MITSUBISHI ELECTRIC RESEARCH LABORATORIES
https://www.merl.com

H-Infinity Loop-Shaped Model Predictive Control with HVAC Application

Bortoff, Scott A.; Schwerdtner, Paul; Danielson, Claus; Di Cairano, Stefano; Burns, Daniel J.

TR2022-028 March 23, 2022

Abstract

We formulate a Model Predictive Control (MPC) for linear time-invariant systems based on H-infinity loop-shaping. The design results in a closed-loop system that includes a state estimator and attains an optimized stability margin. Input and output weights are designed in the frequency domain to satisfy steady-state and transient performance requirements, in lieu of standard MPC plant model augmentations. The H-infinity loopshaping synthesis results in an observer-based state feedback structure. An inverse optimal control problem is solved to construct the MPC cost function, so that the control input computed by MPC is equal to the H-infinity control input when the constraints are inactive. The MPC inherits the closed-loop performance and stability margin of the loop-shaped design when constraints are inactive. We apply the methodology to a multizone heat pump, and validate the results in simulations and laboratory experiments. The design rejects constant unmeasured disturbances, tracks constant references with zero steady-state error, meets transient performance requirements, provides an excellent stability margin, and enforces input and output constraints.

IEEE Transactions on Control Systems Technology 2022

Mitsubishi Electric Research Laboratories, Inc.
201 Broadway, Cambridge, Massachusetts 02139

H-Infinity Loop-Shaped Model Predictive Control with HVAC Application

Scott A. Bortoff[1*], Paul Schwerdtner[2], Claus Danielson[3], Stefano Di Cairano[1] and Daniel J. Burns[1]

***Abstract*—We formulate a Model Predictive Control (MPC) for linear time-invariant systems based on H-infinity loop-shaping. The design results in a closed-loop system that includes a state estimator and attains an optimized stability margin. Input and output weights are designed in the frequency domain to satisfy steady-state and transient performance requirements, in lieu of standard MPC plant model augmentations. The H-infinity loop-shaping synthesis results in an observer-based state feedback structure. An inverse optimal control problem is solved to construct the MPC cost function, so that the control input computed by MPC is equal to the H-infinity control input when the constraints are inactive. The MPC inherits the closed-loop performance and stability margin of the loop-shaped design when constraints are inactive. We apply the methodology to a multi-zone heat pump, and validate the results in simulations and laboratory experiments. The design rejects constant unmeasured disturbances, tracks constant references with zero steady-state error, meets transient performance requirements, provides an excellent stability margin, and enforces input and output constraints.**

I. Introduction

Because Heating, Ventilation and Air Conditioning (HVAC) systems are the largest consumer of electric power in commercial and residential buildings [1], [2], their energy efficient operation is of growing importance. Of course, comfortable and efficient operation of HVAC equipment is dependent on proper design and implementation of the control system. Generally speaking, HVAC systems are multi-input multi-output, are subject to actuator and process variable constraints, exhibit nonlinearity of different types, and tend to use a minimal number of low quality sensors, primarily for economic reasons. Further, because of the diverse nature of building applications and the fact that the dynamics of the equipment and the building are coupled, plant models have significant uncertainty.

Building HVAC systems may be classified depending on the building type, size, and location. A large commercial building typically will have a so-called "built-up" HVAC system, consisting of a chilled water plant (chillers, water distribution, cooling towers), Variable Air Volume (VAV) air handlers (heat exchangers, valves, ducts, fans), and ducting with controlled dampers for distribution of conditioned air among its zones. For smaller commercial buildings, the HVAC system may be based on an integrated unitary rooftop unit containing an air-source vapor compression system and air handlers which distribute conditioned air via ducts to individual building zones. Large residential apartment buildings are similar, while smaller, single-family residential buildings are typically conditioned by an air-source vapor compression system with indoor heat exchangers located in a supply air duct or directly in one or more rooms. Small residential applications lack active ventilation (controlled introduction of outside air) except in extreme climate zones.

Control architectures have a similar classification. Large commercial building HVAC control systems are organized in a hierarchy that might include a Building Management System (BMS) at the highest level, system-level controls at mid-levels (such as for a chilled water plant), and product-inserted controllers within individual pieces of equipment at the lowest level. For these systems, control algorithms at the higher levels are custom designed and commissioned for each individual application, while the product-inserted controllers at the lowest level are designed to be programmed at commissioning time, and interface to higher levels using industry standards. At the other end of the spectrum, a residential air conditioner might be "controlled" by a one or more thermostats, with no external information connection. The control functionality for these products is embedded in the factory-built equipment, and is turn-key, requiring little to no commissioning at installation time.

In this paper we consider the control of a specific factory-built HVAC product belonging to the latter class: A multi-zone, Variable Refrigerant Flow (VRF) air-source heat pump intended for residential or small commercial building application. Control for this product is challenging because it must provide some building-level function (zone temperature control) in addition to regulating and/or enforcing constraints on a set of internal process variables, over a broad range of operating conditions, and for a diverse range of potential applications, with little to no commissioning at installation time. The control requirements include enforcement of various process variable constraints and actuator constraints, regulation of other internal process variables, and offset-free room temperature setpoint regulation and disturbance rejection, with specified transient response characteristics, all while maximizing energy efficiency.

The baseline production controller for this product is a proprietary, multivariable PID, with anti-windup, selector and override logic to enforce actuator and process variable constraints. This is augmented with a layer of ad hoc protection logic that is intended to prevent the equipment from operating in potentially damaging conditions, or to prevent it from

[1]Mitsubishi Electric Research Laboratories, Cambridge, MA, 02139 USA (e-mail: {bortoff,dicairano,burns}@merl.com)

[2]Technical University of Berlin, Straße des 17, Juni 135, 10623, Berlin, Germany, (e-mail: schwerdt@math.tu-berlin.de)

[3]Department of Mechanical Engineering, University of New Mexico Albuquerque, NM, USA (e-mail: cdanielson@unm.edu)

*Corresponding author

exhibiting undesirable behaviors. Although the architecture is based upon well-founded theory [3], it suffers from two drawbacks. First it is inflexible and difficult to adapt as the product evolves. If a new constraint must be added to the design during product development, or over the product lifecycle, or constraint priorities change, then the entire architecture may need to be redesigned. Second, it lacks a unified mathematical framework for analysis and design. Protection logic and its interaction with the conventional feedback is especially difficult to analyze beyond time-domain simulation. What is needed is a standard and flexible design methodology that provides rigorous constraint enforcement and realization of the intended function of protection logic, and that can adapt as the product evolves over its lifecycle.

Model Predictive Control (MPC) is a potential solution. MPC for HVAC systems, ranging from the equipment level to the building level, has received considerable attention for more than a decade [4]-[22], with [21] and [22] providing thorough literature reviews. Much of the research is intended to be applied to large commercial buildings at the BMS level of the control hierarchy, where the objective of minimizing energy consumption, energy cost, or greenhouse gas production, etc., while enforcing constraints associated with equipment, occupant comfort and ventilation requirements, achieved by actuating setpoints at lower levels of the control hierarchy, such as supply air temperatures or zone temperature setpoints, is well-suited to the MPC paradigm [22].

However, MPC at the building level is fundamentally different than MPC at the equipment level. At the building level, the cost function is generally directly meaningful, such as the cost of energy, and its minimization is the primary objective. Building models are custom built for each application. Time scales span from minutes to hours. Robustness of the feedback loop that is closed by the actions of receding horizon control is seldom a concern. On the other hand, MPC at the equipment level is primarily concerned with enforcing constraints, and a cost function is formulated to meet other objectives such as setpoint tracking and disturbance rejection. Minimizing the cost is only a means to an end. Equipment models are central, with timescales ranging from seconds to minutes. A single MPC, derived from a single model, is embedded in a factory-installed controller, and needs to operate in a broad range of conditions and in a diverse range of applications. As such, robustness of the feedback loop that is closed by the actions of receding horizon control is of primary concern.

At the equipment level, MPC has seen some success in laboratory and simulation demonstrations. An MPC was designed for a multi-evaporator, water source heat pump in [6], but the plant included refrigerant pressure sensors so that the thermodynamic variables used by MPC were effectively measured, and no state estimator was required. An offset-free MPC for a heat pump was considered in [16]. The MPC augmented the model to estimate disturbances, which could compensate some plant-model mismatch, and the design incorporated a Luenberger observer to estimate the state. The authors manually tuned the gains by trial-and-error, modified the cost to include damping, and noted oscillatory behavior for aggressive tuning. An offset-free MPC design for a multi-zone heat pump in cooling mode was also considered in [19], where offset free tracking and disturbance rejection was achieved after laborious manual tuning of the estimator gain and weights.

Although the state-of-the-art has progressed considerably, the control designer faces multiple challenges in the design and implementation of an MPC for factory-produced HVAC equipment. First of all, MPC is a state-based feedback control, and the full state is not usually measured in this type of equipment. In fact, for many products, only a limited number of low precision temperature sensors is used. Refrigerant pressure sensors or flow meters, which may be available in the laboratory e.g. [6], are not economically viable for use in most products. This implies that a state estimator of some type must be used, which is typically designed using the separation principle [22]. However, such an estimated state when used in an MPC provides no guarantee of closed-loop stability margin, and this fact, combined with the presence of plant model uncertainty, can lead to our first problem: Poor closed-loop performance, often manifested by oscillatory behavior.

In the literature, oscillatory behavior of MPC is often attributed to model mismatch or model uncertainty e.g., [14], [16], [22]. However, this is only half of the story: An MPC that uses a state estimator is a feedback loop that may possess a high sensitivity to model uncertainty. The root cause of oscillations is not only model mismatch, but also high sensitivity of this feedback loop to model uncertainty. High sensitivity can be caused by independently designing the estimator gain and cost function weight, i.e., applying the separation principle, without considering the stability margin of the closed loop.

The second related problem is the practical one of tuning a large number of parameters that are commonly associated with MPC (i.e., the cost function weights) and the state estimator gain. In our laboratory experience, the number of scalar parameters that must be tuned for a moderately-sized vapor compression system may exceed 1000. Setting most to zero for convenience, which is common in the literature, exasperates the first problem.

This paper addresses both problems with a design methodology that we call Loop-Shaped Model Predictive Control (LSMPC), in which the MPC design is based not on the separation principle and LQR/LQG, but rather on robust $\mathcal{H}_\infty$ loop-shaping [23], [24]. This results in a control system with an optimized stability margin, and provides a rigorous way to *design* all of the plant augmentations in both the prediction model and the state estimator that are typically used to meet performance requirements for closed-loop bandwidth, transient response, disturbance rejection, and reference tracking. The methodology provides values for most of the design parameters of the MPC and state estimator. It consists of the following steps.

1) Design input and output weights for the plant to meet closed-loop performance specifications. This step *replaces* standard MPC plant augmentations.
2) Synthesize the $\mathcal{H}_\infty$ loop-shaping compensator, which has an observer-based state feedback structure like LQG, but

provides an optimized stability margin.[1]

3) Construct the MPC prediction model and cost function weights by solving an inverse-optimal control problem for the state feedback from Step 2.
4) Use the $\mathcal{H}_\infty$ state estimator to initialize the MPC prediction model at each time step. An additional model augmentation is needed to enforce output constraints if any are present in the problem.

The closed-loop LSMPC inherits the reference tracking, disturbance rejection, transient response, and stability margin of the $\mathcal{H}_\infty$ loop-shaped controller in the region where constraints are inactive. The state estimator gain is computed directly by $\mathcal{H}_\infty$ synthesis and therefore does not require additional tuning. The key synthesis step is to compute solutions to two decoupled Riccati equations, or alternatively to solve a Linear Matrix Inequality (LMI). The latter is straightforward with available tools such as SeDuMi [25] and YALMIP [26]. LSMPC can also enforce constraints on inputs and outputs.

This paper extends some previously published results. Rowe and Maciejowski [27] apply $\mathcal{H}_\infty$ loop-shaping and inverse optimality to compute the state estimator and MPC cost. However, the authors do not consider output constraints, and they augment models of the reference and disturbance into the plant model, implying that the disturbance must be measured or estimated. In this paper, plant augmentations are the result of $\mathcal{H}_\infty$ loop-shaping, we consider input and output constraints, and we consider unmeasured disturbances. Note that some standard plant augmentations, such as adding a constant offset vector to the output and estimating its value in order to achieve offset-free tracking, make the augmented plant uncontrollable, which violates a sufficient condition for the $\mathcal{H}_\infty$ loop-shaping compensator. Maciejowski [28] and Di Cairano and Bemporad [29] consider the problem of controller matching, or finding a matching cost function, assuming an output feedback controller or state feedback controller, respectively, is given *a priori*. In this paper we *design* both the state feedback and state estimator to meet performance requirements and also to optimize a stability margin. This paper improves our prior formulation of LSMPC [30] by ensuring controller anti-windup and also including experimental validation.

