Comput Manag Sci (2018) 15:87–110
https://doi.org/10.1007/s10287-017-0297-2

ORIGINAL PAPER

Stochastic dynamic programming approach to managing power system uncertainty with distributed storage

Luckny Zéphyr[1] · **C. Lindsay Anderson**[1]

Received: 27 September 2016 / Accepted: 22 December 2017 / Published online: 30 December 2017

Abstract Wind integration in power grids is challenging because of the uncertain nature of wind speed. Forecasting errors may have costly consequences. Indeed, power might be purchased at highest prices to meet the load, and in case of surplus, power may be wasted. Energy storage may provide some recourse against the uncertainty of wind generation. Because of their sequential nature, in theory, power scheduling problems may be solved via stochastic dynamic programming. However, this scheme is limited to small networks by the so-called curse of dimensionality. This paper analyzes the management of a network composed of conventional power units and wind turbines through approximate dynamic programming, more precisely stochastic dual dynamic programming. A general power network model with ramping constraints on the conventional generators is considered. The approximate method is tested on several networks of different sizes. The numerical experiments also include comparisons with classical dynamic programming on a small network. The results show that the combination of approximation techniques enables to solve the problem in reasonable time.

Keywords Power grid management · Energy storage · Stochastic dynamic programming · Stochastic dual dynamic programming · Approximate dynamic programming · Generalized linear programming

✉ Luckny Zéphyr
lz395@cornell.edu

C. Lindsay Anderson
cla28@cornell.edu

[1] Department of Biological and Environmental Engineering, Cornell University, Riley Robb Hall, Ithaca, NY 14853, USA

1 Introduction

Cost-effective management of power units is a very challenging task. Generating units must be committed such that demand for electricity is met. This is more difficult to carry out because of significant variability among loads and generation sources. Commitment of traditional units such as nuclear and coal-fired plants involve cost, mostly due to fuel consumption, and is source of environmental concern. Driven by both increasing environmental awareness and technological advances, over the last decade, wind-based electricity generation has been widely promoted (Zhou et al. 2014). For instance, in 2014, the European Council set three new targets for 2030, namely: "(i) a target to reduce the EU greenhouse gas emissions by 40% relative to emissions in 1990, (ii) a renewable energy target of at least 27% at Union level; and (iii) an indicative target for energy efficiency of at least 27% at Union level." (European Commission 2016). As a result of political emphasis, in 2006, on the need to increase the United States energy efficiency and to diversify the energy portfolio, a collaborative initiative was created to explore the requirements for a 20% wind integration scenario (share of wind power in the total power generation) by 2030 (Lindenberg 2009). In the United States, the renewable energy share of total electricity generation grew from 13.7% in 2015 to over 15% in 2016 (Renewable Energy Policy Network for the 21st Century, 2017).

However, unlike conventional power sources, output from wind turbines is unpredictable. Thus, these resources cannot be relied upon to serve both supply and reliability needs in situations of network stress, such as generator or line failures. As a consequence, wind turbines cannot reliably replace conventional generators to meet all loads (Moura and De Almeida 2010). The intermittency of wind generation may create an imbalance between the supply and demand for power. In situations of very high generation, combined with inaccurate forecasting, wind power may be curtailed to maintain balance. Counter-intuitively, in cases of transmission congestion, wind generation may be curtailed while traditional resources in other locations are increased to meed the demand.

Energy storage devices may serve as recourse to circumvent the uncertainty of wind power. Such devices may be used to store excess generation, or for arbitrage profits. Indeed power might be purchased at a lower price during off-peak hours, to be stored and sold at a higher price during peak demand for power hours (Mokrian and Stephen 2006). For an in-depth description of existing and in-development storage technology, see Tan et al. (2013), Wee (2013) and Mahlia et al. (2014).

A significant stream of research has been focusing on harvesting wind power in the presence of storage. The benefits of coupling intermittent renewable energy with storage are discussed in Grillo et al. (2012), Heussen et al. (2012), Castillo and Gayme (2014) and Suberu et al. (2014). Prior studies have also analyzed the coupling of wind generation with storage via stochastic programming, e.g. Garcia-Gonzalez et al. (2008), Abbey and Joós (2009) and Meibom et al. (2011). Unlike other research where a fixed wind generation curve is used, Succar et al. (2012) analyze the optimal wind turbine rating and the storage configuration of a wind farm coupled to compressed air energy storage. Scott and Powell (2012) compare approximate dynamic programming schemes for energy storage management based on instrumental variables and projected

Bellman errors for a model ignoring transmission lines in the presence of a single storage device. In Zhou et al. (2014), the valuation of a single storage unit provided by a stochastic approach is compared with several valuation heuristics for a model with limited transmission capacity. Storage of energy for arbitrage opportunities is studied by Thatte et al. (2013), Bradbury et al. (2014) and Khani and Zadeh (2015).

The foregoing studies are limited, since they ignore network constraints, and only consider a single storage facility and a single wind turbine. Indeed, power flow equations increase the size of the problem, and both number of wind turbines as well as storage facilities increase the dimension of the *state space*. The dimensionality is a critical issue particularly in the context of stochastic dynamic programming (SDP). Several approximate schemes have been proposed to tackle the inherent dimensionality issue of SDP. One such approach is the stochastic dual dynamic programming (SDDP), particularly widespread in the community of mid- and long-term hydropower management (e.g.Pereira and Pinto 1985; Goor et al. 2010; Maceira et al. 2008; Homem-de Mello et al. 2011; Löhndorf et al. 2013). The convergence of this algorithm is proved by Philpott and Guan (2008), and more recently, Philpott and de Matos (2012) and Shapiro et al. (2013) extended SDDP to embed risk measures. The SDDP algorithm is implemented in this paper.

This paper addresses the shortcomings of these previous studies by: (i) examining the management of a power network composed of conventional units, wind turbines, as well as storage devices, taking into account network constraints; (ii) exploiting the information-decision structure of the problem to derive an SDP formulation, and an approximate dynamic programming scheme using ideas from generalized linear programming, cutting plane methods, and SDDP. In addition, we (iii) provide a detailed description of the problem as well as discuss approaches for its solution, (iv) compare the approximate scheme to classical stochastic dynamic programming, and (v) numerically analyze the scalability of the approximate scheme.

In Sect. 2, the operation of a power system equipped with energy storage facilities is detailed. The problem is formulated under the framework of SDP in Sect. 3. In Sect. 4, some general approximate dynamic programming techniques are reviewed, while Sect. 5 deals with the approximation techniques implemented in this paper. Results of numerical experiments are reported in Sect. 6.

2 Operation and control of a power system with storage

Optimization models are often used to determine an optimal generation schedule for conventional generators over a finite planning horizon, T, in order to meet the demand in each time step t at minimal cost. This is usually achieved by formulating the problem as a two-stage optimization problem, wherein the first stage decisions are made about unit commitment, or the on/off status, of the conventional generators based on forecasts of electricity demand and wind generator output (see Ostrowski et al. 2012; Morales-España et al. 2013). The second stage decisions are concerned with *economic dispatch*, which is the determination of optimal generation output of the committed generators, as well as recourse actions such as reserve requirements. Reserves are pre-determined excess generation capability, held to ensure real-time balance between the supply and

demand when imbalances arise from forecasting errors, or other unforseen events such as generator failure or transmission line disruption. The time step is typically an hour or a 15-min interval, and the planning usually spans 24, 48 or 168 h (a week).

Based on load (energy demand) and wind generation forecasts, the operator of the network makes the decisions about the commitment of the conventional generators in each time period of the horizon. We assume that decisions are made before observation of the random variables (wind generator output). Note that this assumption will be important for the formulation of the objective function as presented below. Here, the focus is on the economic dispatch problem alone, assuming that the conventional generators are previously committed.

2.1 Problem formulation

A power network may be represented by a graph $\mathcal{G} := (N, L)$, where each node $n \in N$ represents a bus, where components such as generators and loads are connected to the system, and each link a transmission line. In each node a set of conventional generators G_n, a wind farm, and a storage device may be located. The different components of the graph are distinguished via the following set of indices: (i) G is the set of conventional generators, (ii) $G_n \subseteq G$ is the subset of conventional generators located at bus n, (iii) $M \subseteq N$ is the set of wind farms, and (iv) $S \subseteq N$ denotes the set of storage facilities.

The generators are mechanical devices, and have to operate within finite generating limits. While in motion, a minimal output may be required for a generator to be in steady state. Similarly, a threshold is imposed on the maximal output to avoid mechanical damages. For any generator $g \in G$, $\underline{p}_g$ and $\bar{p}_g$ (in MW) denote lower and upper bounds, respectively. In addition, there usually are minimum and maximum allowable changes in generation between two consecutive periods. For any generator $g \in G$, the following defines ramp down and ramp up limits: $-\underline{\lambda}_g \leq p_{gt} - p_{g,t-1} \leq \bar{\lambda}_g$. In addition to power generated in a node, power can flow between two nodes through transmission lines defined as the set of undirected pairs $L \subseteq N \times N$ (assuming power can flow in either direction). For any node $n \in N$, define $O_n := \{(n, j) \in L\}$ to be the set of transmission lines that leave it; similarly, let $I_n := \{(j, n) \in L\}$ be the set of transmission lines that enter node n.

