Balancing Wind and Batteries: Towards Predictive Verification of Smart Grids

Thom S. Badings[1(✉)] , Arnd Hartmanns[2] , Nils Jansen[1], and Marnix Suilen[1]

[1] Department of Software Science, Radboud University, Nijmegen, The Netherlands
thom.badings@ru.nl
[2] University of Twente, Enschede, The Netherlands

Abstract. We study a smart grid with wind power and battery storage. Traditionally, day-ahead planning aims to balance demand and wind power, yet actual wind conditions often deviate from forecasts. Short-term flexibility in storage and generation fills potential gaps, planned on a minutes time scale for 30–60 min horizons. Finding the optimal flexibility deployment requires solving a semi-infinite non-convex stochastic program, which is generally intractable to do exactly. Previous approaches rely on sampling, yet such critical problems call for rigorous approaches with stronger guarantees. Our method employs probabilistic model checking techniques. First, we cast the problem as a continuous-space Markov decision process with discretized control, for which an optimal deployment strategy minimizes the expected grid frequency deviation. To mitigate state space explosion, we exploit specific structural properties of the model to implement an iterative exploration method that reuses pre-computed values as wind data is updated. Our experiments show the method's feasibility and versatility across grid configurations and time scales.

1 Introduction

Electricity grids need to constantly maintain a balance between power supply and demand; imbalances result in frequency deviations, which ultimately lead to critical events like blackouts [23]. The increasing deployment of renewable energy sources such as solar and wind power—which react sharply to hard-to-predict weather conditions—makes maintaining the balance increasingly difficult. In day-to-day operation, the balancing is managed at two time scales. One day ahead, the transmission system operator (TSO) schedules conventional generators to match the predicted demand minus the expected renewable generation based on weather forecasts. During the day, the TSO fills potential gaps introduced by any mismatch between forecast and actual weather conditions by *ancillary services*, which run on a short-term schedule that is updated every few minutes.

On the supply side, the most prominent ancillary service are *spinning reserves* from generators, which compensate for contingencies (such as generator failures)

and deviations from the predicted demand. To free capacity for spinning reserves, some generators operate below their rated capacity, making them a costly service. The TSO's day-ahead plan thus makes a tradeoff between allocating sufficient reserves to mitigate any potential imbalance and minimizing unused generator capacity [36]. *Demand-side flexibility* [3], on the other hand, is provided by various assets connected to the grid, including batteries [19] and HVAC systems [12,40]. They can reduce or increase the overall power consumption at some time point by injecting or withdrawing electricity into or from the grid. Such flexibility-based services *shift* power consumption in time rather than changing the total [22]. Examples today include the ODFM service in the UK [27] and various applications of *demand response* worldwide [1]. In this paper, we approach the fundamental challenge of short-term scheduling for flexibility-based ancillary services:

> Given a power grid with significant uncertain wind power generation, optimally schedule the deployment of the available ancillary services over a finite horizon to minimize the expected total grid frequency deviation without violating any hard constraints on grid stability and operation.

Hard constraints include a maximum frequency deviation (we use $\pm 0.1\,\mathrm{Hz}$) as well as all generation, transmission, battery, and ramping capacities. Repeating the optimization every few minutes with new wind measurements leads to a model predictive control (MPC) loop covering a full day of short-term scheduling.

The task can be expressed as a semi-infinite nonconvex stochastic optimization problem, with the potential deviation of wind conditions from the forecast given by some stochastic process. Already a finite version of this problem is NP-hard and infeasible to solve in practice [8,26]. Previous work instantiated the stochastic process by a black-box discrete-time Markov chain (DTMC) and then resorted to a sampling-based *scenario optimization* approach [24,35], which linearizes the nonconvex constraints and solves the resulting linear (but still semi-infinite) stochastic optimization problem up to some statistical confidence and error. The drawbacks of this approach are the *approximation error* introduced by the linearization step and the *statistical error* due to the use of sampling [10].

Our Contribution. To overcome the need for both sampling and linearization, we model the problem as a Markov decision process (MDP) [32]. MDPs combine probabilistic choices, which we use to follow a white-box DTMC for the wind errors, and nondeterminism, which we use to capture the service deployment decisions to optimize over. A direct cast of the problem into an MDP would yield continuous state and action spaces: state variables would represent continuous quantities (e.g., grid frequency), and control decisions would range over real-valued intervals (e.g., charge current applied to a battery). We thus (i) discretize the controllable values into finitely many control *actions*. Since control decisions are only made every few minutes, (ii) the model is discrete-time. Further, wind conditions at the current time are known, so (iii) there is a single initial state for every (iv) finite horizon. In combination, (i)–(iv) entail that we can only reach a finite subset of the continuous state space. Thus we can build an MDP with

finitely many states and actions. We use the original (non-linearized) continuous dynamics of the power grid to compute the successor state following an action. A cost function penalizes frequency deviations; violations of hard constraints lead to absorbing non-goal states. An action selection strategy that minimizes the expected accumulated cost to reach the time horizon then defines an optimal deployment of ancillary services. We track the time, making the MDP a directed acyclic graph in theory and a tree in practice. In the MPC loop, we only need the action selected in the initial (current) state; when time advances to the next control decision, we get a new initial state (based on measured wind conditions in reality and sampled from the wind error DTMC in our experiments) from which to repeat the procedure. Thus a large part of the new iteration's MDP can be reused from the previous iteration; we only need to add one more layer for the advanced time horizon. We present the continuous-state dynamics of the power system in Sect. 3, explain the formal and technical details of our MDP-based approach in Sect. 4, and report on an experimental evaluation in Sect. 5.

Our approach has three key advantages: We (1) obtain a *strategy that is sound* w.r.t. the physical constraints, i.e. it is guaranteed to satisfy the battery, generator, and ramping capacities (but not necessarily the frequency and transmission limits). The same cannot be guaranteed with scenario optimization due to the linearization and statistical error. We (2) *exploit the tree structure* of the MDP to speed up computations in the MPC loop; and (3) by relying on existing probabilistic model checking technology, the approach is *easy to extend*, for example with multiple objectives, unreliable communication, or demand uncertainty. Its main drawbacks are that, while sound and optimal for the discrete MDP, the computed strategy is sound but *may not be optimal for the continuous model*: an optimal strategy in the continuous model may require a control input that lies between the discrete options of the MDP. Moreover, the MDP's *state space grows exponentially* with the time horizon and precision of the discretization. We investigate the effects of varying degrees of discretization and time horizons on the quality of the schedule and the tractability of the problem in our experimental evaluation.

