Received: Added at production | Revised: Added at production | Accepted: Added at production

DOI: xxx/xxxx

**ARTICLE TYPE**

# Computationally Efficient Integrated Design and Predictive Control of Flexible Energy Systems using Multi-fidelity Simulation-based Bayesian Optimization

Farshud Sorourifar† | Naitik Choksi† | Joel A. Paulson*

Department of Chemical & Biomolecular Engineering, The Ohio State University, Ohio, USA

†F. Sorourifar and N. Choksi contributed equally to this work.

**Correspondence**
*J. A. Paulson, 416 CBEC, 152 W. Woodruff Avenue, Columbus, OH 43210. Email: paulson.82@osu.edu

**Abstract**

We present a multi-fidelity black-box optimization approach for integrated design and control (IDC) of constrained nonlinear systems in the presence of uncertainty. The IDC framework is becoming increasingly important for the systematic design of next-generation (flexible) manufacturing and energy systems. However, identifying optimal solutions to realistic IDC problems is intractable when (i) the dynamics occur on much shorter timescales than the system lifetime, (ii) the uncertainties are described by continuous random variables with high variance, and (iii) operational decisions involve a mixture of discrete and continuous variables. Instead of aggressively simplifying the problem to improve tractability, we develop a simulation-based optimization procedure using high-quality decision rules that map information that can be measured online to optimal control actions. In particular, we rely on the Bayesian optimization (BO) framework that has been shown to perform very well on noisy and expensive-to-evaluate objective functions. We also discuss how BO can be extended to take advantage of computationally cheaper low-fidelity approximations to the high-fidelity IDC cost function. Three major low-fidelity approximation strategies are described in this work, which are related to the simplification of the system simulator, decision rule solution method, and time grid. Lastly, we demonstrate the advantages of multi-fidelity BO on the design of a solar-powered building heating/-cooling system (with battery and grid support) under uncertain weather and demand conditions with hourly variation over a year-long planning horizon.

**KEYWORDS:**
Optimization methods; uncertain systems; predictive control

# 1 | INTRODUCTION

With the significant increase in global energy demand and the advent of next-generation manufacturing and energy systems, such as combined heat and power (CHP) plants [1], smart grids [2], and multi-product chemical plants [3], the necessity for a paradigm shift in energy production and operations is paramount. This challenge is intensified by environmental concerns resulting from the commonplace use of traditional energy sources, like fossil fuels and its derivatives, which are large contributors to climate change [4]. To enable this shift into the next-generation, we need to able to design flexible energy production and manufacturing processes that are to respond/adapt to highly dynamic and uncertain conditions in an optimal manner. An effective design strategy must be able to accommodate the following features that are known to complicate optimal design in the presence of realistic models of flexible energy and manufacturing systems:

1. Relevant nonlinear dynamics and uncertainties can occur on much shorter timescales than the lifetime of the system. For example, solar and wind energy profiles, electricity prices, and load demand are unpredictable quantities that can substantially vary on the minute to hour timescale, whereas design decisions last for years or even decades.

2. Certain uncertainties are best described by continuous random variables with large variance including product demand, processing times, and market prices. Considering only a few discrete scenarios may lead to overly optimistic predictions of the performance (depending on how recourse decisions are represented).

3. Many key operational decisions are discrete including adaptive scheduling, unit commitment, and product selection. Since these control actions can impact many variables, especially as processes become more integrated, this implies we need to rely on advanced optimization-based control strategies to derive reasonably accurate predictions of overall system performance.

Systematically accounting for these features requires an integrated design and control (IDC) approach. In current practice, design is typically tackled first, which has been known to produce heavily constrained processes with few degrees of freedom left for control purposes [5]. Although it has been established that optimal IDC problems can be generally formulated as multistage stochastic programs (MSPs) [6], the MSP will be intractable whenever the features presented above are present in the formulation. This is due to the fact that features 1 and 2 imply that a large number of operational periods and uncertainty scenarios need to be considered in the MSP, which are known to suffer from the curse-of-dimensionality [7]. Feature 3, on the other hand, implies that the system model is nonlinear and non-smooth, which generally results in a mixed-integer nonlinear programming (MINLP). Putting these features together, we can deduce that scenario-based approximations to the MSP result in extremely large-scale MINLPs (that co-optimize the design and control decisions for every time stage under every uncertainty scenario) that are far beyond the capabilities of existing solvers (and even specialized solution methods).

There is a large body of literature on different ways of addressing the IDC problem; a detailed summary of a variety of different contributions is presented in Appendix A. These methods have certain shortcomings which prevent them from considering long operational periods or multiple uncertainty scenarios and allow only linear or steady-state models to approximate control metrics. To overcome the limitations of previous methods, recent work has proposed the use of derivative-free simulation optimization (SO) methods to tackle IDC formulations that consider all three features discussed above. The SO paradigm relies on an *outer* optimization of the relevant design variables and an *inner* stochastic dynamic simulation that is used to evaluate the expected long-term operating costs appearing in the outer problem. To ensure a well-defined inner simulation, we must specify a so-called *decision rule* (DR), which can be loosely interpreted as a function that maps the measured data to the control inputs in every operational period. In light of feature 3, it is critical to rely on high-quality DR that can flexibly handle scheduling and control decisions. Model predictive control (MPC) [8] is an optimization-based control strategy that is widely used for controlling complex, large-scale systems that have significant multivariable interactions and safety-critical and/or quality constraints. There has been a significant amount of work on the development of specialized MPC methods in the context of energy systems [9,10,11,12], including recent contributions that directly incorporate nonlinear effects [13,14] and online adaptation of the internal model [15]. In our previous work [16], we showed how a mixed-integer extension of MPC could be used as a much higher-quality DR than simplified PID or linear MPC structures. Regardless of the specific controller structure that is selected, there remain some "tuning parameters" that can have a strong influence on performance and must be adequately chosen to ensure good closed-loop performance. We have demonstrated that the Bayesian optimization (BO) methodology is able to effectively *co-optimize the design and controller tuning variables* using significantly fewer closed-loop simulations than alternative SO methods, even when we can only obtain noisy observations of the cost function. The key advantage of this proposed BO framework is that it can be generally applied to any system model (including computationally expensive models defined over multiple space/time scales) and any controller structure (including any of the advanced forms of MPC mentioned above or innovations developed in the future).

Although we obtained good results with the standard BO approach on problems with a reasonable number of time steps, it still required on the order of 30 high-fidelity evaluations on an IDC problem with a 4-dimensional design space. When only a limited computational budget is available, the maximum number of evaluations may be as few as 5-10, especially as the planning horizon increases to the yearly scale. As discussed in Appendix A, we can develop a wide-variety of approximations to this high-fidelity simulator that can be evaluated in significantly less time. Thus, our main motivation in this work is the following question: *Can we use these computationally cheaper approximations to more effectively guide the BO search process?* It turns out the answer to this question is yes, as we demonstrate in this paper using recent developments in *multi-fidelity* BO. In particular, we study the multi-fidelity Gaussian process upper confidence bound (MF-GP-UCB) algorithm recently proposed by Kandasamy et al. [17]. To apply MF-GP-UCB, we need to develop a sequence of lower-fidelity approximations to our high-fidelity MPC-based IDC problem. We propose three novel strategies for deriving such approximations that can be categorized based on how they

reduce computational cost. The first approach involves building a surrogate model for the system dynamics, which is valuable whenever the process simulator is itself expensive. The second approach is related to simplification of the MPC-based DR that can be quite expensive depending on the complexity of the underlying optimization problem. The third and final approach is based on coarsening the time grid, so that the simulation can be completed in fewer time steps. To demonstrate the advantages of MF-GP-UCB over single-fidelity BO, we apply these methods to the design of a solar-powered building heating and cooling system, with grid and battery support, under uncertainty weather and demand conditions. To ensure the system is realistically modeled, we consider disturbance profiles for Columbus, Ohio that vary on the hour scale over a year-long horizon.

The rest of the paper is organized as follows. In Section 2, we introduce our proposed formulation of the IDC problem that is based on a high-quality estimate of the operating cost found by running detailed closed-loop simulations with control decisions made by an advanced MPC controller. In Section 3, we summarize our previous work on efficiently tackling this challenging problem using BO. Section 4 provides an overview of the MF-GP-UCB algorithm, while Section 5 presents the three different proposed approximation strategies for MPC-based IDC. The proposed algorithm is applied to the solar-powered building heating/cooling system case study in Section 6 and we conclude the article in Section 7.

## 2 | PROBLEM FORMULATION: INTEGRATED DESIGN AND PREDICTIVE CONTROL

In this work, we are interested in integrated design and control (IDC) problems that can be generally stated in terms of the following stochastic optimization problem

$$\min_{(d,z)\in D\times\mathcal{Z}} \; C(d) + \mathbb{E}_{\boldsymbol{\omega}}\{O(d,z,\boldsymbol{\omega})\}, \tag{1}$$

where $d \in \mathbb{R}^{n_d}$ are the design variables that are restricted to the known compact set $\mathcal{D} \subset \mathbb{R}^{n_d}$; $z \in \mathbb{R}^{n_z}$ are parameters related to the control decisions that are restricted to the known compact set $\mathcal{Z} \subset \mathbb{R}^{n_z}$; $\boldsymbol{\omega} \in \mathbb{R}^{n_\omega}$ is the set of uncertain variables that are modeled as a random vector with known probability density function $p_{\boldsymbol{\omega}} : \Omega \to \mathbb{R}_+$ and finite support $\Omega$; $C : \mathcal{D} \to \mathbb{R}$ is the capital cost function that depends only on the design variables; $O : \mathcal{D} \times \mathcal{Z} \times \Omega \to \mathbb{R}$ is the operating cost function that can generally depend on the realization of the design variables, control parameters, and uncertainties; and $\mathbb{E}_{\boldsymbol{\omega}}\{\cdot\}$ denotes the expected value with respect to $\boldsymbol{\omega}$. The high-level representation (1) can represent a large class of IDC problems through the proper specification of the variables $\{d, z, \boldsymbol{\omega}\}$ and functions $\{C(\cdot), O(\cdot), p_{\boldsymbol{\omega}}(\cdot)\}$. Furthermore, the best choice of algorithm will heavily depend on the properties of these variables and functions. Two of the most aggressive simplifications would involve assuming the uncertain variables are deterministic with known values, i.e., $p_{\boldsymbol{\omega}}(\boldsymbol{\omega}) = \delta(\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})$ and assuming the system remains at a fixed steady state so that the operating cost could easily be computed by multiplying the steady-state cost times the number of time stages. Although this would greatly simplify the problem and allow us to use efficient solvers, it clearly would miss the

key aspects discussed in the introduction. In particular, our model would completely miss out on the adaptation that is inherent to any flexible system due to unexpected changes in the market, environment, and equipment. In other words, the estimated operating cost under these approximations is likely to be highly unrealistic, which leads to potentially poor design choices, even when a global solution to the problem can be found.

Instead of pursuing other more detailed approximations, we would like to develop an optimization algorithm that can account for as accurate evaluations of the operating cost $\mathbb{E}_{\boldsymbol{\omega}}\{O(d, z, \boldsymbol{\omega})\}$ as possible – at least as accurate as we can expect given our most detailed model of the system dynamics. We assume the system dynamics can be represented by the following general time-varying nonlinear system under uncertainty

$$x_{t+1}(d, z, \boldsymbol{\omega}) = h_t(x_t(d, z, \boldsymbol{\omega}), u_t(d, z, \boldsymbol{\omega}), w_t, d), \qquad x_0(d, z, \boldsymbol{\omega}) = b_0(d), \tag{2}$$

where $t \in \mathcal{T} := \{0, \dots, T-1\}$ is the discrete time index that ranges over a finite number of $T$ time stages; $\boldsymbol{\omega} = \{w_0, \dots, w_{T-1}\}$ is the sequence of random disturbances $w_t \in \mathcal{W} \subset \mathbb{R}^{n_w}$ that are assumed to be measurable at each $t \in \mathcal{T}$; $u_t(d, z, \boldsymbol{\omega}) \in \mathcal{U} \subset \mathbb{R}^{n_u}$ are the control inputs that are assumed to be parametrized by $z$; and $x_t(d, z, \boldsymbol{\omega}) \in \mathcal{X} \subset \mathbb{R}^{n_x}$ are the system state variables with initial condition $b_0(d)$ that can depend on the design variables. It is important to note we do not make any assumptions about the function $h_t : \mathcal{X} \times \mathcal{U} \times \mathcal{W} \times \mathcal{D} \to \mathcal{X}$. Our main assumption is that we are able to *simulate* the system for any particular $\{d, z, \boldsymbol{\omega}\}$ to compute the future state sequence $\{x_0(d, z, \boldsymbol{\omega}), \dots, x_T(d, z, \boldsymbol{\omega})\}$. This implies that (2) can be any *black-box simulator* derived from, e.g., open-source or proprietary computer codes that are being increasingly used to reduce model development time and to account for multi-scale phenomena. To achieve the goal of an accurate representation of the operating cost, we clearly need to incorporate a *realistic* control policy into our system model that is able to make "optimal" decisions at any given time given the most recently available data. In this work, we represent the control policy as follows

$$u_t(d, z, \boldsymbol{\omega}) = \kappa_t(x_t(d, z, \boldsymbol{\omega}), w_t, z), \tag{3}$$

where $\kappa_t : \mathcal{X} \times \mathcal{W} \times \mathcal{Z} \to \mathcal{U}$ has a fixed structure (with tuning parameters $z$) for all $t \in \mathcal{T}$. We can interpret (3) as a parametrized *decision rule* (DR) that directly enforces causality (i.e., cannot make use of knowledge of future disturbances). For the structure of $\kappa_t(\cdot)$, we make use of model predictive control (MPC), which is the most widely-used methods for advanced control of complex systems with multivariate interactions and safety-critical and/or quality constraints. Even though MPC has been traditionally applied to steady-state tracking problems, recent advances have enabled its application to a significantly broader class of systems including those with nonlinear models, constraints, integer variables, and objectives that can be used to handle much more complicated physical equations as well as economic cost functions. Due to the black-box nature of the methodology pursued in this paper, we can straightforwardly accommodate any of these advanced forms of MPC. A complete mathematical description of the specific MPC formulation used in our case study is provided in Section 6.

Given the closed-loop system in (2) and (3), we can develop a much more accurate representation of the operating cost over the system lifetime $T$ as follows

$$O(d, z, \boldsymbol{\omega}) = \sum_{t=0}^{T-1} \ell_t(x_t(d, z, \boldsymbol{\omega}), u_t(d, z, \boldsymbol{\omega}), w_t, d) + \ell_T(x_T(d, z, \boldsymbol{\omega}, \theta), d), \tag{4}$$

where $\ell_t : \mathcal{X} \times \mathcal{U} \times \mathcal{W} \times \mathcal{D} \to \mathbb{R}$ is the time-varying stage cost function that represents the individual contribution to the operating cost from time step $t \in \mathcal{T}$ and $\ell_T : \mathcal{X} \times \mathcal{D} \to \mathbb{R}$ is the final cost function that represents the cost of the state at the end of the system lifetime. Under this choice of operating cost (4), the IDC problem (1) is intractable for practical problems of interest in this work for three reasons. First, the expectation operator $\mathbb{E}_{\boldsymbol{\omega}}\{\cdot\}$ cannot be exactly computed for general nonlinear functions and continuous random variables. Second, relevant operational details and realizations of the uncertainty often occur at timescales that are significantly shorter than the lifetime of the system, which implies that $T$ can be very large (e.g., $T = 219,000$ when decisions need to be made at each hour for 25 years). Lastly, the dynamics may involve a mixture of continuous and discrete decisions, which implies (2) is a hybrid system that is not continuously differentiable and thus not amenable to derivative-based solvers. In our recent work [16], we demonstrated how Bayesian optimization can help overcome these limitations by simultaneously sampling values of $\{d, z\}$ (co-optimizing design and MPC tuning) in an intelligent fashion to search for the global solution to (1) as quickly as possible. After summarizing this idea in the next section, we extend the concept to account for multiple cheaper approximate models to better guide the search in Sections 4 and 5.