For any node n, $\hat{D}_{nt}$ denotes a "reliable" forecast (in MW) of the load (demand for power), in period t at bus n. This demand is supplied by the generation in the node, or power transported to the node through the transmission lines, or the stored energy, if any, or some combination of power from the three; thus, p_{gt} denotes power generation, in MW, from generator g in period t at the cost $CP_{gt}(p_{gt})$. The output from the wind turbines is random. Indeed, generation from such units is conditioned on wind speed, which depends on uncontrollable meteorological conditions. Therefore, $\tilde{W}_t = (\tilde{W}_{1t}, \ldots, \tilde{W}_{|M|t})$ denotes the random vector of outputs from the wind turbines in period t; w_{mt}, in MW, is a particular realization of the *stochastic process* $\{\tilde{W}_{mt}\}$. For any set X, |X| denotes its cardinality.

We note by e_{lt} the power, in MW, flowing through transmission line l in period t. The transmission lines having limited capacity, upper bounds are imposed on the power flowing through to prevent disruption; $\bar{e}_l, l \in L$, denote such bounds. Power can flow

in either direction of the line, therefore, for any transmission line $l := (n, n') \in L$, in period t, if e_{lt} is positive, power is transmitted from node n to node n'; a negative value indicates the opposite. The power flowing through a transmission line is proportional to the difference between the phase angles, in radians, of the two end buses, i.e., $e_{lt} = B_l(\theta_{n't} - \theta_{nt})$ (Papavasiliou and Oren 2013), where B_l is the susceptance of line l, i.e. "the measure of how much a circuit is susceptible to conducting a changing current."[1] The system includes a reference/slack bus, which can absorb and emit power (e.g. in case of losses), whose voltage angle is usually set to zero to guarantee that the system is not overdetermined.

In each period t, in addition to generation decisions, the network operator also makes decisions on the use of the storage. In each node where a storage device is located, in each period, the latter is charged or discharged to compensate for imbalance between the supply and demand. Such imbalance is more likely to occur in peak demand hours, i.e., periods where electricity consumption is highest. Peak demand hours vary by both geographic region and season. Operators usually prepare for peak demand by committing extra power plants that may be called upon quickly in periods of higher demand or equipment failures.

Define s_{nt} to be the level of stored energy (MW h) at bus n in the beginning of period t (or the end of period $t-1$). This energy is converted into power (MW) via the simple equation Energy = Power × Time. Δ_{nt}^{+} (resp. Δ_{nt}^{-}) denotes the positive (resp. negative) variation in the level of charge from the beginning through the end of period t. The storage units have limited capacity, and in order to last, cannot be completely depleted. Consequently, the level of charge can only be varied progressively over time. Therefore denote by $\underline{s}_n$ and $\bar{s}_n$, lower and upper bounds, respectively, on the level of the storage at bus n. Assume that the discharge and charge maximum capacity is the same, in period t, the variation in the level of the storage device at bus n then obeys:

$$0 \leq \Delta_{nt}^{+} \leq \bar{\Delta}_n$$
$$0 \leq \Delta_{nt}^{-} \leq \bar{\Delta}_n$$

Lastly, each storage facility has input and output efficiency; energy is lost both at charging and discharging. Let c_n (resp. d_n) be the efficiency coefficient of charging (resp. discharging) of the storage device at bus n, where $0 < c_n \leq 1$, and $0 < d_n \leq 1$.

Imbalance between the load and supply may result because of wind power forecasting errors. Since the storage units have limited capacity, they may not be able to absorb the total production surplus (charging) or to deliver the difference (discharging) in case of power shortage. For any node $n \in N$, $\Delta_{nt} \in \mathbf{R}$ denotes the power absorbed or delivered by the storage unit at bus n. In case of imbalance, we assume that power excess (shortage) is absorbed (delivered) at a high rate. For each node $n \in N$, $\kappa_{nt}^{+} \in \mathbf{R}^{+}$ ($\kappa_{nt}^{-} \in \mathbf{R}^{+}$) denotes such excess (shortage); $\mathbf{R}$ and $\mathbf{R}^{+}$ are the set of real and non-negative real numbers, respectively. Thus, in each period t, the power balance equations read

[1] http://www.allaboutcircuits.com/textbook/alternating-current/chpt-5/susceptance-and-admittance/.

$$\sum_{k=1}^{|G_n|} p_{kt} + \sum_{l \in I_n} e_{lt} + w_{nt} - \delta_n \Delta_{nt} - \kappa_{nt}^+ + \kappa_{nt}^- = \sum_{l \in O_n} e_{lt} + \hat{D}_{nt}, \forall n \in N,$$

with

$$\delta_n = \begin{cases} 1 \text{ if there is a storage facility at node n,} \\ 0 \text{ otherwise.} \end{cases}$$

By convention, $w_{nt} = 0$, if there is no wind farm at node n. Note that $\Delta_{nt}^+ = \max\{0, c_n \Delta_{nt}\}$, and $\Delta_{nt}^- = \max\left\{0, -\frac{\Delta_{nt}}{d_n}\right\}$. The storage level then evolves according to

$$s_{n,t+1} = \alpha_n s_{nt} + \Delta_{nt}^+ - \Delta_{nt}^-, \forall n \in N,$$

where $\alpha_n \in (0, 1]$ is the storage efficiency of the device located at bus n. There may be a cost, $CS_{nt}(\Delta_{nt})$, associated with varying the stored energy level. Such cost is proportional to the amount of energy stored or discharged (Schoenung 2011).

A summary of the notation is provided in the "Appendix".

2.2 Optimization program

We aim to find the *policy* $(P^*, \Delta^*) = \left[(P_1^*, \Delta_1^*), \ldots, (P_T^*, \Delta_T^*)\right]$ that minimizes the expected operating cost over the entire planning horizon, where $P_t^* = (p_{1t}^*, \ldots, p_{|G|t}^*)$, and $\Delta_t^* = (\Delta_{1t}^*, \ldots, \Delta_{|N|t}^*)$, $1 \leq t \leq T$. (P^*, Δ^*) then solves

$$\min_{X_t} \left\{ \mathbb{E}\left[\sum_{t=1}^{T} \left(\sum_{g=1}^{|G|} CP_{gt}(p_{gt}) + \sum_{j=1}^{|N|} \delta_n CS_{nt}(\Delta_{nt}) + M \sum_{n=1}^{|N|} (\kappa_{nt}^+ + \kappa_{nt}^-) \right) \right] \right\} \quad (1)$$

S.t., for $1 \leq t \leq T$:

$$\sum_{k=1}^{|G_n|} p_{kt} + \sum_{l \in I_n} e_{lt} + w_{nt} - \delta_n \Delta_{nt} - \kappa_{nt}^+ + \kappa_{nt}^- = \sum_{l \in O_n} e_{lt} + \hat{D}_{nt}, \; n \in N \quad (2)$$

$$e_{lt} = B_l(\theta_{n't} - \theta_{nt}), \qquad l = (n, n') \in L \quad (3)$$

$$-\bar{e}_l \leq e_{lt} \leq \bar{e}_l, \qquad l \in L \quad (4)$$

$$s_{n,t+1} = \alpha_n s_{nt} + \Delta_{nt}^+ - \Delta_{nt}^-, \qquad n \in N \quad (5)$$

$$\underline{s}_n \leq s_{n,t+1} \leq \bar{s}_n, \qquad n \in N \quad (6)$$

$$0 \leq \Delta_{nt}^+ \leq \bar{\Delta}_n, \qquad n \in N \quad (7)$$

$$0 \leq \Delta_{nt}^- \leq \bar{\Delta}_n, \qquad n \in N \quad (8)$$

$$\Delta_{nt} = \frac{\Delta_{nt}^+}{c_n} - d_n \Delta_{nt}^-, \qquad n \in N \quad (9)$$

$$\kappa_{nt}^+, \kappa_{nt}^-, \geq 0, \qquad n \in N \quad (10)$$

$$-\underline{\lambda}_g \leq p_{gt} - p_{g,t-1} \leq \bar{\lambda}_g, \qquad g \in G \tag{11}$$

$$\underline{p}_g \leq p_{gt} \leq \bar{p}_g, \qquad g \in G \tag{12}$$

where $X_t := (P_t, \Delta_t, \kappa_t, s_{t+1})$, $\mathbb{E}$ is the expectation operator, which is taken over $\tilde{W}_t$. M is a big number.

The multi-period problem (1–12) can theoretically be solved to optimality if functions CP_{gt} are convex in p_{gt}, CS_{nt} being proportional to Δ_{nt}. Indeed, the set $\Psi_t = \left\{(P_t, e_t, \tilde{W}_t, \theta_t, s_t, \Delta_t^+, \Delta_t^-, \Delta_t, \kappa_t^+, \kappa_t^-)|(2-12)\right\}$ is a polyhedron. Therefore, by convexity of the cost functions, (1–12) is a convex problem. Note however that due to change of fuel or *valve-point effects*, the cost functions may be non-smooth (e.g. Park et al. 2005; Sayah and Zehar 2008), though convex polynomial approximations are often used (Pereira-Neto et al. 2005; dos Santos Coelho and Mariani 2006).

The problem may not be tractable numerically if we want a detailed representation for the underlying process of $\{\tilde{W}_t\}$. Wind speed (wind power) are often assumed to be serially correlated. For simplicity, authors (e.g. Luh et al. 2014; Yu et al. 2015) assume a lag-one Markovian process, in the sense that the wind in the current time period is only conditioned on realizations at the previous time period. This assumption is removed in Papaefthymiou and Klockl (2008), where the wind process is assumed to obey a general order-n Markovian process. In the same vein, general ARMA models, where current wind realization are assumed to be conditioned on p previous realizations and random shocks, are among the most popular proposals in the literature of wind modeling/forecasting (see Pinson et al. 2013).

