On the complexity of energy storage problems

ARTICLE IN PRESS
Discrete Optimization ▮ (▮▮▮▮) ▮▮▮–▮▮▮

Contents lists available at ScienceDirect
Discrete Optimization
www.elsevier.com/locate/disopt

On the complexity of energy storage problems

Nir Halman[a], Giacomo Nannicini[b,*], James Orlin[c]

[a] Hebrew University of Jerusalem, Israel
[b] IBM T. J. Watson Research Center, Yorktown Heights, NY, United States
[c] Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, United States

ARTICLE INFO

Article history:
Received 31 October 2016
Received in revised form 9 August 2017
Accepted 30 November 2017
Available online xxxx

Keywords:
Approximation algorithms
Energy storage
Dynamic programming
K-approximation sets

ABSTRACT

We analyze the computational complexity of the problem of optimally managing a storage device connected to a source of renewable energy, the power grid, and a household (or some other form of energy demand) in the presence of uncertainty. We provide a mathematical formulation for the problem as a Markov decision process following other models appearing in the literature, and study the complexity of determining a policy to achieve the maximum profit that can be attained over a finite time horizon, or simply the value of such profit. We show that if the problem is deterministic, i.e. there is no uncertainty on prices, energy production, or demand, the problem can be solved in strongly polynomial time. This is also the case in the stochastic setting if energy can be sold and bought for the same price on the spot market. If the sale and buying price are allowed to be different, the stochastic version of the problem is #P-hard, even if we are only interested in determining whether there exists a policy that achieves positive profit. Furthermore, no constant-factor approximation algorithm is possible in general unless P = NP. However, we provide a Fully Polynomial-Time Approximation Scheme (FPTAS) for the variant of the problem in which energy can only be bought from the grid, which is #P-hard.

1. Introduction

A problem that has been receiving increasing attention in recent years (see e.g. [1–4]) is that of optimally managing sources of renewable energy connected to the power grid, batteries, and potentially a household or some other form of energy sink. The presence of a battery is extremely valuable for peak shaving, time shifting, reduction of electricity price arbitrage, and to provide operating reserve, see [4–6]. The *energy storage* problem is that of deciding when to store, release, buy, and sell energy in this context. In this paper we focus on a basic subproblem of the complex decision problem introduced above. We call the basic subproblem *single-node* because it corresponds to the decision problem faced by a single energy-producing node in a smart grid, and does not explicitly take into account the goals of the system operator. The problem can

* Corresponding author.
E-mail addresses: halman@huji.ac.il (N. Halman), nannicini@us.ibm.com (G. Nannicini), jorlin@mit.edu (J. Orlin).

https://doi.org/10.1016/j.disopt.2017.11.001
1572-5286/

naturally be modeled as a stochastic dynamic program. Existing works in the literature focus on the study of practically viable solution methods, typically through approximate dynamic programming approaches, and on the analysis of optimal or heuristic policies. To the best of our knowledge, no analysis of the complexity of this problem has been carried out. This work aims to fill this gap, identifying which characteristics of the problem make it difficult to solve from a theoretical point of view, i.e. no polynomial-time algorithm exists unless P = NP.

In this paper we give a general model for the single-node energy storage problem as a Markov decision process, and study its complexity. Our formulation is consistent with models in the recent literature. We show that the deterministic version of the problem can be cast as a minimum-cost flow problem and solved in strongly polynomial time, but the problem becomes #P-hard as soon as uncertainty is introduced, even for a restricted version of the problem in which energy can only be bought from the grid, i.e. selling on the spot market is not allowed. This hardness result holds for both determining the optimal policy and just determining the optimal policy value. In the general case, we show that even deciding whether the problem admits any policy achieving positive profit is #P-hard. The freedom of distributing energy from the renewable source or the battery to different devices during the same time period may reduce the difficulty of the problem, in the sense that if this possibility is removed or if it comes at a cost, the problem is weakly NP-hard already in the deterministic setting. Restricting the sale price on the spot market to be equal to the buying price makes the problem easier, because we provide a polynomial-time dynamic programming scheme for the stochastic case. This restriction on the price is well-studied in the literature (e.g. [4]), and it implies that the market is arbitrage-free. Our analysis suggests that the difficulty of numerically solving this restricted version of the problem originates from the complexity of handling the random processes, which can be very complicated in practice, rather than from the structure of the decision problem itself. The situation is considerably different than in the general case: the unrestricted version of the problem is intrinsically hard, i.e. it is #P-hard even when the random processes are very simple (independent random variables with support of size two). Finally, we show that no deterministic constant-factor approximation algorithm is possible in general, but we provide a Fully Polynomial-Time Approximation Scheme (FPTAS) for the case where selling to the grid is not allowed, indicating that solving the problem in a context with no sales may be easier than in the general case. An FPTAS is a deterministic ϵ-approximation algorithm that runs in polynomial time in the binary input size and $1/\epsilon$ for any $\epsilon > 0$, and it is considered the strongest possible result in terms of approximation.

Our model considers a single storage device, and models with multiple storage devices such as in [7] are at least as hard. This paper assumes that the states and transition probabilities of the Markov decision process are fully known: from a theoretical point of view this is a necessary assumption for exact computations (if we do not have access to such data, the problem is intractable in general), but it limits the practical relevance of the proposed algorithms. We show that if the probabilities are given implicitly by means of an oracle, instead of explicitly, even a version of the problem that is usually polynomial-time solvable may require an exponential number of oracle calls in the worst case, indicating that attempts at finding efficient algorithms are likely to fail unless specific structure is imposed on the stochastic processes.

From a methodological point of view, our FPTAS is a problem-specific extension of the framework of K-approximation sets and functions introduced in [8] to continuous (as opposed to discrete) state and action spaces, exploiting the piecewise linear structure of the objective function. This extension can easily be adapted to other continuous stochastic dynamic programs that share the same structure, i.e. piecewise linear convex costs (for a minimization problem) and affine transition function, yielding an FPTAS for those problems as well. The input data of the problem can be rational.

This paper is organized as follows. Section 2 is dedicated to a literature review. Section 3 introduces the mathematical model for the energy storage problem. Section 4 discusses our notation and a structural property of optimal policies for the problem, giving simplified models for two restrictions of the problem

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

that are well studied in the literature. Section 5 shows that if the sale price is equal to the buying price, the problem can be solved in polynomial time, and discusses the implications. Section 6 shows that the deterministic problem can be solved in polynomial time even without the above restriction on the sale and buying price, but if restrictions are introduced in distributing energy among the devices, or in the stochastic case, the problem becomes difficult and no constant-factor approximation algorithm is possible in general. Section 7 constructs an FPTAS for the restriction of the problem where selling energy to the grid is not allowed. Section 8 contains concluding remarks.

2. Literature review

The energy storage problem is related to inventory theory, because energy can be seen as a stock that has to be optimally managed. The typical source of uncertainty in the inventory theory literature is demand, but sometimes supply is also considered nondeterministic. For comprehensive references in inventory theory, we refer to [9,10]. The problem also bears similarities with the commodity trading literature; from this point of view, particularly relevant to our work are [11,12], which consider an arbitrage setting with bounds on the trades and a limited warehouse capacity, as is the case in our problem. Note that the single-node decision problem studied here is substantially different from the problem faced by system operators that employ or plan to include storage devices (see e.g. [13]), and we only review previous work that relates to the single-node decision problem.

Existing literature on the single-node version is concerned with the formulation, analysis and/or numerical solution of mathematical models for the problem, typically based on dynamic programming (DP). Zhou et al. [14] study under which conditions it is more profitable to dispose of the surplus of electricity produced rather than store it for future usage in the presence of negative electricity prices, which can be found in several energy markets due to costs and constraints involved in switching on and off power generators. Zhou et al. [4] analyze the management of wind energy in the presence of transmission and storage capacities, which is very similar to the setting of the present paper. Zhou et al. [4] describe the problem as a discrete-time finite-horizon Markov decision process and computes the structure of an optimal policy. Furthermore, they study numerically the relative impact of the different sources of profit that are due to the storage device (reduction of curtailment, time-shifting generation, arbitrage). Natarajan et al. [2] propose and analyze an approximate dynamic programming (ADP) algorithm and several heuristics for the single-node energy storage problem as discussed in this paper, providing numerical results. Harsha and Dahleh [15] formulate an infinite horizon stochastic dynamic program to minimize the average cost derived from installing and managing a storage device to satisfy uncertain demand, in the presence of uncertainty on energy prices. They show that the problem admits a dual threshold (also known as limit) optimal policy under mild assumptions, i.e. a policy with two thresholds (ℓ, u) such that it is optimal to buy (sell) up to ℓ (down to u) if inventory is below ℓ (above u), and do nothing otherwise. Optimality of dual threshold policies was already discussed in a trading context by [11] in a restricted case, and [12] in a general case. The papers [2,4,15] discuss the value of installing a storage device under different assumptions, showing that it can often be a profitable investment. Moazeni et al. [16] consider the effect of risk on the performance of a policy for the energy storage problem, showing that because of fat tails and spikes in prices, the profit attained by a deterministic risk-neutral policy or a myopic policy may notably differ from the expected profit with considerable probability.

Other papers focus mostly on solution methodologies. Nascimento [1] discusses an ADP approach for a class of storage problems with convexity properties, one of these problems being the one discussed in this paper. A method closely connected to the one proposed in [1] is shown to be convergent to the optimal value function under certain technical conditions [17]. Scott and Powell [3] study ADP algorithms based on approximate policy iteration for the single-node energy storage problems using realistic data. Salas and Powell [7] propose a finite-horizon ADP algorithm to design near-optimal policies for energy storage problems with multiple-storage devices, i.e. multiple batteries with different characteristics. The method of [7] does not necessarily converge to an optimal policy, but is shown to be effective on realistic data. Jiang et al. [18] compare several ADP approaches (policy iteration, value iteration, direct policy search) on benchmark instances of the energy storage problem. An approach not based on DP is taken by [19], where a solution to the problem that does not require statistical knowledge of the system dynamics is sought. Under some assumptions, the system can be kept stable and the amount by which the computed policy is suboptimal is bounded.

It should be noted that in the vast majority of the works cited above, the price for selling energy on the spot market is the same as for buying. However, this is not always the case in practice: for example, in the Singaporean electricity market independent energy providers may sell on the spot market at a different price than the regulated retail price determined by the Energy Market Authority.

The construction of the approximation scheme described in this paper uses the framework introduced in [8], which gives sufficient conditions for the existence of an FPTAS for discrete stochastic dynamic programs with special structure. This work provides an extension to continuous state and action spaces, exploiting piecewise linearity of the objective function. A more general extension to continuous state and action spaces of the framework is discussed in [20]: Halman and Nannicini [20] introduce the concept of a (Σ, Π)-FPTAS, that is, an approximation scheme that provides an approximation to the optimal solution up to both an absolute error Σ and a multiplicative error Π for any $\Sigma > 0$ and $\Pi > 1$, and they show that continuous dynamic programs with a certain structure do not necessarily admit a "regular" (purely multiplicative) FPTAS, but admit a (Σ, Π)-FPTAS. In the present paper, we describe an FPTAS for the problem in the traditional sense, that is, an approximation scheme that provides a solution with relative error up to ϵ for any $\epsilon > 0$. Notice that a traditional FPTAS can be seen as a (Σ, Π)-FPTAS with $\Sigma = 0$, and it therefore provides stronger approximation guarantees than a (Σ, Π)-FPTAS. However, to achieve this result we require a piecewise linear convex objective function (for a minimization problem), whereas the framework of [20] applies to any Lipschitz-continuous convex function.

3. Problem description and notation

We now provide a mathematical formulation for the single-node energy storage problem. The high-level description of the problem is the following: we are serving a household with uncertain demand; energy can be provisioned via the renewable energy source, the battery, or the power grid. Buying energy on the spot market incurs a cost while selling yields a revenue, depending on prices. The entirety of the demand must be satisfied in every time period. The objective is to maximize profit over a finite time horizon.