## 3 | PRACTICAL IDC WITH MPC USING SIMULATION-BASED BAYESIAN OPTIMIZATION

### 3.1 | Overview of Derivative-free Simulation Optimization

Even though the IDC problem (1), with detailed operating cost (4) evaluated via a closed-loop simulation with an embedded MPC controller, has a relatively small number of time-invariant decision variables, it remains intractable due to the stochastic uncertainties as well as the complexity and potential non-smoothness in the system model (2) for which we may not be able to estimate derivatives of the cost with respect to $d$ and $z$, if they even exist. One of the main alternatives, when gradient information is not readily available, is to rely on so-called derivative-free optimization (DFO) methods [18,19] (also known as black-box optimization) that can be broadly categorized as stochastic or deterministic. Before selecting a specific DFO method, let us first convert (1) into a more standard form, as follows

$$f^\star = \max_{\theta \in \Theta} f(\theta) := \mathbb{E}_{\boldsymbol{\omega}}\{J(\theta, \boldsymbol{\omega})\}, \tag{5}$$

where $\theta = [d^\top, z^\top]^\top \in \Theta = \mathcal{D} \times \mathcal{Z} \subset \mathbb{R}^{n_\theta}$ is the concatenated vector of design and MPC tuning parameters with $n_\theta = n_d + n_z$ being the total number of decision variables; $J(\theta, \boldsymbol{\omega}) = -(C(d) + O(d, z, \boldsymbol{\omega}))$ is the performance function defined as the negative

of the total capital plus operating cost of the system; and $f^\star = f(\theta^\star)$ is the true (unknown) maximum performance value corresponding to the (unknown) global solution $\theta^\star \in \text{argmax}_{\theta\in\Theta} f(\theta)$. This problem can be interpreted in the simulation-optimization context as an "outer" optimization over $\theta$ and an "inner" stochastic simulation needed to approximate $f(\theta)$.

Due to the expensiveness of the long-time closed-loop simulations, we are interested in finding $\theta^\star$ using as few simulations as possible, which requires careful selection of a DFO method. Stochastic DFO methods are based on evolutionary-like algorithms, such as genetic algorithm (GA)[20] and particle swarm optimization (PSO)[21], which are known to require a large number of function evaluations as they search for the global optimum. This suggests stochastic DFO is not a viable choice for MPC-based IDC problems since each function evaluation is computationally intensive. Deterministic DFO methods, on the other hand, are often motivated by the optimization of expensive objective functions and can be divided into *direct search* or *model-based* approaches. The main difference between these categories are that direct search methods use the evaluations of the objective directly to determine the next search direction whereas model-based methods construct a surrogate model to better guide the search process. Whenever an accurate surrogate model can be constructed, model-based DFO methods have the ability to converge to nearly global solutions in a limited number of iterations.

A wide-variety of model-based DFO methods have been proposed in the literature, which mainly differ by the choice of *scale* (local versus global approximation) and the *type* of surrogate model (e.g., polynomial, neural network, linear interpolation). A key challenge in these approaches is the selection of the "right" type of surrogate model when little is known about the structure of the objective. Gaussian process (GP) models[22] are a particularly attractive class of surrogates since they are both *probabilistic* and *non-parametric*. Given a set of function evaluations, a GP model can be easily derived by placing a prior over the set of possible objective functions and updating this prior with the available data using Bayes' rule. The Bayesian optimization (BO) framework[23] leverages a GP model of the objective function and an expected utility (or acquisition) function, defined in terms of the GP-predicted posterior distribution, to sequentially select the next evaluation point as the one that maximizes the chosen acquisition function. As discussed next, this is our preferred method since it not only addresses the exploration/exploitation tradeoff, but also handles noisy estimates of the objective function, which is of critical importance in IDC problems due to the presence of the high-dimensional uncertainty $\boldsymbol{\omega}$.

## 3.2 | The Bayesian Optimization Framework

We now introduce the BO framework, which is our preferred DFO method. The main idea underlying BO is to place a GP prior on the objective function $f(\cdot) \sim \mathcal{GP}(m(\cdot), k(\cdot,\cdot))$, which is fully specified by its mean function $m(\cdot)$ and covariance kernel $k(\cdot,\cdot)$. In essence, GPs are an uncountable collection of random variables of which any finite subset have a joint Gaussian distribution, i.e., generalizes the notion of multivariate Gaussian distributions to "distributions over functions"[22]. The mean function and

covariance kernel are defined such that, for any pairs of input points $\theta, \theta' \in \mathbb{R}^{n_\theta}$, we have

$$m(\theta) = \mathbb{E}_f\{f(\theta)\}, \tag{6a}$$

$$k(\theta, \theta') = \mathbb{E}_f\{(f(\theta) - m(\theta))(f(\theta') - m(\theta'))\}, \tag{6b}$$

where the expectation is taken over the function space. To update the GP prior, we must evaluate the objective function at specific values of $\theta$. Since we cannot evaluate the expectation in (5) exactly, we must rely on Monte Carlo (MC) estimates as follows

$$y = \frac{1}{K}\sum_{i=1}^{K} \boldsymbol{J}(\theta, \boldsymbol{\omega}^{(i)}), \tag{7}$$

where $\mathbf{W} = \{\boldsymbol{\omega}^{(1)}, \dots, \boldsymbol{\omega}^{(K)}\}$ is a set of $K$ independent and indentically distributed (i.i.d.) samples of the random vector $\boldsymbol{\omega} \sim p_{\boldsymbol{\omega}}$. The estimator (7) is unbiased for any positive integer value of $K$, which implies that $\mathbb{E}_{\mathbf{W}}\{y\} = f(\theta)$ such that we can develop a relatively simple model the noisy measurements of the objective as follows

$$y = f(\theta) + \epsilon, \tag{8}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ is a normal random variables with mean zero and variance $\sigma_\epsilon^2$. Assume we have $n$ noisy objective function evaluations from (8) denoted by $y_{1:n} = \{y_1, \dots, y_n\}$ computed at corresponding inputs $\theta_{1:n} = \{\theta_1, \dots, \theta_n\}$. Based on the GP prior assumption, $y_{1:n}$ and $f(\theta)$ at any test point $\theta$ are jointly Gaussian with the following distribution

$$\begin{bmatrix} y_{1:n} \\ f(\theta) \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} m(\theta_{1:n}) \\ m(\theta) \end{bmatrix}, \begin{bmatrix} k(\theta_{1:n}, \theta_{1:n}) + \sigma_\epsilon^2 I_n & k(\theta_{1:n}, \theta) \\ k(\theta, \theta_{1:n}) & k(\theta, \theta) \end{bmatrix}\right), \tag{9}$$

where, in a standard fashion, we have overloaded the functions $m(\cdot)$ and $k(\cdot, \cdot)$ to include element-wise operations across their inputs. Using the Gaussian conditional formula, the posterior distribution of $f(\theta)$ at any test point $\theta$, conditioned on the data $\{y_{1:n}, \theta_{1:n}\}$, can be analytically determined to be[22]

$$f(\theta)|y_{1:n}, \theta_{1:n} \sim \mathcal{N}(\mu_n(\theta), \sigma_n^2(\theta)), \tag{10}$$

with

$$\mu_n(\theta) = m(\theta) + k(\theta_{1:n}, \theta)\left(k(\theta_{1:n}, \theta_{1:n}) + \sigma_\epsilon^2 I_n\right)^{-1}(y_{1:n} - m(\theta_{1:n})), \tag{11a}$$

$$\sigma_n^2(\theta) = k(\theta, \theta) - k(\theta, \theta_{1:n})\left(k(\theta_{1:n}, \theta_{1:n}) + \sigma_\epsilon^2 I_n\right)^{-1} k(\theta_{1:n}, \theta). \tag{11b}$$

In practice, the mean function and covariance kernel are parametrized by some unknown hyperparameters. These hyperparameters, along with the unknown measurement noise variance $\sigma_\epsilon^2$, can be calibrated to the available data $\{y_{1:n}, \theta_{1:n}\}$ using maximum likelihood estimation (MLE), as discussed by, e.g., Bradford et al.[24].

Given this posterior distribution, we can define the acquisition function $\alpha_n : \Theta \to \mathbb{R}$ that plays a key role in BO, as it must be designed in a way that tradeoffs *exploration* of regions where the surrogate model is uncertain and *exploitation* of the model's high-confidence predictions of good solutions. A popular choice is the expected improvement (EI) over a candidate point[25], which can be expressed analytically in terms of the posterior mean and variance in (11). A recent alternative, that will be particularly relevant in this work, is the upper confidence bound (UCB) that can be defined as follows[26]

$$\alpha_n(\theta) = \text{UCB}_n(\theta) = \mu_n(\theta) + \sqrt{\beta_n}\sigma_n(\theta), \tag{12}$$

where $\beta_n > 0$ is some user-specified constant at each iteration $n$ (typically chosen to be $\beta_n = 0.2n_\theta \log(2(n-1))$[27]). Given these choices, we can summarize the GP-UCB procedure (a particular instance of the BO framework) in Algorithm 1. The main intuition behind this algorithm is that the mean $\mu_n(\theta)$ encourages querying where we know $f(\theta)$ is high, while the standard deviation $\sigma_n(\theta)$ encourages querying at regions that we are most uncertain about $f(\theta)$. As such, the factor $\beta_n$ directly controls the tradeoff between exploration and exploitation. The performance of this algorithm has been studied on a variety of problems and its convergence has been analyzed in Srinivas et al.[26]. In particular, bounds on the simple regret $S_n$

$$S_n = \min_{i\in\{1,\dots,n\}} (f^\star - f(\theta_i)) = f^\star - \max_{i\in\{1,\dots,n\}} f(\theta_i), \tag{13}$$

have been established that show $S_n \sim O(1/\sqrt{n})$, which implies GP-UCB converges to function evaluations that are close to $f^\star$ as the number of iterations increases.

**Algorithm 1** The GP-UCB algorithm under the BO framework[26]

1: **Initialize:** Input space $\Theta$; GP prior $m(\cdot)$ and $k(\cdot,\cdot)$; and maximum number of iterations $N$.
2: **for** $n = 0$ to $N-1$ **do**
3: Construct GP surrogate model for $f(\theta)$ given available data $\{y_{1:n}, \theta_{1:n}\}$ via (10) and (11).
4: Maximize the acquisition function to find $\theta_{n+1} = \text{argmax}_{\theta\in\Theta}\, \mu_n(\theta) + \sqrt{\beta_n}\sigma_n(\theta)$.
5: Perform expensive closed-loop simulations to evaluate objective $y_{n+1} = f(\theta_{n+1}) + \epsilon_{n+1}$ using (7).
6: **end for**

*Remark 1.* It should be noted that (8) is a simplified representation of the measurement process – although the random variable $\epsilon$ does have zero mean, it may not be normally distributed with constant variance for finite $K$. The estimator (7) does, however, satisfy the central limit theorem (CLT):

$$\sqrt{K}(y - f(\theta)) \Rightarrow \mathcal{N}(0, \sigma_f^2(\theta)), \tag{14}$$

where $\sigma_f^2(\theta)$ denotes the variance of the objective function in (5) for any fixed $\theta \in \Theta$ and $\Rightarrow$ denotes convergence in distribution. Thus, we see that the measurement error process is *asymptotically* normal with variance that converges to zero as $K \to \infty$. Due to the expensiveness of the closed-loop simulations, however, we must select $K$ to be very low, meaning we are not able to

invoke the CLT (14) in practice. As long as the effect of the uncertainty $\boldsymbol{\omega}$ is not too large, we have found that the assumption (8) works well in practice, which was the case for our case study. In our future work, we plan to more systematically address this issue by developing more detailed input-dependent (also known as heteroscedastic) noise models[28].

## 4 | REDUCING SIMULATION COST USING MULTI-FIDELITY BAYESIAN OPTIMIZATION

The main bottleneck in Algorithm 1 is that $K$ computationally expensive closed-loop simulations are required at every iteration. An interesting idea for alleviating this bottleneck is the use of series of approximate models for the objective, as opposed to a single high-fidelity model. This is often referred to as *multi-fidelity* optimization in the literature and has been successfully applied in the context of automated machine learning[29,30,31]. Since the approximations are still correlated to the high-fidelity evaluation, we can reasonably expect that these models may provide valuable information at a fraction of the cost, which can be used to avoid "wasting" very expensive function evaluations on particularly poor designs. This discussion motivates the following two questions, which we look to (partially) answer in the next two sections:

1. Can we come up with a systematic procedure to decide at what fidelity and location we should sample?
2. How should we construct lower-fidelity representations of the IDC performance function in (5)?

We address the first question using a recent extension to the GP-UCB algorithm, which we summarize in this section; the second question is discussed further in Section 5.

### 4.1 | Setting up the Multi-Fidelity Problem

The main difference from the traditional setup in Section 3.2 is that we assume access to a set of $M-1$ successively accurate approximations, which we denote by $f^{(1)}, f^{(2)}, \dots, f^{(M-1)}$, to the true function of interest $f^{(M)} = f$. Following Kandasamy et al.[17], these approximations (also known as fidelities) must satisfy two important conditions:

1. The functions $f^{(1)}, \dots, f^{(M-1)}$are approximations of $f^{(M)}$ with bounded error that successively improves, i.e.,

$$\| f^{(m)} - f^{(M)} \|_\infty \leq \zeta^{(m)}, \quad \forall m \in \{1, \dots, M\}, \tag{15}$$

   where the bounds $\zeta^{(1)} > \zeta^{(2)} > \cdots > \zeta^{(M)} = 0$ are known.

2. The functions $f^{(1)}, \dots, f^{(M-1)}$ are cheaper to evaluate than $f^{(M)}$, i.e., $0 < \lambda^{(1)} < \lambda^{(2)} < \cdots < \lambda^{(M)}$ where $\lambda^{(m)}$ denotes the computational cost of querying at fidelity $m \in \{1, \dots, M\}$.

Roughly speaking, these two conditions state that the approximations should become both more accurate and more costly, as the level $m$ increases. As opposed to just sampling $\{\theta_n\}_{n\geq 0}$, the multi-fidelity version of the algorithm must determine a sequence

of query-fidelity pairs $\{(\theta_n, m_n)\}_{n\geq 0}$ where, at any given time $n$, the algorithm can use information from the previous $n-1$ query-observation-fidelity triples, i.e., $\{(\theta_i, m_i, y_i)\}_{i=1}^{n-1}$. Note that, similarly to (8), the observations are modeled as

$$y_i = f^{(m_i)}(\theta_i) + \epsilon_i, \tag{16}$$

where $\epsilon_i$ are independent noise realizations at every iteration $i$, with $\mathbb{E}\{\epsilon_i\} = 0$. Note that the cases of interest in this paper are for $M$ values that are fixed and relatively small (in the range of 2 to 4) and $\lambda^{(1)}$ values that are comparable to $\lambda^{(M)}$ – this implies that the approximations to $f(\theta)$ remain fairly expensive and still require an intelligent BO-like procedure to optimize.

