This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

Efficient Uncertainty Quantification in Stochastic Economic Dispatch

Cosmin Safta, Richard L.-Y. Chen, Habib N. Najm, Ali Pinar, Jean-Paul Watson
Sandia National Laboratories
Email: {csafta,rlchen,hnnajm,apinar,jwatson}@sandia.gov

***Abstract*—Stochastic economic dispatch models address uncertainties in forecasts of renewable generation output by considering a finite number of realizations drawn from a stochastic process model, typically via Monte Carlo sampling. Accurate evaluations of expectations or higher-order moments for quantities of interest, e.g., generating cost, can require a prohibitively large number of samples. We propose an alternative to Monte Carlo sampling based on Polynomial Chaos expansions. These representations enable efficient and accurate propagation of uncertainties in model parameters, using sparse quadrature methods. We also use Karhunen-Loève expansions for efficient representation of uncertain renewable energy generation that follows geographical and temporal correlations derived from historical data at each wind farm. Considering expected production cost, we demonstrate that the proposed approach can yield several orders of magnitude reduction in computational cost for solving stochastic economic dispatch relative to Monte Carlo sampling, for a given target error threshold.**

***Index Terms*—Stochastic economic dispatch, Monte Carlo sampling, Polynomial Chaos Expansion, Karhunen-Loève Expansion.**

I. Introduction

UNIT commitment (UC) deals with scheduling thermal generating units in advance of operations in the electric power grid [1]. The objective is to minimize overall production costs to satisfy forecasted electricity load, while respecting operational constraints of transmission elements (e.g., thermal limits) and generators (e.g., ramping limits). Economic dispatch (ED) is a related problem, in which cost minimization is performed to identify an optimal set of power output levels for a fixed set of on-line thermal generating units. Typically, UC and ED are respectively formulated as mixed-integer and linear optimization problems, and they are solved using commercial solvers.

Uncertainties are ubiquitous in both power systems operations and planning, and the importance of credibly accounting for them is well-recognized. Despite recent improvements in load forecasting technology, next-day predictions are imperfect, with errors on average in the 1-3% range and exceeding 10% on specific days [2]. To account for such inaccuracies, reserve margins are universally imposed. These margins implicitly deal with uncertainty in load forecasts, by ensuring there is sufficient generation capacity available to meet unexpectedly high load during operations.

An alternative to dealing with forecast errors is to *explicitly* model the uncertainty, typically via a finite set of sampled realizations from a stochastic process model. This approach results in a stochastic ED (SED) model, in which the objective is to minimize the expected production cost across a set of load scenarios [3], [4]. By explicitly representing the inherent uncertainty in forecasts, a SED solution ensures sufficient flexibility to meet a range of potential realizations during next-day operations. Further, by explicitly representing uncertainty, reliance on reserve margins is reduced, yielding less costly solutions than those obtained under deterministic ED models. Increasing penetration of renewables (e.g., wind and solar) generation accentuates the differences between stochastic and deterministic problems, due to increased errors in next-day forecasts (relative to load, accurate prediction of next-day meteorological conditions is very difficult).

Despite their conceptual appeal, stochastic variants of grid operations models are not yet used in practice due to their computational difficulty [5]. This difficulty is primarily driven by the number of samples required to achieve high-quality, robust solutions to models such as stochastic UC and ED. In their seminal work, Takriti *et al.* [4] studied the problem of generating electric power when loads are uncertain. More recent work focused on various sources of and frameworks for modeling uncertainty (e.g., stochastic programming, chance constraints, and robust optimization). For example, Wang *et al.* [6] considers a stochastic UC where availability of demand response (DR) is uncertain. DR uncertainty was modeled using a set of scenarios, and a chance constraint was imposed to ensure loss-of-load probability is below a specified level. Chen *et al.* [7] considered the combined UC and ED problem under both random and targeted component failures, where allowable loss-of-load was parameterized by the contingency size. Bertsimas *et al.* [8] proposed a two-stage adaptive robust UC model given uncertainty in nodal net injections, and developed a solution based on a combination of Benders decomposition and outer approximation.

Despite recent advances, lack of advanced methods to accurately model and represent uncertainty and the inability of scenario-based approaches to solve industrial sized problems have led researchers to seek alternatives, in both grid operations and planning contexts. For example, Thiam and DeMarco [9] argue: "*Simply put, when uncertainty is credibly accounted for such methods yield solutions for economic*

This work was funded by the Laboratory Directed Research & Development (LDRD) program at Sandia National Laboratories. Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

benefit of a transmission expansion in which the "error bars" are often larger than the nominal predicted benefit." In other words, only a limited number of samples could be considered while sustaining computational tractability, which in turn impacts the ultimate utility of a solution. Of course, it is not possible to change the nature of uncertainties; thus, when uncertainties are sufficiently large, existing methods can fail to provide useful information. However, as our experiments demonstrate, it is possible to significantly reduce the impact of errors introduced by modeling and sampling of uncertainty.

In this paper, we adopt advanced modeling and sampling techniques from the field of uncertainty quantification (UQ), and leverage them to impact power systems operations problems such as stochastic UC and ED. Such techniques have been successfully applied in many areas of computational science and engineering (see e.g., [10] and references therein). The need for accurate estimation of uncertain model outputs, along with the prohibitive cost of requisite large numbers of Monte Carlo (MC) samples, have led to the development of more *efficient alternatives*. Specifically, we employ Polynomial Chaos Expansions (PCEs) [11] to represent uncertain model inputs in terms of orthogonal polynomials of standard random variables. The task of propagating this functional representation to model outputs can be achieved via several means. In this study, we employ a Galerkin projection technique in conjunction with sparse quadrature methods. We demonstrate that our approach yields a *one to two orders of magnitude reduction* in the number of samples (scenarios) required to estimate expected production cost, relative to MC, depending on given target error thresholds. Consequently, our approach has the potential to dramatically reduce the computational difficulty of stochastic grid operations problems, significantly reducing a major barrier to their use in practice. The PCE approach is implemented in conjunction with a novel renewable power scenario generation technique based on Karhunen-Loève expansions (KLEs). This approach takes into account spatial and temporal correlations in the wind magnitude at adjacent wind farms. The practical implementation of the algorithms presented in this paper was done with the help of the UQ Toolkit (UQTk) [12]. UQTk is a collection of software libraries and tools for the quantification of uncertainty in numerical models.