Related Work. Previous studies of demand-side flexibility consider, for example, *vehicle-to-grid* [19,42] and *buildings-to-grid* integration [21,33,34,40]. As renewable generation is primarily decentralized, regional congestion is an issue [7,31]; congestion management under uncertain generation was studied in [15,18,28]. The majority of the previously cited works use continuous-state models. Several also apply the MPC pattern [20,35,40] and generally state the optimization as an *optimal power flow* problem [23]. As mentioned, scenario optimization [9,10,24] is sampling-based; in a different approach to sampling, [2] uses Monte Carlo tree search for optimal power flow in the presence of many distributed energy resources.

When it comes to Markov models, [16] uses MDPs for optimal storage scheduling, while [13] computes MDP-based optimal charging strategies for electric vehicles. Probabilistic safety guarantees have been formally verified on DTMC models [30,38], i.e. MDPs where a fixed control strategy is embedded in the model.

[17] studies decentralized protocols in solar panels to stabilize grid frequency. Finally, we mention that our way of deriving the MDP is similar to the approach of the StocHy tool [11] and its predecessor FAUST[2] [39], which however lack support for costs/rewards and do not implement our efficient MPC loop.

2   Preliminaries

A *discrete probability distribution* over a finite set X is a function $\mu\colon X \to [0,1]$ with $\sum_{x \in X} \mu(x) = 1$. The set of all distributions over X is $Dist(X)$. We write $|X|$ for the number of elements in X. Notation $x_{1:n}$ introduces a vector $[x_1, \ldots, x_n]$.

Definition 1. *A* Markov decision process *(MDP) is a tuple* $\mathcal{M} = (S, A, s_I, T, c)$ *where S is a finite set of states, A is a finite set of actions, $s_I \in S$ is the initial state, $T\colon S \times A \rightharpoonup Dist(S)$ is the (partial) probabilistic transition function, and $c\colon S \to \mathbb{R}$ is the state-based cost function. We assume deadlock-free MDPs.*

A *discrete-time Markov chain* (DTMC) is an MDP with only one action at every state. For DTMCs, we omit the set of actions A by simply typing the transition function $T\colon S \to Dist(S)$. To define an expected cost measure on MDPs, the nondeterministic choices of actions are resolved by *strategies*. A memoryless deterministic *strategy* for an MDP is a function $\sigma\colon S \to A$. For other types of strategies, we refer to [6]. Applying strategy σ to an MDP $\mathcal{M}$ resolves all nondeterministic choices and yields an *induced DTMC* $\mathcal{M}^\sigma$. The expected cost of reaching a set of goal states $G \subseteq S$ in this induced DTMC is denoted by $\mathrm{EC}^{\mathcal{M}^\sigma}(\lozenge\, G)$. The goal is to compute a strategy that minimizes the expected cost.

3   Continuous-State Power System Modelling

The continuous power system model is a system of nonlinear differential equations. We explain its setup and components in this section. We then discretize the model w.r.t. time, obtaining continuous-state dynamics as a set of nonlinear equations.

3.1   Grid Model Dynamics

We model a power grid as an undirected graph of interconnected nodes, to which generators and loads are connected. The example grid in Fig. 1 has one generator, three nodes with connected loads, plus one wind farm and one battery. We adopt the grid model dynamics proposed in [5,35]. The dynamics in every node are given by the *active power swing equation*, which describes the balance between electrical and kinetic energy at that node in the grid [41]. The state at time t for node n is determined by the voltage angle and frequency, yielding the dynamics:

$$m_n \ddot{\delta}_n(t) + d_n \dot{\delta}_n(t) = \bar{P}_n(t) - \sum_{p \in \mathcal{N}} b_{n,p} \sin(\delta_n(t) - \delta_p(t)) \tag{1}$$
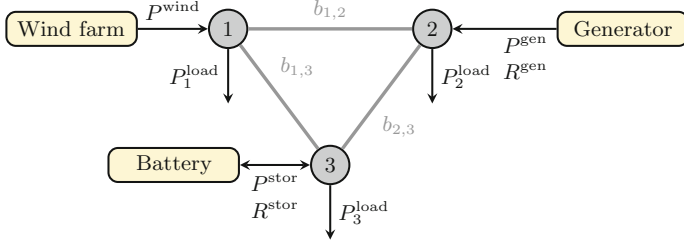
Fig. 1. 3-node example electricity grid.

where $\delta_n(t)$, $\dot{\delta}_n(t)$, $\ddot{\delta}_n(t) \in \mathbb{R}$ are the voltage angle, angular velocity (frequency), and angular acceleration of node n, and m_n and d_n are inertia and damping coefficients, respectively. $\bar{P}_n(t) \in \mathbb{R}$ is the power balance at node n, and is given by the sum of the generation and loads at that node. The power flow between node n and all connected nodes $p \in \mathcal{N}$ is assumed to be purely reactive and characterized by the line susceptance, $b_{n,p}$, and the difference in voltage angle between the connected nodes. A detailed description is available in [3,35].

3.2 Grid Frequency Control

The *grid frequency deviation* for node n, denoted $\omega_n(t)$, is the difference between the absolute frequency, $\dot{\delta}_n(t)$ in Eq. 1, and the desired frequency (for example, 50 Hz in Europe). The value of $\omega_n(t)$ can be controlled by the injection or consumption of electrical energy; any mismatch between power supply and demand results in a deviation of the frequency. As shown in Fig. 1, we distinguish five power generating or consuming assets:

- $P^{\mathrm{gen}}(t)$ is the *conventional power dispatch*. Conventional generators are subject to ramping limits, so the derivative $\dot{P}^{\mathrm{gen}}(t)$ is restricted to certain bounds.
- $P^{\mathrm{load}}(t)$, the *consumer load*, represents the known and uncontrollable demand. We can readily extend the model to controllable or uncertain demand.
- $R^{\mathrm{gen}}(t)$ is the deployment of *spinning reserves*. It is a control variable in the optimal power flow problem.
- $P^{\mathrm{wind}}(t)$, the *wind power generation*, is a random variable [29] due to its limited predictability. We define $P^{\mathrm{wind,fc}}(t)$ as the forecast wind power at time t; then the *forecast error* is $\Delta P^{\mathrm{wind}}(t) = P^{\mathrm{wind}}(t) - P^{\mathrm{wind,fc}}(t)$.
- A *battery* is an energy storage buffer whose *state of charge* (SoC) $q(t)$ follows the injection or consumption of energy. $P_{\mathrm{stor}}(t)$ is the uncontrollable power input variable, known from the day-ahead plan, and $R_{\mathrm{stor}}(t)$ is the demand-side flexibility power rate, which is a control variable in the power flow problem.