The problem has the following parameters:

- T: number of time periods ($T + 1$ is the *terminal stage*).
- $R^{\max}$: maximum capacity of the storage device in MWh.
- η^c, η^d: charging and discharging efficiencies of the device (≤ 1).
- γ^c, γ^d: maximum charging and discharging rates of the device in MWh per time period.
- c^h: holding cost of the storage device, in $ per MWh per time step.

The exogenous information is represented by the following random variables:

- E_t: amount of energy produced by the renewable source at time t.
- D_t: energy demand of the household (or some other type of energy sink) at time t.
- C_t: buying price of electricity at time t.
- P_t: sale price of electricity at time t.

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

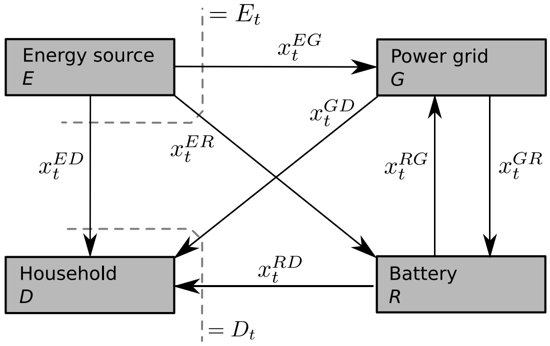

Fig. 1. Main components of the Single-Node Energy Storage problem. At time t, the total energy produced by the energy source is E_t, and the household has a demand of D_t.

Collectively, we denote the exogenous information by a vector $W_t = (E_t, D_t, C_t, P_t)$. The set of possible realizations of W_t is denoted $\mathcal{W}_t$. We assume that the stochastic process $\{W_t : t = 1, \ldots, T\}$ is a finite discrete-time Markov process, and its support and the transition probabilities are given explicitly in binary encoding (relaxing this assumption makes the problem much more difficult, see Theorem 5.2 and the corresponding discussion). In this paper we adopt the convention that W_t is the exogenous information that realizes at time t, i.e. it is $\mathcal{F}_t$-measurable, following [21].

The internal state of the system is given by:

- R_t: amount of energy stored in the storage device at time t.

The set of possible battery levels is $[0, R^{\max}]$ and the *state space* is given by $\mathcal{S}_t := [0, R^{\max}] \times \mathcal{W}_t$ for all t. Here we follow the convention that the state should contain all the necessary information to take decisions regardless of past history, and for this reason the current realization of W_t is included.

As is customary, we assume that $P_t \leq C_t$ for all t (no arbitrage). We further assume that there is no limit to the amount of energy that can be bought from or sold to power grid, and that $C_t \geq 0$. Negative prices can occur in practice in rare situations, but the experimental analysis of Zhou et al. [4] reports that one can still achieve 99% of the optimal value ignoring negative prices for a comprehensive range of system configurations. Transmission capacity constraints that effectively limit the trades are usually not considered in the literature [an exception is 4], but we remark that all the results discussed in this paper except Theorem 5.1 can easily be extended to the case where there are transmission capacities U^b, U^s for the amount that can be bought from or sold to the grid, and inefficiencies in the transmission to and from the grid (a constant-factor loss $\eta^t \leq 1$).

At every time step, we must decide how to manage the system, and in particular, how much energy to transfer to and from each element of the system. At the end of the time horizon, all energy leftover in the battery is discarded (incorporating a positive salvage value p^s for leftover energy at the end of the time horizon is possible by adding an extra time period with deterministic sale price p^s). We identify the grid by the letter G, the battery by R, the household demand by D, and the energy source by E; we are then interested in determining a nonnegative vector $x_t = (x_t^{GR}, x_t^{GD}, x_t^{EG}, x_t^{ER}, x_t^{ED}, x_t^{RG}, x_t^{RD})$ where x_t^{ij} indicates the amount of energy transferred from device i to device j at time t. A sketch of the main components of the system, and the possible energy transfers, is given in Fig. 1.

We can model the problem as a stochastic dynamic program as follows. The internal state transition is given by:

$$R_{t+1} = R_t - x_t^{RD} - x_t^{RG} + \eta^c (x_t^{GR} + x_t^{ER}).$$

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

The action vector must satisfy the following constraints:

$$x_t^{ER} + x_t^{GR} \leq \min\{\gamma^c, R^{\max} - R_t\} \qquad x_t^{RD} + x_t^{RG} \leq \min\{\gamma^d, R_t\}$$
$$x_t^{EG} + x_t^{ER} + x_t^{ED} \leq E_t \qquad x_t^{ED} + \eta^d x_t^{RD} + x_t^{GD} = D_t,$$

and as noted above, $x_t \geq 0$. The set of vectors that satisfy these constraints given the state (R_t, W_t) is the *action space* at stage t, which we denote by $\mathcal{A}_t(R_t, W_t)$. The single-period profit is given by:

$$g_t(R_t, x_t, W_t) := P_t(\eta^d x_t^{RG} + x_t^{EG}) - C_t(x_t^{GR} + x_t^{GD}) - c^h R_{t+1}.$$

The goal of the problem is to determine a *policy* that maximizes profit, i.e. functions $\pi_t : \mathcal{S}_t \to \mathcal{A}_t(R_t, W_t)$ that achieve

$$z^*(R_1^*|W_0) := \max_{\pi_1,\ldots,\pi_T} \mathbb{E}\left[\sum_{t=1}^{T} g_t(R_t, \pi_t(R_t, W_t), W_t)\right],$$

where R_1^* is the initial battery state, W_0 the initial state of the Markov process, the state of the system evolves according to the dynamics previously described, and the expectation is taken with respect to the joint probability distribution of the random variables W_t. We call this the *Single-Node Energy Storage* (SNES) problem. Notice that because this is a stochastic optimization problem, a solution requires the computation of functions that (implicitly or explicitly) define the optimal sequence of actions depending on the state of the problem, which in turn depends on the realization of the random variables. An alternative definition of the problem is to determine the optimal profit $z^*(R_1^*|W_0)$, but not necessarily the optimal policy; in the remainder of the paper we will prove that for our cases of interest, computing the optimal profit is at least as hard as computing the optimal policy, and all polynomial-time algorithms discussed in this paper yield both the (optimal or approximate) policy and the (optimal or approximate) profit.

Our formulation is consistent with other models existing in the literature, e.g. [2]. Salas and Powell [3] and Scott and Powell [7] study a variant where energy cannot be transferred directly from the wind/solar farm to the grid ($x_t^{EG} = 0$), and the system operator is collecting a revenue of $P_t D_t$ in every time period as a reward for satisfying demand. We note that because of this revenue, the value function is nonnegative for all stages of the DP, and this could make the problem easier from an approximation standpoint; in particular, it is easy to show by algebraic manipulations that in this case the single-period profit no longer depends on D_t.

4. Preliminaries

We assume that all data are rational numbers. The support of each random variable is the set of values it can take with positive probability. We denote by $E_t^{\min}, E_t^{\max}$ the minimum and maximum element in the support of E_t, and similarly for D_t, C_t. We denote $q := \max_{t=1,\ldots,T}\{|\mathcal{W}_t|\}$, i.e. an upper bound on the number of states of the Markov process $\{W_t : t = 1, \ldots, T\}$ in each time period. We denote by $(y)^+ := \max\{0, y\}$ and by $(y)^- := \max\{0, -y\}$. We denote intervals on the real line with the usual notation $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$, and intervals of integers by $[a, \ldots, b] := \{x \in \mathbb{Z} : a \leq x \leq b\}$, which we call *contiguous* to emphasize that there are no holes. When referring to a function, for clarification purposes we sometimes use the $(\cdot)$ notation to emphasize which variable is under consideration, e.g. $f(\cdot, x)$ denotes a function of two variables in which the first argument is varying, and the second is fixed at x.

We now state a basic property of optimal policies for (SNES), and describe two restrictions of the problem well-discussed in the literature because they originate from energy market regulations frequently encountered in the real world. These restrictions will be discussed throughout the remainder of the paper.

Lemma 4.1. *There exists an optimal policy for the dynamic program satisfying* $x_t^{ED} = \min\{E_t, D_t\}$ *for any time period* $t = 1, \ldots, T$.

Intuitively, the proof follows from the fact that energy is produced at no cost, and therefore it is always optimal to use as much energy from the renewable source as possible, before buying from the grid. A formal proof is provided in Appendix. We remark that Lemma 4.1 and its proof remain valid even if $P_t = 0$ for all 0, which is equivalent to saying that energy can be discarded at no cost, but no revenue is generated by selling to the grid. If energy cannot be discarded free of charge, Lemma 4.1 no longer holds true. For example, consider an instance with $T = 2$, $D_1 = 1, D_2 = 0$, $E_1 = E_2 = 1$, $P_1 = P_2 = 0$, the starting level for the battery is 1, and we cannot discard energy. The optimal policy for this instance is to satisfy demand from the battery, yielding zero profit, whereas setting $x_1^{ED} = 1$ at the first stage incurs a cost of c^h. In the following, we assume that energy can always be discarded at no cost.

The first restriction of the problem that we discuss is that in which energy cannot be sold to the grid, i.e. the energy system operator is not allowed to sell on the spot market, for example because the energy market is not liberalized. This corresponds to an instance of (SNES) in which $P_t = 0$ for all $t = 1, \dots, T$. We call this restriction (SNES-NS) (NS stands for “no sales”). Notice that in this case we formulate the problem as a cost minimization problem, since there is no revenue. All instances of (SNES-NS) will be intended to be in cost minimization form. It is not difficult to notice that in this case the problem can be simplified because the only decision to be taken in each time period is the net energy balance of the storage device: once the net balance is determined, an optimal energy flow between grid, storage, production and demand can be computed in closed form. This is stated in the next proposition. A full proof of the validity of such reformulation is given in Appendix. We remark that a similar reformulation is described in [2] without a formal proof.

Proposition 4.2. *An instance of (SNES-NS) can be modeled as a dynamic program with scalar action x_t^B as follows:*

$$R_{t+1} = R_t + x_t^B,$$

single-period cost function:

$$g_t(R_t, x_t^B, W_t) := \begin{cases} C_t(\frac{1}{\eta^c} x_t^B + D_t - E_t)^+ + c^h(R_t + x_t^B) & \text{if } x_t^B \geq 0 \\ C_t(D_t - E_t + \eta^d x_t^B)^+ + c^h(R_t + x_t^B) & \text{otherwise,} \end{cases}$$

constraints:

$$-\min\{\gamma^d, R_t\} \leq x_t^B \leq \min\{\gamma^c, R^{\max} - R_t\},$$

and objective function:

$$z^*(R_1^*) := \min_{\pi_1^B, \dots, \pi_T^B} \mathbb{E}\left[\sum_{t=1}^{T} g_t(R_t, \pi_t^B(R_t, W_t), W_t)\right],$$

where $\pi_t^B : \mathcal{S}_t \to \mathcal{A}_t(R_t, W_t)$ describes the policy at stage t.

The second restriction of (SNES) that we consider is that in which the sale price is the same as the buying price, i.e. $P_t \equiv C_t$ for all $t = 1, \dots, T$. We use $\equiv$ to denote the fact that the two coincide, and are not simply two random variables with the same distribution. We denote this restriction by (SNES-EP) (EP stands for “equal price”). Since the sale price is equal to the buying price, we can always assume that any leftover energy from the renewable source after satisfying the demand is sold to the grid, and the problem is greatly simplified. This is stated in the next proposition, whose full proof is given in Appendix. A reformulation along the lines of Proposition 4.3 is described (without a formal proof) and used in existing work such as [4].

Proposition 4.3. *An instance of (SNES-EP) can be modeled as a dynamic program with scalar action x_t^B as follows:*

$$R_{t+1} = R_t + x_t^B,$$

single-period cost function:

$$g_t(R_t, x_t^B, W_t) := P_t(E_t - D_t - \frac{1}{\eta^c}(x_t^B)^+ + \eta^d(x_t^B)^-) - c^h(R_t + x_t^B)$$

constraints:

$$-\min\{\gamma^d, R_t\} \le x_t^B \le \min\{\gamma^c, R^{\max} - R_t\},$$

and objective function:

$$z^*(R_1^*) := \max_{\pi_1^B,\dots,\pi_T^B} \mathbb{E}\left[\sum_{t=1}^{T} g_t(R_t, \pi_t^B(R_t, W_t), W_t)\right],$$

where $\pi_t^B : \mathcal{S}_t \to \mathcal{A}_t(R_t, W_t)$ describes the policy at stage t.

