-level disturbance by a constant factor. Extending the method to remove these limitations is a topic of future work. Another avenue for future work is to utilize known information on the network topology.

References

[1] S. Tom, "Managing Energy and Comfort: Don't sacrifice comfort when managing energy," *ASHRAE journal*, vol. 50, no. 6, pp. 18–26, June 2008.

[2] W. J. Fisk, "Health and productivity gains from better indoor environments and their relationship with building energy efficiency," *Annual Review of Energy and the Environment*, vol. 25, no. 1, pp. 537–566, 2000.

[3] United States Energy Information Administration, "Annual energy outlook," April 2018. [Online]. Available: https://www.eia.gov/outlooks/aeo

[4] O. Seppänen, W. J. Fisk, and D. Faulkner, "Control of temperature for health and productivity in offices," *ASHRAE transactions*, vol. 111, pp. 680–686, 2005.

[5] N. P. Sensharma, J. E. Woods, and A. K. Goodwin, "Relationships between the indoor environment and productivity: a literature review," *Ashrae Transactions*, vol. 104, pp. 686–700, 1998.

[6] ASHRAE, "The ASHRAE handbook fundamentals (SI Edition)," 2009.

[7] L. Schellen, M. G. Loomans, M. H. de Wit, B. W. Olesen, and W. van Marken Lichtenbelt, "The influence of local effects on thermal sensation under non-uniform environmental conditions – gender differences in thermophysiology, thermal comfort and productivity during convective and radiant cooling," *Physiology & behavior*, vol. 107, no. 2, pp. 252–261, 2012.

[8] J. M. Penman, "Second order system identification in the thermal response of a working school," *Building and Environment*, vol. 25, no. 2, pp. 105–110, 1990.

[9] S. Wang and X. Xu, "Parameter estimation of internal thermal mass of building dynamic models using genetic algorithm," *Energy conversion and management*, vol. 47, no. 13, pp. 1927–1941, 2006.

[10] D. Kim, J. Cai, K. B. Ariyur, and J. E. Braun, "System identification for building thermal systems under the presence of unmeasured disturbances in closed loop operation: Lumped disturbance modeling approach," *Building and Environment*, vol. 107, pp. 169 – 180, 2016.

[11] T. Zeng, J. Brooks, and P. Barooah, "Simultaneous identification of building dynamic model and disturbance using sparsity-promoting optimization," in *5th International Conference on High Performance Buildings*, July 2018, pp. 1–10.

[12] C. Duarte, K. Van Den Wymelenberg, and C. Rieger, "Revealing occupancy patterns in office buildings through the use of annual occupancy sensor data," Idaho National Laboratory (INL), Tech. Rep., 2013.

[13] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, 2014.

[14] D. Hayden, Y. H. Chang, J. Goncalves, and C. J. Tomlin, "Sparse network identifiability via compressed sensing," *Automatica*, vol. 68, pp. 9 – 17, 2016.

[15] Z. Yue, J. Thunberg, W. Pan, L. Ljung, and J. Gonçalves, "Linear dynamic network reconstruction from heterogeneous datasets," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 10 586–10 591, 2017.

[16] E. Atam and L. Helsen, "Control-oriented thermal modeling of multizone buildings: methods and issues: intelligent control of a building system," *IEEE Control Systems*, vol. 36, no. 3, pp. 86–111, 2016.

[17] S. Goyal, C. Liao, and P. Barooah, "Identification of multi-zone building thermal interaction model from data," in *Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE Conference on*. IEEE, 2011, pp. 181–186.

[18] H. Doddi, S. Talukdar, D. Deka, and M. Salapaka, "Data-driven identification of thermal network of multi-zone building," in *Decision and Control (CDC), 2018 IEEE 57th Annual Conference on*. IEEE, 2018, under review.

[19] D. Kim, J. Cai, and J. E. Braun, "Identification approach to alleviate effects of unmeasured heat gains for MIMO building thermal systems," in *American Control Conference (ACC), 2017*, 2017, pp. 50–55.

[20] Z. Guo, A. R. Coffman, J. Munk, P. Im, and P. Barooah, "Identification of aggregate building thermal dynamic model and unmeasured internal heat load from data," in *IEEE Conference on Decision and Control*, Decemner 2019, accepted.

[21] D. Materassi, G. Innocenti, L. Giarré, and M. Salapaka, "Model identification of a network as compressing sensing," *Systems & Control Letters*, vol. 62, no. 8, pp. 664–672, 2013.

[22] E. Nozari, Y. Zhao, and J. Cortés, "Network identification with latent nodes via autoregressive models," *IEEE Transactions on Control of Network Systems*, vol. 5, no. 2, pp. 722–736, June 2018.

[23] T. Zeng and P. Barooah, "Identification of network dynamics and disturbance for a multi-zone building," in *2nd IFAC Conference on Cyber-Physical and Human Systems (CPHS'18)*, December 2018.

[24] R. Diestel, *Graph Theory*, 3rd ed., ser. Graduate Texts in Mathematics. Springer-Verlag, Heidelberg, 2005, vol. 173.

[25] A. Ali and R. J. Tibshirani, "The generalized lasso problem and uniqueness," *arXiv preprint arXiv:1805.07682*, 2018.

[26] J. Duan, C. Soussen, D. Brie, J. Idier, M. Wan, and Y.-P. Wang, "Generalized lasso with under-determined regularization matrices," *Signal processing*, vol. 127, pp. 239–246, 2016.

[27] P. C. Hansen, "Analysis of discrete ill-posed problems by means of the l-curve," *SIAM review*, vol. 34, no. 4, pp. 561–580, 1992.

[28] M. Grant and S. Boyd, "CVX: Matlab software for disciplined convex programming, version 1.21," http://cvxr.com/cvx, Feb. 2011.

[29] H. Madsen and J. Holst, "Estimation of continuous-time models for the heat dynamics of a building," *Energy and Buildings*, vol. 22, no. 1, pp. 67 – 79, 1995.

[30] A. R. Coffman and P. Barooah, "Simultaneous identification of dynamic model and occupant-induced disturbance for commercial buildings," *Building and Environment*, vol. 128, pp. 153–160, 2018.

[31] A. V. Oppenheim, A. Willsky, and S. Nawab, "Signals and systems 2nd ed," *New Jersey: Prentice Hall*, 1997.