3 Representation under the framework of stochastic dynamic programming

Power network management is a sequential decision problem. In each period t, the operator observes the level of the stored energy, and generation in the previous period, and based on updated forecast for the wind turbines outputs (or wind speed) and the observation of the previous outputs, the operator determines the output of the conventional generators and the variation in the storage (charging or discharging) in order to meet the demand in each node. Thus, the tuple (s_t, p_{t-1}, w_{t-1}) will be called the *state of the system*, or state for short.

Let us observe that whereas p_{t-1} is known, from constraints (11–12), $p_{g,t-1} - \underline{\lambda}_g$ and $\underline{p}_g$ are two minorants for p_{gt}, $g \in G$. Similarly, $p_{g,t-1} + \bar{\lambda}_g$ and $\bar{p}_g$ are two majorants for p_{gt}. Consequently,

$$\max\{p_{g,t-1} - \underline{\lambda}_g, \underline{p}_g\} \leq p_{gt} \leq \min\{p_{g,t-1} + \bar{\lambda}_g, \bar{p}_g\}.$$

In principle, SDP is suited for problem (1–12). SDP sequentially decomposes (by period) the overall problem into smaller subproblems in a coordinated way, by seeking the best trade-off between the immediate and future use of the storage. Define $F_t(s_t, p_{t-1}, w_{t-1})$ to be the *cost-to-go function* from the beginning of period t to the end of the horizon. In addition, assume that the process $\{\tilde{W}_t\}$ is Markovian, i.e.,

$\mathbb{P}(\tilde{W}_t = w_t | \tilde{W}_{t-1} = w_{t-1}, \ldots, \tilde{W}_0 = w_0) = \mathbb{P}(\tilde{W}_t = w_t | \tilde{W}_{t-1} = w_{t-1})$. Therefore, for $t = T, T-1, \ldots, 1$, an SDP recursion associated with problem (1–12) is given by

$$F_t(s_t, p_{t-1}, w_{t-1}) := \min_{X_t} \left\{ h_t(p_t, \Delta_t, \kappa_t) + \mathbb{E}_{\tilde{W}_t | w_{t-1}} \left[F_{t+1}(s_{t+1}, p_t, \tilde{W}_t) \right] \right\} \quad (13)$$

$$\text{S.t. } (2-10) \quad (14)$$

$$p_{gt} \geq \max\{p_{g,t-1} - \underline{\lambda}_g, \underline{p}_g\}, \quad g \in G \quad (15)$$

$$p_{gt} \leq \min\{p_{g,t-1} + \bar{\lambda}_g, \bar{p}_g\}, \quad g \in G \quad (16)$$

where $h_t(p_t, \Delta_t, \kappa_t) := \sum_{g=1}^{|G|} CP_{gt}(p_{gt}) + \sum_{j=1}^{|N|} \delta_n CS_{nt}(\Delta_{nt}) + M \sum_{n=1}^{|N|} (\kappa_{nt}^{+} + \kappa_{nt}^{-})$.

The complexity of problem (13–16) stems from two fronts, namely (i) computing an expectation, and (ii) finding an optimal policy $\Pi^*(s_t, p_{t-1}, w_{t-1})$. The complexity of (i) is related to the dimension of the random vector $\tilde{W}_{t-1}$, ranging from one to ten in the numerical experiments, and step (ii) may be prohibitive because of the dimension of the joint state space $(s_t, p_{t-1}, \tilde{W}_{t-1})$. As a result, (13–16) is not tractable even for modest size networks. Approximation schemes must then be used.

4 Approximate dynamic programming approaches

A common approach to solve problem (13–16) is to discretize s_t, p_{t-1}, and $\tilde{W}_{t-1}$ spaces into "partial grids", and to solve the problem over the Cartesian product of these grids. Since its inception, DP has been limited to small instance of problems because of the *curse of dimensionality*, as coined by its author, Bellman. The computation burden increases exponentially as the number of states increases, which significantly limits the practicality of SDP as a solution method for real world problems. This is reinforced by the fact that in certain circumstances, the problem is to be solved periodically within constrained time frame. For instance, in the case of the economic dispatch problem, the operator of the system may need to adjust generating decisions based on signals from the market, or may have to resolve the problem periodically as a result of data changes or updated forecast as he/she gets to receive new observations for the random variables.

The practical limitation of DP paved the way for approximate DP (ADP) schemes. The purpose is usually to strike a balance between solution time and reasonable performance of the prescribed policy by replacing the function $F_t(s_t, p_{t-1}, w_{t-1})$ with some approximation $\hat{F}_t(s_t, p_{t-1}, w_{t-1})$. Though he did not refer to as ADP, Bellman was the first to propose approximations of what he called the *functional equation* [e.g., here Eq. (13)]. Using Lagrange relaxation technique, Bellman (1956) (see also Bellman and Dreyfus 1962) discusses successive approximations (SA) of the functional equation by partitioning the computation of the original sequence of functions into the computation of a sequence of functions of fewer state variables. Korsak and Larson (1970) provide a detailed algorithm for the SA technique, and settle conditions for the convergence to true optimal solution. Examples of applications of such technique

to power production are available in Zurn and Quintana (1975), Turgeon (1980) and Yang and Chen (1989).

As indicated earlier, any DP implementation resorts to discretization of the state space. One straightforward approximation of the cost-to-go function is to select, in each period, a small sample of states instead of a dense grid, and solve problem (13–16) for each point of the sample, assuming that the minimization as well as the expectation computation can be carried out efficiently. Then the cost-to-go function may be approximated for any out-of-grid point, for instance by interpolation (e.g., linear, multilinear) of the neighboring grid states (see Johnson et al. 1993; Keane and Wolpin 1994; Philbrick and Kitanidis 2001). Based on convexity assumptions on the cost-to-go function and assuming that the state space is a hyperrectangle, Zéphyr et al. (2015) propose a simplicial approximation scheme guided by local estimations of the approximation error. However, the complexity of hypercube decomposition limits the scope of the method. The hyperrectangle assumption is dropped in Zéphyr et al. (2017), but the methodology is still limited to modest dimension problems.

Several approximation schemes fall within the broader class of *parametric approximation*, wherein the approximation may be noted $\hat{F}_t(s_t, p_{t-1}, w_{t-1}, \delta)$; δ is a set of parameters or weights, which usually are to be determined. For instance, in Bellman (1957) such parameters are Legendre polynomial coefficients. Other polynomial types traditionally used to approximate the cost-to-go function include orthogonal, Chebyshev, spline, and Hermite polynomials. For more account on polynomial approximations, see Howitt et al. (2002) and Topaloglu and Powell (2006). Interpolations for out-of-grid states are performed in Philbrick and Kitanidis (2001) within hypercubes using weighted sums of the cost-to-go function evaluations as well as derivatives (first and second order) of the function at the vertices. The weights are defined as polynomials in the state variables. This multi-dimensional interpolation approach stems from previous works by Kitanidis (1987), and is an extension of the *gradient dynamic programming* scheme by Foufoula-Georgiou and Kitanidis (1988).

Combining ideas from various fields such as *neural networks*, *artificial intelligence*, cognitive sciences, and so on, *reinforcement learning* iteratively constructs approximations to the optimal cost-to-go function, or its expected value (Bertsekas and Tsitsiklis 1995) through simulations; this methodology is applied by Johri and Filipi (2011) and Momoh et al. (2005), to power control decision problems. Though there exist several reinforcement learning techniques, the most popular is the *Q-learning* algorithm, which in contrast with DP, computes the cost-to-go function for a set of randomly selected decisions considering only "visited states". For further details on contextual applications of this algorithm to power system problems, see Naghibi-Sistani et al. (2006), Lee and Labadie (2007) and Tan et al. (2009). Reinforcement learning is built around *policy iteration*, and *value iteration*, two widespread DP algorithms (Bertsekas and Tsitsiklis 1995). The former algorithm iteratively alternates between *policy evaluation* and *policy improvement*, until no further improvement can be achieved (see Qiu and Pedram 1999; Anderson et al. 2011). The latter algorithm is a later name in the DP literature for successive approximation (see Song et al. 2000; Anderson et al. 2011).

A complete review of ADP methods is beyond the scope of this paper.

5 Solution strategy

In this paper a two step strategy is used to solve problem (13–16). First, in Sect. 5.1, a linearization technique based on generalized linear programming is discussed. Second, in Sect. 5.2, an outer polyhedral approximation to the expected cost-to-go is described. This leads to an algorithmic implementation via SDDP in Sect. 5.3.