3.3 Ancillary Service Deployment

Ancillary service deployment is subject to two restrictions: reserves and storage flexibility must be scheduled a day ahead, and the deployment can never exceed

the scheduled amount. The following constraints ensure these restrictions:

$$- R_{\text{ds}}^{\text{gen}}(t) \leq R^{\text{gen}}(t) \leq R_{\text{us}}^{\text{gen}}(t) \tag{2}$$

ensures that the deployment of spinning reserves is within the scheduled bounds, where $R_{\text{ds}}^{\text{gen}}(t) \geq 0$ ($R_{\text{us}}^{\text{gen}}(t) \geq 0$) is the scheduled amount of down-spinning (up-spinning) reserves. Similarly we have for flexibility that

$$- R_{\text{dd}}^{\text{stor}}(t) \leq R^{\text{stor}}(t) \leq R_{\text{id}}^{\text{stor}}(t), \tag{3}$$

where $R_{\text{dd}}^{\text{stor}}(k) \geq 0$ and $R_{\text{id}}^{\text{stor}}(k) \geq 0$ are the scheduled decreased- and increased-demand flexibility, respectively.

3.4   Discrete-Time Storage-Integrated Power System Model

We discretize the continuous dynamics with respect to time to render the problem of optimal frequency control tractable. The resulting model is a set of nonlinear equations, which albeit discretized w.r.t. time are still defined on continuous state and control spaces. They describe the transition from one continuous state and control input to the resulting continuous state one discrete time step later.

Consider a power grid with n_t nodes, n_g generators, n_f wind farms, and n_s batteries for storage. Its continuous dynamics are given as a system of $2n_t + n_g + n_s$ first-order differential equations. The features of the continuous state space are given by the voltage angles $\delta_{1:n_t}$ and frequencies $\omega_{1:n_t}$ for all n_t nodes, the power generation $P_{1:n_g}^{\text{gen}}$ for all n_g generators, and the state of charge $q_{1:n_s}$ of all n_s batteries. The vector of control variables contains the *change* in generator dispatch $\dot{P}_{1:n_g}^{\text{gen}}$ and the reserve deployment $R_{1:n_g}^{\text{gen}}$ for all n_g generators, plus the flexibility deployment $R_{1:n_s}^{\text{stor}}$ for all n_s batteries. We discretize w.r.t. time via the first-order backward Euler implicit method [37] to obtain the nonlinear function

$$x(k+1) = f\big(x(k), u(k), v(k), w(k)\big), \tag{4}$$

- with $f(\cdot)$ reflecting the dynamics of the considered power system, which are nonlinear due to the sinusoid, $\sin(\delta_n(t) - \delta_p(t))$, in Eq. 1,
- $x(k) = [\delta_{1:n_t}, \omega_{1:n_t}, P_{1:n_g}^{\text{gen}}, q_{1:n_s}] \in \mathbb{R}^{2n_t + n_g + n_s}$ the state vector,
- $u(k) = [\dot{P}_{1:n_g}^{\text{gen}}, R_{1:n_g}^{\text{gen}}, R_{1:n_s}^{\text{stor}}] \in \mathbb{R}^{2n_g + n_s}$ the vector of control variables,
- $v(k) = [P_{1:n_t}^{\text{load}}, P_{1:n_f}^{\text{wind,fc}}, P_{1:n_s}^{\text{stor}}] \in \mathbb{R}^{n_t + n_f + n_s}$ the uncontrollable known inputs, and
- $w(k) = [\Delta P_{1:n_f}^{\text{wind}}] \in \mathbb{R}^{n_f}$ the vector of uncontrollable random variables.

We omit further details for the sake of brevity and refer the interested reader to [5] and [40] for the full derivation and discretization of similar grid models.

Power Balance and Control Variables. The day-ahead generator power dispatch is scheduled such that generation plus wind power forecast matches the

consumer load pattern. We impose the following constraint at every time point k:

$$\sum_{i=1}^{n_g} P_i^{\mathrm{gen}}(k) + \sum_{m=1}^{n_f} P_m^{\mathrm{wind,fc}}(k) = \sum_{n=1}^{n_t} P_n^{\mathrm{load}}(k). \qquad (5)$$

Since $P_m^{\mathrm{wind,fc}}$ and P_n^{load} are known, imposing this equality constraint yields $n_g - 1$ independent control variables. Hence, in a single-generator grid, the day-ahead planning is fixed by the wind forecast and consumer load, while in a grid with multiple generators, the required total dispatch must be divided between the different units. In a similar manner, during the day itself, the reserve power and storage flexibility can be deployed together to restore the mismatch in the power balance caused by forecast errors at any time point:

$$\sum_{i=1}^{n_g} R_i^{\mathrm{gen}}(k) + \sum_{m=1}^{n_f} \Delta P_m^{\mathrm{wind}}(k) = \sum_{n=1}^{n_s} R_n^{\mathrm{stor}}(k), \qquad (6)$$

where $\Delta P_m^{\mathrm{wind}}$ is a random variable, and both R_i^{gen} and R_n^{stor} are control variables. Hence, imposing Eq. 6 yields $n_g + n_s - 1$ independent control variables.

Power System Constraints. The discrete-time power system model in Eq. 4 is subject to a number of constraints. First of all, the equality constraints in Eqs. 5 and 6 are imposed to enforce the balance between power supply and demand. Second, power lines have limited transmission capacity, and generators have limited generation capacity and ramping capability. Third, the deployment of reserve power and storage flexibility is limited by their scheduled values, as described in Eqs. 2 and 3. Finally, the electrical storage units have a limited capacity, and can only be charged or discharged at a given maximum rate. For the explicit formulation of the constraints we refer the interested reader to [35].

4 Discrete-State Receding Horizon Control Problem

Our goal is to overcome the need for sampling and linearization to optimize ancillary service deployment. Directly using the model presented in Eq. 4 would require dealing with continuous state and action spaces. Our approach is to discretize the actions, then explore the resulting finite number of (continuously-valued) successor states up to a given exploration depth. In this way, we obtain a finite MDP that can be solved iteratively using a receding horizon principle. This approach is a discrete-state model predictive control technique. We now present the wind error DTMC, followed by the details of our exploration procedure and the formal definition of the finite-horizon MDP.

