Multi-Stage Stochastic Optimization via Parameterized Stochastic Hybrid Approximation

Tomás Tinoco De Rubira Line Roald Gabriela Hug

May 22, 2017

Abstract

Many important complex decision-making processes under uncertainty can be represented as convex multi-stage stochastic optimization problems. However, despite convexity, solving these problems is in general an intractable task due to the presence of nested expectations. The approaches used in practice and proposed in the literature for solving problems of this type typically consist of forming gradually-improving piecewise linear approximations of nested expectation functions using cutting planes. These approaches have been shown to work well on problems with specific properties, such as having linear functions and finite-support uncertainty, but their effectiveness on more general problems remains a challenge. In this work, a new algorithm for solving convex multi-stage stochastic optimization problems is presented. The algorithm is based on a parameterized extension of the stochastic hybrid approximation procedure, which allows leveraging user-provided deterministic approximations of expected cost-to-go-functions, and can be applied to problems with continuous and stage-wise dependent uncertainty. A theoretical analysis of the algorithm is presented in which conditions that guarantee its convergence are provided. Results obtained from comparing the proposed and benchmark algorithms including the popular stochastic dual dynamic programming algorithm on a power system operations planning problem under renewable energy uncertainty are also reported. The results suggest that the proposed algorithm can exploit accurate initial user-provided expected cost-to-go function approximations and achieve better performance compared to the benchmarks.

1 Introduction

Many important applications, such as financial planning, supply chain management, and power system operations planning, can be modeled as sequential decision-making processes under uncertainty [30]. A typical goal in these applications is to determine a decision policy that optimizes a certain performance measure, such as total expected cost. However, obtaining such a policy is in general an intractable task. The reason for this is that the decisions at each point in time can depend on all the decisions and observations made so far, and are subject to any remaining uncertainty present in the future.

Multi-stage stochastic optimization has been an invaluable tool for obtaining decision policies for this type of applications. By far the most common approach has been to represent or approximate the exogenous uncertainty with a finite number of scenarios, and use cutting planes to construct gradually-improving deterministic approximations of expected cost-to-go functions. For example, the algorithm proposed by Birge for solving problems with linear functions, which can be considered a multi-stage extension of Benders decomposition and the L-shaped method [3] [39], performs forward passes through all scenarios to construct candidate solutions, and backward passes to construct cutting planes and update piecewise linear cost-to-go function approximations [6]. An

issue with approaches such as this one is that the number of scenarios can grow exponentially with the length of the decision-making horizon, and hence processing all of them at each iteration is in general computationally impractical. Pereira & Pinto developed the stochastic dual dynamic programming (SDDP) algorithm for problems with stage-wise independent uncertainty, which works with a relatively small number of scenarios that are randomly sampled at each iteration [23] [32] [33]. This algorithm has become a popular tool used by practitioners, in particular for the optimization of hydro power systems [2] [27] [33] [40], and has also been analyzed and extended by several authors in the literature. For example, Philpott & Guan provided a rigorous analysis of the convergence of a broad class of sampling-based algorithms that includes SDDP, and discussed specific requirements for the sampling procedure [25]. Rebennack adapted SDDP to work with stage-wise dependent uncertainty, and showed the applicability of the algorithm on a hydro-thermal scheduling problem with various types of uncertainty [27]. Asamov & Powell, on the other hand, showed how to improve the convergence speed of SDDP by including quadratic regularization in a way that does not require obtaining exponentially-many incumbent solutions [2]. Another sampling-based decomposition algorithm related to SDDP is the multi-stage stochastic decomposition algorithm proposed by Sen & Zhou [31]. This algorithm, which is a dynamic extension of the sampled regularized algorithm proposed by Higle & Sen for two-stage problems [12], constructs statistical estimates of supports of cost-to-go functions that improve over time [11]. Lastly, a different family of algorithms explored in the literature for solving convex multi-stage stochastic optimization problems is that of algorithms based on the principle of progressive hedging [29]. These algorithms model the non-anticipativity of policies using linear constraints that are relaxed through the use of an approximate augmented Lagrangian function. Rockafellar & Wets were the first to describe an algorithm based on this idea [29]. In their algorithm, non-anticipative policies are obtained at each iteration through a projection operator. Mulvey & Ruszczyński proposed an algorithm based on similar ideas for solving large-scale problems exploiting parallel computing resources [19].

Most of the algorithms proposed in the literature for solving convex multi-stage stochastic optimization problems have specifically targeted problems having linear functions. For such problems, expected cost-to-go functions can be shown to be piecewise linear and hence algorithms that build piecewise linear approximations such as SDDP are a natural choice. Only a relatively few number of authors have studied the applicability of these algorithms to problems having nonlinear functions, or extended them to solve such problems more efficiently. For example, Girardeau, Leclere & Philpott studied the convergence of sampling-based nested decomposition algorithms that include SDDP on problems having general convex objective functions [10]. They showed that these algorithms converge on such problems, but did not perform numerical experiments. From [20], it is known that the performance of cutting-plane algorithms can be very slow on instances of problems with general convex objectives, and hence these algorithms may not be the best choice for problems beyond ones having only linear functions. Another example is the work of Louveaux, which consists of a nested decomposition algorithm for problems with quadratic objectives and linear constraints [18]. The algorithm approximates cost-to-go functions by constructing and updating pieces of polyhedral piecewise quadratic functions. However, the algorithm does not consider sampling and hence suffers from the same computational limitations as other early nested decomposition algorithms that require traversing the complete scenario tree. Lau & Womersley also proposed an algorithm for solving problems with quadratic objectives [17]. The algorithm uses a recursive generalized Newton procedure, and requires recursively solving entire subtrees of the scenario tree.

In this work, a new sampling-based algorithm is proposed for solving convex multi-stage stochastic optimization problems having nonlinear functions for which cutting-plane methods such as SDDP may not be the most efficient. The algorithm utilizes user-provided initial approximations of expected cost-to-go functions that are gradually improved using a stochastic hybrid approxima-

tion procedure. Hence, it can leverage application-specific knowledge and start with approximations that with cutting planes alone could be potentially time-consuming to obtain. The stochastic hybrid approximation procedure on which this algorithm is based consists of using noisy subgradients that are computed at each iteration to make slope corrections to the expected cost-to-go function approximations. This procedure was proposed by Cheung & Powell for solving two-stage stochastic optimization problems, and was motivated by a dynamic resource allocation problem [8]. Tinoco De Rubira & Hug tested this procedure on a generator dispatch problem under renewable energy uncertainty using an initial approximation based on the certainty-equivalent problem and obtained promising results [35]. The same authors later extended this procedure to handle expected-value constraints [36]. To use this procedure in the multi-stage setting, the algorithm proposed in this work parameterizes the step lengths, slope corrections, and the expected cost-to-go function approximations by the observation history of the exogenous uncertainty. Additionally, the slope corrections obtained for a function associated with a specific realization of uncertainty are generalized and applied to other functions associated with similar realizations of uncertainty using radial basis functions [1] [37]. This allows the algorithm to be applied directly to problems with continuous uncertainty without the need of discretization or scenario tree construction. An analysis of the convergence of the algorithm is provided for the case of finite-support uncertainty since the presence of nested expectations makes the continuous case theoretically intractable. Lastly, the performance of the algorithm is compared against that of three benchmark algorithms on several instances of a stochastic optimal generator dispatch problem from power system operations planning. The benchmark algorithms consist of a greedy, a certainty-equivalent, and an SDDP algorithm. The results obtained suggest that the proposed algorithm can leverage accurate initial expected cost-to-go function approximations provided by the user and obtain better policies compared to the benchmarks and in shorter times.

This paper is organized as follows: In Section 2, the class of convex multi-stage stochastic optimization problems considered in this work is described. In Section 3, an SDDP algorithm that is used as the main benchmark algorithm is presented. In Section 4, the proposed parameterized stochastic hybrid approximation algorithm is described. In Section 5, the application on which the algorithms considered in this work are tested is discussed and results of numerical experiments are provided. In Section 6, the work is summarized and next research directions are outlined. Lastly, the convergence analysis of the proposed algorithm can be found in the Appendix.

2 Convex Multi-Stage Stochastic Optimization Problems

The general problem considered consists of the following nested optimization problems:

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + G_t(x_t, \mathcal{W}_t) \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x_{t-1}, w_t), \end{aligned} \tag{1}$$

for $t \in \{1, \ldots, T\}$, where $T \in \mathbb{N}$, $F_t(\,\cdot\,, w_t)$ are continuous convex functions, $\mathcal{X}_t(\,\cdot\,, w_t)$ are convex[1] point-to-set mappings with convex compact outputs, $w_t \in \Omega_t$ is an exogenous random vector observed at time t with w_1 being deterministic, Ω_t are compact sets, $\mathcal{W}_t := \{w_1, \ldots, w_t\}$ is the observation history up to time t, and x_0 is a constant vector. The cost-to-go functions $G_t(\,\cdot\,, \mathcal{W}_t)$ are defined by

$$G_t(x, \mathcal{W}_t) := \begin{cases} \mathbb{E}\,[\, H_{t+1}(x, \mathcal{W}_{t+1}) \mid \mathcal{W}_t \,] & \text{if } t \in \{1, \ldots, T-1\}, \\ 0 & \text{else}, \end{cases} \tag{2}$$

[1]For all inputs x and y, and any $\lambda \in [0, 1]$, $\lambda \mathcal{X}_t(x, w_t) + (1-\lambda)\mathcal{X}_t(y, w_t) \subset \mathcal{X}_t(\lambda x + (1-\lambda)y, w_t)$.

for all x, where the function H_t is such that $H_t(x, \mathcal{W}_t)$ is the optimal objective value of problem (1) with $x_{t-1} = x$, and $\mathbb{E}[\,\cdot\,|\,\mathcal{W}_t\,]$ is expectation with respect to $\mathcal{W}_{t+1}$ conditioned on $\mathcal{W}_t$, which is assumed to be always well defined.

It is assumed that for all $t \in \{1, \ldots, T-1\}$ and $\mathcal{W}_{t+1}$, there exists an $\epsilon > 0$ such that for all $\{x_1, \ldots, x_t\}$ that satisfy $x_\tau \in \mathcal{X}_\tau(x_{\tau-1}, w_\tau)$ for each $\tau \in \{1, \ldots, t\}$, and for all δ that satisfy $\|\delta\|_2 < \epsilon$, it holds that $\mathcal{X}_{t+1}(x_t + \delta, w_{t+1}) \neq \emptyset$. This property is slightly stronger than *relatively complete recourse*, which guarantees the feasibility of problem (1) [2] [9]. In fact, it is similar to the *extended relatively complete recourse* condition introduced in [10].

From the assumed properties of $F_t(\,\cdot\,, w_t)$ and $\mathcal{X}_t(\,\cdot\,, w_t)$, it can be shown that $H_t(\,\cdot\,, \mathcal{W}_t)$ and $G_t(\,\cdot\,, \mathcal{W}_t)$ are real-valued convex continuous functions on the convex sets of stage-$(t-1)$ and stage-t feasible inputs, respectively, for each $t \in \{1, \ldots, T\}$.

3 Stochastic Dual Dynamic Programming

3.1 Background and Overview

The SDDP algorithm is a sampling algorithm based on nested Benders decomposition. It works with a finite scenario tree that represents the exogenous random process. Hence, for problems with continuous uncertainty, it solves and approximation of the problem. Since this algorithm is widely used in practice, especially on linear multi-stage problems that arise in the scheduling of hydro-thermal power systems [2] [27] [33] [40], it is used here as a benchmark to assess the performance of the proposed algorithm described in Section 4.

The first version of SDDP proposed by Pereira & Pinto assumed stage-wise independent uncertainty, which implies that expected cost-to-go functions are identical at all the nodes of the scenario tree that correspond to the same stage [23]. This property can be exploited to improve the efficiency of the algorithm by sharing information, *e.g.*, cutting planes, between cost-to-go functions corresponding to same-stage nodes. The problem described in Section 2 does not assume stage-wise independent uncertainty, and hence the version of the SDDP algorithm considered here treats the expected cost-go-to functions at different nodes of the scenario tree as different functions. This case is also considered by Girardeau, Leclere & Philpott [10], and Rebennack [27].

In the following, the notation used to represent the scenario tree is introduced, and a sample-average approximation (SAA) of problem (1) based on this scenario tree is described. Then, the SDDP algorithm is presented followed by a list of references where its convergence is analyzed.

3.2 Sample-Average Approximation

For problems with continuous uncertainty, a commonly-used approach for constructing a scenario tree is conditional Monte Carlo sampling [13] [33]. The resulting scenario tree can be represented as follows: For each $t \in \{1, \ldots, T\}$, $\mathcal{N}(t) \subset \Omega_t$ denotes the set of nodes associated with stage t. The set $\mathcal{C}(w_t) \subset \mathcal{N}(t+1)$ denotes the children of node $w_t \in \mathcal{N}(t)$. For $w_T \in \mathcal{N}(T)$, $|\mathcal{C}(w_T)| = 0$, *i.e.*, nodes of stage T are leaf nodes, and $w_1 \in \mathcal{N}(1) = \{w_1\}$ is the root or deterministic node of the tree. An arbitrary tree branch of length t is denoted by $\mathcal{W}_t := \{w_1, \ldots, w_t\}$, where $w_{\tau+1} \in \mathcal{C}(w_\tau)$ for all τ such that $1 \leq \tau < t$. The set of all t-length branches is denoted by $\mathcal{B}(t)$.

Using this notation for representing the scenario tree, the SAA of problem (1) then consists of the following nested optimization problems:

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + \widetilde{G}_t(x_t, \mathcal{W}_t) \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x_{t-1}, w_t), \end{aligned} \tag{3}$$

for $t \in \{1, \ldots, T\}$ and $\mathcal{W}_t \in \mathcal{B}(t)$. The cost-to-go functions $\widetilde{G}_t(\,\cdot\,, \mathcal{W}_t)$ are defined by

$$\widetilde{G}_t(x, \mathcal{W}_t) := \begin{cases} |\mathcal{C}(w_t)|^{-1} \sum_{w \in \mathcal{C}(w_t)} \widetilde{H}_{t+1}(x, \mathcal{W}_t \cup \{w\}) & \text{if } t \in \{1, \ldots, T-1\}, \\ 0 & \text{else,} \end{cases} \tag{4}$$

for all x, where the function $\widetilde{H}_t$ is such that $\widetilde{H}_t(x, \mathcal{W}_t)$ is the optimal objective value of problem (3) with $x_{t-1} = x$.