LSMPC may have some minor disadvantages. We have observed that for some plants, inverse optimality may result in a numerically ill-conditioned cost function which can make the real-time optimization problem ill-conditioned, although this was not an issue for the heat pump application. It is possible that this could be addressed by casting the key design equations as an LMI. Also, for some problems, the prediction model can have a larger dimension when compared to a standard MPC design. This is because any constrained output variables require an additional prediction model augmentation. However, in the special cases that the MPC need not enforce output constraints, or the problem does not require rejection of an unmeasured disturbance, the prediction model does not require such additional augmentation and could be of lower dimension compared to a standard MPC design.

[1]In practice the resulting compensator possesses a slightly sub-optimal stability margin.

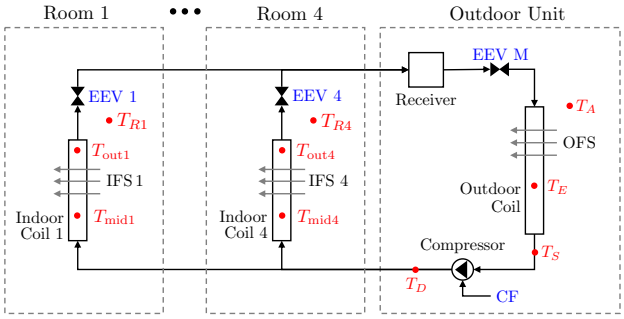

Fig. 1. Vapor compression system showing temperature sensors (red) and control actuators (blue) [31]. Refrigerant flows clockwise in heating mode and counterclockwise in cooling mode, changed by a valve not shown.

Throughout this paper, calligraphic font is used to represent a linear system, e.g. $\mathcal{P}$, and italics roman font is used to represent matrices, e.g. A_s. The positive feedback convention is used since it is standard in $\mathcal{H}_\infty$ loop-shaping literature. In Section II, the multi-zone heat pump is described and control system requirements are listed. In Section IV, the design considerations for applying MPC to the multi-zone heat pump are discussed, and we formally state the problem solved in this paper. In Section V we design the $\mathcal{H}_\infty$ loop-shaping control for the heat pump, and realize it as an MPC in Section VI. The proposition of equivalence between the MPC and the $\mathcal{H}_\infty$ control when constraints are inactive is stated formally in Section VI-E. Simulation and experimental results are presented in Section VII. We suggest some extensions in Section VIII.

II. Heat Pump Description & Requirements

The specific air-source multi-zone vapor compression system we consider is diagrammed in Fig. 1, consisting of one outdoor unit and four indoor units. The outdoor unit contains a receiver, an electronic expansion valve (EEV), denoted EEV M, a heat exchange coil, a compressor and an outdoor fan. The indoor units each contain a heat exchange coil, an EEV and an indoor fan. The system can operate in either cooling mode, where it moves heat from the indoor units to the outdoor unit, or in heating mode, where it runs in reverse. When operating in heating mode, it is commonly referred to as a heat pump. We consider only heating mode in this paper.

The system moves heat from the colder outdoor air to the warmer indoor air as follows. The compressor compresses the refrigerant to hot, super-heated gas, which is distributed to the indoor coils where it condenses, releasing heat. Neglecting the small pressure drop in the distribution pipes, the condensing pressure, and therefore the condensing temperature, is identical in each indoor unit. Therefore, to reject asymmetric heat load disturbances in each room, the EEVs are controlled in a way that modulates the refrigerant mass flow rate, and therefore modifies the amount of refrigerant subcooling (the difference between the condensing temperature and the temperature of the liquid exiting the heat exchanger), that occurs in each indoor heat exchanger, with less mass flow rate resulting in more subcooling, resulting in a lower amount of heat flux [31].

After expanding through an indoor EEV i, $1 \leq i \leq 4$, the refrigerant returns to the outdoor unit, where it is expanded a second time through the common EEV M. The resulting cold refrigerant is typically two-phase on entry into the outdoor coil, where it evaporates completely, absorbing heat from the outdoor air and completing the cycle.

The system has seven control actuators: The compressor frequency CF, the commanded settings for each EEV i, $1 \leq i \leq 4$ and EEV M, and the outdoor unit fan speed OFS. The fan speed for each indoor unit, denoted IFS i, $1 \leq i \leq 4$, is set by the occupant and is considered to be a measured disturbance, assumed constant in this paper. Each zone is subject to an unmeasured heat disturbance Q_i, generated by the occupants, surrounding environment, etc. Note that $Q_i < 0$ in heating mode. The entire system is subject to the outdoor air temperature disturbance T_A. The 16 measurements are: The four room temperatures T_{Ri}, the eight condenser mid-point and outlet temperatures $T_{\text{mid}i}$ and $T_{\text{out}i}$, $1 \leq i \leq 4$, the evaporator temperature T_E, the suction temperature T_S, the compressor discharge temperature T_D,and the outdoor air temperature T_A. The condensing temperature $T_C = T_{\text{mid}i}$, $1 \leq i \leq 4$, is identical in each indoor unit because they are all at the same pressure.

The requirements for the closed-loop system are as follows.

A) Exponential stability of the equilibrium for constant references and disturbances;
B) Offset-free tracking of constant room temperature setpoints, if possible;
C) Reject constant, unmeasured heat load disturbances and the outside air temperature, if possible;
D) Track a desired compressor discharge temperature with zero steady-state error, if possible;
E) Track a desired evaporator temperature with zero steady-state error, if possible;
F) Achieve a closed-loop bandwidth ω_b for room temperature setpoint tracking, if possible;
G) Enforce hard constraints on all control inputs; and
H) Enforce constraints on the following outputs:
   a) A minimum subcooling temperature for $1 \leq i \leq 4$,
   $$T_{SCi} = T_{\text{mid}i} - T_{\text{out}i} > T_{SCmin};$$
   b) A maximum compressor discharge temperature,
   $$T_D < T_{Dmax};$$
   c) A maximum condensing temperature,
   $$T_C < T_{Cmax};$$
   d) A minimum discharge temperature super-heat,
   $$T_{DSH} = T_D - T_{\text{mid}i} > T_{DSHmin}.$$

In Requirement A, *exponential stability* means that the closed-loop system, when subject to constant references and disturbances, possesses a unique equilibrium solution that is locally exponentially stable. In Requirements B-F, *if possible* means when permitted by the system constraints. This implies that the constraint enforcement Requirements G and H are higher priority than the tracking and disturbance rejection requirements B-F. Requirement D is used to achieve energy efficient operation, as described in Section V-A. The input constraints in Requirement G are all hard, meaning there are actuator limits that cannot be exceeded. This has implications on integrator anti-windup that are discussed later. On the other hand, the output constraints in Requirement H are soft, meaning that they may be exceeded by small amounts in transient conditions for short periods of time, where the quantifiers "small" and "short" are determined by proprietary engineering judgement. However, during normal operation, one or more of the input or output constraints may be at or near its limit for extended periods of time, i.e., the system must operate stably at or near the boundary of its feasible region in the steady-state, for any constant disturbance and reference.

III. Dynamic Model

A nonlinear dynamic model of the heat pump was constructed in the Modelica modeling language [32], [33]. A complete description of the full nonlinear model is beyond our scope, and we refer the reader to [34], [35], which also report results of transient model validation experiments. Modelica is a declarative, acausal, equation-based language with a differential algebraic model of computation. (In contrast, Matlab Simulink has a causal, signal-flow model of computation.) This allows for equations to be transcribed into the language through declarations and mathematical statements. In addition, Modelica is component-oriented, with object-oriented features for organization.

As such, the dynamic model was organized by components, with a set of declared differential and algebraic equations for each component. The components were connected together to equate potential, flow and stream variables at each component boundary, resulting in a large set of nonlinear differential-algebraic equations. The Modelica compiler (Dymola) then manipulated these equations, sorting them for computational efficiency, choosing state variables and assigning causality, producing an efficient ordinary differential equation (ODE) representation.

The variable-speed compressor was modeled with a set of coupled, nonlinear algebraic equations that relate the compressor speed, power consumption, refrigerant mass flow rate, inlet pressure, enthalpy and density, and outlet pressure and enthalpy to one another. This model contained a number of coefficients that were fit to laboratory data. The electric expansion valves were modeled algebraically as $\dot{m} = C_v\sqrt{\rho_{in}\Delta p}$, where $\dot{m}$ is the refrigerant mass flow rate, ρ_{in} is the refrigerant inlet density, Δp is the pressure drop across the valve, and C_v depends on the valve opening command and was fit to laboratory data. The receiver was modeled as a lumped, adiabatic control volume with a single inlet and outlet, negligible pressure drop, and assuming an ideal phase separation and thermodynamic equilibrium between the phases.

Three-dimensional, finite volume models of heat and fluid flow were used for the tube-and-fin heat exchangers. Each heat exchanger was divided into a number of segments, one per tube, and each segment was divided into three sections:

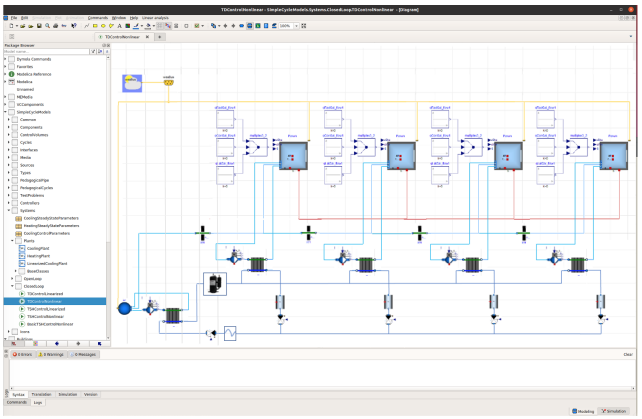

Fig. 2. Dymola integrated development environment for Modelica, showing an iconic view of the four-zone heat pump model with four-room building (right) and component libraries (left).

The refrigerant stream, the finned walls, and the air stream. The refrigerant stream was described by a one-dimensional flow with fluid properties varying only in the direction of flow. Under assumptions described in [34], [35], the conservation of mass, momentum and energy for each segment was modeled as

$$\frac{\partial(\rho A)}{\partial t} + \frac{\partial(\rho A v)}{\partial x} = 0$$

$$\frac{\partial(\rho v A)}{\partial t} + \frac{\partial(\rho v^2 A)}{\partial x} = -A\frac{\partial P}{\partial x} - F_f$$

$$\frac{\partial(\rho u A)}{\partial t} + \frac{\partial(\rho v h A)}{\partial x} = vA\frac{\partial P}{\partial x} + vF_f + \frac{\partial Q}{\partial x},$$

where ρ is the density, A is the cross-sectional area of the flow, v is the velocity, P is the pressure, F_f is the frictional pressure drop, u is the specific internal energy, h is the specific enthalpy, and Q is the heat flow rate into or out of the fluid. For the air stream, the energy balance equation

$$\dot{m}_a c_{p,a}\frac{dT_a}{dy}\Delta y = \alpha_a\left(A_{0,t} + \eta_{\text{fin}}A_{0,\text{fin}}\right)\left(T_w - T_a\right)$$

was used, where $\dot{m}_a$ is the airflow, T_a is the air temperature, T_w is the wall temperature, $c_{p,q}$ is a heat capacity. Finally, the energy balance equation for tube and fin heat conduction was

$$\left(M_t c_{p,t} + M_{\text{fin}}c_{p,\text{fin}}\right)\frac{dT_w}{dt} = q_r + q_a$$

$$q_r = \alpha_r A\left(T_r - T_w\right)$$

$$q_a = \dot{m}_a\left(c_{p,q}(T_{a,\text{in}} - T_{a,\text{out}})\right),$$

where M_t is the tube wall mass, M_{fin} is the fin mass, q_r is the heat flux from the refrigerant, q_a is the heat flux to the air, T_w is the tube wall temperature, $T_{a,\text{in}}$ is the inlet air temperature, and $T_{a,\text{in}}$ is the outlet air temperature. This model allows for complex circuiting of the tubes and also for split-flow heat exchangers which are common features. The balance equations were augmented with a set of empirical closure relations describing the single- and two-phase heat transfer coefficients and frictional pressure drops for the fluid on both the refrigerant and air sides. These equations were discretized using the Reynolds transport theorem using a staggered-grid approach along their respective flow directions, and transcribed into Modelica. The system model was coupled to a four-room building model constructed from the Modelica Buildings Library [36], which captures convective, conductive and radiative heat transfer to and from the environment. An iconic view of the model is shown in Fig. 2.