The remainder of this paper is organized as follows. We discuss models for SED formulation in Section II. Section III reviews key concepts in the representation of uncertainty using PCEs and KLEs. Section IV-A explores the feasibility of using KLEs for representing uncertain wind speed profiles and proposes a model for generating wind power samples that is consistent with uncertainties observed in current forecast models. Section IV-B discusses the use of our PCE-based surrogate models for SED followed by empirical convergence results in Section IV-C. Conclusions are presented Section V.

II. Modeling Uncertainties for Stochastic Economic Dispatch

In this paper, we assume that unit commitment decisions for conventional power generation are fixed; thus, we focus on modeling uncertainties for the Stochastic Economic Dispatch model. We model uncertain time-dependent renewable power generation as a random field, and present, in subsequent sections, an efficient approach to represent such random fields. In principle, one can also consider uncertain loads, and model them similarly as random fields [13]. For conciseness, we focus strictly on uncertain renewable outputs and, without loss of generality, presume deterministic loads. Because uncertain renewables generation is represented using random fields and ultimately, as will be illustrated below, as functions of a vector of random variables $\boldsymbol{\xi}(\omega)$, the corresponding minimum production cost $Q(\boldsymbol{x},\boldsymbol{\xi}(\omega))$, with $\boldsymbol{x} \in \mathcal{X}$ a set of fixed unit commitments decisions, is similarly uncertain or random. For completness, we provide in Appendix A the Stochastic Economic Dispatch formulation employed in this paper.

Several wind forecast models have been developed for wind scenario generation in the context of stochastic UC&ED. Physical models, such as numeric weather prediction (NWP), use real time weather data and meteorological techniques for wind forecasting. These models incorporate detailed physical terrain information such as obstacles, local surface roughness, orographic effect, speed up or down, etc [14], [15]; however, the computational tractability of these models remain a challenge. Other wind forecasting techniques rely on conventional statistical models. For example, autoregressive integrated moving average (ARIMA) and its variants, and other statistical models can also be used to generate wind scenarios. ARIMA models use historical data, pattern identification, and combine a non-stationary model and a wide-sense stationary model within an autoregressive forecasting framework [16]. Unlike conventional statistical models, spatial correlation models further take into consideration the spatial relationship of groups of wind sites. Spatial correlation models were also used, in conjunction with neural networks, for wind forecasting, e.g. [17], [18]. Pinson *et al.* [19] present a probabilistic approach that is based on a recursive computation of a covariance matrix. The multivariate Gaussian estimates are then converted to wind power predictions over 2 to 3 days. A study corresponding to a multi-MW wind farm was presented. Finally, [20] and [21] present recent and in-depth reviews of current methods and advances in wind power forecasting and prediction.

In relation to this body of work, our present approach pursues a general representation of uncertain wind generation as a random field, allowing for *arbitrary* auto/cross-correlation structure and probability distributions. This is rendered feasible by the use of an efficient spectral representation of random fields, namely KLE. The KLE is an optimal random field representation, in that it minimizes the mean square error for a given number of terms in the spectral representation of the random field (see Chapter 2.3 in [11]). In addition to temporal correlations in the wind generation from any given wind farm, this approach can also accomodate spatial correlations among different wind farms.

High quality stochastic UC&ED solution requires an accurate representation of the uncertainty, which necessitates the use of a large number of scenarios. However, an increase in the number of scenarios presents considerable computational challenges. A number of scenario reduction techniques have been proposed to permit solution of stochastic UC&ED in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

practical time frames. Various metrics such as probability, hourly magnitudes, distance, and cost have been proposed as similarity measures in clustering algorithms [22] [23]. Unsupervised clustering method, *k*-means and its variants, [24]–[26] can be used to quickly partition a set of scenarios in a given number of clusters, scenarios with similar features are clustered together, typically an original scenario closed to the centroid is used to represent the cluster. Wind scenarios generated via statistical models can also be combined into a single set of scenarios using an ensemble approach in which the predictions of various models are weighted in a manner that preserves the special features [27].

In this paper the SED algorithm relies on a spectral representation for the uncertain minimum cost $Q(\boldsymbol{x},\boldsymbol{\xi}(\omega))$ resulting from the uncertain wind power. Unlike current approaches that rely on MC sampling and subsequent heuristic scenario reduction, the approach presented here is based on PCEs, and it allows the computation of expectations and other moments with fewer samples compared to traditional sampling approaches. Effectively the PCE approach converts the random sampling of the stochastic space to an optimal (in the mean square sense) deterministic sampling.

The KLE approach mentioned above is used to develop a representation of the stochastic wind output as a function of a low dimensional set of random variables (RVs), which in turn makes the minimum cost resulted from the ED problem a function of the same set of RVs. We then use a PCE approach to represent this dependence and use a sparse quadrature approach to compute the expected value more efficiently compared to several Monte Carlo-based algorithms. The PCE approach converts the random sampling of the stochastic space to a deterministic sampling, with sample points chosen according the quadrature rule employed for computing the expectation or higher order moments. In the following section, we will present these two components for efficient representation of uncertainties in the SED problem.

III. Modeling Uncertainties as Random Variables/Random Fields

In this section we provide a brief description of the theory supporting the probabilistic methods we propose for the SED problem. We employ KLEs to construct optimal representations for uncertain wind scenarios. This approach employs the correlation structure present in the wind data to construct compact representations for time-dependent wind scenarios. This representations are optimal in a mean-square sense, and one can employ this property to discard the uncertain dimensions that are less important, and alleviate some of the difficulties associated with sampling high-dimensional stochastic spaces. Once uncertain wind scenarios are constructed based on KLE methods, we then proceed to represent the uncertain minimum production cost using PCEs. Here we treat the cost as a random variable and construct a spectral representation, i.e. a PCE, for this RV as a function of the standard random variables for the uncertain wind scenarios. This approach leads to more accurate estimates for the mean cost compared to MC-based methods, as it will be shown later in this paper.

A. Polynomial Chaos Expansions

Given the formulation of the SED problem (given by (19) and (20) in Appendix A) with uncertain/random loads leading to uncertain/random production costs, we employ efficient UQ methods, specifically PCEs, that rely on *functional representations* of random variables. A brief description of polynomial chaos concepts is presented below. For in-depth descriptions, see [10], [11], [28]–[30].