5.1 Linear approximation

Problem (13–16) is linearized via inner *generalized linear programming* (GLP) (see Shapiro 1979). For each generator $g \in G$, the cost is computed exactly over a sample of generation points $\left\{\hat{p}_{gj} | j \in \Lambda_g\right\}$, and interpolated elsewhere. As shown in Sect. 5.2, the expectation will be replaced by an outer approximation; therefore, the following is a linear approximation to (13–16):

$$\tilde{F}_t(s_t, p_{t-1}, w_{t-1}) := \min_{X_t, \beta_t, \rho_{t+1}} \left\{ \sum_{g=1}^{|G|} \sum_{j \in \Lambda_g} \beta_{gj,t} CP_g(\hat{p}_{gj}) + g_t(p_t, \Delta_t, \kappa_t) + \rho_{t+1} \right\} \tag{17}$$

$$\text{S.t. } (14)-(16) \tag{18}$$

$$p_{gt} = \sum_{j \in \Lambda_g} \beta_{gj,t} \hat{p}_{gj}, \qquad g \in G \tag{19}$$

$$\sum_{j \in \Lambda_g} \beta_{gj,t} = 1, \qquad g \in G \tag{20}$$

$$\beta_{gj,t} \geq 0, \qquad g \in G, j \in \Lambda_g \tag{21}$$

$$\rho_{t+1} \geq \mathbb{E}_{\tilde{W}_t | w_{t-1}} \left[F_{t+1}(s_{t+1}, p_t, \tilde{W}_t) \right] \tag{22}$$

with $g_t(\Delta_t, \kappa_t) := \sum_{j=1}^{|N|} \delta_n CS_{nt}(\Delta_{nt}) + M \sum_{n=1}^{|N|} (\kappa_{nt}^+ + \kappa_{nt}^-)$. In each period t, for each generator g, Eq. (19), interpolates the production over the sample of generating points $\left\{\hat{p}_{gj} | j \in \Lambda_g\right\}$ using convex combination coefficients $\beta_{gj,t}$ as computed in Eqs. (20), (21). In addition, it is clear that (18–22) is a finite intersection of *closed half-spaces*, hence it is a *polyhedral set*. By construction, $\tilde{F}_t(s_t, p_{t-1}, w_{t-1})$ is piecewise linear, since in the objective, we seek the best interpolation for the cost functions. Hence, $\tilde{F}_t(s_t, p_{t-1}, w_{t-1})$ is convex, and so is its expectation (by the linearity property of the expectation).

Let us now briefly turn to the implementation of the linear approximation scheme. For each generator, a non-linear analytical expression for the cost function is available. One straightforward way to build sample points may be by discretizing the production levels to obtain the samples $\left\{\hat{p}_{gj} | j \in \Lambda_g\right\}$, and evaluate the cost for each point of each sample. Note that the sample points need not be feasible since the production decisions are constrained in the optimization problem. Out-of-sample points $\left\{\hat{p'}_{gj} | j \in \Lambda'_g\right\}$ are

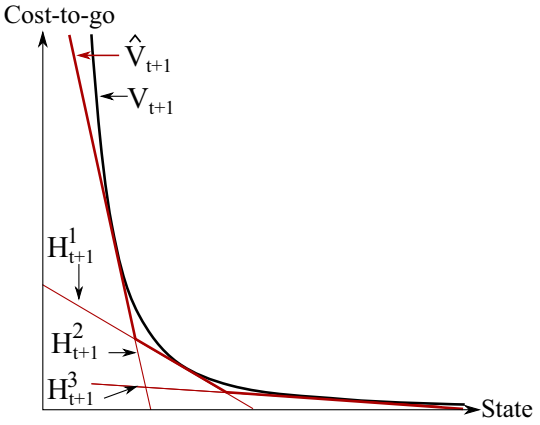

Fig. 1 Illustration of outer approximation of the expected cost-to-go approximation. The approximation is given by the maximum of the supporting hyperplanes of the epigraph of the expected cost-to-go function

then selected and their exact cost (using the true cost functions) is compared to their interpolated cost over the first sample (that will be used to implement GLP). The sample points $\{\hat{p}_{gj} | j \in \Lambda_g\}$ may be densified until the maximum interpolation error is less than a predefined threshold, yet noting that ideally, the sizes of the samples of points should be kept as small as possible as each sample point entails a new decision variable (a convex combination coefficient). Note that the generation of the sample points is carried out once for all, and is independent of the SDP loop.

5.2 Outer approximation to the expected cost-to-go

The next challenge is the computation of the expectation. In general, the expectation cannot be computed exactly, because of the large dimension of the state space. Assuming the expected cost-to-go to be convex, the latter will be replaced with an outer polyhedral approximation.

Let $V_{t+1}(s_{t+1}, p_t, w_t) := \mathbb{E}_{\tilde{W}_t|w_{t-1}}[\tilde{F}_{t+1}(s_{t+1}, p_t, \tilde{W}_t)]$. By the convexity of $V_{t+1}(s_{t+1}, p_t, w_t)$ and the feasible domain, there exist points $(s^i_{t+1}, p^i_t, w^i_t)$, such that $H^i_{t+1}(s_{t+1}, p_t, w_t), i \in I$, are *supporting hyperplanes* of the *epigraph* of $V_{t+1}(s_{t+1}, p_t, \tilde{W}_t)$ (see Theorem 3.2.5 in Bazaraa et al. 2013, p. 107). As a result $\hat{V}_{t+1}(s_{t+1}, p_t, w_t) := \max_i \{H^i_{t+1}(s_{t+1}, p_t, w_t) | i \in I\}$ is a lower bound of the true expected cost-to-go $V_{t+1}(s_{t+1}, p_t, w_t)$. Figure 1 illustrates such an approximation for a hypothetical case.

Recall that for any function f of three variables x, y, and z, the first-order Taylor approximation, $\hat{f}$, in the neighborhood of the point (x_0, y_0, z_0) is given by $\hat{f}(x, y, z) = f(x_0, y_0, z_0) + g'_x(x - x_0) + g'_y(y - y_0) + g'_z(z - z_0)$, which reduces to $\hat{f}(x, y, z) = c + g'_x x + g'_y y + g'_z z$, where $g_x \in \partial_x f(x_0, y_0, z_0)$, ($g_y \in \partial_y f(x_0, y_0, z_0)$, resp. $g_z \in \partial_z f(x_0, y_0, z_0)$) is a partial subgradient vector of f at x (resp. z), and

$$c = f(x_0, y_0, z_0) - g'_x x_0 - g'_y y_0 - g'_z z_0. \tag{23}$$

For any vector a, g_a is a column vector, and g'_a is its transpose. We will henceforth drop the transposition operator for simplicity. For a given set of "trial points" $\Theta_{t+1} := \left\{\left(s^i_{t+1}, p^i_t, w^i_t\right) | 1 \leq i \leq I\right\}$, the equation of each hyperplane i is given by

$$H^i_{t+1}(s_{t+1}, p_t, w_t) = \tilde{c}^i_{t+1} + \tilde{g}_{s^i_{t+1}} s_{t+1} + \tilde{g}_{p^i_t} p_t + \tilde{g}_{w^i_t} w_t, \ 1 \leq i \leq I;$$

$\tilde{c}$ and $\tilde{g}$ are expected values with respect to the random variable $\tilde{W}_{t+1}$. It is readily verified that

$$\hat{V}_{t+1}(s_{t+1}, p_t, w_t) = \min\left\{\rho_{t+1} | \rho_{t+1} \geq \tilde{c}^i_{t+1} + \tilde{g}_{s^i_{t+1}} s_{t+1} + \tilde{g}_{p^i_t} p_t + \tilde{g}_{w^i_t} w_t\right\}.$$

An approximation to function $\tilde{F}_t(s_t, p_{t-1}, w_{t-1})$ then reads:

$$\hat{F}_t(s_t, p_{t-1}, w_{t-1}) := \min_{X_t, \rho_{t+1}, \beta_t} \left\{\sum_{g=1}^{|G|} \sum_{j \in \Lambda_g} \beta_{gj,t} CP_g(\hat{p}_{gj}) + g_t(p_t, \Delta_t, \kappa_t) + \rho_{t+1}\right\} \tag{24}$$

$$\text{S.t. } (18) - (21) \tag{25}$$

$$\rho_{t+1} - \tilde{g}_{s^i_{t+1}} s_{t+1} - \tilde{g}_{p^i_t} p_t \geq \tilde{c}^i_{t+1} + \tilde{g}_{w^i_t} w_t, \ 1 \leq i \leq I \tag{26}$$

One can easily figure out that this approximation scheme is akin to general ideas from Benders' decomposition and *cutting plane algorithms* (e.g. see Kelley 1960).