3.3 Algorithm

At each iteration $k \in \mathbb{Z}_+$, the SDDP algorithm considered here performs the following steps: First, a tree branch $\mathcal{W}_t^k = \{w_1^k, \ldots, w_t^k\} \in \mathcal{B}(t)$ is selected uniformly at random for $t = T$, independently from samples drawn in previous iterations. Then, the problem

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + \widetilde{G}_t^k(x_t, \mathcal{W}_t) \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x_{t-1}, w_t) \end{aligned} \tag{5}$$

is solved *forward in time* for each $t = 1, \ldots, T$ with $x_{t-1} = x_{t-1}^k$ and $\mathcal{W}_t = \mathcal{W}_t^k$ to obtain a solution x_t^k. The vector $x_0^k := x_0$ is a constant and the function $\widetilde{G}_t^k(\,\cdot\,, \mathcal{W}_t)$ is a piecewise linear approximation of the expected cost-to-go function $\widetilde{G}_t(\,\cdot\,, \mathcal{W}_t)$ defined in (4) such that $\widetilde{G}_t^0 := 0$ for all t and $\widetilde{G}_T^k := 0$ for all k. After reaching the end of the branch, the branch is processed *backward in time*. More specifically, for each $t = T-1, \ldots, 1$, a subgradient

$$\nu_{t+1}(w) \in \partial \widetilde{H}_{t+1}^{k+1}\big(x_t^k, \mathcal{W}_t^k \cup \{w\}\big)$$

is computed for each $w \in \mathcal{C}(w_t^k)$, where the function $\widetilde{H}_t^k$ is such that $\widetilde{H}_t^k(x, \mathcal{W}_t)$ is the optimal objective value of problem (5) with $x_{t-1} = x$. The computed subgradients $\big\{\, \nu_{t+1}(w) \;\big|\; w \in \mathcal{C}(w_t^k) \,\big\}$, are used to construct a new function approximation of $\widetilde{G}_t\big(\,\cdot\,, \mathcal{W}_t^k\big)$ as follows:

$$\widetilde{G}_t^{k+1}\big(x, \mathcal{W}_t^k\big) = \begin{cases} h_t^k\big(x, \mathcal{W}_t^k\big) & \text{if } k = 0, \\ \max\Big\{\widetilde{G}_t^k\big(x, \mathcal{W}_t^k\big), h_t^k\big(x, \mathcal{W}_t^k\big)\Big\} & \text{else,} \end{cases}$$

for all x, where

$$h_t^k\big(x, \mathcal{W}_t^k\big) := \big|\mathcal{C}(w_t^k)\big|^{-1} \sum_{w \in \mathcal{C}(w_t^k)} \Big(\widetilde{H}_{t+1}^{k+1}\big(x_t^k, \mathcal{W}_t^k \cup \{w\}\big) + \nu_{t+1}(w)^T\big(x - x_t^k\big)\Big).$$

For all $t \in \{1, \ldots, T-1\}$ and all t-length branches $\mathcal{W}_t$ not sampled during iteration k, *i.e.*, $\mathcal{W}_t \neq \mathcal{W}_t^k$, the cost-to-go function approximations are kept the same, *i.e.*, $\widetilde{G}_t^{k+1}(\,\cdot\,, \mathcal{W}_t) = \widetilde{G}_t^k(\,\cdot\,, \mathcal{W}_t)$.

An important property of this algorithm is that the function $\widetilde{G}_t^k(\,\cdot\,, \mathcal{W}_t)$ is an under-estimator of the expected cost-to-go function $\widetilde{G}_t(\,\cdot\,, \mathcal{W}_t)$ for each t, $\mathcal{W}_t$ and k. Hence, $F_1\big(x_1^k, \mathcal{W}_1\big) + \widetilde{G}_1^k\big(x_1^k, \mathcal{W}_1\big)$ is a lower bound of the optimal value of the nested problem (3) for stage $t = 1$. To obtain upper bounds, (5) can be solved forward for each node of the tree. These bounds can in theory be used to stop the algorithm when a desired accuracy has been reached. However, computing exact upper bounds is typically impractical since a scenario tree can have a very large number of nodes. A more tractable option is to use statistical estimates of upper bounds by solving (5) only for a relatively small number of sampled branches of the tree [33]. Nevertheless, as stated in [2], it is more common in practice to run the algorithm for a fixed number of iterations due to the difficulty of estimating appropriate convergence tolerances a priori.

3.4 Convergence

For the case of problems with linear functions and stage-wise independent uncertainty, Chen & Powell and Philpott & Guan provide convergence proofs for SDDP and related algorithms [7] [25]. For the more general case of problems having convex objective functions and potentially stage-wise dependent uncertainty, this is done by Girardeau, Leclere & Philpott [10]. The interested reader is referred to these works for more details.

4 Parameterized Stochastic Hybrid Approximation

4.1 Background and Overview

In [8], Cheung & Powell proposed a stochastic hybrid approximation procedure for solving two-stage stochastic optimization problems of the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & F(x) + \mathbb{E}\,[\,Q(x,w)\,] \\ \text{subject to} \quad & x \in \mathcal{X}, \end{aligned}$$

where $F + Q(\,\cdot\,, w)$ are convex functions, $\mathcal{X}$ is a convex compact set, and $\mathbb{E}\,[\,\cdot\,]$ denotes expectation with respect to a random vector w. The procedure consists of generating iterates

$$x^k := \underset{x \in \mathcal{X}}{\arg\min}\, F(x) + \mathcal{Q}^k(x),$$

where $\mathcal{Q}^k$ are approximations of $\mathbb{E}\,[\,Q(\,\cdot\,, w)\,]$ such that $F + \mathcal{Q}^k$ are strongly convex and differentiable. The first approximation $\mathcal{Q}^0$ is provided by the user, and the rest are obtained by performing slope corrections using noisy subgradients of $\mathbb{E}\,[\,Q(\,\cdot\,, w)\,]$ as follows:

$$\mathcal{Q}^{k+1}(x) = \mathcal{Q}^k(x) + \alpha_k\big(\xi^k - \nabla\mathcal{Q}^k\big(x^k\big)\big)^T x, \tag{6}$$

for all x, where $\xi^k \in \partial Q(x^k, w^k)$, w^k are independent samples of w, and α_k are step lengths that satisfy conditions that are common in stochastic approximation algorithms [5] [15]. The strengths of this approach are its ability to exploit an accurate initial function approximation $\mathcal{Q}^0$, and the fact that the slope corrections do not change the structure of this approximation.

In [35], Tinoco De Rubira & Hug applied this procedure to a two-stage stochastic generator dispatch problem. Motivated by the fact that in power system operations planning certainty-equivalent models are common, *i.e.*, models with random quantities replaced by their expected values, they considered using such a model for the initial function approximation, *i.e.*, $\mathcal{Q}^0 = Q(\,\cdot\,, \mathbb{E}\,[\,w\,])$. The results obtained with this approach were promising since they were better than those obtained with a stochastic gradient algorithm and two cutting-plane SAA-based methods. The same authors later combined the stochastic hybrid approximation procedure with dual stochastic gradient ascent in order to solve problems of the form

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & F(x) + \mathbb{E}\,[\,Q(x,w)\,] \\ \text{subject to} \quad & \mathbb{E}\,[\,G(x,w)\,] \leq 0 \\ & x \in \mathcal{X}, \end{aligned}$$

where $G(\,\cdot\,, w)$ are vector-valued functions composed of convex functions [36]. The proposed algorithm was tested on a risk-averse version of the two-stage stochastic generator dispatch problem

considered in [35] using initial approximations of $\mathbb{E}\,[\,Q(x,w)\,]$ and $\mathbb{E}\,[\,G(x,w)\,]$ that were also based on the certainty-equivalent problem. The results obtained were again positive, and this suggests that at least for the application of generator dispatch under uncertainty, the stochastic hybrid approximation procedure with initial approximations based on certainty-equivalent models is an effective combination.

Motivated by the results obtained by Tinoco De Rubira & Hug in [35] and [36], the stochastic hybrid procedure is extended here to solve convex multi-stage stochastic optimization problems of the form of (1). Unlike in the two-stage case, continuous uncertainty makes this problem intractable due to the presence of at least one level of nested expectations, which results in an uncountably infinite number of expected cost-to-go functions. For this reason, the aim here is to design an algorithm that can be applied directly to problems with continuous uncertainty, and that it is guaranteed to work on the theoretically-tractable case of finite-support uncertainty under some conditions that are likely to hold in practice.

The main idea of the proposed algorithm is to parameterize the stochastic hybrid approximation procedure by the exogenous random process. More specifically, the step lengths, slope corrections, and cost-to-go function approximations are treated as functions of the exogenous random process, with their values being dependent on how close a given realization of the exogenous random process is to each of the realizations drawn so far during previous iterations. Radial basis functions, which are widely used in machine learning [1] [37], are used to control how slope corrections are shared or generalized for similar realizations of uncertainty.

In the following, a detailed description of the algorithm is provided followed by a discussion of its memory requirements and convergence.

4.2 Algorithm

At each iteration $k \in \mathbb{Z}_+$, the proposed algorithm performs the following steps: First, a realization $\mathcal{W}_t^k := \{w_1^k, \ldots, w_t^k\}$ of the exogenous random process is sampled for $t = T$. Then, the problem

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + \widehat{G}_t(x_t, \mathcal{W}_t) + g_t^k(\mathcal{W}_t)^T x_t \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x_{t-1}, w_t) \end{aligned} \tag{7}$$

is solved *forward in time* for each $t = 1, \ldots, T$ with $x_{t-1} = x_{t-1}^k$ and $\mathcal{W}_t = \mathcal{W}_t^k$ to obtain a solution x_t^k. The vector $x_0^k := x_0$ is a constant and g_t^k is a slope-correction function such that $g_t^0 := 0$ for all t and $g_T^k := 0$ for all k. The function $\widehat{G}_t(\,\cdot\,, \mathcal{W}_t)$ is a user-provided convex differentiable approximation of the expected cost-to-go function $G_t(\,\cdot\,, \mathcal{W}_t)$ defined in (2) such that $\widehat{G}_T := 0$. For each $t \neq T$, after obtaining x_t^k, the vectors

$$\xi_t^k \in \partial H_{t+1}^k\big(x_t^k, \mathcal{W}_{t+1}^k\big) \tag{8}$$

$$\eta_t^k = \nabla \widehat{G}_t\big(x_t^k, \mathcal{W}_t^k\big) \tag{9}$$

are computed, where H_t^k is the function such that $H_t^k(x, \mathcal{W}_t)$ is the optimal objective value of problem (7) with $x_{t-1} = x$. These vectors are then used to update each of the slope-correction functions g_t^k as follows:

$$g_t^{k+1}(\mathcal{W}_t) = g_t^k(\mathcal{W}_t) + \alpha_t^k(\mathcal{W}_t)\left(\xi_t^k - \eta_t^k - g_t^k\big(\mathcal{W}_t^k\big)\right), \tag{10}$$

for all $t \in \{1, \ldots, T-1\}$ and $\mathcal{W}_t$. The step length functions α_t^k are defined by

$$\alpha_t^k(\mathcal{W}_t) := \phi_t^k(\mathcal{W}_t)\beta_k \tag{11}$$

for all $t \in \{1, \ldots, T-1\}$, $\mathcal{W}_t$, and $k \in \mathbb{Z}_+$, where $\beta_k \in (0,1)$, ϕ_t^k is the *radial basis function* [37] defined by

$$\phi_t^k(\mathcal{W}_t) := e^{-\gamma_t \|\mathcal{W}_t - \mathcal{W}_t^k\|_2}, \tag{12}$$

and $\gamma_t \in \mathbb{R}_{++}$ is an algorithm parameter. The role of this radial basis function is to weigh how much a slope correction obtained from a specific realization of the uncertainty, say $\mathcal{W}_t^k$, should count towards the slope correction applied to a function associated with a path $\mathcal{W}_t$ of the exogenous random process. The parameter γ_t controls how this weight decreases as the "distance" between paths of the exogenous random process increases.

4.3 Memory Requirements

From (10), the slope correction function g_t^k corresponding to stage $t \in \{1, \ldots, T-1\}$ and iteration $k \in \mathbb{Z}_+$ can be expressed

$$g_t^k(\mathcal{W}_t) := \begin{cases} \sum_{j=0}^{k-1} \alpha_t^j(\mathcal{W}_t) \left(\xi_t^j - \eta_t^j - g_t^j\left(\mathcal{W}_t^j\right) \right) & \text{if } k > 0, \\ 0 & \text{else,} \end{cases} \tag{13}$$

for all $\mathcal{W}_t$. In contrast to the two-stage case [8] [35], in which g_t^k and is a vector that gets updated at each iteration, here it is a function that determines a slope-correction vector suitable for an approximation associated with a specific path of the exogenous random process. From (13), its output depends on how similar its input is to each of the realizations of the random process sampled so far, *i.e.*, $\left\{ \mathcal{W}_t^j \;\middle|\; 0 \leq j < k \right\}$, as well as the slope-correction vectors computed so far, *i.e.*, $\left\{ \xi_t^j - \eta_t^j - g_t^j\left(\mathcal{W}_t^j\right) \;\middle|\; 0 \leq j < k \right\}$. Hence, the proposed algorithm requires storing all this information, which grows linearly with the number of iterations.

4.4 Convergence

For the case of finite-support uncertainty, for which the exogenous random process can be represented by a finite scenario tree, for specific parameters of the algorithm, namely $\gamma_t = \infty$ for each $t \in \{1, \ldots, T-1\}$, and under certain assumptions, the iterates produced by the proposed algorithm can be shown to have a subsequence that converges to the solutions of the nested problems (1). The details can be found in Appendix A. The main idea behind the proof is to show that the iterates produced by the algorithm, which are obtained using deterministic cost-to-go function approximations and are referred to as "level-0" solutions, "track" the solutions of approximations that have one level of exact expected values, or "level-1" solutions. These approximations with one level of exact expected values, say for stage t, are based on the exact expected value of the cost of stage $t+1$ but assume deterministic approximations of the cost-to-go of the future stages. Then, show inductively that level-1 solutions track level-2 solutions, level-2 solutions track level-3 solutions, and so on. For a given stage t, level-$(T-t)$ solutions are exact, and hence the iterates produced by the proposed algorithm can be shown to track these exact solutions. Aside from mathematical induction, the proof uses techniques based on those used by Cheung & Powell for the two-stage case [8], and by Tinoco De Rubira & Hug for the case with expected-value constraints [36]. These consist (respectively) of showing that level-τ solutions for each τ get closer together as the number of iterations increases, and assuming that an infinite sum of uncontrolled side effects is bounded above almost surely. One condition required for the proof of the parameterized algorithm proposed here, which is not required by the algorithms analyzed in [8] and [36], is that the cost functions $F_t(\,\cdot\,, \mathcal{W}_t)$ need to be strongly convex. However, in practice, not having this property is not likely to be an issue since regularization can be added to the problem without significantly affecting accuracy.

5 Multi-Stage Optimal Generator Dispatch

5.1 Background and Overview

In electric power systems, which are also known as power networks or grids, the operation of generators needs to be planned in advance, *e.g.*, days or hours ahead. One reason for this is that generators with low operating cost typically have limited flexibility [4] [16] [34]. In practice, the on-off states of generators as well as tentative or financially-binding production levels for every hour of the day are determined the day before. Then, adjustments to this schedule are typically made shortly before delivery, *e.g.*, an hour ahead, to ensure that the system remains balanced in an economically efficient way in the event of deviations from predicted conditions. Any last-minute disturbances and deviations are handled automatically through frequency control, with resources such as Automatic Generation Control (AGC) [21]. Until recently, uncertainty in power systems came mainly from demand forecast errors and potential fault occurrences, was relatively low, and hence deterministic models with enough reserve requirements were acceptable. However, with the projected large-scale penetration of renewable energy sources, in particular wind and solar, which are highly variable and unpredictable, uncertainty must be properly considered. If not, operating costs of the system could become excessively high and also security issues could arise [4] [34].

In this section, the performance of the algorithms described in Sections 3 and 4 is compared on a multi-stage day-ahead optimal generator dispatch problem from power systems with large-scale penetration of renewable energy such as wind and solar. The problem considered assumes that generator on-off states have been determined, and seeks to determine power production levels taking into account the limited flexibility of certain generation technologies. This problem is formulated with the goal of capturing important properties of generation scheduling in networks with large-scale penetration of renewable energy, namely uncertainty, the sequential nature of the decision-making process, and generator limitations. In practice, the techniques evaluated here could be used to construct operational policies in a receding-horizon fashion. Lastly, the applicability and performance of the algorithms described in this work on more realistic generator scheduling problems, *e.g.*, that consider unit commitment decisions or market-clearing practices, is subject of future work.