We begin by setting up a requisite theoretical framework as follows. Under certain technical conditions [30], a random variable (RV) X with finite variance can be written as a PCE:

$$X(\omega) = X(\boldsymbol{\xi}(\omega)) = \sum_{k=0}^{\infty} \alpha_k \Psi_k(\boldsymbol{\xi}) \tag{1}$$

where the basis functions Ψ_k are multivariate polynomials[1] that are orthogonal, by construction, with respect to the density of $\boldsymbol{\xi}$, $p_{\boldsymbol{\xi}}(\boldsymbol{\xi})$. Here $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$ is a vector of n identically-distributed independent RVs. The expectation for any RVs $X(\boldsymbol{\xi})$ is defined by

$$\langle X \rangle = \int X(\boldsymbol{\xi}) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi}. \tag{2}$$

Given the orthogonality of the basis functions, we have

$$\langle \Psi_i \Psi_j \rangle = \int \Psi_i(\xi) \Psi_j(\xi) p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) d\boldsymbol{\xi} = \delta_{ij} \langle \Psi_i^2 \rangle \tag{3}$$

and

$$\alpha_k = \langle X \Psi_k \rangle / \langle \Psi_k^2 \rangle, \tag{4}$$

where δ_{ij} is Kronecker's delta.

The multivariate polynomials Ψ_k are products of univariate polynomials, namely $\Psi_k(\boldsymbol{\xi}) = \psi_{k_1}(\xi_1) \cdots \psi_{k_n}(\xi_n)$. Given that the RVs $\{\xi_1, \xi_2, \ldots, \xi_n\}$ are independent, it follows that the density $p_{\boldsymbol{\xi}}(\boldsymbol{\xi})$ is equal to the product of densities of each RV, $p_{\boldsymbol{\xi}}(\boldsymbol{\xi}) = p_{\xi_1}(\xi_1) p_{\xi_1}(\xi_1) \ldots p_{\xi_n}(\xi_n)$. Each family of polynomials corresponds to a given choice of distribution for the ξ_i. The most used choices for $(\boldsymbol{\xi}, \Psi)$ are normal RVs with Hermite polynomials and uniform RVs with Legendre polynomials.

In a practical computational context, one truncates the PCE to order p. The number of terms in the resulting total-order finite PCE

$$X \approx \sum_{k=0}^{P} \alpha_k \Psi_k(\boldsymbol{\xi}) \tag{5}$$

is given by $P+1 = (n+p)!/n!p!$. We dispense with the $\approx$ symbol in the remainder of this paper, employing the equal sign for any truncated PCE.

A key utility of polynomial chaos in UQ is the efficient propagation of uncertainty from model inputs to outputs. Consider a model $Y = f(X)$, where X and Y are random variables, representing uncertain model inputs and outputs, respectively. Given a known specification for X, the *forward propagation of uncertainty*, or forward UQ for short, involves the estimation of the corresponding specification for Y. Clearly, one way to do this is to generate random samples of X from its given density $p(X)$, evaluate the corresponding random samples

[1] Generally, other, non-polynomial basis functions can be used, but here we restrict ourselves, without loss of generality, to the most common polynomial-based usage.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

of Y, and bin them to arrive at an estimate of the density $p(Y)$. This Monte Carlo (MC) based method for forward UQ, while robust, can require prohibitively large numbers of samples for accurate estimation of the moments or density of Y. In principle, polynomial chaos methods, by relying on functional representations of random variables rather than their densities, provide potentially more efficient means of forward propagation of uncertainty in a wide class of problems, as is further discussed below.

Several polynomial chaos-based UQ approaches have been developed. In this paper we employ "non-intrusive" methods where, given a PCE of model inputs $X = \sum \alpha_k \Psi_k(\boldsymbol{\xi})$, the estimation of the PCE coefficients $\{\beta_k\}_{k=0}^{P}$ of the model output, $Y = \sum \beta_k \Psi_k$, relies on the numerical evaluation of the Galerkin projection integrals, namely $\beta_k = \langle Y, \Psi_k \rangle / \langle \Psi_k, \Psi_k \rangle$.

The evaluation of these integrals can be done, of course, using random sampling methods, including MC and a wide range of quasi-MC (QMC) methods. However, the convergence rate of MC/QMC methods is generally slow, resulting in poor computational performance. Often, a more efficient alternative, which relies on a degree of presumed smoothness in the integrand, is the use of sparse quadrature, relying on a number of *deterministic* samples of the forward model. Using N_q quadrature points $\boldsymbol{\xi}_i$, with associated quadrature weights w_i, the PCE coefficients can be computed as

$$\alpha_k = \frac{\langle X \Psi_k \rangle}{\langle \Psi_k^2 \rangle} = \frac{1}{\langle \Psi_k^2 \rangle} \sum_{i=1}^{N_q} w_i X(\boldsymbol{\xi}_i) \Psi_k(\boldsymbol{\xi}_i) \tag{6}$$

Sparse quadrature methods [31]–[33] provide significant efficiency gains over random sampling methods for low/moderate dimensional systems, and require fewer model evaluations compared to full-tensor quadrature. Fig. 1 shows sample grids corresponding to Clenshaw-Curtis [34] and Gauss-Kronrod-Patterson [35] rules, respectively. The sparse quadrature approach requires fewer model evaluations compared to full-tensor quadrature and helps delay the curse-of-dimensionality.

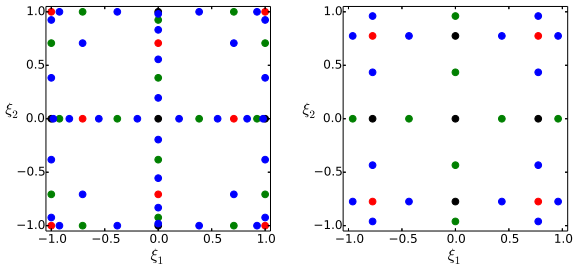

Fig. 1. Sample sparse grid constructions using a Clenshaw-Curtiss rule (left frame) and Gauss-Kronrod-Patterson (right frame). Black, red, green, and blue markers show quadrature points corresponding to progressively higher order quadrature rules.