5.2.1 Computation of the hyperplanes parameters

Constraints (26) are constructed via *cut parameters* ($\tilde{g}_{s_{t+1}}$, $\tilde{g}_{p_t}$, $\tilde{g}_{w_t}$, and $\tilde{c}_{t+1}$) computed in period $t+1$ as follows. Firstly, suppose that in period $t+1$, the wind process is discretized into a finite set of realizations $\Omega_{t+1} := \left\{w^1_{t+1}, \ldots, w^J_{t+1}\right\}$. Therefore, in period $t+1$, for each state value (s_{t+1}, p_t, w_t), the problem is solved for each $w_{t+1} \in \Omega_{t+1}$. For each such value of the random process, let $\pi^j_{s,t+1}$ be the vector of dual multipliers associated with the storage dynamics constraints (see Eq. (5)). An expected partial subgradient $\tilde{g}_{s_{t+1}}$ may then be computed as $\tilde{g}_{s_{t+1}} = \alpha \sum_{j=1}^J \omega_j \pi^j_{s,t+1}$, where ω_j is the probability of the observation w^j_{t+1}, and $\sum_{j=1}^J \omega_j = 1$. Secondly, in period $t+1$, let π^j_{gl} and π^j_{gu} be the dual prices associated with constraints (15), and (16) (for each generator $g \in G$), respectively. Then, $\tilde{g}_{p_t} = \sum_{j=1}^J \omega_j (\pi^j_{gl} + \pi^j_{gu})$ is taken. Third, the computation of the partial subgradients $\tilde{g}_{w_t}$ is not as straightforward as that of the partial subgradients $\tilde{g}_{s_{t+1}}$ and $\tilde{g}_{p_t}$. Indeed, the generation of wind turbines located in the same region are correlated because of similar environmental conditions (wind speed). Consequently, suppose the wind generation is modeled as a lag-p multivariate autoregressive process:

$$w_{t+1} = \mu + \sum_{i=0}^{p-1} \Phi_j (w_{t-i} - \mu) + \tilde{\epsilon}_{t+1}, \tag{27}$$

$\Phi_j, 0 \le j \le p-1$, is an $M \times M$ matrix of coefficients; μ is the mean vector of the process, and $\tilde{\epsilon}_{t+1}$ is a vector of innovations. In period $t+1$, for each $w^j_{t+1} \in \Omega_{t+1}$, and each bus where a wind farm m is located, the power balance equation is defined as

$$\sum_{k=1}^{|G_m|} p_{k,t+1} + \sum_{l \in I_m} e_{l,t+1} - \Delta_{m,t+1} - \kappa^+_{m,t+1} + \kappa^-_{m,t+1} - \sum_{l \in O_m} e_{l,t+1} = \hat{D}_{m,t+1} - w^j_{m,t+1}. \tag{28}$$

Let $\pi^j_{wd,t+1}$ be the vector of dual multipliers associated with these constraints. Also recall that in period $t+1$, function $V_{t+2}(s_{t+2}, p_{t+1}, w_{t+1})$ is approximated through a set of cuts:

$$\rho_{t+2} - \tilde{g}^k_{s_{t+2}} s_{t+2} - \tilde{g}^k_{p_{t+1}} p_{t+1} \ge \tilde{c}^k_{t+2} + \tilde{g}^k_{w_{t+1}} w^j_{t+1}, 1 \le k \le K, \tag{29}$$

(see (24–26)). Let $\pi^{jk}_{wc,t+1}$ be the dual multiplier associated with inequality k. Combining the dual prices associated with Eq. (28) and inequalities (29), respectively, and Eq. (27), by the chain rule, a partial subgradient $g^j_{w_t}$ is computed as

$$g_{w^j_t} = \Phi_1 \left(-\pi^j_{wd,t+1} + \sum_{k=1}^{K} \pi^{jk}_{wc,t+1} \tilde{g}^k_{w_{t+1}} \right).$$

Thus, an expected partial subgradient $\tilde{g}_{w_t}$ is given by $\tilde{g}_{w_t} = \sum_{j=1}^{J} \omega_j g_{w^j_t}$. Lastly, from Eq. (23), we see that

$$\tilde{c}_{t+1} = \sum_{j=1}^{J} \omega_j \hat{F}_{t+1}(s_{t+1}, p_t, w_t) - \tilde{g}_{s_{t+2}} s_{t+2} - \tilde{g}_{p_{t+1}} p_{t+1} - \tilde{g}_{w_{t+1}} w_{t+1},$$

where $(s_{t+2}, p_{t+1}, w_{t+1})$ is the state observed in $t+2$ when (s_{t+1}, p_t, w_t) was observed in $t+1$.

5.3 Algorithmic implementation

It is clear from the above discussions that in each period t, a sample of points is required to approximate the expected cost-to-go. The algorithmic framework of the stochastic dual dynamic programming (SDDP) (see Pereira and Pinto 1985; Pereira 1989; Pereira and Pinto 1991) is used to alternate between backward and forward steps toward solving the approximate problem (24–26).

Initialization: $\Theta_{T+1} = \emptyset$, $\tilde{g}_{s_{T+1}} = 0, \tilde{g}_{p_T} = 0, \tilde{g}_{w_T} = 0$, and $\tilde{c}_{T+1} = 0$.
For $t = T, T-1, \ldots, 1$
  Inputs: sampled states $(s_t^i, p_{t-1}^i, w_{t-1}^i), 1 \leq i \leq I$.
  For $i = 1, \ldots, I$
    Sample J wind generation vectors w_t^{ij}, $1 \leq j \leq J$.
    For $j = 1, \ldots, J$
      Solve the minimization problem (24-26).
      Store the multipliers $\pi_{s,t+1}^j, \pi_{gl}^j, \pi_{gu}^j, \pi_{wd,t+1}^j$ and $\pi_{wc,t+1}^{jk}$, and of $\hat{F}_t^{j*}$.
    End
    Compute the expected subgradients $\tilde{g}_{s_t}, \tilde{g}_{p_{t-1}}, \tilde{g}_{w_{t-1}}$, and $\tilde{c}_t$.
    Construct one supporting hyperplane $H_t^i(\cdot)$ for the problem in period $t-1$.
  End
End
If a termination criterion is attained, stop, otherwise perform a forward step.

Fig. 2 Backward induction procedure

The purpose of the backward recursion phase is to construct the approximations of the cost-to-go function, in each iteration through supporting hyperplanes of the epigraph of the true function. In each period (for each iteration), suppose a set of states, $\Theta_t := \left\{(s_t^i, p_{t-1}^i, w_{t-1}^i) \,|1 \leq i \leq I\right\}$, has been sampled. For each state vector, assume J vectors of wind generation w_t^{ij}, $1 \leq j \leq J$, are sampled with probability ω_i^j each. Then for each wind output value, the minimization problem (24-26) is solved backward in time. From the J minimization problems the appropriate multipliers are retrieved to compute the expected value of the parameters $\tilde{g}_{s_t^i}, \tilde{g}_{p_{t-1}^i}, \tilde{g}_{w_t^i}$, and $\tilde{c}_t^i$. These parameters are used to construct one supporting hyperplane $H_t^i(s_t, p_{t-1}, w_{t-1})$ for the minimization problem in period $t-1$. Thus, at the end of the recursion, I supporting hyperplanes are constructed to pass to period $t-1$. As initial conditions, we set $\Theta_{T+1} = \emptyset$, $\tilde{g}_{s_{T+1}} = 0$, $\tilde{g}_{p_T} = 0$, $\tilde{g}_{w_T} = 0$, and $\tilde{c}_{T+1} = 0$. Thus the terminal value is assumed to be zero. A typical backward iteration is summarized in Fig. 2.

The objective of carrying out a forward step is to sample state points $(s_t^i, p_{t-1}^i, w_{t-1}^i), i \in I$, around which the expected cost-to-go will be approximated in each period. For a given initial state (s_1, p_0, w_0), K series of wind generation vectors $w_t^k, 1 \leq t \leq T, 1 \leq k \leq K$, with probability ω_k each, are simulated. For each such trajectory, problem (24–26) is solved, forward in time, using the last approximation to the expected cost-to-go function constructed in the backward step as well as the simulated wind generations. Thus, K series of admissible trajectories for the storage as well as the generation levels are obtained, which will be used as sampled states in the next backward recursion step. Figure 3 summarizes the forward procedure.

Possible termination criteria may include (i) a fixed number of iterations, (ii) a threshold on the computation time, (iii) the first period cost-to-go function is steady, and (iv) statistical comparison of an upper and a lower bound (see Pereira 1989). For more details, see Shapiro (2011). In this paper, depending on the cases, the number of iterations performed varied between 4 and 11, based on the convergence of the first period cost-to-go function.

Inputs: The value function from the previous backward recursion.
Select an initial state (s_1, p_0, w_0).
Simulate K series of wind generations each of length T.
For $l = 1, \ldots, K$
    For $t = 1, \ldots, T$
        Solve the minimization problem (24-26).
        Store the optimal decisions $s_{t+1}^{k^*}$ and $p_t^{k^*}$ and the wind value w_t^k.
    End
End
Go to the backward step

Fig. 3 Forward simulation procedure

In closing this section, it is worth pointing out that the main shortcoming of SDP is the discretization of the state space; in SDDP the discretization is replaced with sampling; consequently, this is the main advantage of SDDP over SDP. On the other hand, this accelerates the backward phase; parameters are easily computed using the optimal value and corresponding dual multipliers supplied by all common LP softwares along with an optimal solution.

6 Numerical experiments

Two lines of enquiry are pursued. The performance of the approximation scheme is demonstrated over a small network and comparisons with SDP are provided in Sect. 6.1. Results of additional experiments on larger networks are reported in Sect. 6.2.

6.1 Illustration and comparison with classical SDP

IEEE 9-bus configuration (three conventional generators) is used to illustrate and compare SDDP with classical SDP. NYISO (New York Independent System Operator) scaled average hourly load data for January 2016 (see references) are used. One wind farm and one storage facility are located at bus 5. Wind data were obtained from the website of the NREL (National Renewable Energy Laboratory) (see references). A 15% wind integration scenario and a 30% wind capacity factor were used in this case; indeed, a wind turbine may not continuously operate at full capacity due to, for instance, the availability of the wind, its location, regulation requirements, market drivers, and reliability maintenance.