5.2 Problem Formulation

The problem considered consists of determining generator output powers for a sequence of time periods indexed by $t \in \{1, \ldots, T\}$, where $T \in \mathbb{N}$, in the presence of exogenous uncertain output powers of renewable energy sources. The power network is assumed to have $n_p \in \mathbb{N}$ generators with limited flexibility, *e.g.*, "baseload units" such as large nuclear and coal-fired power plants [4] [16], $n_q \in \mathbb{N}$ fast-ramping flexible generators, *e.g.*, "peaking units" such as gas-fired power plants [4] [16], $n_r \in \mathbb{N}$ renewable energy sources, $n_n \in \mathbb{N}$ nodes, $n_e \in \mathbb{N}$ edges, and $n_d \in \mathbb{N}$ loads. At each stage $t \in \{1, \ldots, T\}$, $p_t \in \mathbb{R}^{n_p}$ denotes output powers of generators with limited flexibility, $q_t \in \mathbb{R}^{n_q}$ denotes output powers of fast-ramping flexible generators, $r_t \in \mathbb{R}^{n_r}_+$ denotes output powers of renewable energy sources, $s_t \in \mathbb{R}^{n_r}$ denotes powers of renewable energy sources after curtailments, $\theta_t \in \mathbb{R}^{n_n}$ denotes node voltage angles, where one is treated as a reference, and $d_t \in \mathbb{R}^{n_d}$ denotes powers consumed by loads. With this notation, the problem can be expressed mathematically as the following nested optimization problem:

$$\underset{p_t, q_t, \theta_t, s_t}{\text{minimize}} \quad \varphi(p_t) + \psi(q_t) + G_t(p_t, \mathcal{R}_t) \tag{14a}$$

$$\text{subject to} \quad Pp_t + Qq_t + Ss_t - A\theta_t - Dd_t = 0 \tag{14b}$$

$$y^{\min} \leq p_t - p_{t-1} \leq y^{\max} \tag{14c}$$
$$z^{\min} \leq J\theta_t \leq z^{\max} \tag{14d}$$
$$p^{\min} \leq p_t \leq p^{\max} \tag{14e}$$
$$q^{\min} \leq q_t \leq q^{\max} \tag{14f}$$
$$0 \leq s_t \leq r_t, \tag{14g}$$

for each $t \in \{1, \ldots, T\}$, where φ and ψ are separable strongly convex quadratic functions that quantify power generation costs of generators with limited flexibility and of fast-ramping flexible generators, respectively, $\{P, Q, S, A, D, J\}$ are sparse constant matrices, $\mathcal{R}_t := \{r_1, \ldots, r_t\}$, and p_0 is a constant vector. Constraint (14b) enforces power balance at each node of the network using a DC power flow model [42]. Constraint (14c) imposes ramping constraints on the generators with limited flexibility to model their limited ability to change power levels during operations due to physical, economic, and design reasons [4] [16]. Constraint (14d) enforces edge flow limits due to thermal ratings of transmission lines and transformers, constraint (14e) enforces output power limits of generators with limited flexibility, constraint (14f) enforces output power limits of fast-ramping flexible generators, and constraint (14g) enforces limits of powers of renewable energy sources after curtailments. The cost-to-go functions $G_t(\,\cdot\,, \mathcal{R}_t)$ are defined by

$$G_t(p, \mathcal{R}_t) := \begin{cases} \mathbb{E}\,[\, H_{t+1}(p, \mathcal{R}_{t+1}) \,|\, \mathcal{R}_t \,] & \text{if } t \in \{1, \ldots, T-1\}, \\ 0 & \text{else,} \end{cases} \tag{15}$$

for all p, where the function H_t is such that $H_t(p, \mathcal{R}_t)$ is the optimal objective value of problem (14) with $p_{t-1} = p$, and $\mathbb{E}\,[\,\cdot\,|\,\mathcal{R}_t\,]$ is expectation with respect to $\mathcal{R}_{t+1}$ conditioned on $\mathcal{R}_t$.

Due to the ramping restrictions (14c) on the output powers of generators with limited flexibility, and the output power limits (14e) and (14f), the nested problem (14) may not in general have the extended relatively complete recourse property described in Section 2. However, this can be typically ensured by modifying the problem slightly to allow for emergency load shedding, at an appropriately high cost, or demand response.

5.3 Exogenous Random Process

The exogenous random output powers of renewable energy sources are assumed to be given by

$$r_t = \Pi_c\big(\Pi_\delta(\bar{r}_t + \delta_t, r_{t-1})\big), \tag{16}$$

for each $t \in \{1, \ldots, T\}$, where $\bar{r}_t \in \mathbb{R}^{n_r}_+$ are constant vectors, and $\delta_t \sim \mathcal{N}(0, \Sigma_t)$. Π_c is the projection operator given by

$$\Pi_c(z) := \arg\min \big\{\, \|r - z\|_2 \;\big|\; 0 \leq r \leq r^{\max} \,\big\},$$

for all z, where $r^{\max} \in \mathbb{R}^{n_r}_{++}$ are the power capacities of the renewable energy sources. Π_δ is a randomized operator given by

$$\Pi_\delta(z_1, z_2) := \begin{cases} \arg\min \big\{\, \|r - z_1\|_2 \;\big|\; -\delta^{\max} \leq r - z_2 \leq \delta^{\max} \,\big\} & \text{with probability } 1 - \epsilon, \\ z_1 & \text{with probability } \epsilon, \end{cases} \tag{17}$$

for all (z_1, z_2), where $\delta^{\max} \in \mathbb{R}^{n_r}_{++}$ is a constant vector. Lastly, for $t \in \{1, \ldots, T\}$, the covariance matrix Σ_t is assumed to satisfy $\Sigma_t = (t-1)\Sigma/(T-1)$, for $T > 1$, where Σ is a positive definite matrix. This captures the fact that the output powers of renewable energy sources are known at time $t = 1$, and that the uncertainty increases with t.

The use of Gaussian random variables for modeling renewable energy uncertainty or forecast errors is common in the power systems literature [24] [26] [28]. The model considered here adds Gaussian noise δ_t to "base" power levels $\bar{r}_t$. The projection operator Π_c ensures that the resulting powers are non-negative and not greater than the capacities of the sources. The randomized operator Π_δ guarantees with probability $1-\epsilon$ that the resulting powers have ramps bounded by a given $\delta^{\max}$. This is used to prevent having large ramps that occur unrealistically often in the planning horizon. Lastly, as is done in [35] and [36], spatial correlation between "nearby" renewable energy sources is considered here through a non-diagonal matrix Σ parameterized by a correlation coefficient ρ and a "correlation distance" N.

5.4 Application of Proposed Algorithm

In order to apply the parameterized stochastic hybrid approximation algorithm described in Section 4.2 to the nested problem (14), initial expected cost-to-go function approximations are required. Motivated by the results obtained by Tinoco De Rubira & Hug in [35] and [36], the approximations considered here consist of

$$\widehat{G}_t(p,\mathcal{R}_t) := \begin{cases} H^0_{t+1}(p, \mathbb{E}\left[\,\mathcal{R}_{t+1}\,|\,\mathcal{R}_t\,\right]) & \text{if } t \in \{1,\ldots,T-1\}, \\ 0 & \text{else,} \end{cases} \tag{18}$$

for all p, for each $t \in \{1,\ldots,T\}$ and $\mathcal{R}_t$, where the function H^k_t here is such that $H^k_t(p,\mathcal{R}_t)$ is the optimal objective value of problem (7) with $p_{t-1}=p$. An important observation is that this choice of approximation is not necessarily differentiable everywhere due to the inequality constraints in (14). However, as will be seen in Section 5.6, this does not cause problems possibly due to the fact that the lack of differentiability only occurs at certain points. For instances of problem (7) on which issues do arise due to this lack of differentiability everywhere, one may consider using instead of $H^0_t(p,\mathcal{R}_t)$ an approximation that replaces inequality constraints with barrier terms in the objective.

For this choice of initial expected cost-to-go function approximations, the subproblems (7) solved by the proposed algorithm can be expressed as

$$\underset{\{p_\tau,q_\tau,\theta_\tau,s_\tau\}_{\tau=t}^T}{\text{minimize}} \quad \varphi(p_t)+\psi(q_t)+g_t^m\big(\mathcal{R}_t\big)^T p_t \,+\sum_{\tau=t+1}^{T}\Big(\varphi(p_\tau)+\psi(q_\tau)+g_\tau^0\big(\widehat{\mathcal{R}}_{t,\tau}\big)^T p_\tau\Big) \tag{19a}$$

$$\text{subject to} \quad (p_t,q_t,\theta_t,s_t) \in \mathcal{K}_t\big(p_{t-1},r_t\big) \tag{19b}$$

$$(p_\tau,q_\tau,\theta_\tau,s_\tau) \in \mathcal{K}_\tau\big(p_{\tau-1},\hat{r}_\tau\big),\ \tau=t+1,\ldots,T, \tag{19c}$$

for each $t \in \{1,\ldots,T\}$ and $k \in \mathbb{Z}_+$ with $p_{t-1}=p^k_{t-1}$, $\mathcal{R}_t=\mathcal{R}^k_t$, and $m=k$. Here, p^k_{t-1} denotes the solution of the previous-stage subproblem, $\mathcal{R}^k_t := \big\{r^k_1,\ldots,r^k_t\big\}$ are the sampled realizations of renewable powers up to stage t during iteration k, constraints (19b) and (19c) correspond to constraints (14b)-(14g) for $\tau \in \{t,\ldots,T\}$, and $\widehat{\mathcal{R}}_{t,\tau} := \big\{\hat{r}_1,\ldots,\hat{r}_\tau\big\}$ is defined recursively as

$$\widehat{\mathcal{R}}_{t,\tau} = \begin{cases} \mathbb{E}\left[\,\mathcal{R}_\tau \,\middle|\, \widehat{\mathcal{R}}_{t,\tau-1}\,\right] & \text{if } \tau > t, \\ \mathcal{R}_t & \text{else,} \end{cases} \tag{20}$$

for each $\tau \in \{t,\ldots,T\}$. From the properties of φ, ψ and $\mathcal{K}_\tau$, these optimization problems are deterministic multi-stage convex Quadratic Programming (QP) problems.

From (8), the vector ξ^k_t can be constructed from the dual solution of subproblem (19) for stage $t+1$ with $p_t=p^k_t$, $\mathcal{R}_{t+1}=\mathcal{R}^k_{t+1}$, and $m=k$, which is solved during the "forward pass" of

the algorithm. More specifically, it can be shown that $\xi_t^k = -\mu_{t+1}^k + \pi_{t+1}^k$, where μ_{t+1}^k and π_{t+1}^k denote the Lagrange multipliers associated with the upper and lower ramping constraints (14c), respectively. Similarly, from (9) and (18), the vector η_t^k can be constructed (estimated) using an analogous formula from the dual solution of subproblem (19) for stage $t+1$ with $p_t = p_t^k$, $\mathcal{R}_{t+1} = \mathbb{E}\left[\mathcal{R}_{t+1} \,\middle|\, \mathcal{R}_t^k\right]$, and $m = 0$. This subproblem is also solved during the "forward pass" of the algorithm but only implicitly, *i.e.*, as part of the solution of another subproblem. In particular, its primal solution is obtained during this process, and this can be exploited when computing the dual solution.

As mentioned above, the proposed initial function approximations (18) result in the algorithm having to solve the deterministic multi-stage subproblems (19) during the forward pass. More specifically, at each iteration, the algorithm has to solve a subproblem with $T - \tau + 1$ stages for each $\tau \in \{1, \ldots, T\}$. Hence, this choice of initial function approximations is not suitable for problems with a large number of stages since the computational requirements of the algorithm would be high. For such problems, a "depth-limited" modification of (18) that ignores the cost-go-to after a certain number of stages may be considered. The investigation of this idea will be subject of future work.

5.5 Implementation

The proposed and benchmark algorithms as well as the multi-stage optimal generator dispatch problem were implemented using the Python programming language. More specifically, the SDDP and parameterized stochastic hybrid approximation algorithms were implemented within the Python package `OPTALG` [2] using the scientific-computing packages `Numpy` and `Scipy` [14] [38]. The multi-stage optimal generator dispatch problem was implemented within the Python package `GRIDOPT` [3] using the modeling library `PFNET` [4]. For solving problems (19), the sparse interior-point QP solver IQP of `OPTALG` was used. For obtaining $\Sigma^{1/2}$, where Σ is the spatial covariance matrix defined in Section 5.3, the sparse Cholesky code `CHOLMOD` was used via its Python wrapper [5]. Lastly, to construct a suitable scenario tree for SDDP, the clustering routines from the Python package `scikit-learn` [22] were used.

5.6 Numerical Experiments

The proposed and benchmark algorithms were applied to several instances of the multi-stage optimal generator dispatch problem in order to compare their performance. The performance of these two algorithms was also compared against that of a greedy and a certainty-equivalent algorithm. The greedy algorithm ignores the expected cost-to-go for subsequent stages at each stage, and hence is equivalent to the SDDP algorithm executed for 0 iterations. This follows from the fact that $\widetilde{G}_t^0 = 0$ for all t, as stated in Section 3.3. On the other hand, the certainty-equivalent algorithm solves the problem by replacing expected costs by costs associated with expected realizations of the uncertainty. Due to the particular choice (18) of initial function approximations, the certainty-equivalent algorithm is equivalent to the parameterized stochastic hybrid approximation algorithm before any slope corrections are made, *i.e.*, executed for 0 iterations.

[2] http://optalg.readthedocs.io/
[3] http://gridopt.readthedocs.io/
[4] http://pfnet.readthedocs.io/
[5] http://pythonhosted.org/scikits.sparse/

5.6.1 Test Cases

The power networks used to construct instances of the multi-stage optimal generator dispatch problem consisted of the small `IEEE14` network [6] and two mid-scale networks from North America and Europe. For all of these networks, a 24-hour scheduling period divided into 8 stages was considered. Table 1 shows the properties of the test cases constructed. The labels of the columns correspond to the quantities defined in Section 5.2. In practice, higher resolutions, *e.g.*, hourly, are typically used for generator scheduling but with deterministic models. Multi-stage stochastic generator scheduling problems with such resolution will be considered in future work along with a suitable choice of initial function approximations, such as those suggested at the end of Section 5.4.

Table 1: Test Cases.

name	n_n	n_e	n_p	n_q	n_r	n_d	T
Case A	14	20	5	5	5	11	8
Case B	2454	2781	211	211	236	1149	8
Case C	3120	3693	278	278	248	2314	8

The generators with limited flexibility were taken as the adjustable generators originally present in the network data. Due to the absence of cost data, the separable cost function φ for these generators was constructed using uniformly random coefficients in $[0.01, 0.05]$ for the quadratic terms, and uniformly random coefficients in $[10, 50]$ for the linear terms. These coefficients assume that power quantities are in units of MW, and are consistent with those found in several `MATPOWER` cases [43]. For these generators, the maximum ramping rates were set to 1% of their power capacities per hour in order to model their limited ability to change power levels during operations due to physical, economic, and design reasons. The fast-ramping flexible generators were new generators added to the networks in the same locations as the generators with limited flexibility. The separable cost function ψ for these generators was set to satisfy $\psi = 10\varphi$. This cost difference is consistent with the difference in marginal costs between baseload and peaking generators analyzed in [41]. Lastly, the known constant output powers of generators with limited flexibility before stage 1, *i.e.*, p_0, were set to the least-cost powers that balanced the network for stage 1.