We discuss below a simple illustrative 2D example, with inputs (ξ_1, ξ_2), and output $X(\xi_1, \xi_2)$. Here, both ξ_1 and ξ_2 are considered uniform in $[-1, 1]$. We employ the function $X(\xi_1, \xi_2) = (1 + (\xi_1 - 0.1)^2 + (\xi_2 + 0.1)^2)^{-1}$, as a black-box input-output map. The function shape over the range for the input random variables ξ_1 and ξ_2 is shown in the left frame of Fig. 2, and the PCE coefficients up to 10-th order are shown in the right frame. These coefficients are computed via sparse quadrature using a Clenshaw-Curtis rule. Approximately 1500 model evaluations are necessary to compute the coefficients for a 10-th order expansion with relative error of $o(10^{-6})$ for each coefficient, while, for a 6-th order expansion, the number of model evaluations for the same order of magnitude is about 700. We now illustrate a convergence rate check for the first

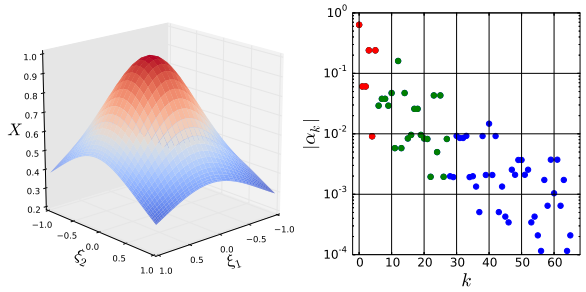

Fig. 2. Sample 2D test function (left frame) and absolute values of PCE coefficients α_k (right frame): 2nd order expansion (red), 6-th order (red+green), and 10-th order (all).

two moments, the expectation and variance, of X, given by

$$E[X] = \alpha_0, \quad V[X] = \sum_{k=1}^{P} \alpha_k^2 \langle \Psi_k^2 \rangle.$$

We estimate the relative error in the estimate of $E[X]$ as

$$E_{\mathrm{PCE},i} = |\alpha_{0,i} - \alpha_{0,i+1}| / \alpha_{0,i+1} \tag{7}$$

where index i indicates the sparse quadrature level. A similar estimate is employed to check for the accuracy of $V[X]$. The left frame in Fig. 3 shows a near exponential convergence rate for the $E[X]$. As might be expected, a larger number of model evaluations is required for the same requisite error in the second moment of X compared to its first moment. In the right frame of the same figure we present a comparison between the "true" PDF of X, $p(X)$, and PDF profiles obtained with different PCE orders. These PDFs were computed via MC sampling, using a large number, 10^7, of either model evaluations for the "truth" (in black), or PCE evaluations. The PCE results are converged, with no discernible change in the PDF shape, at sufficiently high order. We required about 65 sparse Clenshaw-Curtis quadrature points for 2^{nd} order PCE, ~ 700 for 6^{th} order, and ~ 1500 for 10^{th} order. Comparing the two plots, we note that while 65 model evaluations are sufficient to capture the expectation and variance within $o(0.1-1)\%$, the PDF for the equivalent 2^{nd} order PCE is very different from the true PDF. While low order moments can be estimated employing lower-order PCEs, depending on the model at hand, higher-order PCEs may be needed for an accurate picture of higher-order moments.

B. Karhunen-Loève Expansions

In this section we expand the theory fundamentals from spectral representations of random variables to representations tailored to random fields, or stochastic processes. Generally, a random field/stochastic process is a function that depends on both random and deterministic inputs. Thus, in the present

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

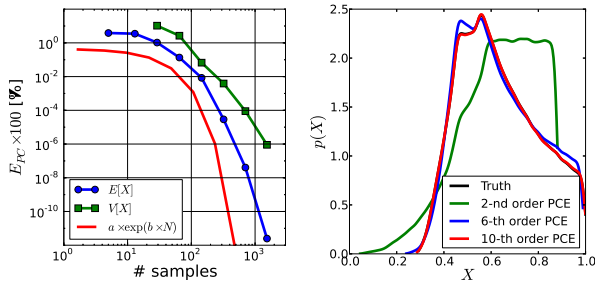

Fig. 3. Left frame: Convergence of $E[X]$ and $V[X]$ with increasing number of model evaluations; Right frame: convergence of probability density function of X, $p(X)$, with increasing PCE order.

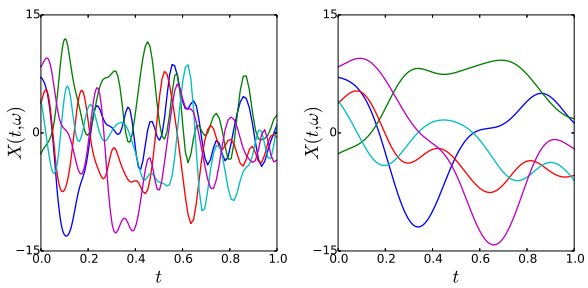

Fig. 4. Sample 1D stochastic process realizations corresponding to a square exponential auto-correlation function with $l_c = 0.05$ (left frame) and $l_c = 0.2$ (right frame).

context, uncertain wind generation is modeled as random function of time. More formally, consider a bounded domain $D \subseteq \mathbb{R}^n$ and the sample space Ω, then a stochastic process X, defined as the mapping $X : D \times \Omega \to \mathbb{R}$, is essentially an infinite collection of random variables X_t, with $t \in D$. Similar to the previous section, we presume all X_t have finite variance.

The stochastic process X is centered if the expectation $E[X_t] = 0$ for all $t \in D$. If X is not centered, then one can always subtract the expectation $(X_t - E[X_t])$ and follow the derivations below for the resulting centered stochastic process. The autocorrelation function C of the stochastic process X is defined as

$$C(t,s) = E[X_t X_s], \quad \forall t,s \in D.$$

It can be shown [29] that C admits the spectral decomposition

$$C(t,s) = \sum_{i=1}^{\infty} \lambda_i u_i(t) u_i(s)$$

with eigenvalues λ_i and eigenfunctions u_i.

The KLE for the stochastic process $X(t,\omega)$ is given by

$$X(t,\omega) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} u_i(t) \xi_i(\omega) \tag{8}$$

where the random variables $\xi_i(\omega)$ have zero mean, unit variance, and are mutually uncorrelated, $E[\xi_i \xi_j] = \delta_{ij}$. Further, if X is a Gaussian process, then ξ_i are normal random variables and hence independent. In general, when ξ_i are non-standard (*e.g.* when not uniform, normal, etc) one can express these random variables as PCEs. The KLE representation is *optimal*, corresponding to minimal least-squares error among all possible choices for random fields representation.