4.1 Stochastic Wind Power Model

Recall that the wind power forecast error, $\Delta P_m^{\mathrm{wind}}(k) \in \mathbb{R}$, of every individual wind farm $m \in \{1, \ldots, n_f\}$ is a continuous random variable in Eq. 4. Exploiting the time discretization of Sect. 3.4, we construct a DTMC for the forecast error of every wind farm with time resolution equal to the discretization level of the

(a) Using 5-minute based data.   (b) Using 15-minute based data.
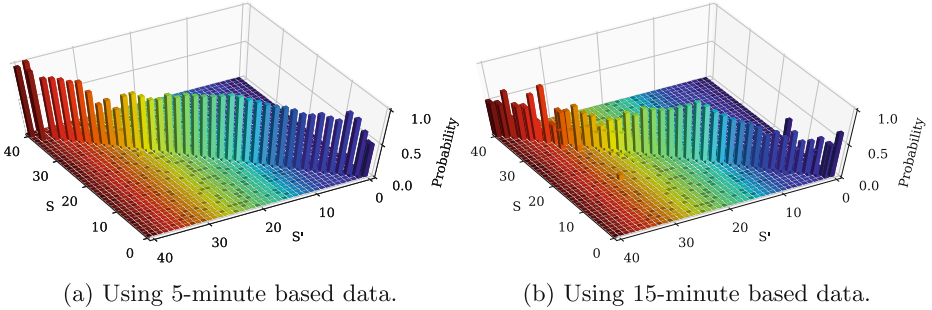
Fig. 2. Transition probability function of the wind error DTMC.

dynamics. For this section, we assume the presence of only one wind farm to simplify notation. Let S be the finite state space of the DTMC, and let function $M \colon \mathbb{R} \to S$ map the wind power forecast error $\Delta P^{\mathrm{wind}} \in \mathbb{R}$ to a state $s_w \in S$: every value of the continuous random variable ΔP^{wind} is approximated by the value of a discrete state $s_w \in S$. The DTMC's transition function $T \colon S \to Dist(S)$ describes the probability that a transition occurs from one state s_w at any time k to another state s'_w at time $k+1$ with probability $T(s_w)(s'_w)$.

Historical Wind Power Data. We follow the method proposed in [25,29] to construct the DTMC based on the historical wind power forecast error. The states S are based on a uniform discretization of the historical forecast error. The transition probabilities in T are determined using a maximum likelihood estimation, by counting the number of transitions from one state s_w at time k to a successor state s'_w at time $k+1$. We use five years (2015–2019) of on-shore wind power data measured every 15 min from the TenneT region of the German transmission grid, obtained from the *ENTSO-E Transparency Platform* [14]. We interpolate the data to a 5-min basis, to match the time resolution we employ in the numerical demonstration in Sect. 5. The data set contains the wind power forecast, $P^{\mathrm{wind,fc}}$, and the actual power, P^{wind}, thus providing a five-year time series of the observed wind power error, ΔP^{wind}. M then maps every continuous value of ΔP^{wind} to one of 41 discrete states, i.e. $S = \{s_w^1, \ldots, s_w^{41}\}$, as in [25]. We show the resulting transition matrices in Fig. 2 for the original 15-min and the interpolated 5-min data. The matrix is diagonally dominant, reflecting the strong auto-correlation of the forecast error.

4.2 State Space Exploration Procedure

Next, we present our method to explore the continuous state space of the model up to a given exploration horizon. We discretize the continuous control variables such that their potential values define a set of actions for an MDP, then a priori eliminate those actions that lead to states violating the constraints of Sect. 3.4.

Let $\mathsf{X}_k = (x(k), s_w)$ denote the continuous state vector $x(k)$ at time k, and the DTMC state s_w associated with the current wind power forecast error. To define an initial state, we use a concrete measurement at the first time step. Then, we use the dynamics in Eq. 4 in combination with transitions of the wind error DTMC to compute the possible successor states for each time step. At time step k we select a control input $u(k)$ and, based on the dynamics in Eq. 4, receive a set of successor states with different $x(k+1)^1, \ldots, x(k+1)^n$, one associated with every possible wind error successor state, $s_w^1, \ldots, s_w^n$ with $T(s_w)(s_w^i) > 0$ for all $1 \le i \le n$. The features in $x(k+1)^i$ related to power generation ($P_{1:n_g}^{\mathrm{gen}}$) and battery SoC ($q_{1:n_s}$) are equal for all $1 \le i \le n$, while the grid features ($\delta_{1:n_t}, \omega_{1:n_t}$) depend on the wind successor state, s_w^i. As the probabilities to reach these successor states depend exclusively on the probabilities defined by the DTMC, we reach state $\mathsf{X}_{k+1} = (x(k+1)^i, s_w^i))$ with probability $T(s_w)(s_w^i)$.

Feasible Control Space. Under the two balance constraints in Eqs. 5 and 6, the vector of control variables $u(k)$ in Eq. 4 contains $2n_g + n_s - 2$ independent variables. Since the dynamics in Eq. 4 and the constraints imposed on the system are known, we can determine the continuous subset of control inputs $u \in \mathcal{U}_{\mathsf{X}_k} \subset \mathbb{R}^{2n_g+n_s-2}$ that do not lead to a violation of any of the constraints at time $k+1$. Note that this set depends on the state X_k at time k. Given the subset of feasible continuous control inputs, we apply a grid-based discretization in all $2n_g+n_s-2$ dimensions, to obtain a set A_{X_k} of *feasible and discrete actions* at time k, where the subscript denotes the dependency on X_k. From this discretization of the control actions, an important property follows. Given the current state X_k at time k, every action $a \in A_{\mathsf{X}_k}$ has a different set of *discrete successor states* X_{k+1} at time $k+1$. By only exploring the continuous state-space for the feasible actions in A_{X_k}, we minimize the size of the resulting finite-state model.

A schematic example of this procedure for two state features (one node frequency and one battery SoC) is shown in Fig. 3. The blue dot corresponds to current state X_k at time k, and the straight arrows show the discrete actions $a \in A_{\mathsf{X}_k}$. The curved arrows show the effect of the forecast error to the grid frequency, which depends directly on the actual successor state in the DTMC. Because all successor states of a_1 violate the maximum SoC constraint ($q(k) \le q^{\mathrm{max}}$), this action can be *eliminated a priori*. Similarly, all successors for a_5 violate the maximum frequency deviation limit ($\omega(k) \le \omega^{\mathrm{max}}$). Actions a_2 and a_3 will not lead to any violation, and are, therefore, included as feasible discrete actions. Action a_4 may or may not violate the constraints depending on the wind power forecast error. Therefore, this action cannot be eliminated a priori.

Exploration. The finite exploration horizon with starting time k and look-ahead of K_h steps is the set $\{k, \ldots, k + K_h\}$. The exploration procedure can be performed recursively for all discrete actions $a \in A_{\mathsf{X}_k}$ for every k, until the desired horizon is reached. We obtain a tree-structured model as in Fig. 4, where the *branching factor* depends on the number of actions and possible wind successor states, and the *depth* is given by the horizon length. For brevity, the
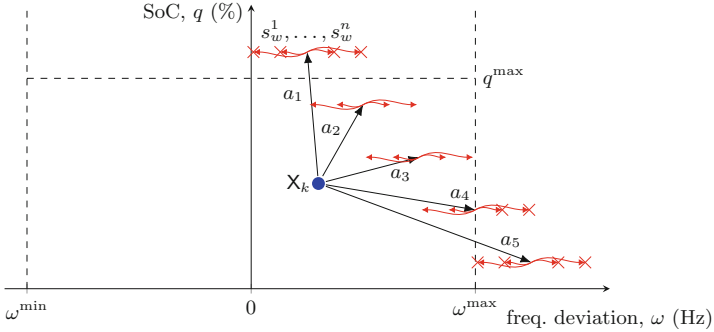
Fig. 3. Discrete actions $a_1, \ldots, a_5$ and their mapping to two continuous state-space features: a battery SoC and the frequency in one node. Curved arrows show the effect of wind successor states, and dashed lines are system constraints.

dependency of the set of discrete actions on the current state is omitted in this figure.