The complete ODE model produced by Dymola consisted of 739 differential equations and states, with a time scale spanning from milliseconds to about one day. These were linearized at a nominal steady-state operating condition, and inputs and outputs scaled so that the steady-state gain from control inputs and disturbances to measured outputs was $O(1)$. (Specifically, the temperature outputs were all scaled by 1, while the disturbance inputs, compressor speed and valve inputs were scaled so the steady-state gains to room temperatures was 1, and the gain from the common EEV to discharge temperature was 1.) The linearized model was then reduced by balanced truncation and singular perturbation. The resulting linear model was discretized, giving the 24-dimensional, scaled, minimal, strictly proper, stable discrete-time linear model

$$x(k+1) = Ax(k) + Bu(k) + B_d q(k) \tag{1a}$$

$$y(k) = Cx(k) \tag{1b}$$

$$v(k) = Fx(k) + Gu(k), \tag{1c}$$

where $x(k) \in \mathbb{R}^n$ is the state, $u(k) \in \mathbb{R}^m$ is the scaled control input, $q(k) \in \mathbb{R}^d$ is the scaled unmeasured heat disturbance input, $y(k) \in \mathbb{R}^p$ is the scaled measured output, $v(k) \in \mathbb{R}^{l+2m}$ is the scaled constrained output (including the l-dimensional constrained output, and the $2m$-dimensional constrained input, accounting for upper and lower limits on all of the actuators). The measurement y was arranged as $y = [y_1^T \;\; y_2^T]^T$, where

$$y_1 = [T_{R1} \ldots T_{R4} \; T_D \; T_E]^T \tag{2a}$$

$$y_2 = [T_C \; T_S \; T_{SC1} \ldots T_{SC4}]^T, \tag{2b}$$

so that y_1 contained the *regulated* outputs (those which have tracking requirements) and y_2 contained all other measurements. The input vector u was arranged as

$$u = [\text{OFS} \;\; \text{CF} \;\; \text{EEV 1} \ldots \text{EEV 4} \;\; \text{EEV M}]^T. \tag{3}$$

Denote $\mathcal{P}_u$ as the transfer function from u to y_1, and $\mathcal{P}_q$ as the transfer function from q to y_1. Note that all of the elements of v are elements of y or u, or are computed from elements of y. In other words, the constrained output v is measured. In this representation, u, q, y and v are deviations from a prescribed steady-state operating condition, and are scaled such that they are $O(1)$ in nominal operation.

Although the full model is nonlinear and extremely numerically stiff, the reduced-order linear model (1) is remarkably accurate [35], especially for medium frequencies around crossover, where higher fidelity is important for robust control design. The low-frequency gain does vary as a function of operating condition, as also noted in [37], especially at low compressor frequencies, and probably necessitates gain scheduling, but this is left for future research.

IV. MPC Design Considerations

MPC is an attractive control methodology because of the the totality of requirements, especially Requirement H. MPC computes the control input by solving a constrained finite-time optimal control (CFTOC) problem:

$$\min_{u(k)} x_p^T(N)Px_p(N) + \sum_{k=1}^{N-1} z^T(k)Qz(k) + u^T(k)Ru(k) \tag{4a}$$

$$\text{s.t.} \quad x_p(k+1) = A_p x_p(k) + B_p u(k) \tag{4b}$$

$$v(k) = F_p x_p(k) + G_p u(k) \tag{4c}$$

$$v(k) \in \mathcal{V} \tag{4d}$$

$$z(k) = E_p x_p(k) \tag{4e}$$

$$\text{and} \quad x_p(0) = \widehat{x}_p(k)\,, \tag{4f}$$

where the subscript "p" denotes "prediction model," $x_p(k) \in \mathbb{R}^{n_p}$ and $v(k) \in \mathbb{R}^{l+2m}$ are the predicted augmented state and constrained output, respectively, resulting from the control input $u(k) \in \mathbb{R}^m$, $z(k) \in \mathbb{R}^s$ is the performance output, with $Q \geq 0$, $R > 0$, and E_p tuned to satisfy Requirements A-F, N is the prediction horizon, and the constraint set

$$\mathcal{V} = \{v \in \mathbb{R}^{l+2m} : \; v_{min} \leq v \leq v_{max}\}$$

is derived from Requirements G and H. The prediction model (4b) includes the nominal plant dynamics (1), and also includes additional states to predict the effect of the unmeasured disturbance q, as well as shaping filters and/or reformulations such as incremental inputs, all used to shape the closed-loop dynamics to meet Requirements A-F. Let $U_k^* = [u^*(0) \dots u^*(N-1)]^T$ be the solution of (4) at time k. Then MPC control input is

$$u(k) = u^*(0).$$

Because the full state is not measured, the initial state $\widehat{x}_p(k)$ in (4f) is computed with a state observer of the form

$$\widehat{x}_e(k+1) = A_e\widehat{x}_e(k) + B_e u(k) + H_e\left(\widehat{y}(k) - y(k) + r(k)\right) \tag{5a}$$

$$\widehat{y}(k) = C_e\widehat{x}_e(k) \tag{5b}$$

where $\widehat{x}_e(k)$ is the augmented observer state that includes the estimate $\widehat{x}_p(k)$ and possibly an estimate of the unmeasured disturbance, $\widehat{q}(k)$, or its affect on the output $y(k)$ in order to achieve offset-free tracking for constant disturbances, and $r(k) \in \mathbb{R}^m$ is the setpoint reference for the regulated output y_1. Although r may be included in other ways, this formulation allows the MPC and estimator to be designed as a regulation problem, but will still satisfy tracking requirements. This is elaborated upon in Section VI. Some of the components of $x_p(0)$ in (4f) correspond to the shaping filters, whose values are known to the controller and do not require estimation. It is known that this type of plant and observer augmentation, and separated controller - observer design will result in closed-loop offset-free tracking of the MPC, under some assumptions [38], [39].

However, using an estimated state in a full-state feedback control law such as Linear Quadratic Gaussian (LQG) provides no guarantee of any closed-loop stability margin [40]. This result also holds for our MPC formulation, because the cost (4a) is quadratic and (5) is a Luenberger observer designed using the separation principle. Thus, the designer encounters a significant problem in attempting to apply this formulation of MPC to the multi-zone heat pump: How to compute the matrix parameters Q, R, E_p and H_e, and also how to design the augmentations and shaping filters in both the prediction model and the state observer, that satisfy Requirements A-F *and also provide a meaningful closed-loop stability margin*? This is important because, for poor choices of Q, R, E_p and H_e, even a small amount of model uncertainty can result in poor closed-loop performance or instability. The problem also has a practical aspect: For our multi-zone heat pump model, with $m = 7$, $n = 25$ and $p = 16$, there are more than 1000 scalar parameters to be tuned. Stated formally:

Problem 1: MPC Synthesis. Compute plant model augmentations, shaping filters, values for $Q \geq 0$, $R > 0$ and E_p in the cost function (4a) and performance function (4d), respectively, and a value for the observer gain H_e in (5) that satisfy Requirements A-F, and maximize a stability margin.

We solve Problem 1 by designing *all* MPC plant model augmentations and shaping filters to meet Requirements A-F using $\mathcal{H}_\infty$ loop-shaping, and then use the $\mathcal{H}_\infty$ design to synthesize an MPC that meets requirements G and H, and that is equivalent to the $\mathcal{H}_\infty$ regulator when the constraints are inactive along the prediction horizon. In the next section, we present the $\mathcal{H}_\infty$ loop-shaping design procedure for the heat pump, and in the following section, we realize it as an MPC.

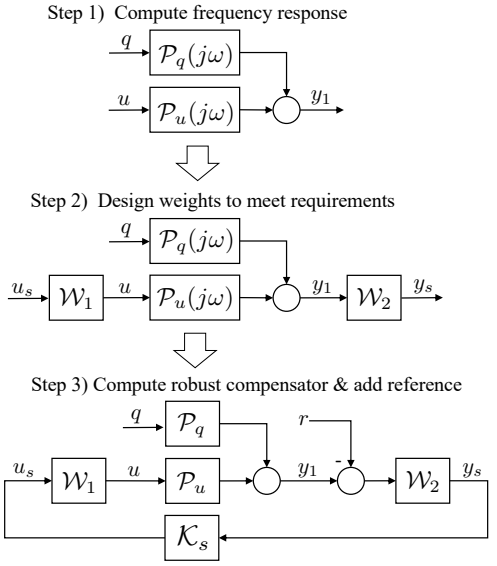

Fig. 3. Design steps for $\mathcal{H}_\infty$ loop-shaping. 1) Compute the frequency response of $\mathcal{P}$. 2) Design weights $\mathcal{W}_1$ and $\mathcal{W}_2$ to shape $|\mathcal{P}_s(j\omega)| = |\mathcal{W}_2(j\omega)\mathcal{P}_u(j\omega)\mathcal{W}_1(j\omega)|$. 3) Compute the robustifying compensator $\mathcal{K}_s$.

V. $\mathcal{H}_\infty$ Loop-Shaped Control

Following [23], [43], [44], $\mathcal{H}_\infty$ loop-shaping proceeds by computing the frequency response for the system $y_1 = \mathcal{P}_u u + \mathcal{P}_q q$, translating the steady-state tracking, disturbance rejection

and transient response requirements into the frequency domain, and designing an input weight (filter) $\mathcal{W}_1$ and an output weight (filter) $\mathcal{W}_2$, in order shape the frequency response of the compensated open-loop system $\mathcal{P}_s = \mathcal{W}_2\mathcal{P}_u\mathcal{W}_1$ to have characteristics that will ensure the nominal closed-loop system satisfies Requirements A-F. Once a suitable $\mathcal{P}_s$ is designed, the robustifying compensator $\mathcal{K}_s$ is computed by solving two decoupled Riccati equations [44]. This compensator stabilizes a family of perturbed plants, maximizing a stability margin. The basic design steps are shown in Fig. 3.

The next subsections describe the design details for the multi-zone heat pump, although we point out that the procedure is general in nature.

A. Compressor Discharge Temperature Schedule

The compressor discharge temperature T_D varies with operating conditions, and Requirement D means it must track a reference that is computed via a schedule. We define the T_D schedule as the output of a first-order lag filter with the compressor frequency CF as input [31]. This serves as a surrogate for the heat pump load. The schedule is tuned to optimize the system power consumption throughout the operating envelope. The full nonlinear model is used for this purpose. The simulation model is driven to a variety of steady-state conditions, and the steady-state gain of the schedule is tuned to give a small amount of superheat at the compressor inlet, which corresponds to an energy efficient and safe operation. In practice, the schedule is a nonlinear curve fit through these points, and depends on some other variables, but these details are beyond our scope. We augment the model (1) with the T_D reference schedule, and redefine to the resulting augmented model as $\bar{\mathcal{P}}$ henceforth, as shown in Fig. 4.

B. Weight Design

As is typical, $\mathcal{W}_1$ is a diagonal system of first-order filters, some of which include integral action to meet tracking, disturbance rejection and transient response Requirements A-F. $\mathcal{W}_2$ is a diagonal matrix of constants. In Section VI we will define constraints for each integral state in $\mathcal{W}_1$ to be the corresponding actuator limits, in order to provide integrator anti-windup. Note that if integral action were incorporated in $\mathcal{W}_2$ instead of $\mathcal{W}_1$, then it is not obvious how to realize anti-windup because each integrator would not correspond directly to an actuator.

For purposes of design, it is convenient to express the diagonal elements of $\mathcal{W}_1$ in a continuous-time Bode form representation so that the parameters have physical meaning during the iterative tuning procedure, and then discretize the result after the design is determined, although this is a slight abuse of notation in Fig. 4 since $\bar{\mathcal{P}}$ is defined in discrete-time.

For six of the inputs, CF, EEV i, $1 \le i \le 4$ and EEV M, we use Proportional Integral (PI) weights of the form

$$\mathcal{W}_{1j}(s) = k_j \frac{1 + s/\omega_j}{s}, \tag{6}$$

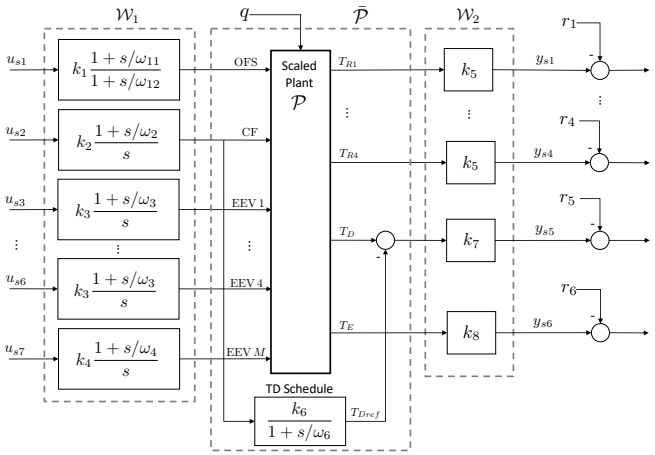

Fig. 4. Shaped plant $\bar{\mathcal{P}}_s$, with scaled plant $\mathcal{P} = \mathcal{P}_u + \mathcal{P}_q$ augmented with the T_D reference schedule to form $\bar{\mathcal{P}}$, input and output weights $\mathcal{W}_1$ and $\mathcal{W}_2$, respectively, and the reference $-r$ added to y_s.