Let $\Lambda$ denote the maximum allowed resources that can be used by the multi-fidelity optimization procedure. The number of iterations taken until the resources have been exhausted can be inferred from the evaluation cost of each fidelity above as follows

$$N = \max\{n \geq 1 : \textstyle\sum_{i=1}^{n} \lambda^{(m_i)} \leq \Lambda\}, \tag{17}$$

for a given set of data $\{(\theta_i, m_i, y_i)\}_{i\geq 0}$ produced by some algorithm. It is important to note that $N$ is an implicit function of the initial data in this case. This means $N$ cannot be computed *a priori* and will be a *random* variable whenever the algorithm is seeded with some randomly selected initial function evaluations.

## 4.2 | Multi-Fidelity Bayesian Optimization with Upper Confidence Bounds

Now, we discuss the extension of the GP-UCB algorithm to the multi-fidelity setting based on the algorithm presented by Kandasamy et al.[17]. The main idea is to maintain an upper confidence bound for $f^{(M)}$ using the data available at *all* fidelity levels. Due to the constraints (15), the posterior for any $f^{(m)}$ conditioned on all available data is not Gaussian. Let $\mu_n^{(m)}(\theta)$ and $\sigma_n^{(m)}(\theta)$ denote the posterior GP mean and standard deviation for $f^{(m)}$ conditioned on *only* the previous data points available at the $m^{\text{th}}$ fidelity. For a reasonably chosen $\beta_n$ value, we know that $\mu_n^{(m)}(\theta) + \sqrt{\beta_n}\sigma_n^{(m)}(\theta)$ will upper bound $f^{(m)}(\theta)$ with high probability. Combining this with the bounds in (15), we know

$$\varphi_n^{(m)}(\theta) = \mu_n^{(m)}(\theta) + \sqrt{\beta_n}\sigma_n^{(m)}(\theta) + \zeta^{(m)}, \quad \forall m \in \{1, \dots, M\}, \tag{18}$$

represent a set of $M$ upper bounds for $f^{(M)}$. The best upper bound is then given by

$$\varphi_n(\theta) = \min_{m\in\{1,\dots,M\}} \varphi_n^{(m)}(\theta). \tag{19}$$

We use this upper confidence bound for our acquisition function, in place of the traditional single-fidelity GP-UCB in (12), meaning that our next query point is at $\theta_{n+1} = \text{argmax}_{\theta\in\Theta}\, \varphi_n(\theta)$. To determine which fidelity to query, we find the smallest fidelity such that the following inequality holds

$$\sqrt{\beta_n}\sigma_n^{(m)}(\theta_{n+1}) \geq \gamma^{(m)}, \tag{20}$$

where $\gamma^{(m)} > 0$ is a threshold value for all $m \in \{1, \dots, M-1\}$. If this is not satisfied for any $m$, then we query at $m_{n+1} = M$. The intuition behind (20) is that it is not worth spending resources in a region where the function $f^{(m)}$ has a small amount of uncertainty since the bound $\zeta^{(m)}$ caps off how much we can learn about the true function. As such, smaller values for $\gamma^{(m)}$ result in a larger number of queries at fidelity $m$ to reduce the variance below this threshold. A summary of the multi-fidelity GP-UCB (MF-GP-UCB) method is provided in Algorithm 2.

**Algorithm 2** The MF-GP-UCB algorithm[17]

**Initialize:** Input space $\Theta$; specify GP priors $\{m^{(m)}(\cdot), k^{(m)}(\cdot,\cdot)\}_{m=1}^{M}$; bounds $\{\zeta^{(m)}\}_{m=1}^{M}$; thresholds $\{\gamma^{(m)}\}_{m=1}^{M}$; initial datasets $\mathcal{D}_0^{(m)} = \emptyset$ for all $m = 1, \dots, M$; and maximum allowed resources $\Lambda$.
**for** $n = 0$ to $N-1$ **do** ▷ $N$ is defined implicitly based on spent resources according to (17).
  **for** $m = 1$ to $M$ **do**
    Construct a GP surrogate model for $f^{(m)}(\theta)$ given available data $\mathcal{D}_n^{(m)}$ similar to (10) and (11). ▷ only needed if new data was added in previous iteration, otherwise can reuse previous mean and covariance functions.
  **end for**
  Maximize the MF acquisition function to find $\theta_{n+1} = \text{argmax}_{\theta\in\Theta}\, \varphi_n(\theta)$.
  Select fidelity level based on $m_{n+1} = \min\{m : \sqrt{\beta_n}\sigma_n^{(m)}(\theta_{n+1}) \geq \gamma^{(m)} \text{ or } m = M\}$.
  Query the function $f^{(m_{n+1})}(\theta_{n+1})$ to get observation $y_{n+1}$.
  Update $\mathcal{D}_{n+1}^{(m)} \leftarrow \mathcal{D}_n^{(m)} \cup \{(\theta_{n+1}, y_{n+1})\}$ and set $\mathcal{D}_{n+1}^{(m)} \leftarrow \mathcal{D}_n^{(m)}$ for all $m \neq m_{n+1}$.
**end for**

The $\gamma^{(1)}, \dots, \gamma^{(M-1)}$ values are tuning parameters of the algorithm, which are needed to ensure too much effort is not spent at the lower fidelities. This is achieved in practice by setting $\gamma^{(m)}$ to small values for all $m \in \{1, \dots, M-1\}$; however, if the algorithm does not query above fidelity $m$ for more than $\lambda^{(m+1)}/\lambda^{(m)}$ iterations, then $\gamma^{(m)} \leftarrow 2\gamma^{(m)}$. All of the $\gamma^{(m)}$ values were initialized to 1% of the range of the observations from the initial queries. Additionally, Algorithm 2 assumes that the bounds $\zeta^{(1)}, \dots \zeta^{(M-1)}$ are given, which is hardly the case in most practical applications. In the available open-source implementation of MF-GP-UCB, these $M-1$ values are converted into a single bound by making the following stronger assumption

$$\|f^{(m)} - f^{(m-1)}\|_\infty \leq \zeta, \quad \forall m \in \{2, \dots, M\}. \tag{21}$$

Note that this satisfies (15) by setting $(\zeta^{(1)}, \dots, \zeta^{(M)}) = ((M-1)\zeta, \dots, \zeta)$. The value of $\zeta$ is initialized to 1% of the range of the observations from the initial queries. In addition, whenever we query at any fidelity $m > 1$, we check if thee following condition holds $|f^{(m)}(\theta_{n+1}) - \mu_n^{(m-1)}(\theta_{n+1})| > \zeta$; if so, then we also query at fidelity $m-1$. If the difference between the evaluation at the two fidelity levels exceeds the current bound, i.e., $|f^{(m)}(\theta_{n+1}) - f^{(m-1)}(\theta_{n+1})| > \zeta$, then the bound is doubled $\zeta \leftarrow 2\zeta$. Lastly, we note that the required resources $\lambda^{(1)}, \dots, \lambda^{(M)}$ may not be known exactly, so that they must also be estimated as the average computational cost for the initial set of queries at each fidelity level.

Similarly to the single-fidelity case, simple regret bounds have been established for the multi-fidelity case shown in Algorithm 2[17]. Since only the highest fidelity model is of interest to us, we need to modify the definition of the simple regret as follows

$$S(\Lambda) = \begin{cases} \min_{n:m_n=M,\forall n\in\{1,\dots,N\}} f^\star - f^{(M)}(\theta_n) & \text{if we have queried at the } M^{\text{th}} \text{ fidelity at least once,} \\ +\infty & \text{otherwise.} \end{cases} \tag{22}$$

Note that this expression reduces to (13) when only $f^{(M)}$ is observed with $n = N = \lfloor \Lambda/\lambda^{(M)} \rfloor$. The main remaining ingredient of the MF-GP-UCB algorithm is how to select the different approximation methods to derive $f^{(1)}, \dots, f^{(M-1)}$. Several different strategies for doing this in the context of DR-based IDC problems are discussed next.

## 5 | PROPOSED LOWER FIDELITY MODELS FOR DR-BASED IDC PROBLEMS

In this section, we discuss three broad approaches for deriving lower fidelity representations of the MPC-based IDC problem in (5), which requires cheaper approximations of the closed-loop cost function. Let $\lambda_x$ and $\lambda_u$ denote the cost of determining the successor state in (2) and optimal control input in (3), respectively, which are assumed to be roughly constant at each time step $t \in \mathcal{T}$. The total cost of a the high-fidelity simulation is then approximately

$$\lambda^{(M)} \sim (\lambda_x + \lambda_u)T. \tag{23}$$

The first approach involves the development of dynamic reduced models for (2) to speed up dynamic simulation of the overall system by reducing $\lambda_x$. The second approach is based on reducing the complexity of the DR (i.e., lowering $\lambda_u$) which, in this work, involves the repeated solution of the MPC optimization problem at every time step of the simulation. The third and final approach derives a simplified representation of the time grid (i.e., reducing $T$) using machine learning-based methods.

Note that the discussions provided in this section are not intended to be a comprehensive list of all possible reductions. The main goal is to highlight the many different approximation avenues that are available in practically relevant MPC-based IDC problems. Furthermore, these different strategies can easily be combined to develop any sequence of models that satisfy bounds above in (15). As such, we are not advocating any particular approximation or sequencing strategy in this paper – we intend to study this topic more in our future work. Here, we mainly want to highlight the value of lower-fidelity approximations and their impact in the context of Algorithm 2 on a case study defined in Section 6.

## 5.1 | Surrogate System Models

Modern simulation models can cover a wide range of length- and time-scales depending on the system of interest. For example, physics-based building energy and micogrid simulators, such as EnergyPlus[32] and HOMER Pro[33], require the solution to complex systems of differential algebraic equations (DAEs) that can involve thousands to millions of state variables. This is especially true whenever the DAEs are derived from the spatial discretization of a partial differential equation that describes the evolution of some states through both space and time. There is a large body of literature on how to construct reasonably fast and accurate dynamic reduced models (D-RMs) from (2); however, the specific choice of method will depend on the characteristics of the system representation. If we have access to the underlying DAE-based representation (also known as a white-box model), then we are able to apply physics-based reductions that effectively project the high-dimensional DAE into a lower-dimensional space[34]. This type of approach, which yields a DAE with many fewer unknown variables, is often referred to as *reduced-order* D-RM where "order" refers to the number of effective states variables. Since the reduced-order DAE is defined in terms of fewer variables, its cost will be lower than $\lambda_x$ – the exact speedup will depend on the choice of solver and the ratio of the number of states in the original and reduced DAE. Other white-box reduction methods include perturbation methods that convert the stiff differential equations for the fast-varying states into simpler algebraic equations[35].

If we do not have access to an equation-oriented representation of (2), then the main alternative is to construct a *data-driven* D-RM from transient input-output data. This data can be generated from repeated simulation of the system under multiple step changes in the input – system identification methods can then be used to build a surrogate model for (2) in the form of, e.g., a Volterra series, neural network, or nonlinear autoregressive moving average with exogenous input (NARMAX) model. An important distinction in the IDC application, compared to traditional system identification, is that the surrogate model must be constructed such that it can approximate the transient behavior under difference choices of the design variables $\theta \in \Theta$. These design variables can be accommodated within standard system identification methods by either building separate models for a finite number of randomly sampled designs and interpolating between them or treating the design variables as an additional time-varying control input that is changed at a slower rate than $u_t$ when constructing the identification dataset.

No matter how the D-RM is built, we expect significant gains in simulation time whenever $\lambda_x$ is large since the evaluation of these simplified models can be done in a very efficient manner. The main challenge in D-RM construction, however, is that we need sufficiently informative identification data. Dynamic optimal experiment design (dOED) methods[36] have been established for DAEs with known structure, which can optimally select control inputs to maximize the information content gained from running a simulation/experiment. The basic dOED framework could be applied to the parametrized black-box model representations to sequentially obtain better parameter estimates, which would likely result in learning an accurate model quicker than random exploration of the control input/design space.

## 5.2 | Reducing Decision Rule Complexity

As systems are becoming more complex and integrated, the control strategies needed to robustly ensure high performance in the face of uncertainties must also increase in complexity. This is the main rationale for our choice of MPC is based on the fact that it re-optimizes the control inputs as new data is collected. The upside of such a scheme is flexibility, as we can directly handle multivariate interactions, multiple unit operations, time delays, general performance measures, and system constraints. The downside, however, is the added complexity of the implicit representation of the control law, defined in terms of a potentially large-scale, non-convex dynamic optimization problem to evaluate (3), that can be difficult to solve in real-time. Although a significant amount of work has been done on the development of algorithms for real-time implementation of MPC under various assumptions, computational cost is still a bottleneck in many large-scale applications.

In cases that $\lambda_u$ is large, there are several strategies that can be employed to approximate (3) in order to develop a lower-fidelity representation of the MPC-based IDC problem (5). One of the most obvious strategies is to reduce the prediction horizon $N_p$, which directly reduces the number of variables that need to be considered in the MPC optimization. To visually demonstrate this effect, we plotted the simple regret and overall closed-loop simulation time versus $N_p \in \{8, \dots, 24\}$ in Figure 1 for the case study presented in Section 6. Note that the results have been averaged over ten random disturbance realizations, with error bars depicting the standard error estimates. The main thing to note is that we see a clear tradeoff emerge between approximation error (measured in terms of simple regret) and simulation cost as the prediction horizon changes – and this relationship is highly nonlinear with respect to quality of the approximation. For example, there is a fairly big reduction in the objective value as $N_p$ decrease from 20 to 16, while there is a negligible change in the objective value when $N_p$ decreases from 12 to 8. This suggests that there likely exists an "optimal" approximation method, though it would be difficult to derive such a method in general, especially since it will depend on the specifics of the multi-fidelity optimization approach (e.g., Algorithm 2). Other general approximation strategies include increasing the solver tolerance (or limiting the maximum number of iterations) as well as replacing the cost, system, and/or constraints with convex approximations, so that the resulting approximate optimization can be solved to global optimality using state-of-the-art convex programming solvers. Another emerging strategy to reduce complexity of MPC laws is to construct a data-driven approximation of $\kappa_t(\cdot)$ using deep neural networks (DNNs)[37,38]. DNNs are a particularly attractive because they are universal function approximators[39] and there have been significant advances in training algorithms/software so that they can effectively scale to high-dimensional problems[40].

Lastly we note that, although the strategies discussed above are meant for generic MPC problems, many problem-specific approximations can also be developed that might be valuable. The one that is likely most relevant in energy and smart-grid applications is relaxing any integer control decisions to be continuous (e.g., relaxing $u_t \in \{0, 1\}$ to $u_t \in [0, 1]$). In this case, the MPC optimization would reduce to either a nonlinear or quadratic program for which efficient solvers are readily available. The solution to this continuous optimization can then be rounded to satisfy the integer input constraints, which is a commonly

used heuristic. Similarly, distributed optimization methods, such as alternating direction method of multipliers (ADMM), can be applied to structured nonlinear programs as an inexact version of the method of multipliers[41], which is expected to be particularly useful in reducing $\lambda_u$ in systems composed of many weakly interacting subsystems.