Three daily load profiles were obtained from a North American power marketing administration. Each load profile was normalized and used to modulate all the (static) loads of one of the power networks. Figure 1 shows the resulting aggregate load profiles for each of the test cases.

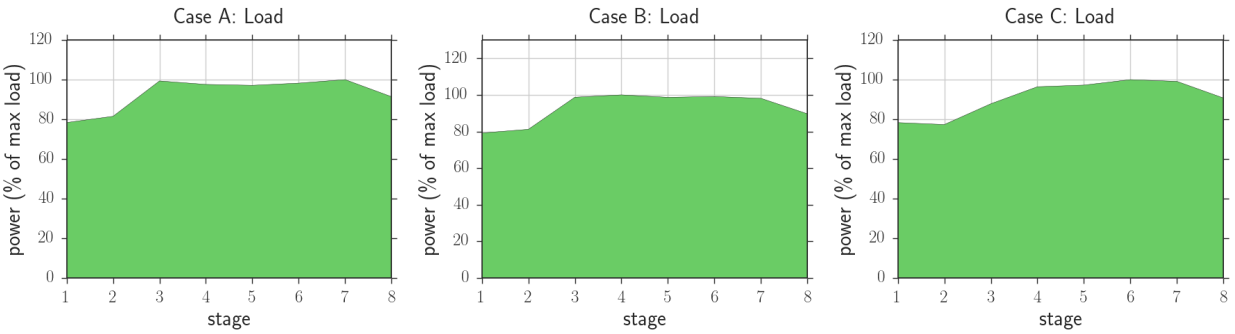

Figure 1: Load profiles.

[6] http://www2.ee.washington.edu/research/pstca/

The renewable energy sources used to construct the test cases of Table 1 were added manually to each of the power networks at the nodes with generators. The capacity $r_i^{\max}$ of each source was set to $\max_t 1^T d_t/n_r$, *i.e.*, to the maximum aggregate load divided by the number of renewable energy sources. The base power profiles $\{\bar{r}_t\}$ were obtained by multiplying $0.5r^{\max}$ by a normalized daily wind power profile obtained from the same North American power marketing administration from which the load data was obtained. This resulted in a maximum base renewable energy penetration of 50% of the maximum aggregate load, which is a high penetration scenario. A different normalized wind power profile was used for each network. For the covariance matrix Σ, its diagonal was set to the element-wise square of $r^{\max}$, and its off-diagonal elements were set in such a way that powers from sources connected to nodes that are at most $N = 10$ edges ways had a correlation coefficient of $\rho = 0.1$, and zero otherwise. For these sources, the maximum ramping rates were set to 30% of their power capacities per hour, and these limits were enforced with a probability of 0.7, *i.e.*, $\epsilon = 0.3$ in (17). Figure 2 shows sample realizations of the aggregate powers from the renewable energy sources for each of the test cases. An important observation is that the aggregate power from renewable sources is much more variable for Case A than for the other cases. The reason for this is that Cases B and C have many more renewable energy sources that are uncorrelated.

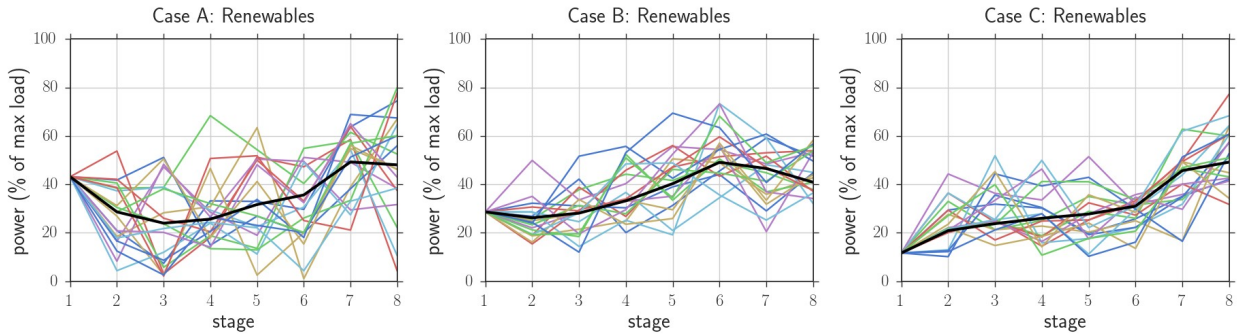

Figure 2: Sample realizations of powers from renewable energy sources.

5.6.2 SDDP Validation

In order to validate the implementation of the SDDP algorithm, convergence to zero of the difference between upper and lower bounds computed at every iteration was checked. This was done for the smallest case, namely, Case A, using a small scenario tree obtained using Monte Carlo sampling with branching factors given by $(4, 3, 2, 1, 1, 1, 1)$, which gives 137 nodes and 24 scenarios. The upper and lower bounds were obtained at each iteration $k \in \mathbb{Z}_+$ of the algorithm by evaluating $F_1(x_1^k, w_1) + \widetilde{G}_1(x_1^k, \mathcal{W}_1)$ and $F_1(x_1^k, w_1) + \widetilde{G}_1^k(x_1^k, \mathcal{W}_1)$, respectively. The results obtained are shown in Figure 3.

5.6.3 Algorithm Comparison

The algorithms were executed for various numbers of iterations and the policies obtained were evaluated. More specifically, the parameterized stochastic hybrid approximation was executed for 0, 25, 50 and 100 iterations. As already noted, for 0 iterations the policy obtained with this algorithm is the same as the obtained from solving certainty-equivalent problems at each stage. For the radial basis functions used by the algorithm, the parameter γ_t was set to $\gamma(n_r t)^{-1/2}$ for each $t \in \{1, \ldots, T\}$, where $\gamma = 1$, and n_r is the number of renewable energy sources. On the

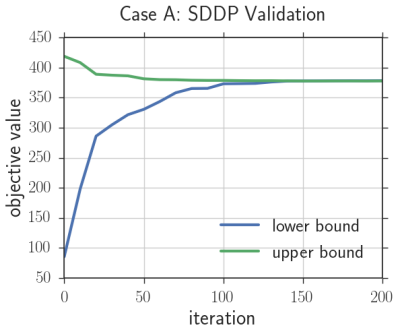

Figure 3: SDDP validation.

other hand, the SDDP algorithm was executed for 0, 25, 100 and 400 iterations. As already noted, for 0 iterations the policy obtained with this algorithm is the same as the greedy policy obtained by solving problem (5) at each stage ignoring the cost-to-go function approximation. The scenario trees used for this algorithm were constructed using conditional Monte Carlo sampling and K-means clustering at each stage. Information about these scenario trees including their construction time is shown in Table 2. To evaluate the policies produced by the algorithms, the policies were simulated using 500 sampled realizations of the exogenous random process. The average total costs obtained are shown Figure 4 as a function of the execution time of the algorithm, where GR denotes greedy, CE denotes certainty-equivalent, and SHA denotes stochastic hybrid approximation.

Table 2: Scenario trees for SDDP.

case	stages	samples/stage	clusters/stage	nodes	scenarios	time (min)
Case A	8	1000	3	3280	2187	1.65
Case B	8	1000	3	3280	2187	7.02
Case C	8	1000	3	3280	2187	7.26

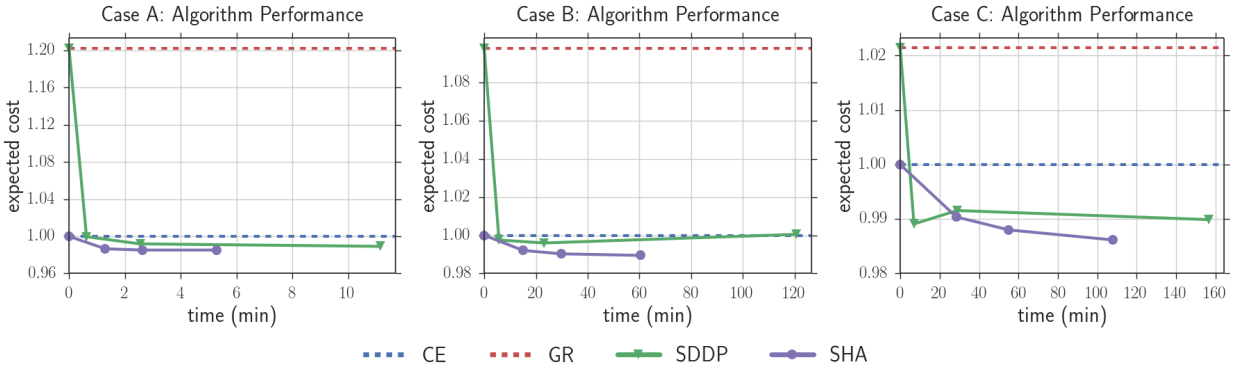

Figure 4: Algorithm performance.

As the plots of Figure 4 suggest, the proposed algorithm is able to exploit the accuracy of the initial expected cost-to-go function approximations provided, which in this case are the certainty-equivalent models, and quickly generate policies that achieve lower expected costs compared to the

ones obtained using SDDP. Another important observation is that for Cases A and B, the greedy model used by SDDP at the beginning of its execution can be much worse than the certainty-equivalent model. For Case B, the expected cost associated with the SDDP policy obtained after 400 iterations is slightly worse than with the one obtained after 100 iterations. This phenomenon may be attributed to a slight overfitting of the policy to the scenario tree or to the small errors obtained by computing expected costs using a finite number of sampled realizations. We note here that the execution times of SDDP shown in these plots do not include the times spent constructing the scenario trees, whose details are provided in Table 2.

In addition to the average total costs obtained by simulating the policies obtained with the algorithms, the average aggregate power generation profiles were also plotted. This was done to obtain a simple visualization of the strategy used by each policy. The policies considered were those obtained with the proposed algorithm after 0 iterations, *i.e.*, certainty-equivalent, and 100 iterations, and with SDDP after 0 iterations, *i.e.*, greedy, and 400 iterations. The results obtained are shown in Figure 5, where again GR denotes greedy, CE denotes certainty-equivalent, and SHA denotes stochastic hybrid approximation. From the plots, it can be seen that the greedy policy starts by using all the renewable energy available, and has in general less renewable energy left on average during all stages. The other three policies are visually similar but the ones obtained with the proposed algorithm after 100 iterations and with SDDP after 400 iterations have in general a smaller "red area" than the certainty-equivalent policy. This corresponds to using less energy on average from fast-ramping flexible generators, which have a higher operating cost. In other words, the policies obtained using stochastic optimization exploit better the limited flexibility of the resources that have lower operating cost.

5.6.4 Generalization Analysis

An experiment was also performed to determine the effects of the choice of generalization parameters γ_t used by the radial basis functions on the performance of the parameterized stochastic hybrid approximation algorithm. This experiment consisted of executing the algorithm for 0, 25, 50, and 100 iterations, as in the previous experiment, using $\gamma_t = \gamma(n_r t)^{-1/2}$ for each $t \in \{1, \ldots, T\}$ and $\gamma \in \left\{10^{-2}, 10^{-1}, 10^{0}, 10^{1}, 10^{2}\right\}$. The results obtained are shown in Figure 6.

As the plots of Figure 6 show, for large values of γ, *i.e.*, for little or no generalization, the performance of the algorithm is similar to that of the certainty-equivalent algorithm. This is expected since for γ large, the same path of the random process needs to be sampled again in order to see the effects of a slope correction. Another important and perhaps surprising observation is that for all $\gamma \in \left\{10^{-2}, 10^{-1}, 10^{0}\right\}$, the performance of the algorithm is equally promising. This shows that the algorithm is benefiting similarly from generalizing slope corrections to neighboring paths of the random process for a wide range of neighborhood "sizes", making the tuning of the generalization parameters less difficult.

6 Conclusions

In this work, a new algorithm for solving convex multi-stage stochastic optimization problems has been presented. The algorithm is based on a parameterized extension of a stochastic hybrid approximation procedure that allows it to exploit accurate initial expected cost-to-go function approximations provided by the user. The algorithm can be applied directly to problems with continuous uncertainty since it uses radial basis functions for generalizing information learned from specific realizations of the random process. An analysis of the convergence of the algorithm has been provided for the tractable case of finite-support uncertainty. In particular, conditions have

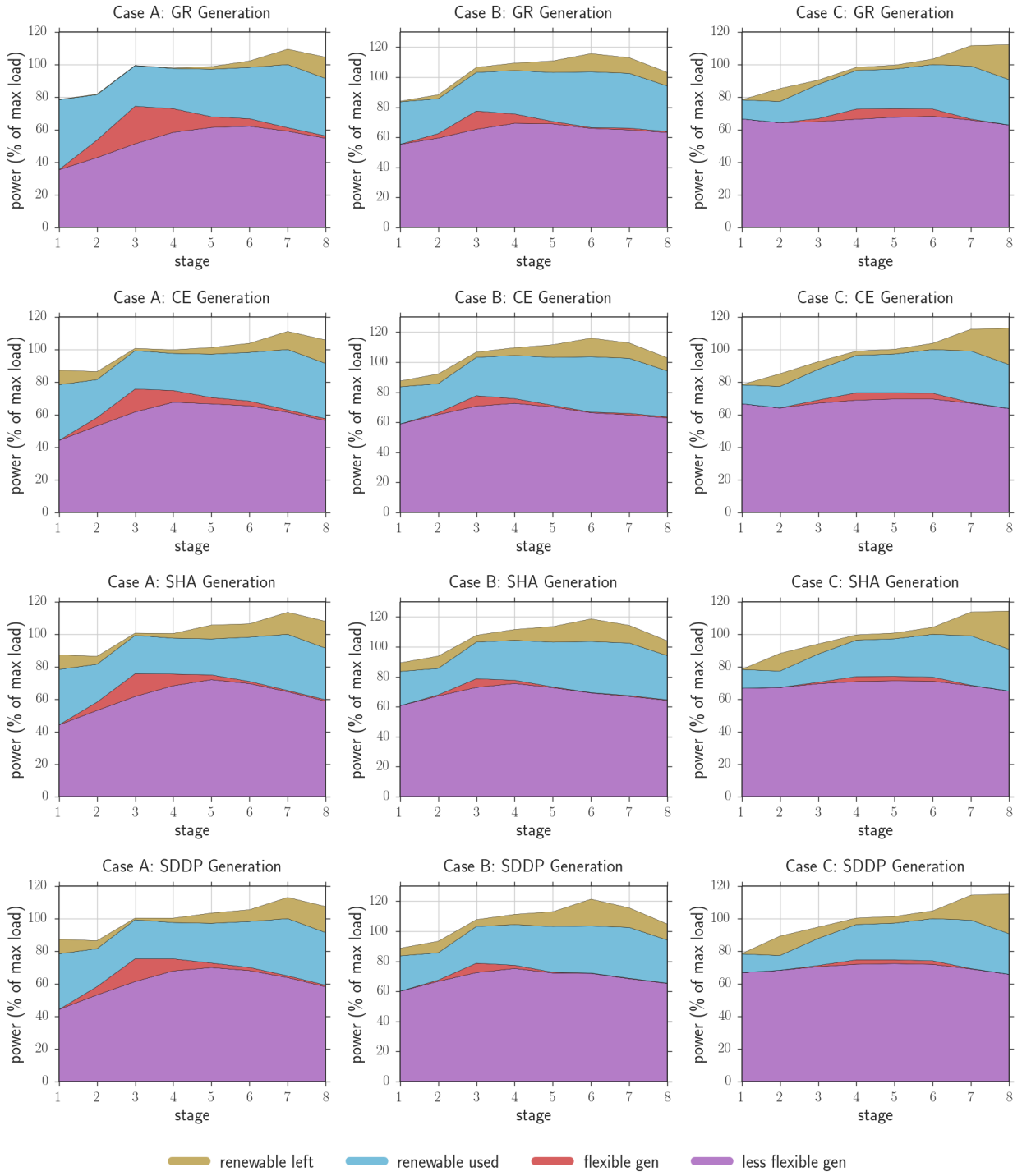

Figure 5: Generation profiles.