5. Complexity of (SNES-EP)

In this section we discuss the complexity of (SNES-EP), the restricted version of the problem in which the sale price is equal to the buying price. Our main result is that stochastic (SNES-EP) can be solved in polynomial time by dynamic programming. The seminal work in this area is [22], which provides an analytic solution to a dynamic programming formulation for a linear planning problem under uncertainty. Our model is a variation of that of Charnes et al. [22] and of Secomandi [12] where the additional sources of difficulty are the inefficiency of the storage device, and the capacity constraints (considered in the latter paper but not in the former). While it is not too difficult to adapt the dynamic programming approach of Secomandi [12] to (SNES-EP), we remark that our goal here is not simply to provide a solution procedure, but to show an algorithm that runs in polynomial time for any choice of rational input data, rather than pseudopolynomial time as in [12]. Recall that $q := \max_{t=1,\dots,T}\{|\mathcal{W}_t|\}$.

Theorem 5.1. *(SNES-EP) can be solved in $O(q^2T^3)$ time by dynamic programming.*

Proof of Theorem 5.1. First, notice that by expanding the value function z^*, the constant factor $\mathbb{E}[\sum_{t=1}^{T} P_t(E_t - D_t)]$ can be taken out, leaving the single-period cost function as:

$$g_t(R_t, x_t^B, W_t) := P_t(-\frac{1}{\eta^c}(x_t^B)^+ + \eta^d(x_t^B)^-) - c^h(R_t + x_t^B).$$

This implies that the optimal policy is independent of E_t and D_t. Intuitively, this follows from the proof of Proposition 4.3, because one can always assume that all the renewable energy is sold to the grid and all the demand is satisfied from the grid for an expected profit of $\mathbb{E}[\sum_{t=1}^{T} P_t(E_t - D_t)]$. The remaining part of the problem concerns how to optimally manage the battery, given the price fluctuations. Hence, the policy depends on R_t and C_t, which is in turn fully characterized given W_{t-1}.

We claim that the problem can be solved as a dynamic program and the value function at each stage is piecewise linear concave with a polynomially bounded number of breakpoints, which implies the desired result. We now show by backward induction that the number of breakpoints of the value function $z_t(\cdot|W_{t-1})$

is a subset of $\Delta_t = \{a\gamma^d - b\gamma^c : 0 \leq a\gamma^d - b\gamma^c \leq R^{\max}, a+b \leq T+1-t, a,b \in \mathbb{N}\}$, for every realization of W_{t-1}. This implies that the same claim holds for $\mathbb{E}_{W_{t-1}} z_t(R_t|W_{t-1})$ because the breakpoints of the expectation are a subset of the union of the breakpoints of $z_t(R_t|W_{t-1})$.

The first induction step is trivial because at stage $T+1$ the value function is identically zero, and at stage T it is optimal to sell all the remaining energy, so there is at most one breakpoint at γ^d. Assume now that the value function $z_{t+1}(\cdot|W_t)$ satisfies the induction hypothesis, and consider z_t. We have:

$$z_t(R_t|W_{t-1}) = \mathbb{E}\left[\max_{-\min\{\gamma^d, R_t\} \leq x_t^B \leq \min\{\gamma^c, R^{\max} - R_t\}} g_t(R_t, x_t^B, W_t) + z_{t+1}(R_t + x_t^B|W_t) \middle| W_{t-1}\right]. \tag{1}$$

It is well known that a problem with piecewise linear concave single-stage cost function and affine transition function has a piecewise linear concave value function at every stage, and it admits an optimal dual threshold policy at every stage for every realization of W_t; see e.g. [12], [23, Prop. 10.3]. As defined in Section 2, a dual threshold policy for given W_t consists of two levels $\ell(W_t), u(W_t)$ such that it is optimal to buy up to $\ell(W_t)$ if the current storage level is below $\ell(W_t)$, sell down to $u(W_t)$ if the current storage level is above $u(W_t)$, and do nothing otherwise. Optimality of such policies for piecewise linear concave value functions (for a maximization problem) has been first proven in [22], to the best of our knowledge. The lower threshold $\ell(W_t)$ can be computed by solving:

$$\max_{0 \leq x_t^B \leq R^{\max}} g_t(0, x_t^B, W_t) + z_{t+1}(x_t^B|W_t).$$

Such a value $\ell(W_t)$ is attained at one of the breakpoints of $z_{t+1}(\cdot|W_t)$, since $g_t(0, x_t^B, W_t)$ is linear in x_t^B over $0 \leq x_t^B \leq R^{\max}$. A similar argument, substituting $R^{\max}$ for R_t in (1), shows that we can easily compute the upper threshold $u(W_t)$ and it is attained at one of the breakpoints of $z_{t+1}(\cdot|W_t)$.

Now for every value of W_t consider the right-hand side of (1) separately for $R_t \leq \ell(W_t)$, $\ell(W_t) < R_t < u(W_t)$, and $R_t \geq u(W_t)$. In the first case, the optimal policy is $\bar{x}_t^B = \min\{\gamma^c, \ell(W_t) - R_t\}$. If $x_t^B = \gamma^c$, the breakpoints of (1) are those of $g_t(R_t, \gamma^c, W_t)$ or those of $z_{t+1}(R_t|W_t)$ shifted by γ^c. Since $g_t(R_t, \gamma^c, W_t)$ is linear in R_t, the breakpoints are those of $z_{t+1}(\cdot|W_t)$ shifted to the left by γ^c. If $\bar{x}_t^B = \ell(W_t) - R_t$, the value function is $g_t(R_t, \ell(W_t) - R_t, W_t) + z_{t+1}(\ell(W_t) - R_t|W_t)$, and the breakpoints are those of $z_{t+1}(\cdot|W_t)$. When $\ell(W_t) < R_t < u(W_t)$, the value function is $g_t(R_t, 0, W_t) + z_{t+1}(R_t|W_t)$, so the breakpoints are those of $z_{t+1}(\cdot|W_t)$. Finally, when $R_t \geq u(W_t)$, we have $\bar{x}_t^B = \max\{-\gamma^d, u(W_t) - R_t\}$, and we can apply a similar argument to the case $R_t \leq \ell(W_t)$: the breakpoints are either those of $z_{t+1}(\cdot|W_t)$ shifted to the right by γ^d, or those of $z_{t+1}(\cdot|W_t)$ unmodified. By the induction hypothesis, the breakpoints of $z_{t+1}(\cdot|W_t)$ belong to the set Δ_{t+1}. Shifting them to the left by γ^c or to the right by γ^d yields a subset of the set Δ_t, by definition of Δ_t. This concludes our inductive claim.

To solve the problem by backward recursion, we need to compute the value function at each stage. Computing the thresholds $(\ell(W_t), u(W_t))$ for every value of W_t takes $O(q \log T)$ time because there are at most q values of W_t, the computation can be performed by binary search over Δ_{t+1}, and $\max_t |\Delta_t|$ is $O(T^2)$. To compute the value of $z_t(R_t|W_{t-1})$ using (1), we must take the expected value and compute a slope for each breakpoint, and the running time is bounded $O(qT^2)$. Since there are at most q values of W_{t-1}, processing one stage of the dynamic program requires $O(q^2T^2)$, and the T stages of the problem increase the total running time to $O(q^2T^3)$. The time to compute the shift of the objective function $\mathbb{E}[\sum_{t=1}^{T} P_t(E_t - D_t)]$ is $O(q^2T)$, which completes the proof. □

Theorem 5.1 may seem to dispute the claimed difficulty of solving (SNES-EP) in practice, but it is not so. Indeed, the polynomial-time algorithm described above relies on the assumption that the Markov decision process has a polynomially bounded number of states, and the transition probabilities are given explicitly. This is rarely the case in practice, and it begs the question of how to solve the problem if the probabilities are not given explicitly. Theorem 5.1 suggests that the difficulty of (SNES-EP) is mainly due to the complexity of handling the stochastic processes, rather than the structure of the problem itself. This interpretation is supported by our next result, in which we analyze a less restrictive representation for the random variables.

A natural model to implicitly represent a stochastic process on a computer is to provide its support, and an oracle that computes the cumulative distribution function or the probability mass function (PMF); for example, if a random variable is distributed according to a (possibly truncated) known distribution with given parameters, it can easily be represented in the way mentioned above. Theorem 5.2 shows that under such a representation, there is no deterministic algorithm that can compute the expected profit of a very simple instance of (SNES-EP) in polynomial time. The proof technique is adapted from [24].

Theorem 5.2. *Given an instance of (SNES-EP) with two time periods, suppose that* $E_1 = D_1 = P_1 = 0$, *and all among* E_2, D_2, P_2 *are deterministic, except for a single random variable that has support* $[1, \dots, U]$ *and its PMF is given by an oracle* $f(x)$ *computable in polynomial time. Then computing the optimal value* $z^*(R_1|W_0)$ *of (SNES-EP) requires* $\Omega(U)$ *evaluations of* $f(x)$ *in the worst case, which is exponential in the binary input size* $\log U$.

Proof of Theorem 5.2. The idea for this proof is to construct a family of $U - 1$ probability mass functions that are all equal to a given distribution except at a single point $k \in [1, \dots, U-1]$, and an instance of (SNES-EP) such that computing the corresponding optimal profit $z^*(R_1|W_0)$ requires exact knowledge of the PMF. Any deterministic algorithm that aims to compute $z^*(R_1|W_0)$ must therefore determine the value of k, and this may require up to $U - 1$ evaluations of $f(x)$.

Define:

$$\alpha := \frac{2}{U(U+1)} - \frac{2}{U^2(U+1)},$$

and consider the following random variable X_k parametrized by $k \in [1, \dots, U-1]$:

$$\Pr(X_k = x) \begin{cases} \frac{1}{U^2} & \text{if } x = k \\ \alpha(2k+1) + \frac{1}{U^2} & \text{if } x = k+1 \\ \alpha x + \frac{1}{U^2} & \text{if } x \in [1 \dots, U] \setminus \{k, k+1\} \\ 0 & \text{otherwise.} \end{cases}$$

Notice that the support of X_k is exactly $[1, \dots, U]$, and we can construct a polynomial-time oracle f for its PMF using the closed-form expression given above.

Construct an instance of (SNES-EP) with two time periods as follows: $P_1 = 0, P_2 = X_k, E_1 = E_2 = D_1 = D_2 = 0$, $\eta^d = \eta^c = \gamma^d = \gamma^d = R^{\max} = 1$, $c^h = 0$ (it is easy to amend this construction to allow nontrivial values for these parameters). Clearly in this case the optimal decision is to fill-up the battery in stage 1, and sell all stored energy (a single unit) in stage 2. The optimal profit is therefore given by $\mathbb{E}[P_2] = \mathbb{E}[X_k]$. Carrying out the calculations yields:

$$\mathbb{E}[X_k] = \frac{\alpha U(U+1)(2U+1)}{6} + \frac{U+1}{2U} + \alpha k.$$

Thus, to determine the optimal profit one has to determine k, and any deterministic algorithm may require up to $U - 1$ calls to f for this task. This concludes the proof for the case in which P_2 is the chosen random variable.

For the second case, construct a similar instance of (SNES-EP) with: $P_1 = E_1 = D_1 = D_2 = 0, P_2 = 1, E_2 = X_k$. It is optimal to buy one unit in the first period and sell all energy in the second time period, yielding a profit of $\mathbb{E}[E_2] + 1 = \mathbb{E}[X_k] + 1$. The rest of the proof proceeds as above.