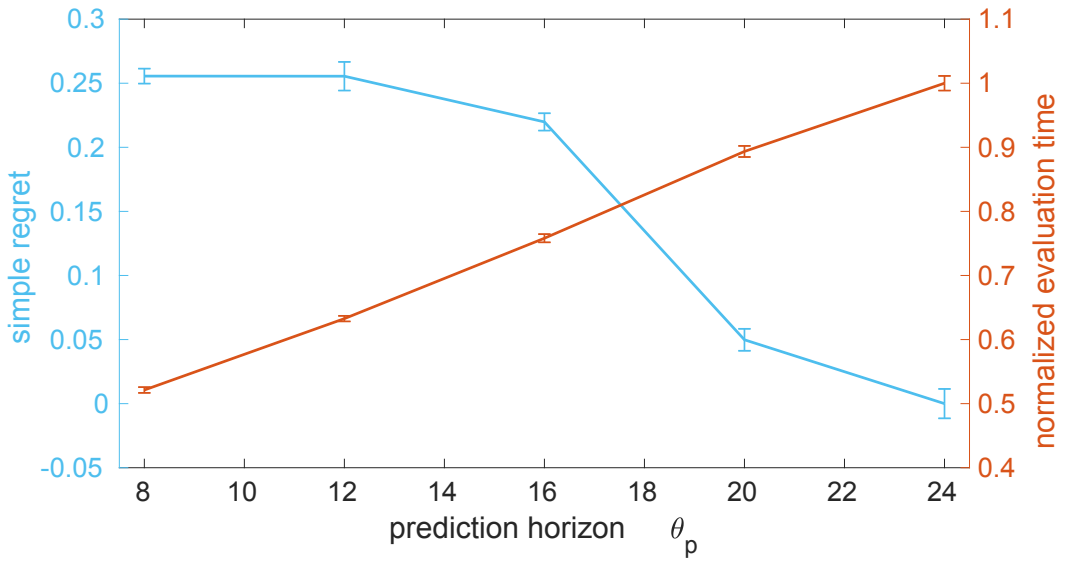



**FIGURE 1** Simple regret and normalized simulation time (averaged over 10 random disturbance realizations) versus MPC prediction horizon for the case study presented in Section 6. We clearly see a tradeoff between regret and computational cost as the prediction horizon changes.

## 5.3 | Coarsening the Time Grid Representation

As mentioned previously, one of the key advantages of IDC is considering the transient response of the system to disturbances that are inherently multi-scale in nature. This is an especially large challenge in the context of energy systems where the design decisions last for years (or even decades), while the operational decisions occur on the order of minutes to hours. As renewable energy integration continues to increase in scope, the coupling between these different time scales will only increase in importance. An emerging strategy to include short time scale phenomena in long-term planning problems is to aggregate time-series data into *representative* periods, which directly reduces the number of time steps $T$ needed to complete one closed-loop simulation. It is not uncommon to be able to represent an entire year (365 days) with a set of 5-10 days, which easily produces 1-2 orders of magnitude reduction in computational cost[42]. Representative periods can be created using time series-based clustering methods that have been developed within the machine learning community – the basic idea is to group periods (usually days) into a small number of groups that are similar. A wide-variety of clustering methods have been used for deriving representative

periods, which includes *k*-means, *k*-medoids, hierarchical, and dynamic time warping barycenter averaging (DBA) clustering. A detailed comparison between these different clustering methods was recently performed by Teichgraeber et al.[42].

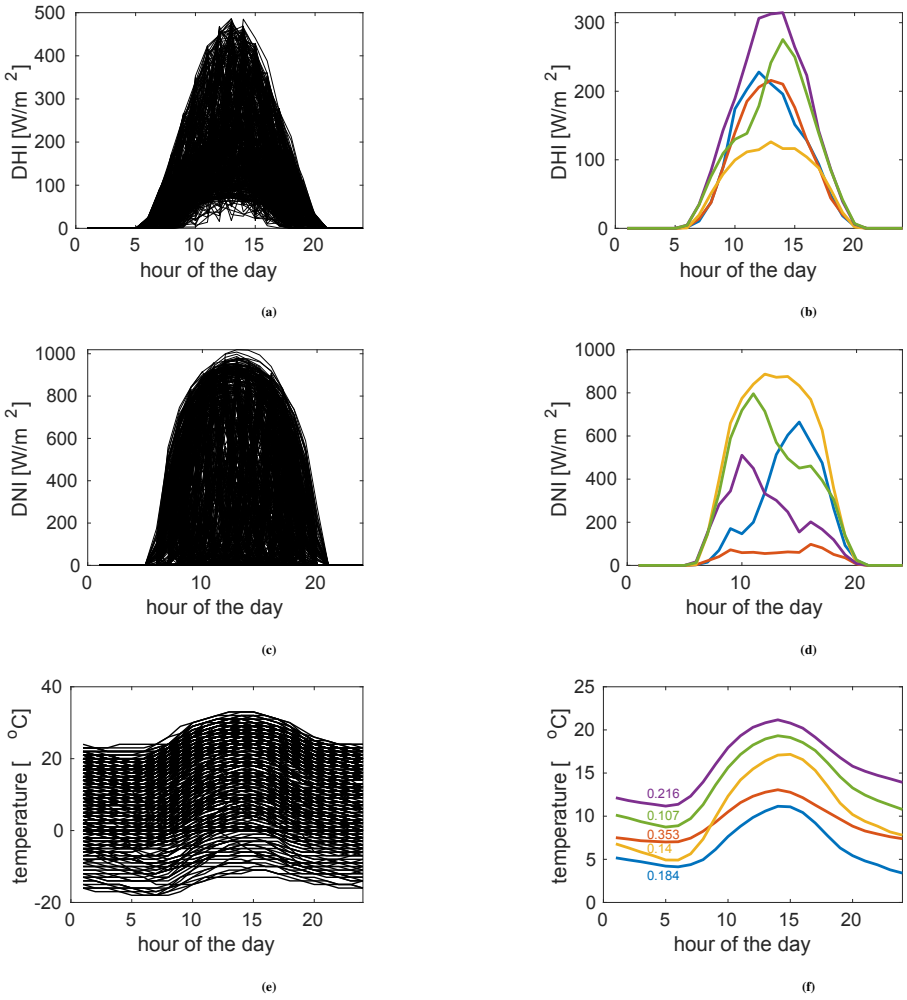



**FIGURE 2** Clustering one year of weather data into 5 representative days using k-means clustering. The top, middle, and bottom rows represent DHI, DNI, and temperature, respectively, while the left and right columns represent the year-long and clustered data, respectively. The fraction of occurrence for each representative day in the year is shown next to each curve in (f).

Here, we focus on *k*-mean clustering approach for simplicity. To illustrate the method, we consider one year worth of ambient temperature and solar irradiation data for Columbus, Ohio from January 2015 to December 2015. Solar irradiation was represented by the direct normal irradiance (DNI) and direct horizontal irradiance (DHI) values. Temperature, DNI, and DHI profiles over this time period were obtained from the publicly available National Solar Radiation Data Base[43] maintained by by the National Renewable Energy Laboratory. Daily profiles for these quantities over this two-year period are shown in Figure 2ace. To preserve any existing correlation between these quantities, the three datasets are combined together into a 72-element vector before applying the *k*-means clustering algorithm. To determine the "best" number of representative days, we successively run

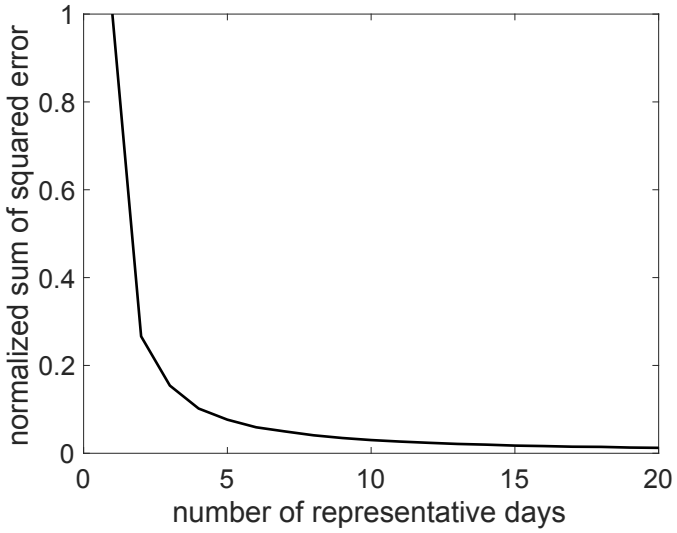



**FIGURE 3** Elbow plot showing how well the disturbance dataset for the case study in Section 6 is represented as a function of the number of clusters.

the algorithm for increasing number of clusters $k = 2, 3, \ldots, 20$ and calculate the normalized sum of squared errors between the clustered and actual days, as shown in Figure 3. In this case, we see that 5 representative days reduces the error to acceptable levels (below 10%). The resulting temperature, DNI, and DHI profiles for the 5 representative days (corresponding to the centroid of the 5 clusters) are shown in Figures 2bdf, along with the fraction of the year that each day represents. We are now able to use these representative days, weighted by their fraction of occurrence, as a lower-fidelity approximation of (5).

# 6 | CASE STUDY: DESIGN OF A BUILDING HEATING AND COOLING SYSTEM WITH PHOTOVOLTAIC POWER GENERATION, BATTERY STORAGE, AND GRID SUPPORT

## 6.1 | Description of System Model and IDC Problem

We consider the design of a building heating and cooling (HC) system that is connected to a photovoltaic (PV) array and battery energy storage device, with grid support, as depicted in Figure 4. The main system design variables of interest are the battery capacity $d^B$ and the PV area $d^{PV}$. We also have three key control inputs $u_t = \{u_t^{HC}, u_t^B, u_t^G\}$ where $u_t^{HC}$ is the net heating energy supplied to the building (which will be positive when heating and negative when cooling), $u_t^B$ is the energy sourced from the battery, and $u_t^G$ is the energy sourced from the grid. As depicted in Figure 4, these variables must always satisfy an energy balance $u_t^{HC} = u_t^G + u_t^B$. We model the grid as an infinite reservoir, meaning that energy may purchased and stored in the battery, used directly for the HC load, or drawn from the battery and sold for profit. Positive values $u_t^G > 0$ indicate electricity has been purchased from the grid, while negative values $u_t^G < 0$ indicate electricity is being sold to the grid. All electricity lines are assumed to be limited to 1500 kWh.

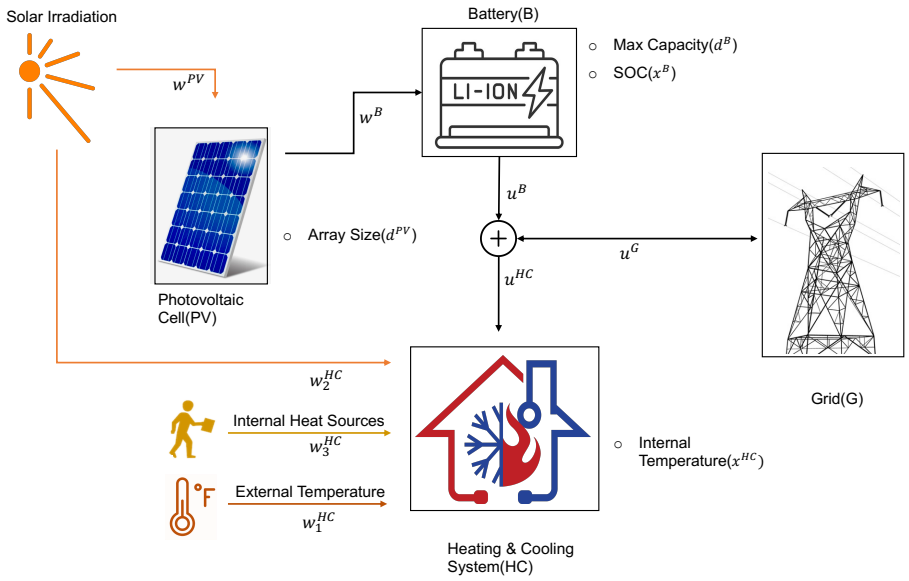



**FIGURE 4** Schematic overview of solar-powered building heating and cooling system with battery storage and grid support.

The building and battery models are represented in terms of discretized differential equations that are provided in detail in Appendix B.1. The building model is adapted from Gondhalekar et al.[44] that describes the dynamic evolution of temperature inside of a single room in a larger office building and is composed of three states $x_t^{HC} = \{x_{t,1}^{HC}, x_{t,2}^{HC}, x_{t,3}^{HC}\}$ that represent the indoor building temperature, average interior wall temperature and exterior wall temperature, respectively. The battery model is composted of a single state $x_t^B$ that represents the state-of-charge (SOC) of the battery. Although a simplified battery model was used here to reduce computational burden in the extensive testing performed, the proposed DR-based IDC and MF-GP-UCB approaches are generally applicable to sophisticated physics-based models including those available in the Simulink toolbox that explicitly model chemistry-specific degradation rates. All relevant parameters in the building and battery models, including the bounds on the design variables and control inputs, are summarized in Appendix B.2 in Table B2.

Several disturbances also enter into different components of the model . The PV can only absorb so much energy from the sunlight based on the direct horizontal irradiance (DHI), denoted by $w_t^{PV}$. Only a fraction of the DHI can then be stored by the battery in every time instance (assuming capacity is available), which is given by $w_t^B = d^{PV} w_t^{PV} \eta^{PV}$ where $\eta^{PV}$ denotes the PV efficiency. The external temperature $w_{t,1}^{HC}$, direct normal irradiance (DNI) $w_{t,2}^{HC}$, and internal heat sources $w_{t,3}^{HC}$ also impact the system. Historical weather data for Columbus, Ohio was used to represent $w_t^{PV}$, $w_{t,1}^{HC}$, and $w_{t,2}^{HC}$, as discussed in Section 5.3. The internal heat sources $w_{t,3}^{HC}$, on the other hand, were modeled as a uniform random variable between the bounds given in Table B2 whenever the time period satisfies $t \mod 24 \in \{8, \ldots, 18\}$ (represents business hours from 8am to 6pm) and zero otherwise. The overall system model can then be cast in the form of (2) with states $x_t = \{x_t^{HC}, x_t^B\}$, control inputs $u_t = \{u_t^{HC}, u_t^B, u_t^G\}$,

disturbances $w_t = \{w_t^{PV}, w_t^B, w_{t,1}^{HC}, w_{t,2}^{HC}, w_{t,3}^{HC}\}$, and design variables $d = \{d^B, d^{PV}\}$ using a forward Euler discretization scheme with a 1 hour sampling time. We also assume a fixed initial condition $b_0(d) = \{21°\text{C},\ 20°\text{C},\ 4°\text{C},\ 50\%\}$.

We now formulate our cost function in the IDC problem (1), which is composed of the capital cost $C(d)$ and operating cost $O(d, z, \boldsymbol{\omega})$. The capital cost function is the sum of the battery and PV costs, which is given by

$$C(d) = \pi^B d^B + \pi^{PV} d^{PV}, \tag{24}$$

where $\pi^B$ and $\pi^{PV}$ are the per unit costs of the battery and PV, respectively, reported in Table B2. The operating costs, on the other hand, are computed using (4) for a one-year planning horizon corresponding to $T = 8760$ time steps. The stage cost is assumed to be a weighted combination of the electricity cost and violation of time-varying temperature and SOC constraints. Before defining $\ell_t(\cdot)$, let us first define the desired state constraints in the form of $g_t(x_t, u_t, w_t) \leq 0$ with $g_t(\cdot)$ defined as follows

$$g_t(x_t, u_t, w_t) = \begin{bmatrix} x_{t,1}^{HC} - \overline{x}_{t,1}^{HC} \\ \underline{x}_{t,1}^{HC} - x_{t,1}^{HC} \\ x_t^B - \overline{x}^B \\ \underline{x}^B - x_t^B \end{bmatrix}, \tag{25}$$

where $\overline{x}_{t,1}^{HC}$ and $\underline{x}_{t,1}^{HC}$ are the time-varying upper and lower bounds for the internal building temperature and $\overline{x}^B$ and $\underline{x}^B$ are the fixed upper and lower bounds for the SOC. These constraints will be directly enforced in the MPC problem defined later, but are simply penalized with an appropriate cost factor when defining the operating cost. As such, the stage cost in (4) is given by the following expression in this example

$$\ell_t(x_t, u_t, w_t, d) = \pi_t^G u_t^G + \pi_t^{HC} \left( [x_{t,1}^{HC} - \overline{x}_{t,1}^{HC}]^+ + [\underline{x}_{t,1}^{HC} - x_{t,1}^{HC}]^+ \right) + \pi_t^{SOC} \left( [x_t^B - \overline{x}^B]^+ + [\underline{x}^B - x_t^B]^+ \right), \tag{26}$$

where $[a]^+ = \max\{a, 0\}$ and $\pi_t^G$, $\pi_t^{HC}$ and $\pi_t^{SOC}$ are the time-varying per unit time costs of electricity, temperature constraint violations, and SOC constraint violations, respectively. We assume the terminal cost is equal to zero, i.e., $\ell_T(\cdot) = 0$. All relevant price and constraint values are reported in Tables B2 and B3. Note that $\pi_t^{HC}$ is chosen so that higher penalties are incurred during business hours. We have not specified a control law of the form (3) yet, as we will explore two different cases below.