We present below KLE examples for zero-mean 1D Gaussian processes. These cases were constructed using correlation functions given by $C(t,s) = \exp(-(t-s)^2/l_c^2)$, where l_c is the correlation length. One-dimensional stochastic process realizations for $D = [0,1]$ are presented in Fig. 4 for $l_c = 0.05$ and 0.2, respectively. Fig. 5 shows, in the left frame, the first few scaled eigenmodes, $\sqrt{\lambda}u$ for $l_c = 0.2$. The eigenvalue decay is shown in the right frame for several values of the correlation length. The sharp decay for $l_c = 0.2$ is consistent with the fact that, for relatively large correlation lengths, the first few terms on the right-hand side of Eq. (8) are expected to be sufficient to represent the stochastic process accurately. Conversely, for smaller correlation lengths, *e.g.* $l_c = 0.05$ in this series of examples, the number of terms in the expansion is much larger for the same requisite accuracy.

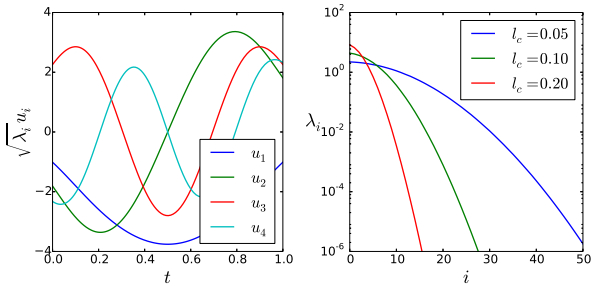

Fig. 5. Illustration of the first 4 scaled KL modes for a stochastic process with a square exponential covariance matrix with $l_c = 0.2$ (left frame), and decay of eigenvalue magnitude for several correlation lengths (right frame).

Here, we presented results for examples which employed analytical autocorrelation functions. In Section IV-A we will examine, based on available historical wind data, the actual correlation structure at several wind sites. In this context, with correlation structure learned from data, it is clear that the KLE eigenmodes are the same as those evaluated from principal component analysis (PCA). With this shared eigenstructure, the PCA is a compact representation of the *data*, while the KLE is a compact representation of the *random process* behind the data.

IV. Results

Accurate, efficient, and low-dimensional representations of uncertainties are essential for the success of stochastic grid operations models. We now discuss how to construct a lower dimensional representation of wind power uncertainty. Towards this goal, we explore representing uncertain wind production time profiles $p_r^t(\boldsymbol{\xi})$ via KLEs [29] in Section IV-A. Section IV-B discusses the use of our PCE-based surrogate models for SED given the stochastic space spanned by the model for wind renewables. We conclude the results section with empirical convergence results in Section IV-C.

A. Wind Power Forecast Models

In order to assess the feasibility of the KLE approach for representing wind generator output, we consider data for three wind sites extracted from NREL's Western Wind Dataset [36]. Two of these sites, #15414 and #16238, are located in Wyoming and are geographically close. The third

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

site, #3560, located in California, was selected to be far from the other two. The proximity of the two Wyoming sites allows us to explore means of accounting for correlations in wind power, while the third site is sufficiently distant from them such that any correlation in their production is negligible. For each site, wind speed and power data is available at

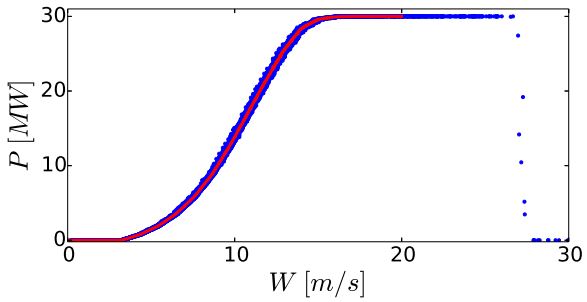

Fig. 6. Rated power output P vs. wind speed W at site #15414 for year 2004. Blue dots represent instantaneous measurements while the red curve represents power data filtered with a top-hat filter of width $\Delta_W = 0.05$ [m/s], and interpolated via cubic splines.

10 minute intervals for the years 2004 through 2006. Fig. 6 shows a scatter plot of rated power output vs. wind speed at site #15414. The rated power output is zero for wind speeds smaller than a threshold value of approximately 3.2 m/s. At speeds greater then approximately 26 m/s the generators are turned off, for safety reasons.

We take the following steps to post-process the NREL wind data into wind samples for use in our SED model. We first construct hourly averages for the wind speed; and consider the resulting hourly-average wind speed data for each day, over a specified seasonal/monthly date-range, to be an independent sample from a 24-dimensional random field. Choosing the date-range, we assemble an associated set of 24-dimensional samples. For example, considering the month of January for 2004 through 2006 leads to 93 samples. This approach was adopted to account for seasonal changes in wind patterns. Fig. 7 shows select daily log-transformed ($W_L = \log(W)$) wind speed samples at site #15414. We adopted this transformation to ensure positivity of wind speed samples, generated via the algorithm described below.

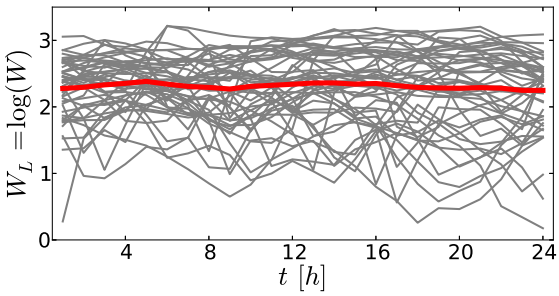

Fig. 7. Select daily wind data from January 2004-2006 at site #15414. The red curve corresponds to the mean of these daily samples.

We subtract the mean values and represent the resulting centered process using the KLE, which, given the 24-long discrete structure of the time-data is at most 24-dimensional,

$$W_L(t,\omega) - \langle W_L(t,\omega)\rangle = \sum_{i=1}^{24} \sqrt{\lambda_k} u_i(t)\xi_i(\omega) \qquad (9)$$

where $\langle W_L(t,\omega)\rangle$ denotes the mean of $W_L(t,\omega)$.

Were it known, the covariance matrix of $W_L(t,\omega)$, Σ_{W_L}, could have been specified analytically, as in the examples provided in Section III-B section. Alternately, if sufficient 24-h samples are available, the 24^2 covariance matrix can be estimated from these realizations. Here, we estimate the covariance matrix using the daily samples for the select data range in $2004-2006$. Once the covariance matrix is available, one can compute [29] the eigenvalues and eigenvectors in Eq. (9).