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

For the last case, construct an instance of (SNES-EP) with: $P_1 = E_1 = E_2 = D_1 = 0, P_2 = 1, D_2 = X_k$, $\eta^d = \eta^c = \gamma^d = \gamma^d = R^{\max} = U$, $c^h = 0$. It is optimal to store U units in the first time period, use it in the second period to satisfy demand, and sell leftover energy. This yields a profit of $\mathbb{E}[U - D_2] = U - \mathbb{E}[X_k]$, and the rest of the proof proceeds as above. □

Due to this result, it is unlikely that any efficient algorithm can compute the optimal profit for (SNES-EP) for any distribution of the random variables; of course, this leaves open the possibility of specialized algorithms that can take advantage of specific structures. We remark that in the proof, the optimal policies are trivial because we specifically construct problem instances with the desired characteristics: the above hardness result applies to the computation of the optimal value only, in general.

We discuss an additional result that may initially sound counterintuitive: the worst-case running time to determine an optimal policy for a variant of (SNES-EP) in which W_t is $\mathcal{F}_{t+1}$-measurable, rather than $\mathcal{F}_t$-measurable, is significantly better. This is because if we are not allowed to see W_t before taking a decision, the optimal policy is simpler and so is the solution process. In fact, in this case the problem is solvable in $O(T^2 \log T + q^2 T)$ time rather than $O(q^2 T^3)$, which could yield a significant improvement since the q^2 factor increases only linearly with T. This corroborates our observation that the structure of (SNES-EP) is not too difficult, but uncertainty makes the problem difficult in practice, and simplifications in the uncertainty model can have a large impact on the effectiveness of solution algorithms. Let (SNES-EP-$(t+1)$) be the variant of (SNES-EP) in which W_t is only revealed at stage $t+1$.

Theorem 5.3. *(SNES-EP-$(t+1)$) can be transformed into a minimum cost flow problem over a network with $O(T)$ nodes and arcs, hence it can be solved in $O(T^2 \log T + q^2 T)$ time.*

Proof of Theorem 5.3. The first part of the proof of Theorem 5.1 shows that we are only concerned with optimally managing the battery, given the price fluctuations.

For every stage t, since the battery state R_t is determined solely by R_{t-1} and x^B_{t-1}, and the latter cannot depend on W_t, nonanticipativity implies that the decisions are the same for all sample paths (i.e. with the same past history, including the battery state, we must take the same decision): $x^B_t = \pi_t(R_t, W_t)$ must be the same for all $W_t \in \mathcal{W}_t$. Therefore, it is straightforward to see that there is a single decision to be taken at each stage, and maximizing the profit is equivalent to finding a minimum cost flow on a network constructed as follows: there is a node v^G and nodes v^R_t for $t = 1, \ldots, T$; there is an arc (v^R_t, v^R_{t+1}) for $t = 1, \ldots, T-1$ with cost c^h and capacity $R^{\max}$; there is an arc (v^G, v^R_t) with cost $\mathbb{E}[P_t]/\eta^c$ and capacity γ^c for all $t = 1, \ldots, T$; there is an arc (v^R_t, v^G) with cost $-\eta^d \mathbb{E}[P_t]$ and capacity γ^d for all $t = 1, \ldots, T$. If the initial battery level is nonzero, we introduce a corresponding supply for node v^R_1. All other supplies/demands are zero. A sketch of the corresponding network is given in Fig. 2. Such a network has $O(T)$ nodes and arcs, and using the algorithm of [25], a minimum cost flow can be determined in $O(T^2 \log T)$ time; there is an improvement of $\log T$ over the usual running time because the shortest path subproblems are each solvable in $O(T)$ time. The time to compute the data for the network, i.e. the expected values $\mathbb{E}[P_t]$, and the shift $\sum_{t=1}^{T} \mathbb{E}[P_t(E_t - D_t)]$, is $O(q^2 T)$. Determining the minimum cost flow and adding the shift $\sum_{t=1}^{T} \mathbb{E}[P_t(E_t - D_t)]$ yields the optimal profit and the optimal policy. □

We conclude this section with a remark about the dependence of the running times on q. Throughout the paper it is assumed, as is customary, that the states of the Markov process $\{W_t : t = 1, \ldots, T\}$ in each time period are the possible realizations of W_t: this may yield a large q and therefore a large input size. A considerable simplification may occur when W_t is Markov-modulated and the number of states of the underlying Markov process is much smaller than the support of W_t. For example, there may be a "world state" (e.g. the weather) with a small number of states q_w and each of E_t, D_t, C_t, P_t takes at most q_s values: then the running times in Theorems 5.1 and 5.3 can be reduced replacing q^2 with the smaller factor $q_w q_s$. A similar consideration applies to the approximation algorithm discussed in Section 7, see also [23, Sec. 10.3].

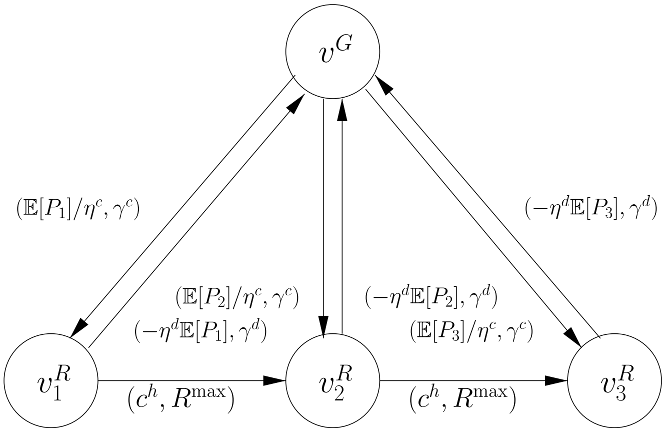

Fig. 2. An example of the network flow problem that can be used to solve an instance of (SNES-EP-$(t+1)$). In this case, we consider only 3 time periods and $R_1^* = 0$. All nodes have zero supply. Arc labels should be interpreted as (cost, capacity).

6. Complexity of (SNES-NS) and (SNES)

In this section we analyze the complexity of (SNES) in the general case. Our main goal is to show that (SNES) is #P-hard and it does not admit any constant-factor approximation unless P = NP.

6.1. Polynomial-time solvability in the deterministic case

The first main result for this section is that in the absence of uncertainty, (SNES) can be solved in strongly polynomial time as a network flow problem. Intuitively, this follows from the fact that the problem is a slightly more sophisticated version of a continuous inventory control problem (see e.g. [10]): the main differences are that our problem has two types of procurement sources, two types of demand, the warehouse has capacities on the amount that can be deposited/withdrawn in any time period, and the warehouse is inefficient (i.e. there can be losses every time storage is used).

Theorem 6.1. *The deterministic version of (SNES) can be transformed into a minimum cost flow problem over a network with $O(T)$ nodes and arcs, hence it can be solved in $O(T^2 \log^2 T)$ time.*

Proof of Theorem 6.1. We reformulate a deterministic (SNES) into an equivalent network flow problem. In this proof, the capacity of an arc is assumed to be infinite and the supply/demand of a node is assumed to be zero if not otherwise specified.

We introduce a node v^G to represent the grid, a node $v_t^R, t = 1, \ldots, T+1$ to represent storage, a node v_t^O to represent output from storage, and a node v_t^I to represent input to storage. For all $t = 1, \ldots, T$, we introduce arcs (v_t^R, v_{t+1}^R) with capacity $R^{\max}$ and cost c^h, arcs (v_t^R, v_t^O) with capacity γ^d and cost 0, arcs (v_t^I, v_t^R) with capacity γ^c and cost 0, arcs (v^G, v_t^I) with cost $\frac{1}{\eta^c}C_t$, and arcs (v_t^O, v^G) with cost $-\eta^d P_t$. We connect the terminal node v_{T+1}^R to v^G with an arc of cost 0. Node v_1^R has a supply of R_1^* (the initial battery level).

Let $B_t = E_t - D_t$. For all $t = 1, \ldots, T$ such that $B_t > 0$, we introduce a node v_t^B with supply $\eta^c B_t$ and arcs (v_t^B, v_t^I) with cost 0 and (v_t^B, v^G) with cost $-P_t/\eta^c$. For all $t = 1, \ldots, T$ such that $B_t < 0$, we introduce a node v_t^B with demand B_t/η^d and arcs (v_t^O, v_t^B) with cost 0 and (v^G, v_t^B) with cost $\eta^d C_t$. Finally, we set the supply/demand of v^G to $\sum_{t=1}^{T} - \left(\eta^c (B_t)^+ - \frac{1}{\eta^d}(B_t)^-\right) - R_1^*$. The resulting min-cost flow problem has

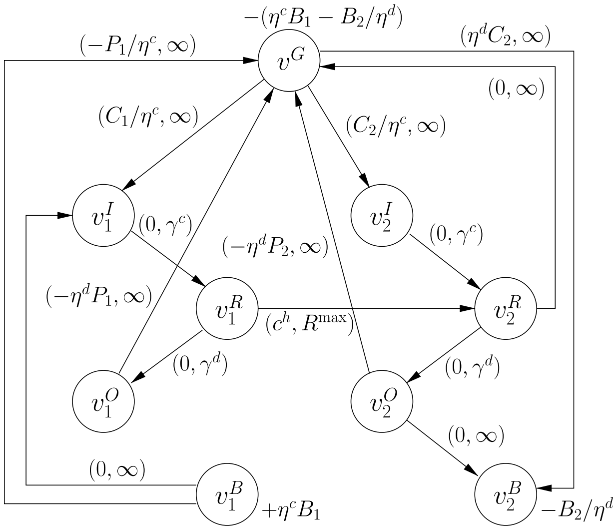

Fig. 3. An example of the network flow problem that can be used to solve a deterministic instance of (SNES). In this case, we consider only 2 time periods and $R_1^* = 0$. Node supplies are indicated with a "+", demands with a "−". Arc labels should be interpreted as (cost, capacity).

rational supply/demand values (as well as rational costs), but we can rescale the data to be integer and this transformation is polynomial in the size of the (SNES) instance. A sketch of the network is given in Fig. 3.

We now show that this transformation solves the original (SNES) instance. We show a one-to-one correspondence between a feasible solution $\bar{x}_t, t = 1, \ldots, T$ to (SNES) and a feasible network flow. By Lemma 4.1, we can assume $\bar{x}_t^{ED} = \min\{E_t, D_t\}$. To avoid cumbersome notation, we indicate the decision variable corresponding to the flow value over an arc by the name of the arc itself. Set $(v_t^B, v_t^I) = \eta^c \bar{x}_t^{ER}, (v_t^B, v^G) = \eta^c \bar{x}_t^{EG}, (v_t^O, v_t^B) = \bar{x}_t^{RD}, (v^G, v_t^B) = \frac{1}{\eta^d} \bar{x}_t^{GD}, (v^G, v_t^I) = \eta^c \bar{x}_t^{GR}, (v_t^O, v^G) = \bar{x}_t^{RG}$. By plugging into the objective function, the total cost of such a flow is:

$$\sum_{t=1}^{T} \left(-P_t(\eta^d \bar{x}_t^{RG} + \bar{x}_t^{EG}) + C_t(\bar{x}_t^{GR} + \bar{x}_t^{GD}) + c^h \bar{R}_{t+1} \right),$$

where

$$\bar{R}_{t+1} = \bar{R}_t - x_t^{RD} - x_t^{RG} + \eta^c (x_t^{GR} + x_t^{ER})$$

and $\bar{R}_1 = R_1^*$. This is exactly the optimal cost of the dynamic program. Furthermore, one can verify that the constraints of the dynamic program are mapped one-to-one to the capacity constraints of the network flow. The network constructed this way has $O(T)$ nodes and arcs, and we can apply the algorithm described in [25] to obtain the claimed running time. This concludes the proof. Note that, in general, the optimal solution to the deterministic (SNES) will be rational. □

6.2. Hardness results

The result in the previous section shows that (SNES) is solvable in polynomial time in the deterministic case. We will now show that this is the only case in which (SNES) is easy: introducing further difficulty makes the problem transition from easy to hard, i.e. not solvable in polynomial time unless P = NP.