been identified that guarantee that the iterates produced by the algorithm have a subsequence that convergences to the solution of the problem. The performance of the algorithm has been compared against that of a greedy, a certainty-equivalent, and an SDDP algorithm on a multi-stage optimal generator dispatch problem from power system operations planning under high penetration of

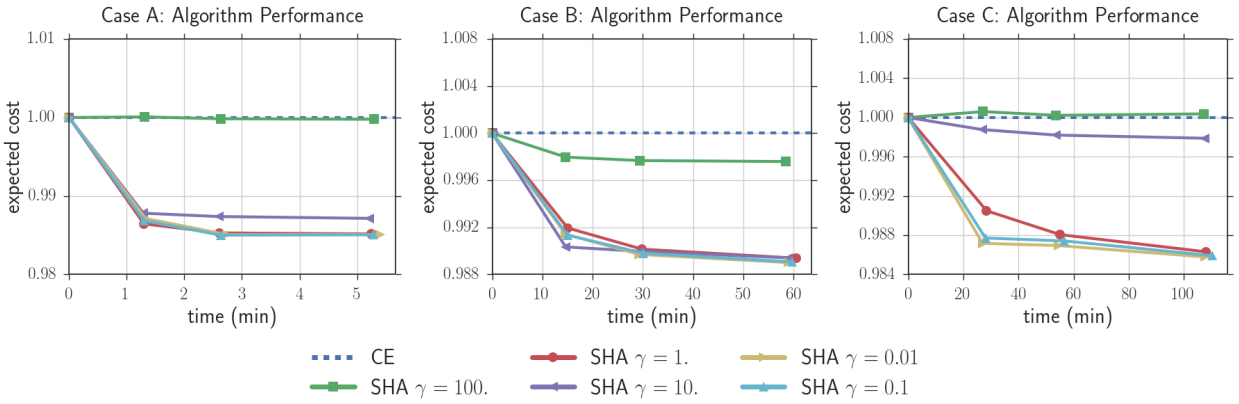

Figure 6: Generalization analysis.

renewable energy. The results obtained showed that for this problem the proposed algorithm was able to exploit the accuracy of the initial approximations of expected cost-to-go functions and outperform the benchmark algorithms.

A future research direction consists of exploring "depth-limited" initial function approximations based on certainty-equivalent models to handle problems with a large number of stages. Other directions include extending the algorithm to handle discrete variables in the first stage and non-convex constraints. These last two research directions are motivated from the fact that more accurate generator dispatch problems need to consider AC network models as well as generator properties such as start-up costs and minimum up and down times.

A Convergence Analysis

In this section, a convergence analysis of the parameterized stochastic hybrid approximation algorithm described in Section 4.2 is presented for the theoretically tractable case of finite-support uncertainty. In particular, conditions are identified that guarantee that the iterates produced by the algorithm have a subsequence that converges to the solutions of the nested optimization problems (1).

This section is structured as follows: First, key functions and notation needed for the analysis are introduced. Then, the assumptions under which convergence is analyzed are described followed by a discussion regarding their justification and implications. Lastly, Lemmas, Corollaries and Theorems are presented that gradually derive the theoretical results and properties of the algorithm.

For $t \in \{1, \ldots, T\}$ and $k \in \mathbb{Z}_+$, let $H^k_{t,0}(x, \mathcal{W}_t)$ and $\mathcal{H}^k_{t,0}(x, \mathcal{W}_t)$ be the optimal value and an optimal point of

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + \widehat{G}^k_t(x_t, \mathcal{W}_t) \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x, \mathcal{W}_t), \end{aligned}$$

respectively, where the function $\widehat{G}^k_t(\,\cdot\,, \mathcal{W}_t)$ is defined by

$$\widehat{G}^k_t(x, \mathcal{W}_t) := \widehat{G}_t(x, \mathcal{W}_t) + g^k_t(\mathcal{W}_t)^T x, \ \forall x. \tag{21}$$

For $\tau \in \{1, \ldots, T-t\}$, let $H^k_{t,\tau}(x, \mathcal{W}_t)$ and $\mathcal{H}^k_{t,\tau}(x, \mathcal{W}_t)$ be the optimal value and an optimal point of

$$\begin{aligned} \underset{x_t}{\text{minimize}} \quad & F_t(x_t, w_t) + \mathbb{E}\left[H^k_{t+1,\tau-1}(x_t, \mathcal{W}_{t+1}) \,\middle|\, \mathcal{W}_t \right] \\ \text{subject to} \quad & x_t \in \mathcal{X}_t(x, \mathcal{W}_t), \end{aligned}$$

respectively. Hence, the parameter τ denotes the number of exact expected values present in the nested problem, or the "depth" of the approximation. It follows that for $t \in \{1, \ldots, T\}$ and $\tau = T - t$, $H^k_{t,\tau}(x, \mathcal{W}_t)$ and $\mathcal{H}^k_{t,\tau}(x, \mathcal{W}_t)$ are independent of k and are equal to the optimal value and an optimal point of problem (1), respectively.

Also, let $\widehat{\mathcal{H}}^k_{t,\tau}(x, \mathcal{W}_t)$ be an optimal point of the "expanded" problem

$$\begin{aligned} \underset{y}{\text{minimize}} \quad & L^k_{t,\tau}(y, \mathcal{W}_t) \\ \text{subject to} \quad & y \in \mathcal{Y}_t(x, w_t), \end{aligned}$$

where the objective function $L^k_{t,\tau}(\,\cdot\,, \mathcal{W}_t)$ is given by

$$\begin{aligned} L^k_{t,\tau}(y, \mathcal{W}_t) = \; & F_t(y_t(w_t), w_t) + \\ & \mathbb{E}\Big[F_{t+1}(y_{t+1}(w_{t+1}), w_{t+1}) + \cdots \\ & \mathbb{E}\Big[F_{t+\tau}(y_{t+\tau}(w_{t+\tau}), w_{t+\tau}) + \widehat{G}^k_{t+\tau}(y_{t+\tau}(w_{t+\tau}), \mathcal{W}_{t+\tau}) \Big| \mathcal{W}_{t+\tau-1} \Big] \cdots \Big| \mathcal{W}_t \Big], \end{aligned} \tag{22}$$

$$y = \left\{ \, y_\varsigma(w_\varsigma) \;\middle|\; \forall \varsigma \in \{t, \ldots, t+\tau\}, \forall w_\varsigma \in \Omega_\varsigma \, \right\},$$

and the constraint $y \in \mathcal{Y}_t(x, w_t)$ enforces $y_t(w_t) \in \mathcal{X}_t(x, w_t)$, $y_{t+1}(w_{t+1}) \in \mathcal{X}_{t+1}(y_t(w_t), w_{t+1})$ and so on. For convenience, let

$$y^k_{t,\tau} := \widehat{\mathcal{H}}^k_{t,\tau}\big(x^k_{t-1}, \mathcal{W}^k_t\big), \tag{23}$$

where x_t^k denotes the iterate produced by the algorithm for stage t and sample path $\mathcal{W}_t^k$ during iteration k. An important relationship between x_t^k and $y_{t,\tau}^k$ is that $x_t^k = y_{t,0}^k$ for all $t \in \{1,\dots,T\}$ and $k \in \mathbb{Z}_+$. For each $t \in \{1,\dots,T\}$, $\tau \in \{0,\dots,T-t\}$ and $k \in \mathbb{Z}_+$, $y_{t,\tau}^k$ is referred to as the "level-τ" solution for the path $\mathcal{W}_t^k$ sampled during iteration k. Using this terminology, for each stage $t \in \{1,\dots,T\}$, level-0 solutions are the ones produced by the algorithm while level-$(T-t)$ are exact solutions of the corresponding nested problem (1).

To prevent the notation from becoming more complicated, expressions of the form $\widehat{\mathcal{H}}_{t,\tau_1}^{k_1}(x,\mathcal{W}_t) - \widehat{\mathcal{H}}_{t,\tau_2}^{k_2}(x,\mathcal{W}_t)$ for $\tau_1 \neq \tau_2$ are assumed to be valid despite the difference in dimension. In this case the subtraction is assumed to over the vector of smallest size. Similarly, expressions of the form $L_{t,\tau_1}^{k_1}\big(\widehat{\mathcal{H}}_{t,\tau_2}^{k_2}(x,\mathcal{W}_t),\mathcal{W}_t\big)$ for $\tau_1 \neq \tau_2$ are also assumed to be valid. In this case, the level of expansion is taken to be the smallest of the two. Lastly, the notation $\mathcal{O}(a_k)$ is used to express that the elements of a sequence of scalars indexed by k are uniformly bounded by a constant multiple of a_k over all $k \in \mathbb{Z}_+$.

In addition to the assumptions made in Section 2 about the properties of problem (1), the following assumptions are also made in order to analyze the convergence of the algorithm:

Assumption 1 (Strong Convexity)**.** *The function $F_t(\,\cdot\,, w_t)$ is strongly convex for each $t \in \{1,\dots,T\}$ and w_t.*

Assumption 2 (Uncertainty)**.** *The exogenous random process $\{w_t\}_{t=1}^T$ has finite support and hence can be represented by a finite scenario tree. The children nodes of each node of the scenario tree are assumed to occur with equal probability.*

Assumption 3 (Parameters)**.** *The algorithm parameter γ_t used in the radial basis function* (12) *satisfies $\gamma_t = \infty$ for all $t \in \{1,\dots,T\}$.*

Assumption 4 (Initial Function Approximations)**.** *The initial function approximation $\widehat{G}_t(\,\cdot\,,\mathcal{W}_t)$ is continuously differentiable and convex for all $t \in \{1,\dots,T-1\}$ and $\mathcal{W}_t$.*

Assumption 5 (Step Lengths)**.** *The step lengths β_k satisfy $\beta_k \in (0,1)$ for all $k \in \mathbb{Z}_+$, $\sum_{k=0}^{\infty}\beta_k = \infty$ almost surely, and $\sum_{k=0}^{\infty}\mathbb{E}\left[\,\beta_k^2\,\right] < \infty$.*

Assumption 6 (Sampled Subgradients)**.** *The sampled subgradients ξ_t^k defined in* (8) *have norms that are bounded above by a constant multiple of $1 + \big\|g_{t+1}^k\big(\mathcal{W}_{t+1}^k\big)\big\|_2$ for all $t \in \{1,\dots,T-1\}$ and $k \in \mathbb{Z}_+$.*

Assumption 7 (Neighbor Subgradients)**.** *For each $t \in \{1,\dots,T\}$ and $\tau \in \{0,\dots,T-t\}$, there exists an M such that for each $k \in \mathbb{Z}_+$ there is a $\bar{\xi}_{t+\tau}^k \in \partial H_{t+\tau+1}^k\big(z_{t,\tau}^k\big(\mathcal{W}_{t+\tau}^k\big),\mathcal{W}_{t+\tau+1}^k\big)$ such that*

$$\big\|\bar{\xi}_{t+\tau}^k - \xi_{t+\tau}^k\big\|_2 \leq M\big\|x_{t+\tau}^k - z_{t,\tau}^k\big(\mathcal{W}_{t+\tau}^k\big)\big\|_2,$$

where $z_{t,\tau}^k\big(\mathcal{W}_{t+\tau}^k\big)$ denotes the component of $y_{t,\tau}^k$ associated with stage $t+\tau$ and sample path $\mathcal{W}_{t+\tau}^k$.

Assumption 8 (Inter-Stage Drift)**.** *For each $t \in \{1,\dots,T\}$ and $\tau \in \{0,\dots,T-t\}$, if the quantities $\big\|x_{t-1}^{k+1} - x_{t-1}^k\big\|_2$ and $\big\|\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^k,\mathcal{W}_t^k\big) - \widehat{\mathcal{H}}_{t,\tau+1}^k\big(x_{t-1}^k,\mathcal{W}_t^k\big)\big\|_2$ are $\mathcal{O}(\beta_k)$, then the sequences of partial sums $\sum_{k=0}^K \Gamma_{t,\tau}^k$, $\sum_{k=0}^K \Upsilon_{t,\tau}^k$, and $\sum_{k=0}^K \Psi_{t,\tau}^k$ are almost-surely bounded above, where*

$$\Gamma_{t,\tau}^k := L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^{k+1},\mathcal{W}_t^k\big),\mathcal{W}_t^k\big) - L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^k,\mathcal{W}_t^k\big),\mathcal{W}_t^k\big) \tag{24}$$

$$\Upsilon_{t,\tau}^k := L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^k,\mathcal{W}_t^k\big),\mathcal{W}_t^k\big) - L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau+1}^k\big(x_{t-1}^k,\mathcal{W}_t^k\big),\mathcal{W}_t^k\big) \tag{25}$$

$$\Psi_{t,\tau}^k := L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^k,\mathcal{W}_t^k\big),\mathcal{W}_t^k\big) - L_{t,\tau}^k\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^{k+1},\mathcal{W}_t^k\big),\mathcal{W}_t^k\big). \tag{26}$$

Assumption 9 (Slope Drift)**.** *For each $t \in \{1,\dots,T\}$ and $\tau \in \{0,\dots,T-t\}$, if the sums $\sum_{k=0}^{\infty}\beta_k\|\bar{\xi}_{t+\tau}^k - \xi_{t+\tau}^k\|_2^2$ and $\sum_{k=0}^{\infty}\beta_k\|\nabla\widehat{G}_{t+\tau}^k(z_{t,\tau}^k(\mathcal{W}_{t+\tau}^k),\mathcal{W}_{t+\tau}^k) - \nabla\widehat{G}_{t+\tau}^k(x_{t+\tau}^k,\mathcal{W}_{t+\tau}^k)\|_2^2$ are finite almost surely, then the sequences of partial sums $\sum_{k=0}^{K}\Theta_{t,\tau}^k$ and $\sum_{k=0}^{K}\Xi_{t,\tau}^k$ are almost-surely bounded above, where*

$$\Theta_{t,\tau}^k := p_{t,\tau}^k\beta_k\big(\xi_{t+\tau}^k - \bar{\xi}_{t+\tau}^k\big)^T\Delta z_{t,\tau}^k \tag{27}$$

$$\Xi_{t,\tau}^k := p_{t,\tau}^k\beta_k\Big(\nabla\widehat{G}_{t+\tau}^k(z_{t,\tau}^k(\mathcal{W}_{t+\tau}^k),\mathcal{W}_{t+\tau}^k) - \nabla\widehat{G}_{t+\tau}^k(x_{t+\tau}^k,\mathcal{W}_{t+\tau}^k)\Big)^T\Delta z_{t,\tau}^k, \tag{28}$$

$p_{t,\tau}^k$ denotes the probability that $\mathcal{W}_{t+\tau} = \mathcal{W}_{t+\tau}^k$ given $\mathcal{W}_t^k$, and $\Delta z_{t,\tau}^k := z_{t,\tau+1}^k(\mathcal{W}_{t+\tau}^k) - z_{t,\tau}^k(\mathcal{W}_{t+\tau}^k)$.