Fig. 8 shows the first two KLE modes corresponding to the month of January at the three wind sites considered in this study. The large degree of similarity for these modes between the three sites is reflective of diurnal wind patterns. We will further explore later in this section the structure of the covariance matrices Σ_{W_L}. Next, we examine the degree

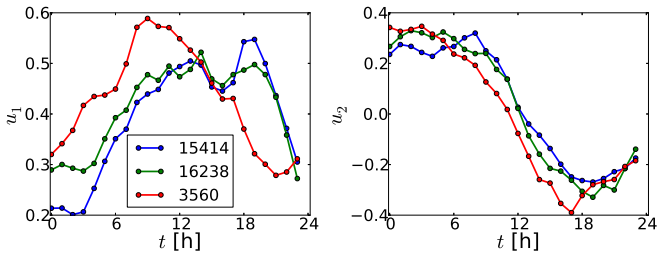

Fig. 8. The first two KLE modes, u_1 and u_2, at several wind sites based on data from January 2004-2006.

of dependence between KLE random variables at the same site. These RVs are uncorrelated by construction, but not necessarily independent. In a general setting one can further model these RVs as polynomial chaos expansions in terms of standard RVs to enable standard techniques for quantifying uncertainties. In the case at hand, we first construct empirical cumulative distribution functions (CDFs) from RV samples. These samples are obtained by projecting each daily wind speed sample onto the corresponding KLE mode. Fig. 9 shows the empirical CDFs for ξ_1 through ξ_{15} at two select wind sites. Visual comparison of these results with the CDF of a standard normal RV indicate strong similarities. To simplify the demonstration of the KLE approach for modeling wind uncertainties and without loss of generality, we model these RVs as standard normal RVs. Since they are uncorrelated by construction, this implies they are also independent.

We now investigate relationships among the standard normal RVs across the sites we considered here. To this end we employ distance correlation factors [37], [38] between pairs of RVs corresponding to the same KLE mode at different sites. These factors are shown in Table I. Smaller values, towards 0, indicate negligible dependence between pairs of RVs, while larger values, close to 1, indicate a strong dependence. These results indicate a strong dependence between sites #15414 and #16238 for the first two KLE modes, while the same KLE mode at site #3560 shows little dependence with the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

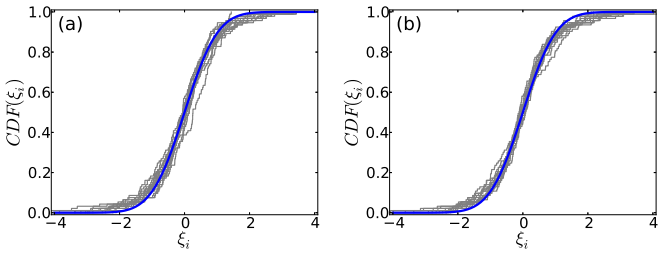

Fig. 9. Empirical CDFs (depicted in gray) for ξ_1 through ξ_{15} at (a) site 16238 and (b) site 3560. The CDF for a standard normal random variable is shown in blue, for reference.

TABLE I
DISTANCE CORRELATION FACTORS FOR ξ_1 - ξ_4 BETWEEN WIND SITES.

RV	16238-15414	16238-3560	15414-3560
ξ_1	0.91	0.24	0.24
ξ_2	0.84	0.13	0.18
ξ_3	0.67	0.18	0.19
ξ_4	0.53	0.23	0.22

first two sites. The third and fourth RVs, as well as the other RVs (not shown), show little dependency between sites. This is somewhat to be expected given the nature of turbulence. Specifically, low order KLE modes are associated with large scale structures which are likely similar if sites are geographically situated, while higher order KLE modes are associated with smaller scale structures with much faster eddy turnover times.

Because the stochastic dimensions are uncorrelated, the total variance of the stochastic process is given as the sum of variances from individual KLE modes. Given that eigenvectors are orthonormal with respect to the deterministic space (the discretized time axis in our study), and that random variables ξ_i have unit variance, it follows that:

$$\int_T \mathrm{Var}[W_L(t,\omega)]dt = \sum_{i=1}^{24} \lambda_i. \tag{10}$$

As a result, an N-truncated expansion, $N \leq 24$, will explain

$$100 \times \left(\sum_{i=1}^{N} \lambda_i\right) \Big/ \left(\sum_{i=1}^{24} \lambda_i\right) \tag{11}$$

of the total variance of the random field. Fig. 10 shows the dependence of the fractional variance given in Eq. (11) on the number of terms N in the truncated KLEs at the three wind sites. It is evident that for all sites, $N = 6$ modes are sufficient to capture approximately 95% of the total variance in the daily wind profiles.

In experiments described below, we consider a stochastic space with $N = 6$ dimensions at each site. Given that we model the first two modes at two sites as dependent, the total dimensionality of the stochastic space is $3 \times 6 - 2 = 16$. Given the representation of daily wind profiles as truncated KLEs, it follows that renewable generation $p_r^t(\boldsymbol{\xi}) = f(\exp(W_L(t,\boldsymbol{\xi}))$ is a function on the same stochastic space. We neglect the (small) noise in the conversion of wind speed into power and approximate $f(\exp(W_L(t,\boldsymbol{\xi})))$ as a cubic spline interpolation through the filtered rated power data. This approximation is

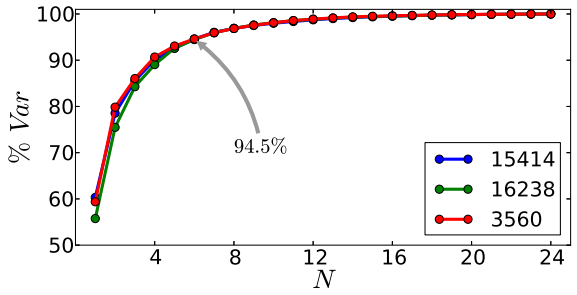

Fig. 10. Percentages of the total variance explained by truncated KLEs with an increasing number of modes, for our three wind sites.

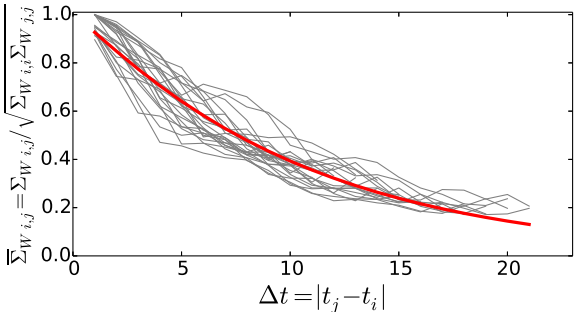

Fig. 11. Decay of covariance matrix Σ_W components, shown with thin grey lines, with increasing time lag $|t_j - t_i|$, anchored at several time instances i for site #16238. Values are normalized by the diagonal entries. Matérn covariance model, fitted by least squares, is shown with red line.

depicted by the red line in Fig. 6.