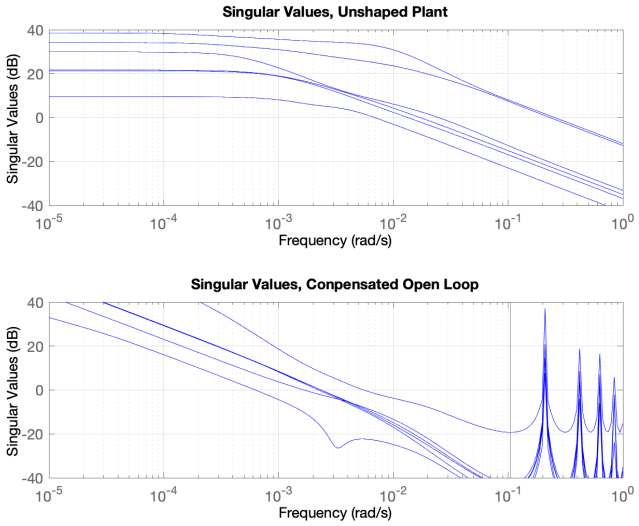

Fig. 5. Frequency response of $\bar{\mathcal{P}}_u$ (top), and $\bar{\mathcal{P}}_s\mathcal{K}_s$ (bottom).

for $2 \le j \le 7$. By symmetry, the gains and zeros of the EEV i PI weights, $1 \le i \le 4$, are all identical, as shown in Fig. 4. For the OFS input, we use a Lead-Lag (LL) weight

$$\mathcal{W}_{11}(s) = k_1 \frac{1 + s/\omega_{11}}{1 + s/\omega_{21}}, \tag{7}$$

with $\omega_{21} \ll \omega_{11}$ so that the gain is high at low frequencies. The integral action in the six compensators (6) ensures that the six tracking and disturbance rejection Requirements B-E are satisfied. Integral action on all seven inputs would result in an ill-posed loop-shaping problem, since we have only six regulated outputs. The LL weight (7) is used at the OFS input because it is the lowest gain input. In total, there are only 12 parameters to tune.

Frequency responses are plotted in Fig. 5. In the plot of $|\bar{\mathcal{P}}(j\omega)|$ (top) we see that there is one singular value with a faster bandwidth. It is aligned strongly with the CF input and the T_D output. The weakest (lowest gain) direction is

aligned in the direction of *differences* in room temperatures, which is a consequence of the heat pump architecture in which all the condensers are at the same pressure (neglecting pipe losses). The zero locations in (6)-(7) are tuned iteratively using the shaped frequency response (singular values) of $\bar{\mathcal{P}}_s$ to achieve a specified cross-over frequency $\omega_b \approx 0.005\,\mathrm{rad/s}$ corresponding to a rise-time of an $20-30$ minutes, a small amount of overshoot, and a desirable cross-over phase in order to meet Requirements A-F. Linear simulations are also done to tune the gains to achieve an acceptable transient response. At this point, the unconstrained design is complete and the subsequent steps are algorithmic.

C. Robustifying Compensator and Closed-Loop Properties

Once $\mathcal{W}_1$ and $\mathcal{W}_2$ are designed and discretized, the robustifying compensator $\mathcal{K}_s$ is computed for the shaped plant $\bar{\mathcal{P}}_s$. This involves solving two decoupled Riccati equations, for which solutions exist under the mild conditions that $\bar{\mathcal{P}}$ is stabilizable and detectable, and that the weights and plant do not possess any common pole-zero cancellations. The algorithm, provided in the Appendix, computes the compensator $\mathcal{K}_s$ that stabilizes the family of perturbed shaped plants

$$\widehat{\mathcal{P}}_s = \{(\mathcal{M}_s+\Delta_M)^{-1}(\mathcal{N}_s+\Delta_N) : \|\Delta_N \quad \Delta_M\|_\infty < \epsilon\} \quad (8)$$

for maximum $\epsilon > 0$, defining the stability margin ϵ, where $\bar{\mathcal{P}}_s = \mathcal{M}_s^{-1}\mathcal{N}_s$ is a normalized left coprime factorization of the shaped plant $\bar{\mathcal{P}}_s$ [23], [43], [44]. For the multi-zone heat pump, we have $\epsilon = 0.56$. This is good considering the rule-of-thumb that any $\epsilon > 0.25$ is acceptable [44].

D. $\mathcal{H}_\infty$ Controller Realizations

The $\mathcal{H}_\infty$ loop-shaping compensator $\mathcal{K}_s$ has an observer-based state feedback structure [24], [44]. The controller can be realized in different forms depending on how the reference r enters the feedback loop. One realization is

$$\begin{aligned}\widehat{x}_s(k+1) &= A_s\widehat{x}_s(k) + B_s u_s(k) \\ &\quad + H_s\left(\widehat{y}_s(k) - y_s(k) + r(k)\right) && (9a)\\ \widehat{y}_s(k) &= C_s\widehat{x}_s(k) && (9b)\\ u_s(k) &= K_s\widehat{x}_s(k)\,, && (9c)\end{aligned}$$

where the shaped plant, including disturbance, is realized as

$$\begin{aligned}x_s(k+1) &= A_s x_s(k) + B_s u_s(k) + B_{qs} q(k) && (10a)\\ y_s(k) &= C_s x_s(k)\,, && (10b)\end{aligned}$$

so that x_s includes x from (1a) and states from $\mathcal{W}_1$, and B, B_q and C in (1a)-(1b) are suitably augmented to form B_s, B_{qs} and C_s, respectively. Equations (9a)-(9b) are the state estimator with H_s being the robust estimator gain, and (9c) is the state feedback. This is diagrammed in Fig. 4, where the reference r enters the controller at the input of the estimator in (9a). In this realization, the compensator has the property that for constant values of disturbance q and reference r, the controller state $\widehat{x}_s$ converges to zero, when no constraints are active. This is a consequence of the stability of the closed-loop and observability of the controller $\mathcal{K}_s$.

An alternative realization, which attenuates the so-called derivative-kick, is

$$\begin{aligned}\widehat{x}_s(k+1) &= A_s\widehat{x}_s(k) + B_s u_s(k) \\ &\quad + H_s\left(\widehat{y}_s(k) - y_s(k)\right) && (11a)\\ \widehat{y}_s(k) &= C_s\widehat{x}_s(k) && (11b)\\ u_s(k) &= K_s\widehat{x}_s(k) + K_0 r(k). && (11c)\end{aligned}$$

In (11c), the $m \times m$ matrix

$$K_0 = K_s(I - A_s - B_s K_s - H_s C_s)^{-1} H_s W_2$$

is defined so that $r = y_1$ for constant r in steady-state, where y_1 corresponds to the six regulated outputs, and W_2 is the steady-state gain of $\mathcal{W}_2$.

VI. $\mathcal{H}_\infty$ Loop-Shaped MPC

One important feature of $\mathcal{H}_\infty$ loop-shaping is that the resulting $\mathcal{K}_s$ has an estimated state-feedback structure. For the LSMPC realization, the state estimator is used as (5), and the state feedback gain is used to compute Q, R and E_p via inverse optimality. This step is computational so that many design parameters for LSMPC are automatically determined by the $\mathcal{H}_\infty$ loop-shaping procedure. However, some MPC design parameters such as prediction horizon length and gains for soft constraints are not affected by the $\mathcal{H}_\infty$ design and hence still must be selected.

A. Inverse Optimality and the MPC Cost Function

Our objective is to construct an MPC that implements the loop-shaping controller $\mathcal{K}_s$ exactly when the constraints are inactive. We begin with the cost function. The matrices Q, R, and associated P in (4a) are computed by solving an inverse-optimal control problem, which can be stated as follows. Given the shaped plant state-space model for either of the controller realizations from Section V-D,

$$x_s(k+1) = A_s x_s(k) + B_s u_s(k)\,, \quad (12)$$

and the stabilizing full state feedback

$$u_s(k) = K_s x_s(k)\,, \quad (13)$$

find a quadratic cost function

$$J(u_s) = \sum_{k=1}^{\infty} x_s^T(k) Q_s x_s(k) + u_s^T(k) R_s u_s(k)\,, \quad (14)$$

where $R_s > 0$ and $Q_s \geq 0$ such that (13) minimizes (14) subject to (12). In other words, given A_s, B_s and K_s, compute $Q_s \geq 0$ and $R_s > 0$ such that P_s and K_s satisfy the associated Riccati equation

$$A_s^T P_s A_s - A_s^T P_s B_s (R_s + B_s^T P_s B_s)^{-1} B_s^T P_s A_s + Q_s = P_s \quad (15)$$

where

$$K_s = -(B_s^T P_s B_s + R_s)^{-1} B_s^T P_s A_s. \quad (16)$$

In this formulation, $E_p = I$ so $z = x_p = x_s$ in (4d).

Solutions to this problem are published [27], [45], with one solution being: Set $R_s = I$, and compute Q_s by solving the

same Riccati equations that are used to compute K_s and H_s in $\mathcal{K}_s$. Then (15) is used to compute P_s. For some plants, this approach may result in a numerically ill-conditioned P_s, which will make the resulting MPC optimization problem ill-conditioned, e.g. [42]. For the multi-zone heat pump, the condition number of P_s is 10^8, which is acceptable. An alternative approach is to solve an LMI for $Q_s \geq 0$, $R_s > 0$ and $P_s > 0$, subject to (15) and (16), that numerically minimizes the condition number of P_s. Also note that for the heat pump, Q_s is 32×32, meaning 528 scalar parameters would need to be tuned in the cost function using an LQR/LQG separation-principle approach. However, using loop-shaping, Q_s is synthesized using only 12 tuning parameters, all of which have a physical interpretation.

The basic idea in LSMPC is to use the state estimator from $\mathcal{K}_s$ to initialize the states of the prediction model $\bar{\mathcal{P}}_s$, realized as (12). However, there is a complication: The $\mathcal{H}_\infty$ estimator state $\widehat{x}_s(k)$ does not converge to the plant state $x_s(k)$ as $k \to \infty$ for non-zero, constant values of q. As a consequence, the constrained output v, computed using (12) as a prediction model, will be biased, leading to errors in constraint enforcement.

It is not feasible to augment the original plant state x with q and design the $\mathcal{H}_\infty$ estimator for the augmented plant in an attempt to integrate disturbance estimation into the $\mathcal{H}_\infty$ estimator, because such an augmented plant is not stabilizable and one of the algebraic Riccati equations in the $\mathcal{H}_\infty$ loop-shaping design cannot be solved. Furthermore, it is not feasible to add the term $B_{qs}\widehat{q}$, where $\widehat{q}$ is an estimate of the unmeasured q, to either (11b) or (9a) with the intention of removing the bias from the estimate $\widehat{x}_s$, because this will result in loss of tracking of r.

Instead, we construct an *augmented* prediction model that *includes* the shaped plant $\mathcal{P}_s$ to predict the performance output z, but is augmented with additional states to predict the constrained output v without any bias caused by q. Toward this end, we construct an observer for q and v, which we denote the *disturbance observer*, with two alternative approaches outlined in the next two subsections

B. Full-Order Disturbance Observer

Rewrite (1) to include q as an additional state, assuming that q is constant over the prediction horizon,

$$\bar{x}(k+1) = A_q\bar{x}(k) + B_q u(k) \tag{17a}$$
$$y(k) = C_q\bar{x}(k) \tag{17b}$$
$$v(k) = F_q\bar{x}(k) + Gu(k) \tag{17c}$$

where

$$A_q = \begin{bmatrix} A & B_d \\ 0 & I \end{bmatrix} \qquad B_q = \begin{bmatrix} B \\ 0 \end{bmatrix}, \tag{18}$$

$C_q = [C \;\; 0]$, $F_q = [F \;\; 0]$, and $\bar{x} = [x \;\; q]^T$. If (A_q, C_q) is detectable, which holds for our system, we construct a Luenberger observer, defining the process and measurement weights as

$$Q_q = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} \qquad \text{and} \qquad R_q = \rho I, \tag{19}$$

where $\rho > 0$ is a single tuning parameter (to reduce the number of tuning parameters), and computing the observer gain L_q by solving the Algebraic Riccati Equation, resulting in

$$\widehat{\bar{x}}(k+1) = A_q\widehat{\bar{x}}(k) + B_q u(k) + L_q\left(\widehat{y}(k) - y(k)\right) \tag{20a}$$
$$\widehat{y}(k) = C_q\widehat{\bar{x}}(k) \tag{20b}$$
$$\widehat{v}(k) = F_q\widehat{\bar{x}}(k) + Gu(k) \tag{20c}$$

This observer is appropriate for the heat pump application, but other types of observers are also possible.

C. Reduced-Order Disturbance Observer

Observer (20) uses the open-loop plant model $\mathcal{P}$, and therefore the estimate error $v - \widehat{v}$ is sensitive to model uncertainty. In particular, it can produce estimates $\widehat{v}$ that have steady-state bias for constant q. However, in many applications including the multi-zone heat pump, v is directly measured, and a reduced-order observer reduces this sensitivity and eliminates the bias in the steady-state.