Assumption 1 is a strong assumption that is not present in the convergence analyses of stochastic hybrid approximation algorithms for two-stage stochastic problems [8] and problems with expected-value constraints [36]. It is needed here in order to obtain that the "expanded" objective functions $L_{t,\tau}^k(\,\cdot\,,\mathcal{W}_t)$ are strongly convex, and that the "expanded" iterates $y_{t,\tau}^k$ defined in (23) get closer together as k increases. In practice, for problems of the form of (1) for which $F_t(\,\cdot\,,w_t)$ is not strongly convex, small regularization terms can be added to the objective without affecting the solution quality significantly.

From Assumption 2, the exogenous random process can be represented by a finite scenario tree. This tree can be characterized by some functions $\mathcal{N}$, $\mathcal{C}$ and $\mathcal{B}$ that give the nodes for each stage, the child nodes of each node, and the branches of a given length, respectively. The finiteness of the tree gives that for each $t \in \{1,\dots,T\}$, there is a finite number of cost-to-go functions $G_t(\,\cdot\,,\mathcal{W}_t)$ (see equation (2)) parameterized by $\mathcal{W}_t \in \mathcal{B}(t)$ that need to be approximated. Since in a scenario tree, a t-length branch $\mathcal{W}_t = \{w_1,\dots,w_t\}$ can be uniquely associated with its stage-t node w_t, the notation is simplified in the remaining part of this section and w_t is used instead of $\mathcal{W}_t$ as the argument for functions that depend on the realization of the random process up to stage t.

Assumption 3 states that there is no generalization of information associated with a sample path to neighboring paths of the random process. This is justified since the scenario tree is finite and each scenario has a positive probability of being sampled. It follows from this assumption and (12) that the radial basis functions satisfy

$$\phi_t^k(w_t) = \begin{cases} 1 & \text{if } w_t = w_t^k, \\ 0 & \text{else,} \end{cases}$$

for each $t \in \{1,\dots,T\}$, $k \in \mathbb{Z}_+$, and $w_t \in \mathcal{N}(t)$, where w_t^k is the stage-t node of the tree branch sampled at the beginning of iteration k. This and (11) give that the step lengths can be expressed as

$$\alpha_t^k(w_t) = \begin{cases} \beta_k & \text{if } w_t = w_t^k, \\ 0 & \text{else.} \end{cases} \tag{29}$$

Assumptions 4, 5 and 6 are similar to those made in the convergence analysis of other stochastic hybrid approximation algorithms [8] [36]. One difference is that here convexity and not strong convexity is assumed for the initial function approximations $\widehat{G}_t(\,\cdot\,,w_t)$. The reason for this is that strong convexity here is already provided by $F_t(\,\cdot\,,w_t)$. The second difference is that the norm of the sampled subgradients ξ_t^k are allowed to increase with increasing $\big\|g_{t+1}^k(w_{t+1}^k)\big\|_2$. The reason for this is that $\xi_t^k \in \partial H_{t+1}^k\big(x_t^k, w_{t+1}^k\big)$, and $g_{t+1}^k\big(w_{t+1}^k\big)$ is a varying component that affects the objective function of the optimization problem that defines $H_{t+1}^k\big(\,\cdot\,,w_{t+1}^k\big)$.

Assumption 7 is a technical condition that ensures that if the iterates produced by the algorithm get close to the components of $y_{t,\tau}^k$, then subgradients associated with these points can be found that also get close between them.

Lastly, Assumptions 8 and 9 are similar to the "drift" assumption used in the convergence analysis of the primal-dual stochastic hybrid approximation algorithm described in [36]. Both assumptions involve obscure quantities but can be interpreted in simple terms: Assumption 8 can be interpreted as assuming that sums of diminishing $\mathcal{O}(\beta_k)$ changes in past and future stages, as measured with respect to the present stage, are bounded. In practice, a lack of synchronization or correlation among these uncontrolled changes at different iterations would suffice to make the sums bounded. Similarly, assumption 9 can be interpreted as assuming that slope errors when they approach zero they do so either fast enough or are not fully synchronized or directionally correlated to solution errors.

In the proofs of the following Lemmas, Corollaries and Theorems, the notation is further simplified by dropping node arguments w_t from the notation of functions whenever possible.

Lemma 1. *The slope corrections $g_t^k(w_t)$ are uniformly bounded over all $t \in \{1, \ldots, T\}$, $w_t \in \mathcal{N}(t)$, and $k \in \mathbb{Z}_+$.*

Proof. From (10) it holds that

$$\begin{aligned} g_t^{k+1} &= g_t^k + \alpha_t^k(\xi_t^k - \eta_t^k - g_t^k) \\ &= (1 - \alpha_t^k) g_t^k + \alpha_t^k(\xi_t^k - \eta_t^k). \end{aligned}$$

Hence,

$$\left\|g_t^{k+1}\right\|_2 \leq (1 - \alpha_t^k)\left\|g_t^k\right\|_2 + \alpha_t^k \left(\left\|\xi_t^k\right\|_2 + \left\|\eta_t^k\right\|_2\right).$$

Assumptions 4 and 6, and the compactness of the feasible sets then imply that there exist M_1 and M_2 such that

$$\left\|g_t^{k+1}\right\|_2 \leq (1 - \alpha_t^k)\left\|g_t^k\right\|_2 + \alpha_t^k \left(M_1 + M_1\left\|g_{t+1}^k(w_{t+1}^k)\right\|_2 + M_2\right)$$

for all t and k. Since $g_T^k = 0$ for all k, it holds that

$$\left\|g_{T-1}^{k+1}\right\|_2 \leq (1 - \alpha_{T-1}^k)\left\|g_{T-1}^k\right\|_2 + \alpha_{T-1}^k N_{T-1}$$

for all k, where $N_{T-1} := M_1 + M_2$. It follows from this and $0 \leq \alpha_t^k \leq 1$ that

$$\left\|g_{T-1}^k\right\|_2 \leq N_{T-1} \implies \left\|g_{T-1}^{k+1}\right\|_2 \leq N_{T-1}$$

for all k. Hence, $g_{T-1}^0 = 0$ and induction on k give that $\left\|g_{T-1}^k\right\|_2 \leq N_{T-1}$ for all k. Using this bound and repeating the analysis recursively for decreasing t gives that

$$\left\|g_t^k\right\|_2 \leq N_t \implies \left\|g_t^{k+1}\right\|_2 \leq N_t$$

for all k, where N_t is defined recursively by

$$N_t = \begin{cases} 0 & \text{if } t = T, \\ M_1 + M_1 N_{t+1} + M_2 & \text{else,} \end{cases}$$

for $t \in \{1, \ldots, T\}$. Hence, $g_t^0 = 0$ and induction on k give that $\left\|g_t^k\right\|_2 \leq N_t$ for all k and t. It can be concluded from this that $\left\|g_t^k\right\|_2 \leq \max\left\{ N_\tau \mid \tau = 1, \ldots, T \right\}$ for all t and k. □

Lemma 2. *The function approximations $\widehat{G}_t^k(\,\cdot\,, w_t)$ defined in* (21) *and their gradients are uniformly Lipschitz over all $t \in \{1, \ldots, T\}$, $w_t \in \mathcal{N}(t)$, and $k \in \mathbb{Z}_+$.*

Proof. This follows directly from Assumption 4 and Lemma 1. □

Corollary 1. *The slope correction vectors defined by* $\Delta g_t^k := \xi_t^k - \eta_t^k - g_t^k(w_t^k)$ *are uniformly bounded over all* $t \in \{1, \ldots, T\}$ *and* $k \in \mathbb{Z}_+$.

Proof. This follows directly from the compactness of the feasible sets, Assumptions 4 and 6, and Lemma 1. □

Lemma 3. *The functions* $L_{t,\tau}^k(\,\cdot\,, w_t)$ *are uniformly strongly convex over all* $t \in \{1, \ldots, T\}$, $\tau \in \{0, \ldots, T-t\}$, $w_t \in \mathcal{N}(t)$, *and* $k \in \mathbb{Z}_+$.

Proof. This follows directly from Assumptions 1, 2 and 4, and Lemma 1. □

Lemma 4. *There exists an* M *such that for all* $t \in \{1, \ldots, T\}$, $\tau \in \{0, \ldots, T-t\}$, $w_t \in \mathcal{N}(t)$, $k \in \mathbb{Z}_+$, *and stage-*$(t-1)$ *feasible* x,

$$\left\|\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x, w_t) - \widehat{\mathcal{H}}_{t,\tau}^{k}(x, w_t)\right\|_2 \leq M\beta_k.$$

Proof. Letting $y_0 := \widehat{\mathcal{H}}_{t,\tau}^{k}(x, w_t)$ and $y_1 := \widehat{\mathcal{H}}_{t,\tau}^{k+1}(x, w_t)$, and using the definition of $\widehat{\mathcal{H}}_{t,\tau}^{k}$ and $L_{t,\tau}^k$ gives that there exist $\zeta_0 \in \partial L_{t,\tau}^k(y_0, w_t)$ and $\zeta_1 \in \partial L_{t,\tau}^{k+1}(y_1, w_t)$ such that $\zeta_0^T(y_1 - y_0) \geq 0$ and $\zeta_1^T(y_0 - y_1) \geq 0$. From (10), (21), and (22), it holds that

$$0 \leq \zeta_1^T(y_0 - y_1) = \bar{\zeta_1}^T(y_0 - y_1) + \mathbb{E}\left[\, \alpha_t^k(w_{t+\tau})\big(\Delta g_{t+\tau}^k\big)^T\big(z_0(w_{t+\tau}) - z_1(w_{t+\tau})\big) \,\middle|\, w_t \,\right],$$

where $\bar{\zeta_1} \in \partial L_{t,\tau}^k(y_1, w_t)$, $z_0(w_{t+\tau})$ and $z_1(w_{t+\tau})$ denote the components of y_0 and y_1, respectively, associated with node $w_{t+\tau}$, and the expectation is with respect to $w_{t+\tau}$ given w_t. Combining this and $\zeta_0^T(y_1 - y_0) \geq 0$ gives

$$\big(\bar{\zeta_1} - \zeta_0\big)^T(y_1 - y_0) \leq \mathbb{E}\left[\, \alpha_t^k(w_{t+\tau})\big(\Delta g_{t+\tau}^k\big)^T\big(z_0(w_{t+\tau}) - z_1(w_{t+\tau})\big) \,\middle|\, w_t \,\right].$$

From this, Corollary 1, (29), Lemma 3, and the compactness of the feasible sets, it follows that there exist positive constants C and M_1 independent of x, t, w_t, k and τ such that

$$C\|y_1 - y_0\|_2^2 \leq M_1\beta_k\|y_1 - y_0\|_2,$$

and hence that

$$\left\|\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x, w_t) - \widehat{\mathcal{H}}_{t,\tau}^{k}(x, w_t)\right\|_2 \leq M_1\beta_k/C.$$

□

Lemma 5. *The function* $\widehat{\mathcal{H}}_{t,\tau}^{k}(\,\cdot\,, w_t)$ *is uniformly Lipschitz continuous for stage-*$(t-1)$ *feasible inputs over all* $t \in \{1, \ldots, T\}$, $\tau \in \{0, \ldots, T-1\}$, $w_t \in \mathcal{N}(t)$, *and* $k \in \mathbb{Z}_+$.

Proof. From the assumptions of Section 2 and Lemma 1, the function $\widehat{\mathcal{H}}_{t,\tau}^{k}(\,\cdot\,, w_t)$ is real-valued and convex on an open convex set that contains all stage-$(t-1)$ feasible points. Hence, it is continuous. Since the set of stage-$(t-1)$ feasible inputs is compact, the function is Lipschitz continuous on this set. From this, Lemma 1, and the fact that there are finitely many t, τ, and w_t, there exists an M_1 such that for all t, τ, k, w_t, and stage-$(t-1)$ feasible inputs x_0 and x_1, it holds that

$$\left\|\widehat{\mathcal{H}}_{t,\tau}^{k}(x_1, w_t) - \widehat{\mathcal{H}}_{t,\tau}^{k}(x_0, w_t)\right\|_2 \leq M_1\|x_1 - x_0\|_2.$$

□

Lemma 6. *There exists an M such that for any $t \in \{1, \dots, T\}$, $\tau \in \{0, \dots, T-t\}$, and $k \in \mathbb{Z}_+$,*

$$\left\|y_{t,\tau}^{k+1} - y_{t,\tau}^{k}\right\|_2 \leq M\beta_k,$$

where $y_{t,\tau}^k$ is as defined in (23).

Proof. For all $t \in \{1, \dots, T\}$, $\tau \in \{0, \dots, T-t\}$, and $k \in \mathbb{Z}_+$, it holds that

$$\begin{aligned}\left\|y_{t,\tau}^{k+1} - y_{t,\tau}^{k}\right\|_2 &= \left\|\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x_{t-1}^{k+1}) - \widehat{\mathcal{H}}_{t,\tau}^{k}(x_{t-1}^{k})\right\|_2 \\ &\leq \left\|\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x_{t-1}^{k+1}) - \widehat{\mathcal{H}}_{t,\tau}^{k+1}(x_{t-1}^{k})\right\|_2 + \left\|\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x_{t-1}^{k}) - \widehat{\mathcal{H}}_{t,\tau}^{k}(x_{t-1}^{k})\right\|_2.\end{aligned}$$

From this and Lemmas 4 and 5, there exist positive constants M_1 and M_2 such that

$$\left\|y_{t,\tau}^{k+1} - y_{t,\tau}^{k}\right\|_2 \leq M_1\|x_{t-1}^{k+1} - x_{t-1}^{k}\|_2 + M_2\beta_k.$$

Since $y_{t,0}^k = x_t^k$ for all k and t by definition, the above inequality gives that

$$\|x_t^{k+1} - x_t^k\|_2 \leq M_1\|x_{t-1}^{k+1} - x_{t-1}^{k}\|_2 + M_2\beta_k.$$

Since, $x_0^{k+1} = x_0^k$ for all k, the above inequality implies that

$$\|x_t^{k+1} - x_t^k\|_2 \leq \sum_{\varsigma=0}^{t-1} M_1^{\varsigma} M_2\beta_k \leq \sum_{\varsigma=0}^{T-1} M_1^{\varsigma} M_2\beta_k.$$

It follows that

$$\left\|y_{t,\tau}^{k+1} - y_{t,\tau}^{k}\right\|_2 \leq \sum_{\varsigma=0}^{T} M_1^{\varsigma} M_2\beta_k$$

for all t, τ, and k. □

Lemma 7. *For each $t \in \{1, \dots, T-1\}$ and $\tau \in \{0, \dots, T-t-1\}$, the sequence $\{T_{t,\tau}^k\}_{k\in\mathbb{Z}_+}$ defined by*

$$T_{t,\tau}^k := L_{t,\tau}^k\big(y_{t,\tau+1}^k, w_t^k\big) - L_{t,\tau}^k\big(y_{t,\tau}^k, w_t^k\big)$$

satisfies

$$\begin{aligned}\mathbb{E}\left[T_{t,\tau}^{k+1} \,\middle|\, w_t^k\right] - T_{t,\tau}^k \leq \mathbb{E}\Big[\Gamma_{t,\tau}^k + \Upsilon_{t,\tau}^k + \Psi_{t,\tau}^k + \Theta_{t,\tau}^k + \Xi_{t,\tau}^k \,\Big|\, w_t^k\Big] \; + \\ p_{t,\tau}^k\beta_k\big(L_{t,\tau+1}^k\big(y_{t,\tau+1}^k, w_t^k\big) - L_{t,\tau+1}^k\big(y_{t,\tau}^k, w_t^k\big)\big) \; + \\ \mathcal{O}\left(\mathbb{E}\left[\beta_k^2 \,\middle|\, w_t^k\right]\right)\end{aligned}$$

for all $k \in \mathbb{Z}_+$, where $p_{t,\tau}^k > 0$ denotes the probability that $w_{t+\tau} = w_{t+\tau}^k$ given w_t^k.