So far in this section we proposed an approach for representing daily wind profiles via KLE based on historical records. However, this leads to a range of uncertainties that encompasses past observations, and this range is significantly larger than typical uncertainties associated with short-term weather forecasts, e.g. day-ahead predictions. Below, we propose to adapt the KLE to be consistent with level of uncertainties consistent with typical weather forecasts.

We first start by parameterizing the covariance matrices corresponding to each wind farm site. Fig. 11 shows, with thin grey lines, a typical decay in the components of Σ_W with increasing time lag. We normalize these matrices by the magnitude of the diagonal elements and use a Matérn covariance model to fit the normalized covariance matrices at each wind site kernel [39]

$$\overline{\Sigma}_W(\Delta t) = \frac{2^{1-\nu}}{\Gamma(\nu)} \left(\frac{\sqrt{2\nu}\Delta t}{l_t}\right) K_\nu \left(\frac{\sqrt{2\nu}\Delta t}{l_t}\right), \tag{12}$$

where ν and l_t are positive parameters, $\Gamma(\cdot)$ is the gamma function, and $K_\nu(\cdot)$ is the modified Bessel function of the second kind. The Matérn kernel offers more flexibility for modeling covariances compared to, for example, exponential and square exponential forms which are particular cases of the Matérn kernel, for $\nu = 1/2$ and $\nu \to \infty$, respectively. Table II shows the parameters of the Matérn kernels for the three sites employed in this study. These values were obtained by least-squares fitting, using the empirical correlation data from 2004-2006. We obtain similar values for the two sites (15414 and

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

TABLE II
PARAMETERS OF MATÉRN KERNELS FOR SELECTED WIND SITES.

Wind Site	l_t	ν
15414	11.40	0.56
16238	11.15	0.57
3560	9.79	0.78

16238) that are geographically close.

Based on the formulations presented in this section we propose the following algorithm to generate wind power samples that are consistent with historical wind characteristics at the selected wind sites:

1) Compute normalized covariance matrices based on historical data at each site.
2) Fit Matérn covariance models through the normalized covariance kernels.
3) Forecast day-ahead wind profiles at selected wind sites and estimate the uncertainty in the day-ahead forecast.
4) Scale the normalized covariance matrices, perform the eigen-decomposition and generate wind samples via Eq. (8) for select samples of $\boldsymbol{\xi}$. Here $\boldsymbol{\xi}$ is a vector of standard normal random variables as estimated above in this section.
5) Convert wind samples into wind power values, see Fig. 6.

We included step (2) in the algorithm above for the purpose of comparing the characteristics of wind profiles across different sites. In general, this step can be skipped and one can use the normalized covariance matrices computed from data. In the results presented in the next section, we randomly picked daily wind profiles, from the data available at each site, as surrogates for the day-ahead forecast. Since data on typical uncertainties in the day-ahead weather forecasts at the above sites was not immediately available to us, we proceeded to estimate uncertainties in the wind forecasts given typical available data for wind power from alternative data sources. Specifically, we used data for day-ahead forecast and actual realizations obtained from the Belgium Electricity Grid Operator ELIA [40]. This data is available only for the aggregated wind power, for either onshore or offshore wind farms. We considered the aggregate wind power values for the land-based wind farms and computed the coefficient of variation (CV) for hourly wind power output based on data for years 2012 through 2015. We found a nearly constant value, $CV(P_W) \approx 0.35$ for the day-ahead predictions. We then employ this value and the rated power curve for each site to backtrack uncertainties in day-ahead wind speed forecasts $\sigma_W^2 = \sigma_W^2(CV(P_W))$. These, in turn, are used as scale factors for covariance matrices at each wind site.

B. Accurate estimation of expected cost with limited samples

We now return to the evaluation of the expected cost (Eq. (18) in Appendix A), corresponding to the SED problem defined by (19) and (20), given a KLE representation for renewable generation. We can *estimate* the expected production cost by using a finite number of renewable power realizations (i.e., scenarios) $\boldsymbol{\xi}_s \in \mathcal{S}$ sampled from the joint density $PDF(\boldsymbol{\xi})$. For our current example, each component $\boldsymbol{\xi}$ is a standard normal RV, hence the sampling can be done independently in each stochastic direction. Defining $\rho \equiv 1/|\mathcal{S}|$, where $|\mathcal{S}|$ is the cardinality of $\mathcal{S}$, the SED in Eq. (18) can be rewritten as:

$$\min_{\boldsymbol{f},\boldsymbol{p},\boldsymbol{q},\boldsymbol{\theta}} \quad \rho \sum_{s\in\mathcal{S}} Q(\boldsymbol{x},\boldsymbol{\xi}_s) \tag{13}$$

where $Q(\boldsymbol{x},\boldsymbol{\xi}_s)$ is the solution of Eqs. (19) and (20) for a particular instance of the renewable generation $p_r^t(\boldsymbol{\xi}_s)$.

Formulation (13) represents an *extensive form* of the SED problem, based on $|\mathcal{S}|$ sampled scenarios from the stochastic space corresponding to wind power generators. The typical scenario sampling approach described above uses MC sampling to approximate an integration, thereby estimating an expectation. While MC algorithms are commonly used for their convenience and robustness, their slow convergence rate is well-known. The MC estimate of the expectation has error

$$\mathrm{V}[Q(\boldsymbol{x},\boldsymbol{\xi})]/\sqrt{|\mathcal{S}|}, \tag{14}$$

where V[Q] denotes the variance of the RV Q. Given the significant additional complexity incurred by including stochasticity in the optimization problem, a stochastic formulation becomes advantageous relative to a deterministic formulation when the variance is large. Hence, accurate estimation of the expectation is not only an academic exercise but is important in practice.

According to Eq. (14), accuracy can be achieved with more samples. However, a linear decrease in error requires a quadratic increase in the number of samples, which can quickly render the stochastic optimization problem intractable. This illustrates the limitation of MC algorithms for accurate estimations; while they are convenient, they are not efficient.