The construction of a reduced-order observer for v is conventional [46]. From the observability matrix

$$\Omega = \begin{bmatrix} C_q & C_qA_q & \ldots & C_qA_q^{n-1} \end{bmatrix}^T, \tag{21}$$

choose the first $n+d$ linearly independent rows to form the matrix Φ, defining the change of coordinates $\xi = \Phi\bar{x}$. To express (17) in the ξ-coordinates, compute $\bar{A} = \Phi A_q \Phi^{-1}$, $\bar{B} = \Phi B_q$ and $\bar{C} = C_q\Phi^{-1}$, which by construction have the form

$$\bar{A} = \begin{bmatrix} \bar{A}_{11} & \bar{A}_{12} \\ \bar{A}_{21} & \bar{A}_{22} \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} \bar{B}_1 \\ \bar{B}_2 \end{bmatrix}, \quad \bar{C} = [I \;\; 0], \tag{22}$$

meaning the first p elements of ξ are y, which contains v and which does not require estimation. Choose $(n+d-p)\times(n+d-p)$ matrix $\bar{Q} \geq 0$ and $p \times p$ matrix $\bar{R} > 0$ and solve the Riccati equation

$$\bar{A}_{22}\bar{S}\bar{A}_{22}^T - \bar{S} - \bar{A}_{22}\bar{S}\bar{A}_{12}^T(\bar{R} + \bar{A}_{12}\bar{S}\bar{A}_{12}^T)^{-1}\bar{A}_{12}\bar{S}\bar{A}_{22}^T + \bar{Q} = 0$$

for $\bar{S} > 0$, which exists if (A_q, C_q) is detectable. Then the reduced-order observer is

$$\widehat{w}(k) = L_w y(k) + w(k) \tag{23a}$$
$$\widehat{\bar{x}}(k) = \Phi^{-1}\begin{bmatrix} y(k) & \widehat{w}(k) \end{bmatrix}^T \tag{23b}$$
$$w(k+1) = F_w w(k) + G_w y(k) + H_w u(k) \tag{23c}$$
$$\widehat{v}(k) = F_q\widehat{\bar{x}}(k) + Gu(k) \tag{23d}$$

where

$$L_w = \left(\bar{A}_{12}\bar{S}\bar{A}_{12}^T + \bar{R}\right)^{-1}\bar{A}_{12}\bar{S}\bar{A}_{22}^T \tag{24a}$$
$$F_w = \bar{A}_{22} - L_w\bar{A}_{12} \tag{24b}$$
$$H_w = \bar{B}_2 - L_w\bar{B}_1 \tag{24c}$$
$$G_w = \bar{A}_{21} - L_w\bar{A}_{11} + F_wL_w. \tag{24d}$$

The reduced-order observer has the advantage of producing an estimate $\widehat{v}(k)$ at sample time kT without bias due to q or model uncertainty, although it may be more sensitive to sensor noise than the full-order observer.

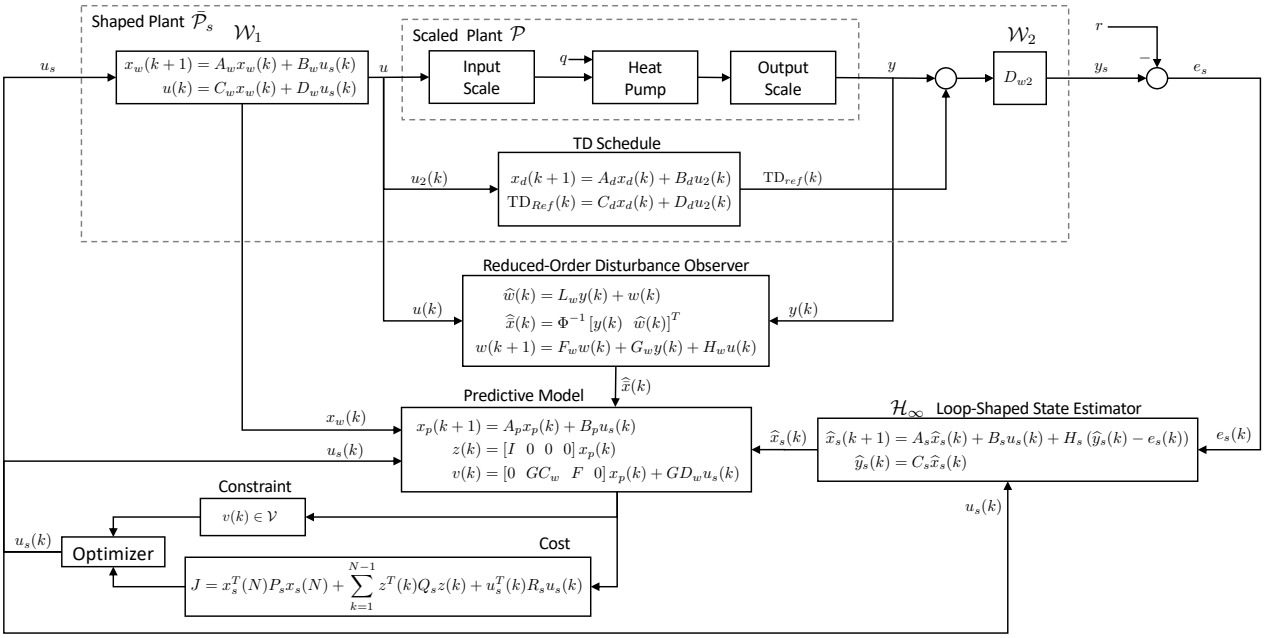

Fig. 6. Block diagram of LSMPC for the multi-zone heat pump, using the reduced-order disturbance observer.

D. LSMPC Realization

To construct the LSMPC, we define the cost function using the inverse-optimal Q_s, R_s and P_s from Subsection VI-A, and construct the prediction model to include both the shaped plant model (A_s, B_s), which predicts the performance output $z(k)$, and the disturbance observer (20) or (23), which predicts the effect of the disturbance q. First, we realize the input weight $\mathcal{W}_1$ in state-space as

$$x_w(k+1) = A_w x_w(k) + B_w u_s(k) \tag{25a}$$
$$u(k) = C_w x_w(k) + D_w u_s(k), \tag{25b}$$

which includes the integral states and lead-lag states from (6) and (7). Define the prediction model state as

$$x_p(k) = [\widehat{x}_s^T(k) \;\; x_w^T(k) \;\; \widehat{\bar{x}}^T(k)]^T, \tag{26}$$

where $\widehat{\bar{x}}(k)$ is from either (20) or (23). Then the LSMPC prediction model is

$$x_p(k+1) = \begin{bmatrix} A_s & 0 & 0 & 0 \\ 0 & A_w & 0 & 0 \\ 0 & BC_w & A & B_d \\ 0 & 0 & 0 & I \end{bmatrix} x_p(k) + \begin{bmatrix} B_s \\ B_w \\ BD_w \\ 0 \end{bmatrix} u_s(k) \tag{27a}$$
$$z(k) = \begin{bmatrix} I & 0 & 0 & 0 \end{bmatrix} x_p(k) \tag{27b}$$
$$v(k) = \begin{bmatrix} 0 & GC_w & F & 0 \end{bmatrix} x_p(k) + GD_w u_s(k) \tag{27c}$$

and the cost function is

$$J(u_s) = x_s^T(N) P_s x_s(N) + \sum_{k=1}^{N-1} z^T(k) Q_s z(k) + u_s^T(k) R_s u_s(k). \tag{28}$$

At sample time kT, the prediction model (27) is initialized with $\widehat{x}_s(k)$, which is computed by the $\mathcal{H}_\infty$ estimator (9a)-(9b) or (11a)-(11b), and with $\widehat{\bar{x}}(k)$, which is computed by the disturbance observer (20) or (23), and with $x_w(k)$. Note that the performance variable z depends on only the $\mathcal{H}_\infty$ estimator state $\widehat{x}_s$, due to the structure of (27b). The disturbance estimate $\widehat{q}$ affects the predicted constrained output v, but not the predicted performance output z. Therefore the LSMPC is identical to the $\mathcal{H}_\infty$ loop-shaped controller when constraints are inactive. A block diagram of the controller is shown in Fig. 6.

If the problem lacks output constraints, then the prediction model (27) can be simplified. In this case, a disturbance observer is not required, and (26) includes only $\widehat{x}_s(k)$ and $x_w(k)$. Any input constraints will be properly enforced, and any unmeasured disturbance $q(k)$ will be properly rejected, because $q(k)$ does not directly affect $x_w(k)$.

E. LSMPC Properties

Proposition 1: For the set of perturbed plants (8) and the model predictive control with prediction model (27), cost (28), observer (20) or (23), and $\mathcal{H}_\infty$ estimator (9a)-(9b), let $\mathbb{X}_{ia}$ be the set of states where the constraints (4c) are strictly satisfied by the MPC input sequence along the entire prediction horizon. For every equilibrium $\hat{x}_p^e$ in the interior of $\mathbb{X}_{ia}$ there exists a set $\mathbb{S} \subseteq \mathbb{X}_{ia}$ with $\hat{x}_p^e$ in the interior of $\mathbb{S}$ such that $\mathbb{S}$ is invariant for the plant in closed-loop with MPC, and the closed-loop is stable in $\mathbb{S}$ with the same stability margin ϵ provided by $\mathcal{H}_\infty$ loop-shaped controller.

Proof: Since the MPC cost function is designed by inverse optimality, in $\mathbb{X}_{ia}$ the MPC command is equal to the command of the $\mathcal{H}_\infty$ loop-shaped controller (9). Note that $\mathbb{X}_{ia}$ has a non-empty interior since the $\mathcal{H}_\infty$ loop-shaped controller is stabilizing. Let $\mathbb{S} \subseteq \mathbb{X}_{ia}$ be any invariant set of the closed-loop between the plant (8) and the $\mathcal{H}_\infty$ loop-shaped controller (9) with $\hat{x}_p^e$ in the interior of $\mathbb{S}$, so that starting from within $\mathbb{S}$, the closed-loop trajectory remains in $\mathbb{S}$. Since $\mathbb{S} \subseteq \mathbb{X}_{ia}$, the MPC is equal to the $\mathcal{H}_\infty$ loop-shaped controller in $\mathbb{S}$, so that $\mathbb{S}$ is invariant also for the closed-loop between the plant (8) and

the MPC. Finally, the existence of $\mathbb{S}$ with the above properties follows from the closed-loop between the plant (8) and the $\mathcal{H}_\infty$ loop-shaped controller being exponentially stable, which ensures the existence of a Lyapunov function, whose sublevel sets are invariant and contain $\hat{x}_p^e$ in their interior. Hence, any such sublevel set is entirely contained in $\mathbb{X}_{ia}$.

Summarizing, MPC is equivalent to the $\mathcal{H}_\infty$ loop-shaped controller in the invariant set $\mathbb{S}$ which contains $\hat{x}_p^e$ in its interior, and hence in such set, the closed-loop system has the same properties as the closed-loop system obtained with the $\mathcal{H}_\infty$ controller, including the same stability margin ϵ. ■

VII. Simulation & Experimental Results

We present both simulation and experimental results. Simulations of the linear plant $\mathcal{P}$ with constraints show the ideal performance of LSMPC and are also included because the MERL laboratory operates within certain constraints, e.g. the outdoor air temperature must be above freezing. For both, we used the same LSMPC controller realization (9) and formed the prediction model, cost and constraints as described in Section VI-D using the reduced-order disturbance observer (23). There were hard upper and lower limits on all of the actuators, and soft constraints on the following outputs: $T_{SCi} \geq 0°\mathrm{C}$, $T_D \leq 80°\mathrm{C}$, $T_{DSH} \geq 10°\mathrm{C}$ and $T_C \leq 50°\mathrm{C}$. Constraint softening on the outputs guaranteed that the optimization problem was feasible at every sample time. The hard constraint on the inputs u ensured that the integrators in $\mathcal{W}_1$ did not wind-up. The LSMPC horizon was set to $N = 5$, the sample period was $T = 30\,\mathrm{s}$, making the horizon $2.5\,\mathrm{min}$ in length. Longer horizons provided limited benefit. This was because the bandwidth of the system from inputs to constrained outputs is faster than that from inputs to the room temperatures, and also because output constraints were softly enforced.

At each sample time, the constrained convex quadratic program was solved using two methods: ADMM [47] and PQP [48]. ADMM was found to be more consistent in terms of computation time across different initial conditions. Furthermore, the ADMM algorithm in [47] allows for implementing "exact soft constraint penalties" [49] based on the 1-norm, which is not allowed by methods requiring strictly positive definite Hessian in the quadratic program.