Both algorithms were implemented in Python 2.7.10, and the GLP problems were solved with Cplex 12.5.1.0 on a Toshiba Portégé Z930-10 Ultrabook (64-bit, Core i7-3687U CPU @2.10GHz, 8.00 Go RAM). At each iteration of the approximate dynamic backward pass, 25 vectors of wind outputs were simulated to compute the expected values of the cost-to-go functions as well as the expected values of the partial subgradients $(\tilde{g}_{s_t}, \tilde{g}_{p_{t-1}}, \tilde{g}_{w_{t-1}})$, and the expected values of the intercepts $(\tilde{c}_t^i)$. A set of 25 series of state values was used in each iteration of the forward pass. As a result, in each iteration of the backward pass, 25 new cuts were added to the GLP problems. Note that there is no theoretical rule to select the number of cuts in each iteration; the

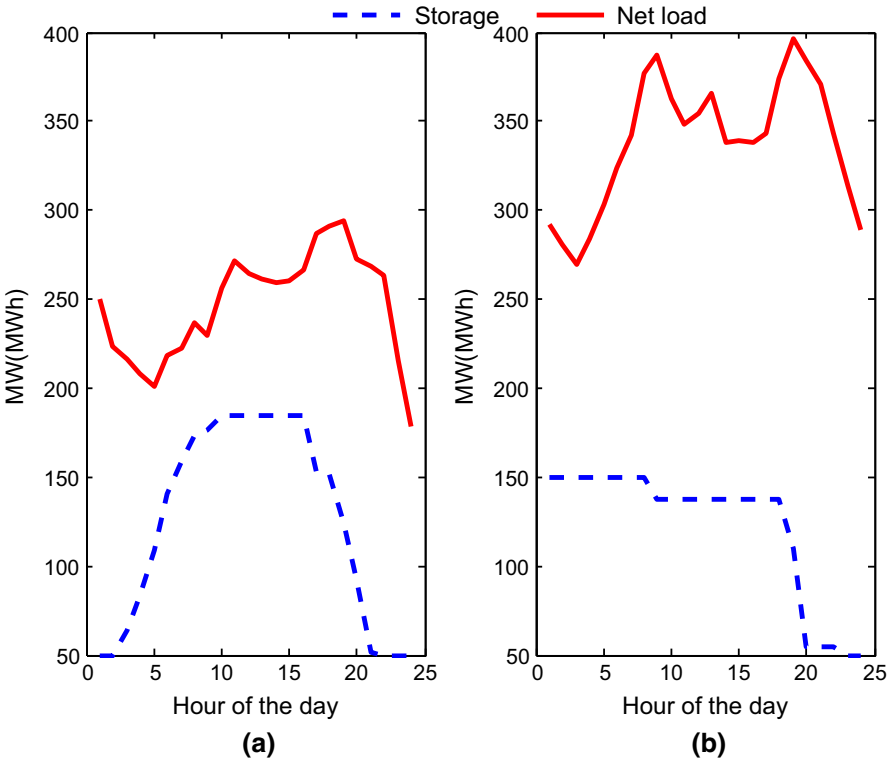

Fig. 4 Example of storage trajectory in which the charging (discharging) cost is low. In case **a** the battery is at the minimum required level in the beginning of the first hour; in case **b** the battery is at 60% of the maximum level of charge in the beginning. In both cases, the optimal strategy consists in using all the available energy over the planning horizon based upon the level of charge in the beginning of the first hour, the net load and the peak hour

number used here was adopted based on empirical trials. In each of the 24 h, this small problem comprises 221 decision variables, and (45+the number of cuts) constraints; thus, the smallest subproblem contains 70 constraints. Results for much larger size problems (thousands constraints) are presented in the subsequent subsection. Note that in SDP, the main challenge is not related to the size of the problem, but to the state space dimension. In this paper, the state space dimension is related to the number of conventional generators, wind farm an storage devices.

Figure 4 depicts the trajectories of the storage (obtained with SDDP) facility over a 24 h span for two cases. In case (a) the stored energy is at the minimum allowable level in the beginning of the first hour; in case (b) the level of charge is at 60% in the beginning of the first hour. It was also assumed that varying the level of energy was cost-free in this example. The net load (demand - wind power) is also depicted in each case.

In the first case, the battery is charged progressively in the beginning as the net load is low, then is steady as of the first peak hour, and up to the second peak hour. The battery is then discharged progressively to its minimum level, since the model

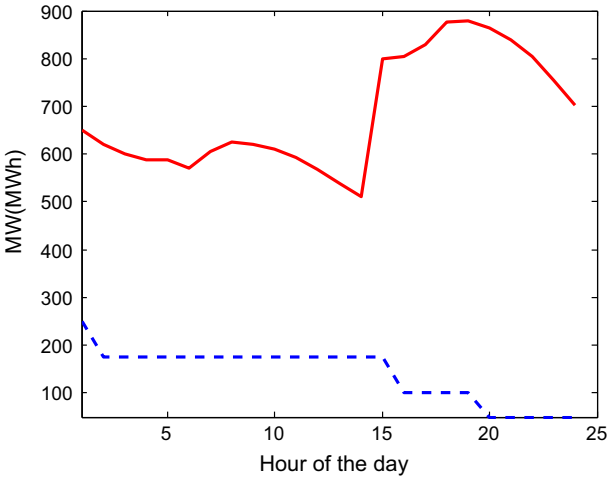

Fig. 5 Example of storage trajectory in which the charging (discharging) cost is high. The storage is used only when the load cannot be met with the conventional generators

considered no "terminal value" for the storage. In the second case, the level of charge is steady until the first peak hour, when the battery is depleted until the end of the horizon. This is exactly what a good algorithm should do. Observe that in case (a) in contrast with case (b), not only the battery is at its minimum allowable level in the beginning, but also the net load is lower than in case (b). It is then clear that case (a) is more favorable to charging than is case (b). In addition in both cases, not only is the stored energy completely used, discharging occurs almost entirely at the peak hour.

Figure 5 illustrates a situation in which varying the level of charge of the battery is very expensive as compared to the cost of operating the conventional units. Since utilizing the battery is very costly, the stored energy is used only in hours where the conventional generators cannot meet the demand.

Of course, comparison of the approximate scheme (SDDP) and SDP performance is an important metric. In general, this is difficult due to the prohibitiveness of the computational burden of SDP, but the performance of both algorithms are compared on the 9-bus network example based on two criteria, namely (i) the CPU time, and (ii) the solution performance (total cost). First the cost functions were approximated with both approaches. For SDP, in each hour, the storage, and the generators outputs were discretized into six levels each, and the wind farm output into seven levels. This was found to be the finest solvable grid for the problems in reasonable time. Since we assumed the wind process to be Markovian, in each hour, it is necessary to consider all the transitions from the previous wind values (i.e. 7×7 transitions). As a result, in each hour, the SDP problem was solved over a grid of $6 \times 6^3 \times 7^2$ levels (the network comprising three conventional generators), which resulted in 63,504 evaluations of the cost-to-go function in each hour. Therefore, 1,524,096 evaluations were performed over the 24-h horizon.

Table 1 Comparison of the CPU time in seconds: SDDP and SDP

Method	Run 1	Run 2
SDDP	499.35	1 876.22
SDP	13,038.16	12,726.88

Table 2 Comparison of solution cost: SDDP and SDP

Method	Min	Max	Mean	Standard deviation
Run 1				
SDDP	78,622.29	162,215.38	126,872.83	21,812.53
SDP	75,392.76	161,872.07	125,968.26	22,922.10
Run 2				
SDDP	88,691.21	164,333.09	134,267.42	19,210.77
SDP	85,882.05	164,333.09	133,981.13	19,657.52

Following the construction of the cost-to-go functions with both methods, the operation of the network was simulated for two runs, each comprising one hundred 24-h wind scenarios. With SDP, in both runs, the cost-to-go functions were constructed over the same state space. With SDDP, in the first run, we performed only four iterations to approximate the value functions. Since in each hour the expected cost-to-go was approximated through 25 supporting hyperplanes (state values) and 20 wind values were used to compute the expectations, with the approximate scheme, the total number of function evaluations then were 500, for a total of 12,000 function evaluations over the 24 h horizon. Therefore, 480,000 function evaluations were performed overall in the first simulation. Those numbers do not include the SDDP forward step, since this is very fast. In the second simulation, we tried to improve the quality of the approximation by carrying out ten iterations. Consequently, 120,000 function evaluations were performed in that run.

Table 1 presents the CPU time for each method and each run. In the first case, the computation time for the approximate dynamic scheme (ADS) was nearly 4% of that of SDP. In the second case, that proportion was about 15% as better approximations were sought. Table 2 reports descriptive statistics on the performance (total cost) of each method for each run. In the first case, on average, using SDP, the total cost was only improved by 0.71% as compared to ADS. In the second case, on average, the total cost difference decreased to 0.21%. This suggests that, overall, the ADS allowed for a fair trade-off between solution time and accuracy.