Recall that we consider only those states $X_k = (x(k), s_w)$ of the continuous state space that are visited during the exploration. We provide the full definition of the MDP $\mathcal{M} = (S, A, s_I, T, c)$ with exploration horizon $\{k, \ldots k + K_h\}$, where

– every state $s \in S$ is associated with an X_k from the continuous state space;
– the set of actions A is the union of all feasible action sets A_{X_k} for all X_k;
– the initial state s_I is given by a concrete measurement (x_0, s_w);
– the partial probabilistic transition function $T \colon S \times A \rightharpoonup Dist(S)$ maps state-action pairs to the corresponding distributions over successor states according to the wind error DTMC;
– the cost function $c \colon S \to \mathbb{R}$ assigns the immediate cost given by the sum of the absolute value of the frequency deviation in every grid node.

Finally, we define the set $G \subseteq S$ of *goal states* as the states that are reached at the end of the horizon $k + K_h$ and satisfy the constraints.

Receding Horizon and Tree Structure. We compute the strategy σ that induces the minimal expected cost $\text{EC}^{\mathcal{M}^\sigma}(\lozenge G)$ for the MDP with horizon $\{k, \ldots k + K_h\}$. Using this strategy, we implement the (first) optimal action $a_0 = \sigma(s_0)$ where s is associated with X_k. Having executed the optimal action at time k, we then apply the so-called *receding horizon principle*, meaning that we *shift* the exploration horizon one step forward in time. We update the MDP for the shifted horizon, which is now given as $\{k + 1, \ldots k + K_h + 1\}$, and again compute an optimal strategy. Hence, with every shift of the exploration horizon in time, we *reveal small bits of new information* near the end of the horizon.

As an example, consider an initial exploration horizon defined as the time frame between 9:00–9:15 AM. Computing an optimal strategy for the corresponding MDP means that we take all the information within that time frame into
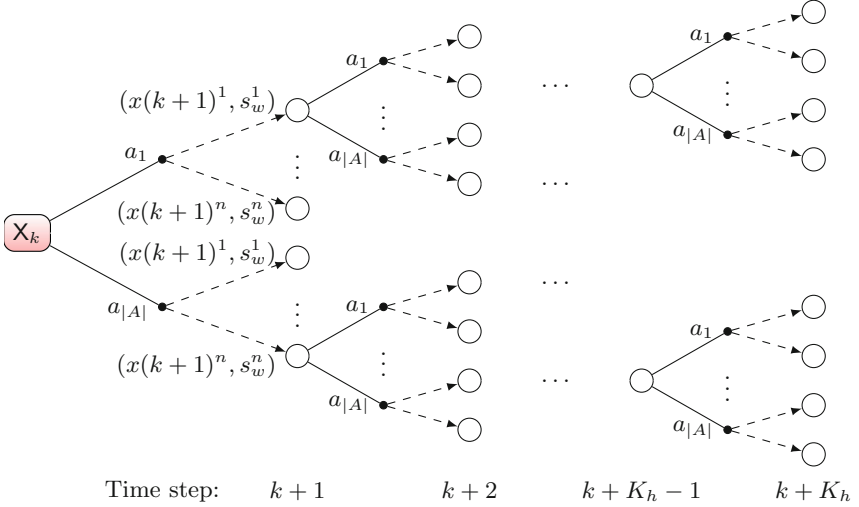
Fig. 4. Visualization of the state-space exploration procedure, with initial measurement X_k at time k, and finite horizon with end time $k + K_h$. Solid edges represent discrete actions, whereas dashed lines reflect different successor states.

account, i.e. we have perfect knowledge within the horizon. However, a possible change in the power demand (or any other uncontrollable input) that is forecast at 9:30 AM is not revealed to the model, until the exploration horizon also spans that time step. Defining an adequate length of the exploration horizon reflects a trade-off between model size and the optimality of the solution.

When we shift the exploration horizon from $\{k, \ldots, k+K_h\}$ to $\{k+1, \ldots, k+K_h+1\}$, another layer is added to the MDP. As the starting time of the horizon also shifts, we gain new information via a new measurement at $k + 1$, and then obtain a single new initial state for the MDP with the new horizon. By exploiting the tree structure of the MDP, we extend the current MDP with a new layer of successor states and take the subtree starting at the newly found initial state at time $k + 1$. Simultaneously, we shrink the layer at the top, by pruning all states that cannot be reached anymore from the new initial state.

The approach of iteratively solving and updating the MDP for shifting horizons is similar to a discrete-state model predictive control technique. This iterative method is applied until a desired simulation horizon is reached (such as 24 h).

5   Numerical Study

We demonstrate the performance of our approach on multiple variants of the 3-node example grid already introduced in Fig. 1. To this end, we *simulate the wind error DTMC*, and apply our method to build the MDP using the sampled wind
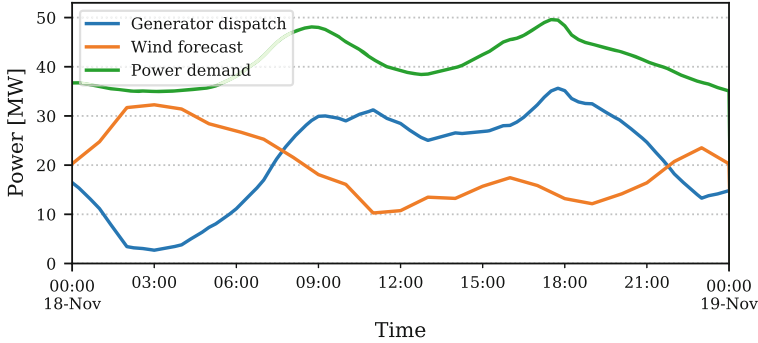
Fig. 5. The day-ahead power balance between generator dispatch, wind power forecast, and the power demand, which is equivalent for all performed simulations.

power forecast error at every time step. As such, every wind power trajectory sampled from the DTMC results in a (potentially different) run of our approach. We solve the MDP to obtain the first action in the sequence of optimal decisions over the exploration horizon, which is then executed to obtain the initial state at the next time step. By applying the receding horizon principle, we iteratively follow this procedure, until a final simulation horizon of 24 h is reached.