We start by considering a variant of the deterministic problem in which energy distribution among the different devices is restricted, and more specifically: in every time period energy can be transferred only between two selected devices, or a fee has to be paid whenever the destination device of an energy transfer changes. These characteristics make the problem NP-hard already in the deterministic setting, and they suggest that the freedom to distribute energy between any device at any time is a valuable property that may make the problem easier to solve in practice. Intuitively, hardness stems from the fact that the problem can no longer be modeled as a simple network flow problem, but becomes similar to a network design problem, which is much more difficult. The proof of this result will use the well-known NP-hard problem Partition.

Problem (Partition): Determining a partition of a set of numbers.
Input: integers $\{a_1, \dots, a_n\}$, and $b = \sum_i a_i/2$.
Output: Does there exist $K \subset \{1, \dots, n\}$ such that $\sum_{i \in K} a_i = \sum_{i \notin K} a_i = b$?

Theorem 6.2. *If one of the following two conditions holds, determining the optimal cost of deterministic (SNES) is weakly NP-complete:*

(i) *In every time period, the system operator is allowed to transfer energy only between a chosen pair of devices and in a single direction;*
(ii) *The battery does not have perfect efficiency ($\eta^c < 1$) and the system operator must pay a cost whenever the destination device of an energy transfer changes, even if multiple such changes can take place in the same time period.*

Proof of Theorem 6.2. This proof uses reductions from Partition with data as described above.

(i) Consider an instance of the deterministic (SNES) with $T = n + 1$ time periods, energy production $E_t = a_t$ for $t = 1, \dots, n$, $E_T = 0$, $D_t = 0$ for $t = 1, \dots, n$, $D_T = b$. The storage efficiency is $\eta^c = \eta^d = 1$, the battery capacity is $2b$, the price for selling to the grid is $P_t = 1$ for $t = 1, \dots, T$, and the cost for buying is $C_t = 2b$. Suppose this instance of (SNES) admits a solution with profit b. This is clearly the largest value that can be obtained, because $2b$ units are produced and b are necessary to satisfy demand. Notice that energy cannot be bought from the grid because buying a single unit negates any profit. Thus, energy only comes from the renewable source, and because the total demand is b and it must be satisfied, there are b units that are produced and sold to achieve a profit b in periods $1, \dots, n$, while b units are stored in periods $1, \dots, n$ and employed to satisfy demand at time T. By assumption, in every time period we can only transfer energy from the renewable source to the storage device or to the grid. This implies that we found a partition of the set $\{a_1, \dots, a_n\}$ into two sets with sum b. Suppose now that there exists a solution to Partition. Then we can construct a solution to (SNES) with profit b in the way indicated above.

(ii) Consider an instance of deterministic (SNES) with $T = 2n+1$ time periods, energy production $E_t = 1$ for $t = 1, 3, \dots, 2n-1$, $E_t = 1 + 2a_{t/2}$ for $t = 2, 4, \dots, 2n$, $E_T = 0$, energy demand $D_t = 1$ for $t = 1, \dots, 2n$, $D_T = b$. The storage efficiency is $\eta^c = 1/2, \eta^d = 1$, the battery capacity is $2b$, the price for selling to the grid is $P_t = 1$ for $t = 1, \dots, T$, and the cost for buying is $C_t = 4b + 4n$. Furthermore, there is a cost of $1/3$ every time the destination of an energy transfer changes (both between time periods or within a single time period). Suppose this instance of (SNES) admits a solution with profit $2b - 2n/3$. Observe that the total demand is $b + 2n$ and the total production is $4b + 2n$, so there is a surplus of $3b$ units. However, no energy can be bought from the grid, because the cost of a single unit would yield negative profit overall. Furthermore, in every time period 1 unit of energy must go to the household, and is easy to show that it is optimal to take this unit from the renewable source directly because the battery has charging efficiency $\eta^c = 1/2$. It follows that redirecting the surplus $2a_{t/2}$ when t is even requires at least $2n$ changes of the energy destination: one to change from the household to the grid or battery, and one to change back to satisfy the household demand in the subsequent time period. A solution with profit $2b - 2n/3$ implies that exactly $2n$ changes are performed and $2b$ units are sold; selling from the battery to the grid would require

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

at least an extra change of destination, hence nothing is sold from the battery and in exactly n time periods the surplus energy $2a_{t/2}$ is sent to the grid. This implies that there exists a Partition. Suppose now that Partition has a solution with index set $K \subset \{1,\dots,n\}$. We can construct a solution to (SNES) with profit $2b-2n/3$ as follows: in every time period t with t odd, send 1 unit from the energy source to the household, and in every time period t with t even, send 1 unit from the energy source to the household, while storing the remaining $2a_{t/2}$ in the battery if $t/2 \in K$, and selling it to the grid if $t/2 \notin K$. This way, exactly $2b$ units are sold for a revenue of b, and we need to pay $2n/3$ in transition costs.

To conclude the proof, we notice that the deterministic version of (SNES) is clearly in NP because we can provide a certificate for a solution of given value by indicating how much energy is transferred to each device in every time period, and the problem can be solved in pseudopolynomial time by dynamic programming over the possible storage level, discretized to the least common denominator of the input data. □

We now analyze the difficulty of stochastic (SNES). We will first show that (SNES-NS) is hard in the stochastic setting, and use this to prove hardness of approximation of (SNES) in the general case. This requires some preliminary results. In the context of this paper, a function h defined over a contiguous interval $S \subseteq \mathbb{Z}$ is said to be *discrete convex* if $h(x+1)+h(x-1) \geq 2h(x)$ for all $x \in S$. As a consequence, local optimality implies global optimality. Note that this is a considerable simplification of the treatment of discrete convexity given in the literature, see [26–28]. However, since we only employ this concept for univariate functions over the integers, the simplification will suffice for our purposes.

Lemma 6.3. *Let D be a discrete integer-valued random variable with domain $[0,u]$. Then, the function $\mathbb{E}[(D-x)^+]$ is discrete convex nonincreasing in x.*

Proof of Lemma 6.3. Clearly $\mathbb{E}[(D-x)^+]$ is a nonincreasing function. For convexity, we show that $\mathbb{E}[(D-x)^+] + \mathbb{E}[(D-(x+2))^+] \geq 2\mathbb{E}[(D-(x+1))^+]$ for all $x \in \mathbb{Z}$. We calculate the expectations via the law of total expectation with the partition $(D \leq x) \uplus (D = x+1) \uplus (D \geq x+2)$ (x is fixed). We note that $E[(D-(x+i))^+ | D \leq x] = 0$, for $i = 0,1,2$, so the inequality holds at equality for the case of $D \leq x$. For the case of $D \geq x+2$ we have $E[(D-(x+i))^+ | D \geq x+2] = E[D-(x+i) | D \geq x+2]$ for $i = 0,1,2$, so due to linearity of expectation the inequality is satisfied again as equality. For the remaining case $D = x+1$, $E[(D-x)^+ | D = x+1] = 1$, $E[(D-(x+2))^+ | D = x+1] = 0$ and $E[(D-(x+1))^+ | D = x+1] = 0$, so the inequality is satisfied as a strict inequality. We conclude that $E[(D-x)^+] + E[(D-(x+2))^+] - 2E[(D-(x+1))^+] = \Pr(D = x+1) \geq 0$, as needed. □

We will use reductions from the following problem.

Problem (CDF): Evaluating the CDF of the convolution of discrete random variables.
Input: Discrete random variables $X_1,\dots,X_n$, support $a_{i,j} \in \mathbb{N}^+$ for $i = 1,\dots,n, j = 1,\dots,m$, probabilities $p_{i,j} = \text{Prob}(X_i = a_{i,j})$, values $\Lambda \in \mathbb{N}^+$ and $\lambda \in Q^+$ with $0 < \lambda \leq 1$.
Output: Is $\text{Prob}(\sum_{i=1}^n X_i \leq \Lambda) \geq \lambda$?

Proposition 6.4 ([8]). *(CDF) is #P-hard even when $m = 2, p_{i,j} = \frac{1}{2}\ \forall i,j$, and $a_{i,2} = 0\ \forall i$.*

The class #P contains solution counting problems associated with decision problems in NP; two famous examples of #P-complete problems are: counting the number of Hamiltonian circuits in graphs, counting the number of solutions to a knapsack problem. Notice that the counting version of a problem can be hard even if the corresponding decision problem is solvable in polynomial time: for example, one can find a perfect matching in a bipartite graph in polynomial time, but counting the number of distinct perfect matchings is

#P-complete [29]. Solving #P-complete problems exactly is considered very difficult, because all problems in the polynomial-time hierarchy can be solved in a polynomial number of calls to a #P-oracle [30].

We are now ready to state our first hardness result for stochastic (SNES-NS).

Theorem 6.5. *It is #P-hard to determine either the optimal policy or the optimal profit (cost) for the stochastic (SNES-NS) problem, even if:*

- *prices C_t are deterministic;*
- *either one among E_t and D_t is deterministic and the other is independently distributed across time periods, with a support of size 2 and equal probabilities.*

Proof of Theorem 6.5. We consider a reduction from (CDF) with n discrete random variables, $m = 2, p_{i,j} = \frac{1}{2}\ \forall i, j$, and $a_{i,2} = 0\ \forall i$. Let $M = \max_{i,j} a_{i,j}$. In this proof, all instances of (SNES) have $\gamma^c = \gamma^d = R^{\max}$, $\eta^c = \eta^d = 1$, $c^h = 0$, $P_t = 0$ for all t. We will start by showing that determining the optimal policy is #P-hard, and then show that determining the optimal profit is at least as hard.

Define an instance of (SNES-NS) with the following data: $T = n + 1$, $D_1 = 0, D_{t+1} = X_t\ \forall t = 1, \ldots, n$, $C_1 = (1 - \lambda), C_t = 1\ \forall t = 2, \ldots, n + 1$, $E_t = 0\ \forall t$, $R^{\max} = Mn$. The random variables are independent since the X_t's are. This is clearly a polynomial reduction of (CDF). The objective of (SNES-NS) is to minimize the expected cost to satisfy the demand, which is in total equal to $D = \sum_{i=1}^{n} X_i$. At the first stage of the dynamic program we have the option of buying energy at unit price $(1 - \lambda)$ and storing it at no cost. In the following stages we have to satisfy demand using stored energy from the battery or buying from the grid at cost 1. Note that as selling to the grid yields no profit and costs do not vary after the first stage, it is always optimal to use stored energy as much as possible, then buy from the grid after the battery is depleted. Hence, there is only one decision to take: the amount x of energy that is bought at stage 1 with unit cost $(1 - \lambda)$, and the remaining amount $(D - x)^+$ must be bought in subsequent stages at unit cost 1. It follows that this problem is equivalent to a newsvendor problem with demand D where the overage cost is $(1 - \lambda)$ and the underage cost is $1 - (1 - \lambda) = \lambda$. It is well known that in this case, the optimal amount x^* to be bought is:

$$x^* = \arg\min_x \left\{ x : \text{Prob}\Big(\sum_{i=1}^{n} X_t \le x\Big) \ge \frac{\lambda}{\lambda + (1 - \lambda)} = \lambda \right\}. \quad (2)$$

Hence, if we could determine the optimal policy for (SNES) with stochastic demand in polynomial time, we would be able to solve (CDF) in polynomial time. This implies that determining the optimal policy for this restriction of (SNES-NS) is #P-hard.

We now show that determining the optimal cost for (SNES-NS) is as at least as hard as determining the optimal policy. We consider (SNES-NS) instances defined as in the reduction above. Remember that we are minimizing the total cost, which can be split into the procurement cost $(1 - \lambda)x$, and the expected cost of satisfying the demand $\mathbb{E}[(D - x)^+]$, where x is the order quantity. Suppose we could determine the optimal cost $z^*(y)$ for an instance of (SNES) defined as above, except that we set $E_1 = y$, y integer. In other words, we are now looking at a newsvendor problem where y units of inventory are provided for free. Let $x^*(y)$ be the largest optimal order quantity at the first stage when $E_1 = y$. Hence, $x^*(0)$ is the largest optimal amount that should be bought in the first time period starting with empty inventory, and by our discussion above, determining $x^*(0)$ is #P-hard. When $y < x^*(0)$, the optimal decision is $x^*(y) = x^*(0) - y$, i.e. we order up to $x^*(0)$. Therefore,

$$\begin{aligned} z^*(y + 1) - z^*(y) &= (1 - \lambda)(x^* - y - 1) + \mathbb{E}[(D - x^*(0))^+] - (1 - \lambda)(x^* - y) - \mathbb{E}[(D - x^*(0))^+] \\ &= -(1 - \lambda). \end{aligned}$$

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.