6.2 Additional numerical tests

The ultimate goal of this research was to analyze the scalability of the approximate dynamic approach to larger networks. We tested the algorithm on different IEEE networks, with characteristics presented in Table 3. Each network was tested with different number of storage facilities and wind farms due to the importance of these

Table 3 Test networks' characteristics

# Of buses	# Of conv. generators	# Of trans. lines	Total gen. cap. (MW)
30	6	41	335.000
57	7	80	1975.880
89	12	210	9921.230
118	54	186	9966.200
300	69	411	32,678.435

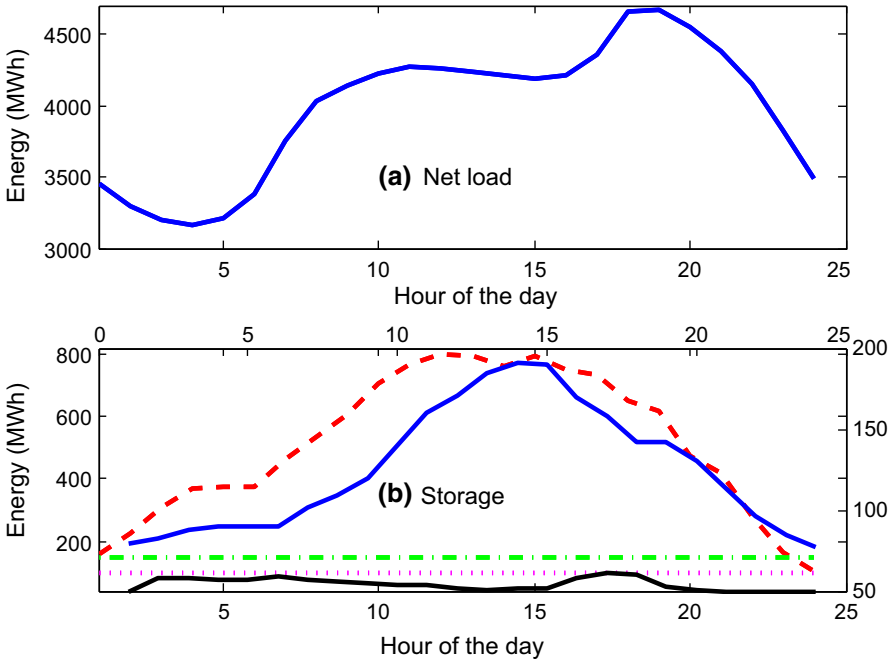

Fig. 6 Mean storage trajectory for the 118-bus network with five storage units and one wind farm. As all the five batteries are not needed to meet the load, only those with the most charging/discharging efficiency are used. The optimal strategy is the same for all utilized batteries: charging when the net load is low and discharging as of the peak hour until the end of the horizon since no terminal value is considered

parameters in the computational effort. Wind and load data were obtained from the same sources as in Sect. 6.1. As previously, the problems are solved over a 24-h horizon.

Figure 6 shows an example of the mean trajectory over 100 simulations for a network composed of 118 buses, five storage facilities and one aggregated wind farm, as well as the net load. Each battery has specific storage, charging, and discharging efficiency, respectively. The three batteries that either have the highest storage or charging/discharging efficiency are utilized to contribute to meet the load, whereas

Table 4 Computation time in seconds for different number of buses, storage facilities and wind farms

# Buses	|S|	|M|	CPU time	# Buses	|S|	|M|	CPU time
30	1	1	1229.80	118	1	1	2399.35
30	5	1	1582.67	118	5	1	2444.99
30	5	5	1323.88	118	5	5	2453.89
57	1	1	1388.09	118	10	5	2179.62
57	5	1	1454.47	118	20	10	2248.39
57	5	5	1396.26	300	1	1	4159.16
57	10	5	1597.71	300	5	1	4234.72
89	1	1	1570.67	300	5	5	4570.01
89	5	1	1709.68	300	10	5	4617.65
89	5	5	1575.09	300	20	10	5036.37
89	10	5	1737.32				

the energy level of the other two batteries (with either the lowest storage or charging/discharging efficiency) is steady over the 24-h span.

For each case, Table 4 reports the computation time for each network, and different number of storage units (|S|) and wind farms (|M|). The number of storage facilities is varied from one to twenty, and the number of wind farms from one to ten. The computation time varied between: (i) 1230 and 1583 s in the case of the 30-bus network, (ii) 1388 and 1598 s for the 57-bus network, (iii) 1570 and 1737 s for the 57-bus network, (iv) 2180 and 2454 s in the case of the 118-bus network, and, (v) 4159.16 and 5036.37 s for the largest network (300 buses). These results clearly suggest that the underlying computational burden of SDP is mitigated as the dimension of the state space increases.

7 Conclusions

This paper analyzed the operation of power networks comprising both conventional and wind generators in the presence of storage. The problem was formulated as a stochastic dynamic programming program and approximated via techniques from generalized linear programming, cutting plane methods, and stochastic dual dynamic programming. The approximation methods were tested on different combinations of network sizes, as well as number of storage facilities and wind generators. The combinations of these techniques allowed handling large dimension state space and enabled to solve the problems in reasonable time. Comparison with stochastic dynamic programming on a small network showed that the approximations are promising both in terms of computation burden, and in accuracy of the solution.

Acknowledgements This work was supported in part by the National Science Foundation under grant ECCS-1453615. The authors acknowledge constructive comments from two anonymous referees that helped improve the paper.

8 Appendix

Nomenclature

Sets

N	Buses
L	Transmission lines
M	Wind farms
G	Conventional generators
G_n	Conventional generators at bus n
S	Storage devices
O_n	Transmission lines that leave node n
I_n	Transmission lines that enter node n
$\mathbf{R}$	Real numbers
$\mathbf{R}^+$	Non negative numbers

Parameters

T	Length of the planning horizon
t	Time period
$\hat{D}_{nt}$	Forecast of the load, in period t at bus n
$\underline{p}_g$	(resp. $\bar{p}_g$) Lower (resp. upper) bound on generator g output
$\underline{\lambda}_g$	(resp. $\bar{\lambda}_g$) Ramp down (resp. ramp up) limit on generator g
$\tilde{W}_t$	Random vector of outputs from the wind turbines in period t
w_{mt}	A particular realization of the stochastic process $\{\tilde{W}_{mt}\}$
$\bar{e}_l$	Bound on power trough line l
B_l	Susceptance of line l
$\underline{s}_n$	(resp. $\bar{s}_n$) Lower (resp. upper) bound on the level of the storage at bus n
c_n	(resp. d_n) Efficiency coefficient of charging (resp. discharging) of the storage device at bus n
$\delta_n =$	$\begin{cases} 1 \text{ if there is a storage facility at node n,} \\ 0 \text{ otherwise.} \end{cases}$

Decision variables

p_{gt}	Power generation from generator g in period t
e_{lt}	Power flowing through transmission line l in period t
θ_{nt}	Phase angle of bus n in period t
s_{nt}	Level of stored energy at bus n in the beginning of period t
Δ_{nt}^+	(resp. Δ_{nt}^-) Positive (resp. negative) variation in the level of charge from the beginning through the end of period t
κ_{nt}^+	(resp. κ_{nt}^-) Power excess absorbed (resp. delivered) by the storage unit located at bus n
Δ_{nt}	Power absorbed or delivered by the storage unit at bus n

Functions

$CP_{gt}(p_{gt})$ Cost function of generator g in period t
$CS_{nt}(\Delta_{nt})$ Cost associated with varying the stored energy at bus n

Operators

$\mathbb{E}$ Mathematical expectation
$|X|$ Cardinality of the set X

References

Abbey C, Joós G (2009) A stochastic optimization approach to rating of energy storage systems in wind-diesel isolated grids. IEEE Trans Power Syst 24:418–426
Anderson RN, Boulanger A, Powell WB, Scott W (2011) Adaptive stochastic control for the smart grid. Proc IEEE 99:1098–1115
Bazaraa MS, Sherali HD, Shetty CM (2013) Nonlinear programming: theory and algorithms. Wiley, Hoboken
Bellman R (1956) Dynamic programming and lagrange multipliers. In: Proceedings of the National Academy of Sciences of the United States of America 42, 767
Bellman RE (1957) Some new techniques in the dynamic-programming solution of variational problems. Quart Appl Math 16:295–305
Bellman RE, Dreyfus SE (1962) Applied dynamic programming. Princeton university press, Princeton
Bertsekas DP, Tsitsiklis JN (1995) Neuro-dynamic programming: an overview. In: Proceedings of the 34th IEEE conference on decision and control. IEEE, vol. 1, pp. 560–564
Bradbury K, Pratson L, Patiño-Echeverri D (2014) Economic viability of energy storage systems based on price arbitrage potential in real-time us electricity markets. Appl Energy 114:512–519
Castillo A, Gayme DF (2014) Grid-scale energy storage applications in renewable energy integration: a survey. Energy Convers Manag 87:885–894
dos Santos Coelho L, Mariani VC (2006) Combining of chaotic differential evolution and quadratic programming for economic dispatch optimization with valve-point effect. IEEE Trans Power Syst 21:989
European Commission Impact Assessment (2016). Sustainability of bioenergy, SWD 418. https://ec.europa.eu/energy/sites/ener/files/documents/1_en_impact_assessment_part4_v4_418.pdf
Foufoula-Georgiou E, Kitanidis PK (1988) Gradient dynamic programming for stochastic optimal control of multidimensional water resources systems. Water Resour Res 24:1345–1359
Garcia-Gonzalez J, la Muela D, Ruiz RM, Santos LM, González AM (2008) Stochastic joint optimization of wind generation and pumped-storage units in an electricity market. IEEE Trans Power Syst 23:460–468
Goor Q, Kelman R, Tilmant A (2010) Optimal multipurpose-multireservoir operation model with variable productivity of hydropower plants. J Water Resour Plan Manag 137:258–267
Grillo S, Marinelli M, Massucco S, Silvestro F (2012) Optimal management strategy of a battery-based storage system to improve renewable energy integration in distribution networks. IEEE Trans Smart Grid 3:950–958
Heussen K, Koch S, Ulbig A, Andersson G (2012) Unified system-level modeling of intermittent renewable energy sources and energy storage for power system operation. IEEE Syst J 6:140–151
Homem-de Mello T, De Matos VL, Finardi EC (2011) Sampling strategies and stopping criteria for stochastic dual dynamic programming: a case study in long-term hydrothermal scheduling. Energy Syst 2:1–31
Howitt R, Msangi S, Reynaud A, Knapp K (2002) Using polynomial approximations to solve stochastic dynamic programming problems: or a 'betty crocker' approach to sdp. University of California, Davis
Johnson SA, Stedinger JR, Shoemaker CA, Li Y, Tejada-Guibert JA (1993) Numerical solution of continuous-state dynamic programs using linear and spline interpolation. Oper Res 41:484–500
Johri R, Filipi ZS (2011) Self-learning neural controller for hybrid power management using neuro-dynamic programming. Presented at the SAE proceedings, Warrendale, PA, USA, Paper 2011-24-0081
Keane M, Wolpin K (1994) The solution and estimation of discrete choice dynamic programming models by simulation and interpolation: Monte Carlo evidence. Rev Econ Stat 76:648–672