A. Simulation Results

Two simulation results are presented. Results from the first are plotted in Fig. 7. The temperature setpoint in Room 1 was increased by $2°\mathrm{C}$ at $t = 10\,\mathrm{min}$, and then the setpoints for the other three rooms were increased by $2°\mathrm{C}$ at $t = 150\,\mathrm{min}$, and finally a $-1\mathrm{kW}$ heat load step was applied to all rooms at $t = 300\,\mathrm{min}$. Only the OFS constraint was active after the disturbance transient, which did not adversely affect tracking. Prior to this, no constraints were active along the entire prediction horizon, so the LSMPC controller was equivalent to the $\mathcal{H}_\infty$ loop-shaping controller. Plots for the $\mathcal{H}_\infty$ loop-shaping controller coincide exactly with the LSMPC when constraints are inactive. They are omitted to simplify the figures. The room temperatures showed little coupling during the first transient, achieved the designed rise-time of $20\,\mathrm{min}$ with little overshoot and acceptable damping, and rejected the disturbance with zero steady-state error.

The second simulation was more aggressive to show constraint enforcement, and also included sensor quantization at the $0.25°\mathrm{C}$ level of production sensors to show that the controller, and specifically the reduced-order observer, is effective despite the noise. Results are plotted in Fig. 8. The setpoint in Room 1 was increased by $5°\mathrm{C}$ at $t = 10\,\mathrm{min}$, and then the setpoints for the other three rooms were increased by $10°\mathrm{C}$ at $t = 150\,\mathrm{min}$, and finally a $-1\mathrm{kW}$ heat load step was applied to all rooms at $t = 300\,\mathrm{min}$. The first setpoint change caused the subcooling constraints to become active, and this prevented the room temperatures from achieving their setpoints. However, the behavior was acceptable since the response was stable (non-oscillatory), all rooms were close to their desired temperature, and the active constraints were "softly" enforced. The second transient caused the discharge and condensing temperatures, T_D and T_C, to hit their respective upper limits, which were "softly" enforced. This mildly slowed the room temperature response. As discussed previously, the small amount of constraint violation in transient is acceptable. Note that after the transient transpired, each room achieved its setpoint, and all actuators were away from their limits. Finally, the load step at $t = 300\,\mathrm{min}$ caused T_C to be at its limit, while both CF and the OFS were at their limits. At this point the equipment was near its maximum capacity. Nevertheless, the room temperatures achieved their setpoints. For this simulation, the ADMM solver, which was compiled from C code using the Matlab MEX interface, took between 0.5 ms and 75 ms, with a mean time of 45 ms, to compute the control on a single core of a 2.5GHz Intel i7-4870HQ, with the longer times associated with the transient at $t = 150\,\mathrm{min}$ when multiple constraints were active.

B. Experimental Results

A Mitsubishi Electric four-zone heat pump with $10\mathrm{kW}$ capacity, sufficient for large residential use, was used to test the LSMPC. The laboratory set-up is diagrammed in Fig. 9. Each indoor unit is installed in an insulated, nearly adiabatic test chamber, together with a hydronic fan coil and electric heater to provide heating and cooling loads. Each fan coil provides a constant "base" heat load, while the electric heaters provide for transients. Two of the indoor chambers are $9\,\mathrm{m}^3$, two are $4.5\,\mathrm{m}^3$, and the outdoor test chamber is $6.3\,\mathrm{m}^3$. The entire system is controlled by a pair of Labview systems. The first operates the balance of plant, including the air-cooled chiller, water pumps, valves, and the electric heaters. It also collects a large amount of performance data, such as thermopile measurements across each heat exchanger. The second Labview system implements the LSMPC controller. See [19] for further details.

When testing in heating mode, the external air-cooled chiller supplies cold water to the fan coils in each indoor chamber to apply a negative heat load, and the electric heaters are used to modulate that load in order to conduct transient experiments. The electric heater in the outdoor test chamber is controlled to regulate the outdoor chamber temperature to a desired value of T_A, which is constant in these experiments.

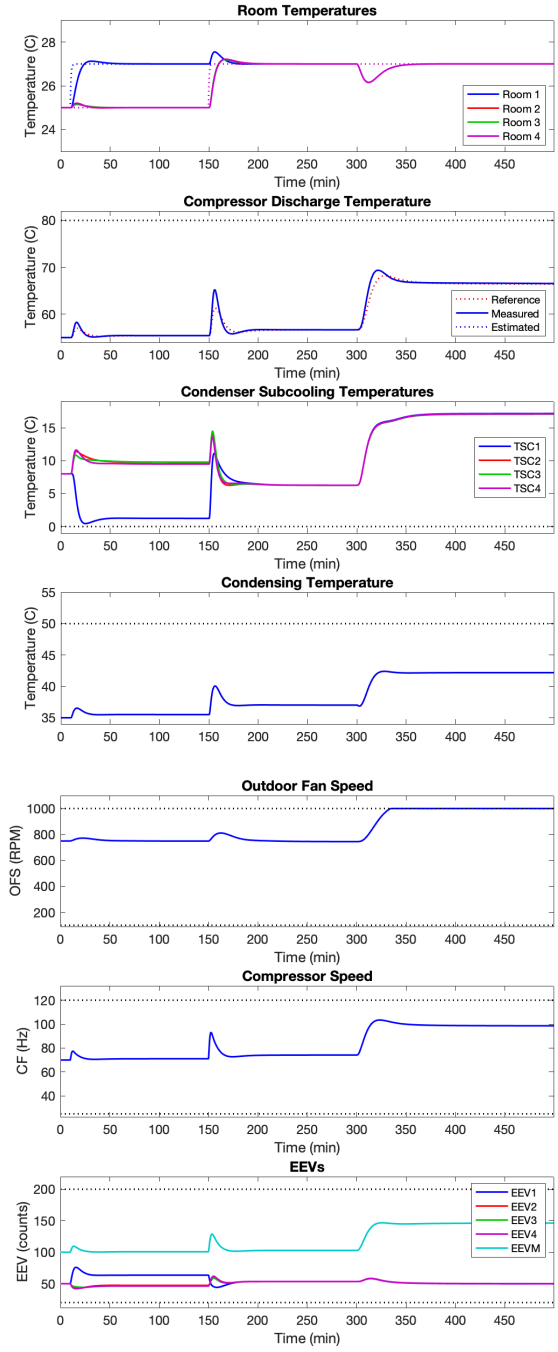

Fig. 7. Simulation results for the linear plant. At $t = 10\,\text{min}$, the setpoint for Room 1 was increased by $2°\text{C}$, then at $t = 150\,\text{min}$, the setpoint for the other three rooms was increased by $2°\text{C}$, and finally at $t = 300\,\text{min}$, a heat load step of $-1\,\text{kW}$ was applied to all rooms. Constraints are indicated by dotted lines. For this simulation, only the OFS constraint was active after the third transient.

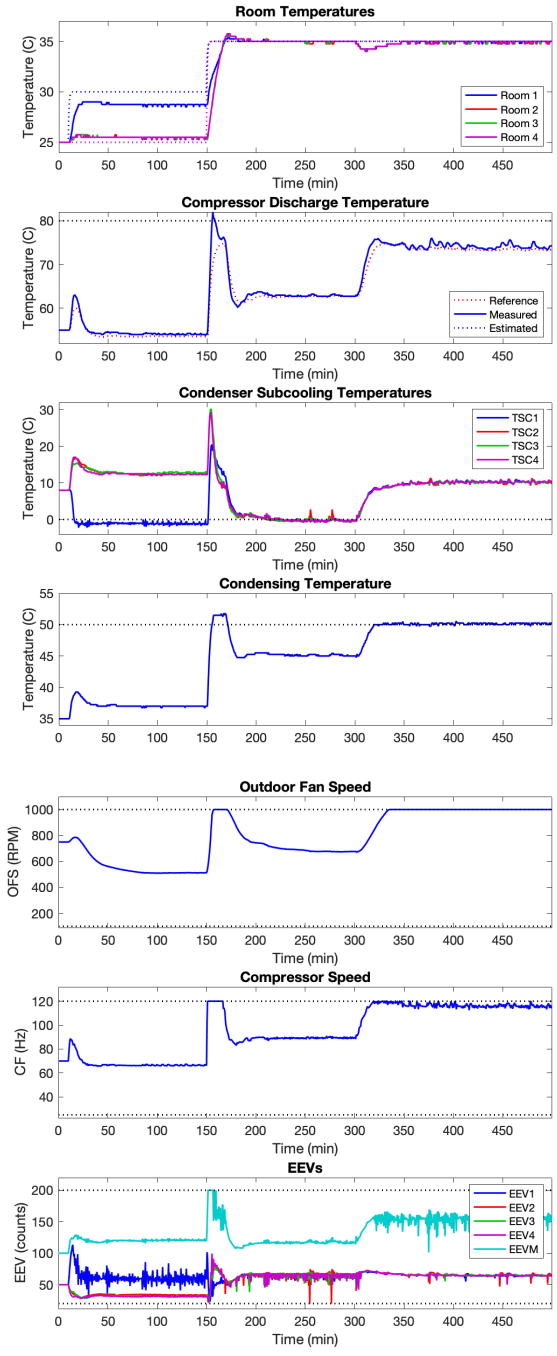

Fig. 8. Simulation results for the linear plant with temperature sensor quantization. At $t = 10\,\text{min}$, the setpoint for Room 1 was increased by $5°\text{C}$, then at $t = 150\,\text{min}$, the setpoint for the other three rooms was increased by $10°\text{C}$, to $35°\text{C}$, and finally at $t = 300\,\text{min}$, a heat load step of $-1\,\text{kW}$ was applied to all rooms. Constraints are indicated by dotted lines. All sensed temperatures were quantized to $0.25°\text{C}$ except T_E, which was quantized to $0.1°\text{C}$. The condenser subcooling constraints were active after the first transient, preventing the room temperatures from achieving their setpoints. Hard actuator constraints were enforced during and after the second and third transients. Soft output constraints were enforced throughout, with a small but acceptable error during the transients.

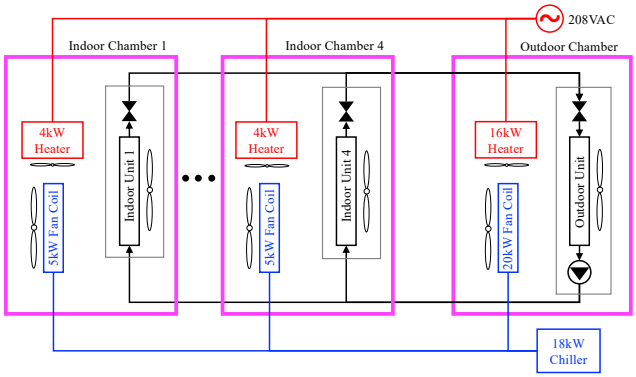

Fig. 9. MERL HVAC laboratory schematic, showing four indoor chambers, one outdoor chamber, and the balance of plant. Each adiabatic test chamber contains an indoor or outdoor unit under test, a hydronic fan coil and an electric heater to modulate the heat load.

The controller was synthesized in Matlab, and a finite-time optimal control solver based on ADMM was implemented in C, compiled and run in real-time within Labview, which interfaced to the heat pump sensors and actuators. The same LSMPC controller was used for both the simulations and the experiments. The startup sequence first turned on the heat pump using its production startup logic, then started the balance of plant, and brought the system to the desired operating condition. The LSMPC was then turned on by first running the two estimators for a few iterations, then turning on the optimizer, and finally closing the loop by actuating the controlled inputs computed by the optimizer.

We present three experimental results. In the first experiment, each of the room temperature setpoints was stepped up by $2°\mathrm{C}$ in succession. Results are plotted in Fig. 10. The room temperatures achieved their setpoints without offset, and with rise-time of approximately $20\,\mathrm{min}$. Rooms 3 and 4 were slightly faster because they are smaller in volume. All temperature sensors were quantized to $0.25°\mathrm{C}$ except T_E, which was quantized to $0.1°\mathrm{C}$. The effect of quantization is especially noticeable in the room temperatures. The compressor discharge temperature and evaporator temperatures both tracked their references without offset. The outdoor fan and compressor speeds remained within limits, but the EEVs hit upper and lower constraints during the transients, which were enforced by the LSMPC. In these experiments, the EEV limits were tighter than in the simulations to protect the equipment.

In the second experiment, all room temperature setpoints were stepped up simultaneously by $3°\mathrm{C}$. Results are plotted in Fig. 11. The rise-time was about $15\,\mathrm{min}$, faster than the first experiment because this output direction (the same temperature in all rooms) has higher plant gain than for the first experiment (different setpoint temperatures in the rooms). All of the regulated outputs achieved their setpoint without offset. For this experiment, no constraints were active. We note that the EEV M limit was 150, beyond the plot range.