5.1   Experimental Setup

We consider a simulation time resolution of 5 min, and a full simulation horizon of 24 h. Using the receding horizon exploration procedure, this means that $\frac{60}{5} \cdot 24 = 288$ MDPs are solved to obtain the results over one 24-h run. The same demand and wind power forecast are used for every simulation, resulting in the day-ahead power dispatch shown in Fig. 5. This figure shows that under the forecast conditions, the power supply and demand are perfectly balanced, leading to a stable grid frequency in the absence of forecast errors. Both the power demand and wind power forecast are based on historical data obtained from the *ENTSO-E Transparency Platform* [14], and are scaled appropriately for the simulation study. The scheduled flexibility deployment limits per battery are set to ± 2 MW, while the scheduled reserve limits per generator are ± 0.25 MW.

Simulation Cases. Simulations are performed on three variants of the example grid in Fig. 1. In the simplest case, we aggregate the three nodes into one, resulting in a grid where all assets are connected to the same node. The second variant is the exact network shown in Fig. 1. The third variant is an extension of the second, where a second storage unit is connected to node 1.

Then, we perform simulation studies for different values of: a) the *exploration horizon* (300, 600, and 900 s), and b) the *levels of discretization for the actions* denoted by λ (with values between 3 and 25 steps). To evaluate the quality of the solution of our technique, we perform a statistical analysis on every case, by

repeating every experiment 100 times. Note that this Monte Carlo type simulation is merely used to evaluate the quality of the obtained solutions, and not required to apply our technique in practice.

Data Availability. The code and data needed to reproduce the results presented below are archived at DOI 10.4121/14185139 [4]. Our prototype implementation is written in Python version 3.8.3, and allows the user to run the cases defined above or simulate with any other parameter setting. Our simulations ran on a Windows laptop with Intel Core i7-1065G7 CPU (1.3–3.0 GHz) and 16 GB RAM.

5.2   Results

Run Times. The observed run times are approximately proportional to the number of MDP states, and grow exponentially with the exploration horizon. For the 3-node system with 1 battery, an action discretization level of $\lambda = 5$ steps, and exploration horizon of 300 s, the average run time *per iteration* of the receding horizon (i.e. for solving one MDP) is 0.01 s, resulting in around 3.51 s per 24 h run. For the same case with longer exploration horizons of 600 and 900 s, average run times are 0.14 and 2.09 s, respectively, per receding horizon iteration.

The strongest increase in run time is observed for the 3-node, 2 battery case, where the average observed times were 0.03, 1.03, and 223.59 s for exploration horizons of 300, 600, and 900 s, respectively. This steeper increase is explained by the exponential growth of the model size with respect to the number of actions, which is higher for the case with 2 batteries.

Failed Runs. As visualized before in Fig. 3, a discrete action can lead to a successor state that violates one of the system constraints. In total, 4.2% of the performed iterations for all cases combined resulted in a violation of either the frequency limits or the power line transmission capacity. No significant differences in the percentage of failed runs is observed across the different cases. Since these runs are incomplete and not representative for the simulated cases, they are not reported for further analysis. Nevertheless, the number of failed runs provides an empirical indication of the adequacy of the scheduling limits for the reserve and flexibility deployment. A high percentage of failed runs means that the ancillary service scheduling limits might be insufficient, and should be enlarged.

Solution Quality. The cost function of the MDP penalizes the sum of the absolute value of the frequency deviations in all grid nodes. Therefore, the quality of the obtained solution can be evaluated by taking the integral of the total observed frequency deviation over the simulation horizon:

$$J = \int_{k_0}^{k_{\text{end}}} \big| \sum_{n=1}^{n_t} \omega_n(k) \big| dk,$$
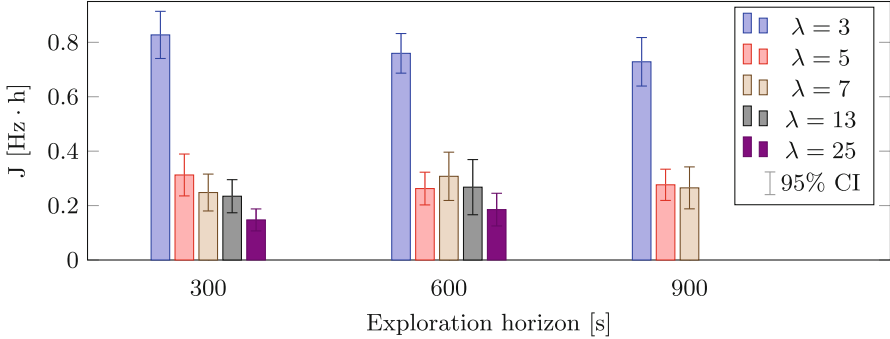
Fig. 6. Average total frequency deviations for the 3-node case for different λ.

where k_0 and k_{end} cover the full 24-h simulation time. The lower the value of J, the better the quality of the solution in terms of total frequency deviations.
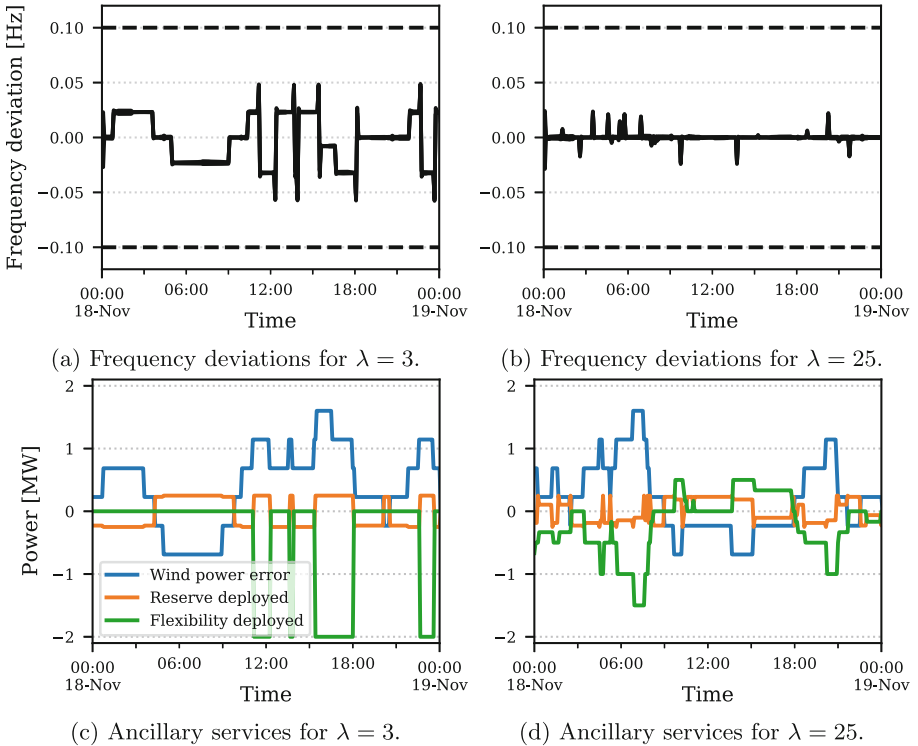
A comparison of the value for J between multiple cases on the 3-node network from Fig. 1, with different levels of $\lambda \in \{3, \ldots, 25\}$, is shown in Fig. 6. Every bar shows the average results and the 95% confidence interval (CI) of the 100 iterations performed for that case. Due to infeasible run times, simulations for the cases with exploration horizon of 900 s and $\lambda \geq 13$ were omitted.