Kelley JE Jr (1960) The cutting-plane method for solving convex programs. J Soc Ind Appl Math 8:703–712
Khani H, Zadeh MRD (2015) Real-time optimal dispatch and economic viability of cryogenic energy storage exploiting arbitrage opportunities in an electricity market. IEEE Trans Smart Grid 6:391–401
Kitanidis P (1987) A first-order approximation to stochastic optimal control of reservoirs. Stoch Hydrol Hydraul 1:169–184
Korsak A, Larson R (1970) A dynamic programming successive approximations technique with convergence proofs. Automatica 6:253–260
Lee J-H, Labadie JW (2007) Stochastic optimization of multireservoir systems via reinforcement learning. Water Resour Res 43:W11408. https://doi.org/10.1029/2006WR005627
Lindenberg S (2009) 20% Wind energy by 2030: increasing wind energy's contribution to US electricity supply. Diane Publishing, Collingdale
Löhndorf N, Wozabal D, Minner S (2013) Optimizing trading decisions for hydro storage systems using approximate dual dynamic programming. Oper Res 61:810–823
Luh PB, Yu Y, Zhang B, Litvinov E, Zheng T, Zhao F, Zhao J, Wang C (2014) Grid integration of intermittent wind generation: a markovian approach. IEEE Trans Smart Grid 5:732–741
Maceira MEP, Duarte V, Penna D, Moraes L, Melo A 2008 Ten years of application of stochastic dual dynamic programming in official and agent studies in Brazil-description of the newave program. In: 16th PSCC, Glasgow, Scotland 14–18
Mahlia T, Saktisahdan T, Jannifar A, Hasan M, Matseelar H (2014) A review of available methods and development on energy storage; technology update. Renew Sustain Energy Rev 33:532–545
Meibom P, Barth R, Hasche B, Brand H, Weber C, O'Malley M (2011) Stochastic optimization model to study the operational impacts of high wind penetrations in ireland. IEEE Trans Power Syst 26:1367–1379
Mokrian P, Stephen M 2006 A stochastic programming framework for the valuation of electricity storage. In: 26th USAEE/IAEE North American Conference. 24–27
Momoh J, Zhang Y et al. (2005) Unit commitment using adaptive dynamic programming. In: Proceedings of the 13th international conference on intelligent systems application to power systems
Morales-España G, Latorre JM, Ramos A (2013) Tight and compact milp formulation for the thermal unit commitment problem. IEEE Trans Power Syst 28:4897–4908
Moura PS, De Almeida AT (2010) The role of demand-side management in the grid integration of wind power. Appl Energy 87:2581–2588
Naghibi-Sistani M, Akbarzadeh-Tootoonchi M, Bayaz MJ-D, Rajabi-Mashhadi H (2006) Application of q-learning with temperature variation for bidding strategies in market based power systems. Energy Convers Manage 47:1529–1538
NREL (National Renewable Energy Laboratory). http://www.nrel.gov/electricity/transmission/eastern_wind_dataset.html
NYISO (New York Independent System Operator). Market and operations. http://www.nyiso.com/public/markets_operations/market_data/load_data/index.jsp
Ostrowski J, Anjos MF, Vannelli A (2012) Tight mixed integer linear programming formulations for the unit commitment problem. IEEE Trans Power Syst 1:39–46
Papaefthymiou G, Klockl B (2008) MCMC for wind power simulation. IEEE Trans Energy Convers 23:234–240
Papavasiliou A, Oren SS (2013) Multiarea stochastic unit commitment for high wind penetration in a transmission constrained network. Oper Res 61:578–592
Park J-B, Lee K-S, Shin J-R, Lee KY (2005) A particle swarm optimization for economic dispatch with nonsmooth cost functions. IEEE Trans Power Syst 20:34–42
Pereira M (1989) Optimal stochastic operations scheduling of large hydroelectric systems. Int J Electr Power Energy Syst 11:161–169
Pereira M, Pinto L (1985) Stochastic optimization of a multireservoir hydroelectric system: a decomposition approach. Water Resour Res 21:779–792
Pereira MV, Pinto LM (1991) Multi-stage stochastic optimization applied to energy planning. Math Progr 52:359–375
Pereira-Neto A, Unsihuay C, Saavedra O (2005) Efficient evolutionary strategy optimisation procedure to solve the nonconvex economic dispatch problem with generator constraints. IEE Proc Gen Transm Distrib 152:653–660
Philbrick CRJ, Kitanidis PK (2001) Improved dynamic programming methods for optimal control of lumped-parameter stochastic systems. Oper Res 49:398–412

Philpott AB, de Matos VL (2012) Dynamic sampling algorithms for multi-stage stochastic programs with risk aversion. Eur J Oper Res 218:470–483
Philpott AB, Guan Z (2008) On the convergence of stochastic dual dynamic programming and related methods. Oper Res Lett 36:450–455
Pinson P et al (2013) Wind energy: forecasting challenges for its operational management. Stat Sci 28:564–585
Qiu Q, Pedram M (1999) Dynamic power management based on continuous-time markov decision processes. In: Proceedings of the 36th annual ACM/IEEE design automation conference. ACM, 555–561
Renewable energy policy network for the 21st century. 2017. Renewables 2017, Global status report. technical report
Sayah S, Zehar K (2008) Modified differential evolution algorithm for optimal power flow with non-smooth cost functions. Energy Convers Manage 49:3036–3042
Schoenung S (2011) Energy storage systems cost update
Scott W, Powell WB (2012) Approximate dynamic programming for energy storage with new results on instrumental variables and projected bellman errors. Submitted to Operations Research (Under Review)
Shapiro A (2011) Analysis of stochastic dual dynamic programming method. Eur J Oper Res 209:63–72
Shapiro A, Tekaya W, da Costa JP, Soares MP (2013) Risk neutral and risk averse stochastic dual dynamic programming method. Eur J Oper Res 224:375–391
Shapiro JF (1979) Mathematical programming: structures and algorithms. Wiley, New York
Song H, Liu C-C, Lawarrée J, Dahlgren RW (2000) Optimal electricity supply bidding by markov decision process. IEEE Trans Power Syst 15:618–624
Suberu MY, Mustafa MW, Bashir N (2014) Energy storage systems for renewable energy power sector integration and mitigation of intermittency. Renew Sustain Energy Rev 35:499–514
Succar S, Denkenberger DC, Williams RH (2012) Optimization of specific rating for wind turbine arrays coupled to compressed air energy storage. Appl Energy 96:222–234
Tan X, Li Q, Wang H (2013) Advances and trends of energy storage technology in microgrid. Int J Electr Power Energy Syst 44:179–191
Tan Y, Liu W, Qiu Q (2009) Adaptive power management using reinforcement learning. In: Proceedings of the 2009 international conference on computer-aided design. ACM, 461–467
Thatte AA, Xie L, Viassolo DE, Singh S (2013) Risk measure based robust bidding strategy for arbitrage using a wind farm and energy storage. IEEE Trans Smart Grid 4:2191–2199
Topaloglu H, Powell WB (2006) Dynamic-programming approximations for stochastic time-staged integer multicommodity-flow problems. INFORMS J Comput 18:31–42
Turgeon A (1980) Optimal operation of multireservoir power systems with stochastic inflows. Water Resour Res 16:275–283
Wee J-H (2013) A review on carbon dioxide capture and storage technology using coal fly ash. Appl Energy 106:143–151
Yang J-S, Chen N (1989) Short term hydrothermal coordination using multi-pass dynamic programming. IEEE Trans Power Syst 4:1050–1056
Yu Y, Luh PB, Litvinov E, Zheng T, Zhao J, Zhao F (2015) Grid integration of distributed wind generation: hybrid markovian and interval unit commitment. IEEE Trans Smart Grid 6:3061–3072
Zéphyr L, Lang P, Lamond BF (2015) Controlled approximation of the value function in stochastic dynamic programming for multi-reservoir systems. CMS 12:539–557
Zéphyr L, Lang P, Lamond BF, Côté P (2017) Approximate stochastic dynamic programming for hydroelectric production planning. Eur J Oper Res 262:586–601
Zhou Y, Scheller-Wolf AA, Secomandi N, Smith S (2014) Managing wind-based electricity generation in the presence of storage and transmission capacity. SSRN 1962414
Zurn H, Quintana V (1975) Generator maintenance scheduling via successive approximations dynamic programming. IEEE Trans Power Appar Syst 94:665–671