The third experiment was a repeat of experiment two, but the limit for the discharge temperature (T_D) was reduced to $62°\mathrm{C}$ to show output constraint enforcement. The results are

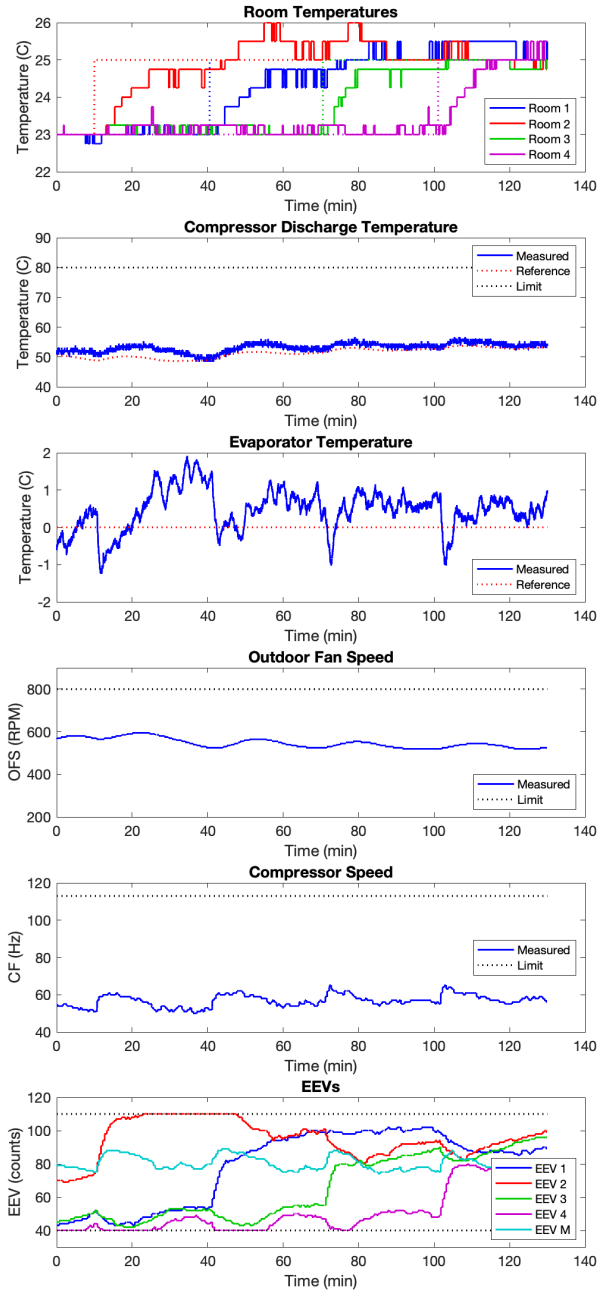

Fig. 10. Experiment 1 of the LSMPC for the four-zone heat pump. The room temperature setpoints were stepped in succession $2°\mathrm{C}$. The EEVs hit upper and lower limits during transients, but the room, compressor discharge and evaporator temperatures all achieved their setpoints in the steady-state.

plotted in Fig. 12. The T_D constraint became active during the transient. To enforce the constraint, the LSMPC reduced the compressor speed (CF) and the room EEVs. As a result the room temperature response was slightly slower in the smaller rooms (3 and 4). After the transient, all regulated variables tracked their references without offset, although EEVs 3 and 4 were at their limits. This occurred because those two rooms are located the farthest from the outdoor unit, so that the EEVs need to be opened wider to deliver the same amount of heat, compared to Rooms 1 and 2.

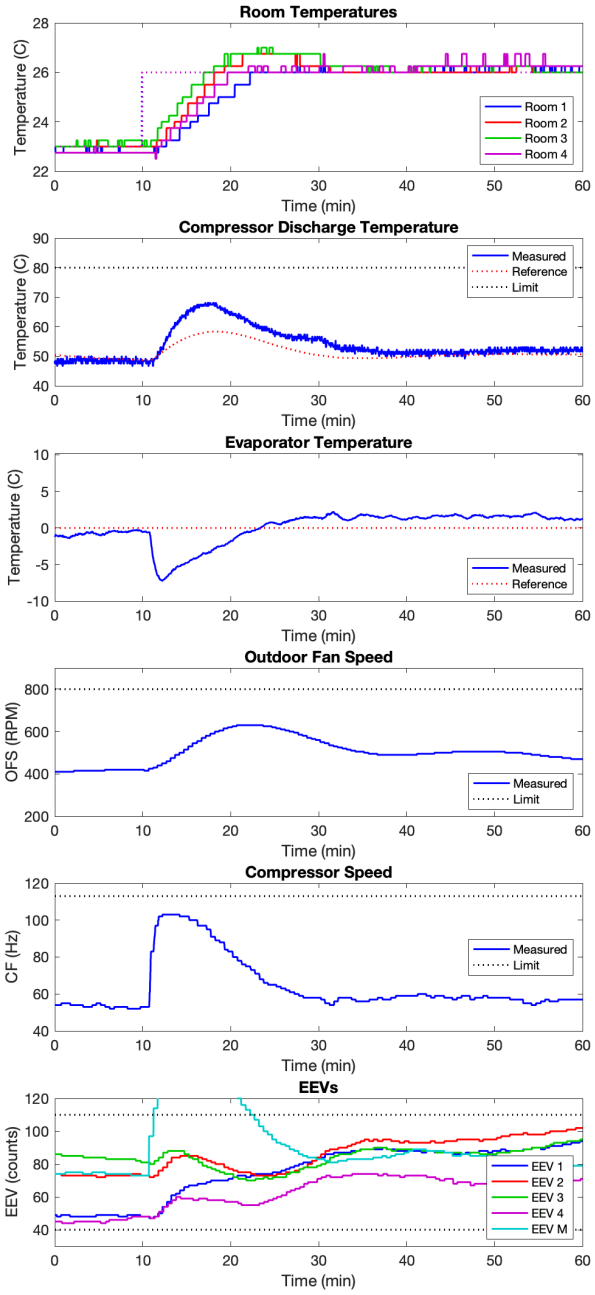

Fig. 11. Experiment 2 of the LSMPC for the four-zone heat pump. The room temperature setpoints were stepped simultaneously by 3°C. The rise-time was faster than for different room temperatures. No constraints were active. Note that the EEV M upper limit was 150, outside the range of the lower plot.

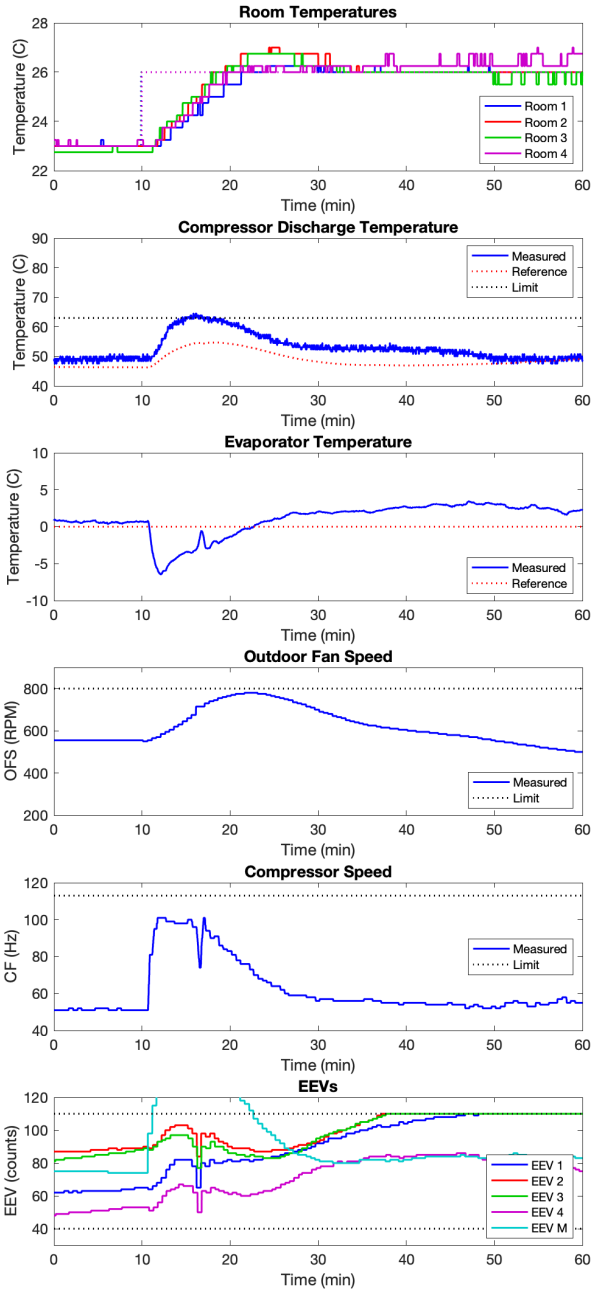

Fig. 12. Experiment 3 of the LSMPC for the four-zone heat pump. The room temperature setpoints were stepped simultaneously by 3°C, and the upper limit for the discharge temperature T_D was set to 62°C to show output constraint enforcement. Compare to Fig. 11.

VIII. Conclusions

An MPC formulation was presented based on $\mathcal{H}_\infty$ loop-shaping. This is a design methodology that uses multivariable frequency domain methods to design plant model augmentations and to compute values for the MPC cost function weights and state estimator gain that meet performance requirements and also maximize the stability margin with respect to normalized coprime factor uncertainty. The method can be applied to general MPC problems, but is most appropriate for those that require state estimation. We used the method to design an MPC for a multi-zone heat pump, and validated the approach in simulation and in experiment. The MPC for this problem has a good stability margin, enforces input and output constraints, tracks constant references and rejects constant disturbances with zero steady-state error, and provides good transient response. The tuning procedure is straightforward with a minimal number of parameters.

A number of extensions is possible. First, alternative architectures of the $\mathcal{H}_\infty$ feedback loop and their realizations as an MPC should be explored. How should reference or

disturbance preview be incorporated? Can the $\mathcal{H}_\infty$ estimator be used to estimate the constrained outputs, which are subject to unmeasured disturbance? This would eliminate the need for the disturbance observer. Secondly, computation of the stability margin when the constraints are active in steady state could be investigated, since the MPC can be realized as a piecewise affine state feedback law for each combination of active constraints [39]. For the heat pump application, when output constraints are active in the steady state, the closed-loop remains stable, but this may not be true for other applications.

ACKNOWLEDGMENT

We gratefully acknowledge Fumiaki Baba, Takaya Yamamoto, Masashi Fujitsuka, Yuki Mori, and Takahiro Nakai of the Advanced Technology Center, Mitsubishi Electric Corporation, Japan, for their strong support of this work.

REFERENCES

[1] A. Afroz, G. Shafiullah, T. Urmee, and G. Higgins, "Modeling techniques used in building HVAC control systems: A review," *Renewable and Sustainable Energy Reviews*, 2017.

[2] "Buildings and climate change: Summary for decision - makers," UNEP's Sustainable Buildings and Climate Initiative (SBCI), 2009.

[3] A. H. Glattfelder and W. Schaufelberger, *Control Systems with Input and Output Constraints*. Springer, 2003.

[4] A. Afram and F. Janabi-Shariffi, "Theory and applications of HVAC control systems - A review of model predictive control (MPC)," *Building and Environment*, pp. 343–355, 2014.

[5] X.-C. Xi, A.-N. Poo, and S.-K. Chou, "Support vector regression model predictive control on an HVAC plant," *Control Engineering Practice*, vol. 15, no. 8, pp. 897–908, 2007.

[6] M. S. Elliott and B. P. Rasmussen, "Model-based predictive control of a multi-evaporator vapor compression cooling cycle," in *Proc. of the American Control Conference*, 2008, pp. 1463–1468.

[7] D. Gyalistras and M. Gwerder, "Use of weather and occupancy forecasts for optimal building climate control (opticontrol): Two years progress report," Terrestrial Systems Ecology ETH Zurich, Tech. Rep., 2010.

[8] A. Kelman and F. Borrelli, "Bilinear model predictive control of an HVAC system using sequential quadratic programming," *IFAC Proc. Volumes*, vol. 44, no. 1, pp. 9869–9874, 2011.

[9] J. Rehrl and M. Horn, "Temperature control for HVAC systems based on exact linearization and model predictive control," in *IEEE International Conference on Control Applications*, 2011.

[10] A. Aswani, N. Master, J. Taneja, D. Culler, and C. Tomlin, "Reducing transient and steady state electricity consumption in HVAC using learning-based model predictive control," *Proc. of the IEEE*, vol. 100, no. 1, pp. 240–253, 2012.

[11] Y. Ma, J. Matusko, and F. Borrelli, "Stochastic model predictive control for building HVAC systems: Complexity and conservatism," *IEEE Trans. on Control Systems Technology*, pp. 101–116, 2012.

[12] M. Maasoumy and A. Sangiovanni-Vincentelli, "Total and peak energy consumption minimization of building HVAC systems using model predictive control," *IEEE Design & Test of Computers*, vol. 29, no. 4, pp. 26–35, 2012.

[13] S. C. Bengea, A. D. Kelman, F. Borrelli, R. Taylor, and S. Narayanan, "Implementation of model predictive control for an HVAC system in a mid-size commercial building," *Journal of HVAC&R Research*, vol. 20, pp. 121–135, 2014.