When $y \geq x^*(0)$, $x^*(y) = 0$, i.e. it is optimal to order nothing at the first stage. Thus,

$$\begin{aligned} z^*(y+1) - z^*(y) &= \mathbb{E}[(D-(y+1))^+] - \mathbb{E}[(D-y)^+] \\ &\geq \mathbb{E}[(D-(x^*(0)+1))^+] - \mathbb{E}[(D-x^*(0))^+] > -(1-\lambda), \end{aligned}$$

where the first inequality follows by convexity of $\mathbb{E}[(D-y)^+]$, see Lemma 6.3, and the second inequality follows because

$$(1-\lambda)(x^*(0)+1) + \mathbb{E}[(D-(x^*(0)+1))^+] > (1-\lambda)x^*(0) + \mathbb{E}[(D-x^*(0))^+]$$

by definition of $x^*(0)$ (recall that $x^*(0)$ is the largest $\arg\min$ of $z^*(0)$). But this shows that if we could compute $z^*(y)$, we would be able to determine whether $y < x^*(0)$ or $y \geq x^*(0)$ by computing $z^*(y+1) - z^*(y)$, and $x^*(0)$ could be determined in polynomial time by binary search on y. It follows that determining $z^*(y)$ must be as difficult as determining $x^*(0)$, i.e. #P-hard.

For the case of stochastic energy production, define an instance of (SNES-NS) with the following data: $T = n+1$, $E_1 = 0, E_{t+1} = (M - X_t)\ \forall t = 1,\dots,n$, $C_1 = (1-\lambda), C_t = Mn\ \forall t = 2,\dots,n$, $C_T = 1$, $D_t = 0\ \forall t = 1,\dots,n$, $D_T = Mn$, $R^{\max} = Mn$. The random variables E_t are independent and this is a polynomial reduction of (CDF). In this instance we must satisfy a demand of Mn at the last stage. The only decision to take is how much energy x should be bought at time $t = 1$ at unit cost $(1-\lambda)$: in subsequent stages until the last one, energy from the grid is too expensive and there is no profit for sales, therefore it is optimal to store all produced energy in the battery up to capacity. The total amount of energy produced over the time horizon is $E = \sum_{t=1}^{T} E_t = Mn - \sum_{i=1}^{n} X_t$. If there are less than Mn units in the battery at the last stage, we must buy the remaining amount $(Mn - E - x)^+ = (\sum_{i=1}^{n} X_t - x)^+$ at unit cost 1. It follows that this problem is equivalent to a newsvendor problem with demand $\sum_{i=1}^{n} X_t$ where the overage cost is $(1-\lambda)$ and the underage cost is $1 - (1-\lambda) = \lambda$, therefore using the same argument as above, we have that determining the optimal policy for this restriction of (SNES-NS) is #P-hard. The proof to show that determining the optimal value for the instance is at least as difficult as determining the optimal policy follows the same argument as for the case of stochastic energy demand. □

We can use Theorem 6.5 to prove that no relative approximation of the optimal value of (SNES) is possible in polynomial time. First, the next corollary shows that one cannot determine efficiently whether (SNES) admits a policy that yields a positive profit. This immediately implies that an approximation of the optimal value in the usual sense (i.e. with bounded relative error) is impossible, because any approximation would answer the question of existence of a policy with positive profit.

Corollary 6.6. *It is #P-hard to determine if (SNES) admits a solution with positive profit, even if the random variables are independent.*

Proof of Corollary 6.6. Suppose we could determine in polynomial time if (SNES) admits a solution with positive profit polynomial time. Consider the instance P of (SNES-NS) with deterministic energy production described in the first part of the proof of Theorem 6.5, which has nonpositive profit because there are no revenues. Let M be a large enough number, e.g. the product of all the nonzero numbers in the description of P. Modify P into an instance P' by adding an initial time period $t = 0$ with deterministic energy production $E_0 = 1/M$, zero demand, and profit for selling to the grid $P_0 = M\delta$. Notice that there is a nonzero sale price, hence P' is an instance of (SNES) rather than (SNES-NS), and the random variables are independent. Since $P_t = 0$ for $t \geq 1$, it is optimal to sell E_0 at stage 0 for a profit of δ. Thus, the modification increases the optimal value of the instance by δ. If we could determine if P' admits a solution with positive profit, we would be able to perform binary search on δ to determine the optimal value of P; by Theorem 6.5, this is #P-hard. □

Corollary 6.7. *There can be no polynomial-time multiplicative approximation algorithm (i.e., with finite approximation ratio) for (SNES), unless P = NP.*

This settles the question of which stochastic variants of (SNES) are solvable in polynomial time: (SNES-EP) can be solved in polynomial time (Theorem 5.1), whereas (SNES-NS) and (SNES) cannot (Theorem 6.5, Corollary 6.6), even when only E_t or D_t are stochastic, and they are independent random variables with support of size two. In contrast, (SNES-EP) is only hard if the stochastic processes are difficult to handle (Theorem 5.2): this indicates that the decision problems associated with (SNES-NS) and (SNES) are intrinsically harder than that of (SNES-EP). Furthermore, (SNES) does not admit any constant-factor approximation (Corollary 6.7). However, the next section shows that (SNES-NS) is somewhat easier than (SNES), because efficient approximation algorithms can be constructed.

7. An approximation algorithm for (SNES-NS)

Corollary 6.7 tells us that there can be no approximation for (SNES) in general, but this does not exclude the possibility of an approximation algorithm for (SNES-NS). We show below that in fact an FPTAS exists. The proof has two steps: first, we show that a discretized version of the reformulation admits an FPTAS using the framework of [23]; then, we prove that there is a polynomial discretization of the problem that yields the optimal solution to (SNES-NS).

Lemma 7.1. *Consider an instance of (SNES-NS) such that the battery levels R_t and the action x_t^B are restricted to take on integer values. Then the instance admits an FPTAS.*

The proof of Lemma 7.1 is given in Appendix. We remark that the assumption of the Markov property appears necessary to provide an approximation: [23, Cor. 10.2] exhibits a stochastic dynamic program with non-Markovian random variables that fits into the framework and is APX-hard, i.e. it does not admit an FPTAS unless P = NP. We use Lemma 7.1 to construct an FPTAS for the original (not discretized) problem.

Theorem 7.2. *Any instance of (SNES-NS) admits an FPTAS.*

Proof of Theorem 7.2. We show that we can apply the FPTAS for integer domains of Lemma 7.1 on a transformation of the problem. In particular, we will show that the set of possible states is finite because all feasible states can be expressed as rational numbers with the same denominator, and we can construct an equivalent problem with integer domain by rescaling.

Let Δ be the least common denominator of $1/\eta^c$, η^d, γ^c, γ^d, and the values in the support of E_t, D_t for all t. Δ is polynomial in the binary input size. Recall that our problem has a piecewise linear convex value function at every stage and it admits an optimal dual threshold policy at every stage for every value of W_t. For fixed W_{t-1}, let $B_x \in \mathbb{R}$ be the set consisting of the projection of the breakpoints of $z_t(\cdot|W_{t-1})$ on the x axis. We show by backward induction that the points in B_x, which we simply call "breakpoints" from now on, have denominator that divides Δ. At stage $T+1$ this is obvious as the value function is identically zero. Assume now that the value function z_{t+1} satisfies the induction hypothesis, and consider $z_t(R_t|W_{t-1})$, defined as in (1). We now switch to a K-approximation of $z_{t+1}(\cdot|W_t)$, denoted by $\hat{z}_{t+1}$, which can be computed following the proof of Lemma 7.1. An approximation of the lower threshold $\ell(W_t)$ that yields a K-optimal policy (see [23, Prop. 10.3]) can be computed by solving:

$$\min_{0 \le x_t^B \le R^{\max}} g_t(0, x_t^B, W_t) + \hat{z}_{t+1}(x_t^B|W_t).$$

The function $g_t(0, x_t^B, W_t)$ is piecewise linear convex with all breakpoints having denominator that divides Δ by construction of Δ. The function $\hat{z}_{t+1}(\cdot|W_t)$ is piecewise linear convex with all breakpoints having denominator that divides Δ by the induction hypothesis. The minimum of a piecewise linear convex function occurs at one of the breakpoints, hence $\ell(W_t)$ has denominator that divides Δ. A similar argument, substituting $R^{\max}$ for R_t in (1) for fixed W_t, shows that the upper threshold $u(W_t)$ has a denominator that divides Δ.

For every value of W_t, we consider (1) separately for $R_t \leq \ell(W_t)$, $\ell(W_t) < R_t < u(W_t)$, and $R_t \geq u(W_t)$. In the first case, a K-approximation of the optimal policy is $\bar{x}_t^B = \min\{\gamma^c, \ell(W_t) - R_t\}$. If $\bar{x}_t^B = \gamma^c$, the breakpoints of $z_t(\cdot|W_{t-1})$ are those of $g_t(\cdot, \gamma^c, w)$ or those of $\hat{z}_{t+1}(\cdot|W_t)$ shifted by γ^c, hence their denominator divides Δ by construction of Δ and the induction hypothesis. If $\bar{x}_t^B = \ell(W_t) - R_t$, the value function reads:

$$g_t(R_t, \ell(W_t) - R_t, w) + \hat{z}_{t+1}(\ell(W_t)|W_t).$$

The breakpoints of $g_t(R_t, \ell(W_t) - R_t, W_t)$ have denominator that divides Δ by construction of Δ, and those of $\hat{z}_{t+1}(\ell(W_t)|W_t)$ have denominator that divides Δ by the induction hypothesis. Hence we are done. When $\ell(W_t) < R_t < u(W_t)$, the value function becomes:

$$g_t(R_t, 0, W_t) + \hat{z}_{t+1}(R_t|W_t),$$

and the same argument applies. Finally, when $R_t \geq u(W_t)$, we have $\bar{x}_t^B = \max\{-\gamma^d, u(W_t) - R_t\}$, and we can follow a similar approach to the case $R_t \leq \ell(W_t)$. In all cases, the approximate value function $\hat{z}_t(R_t|W_{t-1})$ has breakpoints with denominator that divides Δ. As discussed in the proof of Lemma 7.1, to compute the expectation $\mathbb{E}_{W_t}[\min\{g_t(R_t, x_t^B, W_t) + z_{t+1}(R_t + x_t^B|W_t)\}|W_{t-1}]$ we loop over the support of W_t, and the breakpoints of the expectation are contained in the union of the set of breakpoints of $z_{t+1}(R_t|W_t)$. Therefore, the denominator of the breakpoints of the value function $z_t(R_t|W_{t-1})$ divides Δ, and this holds for all W_{t-1}.

To conclude the proof, notice that if $\Delta = 1$ we are done; otherwise, transform the original instance of (SNES-NS) by multiplying $R^{\max}, \gamma^c, \gamma^d, \eta^d, E_t, D_t$ by Δ, and dividing η^c, c^h, C_t by Δ, for all t. By plugging into the objective function, it is easy to verify that this is just a rescaling that blows up the state space and the action space by a factor Δ, but does not change the cost. Hence, there is a one-to-one correspondence between a solution to the transformed instance and a solution of the original instance with the same cost. However, the least common denominator of ΔE_t, ΔD_t, Δ/η^c, $\Delta\eta^d$, $\Delta\gamma^c$, $\Delta\gamma^d$ is 1 by construction of Δ. Thus, the breakpoints of the value function of the transformed instance are integer points, and we can restrict the values of R_t, x_t^B to be integer without loss of generality. By Lemma 7.1, the transformed instance admits an FPTAS, and so does the original instance. □

In the proof above, the size of the state and action space of the transformed instance are possibly much larger than in the original instance, but the worst-case running time of the FPTAS depends on the logarithm of these numbers, and therefore computational efficiency is barely affected, i.e. the algorithm is still polynomial-time. This theoretical observation is consistent with the numerical study of Halman et al. [31], where it is empirically shown that the running time of a discrete FPTAS constructed using the technique of K-approximation sets scales very well with the size of the state and action space.