Proof. From (10) and the definitions of $T_{t,\tau}^k$ and $L_{t,\tau}^k$, it holds that

$$\begin{aligned}T_{t,\tau}^{k+1} - T_{t,\tau}^k &= L_{t,\tau}^{k+1}\big(y_{t,\tau+1}^{k+1}\big) - L_{t,\tau}^{k+1}\big(y_{t,\tau}^{k+1}\big) - L_{t,\tau}^{k}\big(y_{t,\tau+1}^{k}\big) + L_{t,\tau}^{k}\big(y_{t,\tau}^{k}\big) \\ &= L_{t,\tau}^{k}\big(y_{t,\tau+1}^{k+1}\big) - L_{t,\tau}^{k}\big(y_{t,\tau}^{k+1}\big) - L_{t,\tau}^{k}\big(y_{t,\tau+1}^{k}\big) + L_{t,\tau}^{k}\big(y_{t,\tau}^{k}\big) \; + \\ &\quad \mathbb{E}\left[\alpha_{t+\tau}^k(w_{t+\tau})\big(\Delta g_{t+\tau}^k\big)^T\big(z_{t,\tau+1}^{k+1}(w_{t+\tau}) - z_{t,\tau}^{k+1}(w_{t+\tau})\big) \,\middle|\, w_t^k\right],\end{aligned}$$

where $z_{t,\tau+1}^{k+1}(w_{t+\tau})$ and $z_{t,\tau}^{k+1}(w_{t+\tau})$ are the components of $y_{t,\tau+1}^{k+1}$ and $y_{t,\tau}^{k+1}$, respectively, associated with node $w_{t+\tau}$, and the expectation is with respect to $w_{t+\tau}$ given w_t^k. From Corollary 1, Lemma 6, and (29), it follows that

$$\begin{aligned} T_{t,\tau}^{k+1} - T_{t,\tau}^{k} \leq\ & L_{t,\tau}^{k}\big(y_{t,\tau+1}^{k+1}\big) - L_{t,\tau}^{k}\big(y_{t,\tau}^{k+1}\big) - L_{t,\tau}^{k}\big(y_{t,\tau+1}^{k}\big) + L_{t,\tau}^{k}\big(y_{t,\tau}^{k}\big) \ + \\ & \mathbb{E}\Big[\ \alpha_{t+\tau}^{k}(w_{t+\tau})\big(\Delta g_{t+\tau}^{k}\big)^{T}\big(z_{t,\tau+1}^{k}(w_{t+\tau}) - z_{t,\tau}^{k}(w_{t+\tau})\big)\ \Big|\ w_t^k\ \Big]\ + \\ & \mathcal{O}\big(\beta_k^2\big). \end{aligned}$$

Then, using (23) and adding and subtracting terms $L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^{k}\big)\big)$ and $L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^{k}\big)\big)$, it holds that

$$\begin{aligned} T_{t,\tau}^{k+1} - T_{t,\tau}^{k} \leq\ & L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^{k+1}\big)\big) - L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^{k}\big)\big) \ + \\ & L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k+1}\big(x_{t-1}^{k}\big)\big) - L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau+1}^{k}\big(x_{t-1}^{k}\big)\big) \ + \\ & L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau}^{k}\big(x_{t-1}^{k}\big)\big) - L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^{k}\big)\big) \ + \\ & L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^{k}\big)\big) - L_{t,\tau}^{k}\big(\widehat{\mathcal{H}}_{t,\tau}^{k+1}\big(x_{t-1}^{k+1}\big)\big) \ + \\ & \mathbb{E}\Big[\ \alpha_{t+\tau}^{k}(w_{t+\tau})\big(\Delta g_{t+\tau}^{k}\big)^{T}\big(z_{t,\tau+1}^{k}(w_{t+\tau}) - z_{t,\tau}^{k}(w_{t+\tau})\big)\ \Big|\ w_t^k\ \Big]\ + \\ & \mathcal{O}\big(\beta_k^2\big). \end{aligned}$$

Using (24), (25) and (26), and the inequality

$$L_{t,\tau}^{k}(\widehat{\mathcal{H}}_{t,\tau}^{k}(x_{t-1}^{k})) \leq L_{t,\tau}^{k}(\widehat{\mathcal{H}}_{t,\tau}^{k+1}(x_{t-1}^{k})),$$

which follows from the optimality of $\widehat{\mathcal{H}}_{t,\tau}^{k}\big(x_{t-1}^{k}\big)$ with respect to $L_{t,\tau}^{k}\big(\,\cdot\,, w_t^k\big)$, gives

$$\begin{aligned} T_{t,\tau}^{k+1} - T_{t,\tau}^{k} \leq\ & \Gamma_{t,\tau}^{k} + \Upsilon_{t,\tau}^{k} + \Psi_{t,\tau}^{k} \ + \\ & \mathbb{E}\Big[\ \alpha_{t+\tau}^{k}(w_{t+\tau})\big(\Delta g_{t+\tau}^{k}\big)^{T}\big(z_{t,\tau+1}^{k}(w_{t+\tau}) - z_{t,\tau}^{k}(w_{t+\tau})\big)\ \Big|\ w_t^k\ \Big]\ + \\ & \mathcal{O}\big(\beta_k^2\big). \end{aligned}$$

The definition of Δg_t^k and $\widehat{G}_t^k(\,\cdot\,, w_t)$ then gives

$$\begin{aligned} T_{t,\tau}^{k+1} - T_{t,\tau}^{k} \leq\ & \Gamma_{t,\tau}^{k} + \Upsilon_{t,\tau}^{k} + \Psi_{t,\tau}^{k} \ + \\ & \mathbb{E}\Big[\ \alpha_{t+\tau}^{k}(w_{t+\tau})\big(\xi_{t+\tau}^{k}\big)^{T}\big(z_{t,\tau+1}^{k}(w_{t+\tau}) - z_{t,\tau}^{k}(w_{t+\tau})\big)\ \Big|\ w_t^k\ \Big]\ - \\ & \mathbb{E}\Big[\ \alpha_{t+\tau}^{k}(w_{t+\tau})\big(\nabla\widehat{G}_{t+\tau}^{k}\big(x_{t+\tau}^{k}, w_{t+\tau}^{k}\big)\big)^{T}\big(z_{t,\tau+1}^{k}(w_{t+\tau}) - z_{t,\tau}^{k}(w_{t+\tau})\big)\ \Big|\ w_t^k\ \Big]\ + \\ & \mathcal{O}\big(\beta_k^2\big), \end{aligned}$$

where $\xi_{t+\tau}^{k} \in \partial H_{t+\tau+1}^{k}\big(x_{t+\tau}^{k}, w_{t+\tau+1}^{k}\big)$, as defined in (8). From (29) and Assumption 2, it follows that

$$\begin{aligned} T_{t,\tau}^{k+1} - T_{t,\tau}^{k} \leq\ & \Gamma_{t,\tau}^{k} + \Upsilon_{t,\tau}^{k} + \Psi_{t,\tau}^{k} \ + \\ & p_{t,\tau}^{k}\beta_k\big(\xi_{t+\tau}^{k}\big)^{T}\big(z_{t,\tau+1}^{k}\big(w_{t+\tau}^{k}\big) - z_{t,\tau}^{k}\big(w_{t+\tau}^{k}\big)\big) \ - \\ & p_{t,\tau}^{k}\beta_k\big(\nabla\widehat{G}_{t+\tau}^{k}\big(x_{t+\tau}^{k}, w_{t+\tau}^{k}\big)\big)^{T}\big(z_{t,\tau+1}^{k}\big(w_{t+\tau}^{k}\big) - z_{t,\tau}^{k}\big(w_{t+\tau}^{k}\big)\big) \ + \\ & \mathcal{O}\big(\beta_k^2\big), \end{aligned}$$

where $p^k_{t,\tau} > 0$ denotes the probability that $w_{t+\tau} = w^k_{t+\tau}$ given w^k_t. Adding and subtracting terms, it holds that

$$\begin{aligned} T^{k+1}_{t,\tau} - T^k_{t,\tau} \leq \Gamma^k_{t,\tau} + \Upsilon^k_{t,\tau} + \Psi^k_{t,\tau} + \Theta^k_{t,\tau} + \Xi^k_{t,\tau} + \\ p^k_{t,\tau}\beta_k\big(\bar{\xi}^k_{t+\tau}\big)^T\big(z^k_{t,\tau+1}\big(w^k_{t+\tau}\big) - z^k_{t,\tau}\big(w^k_{t+\tau}\big)\big) - \\ p^k_{t,\tau}\beta_k\big(\nabla\widehat{G}^k_{t+\tau}\big(z^k_{t,\tau}\big(w^k_{t+\tau}\big), w^k_{t+\tau}\big)\big)^T\big(z^k_{t,\tau+1}\big(w^k_{t+\tau}\big) - z^k_{t,\tau}\big(w^k_{t+\tau}\big)\big) + \\ \mathcal{O}\big(\beta^2_k\big), \end{aligned}$$

where $\bar{\xi}^k_{t+\tau} \in \partial H^k_{t+\tau+1}\big(z^k_{t,\tau}\big(w^k_{t+\tau}\big), w^k_{t+\tau+1}\big)$, as in Assumption 7. Taking expectation conditioned on w^k_t, and adding and subtracting terms gives

$$\begin{aligned} \mathbb{E}\left[T^{k+1}_{t,\tau} \,\middle|\, w^k_t\right] - T^k_{t,\tau} \leq \mathbb{E}\left[\Gamma^k_{t,\tau} + \Upsilon^k_{t,\tau} + \Psi^k_{t,\tau} + \Theta^k_{t,\tau} + \Xi^k_{t,\tau} \,\middle|\, w^k_t\right] + \\ p^k_{t,\tau}\beta_k\big(\zeta^k_{t,\tau+1}\big)^T\big(y^k_{t,\tau+1} - y^k_{t,\tau}\big) - \\ p^k_{t,\tau}\beta_k\big(\zeta^k_{t,\tau}\big)^T\big(y^k_{t,\tau+1} - y^k_{t,\tau}\big) + \\ \mathcal{O}\left(\mathbb{E}\left[\beta^2_k \,\middle|\, w^k_t\right]\right), \end{aligned}$$

where $\zeta^k_{t,\tau+1} \in \partial L^k_{t,\tau+1}(y^k_{t,\tau}, w^k_t)$ and $\zeta^k_{t,\tau} \in \partial L^k_{t,\tau}(y^k_{t,\tau}, w^k_t)$. We note here that for each term added and subtracted, the former was used to construct $p^k_{t,\tau}\beta_k\big(\zeta^k_{t,\tau+1}\big)^T\big(y^k_{t,\tau+1} - y^k_{t,\tau}\big)$ while the latter was used to construct $-p^k_{t,\tau}\beta_k\big(\zeta^k_{t,\tau}\big)^T\big(y^k_{t,\tau+1} - y^k_{t,\tau}\big)$. The subgradient $\zeta^k_{t,\tau}$ obtained is assumed to one such that, because of the optimality of $y^k_{t,\tau}$, results in $\big(\zeta^k_{t,\tau}\big)^T\big(y^k_{t,\tau+1} - y^k_{t,\tau}\big) \geq 0$. From this and the convexity of $L^k_{t,\tau+1}\big(\,\cdot\,, w^k_t\big)$, it follows that

$$\begin{aligned} \mathbb{E}\left[T^{k+1}_{t,\tau} \,\middle|\, w^k_t\right] - T^k_{t,\tau} \leq \mathbb{E}\left[\Gamma^k_{t,\tau} + \Upsilon^k_{t,\tau} + \Psi^k_{t,\tau} + \Theta^k_{t,\tau} + \Xi^k_{t,\tau} \,\middle|\, w^k_t\right] + \\ p^k_{t,\tau}\beta_k\big(L^k_{t,\tau+1}\big(y^k_{t,\tau+1}, w^k_t\big) - L^k_{t,\tau+1}\big(y^k_{t,\tau}, w^k_t\big)\big) + \\ \mathcal{O}\left(\mathbb{E}\left[\beta^2_k \,\middle|\, w^k_t\right]\right). \end{aligned}$$

□

Lemma 8. *For each $t \in \{1, \ldots, T-1\}$ and $\tau \in \{0, \ldots, T-t-1\}$, if $\sum_{k=0}^{\infty}\beta_k\big\|x^k_{t+\tau} - z^k_{t,\tau}\big(w^k_{t+\tau}\big)\big\|_2^2$ is finite almost surely, where $z^k_{t,\tau}\big(w^k_{t+\tau}\big)$ denotes the components of $y^k_{t,\tau}$ associated with node $w^k_{t+\tau}$, then $\sum_{k=0}^{\infty}\beta_k\big\|y^k_{t,\tau} - y^k_{t,\tau+1}\big\|_2^2$ is finite almost surely.*

Proof. From Lemma 7, it holds that

$$\begin{aligned} \mathbb{E}\left[T^{k+1}_{t,\tau} \,\middle|\, w^k_t\right] - T^k_{t,\tau} \leq \mathbb{E}\left[\Gamma^k_{t,\tau} + \Upsilon^k_{t,\tau} + \Psi^k_{t,\tau} + \Theta^k_{t,\tau} + \Xi^k_{t,\tau} \,\middle|\, w^k_t\right] + \\ p^k_{t,\tau}\beta_k\big(L^k_{t,\tau+1}\big(y^k_{t,\tau+1}, w^k_t\big) - L^k_{t,\tau+1}\big(y^k_{t,\tau}, w^k_t\big)\big) + \\ \mathcal{O}\left(\mathbb{E}\left[\beta^2_k \,\middle|\, w^k_t\right]\right). \end{aligned}$$