[14] J. Cigler, J. Široký, M. Korda, and C. N. Jones, "On the selection of the most appropriate MPC problem formulation for buildings," University Centre for Energy Efficient Buildings, Czech Technical University in Prague, Tech. Rep., 2013.

[15] S. R. West, J. K. Ward, and J. Wall, "Trial results from a model predictive control and optimization system for commercial building HVAC," *Energy and Buildings*, vol. 72, pp. 271–279, 2014.

[16] M. Wallace, P. Mhaskar, J. House, and T. Salsbury, "Offset-free model predictive controller of a heat pump," in *Proc. of the IEEE American Control Conference*, 2014, pp. 2247–2252.

[17] M. Maasoumy, M. Razmara, M. Shakbakhti, and A. Sangiovanni-Vincentelli, "Handling model uncertainty in model predictive control for energy efficient buildings," *Energy and Buildings*, pp. 377–392, 2014.

[18] G. Mantovani and L. Ferrarini, "Temperature control of a commercial building with model predictive control techniques," *IEEE Trans. on Industrial Electronics*, vol. 62, no. 4, pp. 2651–2660, 2015.

[19] D. J. Burns, C. Danielson, S. Di Cairano, C. R. Laughman, and S. A. Bortoff, *Intelligent Building Control Systems: A Survey of Modern Building Control and Sensing Strategies*. Springer, 2017, ch. Model Predictive Control of Multi-Zone Vapor Compression Systems, pp. 105–137.

[20] A. Afram, F. Janabi-Sharifi, A. S. Fung, and K. Raahemifar, "Artificial neural network (ANN) based model predictive control (MPC) and optimization of HVAC systems: A state of the art review and case study of a residential HVAC system," *Energy and Buildings*, vol. 141, no. 15, pp. 96–113, 2017.

[21] G. Serale, M. Fiorentini, A. Capozzoli, D. Bernardini, and A. Bemporad, "Model predictive control (MPC) for enhancing building and HVAC system energy efficiency: Problem formulation, applications and opportunities," *Energies*, vol. 11, no. 3, 2018.

[22] J. Drgoňa, J. Arroyo, I. C. Figueroa, D. Blum, K. Arendt, D. Kim, D. P. Ollé, J. Oravec, M. Wetter, D. L. Vrabie, and L. Helsen, "All you need to know about model predictive control for buildings," *Annual Reviews in Control*, vol. 50, pp. 190–232, 2020.

[23] K. Glover and D. McFarlane, "Robust stabilization of normalized coprime factor plant descriptions with H_∞ bounded uncertainty," *IEEE Trans. on Automatic Control*, vol. 34, no. 8, pp. 821–830, 1989.

[24] J. Sefton and K. Glover, "Pole/zero cancellations in the general H_∞ problem with reference to a two block design," *Systems & Control Letters*, vol. 14, no. 4, pp. 295–306, April 1990.

[25] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999.

[26] J. Löfberg, "YALMIP : A toolbox for modeling and optimization in MATLAB," in *In Proc. of the IEEE CACSD Conference*, Taipei, Taiwan, 2004.

[27] C. Rowe and J. Maciejowski, "Tuning MPC using H_∞ loop shaping," in *Procedings of the American Control Conference*, 2000, pp. 1332–1336.

[28] J. Maciejowski, "Reverse-engineering existing controllers for MPC design," *IFAC Proc.*, vol. 40, no. 20, pp. 436–441, 2007.

[29] S. Di Cairano and A. Bemporad, "Model predictive control tuning by controller matching," *IEEE Trans. on Automatic Control*, vol. 55, no. 1, pp. 185–190, 2010.

[30] S. A. Bortoff, P. Schwerdtner, C. Danielson, and S. Di Cairano, "H-infinity loop-shaped model predictive control with heat pump application," in *Proc. of the European Control Conference*, 2019.

[31] S. A. Bortoff, D. J. Burns, C. R. Laughman, H. Qiao, C. Danielson, A. Goldsmith, and S. Di Cairano, "Power optimizing control of multi-zone heat pumps," in *IEEE Conference on Control Technology and Applications*, Aug. 2018, pp. 826–833.

[32] P. Fritzon, *Principles of Object Oriented Modeling and Simulation with Modelica 3.3: A Cyber-Physical Approach*. Wiley, 2015.

[33] *Modelica Language Specification Version 3.5*, Modelica Association, https://www.Modelica.org/, Feb. 2021.

[34] H. Qiao, V. Aute, and R. Radermacher, "Transient modeling of a flash tank vapor injection heat pump system–Part I: Model development," *International Journal of Refrigeration*, vol. 49, pp. 169–182, 2015.

[35] H. Qiao, C. R. Laughman, S. A. Bortoff, and D. J. Burns, "Dynamic characteristics of an R410a multi-split variable refrigerant flow air conditioning system," in *Proc. of the 12th IEA Heat Pump Conference*, 2017.

[36] M. Wetter, W. Zuo, T. S. Nouidui, and X. Pang, "Modelica buildings library," *Journal of Building Performance Simulation*, vol. 7, no. 4, pp. 253–270, 2014.

[37] B. P. Rasmussen and A. G. Alleyne, "Gain scheduled control of an air conditioning system using the Youla parameterization," *IEEE Trans. on Control Systems Technology*, vol. 18, no. 5, pp. 1216–1225, 2010.

[38] J. B. Rawlings, D. Q. Mayne, and M. M. Diehl, *Model Predictive Control: Theory, Computation and Design, 2nd Edition*. Nob Hill Publishing, LLC, 2017.

[39] F. Borrelli, A. Bemporad, and M. Morari, *Predictive control for linear and hybrid systems*. Cambridge University Press, 2017.

[40] J. Doyle, "Guaranteed margins for LQG regulators," *IEEE Trans. on Automatic Control*, vol. 23, no. 4, pp. 756–757, Aug. 1978.

[41] B. P. Molinari, "The stable regulator problem and its inverse," *IEEE Trans. on Automatic Control*, vol. AC-18, no. 5, pp. 454–459, Oct. 1973.

[42] M. C. Priess, R. Conway, J. Choi, J. John M. Popovich, and C. Radcliffe, "Solutions to the inverse LQR problem with application to biological systems analysis," *IEEE Trans. on Control Systems Technology*, vol. 23, no. 2, pp. 770–777, March 2015.
[43] G. Vinnicombe, *Uncertainty and Feedback: H_∞ Loop-Shaping and the ν-Gap Metric*. Imperial College Press, 2001.
[44] S. Skogestad and I. Postlethwaite, *Multivariable Feedback Control: Analysis and Design*. Wiley, 2005.
[45] P. A. Iglesias, "The strictly proper discrete-time controller for the normalized left-coprime factorization robust stabilization problem," *IEEE Trans. on Automatic Control*, vol. 45, no. 3, pp. 516–520, March 2000.
[46] B. Friedland, *Control System Design: An Introduction to State-Space Methods*. McGraw-Hill, 1986.
[47] A. Raghunathan and S. Di Cairano, "Infeasibility detection in alternating direction method of multipliers for convex quadratic programs," in *IEEE Conference on Decision and Control*, 2014.
[48] S. Di Cairano, M. Brand, and S. A. Bortoff, "Projection-free parallel quadratic programming for linear model predictive control," *International Journal of Control*, vol. 86, no. 8, pp. 1367–1385, Aug. 2013.
[49] R. Fletcher, *Practical Methods of Optimization*. John Wiely & Sons, 2013.

APPENDIX

The design equations for the $\mathcal{H}_\infty$ loop-shaping and inverse optimal controller [27] are included for completeness. For the strictly proper shaped plant

$$x_s(k+1) = A_s x_s(k) + B_s u_s(k)$$
$$y_s(k) = C_s x_s(k)$$

compute solutions X_s and Z_s to the pair of Riccati equations

$$A_s^T X_s B \left(B_s^T X_s B_s + I\right)^{-1} B_s^T X_s A_s - A_s^T X_s A_s + X_s = C_s^T C_s$$
$$A_s Z_s C_s^T \left(C_s Z_s C_s^T + I\right)^{-1} C_s Z_s A_s^T - A_s Z_s A_s^T + Z_s = B_s B_s^T.$$

Stabilizing solutions exist if $(A_s,\, B_s)$ is stabilizable and $(A_s,\, C_s)$ is detectable. Compute $\gamma_{min} = \sqrt{1 + \lambda_{max}(X_s Z_s)}$, where λ_{max} is the maximum eigenvalue. For $\gamma_{rel} > 1$ let $\gamma = \gamma_{rel} \cdot \gamma_{min}$. Then $\epsilon = 1/\gamma$ and

$$W_s = (\gamma^2 - 1) I - Z_s X_s$$
$$K_s = -\gamma^2 B_s^T X_s W_s^{-1} \left(I + \gamma^2 B_s B_s^T X_s W_s^{-1}\right)^{-1} A_s$$
$$H_s = -A_s Z_s C_s^T \left(I + C_s Z_s C_s^T\right)^{-1}$$
$$Q_s = \gamma^2 X_s W_s^{-1} - \gamma A_s^T \left(I + \gamma^2 X_s W_s^{-1} B_s B_s^T\right)^{-1} X_s W_s^{-1} A_s$$
$$R_s = I.$$

Scott A. Bortoff (SM'11) received the B.S. and M.S. degrees from Syracuse University in 1985 and 1986, respectively, and the Ph.D. degree from the University of Illinois at Urbana-Champaign in 1992, all in Electrical Engineering. He held positions of Assistant and Associate Professor of Electrical and Computer Engineering at the University of Toronto, (1992-2000), Group and Project Leaderships in the area of control systems at United Technologies Research Center (2000-2009), and Group Manager of Mechatronics at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA (2009-2016), where he is currently Distinguished Research Scientist and Special Project Leader. His research interests include system-level dynamic modeling and control of thermofluid and mechatronic systems. He is currently Associate Editor of the IEEE Control System Magazine and serves the Modelica community in various capacities.

Paul Schwerdtner received his Master's degree in Engineering Science in '18 from Technical University Berlin, where he is currently working towards his Ph.D. at the chair for numerical mathematics. His principal research interests include robust control and reduced order modeling of large-scale systems. His primary focus is on structure-preserving methods that allow to maintain and exploit underlying physical properties of dynamical systems during controller synthesis and model order reduction.

Claus Danielson (SM'19) is an Assistant Professor in the Department of Mechanical Engineering at the University of New Mexico. He received a Ph.D. (2014) degree from the University of California, Berkeley. He received a master degree (2005) and a bachelor degree (2003) from Rensselaer Polytechnic Institute and the University of Washington, respectively. He was previously a Principle Research Scientist at Mitsubishi Electric Research Laboratories. His research interests are in motion planning and constrained control. His specialty is developing methods for exploiting structure in large-scale or complex planning, control, and optimization problems. He has applied his research to a variety of fields including autonomous vehicles, robotics, spacecraft guidance and control, heating ventilation and air conditioning, energy storage networks, adaptive optics, atomic force microscopy, and cancer treatment.

Stefano Di Cairano (SM'15) received the Master's (Laurea) and the Ph.D. degrees in information engineering in '04 and '08, respectively, from the University of Siena, Italy. During '08-'11, he was with Powertrain Control R&A, Ford Research and Advanced Engineering, Dearborn, MI, USA. Since 2011, he is with Mitsubishi Electric Research Laboratories, Cambridge, MA, USA, where he is currently a Distinguished Research Scientist, and the Senior Team Leader. His research focuses on optimization-based control and decision-making strategies for complex mechatronic systems, in automotive, factory automation, transportation systems, and aerospace. His research interests include model predictive control, constrained control, motion planning, hybrid systems, optimization, and particle filtering. He has authored more than 200 peer-reviewed papers in journals and conference proceedings, and is named an inventor in more than 60 patents. Dr. Di Cairano was the Chair of the IEEE CSS TC Automotive Controls and of the IEEE CSS SC on Standards, and an Associate Editor of the IEEE TRANS. CONTROL SYSTEMS TECHNOLOGY,. He is the inaugural Chair of the IEEE CSS Technology Conferences Editorial Board, and an Associate Editor of the IEEE TRANS. INTELLIGENT VEHICLES.

Daniel J. Burns (M'10-SM'18) Daniel Burns received the M.S. and Ph.D. degrees in mechanical engineering from the Massachusetts Institute of Technology, Cambridge, in 2006 and 2010, respectively. From '10-'19, he was with Mitsubishi Electric Research Laboratories (MERL), Cambridge, MA, as a Senior Principal Research Scientist. At MERL, Dr. Burns developed and prototyped advanced control methods for vapor compression systems. His research interests include multi-physical modeling of mechatronic and thermodynamic systems, extremum seeking control, and applied predictive and adaptive control. Before joining MERL, he worked on flight instrumentation at the Commercial Aviation Systems division of Honeywell, Inc. (Phoenix, AZ) and NASA's Goddard Space Flight Center (Greenbelt, MD).