We observe that (1) a *longer exploration horizon does not improve the quality of the solution* in terms of frequency deviations significantly. On the other hand, (2) *increasing the number of discrete actions* in every dimension yields a significantly better solution quality. This observation suggests that it is more beneficial to have a more fine-grained discretization of the continuous control space, than to invest in a longer optimization horizon.

MDP Model Size. In Fig. 7, the average number of states and actions per MDP are compared for different network configurations, all with $\lambda = 5$. All cases were simulated for 100 runs, except for the case with 2 batteries, which was only simulated for 1 run, due to a too long run time. We see that the number of states and actions is *independent of the number of grid nodes*. The intuition behind this is that an increased number of state features does not yield a larger MDP. In fact, the branching factor of the applied exploration procedure only depends on the number of actions and successor states in the wind error DTMC, and is independent of the number of state features. As expected, increasing the number of batteries to 2 yields a significant increase in the number of MDP states and actions, especially for a longer exploration horizon, due to the additional dimension of the continuous control space.
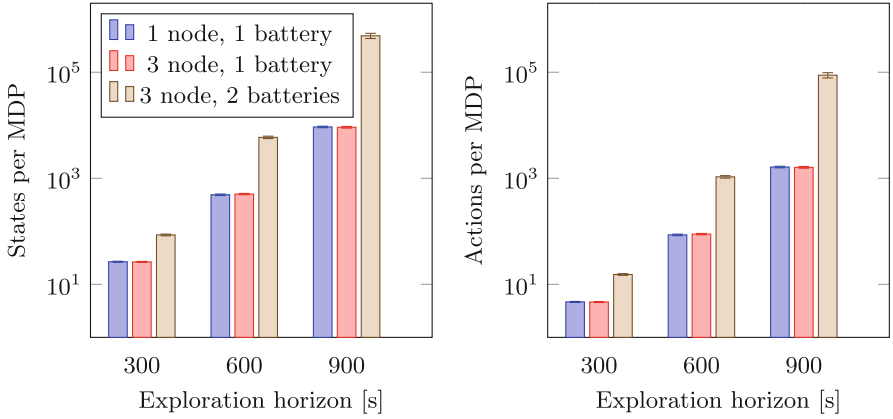
Fig. 7. Average MDP states (left) and actions (right) for cases with $\lambda = 5$, and varying numbers of nodes and batteries (plotted in log scale).

(a) Frequency deviations for $\lambda = 3$.

(b) Frequency deviations for $\lambda = 25$.

(c) Ancillary services for $\lambda = 3$.

(d) Ancillary services for $\lambda = 25$.

Fig. 8. Results for two runs for the 3-node case with 1 battery, for different λ.

Frequency Control and Ancillary Service Deployment. Finally, in Fig. 8, two example runs for the 3-node network with a single battery and exploration horizon of 600 s are shown. Figure 8a presents the frequency deviations for $\lambda = 3$, while Fig. 8b shows the same for $\lambda = 25$. Similarly, Figs. 8c and 8d show the deployment of reserves and flexibility for both cases. Due to the uncertainty in the wind power forecast error, the results in Fig. 8 present only two possible trajectories, and repeating the experiment can lead to different results.

The observed frequency deviations are significantly lower for the case with more fine-grained discretization of the actions (i.e. $\lambda = 25$), thus confirming the results also shown in Fig. 6. This difference in the control precision is clearly depicted between Figs. 8c and 8d. The intuition behind this is that a fine-grained discretization allows for more precise control of the ancillary service deployment, which results in lower frequency deviations.

6   Concluding Remarks

We presented a novel method to solve the problem of short-term scheduling for flexibility-based ancillary services in power systems with uncertain wind power generation. By modelling the problem as an MDP, we overcome the need for both sampling and linearization, as opposed to the continuous-state approaches used by most traditional power system analysis methods. Our experiments show that our approach is feasible for power grids with different levels of complexity and under realistic operating conditions. Furthermore, our results show it is more beneficial to have a more fine-grained discretization of the continuous control space, than to invest in a longer optimization horizon. Since the size of the MDP grows exponentially with both the number of actions and the exploration horizon, making a trade-off between the two is necessary.

In the future, we will exploit the flexibility of our model to incorporate alternative grid configurations and multiple sources of uncertainty, such as imperfect communication between assets in the grid, or demand uncertainty. Moreover, instead of the batteries, other flexible assets can also be modeled, such as flexibility provided by the thermal inertia of large-scale buildings.

References

1. Aghaei, J., Alizadeh, M.I.: Demand response in smart electricity grids equipped with renewable energy sources: a review. Renew. Sustain. Energy Rev. **18**, 64–72 (2013)
2. Al-Saffar, M., Musílek, P.: Distributed optimal power flow for electric power systems with high penetration of distributed energy resources. In: CCECE 2019, pp. 1–5. IEEE (2019)
3. Badings, T.S.: MSc Thesis. Buildings-to-Grid Integration for Demand-Side Flexibility in Power Systems with Uncertain Generation. University of Groningen (2019)
4. Badings, T.S.: Balancing wind and batteries: towards predictive verification of smart grids (artifact). 4TU.ResearchData (2021). https://doi.org/10.4121/14185139