8. Conclusions

This paper studied the computational complexity of the problem of optimally managing an energy storage device connected to a source of renewable energy, a household, and the power grid, under different conditions. We identified some of the characteristics of the problem that make it transition from easy to hard from the point of view of computational complexity. One of our findings is that the version of the problem in which the sale price to the grid is the same as the buying price can be solved in polynomial time via dynamic programming, assuming that the stochastic processes describing the evolution of prices, renewable energy generation, and demand are of manageable size; our analysis suggests that the practical difficulty of solving this problem is mainly due to the necessity of dealing with the stochastic processes, rather than the structure of the decision problem itself. Another finding is that in the general case the problem is #P-hard even when the random variables are independent and with support of size two, and even deciding whether there exists a policy that achieves positive profit is hard. This suggests that the problem is intrinsically very hard, regardless of the intricacy of the random processes, and rules out the possibility of an approximation algorithm. If energy can only be bought from the grid the problem is still #P-hard, but in this case we can provide an FPTAS extending the technique of K-approximation sets introduced in [8]. Our extension concerns the construction of an FPTAS when the state space of the problem is continuous (as opposed to discrete), taking advantage of the piecewise linear convex structure of the value function, and it could easily be applied to other stochastic dynamic programs that share the same structure.

Acknowledgments

Part of this work was carried out at the Singapore University of Technology and Design. The authors are grateful for partial support from the following sources: Israel Science Foundation grant 399/17 (N. Halman), SUTD grant SRES11012 and IDC grants IDSF1200108, IDG21300102 (G. Nannicini), ONR grant N000141410073 (J. Orlin).

Appendix. Proofs omitted from the main text

Proof of Lemma 4.1. At time period t, let (R_t, W_t) be the state and let $\bar{x}_t$ be an optimal action, i.e. $\pi_t(R_t, W_t) = \bar{x}_t$.

Suppose that $\bar{x}_t$ does not satisfy $\bar{x}_t^{ED} = \min\{E_t, D_t\}$; we construct from $\bar{x}_t$ an action vector $\tilde{x}_t$ that satisfies $\tilde{x}_t^{ED} = \min\{E_t, D_t\}$, leads to the same state R_{t+1}, and yields a profit that is at least as large as $\bar{x}_t$. First, notice that for all t there exists an optimal action with $\bar{x}_t^{EG} + \bar{x}_t^{ED} + \bar{x}_t^{ER} = E_t$, because if $\bar{x}_t^{EG} + \bar{x}_t^{ED} + \bar{x}_t^{ER} < E_t$ we could increase $\bar{x}_t^{EG}$ yielding at least the same profit while not violating any constraint.

Consider first the case $E_t \le D_t$.

If $\bar{x}_t^{EG} > 0$, we can set $\tilde{x}_t^{EG} = 0, \tilde{x}_t^{ED} = \bar{x}_t^{ED} + \bar{x}_t^{EG}, \tilde{x}_t^{GD} = (\bar{x}_t^{GD} - \bar{x}_t^{EG})^+, \tilde{x}_t^{RG} = \bar{x}_t^{RG} + \frac{1}{\eta^d}(\bar{x}_t^{EG} - \bar{x}_t^{GD})^+, \tilde{x}_t^{RD} = \bar{x}_t^{RD} - \frac{1}{\eta^d}(\bar{x}_t^{EG} - \bar{x}_t^{GD})^+$. We can verify that this action leads to the same battery state R_{t+1}, yields at least the same profit as $\bar{x}_t$, and satisfies $\tilde{x}_t^{EG} = 0$. So we can assume $\bar{x}_t^{EG} = 0$. We now analyze the case $\bar{x}_t^{ER} > 0$. Let $\beta = \min\{\bar{x}_t^{GD}, \bar{x}_t^{ER}\}$. If $\beta > 0$, we can set $\tilde{x}_t^{ED} = \bar{x}_t^{ED} + \beta, \tilde{x}_t^{GD} = \bar{x}_t^{GD} - \beta, \tilde{x}_t^{ER} = \bar{x}_t^{ER} - \beta, \tilde{x}_t^{GR} = \bar{x}_t^{GR} + \beta$. This action yields exactly the same profit and battery state as $\bar{x}_t$. Observe that if $\beta = \bar{x}_t^{ER}$ the action satisfies $\tilde{x}_t^{ED} = E_t$, hence we are done. Otherwise, we found an optimal action with $\tilde{x}_t^{GD} = 0$, so we now assume $\bar{x}_t^{GD} = 0$. This implies $\bar{x}_t^{RD} = \frac{1}{\eta^d}(D_t - \bar{x}_t^{ED})$. We can set $\tilde{x}_t^{ED} = \bar{x}_t^{ED} + \bar{x}_t^{ER}, \tilde{x}_t^{ER} = 0, \tilde{x}_t^{RD} = \bar{x}_t^{RD} - \frac{1}{\eta^d}\bar{x}_t^{ER}, \tilde{x}_t^{RG} = \bar{x}_t^{RG} + \frac{1}{\eta^d}\bar{x}_t^{ER}$. As before, it is easy to verify that this action leads to the same battery state R_{t+1} and yields at least the same profit as $\bar{x}_t$. Furthermore, it satisfies $\tilde{x}_t^{ED} = E_t$. This implies that there exists an optimal action with $\tilde{x}_t^{ED} = E_t$.

Next, we consider the case $E_t > D_t$.

If $\bar{x}_t^{EG} > 0$, let $\beta = \min\{\bar{x}_t^{EG}, D_t - \bar{x}_t^{ED}\} > 0$. We can set $\tilde{x}_t^{EG} = \bar{x}_t^{EG} - \beta, \tilde{x}_t^{ED} = \bar{x}_t^{ED} + \beta, \tilde{x}_t^{GD} = (\bar{x}_t^{GD} - \beta)^+, \tilde{x}_t^{RG} = \bar{x}_t^{RG} + \frac{1}{\eta^d}(\beta - \bar{x}_t^{GD})^+, \tilde{x}_t^{RD} = \bar{x}_t^{RD} - \frac{1}{\eta^d}(\beta - \bar{x}_t^{GD})^+$. We can verify that this action leads to the same battery state R_{t+1} and yields at least the same profit as $\bar{x}_t$. Furthermore, if $\beta = D_t - \bar{x}_t^{ED}$ we have $\tilde{x}_t^{ED} = D_t$ and we are done, otherwise $\tilde{x}$ satisfies $\tilde{x}_t^{EG} = 0$. So we can assume $\bar{x}_t^{EG} = 0$. We now analyze the case $\bar{x}_t^{ER} > 0$. Notice that $\bar{x}_t^{ED} + \bar{x}_t^{ER} = E_t > D_t$ implies $\bar{x}_t^{ER} > D_t - \bar{x}_t^{ED}$. Let $\beta = \min\{\bar{x}_t^{GD}, D_t - \bar{x}_t^{ED}\}$.

If $\beta > 0$, we can set $\tilde{x}_t^{ED} = \bar{x}_t^{ED} + \beta, \tilde{x}_t^{GD} = \bar{x}_t^{GD} - \beta, \tilde{x}_t^{ER} = \bar{x}_t^{ER} - \beta, \tilde{x}_t^{GR} = \bar{x}_t^{GR} + \beta$. This action yields exactly the same profit and battery state R_{t+1} as $\bar{x}_t$. Observe that if $\beta = D_t - \bar{x}_t^{ED}$ the action satisfies $\tilde{x}_t^{ED} = D_t$, hence we are done. Otherwise, we found an optimal action with $\tilde{x}_t^{GD} = 0$, so we now assume $\bar{x}_t^{GD} = 0$. This implies $\bar{x}_t^{RD} = \frac{1}{\eta^d}(D_t - \bar{x}_t^{ED})$. We can set $\tilde{x}_t^{ED} = D_t \tilde{x}_t^{ER} = \bar{x}_t^{ER} - (D_t - \bar{x}_t^{ED}), \tilde{x}_t^{RD} = 0, \tilde{x}_t^{RG} = \bar{x}_t^{RG} + \frac{1}{\eta^d}(D_t - \bar{x}_t^{ED})$. It can be verified that this action leads to the same battery state R_{t+1} and yields at least the same profit as $\bar{x}_t$. Furthermore, it satisfies $\tilde{x}_t^{ED} = D_t$. This implies that there exists an optimal action with $\tilde{x}_t^{ED} = D_t$, and concludes the proof. □

Proof of Proposition 4.2. We show that the DP described in the statement of the proposition is simply a reformulation of (SNES) that can be obtained under the assumption $P_t = 0$. Define x_t^B as the net energy balance of the battery at time t, that is, $x_t = R_{t+1} - R_t$. Notice that:

$$x_t^B = R_{t+1} - R_t = -x_t^{RD} - x_t^{RG} + \eta^c(x_t^{GR} + x_t^{ER}).$$

One can verify that there exists an optimal action $\bar{x}_t$ such that either $\bar{x}_t^{RD} + \bar{x}_t^{RG} > 0$ or $\bar{x}_t^{GR} + \bar{x}_t^{ER} > 0$, but not both. Indeed, because $P_t \leq C_t$ we can assume $\bar{x}_t^{RG} > 0$ and $\bar{x}_t^{GR} > 0$, but not both. Furthermore, by Lemma 4.1 we can assume $\bar{x}_t^{RD} > 0$ or $\bar{x}_t^{ER} > 0$, but not both. Finally, we observe that if $\bar{x}_t^{RD} > 0, \bar{x}_t^{GR} > 0$ we can construct an action $\tilde{x}_t$ with $\tilde{x}_t^{RD} = \bar{x}_t^{RD} - \min\{\bar{x}_t^{RD}, \frac{1}{\eta^c}\bar{x}_t^{GR}\}, \tilde{x}_t^{GR} = \bar{x}_t^{GR} - \min\{\bar{x}_t^{RD}, \frac{1}{\eta^c}\bar{x}_t^{GR}\}, \tilde{x}_t^{GD} = \bar{x}_t^{GD} + \eta^d \min\{\bar{x}_t^{RD}, \frac{1}{\eta^c}\bar{x}_t^{GR}\}$ that yields the same battery state and at least the same profit as $\bar{x}_t$, while either $\bar{x}_t^{RD} > 0$ or $\bar{x}_t^{GR} > 0$ but not both. A similar argument shows that there exists an optimal action where $x_t^{RG} > 0$ or $x_t^{ER} > 0$ but not both. This shows that $\bar{x}_t^{RD} + \bar{x}_t^{RG} > 0$ or $\bar{x}_t^{GR} + \bar{x}_t^{ER} > 0$ but not both.

From the discussion above we can assume $(x_t^B)^+ = \eta^c(x_t^{GR} + x_t^{ER})$ and $(x_t^B)^- = x_t^{RD} + x_t^{RG}$. From Lemma 4.1, if $(x_t^B)^+ > 0$, it is optimal to set $x_t^{ER} = \min\{\frac{1}{\eta^c}(x_t^B)^+, (E_t - D_t)^+\}$ and $x_t^{GR} = (\frac{1}{\eta^c}(x_t^B)^+ - (E_t - D_t)^+)^+$, and if $(x_t^B)^- > 0$, it is optimal to set $x_t^{RD} = \min\{\eta^d(x_t^B)^-, (D_t - E_t)^+\}, x_t^{RG} = (\eta^d(x_t^B)^- - (D_t - E_t)^+)^+$. Any leftover demand is satisfied by setting $x_t^{GD} = (D_t - E_t - \eta^d(x_t^B)^-)^+$.