Note that we have provided all code, data, and results at: `https://github.com/PaulsonLab/MF-GP-UCB_for_IDC.git` for each of the parts of this case study discussed in detail next.

*Remark 2.* We have implicitly assumed that we have access to an HC system that offers continuous modes of operation between the maximum and minimum net energy input in our formulation of the system model. In practice, several smaller HC units may be required to satisfy the load – if each unit has a fixed duty cycle, then we would need to include discrete/integer decisions

in the model that represent, for example, turning on or off certain units in certain time periods. This assumption was made for simplicity, but our framework can easily accommodate such decision as shown in our previous work[16].

## 6.2 | Comparing MF-GP-UCB to Alternative Black-Box Optimization Methods

In this section, we benchmark the MF-GP-UCB algorithm relative to alternative widely-used black-box optimizers. To avoid the complexity added by the stochastic uncertainties, we initially focus on the simplified case $p_{\boldsymbol{\omega}}(\boldsymbol{\omega}) = \delta(\boldsymbol{\omega} - \hat{\boldsymbol{\omega}})$ where the uncertainties are fixed to their nominal values $\hat{\boldsymbol{\omega}}$. In this case, we can avoid the use of MPC and instead select the optimal control actions from the following large-scale optimization problem

$$\{u_0^\star(d,\hat{\boldsymbol{\omega}}),\dots,u_{T-1}^\star(d,\hat{\boldsymbol{\omega}})\} = \underset{u_0,\dots,u_{T-1}}{\operatorname{argmin}} \; C(d) + \sum_{t=0}^{T-1} \ell_t(x_t(d,\hat{\boldsymbol{\omega}}),u_t,\hat{w_t},d), \tag{27a}$$

$$\text{s.t. } x_{t+1}(d,\hat{\boldsymbol{\omega}}) = h_t(x_t(d,\hat{\boldsymbol{\omega}}),u_t,\hat{w_t},d), \qquad \forall t \in \mathcal{T}, \tag{27b}$$

$$x_0(d,\hat{\boldsymbol{\omega}}) = b_0(d), \tag{27c}$$

$$g_t(x_t(d,\hat{\boldsymbol{\omega}}),u_t,\hat{w_t}) \le 0, \qquad \forall t \in \mathcal{T}, \tag{27d}$$

$$u_t \in \mathcal{U}, \qquad \forall t \in \mathcal{T}. \tag{27e}$$

In this case, $\theta = d$ since there are no additional tuning parameters considered in (27). We then specify our control inputs as $u_0^\star(d,\hat{\boldsymbol{\omega}}),\dots,u_{T-1}^\star(d,\hat{\boldsymbol{\omega}})$ over the entire simulation time, as opposed to using (3), since there is no need to adapt these control actions when fixed disturbance values are considered. For any fixed value of $\theta = d$, (27) can be written as a large-scale linear program based on the assumed system model in Appendix B that can easily be represented in the Yalmip[45] modeling environment and efficiently solved using Gurobi[46]. One high-fidelity evaluation of (27) with $T = 8760$ took approximately 1 minutes on a MacBook Pro with a 2.3 GHz Intel Core i9 and 16 GB of RAM. To develop a low-fidelity model, we used the time series-based clustering strategy presented in Section 5.3 that resulted in 5 representative days (Figure 2). Evaluating the low-fidelity model, which is equivalent to (27) with $T = 120$ and $\hat{\boldsymbol{\omega}}$ replaced with the clustered disturbance sequences, reduced the computational cost by a factor of 50 (estimated as the average over ten separate runs at each fidelity level). As such, we were able to set the cost values to $\lambda^{(1)} = 0.02$ and $\lambda^{(2)} = 1$, with $\lambda^{(2)}$ representing the high-fidelity cost that we normalize to 1. We set our total budget to $\Lambda = 12$, corresponding to 12 equivalent high-fidelity evaluations. Note 4 out of the 12 maximum budget units are allocated to a set of evaluations at randomly sampled points, which are needed to construct an initial GP as well as estimate the bounds in (15).

We compare the MF-GP-UCB algorithm to three alternatives; note that we focus on the BO framework since this has already been demonstrated to be effective in reasonably low-dimensional spaces compared to alternatives (see Section 3 for further discussion). The three alternatives are single-fidelity GP-UCB, expected improvement (EI), and random search. The EI approach was implemented using `bayesopt`[47] within the Matlab Optimization Toolbox, while the others were executed in an open-source

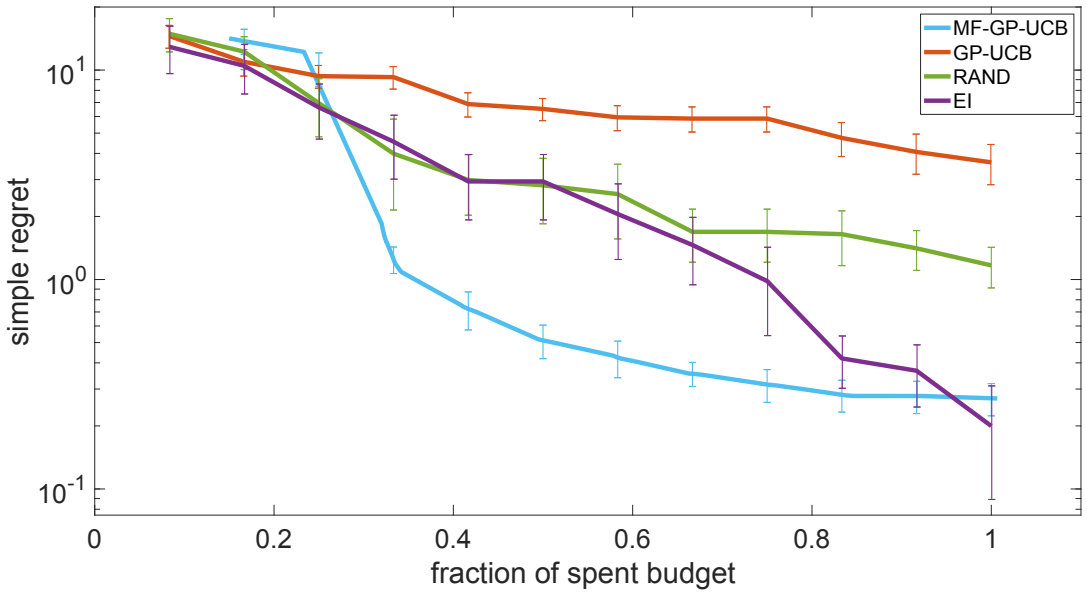



**FIGURE 5** The simple regret $S(\Lambda)$ versus fraction of budget spent for four different optimizers. Each approach was repeated for 10 sets of random seed points. The mean of these 10 runs are shown in bold and confidence intervals shown with error bars.

implementation of MF-GB-UCB available at: `https://github.com/kirthevasank/mf-gp-ucb`. We use simple regret (22) as our performance metric to compare these different methods. Due to the random initialization, simple regret is a random variable and thus it is not informative to show results for a single initialization. Instead, we repeat each experiment 10 times to estimate the average simple regret for each algorithm – error bars are computed by estimating the confidence intervals as 1.96 times the standard deviation divided by the square root of the number of repeats.

The expected simple regret versus fraction of the maximum budget spent for all four algorithms is shown in Figure 5. We clearly see that MF-GB-UCB converges faster than the other methods and does particularly well for budget fractions between 0.5 and 0.8. It is interesting to note that random search actually outperformed the single-fidelity GP-UCB approach in this case. We believe this is due to a relatively small sample size of 10 repeats and the tendency for GP-UCB to over explore. EI, on the other hand, does result in similar quality solutions at the end of the budget; however, it does considerably worse at smaller budgets, which suggests significantly worse anytime performance. To demonstrate that MF-GP-UCB was able to identify good designs, we analyze the solution in more detail for the median of the 10 repeats. The year-long sequence of the day-averaged control trajectories are shown in Figure 6. Although we see that the average net heating is higher during the summer than winter, there is no clear trend in these results, which highlights the need to consider short time-scale phenomena when dealing with highly variable disturbances. Based on these daily averaged profiles, we selected three consecutive summer and winter days to plot at the hourly scale in Figure 7. From this figure, we can see the cyclic nature of the state and control profiles from day to day as well as strong seasonal effects that result in vastly different control strategies.

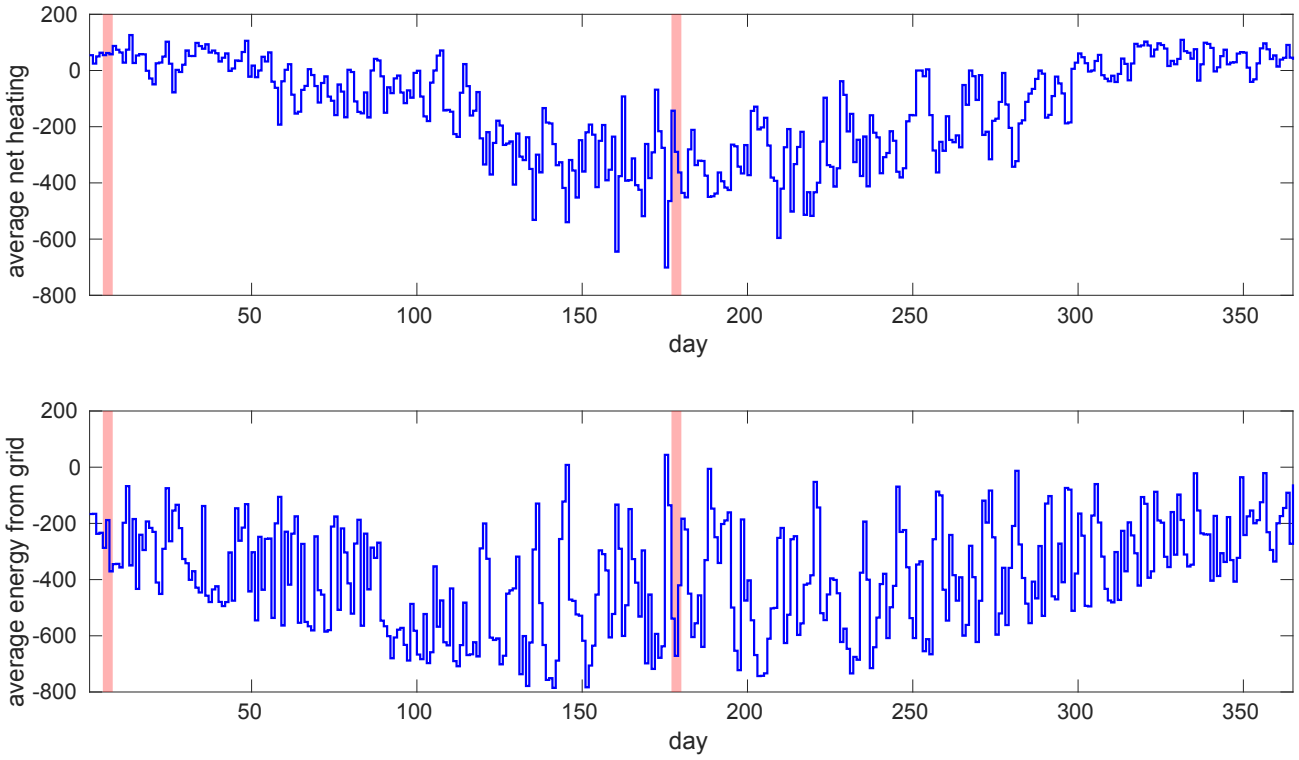



**FIGURE 6** Year-long trajectory of the daily average control strategy during business hours (8am to 6pm) for the MF-GP-UCB solution to (27). Days 5-8 and 177-180 are highlighted in red, and are shown in detail in Figure 7 to highlight the differences in winter versus summer operation

## 6.3 | Applying MF-GP-UCB to MPC-based IDC under Different Levels of Uncertainty

In this section, we utilize a stochastic uncertainty sequence such that we must rely on MPC to adapt the control input under different realizations of the uncertainty. The MPC problem, which we apply to the previously described building HC system with renewable energy storage, can be defined as follows

$$\min_{x_{t+k|t},u_{t+k|t},\epsilon_{t+k|t}} \quad \sum_{k=0}^{N_p-1} \pi^G_{t+k} u^G_{t+k|t} + \rho \|s_{t+k|t}\|_1, \tag{28a}$$

$$\text{s.t.} \quad x_{t+k+1|t} = h_{t+k}(x_{t+k|t}, u_{t+k|t}, \hat{w}_{t+k|t}, d), \qquad \forall k \in \{0, \dots, N_p - 1\}, \tag{28b}$$

$$\hat{w}_{t+k+1|t} = q_{t+k}(\hat{w}_{t+k|t}), \qquad \forall k \in \{0, \dots, N_p - 2\}, \tag{28c}$$

$$\{x_{t|t}, \hat{w}_{t|t}\} = \{x_t(d, z, \boldsymbol{\omega}), w_t\}, \tag{28d}$$

$$g_{t+k}(x_{t+k|t}, u_{t+k|t}, \hat{w}_{t+k|t}) + z^{backoff} \leq s_{t+k|t}, \qquad \forall k \in \{0, \dots, N_p - 1\}, \tag{28e}$$

$$u_{t+k|t} \in \mathcal{U}, \qquad \forall k \in \{0, \dots, N_p - 1\}, \tag{28f}$$

$$s_{t+k|t} \geq 0, \qquad \forall k \in \{0, \dots, N_p - 1\}, \tag{28g}$$

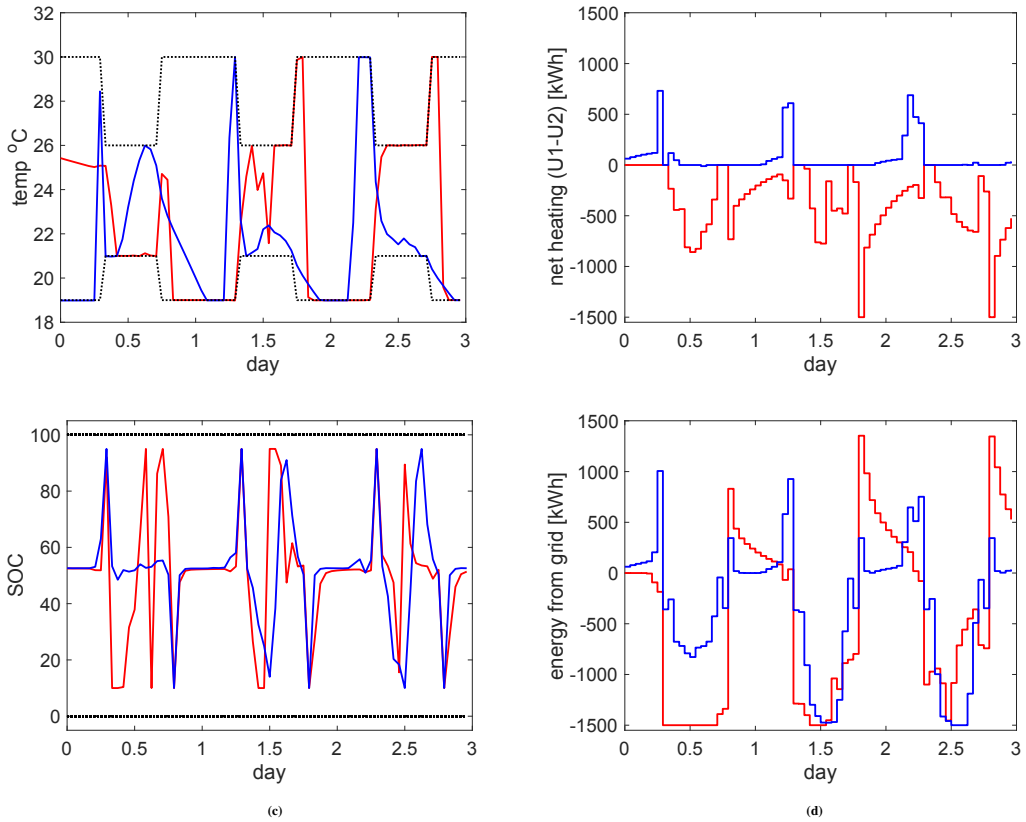



**FIGURE 7** State and control trajectories for three consecutive days in winter (blue) and summer (red) corresponding to Figure 6.