5. Badings, T.S., Rostampour, V., Scherpen, J.M.: Distributed building energy storage units for frequency control service in power systems. IFAC-Pap. OnLine **52**(4), 228–233 (2019)
6. Baier, C., Katoen, J.P.: Principles of Model Checking. MIT Press, Cambridge (2008)
7. Bertsch, J., Hagspiel, S., Just, L.: Congestion management in power systems: long-term modeling framework and large-scale application. J. Regul. Econ. **50**(3), 290–327 (2016)
8. Boyd, S.P., Vandenberghe, L.: Convex Optimization. Cambridge University Press, Cambridge (2014)
9. Calafiore, G.C., Campi, M.C.: The scenario approach to robust control design. IEEE Trans. Autom. Control. **51**(5), 742–753 (2006)
10. Campi, M.C., Garatti, S.: The exact feasibility of randomized solutions of uncertain convex programs. SIAM J. Optim. **19**(3), 1211–1230 (2008)
11. Cauchi, N., Abate, A.: StocHy: Automated verification and synthesis of stochastic processes. In: Vojnar, T., Zhang, L. (eds.) TACAS 2019. LNCS, vol. 11428, pp. 247–264. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-17465-1_14
12. Chertkov, M., Chernyak, V.: Ensemble of thermostatically controlled loads: statistical physics approach. Scientific reports **7**(1), 1–9 (2017)
13. Ding, T., Zeng, Z., Bai, J., Qin, B., Yang, Y., Shahidehpour, M.: Optimal electric vehicle charging strategy with Markov decision process and reinforcement learning technique. IEEE Trans. Ind. Appl. **56**(5), 5811–5823 (2020)
14. ENTSO-e: Transparency Platform - Generation Forecasts for Wind and Solar, Control area Germany (2020)
15. Gerard, H., Rivero Puente, E.I., Six, D.: Coordination between transmission and distribution system operators in the electricity sector: a conceptual framework. Utilities Policy **50**, 40–48 (2018)
16. Grillo, S., Pievatolo, A., Tironi, E.: Optimal storage scheduling using Markov decision processes. IEEE Trans. Sustain. Energy **7**(2), 755–764 (2016)
17. Hartmanns, A., Hermanns, H., Berrang, P.: A comparative analysis of decentralized power grid stabilization strategies. In: Winter Simulation Conference, pp. 158:1–158:13. WSC (2012)
18. Hemmati, R., Saboori, H., Jirdehi, M.A.: Stochastic planning and scheduling of energy storage systems for congestion management in electric power systems including renewable energy resources. Energy **133**, 380–387 (2017)
19. Kempton, W., et al.: A test of Vehicle-to-Grid (V2G) for Energy Storage and Frequency Regulation in the PJM System. Results from an Industry-University Research Partnership (2008)
20. Liu, Y., et al.: Coordinating the operations of smart buildings in smart grids. Appl. Energy **228**(July), 2510–2525 (2018)
21. Lymperopoulos, I., Qureshi, F.A., Nghiem, T., Khatir, A.A., Jones, C.N.: Providing ancillary service with commercial buildings: the Swiss perspective. IFAC-Pap. OnLine **28**(8), 6–13 (2015)
22. MacDougall, P., Roossien, B., Warmer, C., Kok, K.: Quantifying flexibility for smart grid services. In: 2013 IEEE Power Energy Society General Meeting, pp. 1–5. IEEE (2013)
23. Machowski, J., Dong, Z.Y., Member, S., Zhang, P.: Power System Dynamics: Stability and Control. Wiley, Hoboken (2006)
24. Margellos, K., Goulart, P., Lygeros, J.: On the road between robust optimization and the scenario approach for chance constrained optimization problems. IEEE Trans. Autom. Control. **59**(8), 2258–2263 (2014)

25. Margellos, K., Haring, T., Hokayem, P., Schubiger, M., Lygeros, J., Andersson, G.: a robust reserve scheduling technique for power systems with high wind penetration. In: Proceedings of PMAPS, pp. 870–875 (2012)
26. Murty, K.G., Kabadi, S.N.: Some NP-complete problems in quadratic and nonlinear programming. Math. Program. **39**(2), 117–129 (1987)
27. NG ESO: Optional downward flexibility management (ODFM) service documents. National Grid ESO (2020)
28. Nguyen, D.B., Scherpen, J.M.A., Bliek, F.: Distributed optimal control of smart electricity grids with congestion management. IEEE Trans. Autom. Sci. Eng. **14**(2), 494–504 (2017)
29. Papaefthymiou, G., Klöckl, B.: MCMC for wind power simulation. IEEE Trans. Energy Convers. **23**(1), 234–240 (2008)
30. Peruffo, A., Guiu, E., Panciatici, P., Abate, A.: Safety guarantees for the electricity grid with significant renewables generation. In: Parker, D., Wolf, V. (eds.) QEST 2019. LNCS, vol. 11785, pp. 332–349. Springer, Cham (2019). https://doi.org/10.1007/978-3-030-30281-8_19
31. Pillay, A., Prabhakar Karthikeyan, S., Kothari, D.P.: Congestion management in power systems - a review. Int. J. Electr. Power Energy Syst. **70**, 83–90 (2015)
32. Puterman, M.L.: Markov Decision Processes: Discrete Stochastic Dynamic Programming. Wiley Series in Probability and Statistics, Wiley, Hoboken (1994)
33. Razmara, M., Bharati, G.R., Shahbakhti, M., Paudyal, S., Robinett, R.D.: Bilevel optimization framework for smart building-to-grid systems. IEEE Trans. Smart Grid **9**(2), 582–593 (2018)
34. Rostampour, V., Badings, T.S., Scherpen, J.M.A.: Buildings-to-grid integration with high wind power penetration. In: CDC, pp. 2976–2981. IEEE (2019)
35. Rostampour, V., Badings, T.S., Scherpen, J.M.A.: Demand flexibility management for buildings-to-grid integration with uncertain generation. Energies **13**(24) (2020)
36. Rostampour, V., Ter Haar, O., Keviczky, T.: Distributed stochastic reserve scheduling in ac power systems with uncertain generation. IEEE Trans. Power Syst. **34**(2), 1005–1020 (2018)
37. Sincovec, R.F., Erisman, A.M., Yip, E.L., Epton, M.A.: Analysis of descriptor systems using numerical algorithms. IEEE Trans. Autom. Control. **26**(1), 139–147 (1981)
38. Soudjani, S.E.Z., Abate, A.: Aggregation and control of populations of thermostatically controlled loads by formal abstractions. IEEE Trans. Control. Syst. Technol. **23**(3), 975–990 (2015)
39. Soudjani, S.E.Z., Gevaerts, C., Abate, A.: FAUST^2: formal abstractions of uncountable-state stochastic processes. In: TACAS. LNCS, vol. 9035, pp. 272–286. Springer (2015)
40. Taha, A.F., Gatsis, N., Dong, B., Pipri, A., Li, Z.: Buildings-to-grid integration framework. IEEE Trans. Smart Grid **10**(2), 1237–1249 (2019)
41. Trip, S., Bürger, M., Persis, C.D.: An internal model approach to frequency regulation in inverter-based microgrids with time-varying voltages. In: CDC, pp. 223–228. IEEE (2014)
42. Wang, J., Liu, C., Ton, D., Zhou, Y., Kim, J., Vyas, A.: Impact of plug-in hybrid electric vehicles on power systems with demand response and wind power. Energy Policy **39**(7), 4016–4021 (2011)