By plugging these values in the original formulation of (SNES), we obtain the reformulation stated above. The objective function reads:

$$g_t(R_t, x_t^B, W_t) := C_t((\frac{1}{\eta^c}(x_t^B)^+ - (E_t - D_t)^+)^+ + (D_t - E_t - \eta^d(x_t^B)^-)^+) + c^h(R_t + x_t^B),$$

This can be rewritten in the following form:

$$g_t(R_t, x_t^B, W_t) := \begin{cases} C_t(\frac{1}{\eta^c}x_t^B + D_t - E_t)^+ + c^h(R_t + x_t^B) & \text{if } x_t^B \geq 0 \\ C_t(D_t - E_t + \eta^d x_t^B)^+ + c^h(R_t + x_t^B) & \text{otherwise.} \end{cases}$$

Notice that the function is linear in R_t, convex in x_t^B, and continuous. Furthermore, the action space depends on R_t only and could be written as $\mathcal{A}_t(R_t)$. □

Proof of Proposition 4.3. The proof follows the same lines as for Proposition 4.2. Defining x_t^B as in Proposition 4.2, and using Lemma 4.1, it is easy to check that (SNES) can be formulated as above. A simpler, intuitive proof is the following: since the sale price is the same as the buying price, one can always sell the entirety of the energy produced to the grid, and satisfy the demand from the grid. The only decision left to take is by how much the battery should be charged or discharged at the current price P_t. This yields the reformulation. □

Proof of Lemma 7.1. We employ the framework of K-approximation sets and functions (where $K = 1+\epsilon$) described in [23]. Intuitively, a K-approximation set of a discrete monotone increasing function φ is a polylogarithmic-sized subset of its domain such that the value of φ increases by a factor $\leq K$ between successive points; for a formal definition, see [23, Def. 4.2]. Our goal is to show that we can cast the model for (SNES-NS) given in Proposition 4.2 in a form that satisfies the sufficient conditions for convex dynamic programs [23, Condition 3]. We list the sufficient conditions, which are somewhat technical but amount to saying that the problem has a convex structure that preserves convexity by backward induction, even when restricted to integer domains:

(i) The action and state spaces are sets of integers without holes at all time periods.
(ii) The terminal cost function is convex.
(iii) The set $\mathcal{S}_t \otimes \mathcal{A}_t := \{(s,a) \in \mathbb{Z}^2 : s \in \mathcal{S}_t, a \in \mathcal{A}_t(s)\}$, where $\mathcal{S}_t, \mathcal{A}_t$ are respectively the state space and the action space at time period t, and $\mathcal{A}_t(s)$ is the set of feasible actions at state s, is integrally convex for every $t = 1,\dots,T$. (A subset S of the two-dimensional integer lattice is *integrally convex* if there exists a polyhedron P such that $S = P \cap \mathbb{Z}^2$ and the slopes of the edges of P are an integral multiple of 45°.)
(iv) The cost function $g_t(R,x,w)$ can be expressed as $g_t^R(R,w)+g_t^x(x,w)+u_t(f_t(R,x,w))$, and the transition function f_t can be expressed as $f_t(R,x,w) = a(w)R + b(w)x + c(w)$, where w is a random variable, $g_t^R(\cdot,w), g_t^x(\cdot,w), u_t(\cdot)$ are nonnegative univariate convex functions, $a(w) \in \mathbb{Z}, b(w) \in \{-1,0,1\}$, and $c(w) \in \mathbb{Z}$.

Extensions to this framework and more relaxed conditions are given in [32], but the ones above suffice for the present paper. Recall that the value function can be written as:

$$z_t(R_t|W_{t-1}) = \mathbb{E}_{W_t}\left[\min_{x_t^B \in A_t(R_t)} \{g_t(R_t, x_t^B, W_t) + z_{t+1}(R_t + x_t^B|W_t)\}|W_{t-1}\right]. \tag{3}$$

Since the values of W_t and the transition probabilities are given explicitly, it is enough to show that the sufficient conditions hold for fixed W_t: if we can compute a piecewise linear convex approximation of the expression inside the expected value using an approximation of the value function at the subsequent stage z_{t+1}, we can compute $z_t(R_t|W_{t-1})$ looping over the support of W_t in time proportional to q. This process has to be repeated q times at each stage, yielding an algorithm that runs in time proportional to q^2.

Since we now assume that W_t is fixed, condition (i) holds for the set of integer battery levels $[0,\dots,R^{\max}]$, see also [23, Sec. 10.3]. Condition (ii) trivially holds. For condition (iii), notice that the constraints of the DP are $-\min\{\gamma^d, R_t\} \le x_t^B \le \min\{\gamma^c, R^{\max} - R_t\}$. The corresponding $\mathcal{S}_t \otimes \mathcal{A}_t$ is integrally convex. For condition (iv), we see that for fixed value of W_t the transition function $R_{t+1} = R_t + x_t^B$ clearly satisfies the condition, and the cost function $g_t(R_t, x_t^B, W_t)$ of Proposition 4.2 can be expressed as $g_t^x(x_t^B, w) + u_t(f_t(R_t, x_t^B, w))$ by setting $u_t(f_t(R_t, x_t^B, w)) = c^h(R_t + x_t^B)$ and $g_t^x(x_t^B, w) = C_t(D_t - E_t + \frac{1}{\eta^c}(x_t^B)^+ - \eta^d(x_t^B)^-)^+$, with $g_t^x(\cdot, w)$ univariate and convex. Therefore, for every value of W_t, the sufficient conditions hold. This implies that we can compute a piecewise linear convex approximation of the value function at each stage, and apply the FPTAS [23, Alg. 4] to recover an ϵ-approximation of the optimal value for any $\epsilon > 0$. Notice that an approximate policy for a given value of W_t can be computed by storing the (polynomially-many) approximate value functions. □

References

[1] J.M. Nascimento, Approximate Dynamic Programming for Complex Storage Problems, Princeton University, 2008 (Ph.D. thesis).
[2] G. Natarajan, Y. Xu, J. Bradley, Meeting inelastic demand in systems with storage and renewable sources, in: Proceedings of the IEEE Fifth International Conference on Smart Grid Communications (SmartGridComm), 2014.
[3] W.R. Scott, W.B. Powell, Approximate dynamic programming for energy storage with new results on instrumental variables and projected Bellman errors, Technical report, Princeton University, 2012.
[4] Y. Zhou, A. Scheller-Wolf, N. Secomandi, S. Smith, Managing wind-based electricity generation in the presence of storage and transmission capacity. Technical Report 2011-E36, Tepper School of Business, Carnegie Mellon University, 2013.

[5] J.M. Eyer, G.P. Corey, Energy storage for the electricity grid: Benefits and market potential assessment guide. Technical Report SAND2010-0815, Sandia National Laboratories, pp. 69–73, 2010.

[6] J.M. Eyer, J.J. Iannucci, G.P. Corey, Energy storage benefits and market analysis handbook, a study for the DOE energy storage systems program. Technical Report SAND2004-6177, Sandia National Labratories, 2004.

[7] D. Salas, W.B. Powell, Benchmarking a scalable approximation dynamic programming algorithm for stochastic control of multidimensional energy storage problems. Technical report, Princeton University, 2013.

[8] N. Halman, D. Klabjan, M. Mostagir, J. Orlin, D. Simchi-Levi, A fully polynomial time approximation scheme for single-item inventory control with discrete demand, Math. Oper. Res. 34 (3) (2009) 674–685.

[9] E.L. Porteus, Foundations of Stochastic Inventory Theory, Stanford Business Books, Palo Alto, CA, 2002.

[10] P.H. Zipkin, Foundations of Inventory Management, McGraw-Hill, New York, NY, 2000.

[11] R. Rempala, Optimal strategy in a trading problem with stochastic prices, in: J. Henry, J.-P. Yvon (Eds.), System Modelling and Optimization, in: Lecture Notes in Control and Information Sciences, vol. 197, Springer Berlin Heidelberg, 1994, pp. 560–566.

[12] N. Secomandi, Optimal commodity trading with a capacitated storage asset, Manage. Sci. 56 (3) (2010) 449–467.

[13] F. Graves, T. Jenkins, D. Murphy, Opportunities for electricity storage in deregulating markets, Electr. J. 12 (8) (1999) 45–56.

[14] Y. Zhou, A. Scheller-Wolf, N. Secomandi, S. Smith, Electricity trading and negative prices: storage vs. disposal, Manage. Sci. 62 (3) (2016) 880–898.

[15] P. Harsha, M. Dahleh, Optimal management and sizing of energy storage under dynamic pricing for the efficient integration of renewable energy, IEEE Trans. Power Syst. 30 (3) (2015) 1164–1181.

[16] S. Moazeni, W.B. Powell, A.H. Hajimiragha, Mean-conditional value-at-risk optimal energy storage operation in the presence of transaction costs, IEEE Trans. Power Syst. 30 (3) (2015) 1222–1232.

[17] J.M. Nascimento, W.B. Powell, An optimal approximate dynamic programming algorithm for concave, scalar storage problems with vector-valued controls, IEEE Trans. Automat. Control 58 (12) (2013) 2995–3010.

[18] D.R. Jiang, T.V. Pham, W.B. Powell, D.F. Salas, W.R. Scott, A comparison of approximate dynamic programming techniques on benchmark energy storage problems: Does anything work? in: Adaptive Dynamic Programming and Reinforcement Learning (ADPRL), 2014 IEEE Symposium on, 2014, pp. 1–8.

[19] L. Huang, J. Walrand, K. Ramchandran, Optimal demand response with energy storage management. in: Proceedings of the IEEE Third International Conference on Smart Grid Communications (SmartGridComm), 2012, pp. 61–66.

[20] N. Halman, G. Nannicini, Fully polynomial-time (Σ,Π)-approximation schemes for continuous convex stochastic dynamic programs. Technical Report 4990, Optimization Online, 2015.

[21] W.D. Powell, Bridging the fields of stochastic optimization, Optima 96 (2015) 1–5.

[22] A. Charnes, J. Drèze, M. Miller, Decision and horizon rules for stochastic planning problems: a linear example, Econometrica 34 (2) (1966) 307–330.

[23] N. Halman, D. Klabjan, C.-L. Li, J. Orlin, D. Simchi-Levi, Fully polynomial time approximation schemes for stochastic dynamic programs, SIAM J. Discrete Math. 28 (4) (2014) 1725–1796.

[24] N. Halman, J.B. Orlin, D. Simchi-Levi, Approximating the nonlinear newsvendor and single-item stochastic lot-sizing problems when data is given by an oracle, Oper. Res. 60 (2) (2012) 429–446.

[25] J.B. Orlin, A faster strongly polynomial minimum cost flow algorithm, Oper. Res. 41 (2) (1993) 338–350.

[26] P. Favati, F. Tardella, Convexity in nonlinear integer programming, Ric. Oper. 53 (1) (1990) 3–44.

[27] K. Murota, A. Shioura, Relationship of M-/L-convex functions with discrete convex functions by Miller and FavatiTardella, Discrete Appl. Math. 115 (1–3) (2001) 151–176.

[28] K. Murota, Discrete Convex Analysis, SIAM, Philadelphia, PA, 2003.

[29] L. Valiant, The complexity of computing the permanent, Theoret. Comput. Sci. 8 (2) (1979) 189–201.

[30] S. Toda, PP is as hard as the polynomial-time hierarchy, SIAM J. Comput. 20 (5) (1991) 865–877.

[31] N. Halman, G. Nannicini, J. Orlin, A computationally efficient FPTAS for convex stochastic dynamic programs, SIAM J. Optim. 25 (1) (2015) 317–350.

[32] N. Halman, Provably near-optimal approximation schemes for implicit stochastic and for sample-based dynamic programs. Technical Report 4952, Optimization Online, 2015.

Please cite this article in press as: N. Halman, et al., On the complexity of energy storage problems, Discrete Optimization (2017), https://doi.org/10.1016/j.disopt.2017.11.001.