where the notation $a_{t+k|t}$ denotes the predicted value of the variable $a \in \{x, u, \hat{w}\}$ at time $t + k$ given information up until time $t$. Here, (28a) is the total cost to be minimized which is composed of the electricity price and soft constraint penalty over the prediction horizon $N_p$. The constraints (28b) define the predicted state sequence as a function of the predicted input and estimated uncertainty values $\hat{w}$. Although we assumed exact knowledge of the structure of the dynamics $h_{t+k}(\cdot)$ for simplicity, this can easily be replaced with an approximate model derived from physics-based equations or black-box system identification. The recursion (28c) computes the predicted future disturbance sequence $\{\hat{w}_{t+k|t}\}$ given some forecasting function $q_{t+k}(\cdot)$, while (28d) ensures the initial state and disturbance at time $t$ start from their most recently measured values (coming from the detailed closed-loop simulation). Similarly to the system dynamics, we can handle any choice of $q_{t+k}(\cdot)$ in our framework as long as it only uses *previous* data to make its predictions. The set of constraints in (28e) enforce the system constraints (25) for the predicted uncertainty values where $z^{backoff} \geq 0$ denotes backoff tuning parameters that are able to confer strong robustness properties (at the cost of conservative performance) when properly calibrated[48]. To avoid feasibility issues, we still soften these constraints in practice with slack variables $\{s_{t+k|t}\}$ with a relatively large penalty $\rho = 10^3$. Lastly, the desired set of hard input constraints and positivity of the slack variabels are enforced via (28f) and (28g), respectively. We fix the prediction horizon at $N_p = 24$, meaning that the additional MPC tuning parameters are the two backoff values $z = z^{backoff}$ corresponding to the temperature constraints in (28e) – thus the design and tuning variables $\theta = \{d, z\}$ are co-optimized in this section.

To highlight the fact that we can handle any type of disturbance, instead of deriving a specific forecasting function $q_{t+k}(\cdot)$, we replace (28c) with the following perturbed prediction at every time step $t \in \mathcal{T}$

$$\hat{w}_{t+k|t} = w_{t+k} + \varepsilon_{t+k|t}, \quad \forall k \in \{1, \dots, N_p - 1\}, \tag{29}$$

where $\varepsilon_{t+k|t}$ are the forecasting errors that are assumed to be independent and identically (i.i.d.) distributed in terms of a Gaussian distribution with zero mean and covariance $\Sigma_\epsilon$. In this case, we can think of $\boldsymbol{\omega}$ as the collection of all forecasting errors, with larger values on the diagonal of $\Sigma_\epsilon$ implying more challenging to predict disturbances. This will allow us to more easily illustrate the performance of MF-GP-UCB under different levels of uncertainty. We again take the clustered 5 representative day case as our low-fidelity model. Before we can apply the MF-GP-UCB method (Algorithm 2), we need to estimate the difference in computational cost of these two models. One high-fidelity evaluation of the MPC-based IDC objective in (5) was approximately 2 minutes, while the clustered solution took only around 5 seconds. This implies that the cost values should be set to $\lambda^{(1)} = 0.04$ and $\lambda^{(2)} = 1$ where we have again normalized the high-fidelity cost to 1. Due to the increased complexity, we accordingly reduced the maximum budget to $\Lambda = 9$, which includes a random initialization budget of 3.

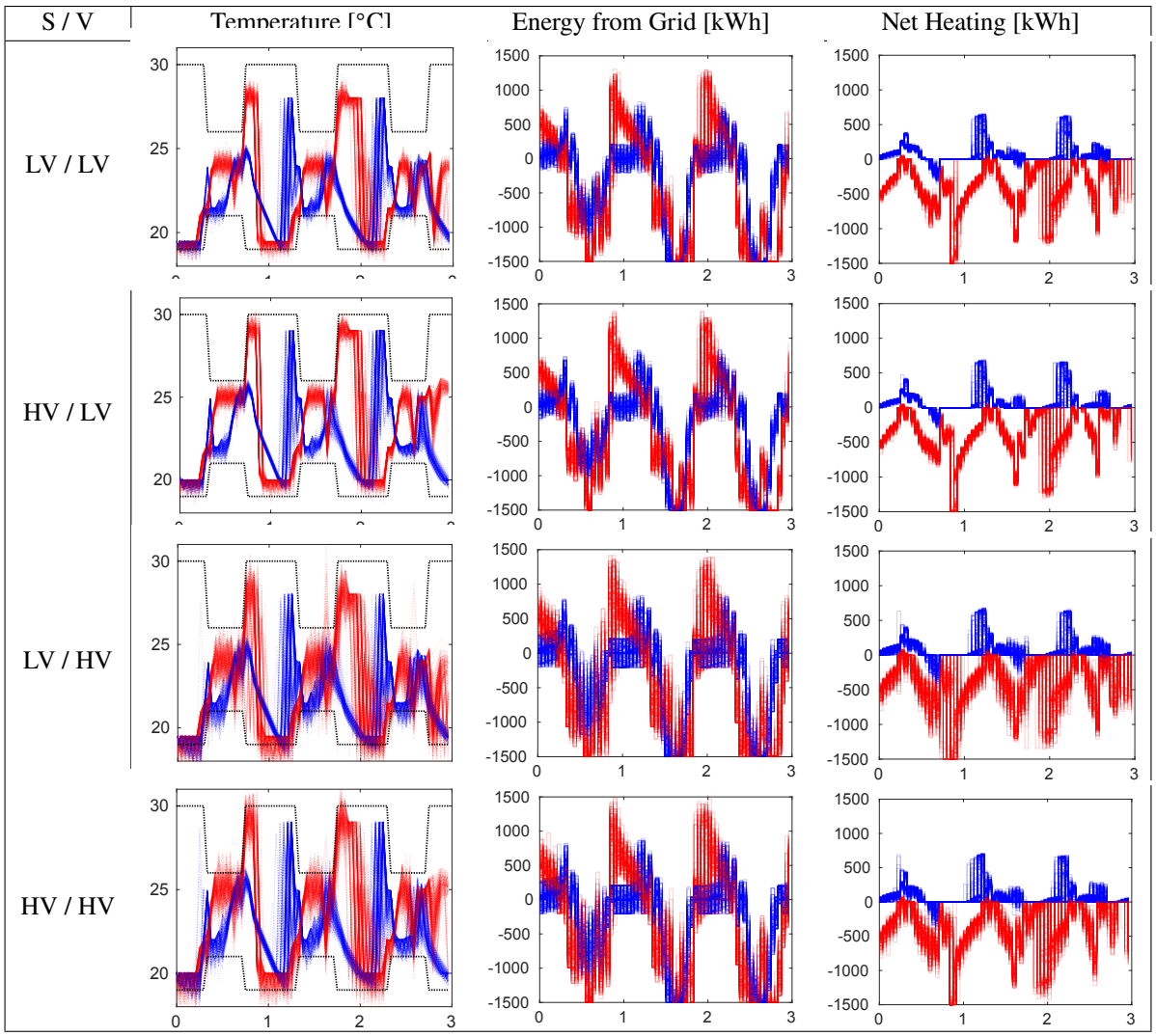
To analyze the robustness of MF-GP-UCB to noisy objective evaluations, we apply the algorithm under two different forecasting error noise levels, which we denote by low variance (LV) $\Sigma_\epsilon = \Sigma_\epsilon^{LV} = \text{diag}(1, 100^2, 10^2)$ and high variance (HV) $\Sigma_\epsilon = \Sigma_\epsilon^{HV} = 2\Sigma_\epsilon^{LV}$. Let $\theta^\star_{LV}$ and $\theta^\star_{HV}$ denote the median optimal design identified by MFBO for 10 different random seed values under low and high variance, respectively. We analyzed the quality of these solutions by performing a set of closed-loop validation runs using a set of i.i.d. samples of $\boldsymbol{\omega}$. Here, we draw two sets of 270 realizations from the LV and HV distributions and test $\theta^\star_{LV}$ and $\theta^\star_{HV}$ on both of these cases. The average performance value (negative of the cost), as defined in (5), is reported in Table 1 for these four cases. It is interesting to note that all four cases were found to have profitable solutions due to the fact that we could sell a significant amount of electricity to the grid to cover our capital cost. Although profit was higher when considering the LV model in the MF-GP-UCB algorithm, this comes at the cost of a greater frequency of temperature constraint violations during business hours, as also shown in Table 1.

**TABLE 1** Estimated performance and constraint violation for the optimal solutions $\theta^\star_{LV}$ and $\theta^\star_{HV}$ (denoted by S) applying a scenario-based verification (denoted by V) under 270 separate LV and HV disturbance realizations.

| S / V | Estimated performance $f(\theta)$ | Probability of constraint violations during business hours |
|---|---|---|
| LV / LV | $6.965290 \pm 0.000948$ | $0.184563 \pm 0.009366$ |
| LV / HV | $6.505025 \pm 0.001489$ | $0.247758 \pm 0.011477$ |
| HV / LV | $6.718205 \pm 0.000773$ | $0.115160 \pm 0.005848$ |
| HV / HV | $6.317560 \pm 0.001288$ | $0.155820 \pm 0.008859$ |

In Table 2, we again zoom in on the three summer and winter days identified in Figure 7 for each of these four cases to better understand the impact of forecasting errors on the results. We can clearly see how the control strategy changes in each season, with summer days incurring more frequent/larger magnitude violations while selling more electricity.

**TABLE 2** State and control trajectories for three consecutive days in winter (blue) and summer (red) corresponding to the red boxes in Figure 6 for optimal solutions $\theta^{\star}_{LV}$ and $\theta^{\star}_{HV}$ (denoted by S) applying a scenario-based verification (denoted by V) under 270 separate LV and HV disturbance realizations.




## 6.4 | Results for MF-GP-UCB on MPC-based IDC with Three Fidelity Levels

The previous sections only considered $M = 2$ fidelity levels. Here, we want to study the impact of the particular sequence of $M = 3$ reductions on the IDC problem with low variance forecasting errors. To do this, we look at two different orderings of the

same type of approximations, mainly reducing the prediction horizon to $N_p = 12$ and time series-based clustering that we refer to as "Case 1" and "Case 2". Both of these cases share the same high-fidelity $m = 3$ and low-fidelity model $m = 1$ models. The main difference is the middle fidelity $m = 2$, which is constructed from the reduced prediction horizon only (full time horizon) and the clustered shortened time horizon (full MPC prediction horizon) for Cases 1 and 2, respectively. The estimated costs for the three different fidelity levels for each of the cases was found to be

1. Case 1: $\{\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}\} = \{0.00998, 0.5794, 1\}$;

2. Case 2: $\{\lambda^{(1)}, \lambda^{(2)}, \lambda^{(3)}\} = \{0.00998, 0.01637, 1\}$.

Thus, we clearly see that Case 2 has a significantly smaller cost for the middle fidelity model compared to Case 1. To observe any potential effects that could arise in later iterations, we set a relatively large maximum budget of $\Lambda = 30$ equivalent high-fidelity evaluations, with a random initialization budget of 3. The MF-GP-UCB algorithm was repeated 20 times for both Cases 1 and 2. The expected simple regret versus fraction of the budget spent for the two cases is shown in Figure 8. We can clearly see that Case 2 begins to query the high-fidelity model earlier on and at much better points on average, though this gap does seem to close as the budget is allowed to increase. This suggests that there can be a substantial benefit in designing particularly cheap approximations to reduce the amount of costly exploration needed at the highest fidelity levels.

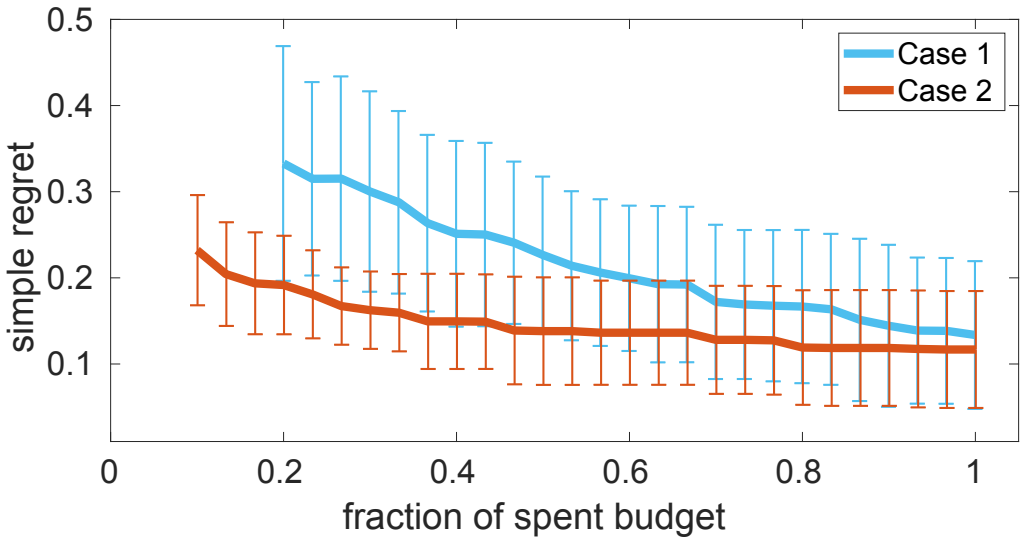



**FIGURE 8** Performance of MF-GP-UCB under two different fidelity sequences for $M = 3$ models. In Case 1, the middle fidelity is constructed by only reducing the prediction horizon ($N_p = 12$) whereas, in Case 2, the middle fidelity is constructed by clustering the disturbance time series into 5 representative days.

## 7 | CONCLUSIONS

In this paper, we present an efficient approach for solving computationally expensive integrated design and control (IDC) problems using multi-fidelity simulation optimization. Such IDC formulations are needed to help enable the shift to next-generation manufacturing and energy systems that can flexibly respond to highly dynamic and uncertain conditions arising from a variety of sources including the increased integration with renewable energy technologies. However, realistic IDC problems are known to be challenging to solve as they are naturally formulated as large-scale, non-convex, and non-smooth optimization problems that cannot be tractably solved using currently available methods. We overcome these challenges in this work using two key ideas: (i) approximating the complex recourse (wait-and-see) decisions using a high-quality decision rule (DR) derived from a model predictive control (MPC) law and (ii) applying a multi-fidelity Bayesian optimization (MFBO) to tackle the MPC-based IDC problem in a simulation-based manner. The MFBO paradigm builds up traditional BO by taking advantage of a sequence of lower-fidelity approximations to the high-fidelity IDC problem that can be evaluated at a fraction of the cost. Using bounds on the quality and estimates of the costs of the approximations, we can use MFBO to sequentially select the next design and fidelity level at which we should evaluate. We also discuss three major ways to derive low-fidelity approximations of the MPC-based IDC cost, which includes simplifications of the system model, decision rule, and time grid – it turns out that different machine learning-based methods can be beneficial in each of these simplification categories.