Taking expectation, rearranging, and summing over k from 0 to K gives

$$p_{t,\tau}\mathbb{E}\left[\sum_{k=0}^{K}\beta_k\big(L^k_{t,\tau+1}\big(y^k_{t,\tau}, w^k_t\big) - L^k_{t,\tau+1}\big(y^k_{t,\tau+1}, w^k_t\big)\big)\right] \leq \mathbb{E}\left[T^0_{t,\tau} - T^{K+1}_{t,\tau}\right] +$$

$$\mathbb{E}\left[\sum_{k=0}^{K}\left(\Gamma_{t,\tau}^{k}+\Upsilon_{t,\tau}^{k}+\Psi_{t,\tau}^{k}\right)\right]+$$
$$\mathbb{E}\left[\sum_{k=0}^{K}\left(\Theta_{t,\tau}^{k}+\Xi_{t,\tau}^{k}\right)\right]+$$
$$\mathcal{O}\left(\mathbb{E}\left[\sum_{k=0}^{K}\beta_{k}^{2}\right]\right),$$

where $p_{t,\tau} := \min_{k\in\mathbb{Z}_+} p_{t,\tau}^k > 0$. If $\sum_{k=0}^{\infty}\beta_k\|x_{t+\tau}^k - z_{t,\tau}^k(w_{t+\tau}^k)\|_2^2$ is finite almost surely then this together with Assumptions 5, 7, 8 and 9, Lemmas 2, 4 and 6, and the uniform boundedness of $T_{t,\tau}^k$ imply that there exists an M_1 such that

$$\sum_{k=0}^{K}\mathbb{E}\left[\beta_k\big(L_{t,\tau+1}^k(y_{t,\tau}^k,w_t^k)-L_{t,\tau+1}^k(y_{t,\tau+1}^k,w_t^k)\big)\right] < M_1 < \infty$$

for all $K\in\mathbb{Z}_+$. From the uniform strong convexity of $L_{t,\tau+1}^k(\,\cdot\,,w_t^k)$ (Lemma 3), and the optimality of $y_{t,\tau+1}^k$ with respect to $L_{t,\tau+1}^k(\,\cdot\,,w_t^k)$, it holds that there exists an $M_2>0$ such that

$$M_2\|y_{t,\tau}^k - y_{t,\tau+1}^k\|_2^2 \le L_{t,\tau+1}^k(y_{t,\tau}^k,w_t^k) - L_{t,\tau+1}^k(y_{t,\tau+1}^k,w_t^k)$$

for all $k\in\mathbb{Z}_+$. It follows that

$$\sum_{k=0}^{\infty}\mathbb{E}\left[\beta_k\|y_{t,\tau}^k - y_{t,\tau+1}^k\|_2^2\right] < M_1/M_2 < \infty,$$

and hence that $\sum_{k=0}^{\infty}\beta_k\|y_{t,\tau}^k - y_{t,\tau+1}^k\|_2^2$ is finite almost surely. □

Lemma 9. *For each $t\in\{1,\dots,T-1\}$ and $\tau\in\{0,\dots,T-t-1\}$, $\sum_{k=0}^{\infty}\beta_k\|y_{t,\tau}^k - y_{t,\tau+1}^k\|_2^2$ is finite almost surely.*

Proof. For $t\in\{1,\dots,T-1\}$, it holds by definition that $x_t^k = z_{t,0}^k(w_t^k)$, where $z_{t,\tau}^k(w)$ denotes the component of $y_{t,\tau}^k$ associated with node w. Hence, $\sum_{k=0}^{\infty}\beta_k\|x_t^k - z_{t,0}^k(w_t^k)\|_2^2 = 0$ and Lemma 8 implies that $\sum_{k=0}^{\infty}\|y_{t,0}^k - y_{t,1}^k\|_2^2$ is finite almost surely. Now let $\tau>0$ and assume that for each ς such that $0\le\varsigma<\tau$, $\sum_{k=0}^{\infty}\|y_{t,\varsigma}^k - y_{t,\varsigma+1}^k\|_2^2$ is finite almost surely. From Lemma 5 and $\|x_t^k - z_{t,0}^k(w_t^k)\|_2 = 0$, it follows that there exists an $M_1>1$ such that

$$\begin{aligned}
\|x_{t+\tau}^k - z_{t,\tau}^k(w_{t+\tau}^k)\|_2 &= \|\widehat{\mathcal{H}}_{t+\tau,0}^k(x_{t+\tau-1}^k,w_{t+\tau}^k) - \widehat{\mathcal{H}}_{t+\tau,0}^k(z_{t,\tau}^k(w_{t+\tau-1}^k),w_{t+\tau}^k)\|_2\\
&\le M_1\|x_{t+\tau-1}^k - z_{t,\tau}^k(w_{t+\tau-1}^k)\|_2\\
&\le M_1\|x_{t+\tau-1}^k - z_{t,\tau-1}^k(w_{t+\tau-1}^k)\|_2 + M_1\|z_{t,\tau-1}^k(w_{t+\tau-1}^k) - z_{t,\tau}^k(w_{t+\tau-1}^k)\|_2\\
&\vdots\\
&\le \sum_{\varsigma=0}^{\tau-1} M_1^{\tau-\varsigma}\|z_{t,\varsigma}^k(w_{t+\varsigma}^k) - z_{t,\varsigma+1}^k(w_{t+\varsigma}^k)\|_2\\
&\le M_1^{\tau}\sum_{\varsigma=0}^{\tau-1}\|z_{t,\varsigma}^k(w_{t+\varsigma}^k) - z_{t,\varsigma+1}^k(w_{t+\varsigma}^k)\|_2
\end{aligned}$$

$$\leq M_1^\tau \sum_{\varsigma=0}^{\tau-1} \left\| y_{t,\varsigma}^k - y_{t,\varsigma+1}^k \right\|_2$$

for all $k \in \mathbb{Z}_+$. From this and the equivalence of norms in finite-dimensional spaces, there exists an $M_2 > 0$ such that

$$\left\| x_{t+\tau}^k - z_{t,\tau}^k\big(w_{t+\tau}^k\big) \right\|_2 \leq M_1^\tau M_2 \sqrt{\sum_{\varsigma=0}^{\tau-1} \left\| y_{t,\varsigma}^k - y_{t,\varsigma+1}^k \right\|_2^2}$$

for all $k \in \mathbb{Z}_+$, and hence that

$$\sum_{k=0}^{\infty} \beta_k \left\| x_{t+\tau}^k - z_{t,\tau}^k\big(w_{t+\tau}^k\big) \right\|_2^2 \leq \big(M_1^\tau M_2\big)^2 \sum_{\varsigma=0}^{\tau-1} \sum_{k=0}^{\infty} \beta_k \left\| y_{t,\varsigma}^k - y_{t,\varsigma+1}^k \right\|_2^2.$$

This and the induction hypothesis then gives that $\sum_{k=0}^{\infty} \beta_k \left\| x_{t+\tau}^k - z_{t,\tau}^k\big(w_{t+\tau}^k\big) \right\|_2^2$ is finite almost surely. Lemma 8 then gives that $\sum_{k=0}^{\infty} \beta_k \left\| y_{t,\tau}^k - y_{t,\tau+1}^k \right\|_2^2$ is finite almost surely. By strong induction, this holds for all $\tau \in \{0, \ldots, T-t-1\}$. □

Theorem 1. *There exists a set of strictly increasing indices* $\{ n_k \in \mathbb{Z}_+ \mid k \in \mathbb{Z}_+ \}$ *such that for all* $t \in \{1, \ldots, T\}$, $\left\| y_{t,0}^{n_k} - y_{t,T-t}^{n_k} \right\|_2 \to 0$ *almost surely as* $k \to \infty$.

Proof. For $t = T$, the statement is trivial. For $t \in \{1, \ldots, T-1\}$, Lemma 9 gives that the sum $\sum_{k=0}^{\infty} \beta_k \left\| y_{t,\tau}^k - y_{t,\tau+1}^k \right\|_2^2$ is finite almost surely for each $\tau \in \{0, \ldots, T-t-1\}$. Hence,

$$\sum_{k=0}^{\infty} \beta_k \sum_{t=1}^{T-1} \sum_{\tau=0}^{T-t-1} \left\| y_{t,\tau}^k - y_{t,\tau+1}^k \right\|_2^2 < \infty$$

almost surely. Since $\sum_{k=0}^{\infty} \beta_k = \infty$ almost surely from Assumption 5, it follows that there exists a set of strictly increasing indices $\{ n_k \in \mathbb{Z}_+ \mid k \in \mathbb{Z}_+ \}$ such that

$$\sum_{t=1}^{T-1} \sum_{\tau=0}^{T-t-1} \left\| y_{t,\tau}^{n_k} - y_{t,\tau+1}^{n_k} \right\|_2^2 \to 0$$

almost surely as $k \to \infty$. It follows from this that $\left\| y_{t,\tau}^{n_k} - y_{t,\tau+1}^{n_k} \right\|_2 \to 0$ almost surely for each $t \in \{1, \ldots, T-1\}$ and $\tau \in \{0, \ldots, T-t-1\}$ as $k \to \infty$. This combined with

$$\left\| y_{t,0}^{n_k} - y_{t,T-t}^{n_k} \right\|_2 \leq \sum_{\tau=0}^{T-t-1} \left\| y_{t,\tau}^{n_k} - y_{t,\tau+1}^{n_k} \right\|_2$$

gives that $\left\| y_{t,0}^{n_k} - y_{t,T-t}^{n_k} \right\|_2 \to 0$ almost surely as $k \to \infty$. □

Acknowledgment

Support for this work was provided by the ETH Zurich Postdoctoral Fellowship FEL-11 15-1.

References

[1] S. Amari and S. Wu. Improving support vector machine classifiers by modifying kernel functions. *Neural Networks*, 12(6):783–789, 1999.

[2] T. Asamov and W. Powell. Regularized decomposition of high-dimensional multistage stochastic programs with Markov uncertainty. *arXiv preprint arXiv:1505.02227*, May 2015.

[3] J. Benders. Partitioning procedures for solving mixed-variables programming problems. *Numerische Mathematik*, 4(1):238–252, December 1962.

[4] R. Bessa, C. Moreira, B. Silva, and M. Matos. Handling renewable energy variability and uncertainty in power systems operation. *Wiley Interdisciplinary Reviews: Energy and Environment*, 3(2):156–178, 2014.

[5] S. Bhatnagar, H. Prasad, and L. Prashanth. *Stochastic Recursive Algorithms for Optimization: Simultaneous Perturbation Methods*, chapter Stochastic Approximation Algorithms, pages 17–28. Springer London, 2013.

[6] J. Birge. Decomposition and partitioning methods for multistage stochastic linear programs. *Operations Research*, 33(5):989–1007, September 1985.

[7] Z. Chen and W. Powell. Convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse. *Journal of Optimization Theory and Applications*, 102(3):497–524, 1999.

[8] R. Cheung and W. Powell. SHAPE - A stochastic hybrid approximation procedure for two-stage stochastic programs. *Operations Research*, 48(1):73–79, 2000.

[9] C. Donohue and J. Birge. The abridged nested decomposition method for multistage stochastic linear programs with relatively complete recourse. *Algorithmic Operations Research*, 1(1), 2006.

[10] P. Girardeau, V. Leclere, and A. Philpott. On the convergence of decomposition methods for multistage stochastic convex programs. *Mathematics of Operations Research*, 40(1):130–145, July 2014.

[11] J. Higle and S. Sen. Stochastic decomposition: An algorithm for two-stage linear programs with recourse. *Mathematics of Operations Research*, 16(3):650–669, 1991.

[12] J. Higle and S. Sen. Finite master programs in regularized stochastic decomposition. *Mathematical Programming*, 67(1):143–168, 1994.

[13] T. Homem De Mello and G. Bayraksan. Monte Carlo sampling-based methods for stochastic optimization. *Surveys in Operations Research and Management Science*, 19(1):56–85, 2014.

[14] E. Jones, T. Oliphant, and P. Peterson. SciPy: Open source scientific tools for Python. `http://www.scipy.org/`, 2001.

[15] H. Kushner and G. Yin. *Stochastic Approximation and Recursive Algorithms and Applications*. Stochastic Modelling and Applied Probability. Springer New York, 2003.

[16] National Renewable Energy Laboratory. The importance of flexible electricity supply. Technical Report DOE/GO-102011-3201, United States Department of Energy, 2011.

[17] K. Lau and R. Womersley. Multistage quadratic stochastic programming. *Journal of Computational and Applied Mathematics*, 129(12):105–138, 2001.

[18] F. Louveaux. A solution method for multistage stochastic programs with recourse with application to an energy investment problem. *Operations Research*, 28(4):889–902, July 1980.

[19] J. Mulvey and A. Ruszczyński. A new scenario decomposition method for large-scale stochastic optimization. *Operations Research*, 43(3):477–490, 1995.

[20] Y. Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. Applied Optimization. Springer US, 2004.

[21] North American Electric Reliability Corporation. Balancing and frequency control. Technical report, prepared by the NERC Resources Subcommittee, January 2011.

[22] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, and E. Duchesnay. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[23] M. Pereira and L. Pinto. Multi-stage stochastic optimization applied to energy planning. *Mathematical Programming*, 52(1):359–375, May 1991.

[24] D. Phan and S. Ghosh. Two-stage stochastic optimization for optimal power flow under renewable generation uncertainty. *ACM Transactions on Modeling and Computer Simulation*, 24(1):1–22, January 2014.

[25] A. Philpott and Z. Guan. On the convergence of stochastic dual dynamic programming and related methods. *Operations Research Letters*, 36(4):450–455, 2008.

[26] R. Rajagopal, E. Bitar, P. Varaiya, and F. Wu. Risk-limiting dispatch for integrating renewable power. *International Journal of Electrical Power and Energy Systems*, 44(1):615–628, January 2013.

[27] S. Rebennack. Combining sampling-based and scenario-based nested Benders decomposition methods: Application to stochastic dual dynamic programming. *Mathematical Programming*, 156(1):343–389, March 2016.

[28] L. Roald, F. Oldewurtel, T. Krause, and G. Andersson. Analytical reformulation of security constrained optimal power flow with probabilistic constraints. In *IEEE PowerTech Conference*, pages 1–6, June 2013.

[29] R. Rockafellar and R. Wets. Scenarios and policy aggregation in optimization under uncertainty. *Mathematics of Operations Research*, 16(1):119–147, February 1991.

[30] S. Sen. Stochastic programming: Computational issues and challenges. *Encyclopedia of operations research and management science*, pages 1–11, 2001.

[31] S. Sen and Z. Zhou. Multistage stochastic decomposition: A bridge between stochastic programming and approximate dynamic programming. *SIAM Journal on Optimization*, 24(1):127–153, 2014.

[32] A. Shapiro. Analysis of stochastic dual dynamic programming method. *European Journal of Operational Research*, 209(1):63–72, February 2011.

[33] A. Shapiro, W. Tekaya, J. Da Costa, and M. Pereira. Risk neutral and risk averse stochastic dual dynamic programming method. *European Journal of Operational Research*, 224(2):375–391, January 2013.

[34] M. Tahanan, W. Van Ackooij, A. Frangioni, and F. Lacalandra. Large-scale unit commitment under uncertainty. *4OR*, 13(2):115–171, 2015.

[35] T. Tinoco De Rubira and G. Hug. Adaptive certainty-equivalent approach for optimal generator dispatch under uncertainty. In *European Control Conference*, June 2016.

[36] T. Tinoco De Rubira and G. Hug. Primal-dual stochastic hybrid approximation algorithm. *Computational Optimization and Applications (under review)*, January 2017.

[37] J. Tsitsiklis and B. Van Roy. Feature-based methods for large scale dynamic programming. *Machine Learning*, 22(1):59–94, January 1996.

[38] S. Van Der Walt, S. Colbert, and G. Varoquaux. The numpy array: A structure for efficient numerical computation. *Computing in Science Engineering*, 13(2):22–30, March 2011.

[39] R. Van Slyke and R. Wets. L-shaped linear programs with applications to optimal control and stochastic programming. *SIAM Journal on Applied Mathematics*, 17(4):638–663, 1969.

[40] S. Wallace and S. Fleten. Stochastic programming models in energy. In *Stochastic Programming*, volume 10 of *Handbooks in Operations Research and Management Science*, pages 637–677. Elsevier, 2003.

[41] R. Yang and G. Hug. Potential and efficient computation of corrective power flow control in cost vs. risk trade-off. *IEEE Transactions on Smart Grid*, 5(4):2033–2043, July 2014.

[42] J. Zhu. *Optimization of Power System Operation*. John Wiley & Sons, Inc., 2009.

[43] R. Zimmerman, C. Murillo-Sánchez, and R. Thomas. MATPOWER: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on Power Systems*, 26(1):12–19, February 2011.