To demonstrate the effectiveness of MFBO, we compare it to traditional BO on the design of a building heating and cooling system with solar power generation, battery storage, and grid support. We consider several sources of uncertainty, including weather and demand conditions, that can vary at the hour scale over a year-long planning horizon; the uncertainty data was pulled from a national database for a region in Columbus, Ohio to emphasize the practical nature of the approach. We found that MFBO consistently found better solutions with fewer expensive function evaluations than alternative methods, especially when the budget is very limited (10 or less total high-fidelity runs). Additionally, we found that MFBO was able to suitably handle random forecast errors in the key disturbances and reduce constraint violations by tuning backoff values in the MPC-based DR. We found that the ability to handle noisy function evaluations, which are inherent in IDC in the context of energy systems, is a key advantage of the MFBO methodology that is difficult to overcome with many alternative derivative-free optimization methods. Lastly, we analyzed the impact on the sequence of approximations impacted the convergence of MFBO. We found that the relative accuracy and computational cost of the fidelities play an important role in the performance of MFBO in the early iterations (with less accurate, cheaper models being preferred), while the differences in performance tended to shrink as the budget increased. Thus, we argue that an important direction for future work is to better quantify the impact of different fidelity models and determine rules for deriving "optimal" low-fidelity approximations. This is especially important in the context of MPC-based IDC since, as discussed in detail in Section 5, there are many ways to derive low-fidelity approximations and some

approaches would clearly be preferred depending on the specific characteristics of the application of interest. In addition, it has been recently shown that significant improvements in the convergence rate of BO can be achieved when the cost function can be broken down into a composite function (i.e., $f(\theta) = g(h(\theta))$ where $g(\cdot)$ is a *known* function and $h(\cdot)$ is an *unknown* vector-valued black-box function)[49]. It is expected that similar performance gains could be obtained in the multi-fidelity case, if this structure (when present) can somehow be exploited by MF-GP-UCB.

## DATA AVAILABILITY STATEMENT

A summary of the model equations is provided in Appendix B. The complete set of data that support the findings of this study are openly available on Github at `https://github.com/PaulsonLab/MF-GP-UCB_for_IDC.git`.

## References

1. Madlener R, Schmid C. Combined heat and power generation in liberalised markets and a carbon-constrained world. *GAIA-Ecological Perspectives for Science and Society* 2003; 12: 114–120.

2. Samad T, Kiliccote S. Smart grid technologies and applications for the industrial sector. *Computers & Chemical Engineering* 2012; 47: 76–84.

3. Tang MC, Chin MWS, Lim KM, et al. Systematic approach for conceptual design of an integrated biorefinery with uncertainties. *Clean Technologies and Environmental Policy* 2013; 15: 783–799.

4. Hegerl, G.C., F. W. Zwiers, P. Braconnot, N.P. Gillett, Y. Luo, J.A. Marengo Orsini, N. Nicholls, J.E. Penner and P.A. Stott, 2007: Under-standing and Attributing Climate Change. In: *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon, S., D. Qin, M. Manning, Z. Chen, M. Marquis,K.B. Averyt, M. Tignor and H.L. Miller (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

5. Sakizlis V, Perkins JD, Pistikopoulos EN. Recent advances in optimization-based simultaneous process and control design. *Computers & Chemical Engineering* 2004; 28(10): 2069–2086.

6. Hakizimana A. Novel Optimization Approaches for Integrated Design and Operation of Smart Manufacturing and Energy Systems. 2019.

7. Pereira MVF, Pinto LMVG. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming* 1991; 52: 359–375.

8. Rawlings JB, Mayne DQ. *Model predictive control: Theory and design*. Nob Hill Publishing, Madison, Wisconsin . 2009.

9. Touretzky CR, Baldea M. Model reduction and nonlinear MPC for energy management in buildings. In: *2013 American Control Conference*IEEE. ; 2013: 461–466.

10. Zhang X, Schildbach G, Sturzenegger D, Morari M. Scenario-based MPC for energy-efficient building climate control under weather and occupancy uncertainty. In: *2013 european control conference (ECC)*IEEE. ; 2013: 1029–1034.

11. Touretzky CR, Baldea M. Nonlinear model reduction and model predictive control of residential buildings with energy recovery. *Journal of Process Control* 2014; 24(6): 723–739.

12. Zhao Y, Lu Y, Yan C, Wang S. MPC-based optimal scheduling of grid-connected low energy buildings with thermal energy storages. *Energy and Buildings* 2015; 86: 415–426.

13. Elsisi M, Aboelela M, Soliman M, Mansour W. Design of optimal model predictive controller for LFC of nonlinear multi-area power system with energy storage devices. *Electric Power Components and Systems* 2018; 46(11-12): 1300–1311.

14. Elsisi M, Bazmohammadi N, Guerrero JM, Ebrahim MA. Energy management of controllable loads in multi-area power systems with wind power penetration based on new supervisor fuzzy nonlinear sliding mode control. *Energy* 2021; 221: 119867.

15. Elsisi M. New design of adaptive model predictive control for energy conversion system with wind torque effect. *Journal of Cleaner Production* 2019; 240: 118265.

16. Choksi N, Paulson J. Simulation-based Integrated Design and Control with Embedded Mixed-Integer MPC using Constrained Bayesian Optimization. In: *Proceedings of the American Control Conference*; 2021.

17. Kandasamy K, Dasarathy G, Oliva J, Schneider J, Poczos B. Multi-fidelity Gaussian process bandit optimisation. *Journal of Artificial Intelligence Research* 2019; 66: 151–196.

18. Conn AR, Scheinberg K, Vicente LN. *Introduction to Derivative-free Optimization*. SIAM . 2009.

19. Rios L, Sahinidis N. Derivative-free optimization: A review of algorithms and comparison of software implementations. *Journal of Global Optimization* 2013; 56(3): 1247–1293.

20. Mukhopadhyay DM, Balitanas MO, Farkhod A, Jeon SH, Bhattacharyya D. Genetic algorithm: A tutorial review. *International Journal of Grid and Distributed Computing* 2009; 2(3): 25–32.

21. Eberhart R, Kennedy J. Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*. 4. ; 1995: 1942–1948.

22. Rasmussen CE, Williams CKI. *Gaussian processes for machine learning.* Cambridge, MA: MIT Press . 2006.

23. Shahriari B, Swersky K, Wang Z, Adams RP, De Freitas N. Taking the human out of the loop: A review of Bayesian optimization. *Proceedings of the IEEE* 2015; 104: 148–175.

24. Bradford E, Schweidtmann AM, Lapkin A. Efficient multiobjective optimization employing Gaussian processes, spectral sampling and a genetic algorithm. *Journal of Global Optimization* 2018; 71(2): 407–438.

25. Jones D, Schonlau M, Welch W. Efficient global optimization of expensive black-box functions. *Journal of Global optimization* 1998; 13(4): 455–492.

26. Srinivas N, Krause A, Kakade SM, Seeger M. Gaussian process optimization in the bandit setting: No regret and experimental design. *arXiv preprint arXiv:0912.3995* 2009.

27. Kandasamy K, Schneider J, Póczos B. High dimensional Bayesian optimisation and bandits via additive models. In: *International Conference on Machine Learning*; 2015: 295–304.

28. Goldberg PW, Williams CKI, Bishop CM. Regression with input-dependent noise: A Gaussian process treatment. *Advances in Neural Information Processing Systems* 1997; 10: 493–499.

29. Li L, Jamieson K, DeSalvo G, Rostamizadeh A, Talwalkar A. Hyperband: A novel bandit-based approach to hyperparameter optimization. *Journal of Machine Learning Research* 2018; 18(185): 1-52.

30. Wu J, Toscano-Palmerin S, Frazier P, Wilson A. Practical multi-fidelity Bayesian optimization for hyperparameter tuning. *arxiv* 2019; abs/1903.04703.

31. Falkner S, Klein A, Hutter F. BOHB: Robust and efficient hyperparameter optimization at scale. In: *International Conference on Machine Learning*; 2018: 1437–1446.

32. Crawley DB, Lawrie LK, Winkelmann FC, et al. EnergyPlus: Creating a new-generation building energy simulation program. *Energy and Buildings* 2001; 33(4): 319–331.

33. Bahramara S, Moghaddam MP, Haghifam MR. Optimal planning of hybrid renewable energy systems using HOMER: A review. *Renewable and Sustainable Energy Reviews* 2016; 62: 609–620.

34. Yu M, Miller DC, Biegler LT. Dynamic reduced order models for simulating bubbling fluidized bed adsorbers. *Industrial & Engineering Chemistry Research* 2015; 54(27): 6959–6974.

35. Chen X, Heidarinejad M, Liu J, Peña d. lDM, Christofides PD. Model predictive control of nonlinear singularly perturbed systems: Application to a large-scale process network. *Journal of Process Control* 2011; 21(9): 1296–1305.

36. Paulson JA, Martin-Casas M, Mesbah A. Optimal Bayesian experiment design for nonlinear dynamic systems with chance constraints. *Journal of Process Control* 2019; 77: 155–171.

37. Paulson JA, Mesbah A. Approximate closed-loop robust model predictive control with guaranteed stability and constraint satisfaction. *IEEE Control Systems Letters* 2020; 4(3): 719–724.

38. Bonzanini AD, Paulson JA, Graves DB, Mesbah A. Toward safe dose delivery in plasma medicine using projected neural network-based fast approximate NMPC. *IFAC-PapersOnLine* 2020; 53(2): 5279–5285.

39. Barron AR. Universal approximation bounds for superpositions of a sigmoidal function. *IEEE Transactions on Information theory* 1993; 39(3): 930–945.

40. Kumar P, Rawlings JB, Wright SJ. Industrial, large-scale model predictive control with structured neural networks. *Computers & Chemical Engineering* 2021; 150: 107291.

41. Rodriguez JS, Nicholson B, Laird C, Zavala VM. Benchmarking ADMM in nonconvex NLPs. *Computers & Chemical Engineering* 2018; 119: 315–325.

42. Teichgraeber H, Brandt A. Clustering methods to find representative periods for the optimization of energy systems: An initial framework and comparison. *Applied Energy* 2019; 239: 1283–1293.

43. Sengupta M, Xie Y, Lopez A, Habte A, Maclaurin G, Shelby J. The national solar radiation data base (NSRDB). *Renewable and Sustainable Energy Reviews* 2018; 89: 51–60.

44. Gondhalekar R, Oldewurtel F, Jones CN. Least-restrictive robust MPC of periodic affine systems with application to building climate control. In: *49th IEEE Conference on Decision and Control (CDC)*IEEE. ; 2010: 5257–5263.

45. Lofberg J. YALMIP: A toolbox for modeling and optimization in MATLAB. In: *2004 IEEE international conference on robotics and automation (IEEE Cat. No. 04CH37508)*IEEE. ; 2004: 284–289.

46. Gurobi Optimization L. Gurobi Optimizer Reference Manual. 2020.

47. Martinez-Cantin R. BayesOpt: A Bayesian Optimization Library for Nonlinear Optimization, Experimental Design and Bandits. *Journal of Machine Learning Research* 2014.

48. Paulson JA, Mesbah A. Nonlinear model predictive control with explicit backoffs for stochastic systems under arbitrary uncertainty. *IFAC-PapersOnLine* 2018; 51(20): 523–534.

49. Paulson JA, Lu C. COBALT: COnstrained Bayesian optimizAtion of computationaLly expensive grey-box models exploiting derivaTive information. *arXiv preprint arXiv:2105.04114* 2021.

50. Swaney RE, Grossmann IE. An index for operational flexibility in chemical process design. Part I: Formulation and theory. *AIChE Journal* 1985; 31(4): 621–630.

51. Mohideen M, Perkins J, Pistikopoulos E. Robust stability considerations in optimal design of dynamic systems under uncertainty. *Journal of Process Control* 1997; 7(5): 371–385.

52. Diangelakis N, Burnak B, Katz J, Pistikopoulos E. Process design and control optimization: A simultaneous approach by multi-parametric programming. *AIChE Journal* 2017; 63(11): 4827–4846.

53. Lenhoff A, Morari M. Design of resilient processing plants—I Process design under consideration of dynamic aspects. *Chemical Engineering Science* 1982; 37(2): 245–258.

54. Narraway L, Perkins J, Barton G. Interaction between process design and process control: economic analysis of process dynamics. *Journal of Process Control* 1991; 1(5): 243–250.

55. Bogle I, Rashid M. An assessment of dynamic operability measures. *Computers & Chemical Engineering* 1989; 13(11-12): 1277–1282.

56. Fraga E, Hagemann J, Estrada-Villagrana A, Bogle I. Incorporation of dynamic behaviour in an automated process synthesis system. *Computers & Chemical Engineering* 2000; 24(2-7): 189–194.

57. Fisher W, Doherty M, Douglas J. The interface between design and control. 1. Process controllability. *Industrial & engineering chemistry research* 1988; 27(4): 597–605.

58. Fisher W, Doherty M, Douglas J. The interface between design and control. 2. Process operability. *Industrial & engineering chemistry research* 1988; 27(4): 606–611.

59. Fisher W, Doherty M, Douglas J. The interface between design and control. 3. Selecting a set of controlled variables. *Industrial & engineering chemistry research* 1988; 27(4): 611–615.

60. Papalexandri K, Pistikopoulos E. Synthesis and retrofit design of operable heat exchanger networks. 1. Flexibility and structural controllability aspects. *Industrial & engineering chemistry research* 1994; 33(7): 1718–1737.

61. Georgiadis M, Schenk M, Gani R, Pistikopoulos E. The interactions of design, control and operability in reactive distillation systems. In: *Computer Aided Chemical Engineering*. 9. Elsevier. 2001 (pp. 997–1002).

62. Luyben M, Floudas C. Analyzing the interaction of design and control—1. A multiobjective framework and application to binary distillation synthesis. *Computers & Chemical Engineering* 1994; 18(10): 933–969.

63. Floudas C. Global optimization in design and control of chemical process systems. *Journal of Process Control* 2000; 10(2-3): 125–134.

64. Floudas CA, Gümüş Z, Ierapetritou M. Global optimization in design under uncertainty: feasibility test and flexibility index problems. *Industrial & Engineering Chemistry Research* 2001; 40(20): 4267–4282.

65. Figueroa J, Bahri P, Bandoni J, Romagnoli J. Economic impact of disturbances and uncertain parameters in chemical processes—A dynamic back-off analysis. *Computers & Chemical Engineering* 1996; 20(4): 453–461.

66. Walsh S, Perkins J. Application of integrated process and control system design to waste water neutralisation. *Computers & Chemical Engineering* 1994; 18: S183–S187.

67. Schweiger C, Floudas C. Interaction of design and control: Optimization with dynamic models. In: *Optimal Control*Springer. 1998 (pp. 388–435).

68. Exler O, Antelo L, Egea J, Alonso A, Banga J. A tabu search-based algorithm for mixed-integer nonlinear problems and its application to integrated process and control system design. *Computers & Chemical Engineering* 2008; 32(8): 1877–1891.

69. Asteasuain M, Bandoni A, Sarmoria C, Brandolin A. Simultaneous process and control system design for grade transition in styrene polymerization. *Chemical Engineering Science* 2006; 61(10): 3362–3378.

70. Dominguez D, Revollar S, Francisco M, Lamanna R, Vega P. Simultaneous synthesis and integrated design of chemical processes using IMC PID tuning methods. *Chemical Engineering Transactions* 2009; 17: 1371–1376.

71. Bansal V, Perkins J, Pistikopoulos E, Ross R, van Schijndel J. Simultaneous design and control optimisation under uncertainty. *Computers & Chemical Engineering* 2000; 24(2): 261 - 266.

72. Bansal V, Perkins JD, Pistikopoulos EN. A Case Study in Simultaneous Design and Control Using Rigorous, Mixed-Integer Dynamic Optimization Models. *Industrial & Engineering Chemistry Research* 2002; 41(4): 760-778.

73. Ricardez-Sandoval L, Budman H, Douglas P. Simultaneous design and control of processes under uncertainty: A robust modelling approach. *Journal of Process Control* 2008; 18(7-8): 735–752.

74. Ricardez-Sandoval L. Optimal design and control of dynamic systems under uncertainty: A probabilistic approach. *Computers & Chemical Engineering* 2012; 43: 91–107.

75. Mehta S, Ricardez-Sandoval L. Integration of design and control of dynamic systems under uncertainty: a new back-off approach. *Industrial & Engineering Chemistry Research* 2016; 55(2): 485–498.

76. Rafiei-Shishavan M, Mehta S, Ricardez-Sandoval L. Simultaneous design and control under uncertainty: A back-off approach using power series expansions. *Computers & Chemical Engineering* 2017; 99: 66–81.

77. Washington ID, Swartz CLE. Design under uncertainty using parallel multiperiod dynamic optimization. *AIChE Journal* 2014; 60: 3151–3168.

78. Flores-Tlacuahuac A, Biegler LT. Integrated control and process design during optimal polymer grade transition operations. *Computers & Chemical Engineering* 2008; 32(11): 2823–2837.

79. Koller RW, Ricardez-Sandoval LA, Biegler LT. Stochastic back-off algorithm for simultaneous design, control, and scheduling of multiproduct systems under uncertainty. *AIChE Journal* 2018; 64(7): 2379–2389.

80. Sakizlis V, Perkins JD, Pistikopoulos EN. Parametric controllers in simultaneous process and control design optimization. *Industrial & Engineering Chemistry Research* 2003; 42(20): 4545–4563.

81. Sakizlis V, Perkins JD, Pistikopoulos E. Simultaneous process and control design using mixed integer dynamic optimization and parametric programming. *Computer Aided Chemical Engineering* 2004; 17: 187–215.

82. Pistikopoulos E, Diangelakis N, Oberdieck R, Papathanasiou M, Nascu I, Sun M. PAROC—An integrated framework and software platform for the optimisation and advanced model-based control of process systems. *Chemical Engineering Science* 2015; 136: 115–138.

83. Avraamidou S, Diangelakis N, Pistikopoulos E. Mixed integer bilevel optimization through multi-parametric programming. *Foundations of computer aided process operations/chemical process control* 2017: In–Press.

84. Gutierrez G, Ricardez L, Budman H, Prada C. Integration of Design and Control using an MPC-based superstructure for control structure selection.. *IFAC Proceedings Volumes* 2011; 44(1): 7648–7653.

85. Sanchez-Sanchez K, Ricardez-Sandoval L. Simultaneous design and control under uncertainty using model predictive control. *Industrial & Engineering Chemistry Research* 2013; 52(13): 4815–4833.

86. Bahakim SS, Ricardez-Sandoval LA. Simultaneous design and MPC-based control for dynamic systems under uncertainty: A stochastic approach. *Computers & Chemical Engineering* 2014; 63: 66–81.

87. Patil B, Maia E, Ricardez-Sandoval L. Integration of scheduling, design, and control of multiproduct chemical processes under uncertainty. *AIChE Journal* 2015; 61(8): 2456–2470.

88. Noeres C, Dadhe K, Gesthuisen R, Engell S, Górak A. Model-based design, control and optimisation of catalytic distillation processes. *Chemical Engineering and Processing: Process Intensification* 2004; 43(3): 421–434.

89. Zhang J, Li K, Wang M, et al. A bi-level program for the planning of an islanded microgrid including CAES. *IEEE Transactions on Industry Applications* 2016; 52(4): 2768–2777.

90. Evins R. Multi-level optimization of building design, energy system sizing and operation. *Energy* 2015; 90: 1775–1789.

91. Lambert T, Gilman P, Lilienthal P. Micropower system modeling with HOMER. *Integration of alternative sources of energy* 2006; 1(1): 379–385.

92. Li B, Roche R, Miraoui A. Microgrid sizing with combined evolutionary algorithm and MILP unit commitment. *Applied energy* 2017; 188: 547–562.

93. Sharifzadeh M. Integration of process design and control: A review. *Chemical Engineering Research and Design* 2013; 91(12): 2515–2549.

94. Yuan Z, Chen B, Sin G, Gani R. State-of-the-art and progress in the optimization-based simultaneous design and control for chemical processes. *AIChE Journal* 2012; 58(6): 1640–1659.

95. Vega P, De Rocco RL, Revollar S, Francisco M. Integrated design and control of chemical processes–Part I: Revision and classification. *Computers & Chemical Engineering* 2014; 71: 602–617.

96. Vega P, Lamanna R, Revollar S, Francisco M. Integrated design and control of chemical processes–Part II: An illustrative example. *Computers & Chemical Engineering* 2014; 71: 618–635.

97. Ricardez-Sandoval L, Budman H, Douglas P. Application of robust control tools to the simultaneous design and control of dynamic systems. *Industrial & Engineering Chemistry Research* 2009; 48(2): 801–813.

98. Pistikopoulos EN, Diangelakis NA. Towards the integration of process design, control and scheduling: Are we getting closer?. *Computers & Chemical Engineering* 2016; 91: 85–92.

99. Burnak B, Diangelakis N, Pistikopoulos E. Towards the grand unification of process design, scheduling, and control—utopia or reality?. *Processes* 2019; 7(7): 461.

# APPENDIX

## A SUMMARY OF CONTRIBUTIONS IN INTEGRATED DESIGN AND CONTROL

Design and control have long been recognized as highly interdependent activities that should be performed simultaneously to readily identify process designs with high *operational flexibility*. Integrated design and control (IDC) refers to a broad collection of methods that systematically account for the effect of the control actions/operational decisions on the process design. Although optimal IDC problems can be generally formulated as multistage stochastic programs, these formulations are practically intractable whenever realistic problem features are considered (long system lifetimes, significant stochastic sources of uncertainty, and mixed discrete-continuous decisions). There has been a significant amount of work on tractable IDC methods is available; the vast majority of these methods can be interpreted in terms of how they go about simplifying features 1 to 3 provided in the introduction. We have attempted to organize these contributions into four main topic areas, starting with the earliest attempts at the top and moving towards more recent solution methods at the bottom. Here, we will just provide a brief discussion on the different alternatives to motivate our selected approach. Interested readers are referred to the several excellent review articles on IDC provided in the last row of Table A1. The earliest work on IDC focused on controllability indices that can be computed for a local linear dynamic approximation (in the form of a state-space or transfer function model) around a particular steady-state operating condition. These metrics could then be included as constraints (or penalized in the objective) within a standard design optimization – the main downside to these type of approaches is that they are only applicable for a limited range of operation and thus do not directly address *flexibility* by considering many different operating conditions depending on the particular realizations of the uncertainties. To address this issue, the notion of a flexibility index was introduced by Swaney and Grossmann[50], which can be formulated as a two-stage robust optimization problem. When only a discrete set of uncertainty scenarios are considered, the flexible design problem reduces to a multiperiod design problem, which assumes perfect recourse to the uncertainties and instantaneous equilibration to a steady-state in each period. To more directly handle the generally nonlinear process dynamics, Mohideen et al.[51] introduced a mixed-integer dynamic optimization (MIDO) formulation of the IDC problem assuming a proportional-integral-derivative (PID) controller structure to counteract a nominal disturbance profile. Since MIDOs take advantage of derivative-based optimization methods, it is not trivial to extend these approaches to more advanced control strategies, such as model predictive control (MPC), in which the input values are represented implicitly as the solution to an underlying optimization problem. Recent work has suggested to use multiparametric MPC (mpMPC) to derive an explicit solution to the control law that can be substituted into the MIDO[52]; however, this is only possible for fairly small-scale linear system models and quadratic cost functions, which limits its applicability to flexible energy systems.

**TABLE A1** Contributions to the field of integrated design and control (IDC).

| Topic | Control policy | Author (Year) | Contribution |
|---|---|---|---|
| Controllability and stability metrics | | Morari & co-workers (1982)[53], Perkins & co-workers (1991)[54], Bogle & co-workers (1989, 2000)[55,56], Douglas & co-workers (1988)[57,58,59] | Controllability metrics included as constraints or objectives within the optimization formulation |
| Two-stage flexible process design | | Pistikopolous & co-workers (1994, 1997, 2001)[60,51,61], Floudas & co-workers (1994, 2000, 2001)[62,63,64], Romagnoli & co-workers (1996)[65], Perkins & co-workers (1994)[66], Floudas & co-workers (1998)[67], Exler & co-workers (2008)[68], Asteasuain & co-workers (2006)[69], Vega & co-workers (2009)[70]. | Flexibility, feasibility, and resilience considerations in steady state (MI)NLP optimization under scenario approximation of uncertainties |
| MIDO | PID | Pistikopoulos & co-workers (2000, 2002)[71,72], Ricardez-Sandoval & co-workers (2008, 2012, 2016, 2017)[73,74,75,76], Swartz & co-workers (2014)[77], Biegler & co-workers (2008, 2018)[78,79] | Substitutes explicit decision rule and nominal disturbance sequence and then converts (MI)DO into (MI)NLP using collocation or decomposition |
| | Linear MPC | Pistikopoulos & co-workers (2003, 2004, 2015g, 2017g)[80,81,82,83,52], Ricardez-Sandoval & co-workers (2011, 2013, 2014, 2015)[84,85,86,87], Engell & co-workers (2004)[88], | |
| DR-SO | Adaptive thresholding | Hakizimana (2019)[6], Zhang & co-workers (2016)[89], Evins & co-workers (2015)[90], Lambert & co-workers (2006)[91] | Substitutes decision rule for control policy to derive single-stage IDC problem tackled via simulation optimization method |
| | Nonlinear MPC | Paulson & co-workers (2021)[16], Li & co-workers (2017)[92] | |
| Review articles | | Sharifzadeh(2013)[93], Gani and co-workers(2012)[94], Vega and co-workers(2014)[95,96], Ricardez Sandoval and co-workers(2009)[97], Pistikopolous and co-workers (2004, 2016, 2019)[5,98,99] | |

# B DETAILED MODEL EQUATIONS FOR BUILDING-BATTERY-GRID SYSTEM

## B.1 Building and Battery Model Equations

The building models the temperature inside a single room in a larger office building and is adapted from Gondhalekar et al.[44]:

$$\dot{x}_1^{HC} = \frac{1}{C_1}[K_3(x_2^{HC} - x_1^{HC}) + K_1(w_1^{HC} - x_1^{HC}) + K_4(x_3^{HC} - x_1^{HC}) + \tau_1 w_2^{HC} + \eta_h u^h + \eta_c u^c + w_3^{HC}], \tag{B1a}$$

$$\dot{x}_2^{HC} = \frac{1}{C_2}[K_2(w_1^{HC} - x_2^{HC}) + K_3(x_1^{HC} - x_2^{HC}) + w_2^{HC}], \tag{B1b}$$

$$\dot{x}_3^{HC} = \frac{1}{C_3}[K_4(x_1^{HC} - x_3^{HC})], \tag{B1c}$$

where $x_1^{HC}, x_2^{HC}, x_3^{HC}$ are the indoor building temperature, average interior wall temperature and exterior wall temperature, respectively. The battery state of charge (SOC) is modeled as

$$\dot{x}^B = \frac{100}{d^B}\left(-u^{HC} + u^G + w^B\right), \tag{B2}$$

where $x^B$ is the SOC in percent. All parameters and variables for these models are defined in Table B2.

## B.2 System, Economic, and Constraint Parameters

The variables appearing in the system model are summarized below in Table B2 and the time-varying constraint, price, and MPC penalty values are listed in below Table B3.

**TABLE B2** System model states, inputs, disturbances, and parameters

| Symbol | Description | Units | Value or Range |
|---|---|---|---|
| $d^{PV}$ | PV array size | [$m^2$] | [0, 540] |
| $d^{B}$ | battery capacity | [kWh] | [0, 1300] |
| $x_1^{HC}$ | room air temperature | [°C] | - |
| $x_2^{HC}$ | exterior wall temperature | [°C] | - |
| $x_3^{HC}$ | interior wall temperature | [°C] | - |
| $x^{B}$ | state of charge | [%] | - |
| $u^{HC}$ | net heating load | [kW] | [-1500,1500] |
| $u^{G}$ | energy from grid | [kW] | [-1500,1500] |
| $u^{B}$ | energy from battery | [kW] | [-1500,1500] |
| $\bar{x}^{B}$ | SOC upper bound | [%] | 95 |
| $\underline{x}^{B}$ | SOC lower bound | [%] | 10 |
| $v^{B}$ | SOC violation penalty | [$] | 0 |
| $\hat{v}^{B}$ | SOC violation penalty | [$] | $2 \cdot 10^6$ |
| $w_1^{HC}$ | outside air temperature | [°C] | [-23,35] |
| $w_2^{HC}$ | horizontal solar radiation | [kW/$m^2$] | [0,1033] |
| $w_3^{HC}$ | internal heat sources | [kW] | [25,35] |
| $\eta^{h}$ | heating efficiency | - | 4 |
| $\eta^{c}$ | cooling efficiency | - | 2 |
| $\tau_1$ | window radiation coefficient | - | 20 |
| $K_1$ | heat conductivity | [kW/° C] | 16.48 |
| $K_2$ | heat conductivity | [kW/° C] | 108.5 |
| $K_3$ | heat conductivity | [kW/° C] | 5 |
| $K_4$ | heat conductivity | [kW/° C] | 30.5 |
| $C_1$ | heat capacity | [kJ/° C] | $9.356 \times 10^5$ |
| $C_2$ | heat capacity | [kJ/° C] | $2.970 \times 10^6$ |
| $C_3$ | heat capacity | [kJ/° C] | $6.695 \times 10^5$ |
| $\pi^{PV}$ | PV price per square meter | [$/$m^2$] | 44 |
| $\pi^{B}$ | battery price per kWh | [$/kWh] | 13.6 |

**TABLE B3** Time-varying constraint, price, and penalty parameter values

| Time-varying parameters | Values in terms of daily time index, $q = t \mod 24$ | | | | Units |
|---|---|---|---|---|---|
| | $q \in \{0, \dots, 7\}$ | $q \in \{8, \dots, 18\}$ | $q \in \{19, 20\}$ | $q \in \{21, \dots, 24\}$ | |
| $\overline{x}_t^{HC}$ | 30 | 26 | 30 | 30 | [°C] |
| $\underline{x}_t^{HC}$ | 19 | 21 | 19 | 19 | [°C] |
| $\pi_t^{HC}$ | $10^{-4}$ | $10^{-2}$ | $10^{-4}$ | $10^{-4}$ | [$ / °C] |
| $\pi^{SOC}$ | 0 | 0 | 0 | 0 | [$ / %] |
| $\pi_t^{G}$ | 0.01 | 0.025 | 0.025 | 0.01 | [$ / kWh] |