Alternatively, we propose an approach based on Polynomial Chaos Expansions that can enable high fidelity uncertainty quantification with fewer samples, as shown in Section III-A. In this context, with uncertain renewable wind generation dependent on a stochastic space of *iid* standard normal random variables, $\boldsymbol{\xi} = \{\xi_1, \xi_2, \ldots, \xi_n\}$, we employ Hermite polynomials to construct a truncated Hermite-Gauss (HG) PCE for the (minimum) cost $Q(\boldsymbol{x},\boldsymbol{\xi})$,

$$Q_{\mathrm{PCE}}(\boldsymbol{x},\boldsymbol{\xi}) = \sum_{k=0}^{P} \alpha_k(\boldsymbol{x})\Psi_k(\boldsymbol{\xi}), \tag{15}$$

where $\Psi_k(\boldsymbol{\xi})$ are n-variate Hermite polynomials.

The coefficients α_k depend on the discrete variable $\boldsymbol{x}$, hence separate PCE approximations for Q would be constructed for each instance of $\boldsymbol{x}$ if this approach is to be extended to a stochastic unit commitment model. We compute these coefficients by Galerkin projection, in Eq. (4). The dimensionality of the PCE in Eq. (15) is the same as the number of stochastic dimensions used to represent the uncertain renewable power. In this paper we used $n = 16$, as mentioned in Section IV-A.

Given $Q_{\mathrm{PCE}}(\boldsymbol{x},\boldsymbol{\xi})$, we have

$$\overline{Q}(\boldsymbol{x}) = \mathrm{E}_{\boldsymbol{\xi}}[Q(\boldsymbol{x},\boldsymbol{\xi})] = \alpha_0(\boldsymbol{x}), \tag{16}$$

being the solution of the SED problem. The PCE in Eq. (15) can also be used to generate higher-order moments for the minimum cost as illustrated in Section III-A.

Several methods can be employed to evaluate the projection

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

integrals, see Eq. (4), for the PCE coefficients α_k in Eq. (15). As discussed in Section III-A, MC methods can be used in principle, but are impractical given their slow convergence rate. Alternatively, for smooth integrands, and particularly in low-moderate dimensional problems, sparse quadrature methods [31], [41] can provide highly accurate results with smaller numbers of deterministic samples. For this study we use a sparse grid based on Gauss-Kronrod-Patterson quadrature [35]. We will show below that projection integrals require much fewer sparse-quadrature evaluations compared to the number of MC samples for a given requisite accuracy.

C. IEEE 118-bus Results

We now present numerical results comparing the expected costs of our SED model using scenarios obtained with both our PCE approach and with existing sampling-based approaches. Specifically, we will employ Monte Carlo, including a scenario reduction method, and quasi Monte Carlo methods. We consider the IEEE 118-bus test system [42] augmented with the three wind generation sites discussed in Section IV-A. These renewable generators replace the conventional generators at buses 89, 69, and 10, respectively. We employ $|T| = 24$ time periods and represent stochastic renewable generation via the KLEs presented in Section III-B.

Given the analysis presented in Section IV-A, the effective dimensionality of the stochastic space is reduced to a fraction $f_t \times |T|$ for each wind site. The aggregate dimensionality for several wind sites can be further reduced by accounting for spatial dependencies for wind farms that are geographically close. For a cluster of n_k wind farms for which a fraction $f_{s,k}$ of KLE modes are dependent, the total number of stochastic dimension would be

$$((n_k - 1)(1 - f_{s,k}) + 1) \times f_t \times |T|$$

This approach is particularly effective when several sites are clustered together, i.e. they share a number of modes in the KLE model for time-dependent wind speed profiles. For the proof-of-concept example presented in this section we have $n_1 = 2$ Wyoming wind farms and one, $n_2 = 1$, California wind farm. The analysis presented in Section IV-A gives $f_t = 1/4$, and $f_{s,1} = 1/3$ ($f_{s,2}$ is not relevant here since $n_2 = 1$). The resulting effective dimension of the stochastic space is 16.

We first explore the dependence of $Q(\boldsymbol{x}, \boldsymbol{\xi})$ on the stochastic space that characterizes renewable generation. Fig. 12 shows 2D slices through the 16-dimensional $\boldsymbol{\xi}$ space. All slices are anchored at the origin $\boldsymbol{\xi} = 0$. RVs ξ_1 through ξ_6 correspond to site 16238, while ξ_1, ξ_2, and ξ_7 - ξ_{10} correspond to site 15414. The remaining random variables, ξ_{11} - ξ_{16}, correspond to site 3560. The expected cost is about \$2.4 million, and the relative change about the mean is about 15% in the numerical tests presented here. The results shown in Fig. 12 indicate that Q is strongly dependent on ξ_1, while the dependence on ξ_2 is weaker – as suggested by smaller contour changes in that direction. The slice in the (ξ_1, ξ_{11}) plane highlights the contribution of ξ_{11} to the variation of Q. Other 2D slices (not shown) confirm the diminishing impact that higher-order modes have on Q. Collectively, these results also indicate that

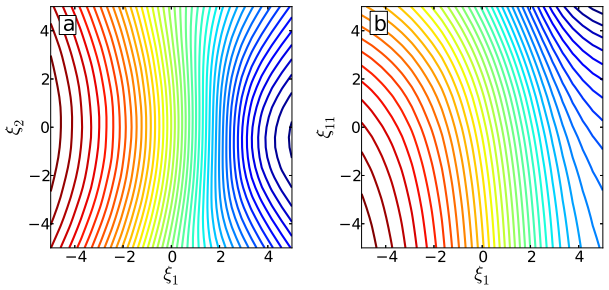

Fig. 12. Two dimensional slices through the center of the stochastic space, $\boldsymbol{\xi} = 0$. Random variables corresponding to (a) first two modes for Wyoming sites and (b) first modes for WY and CA sites. The contour lines correspond to iso-levels of Q, increasing from blue to yellow to red.

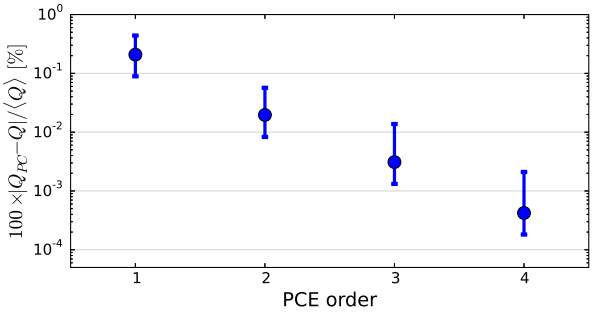

Fig. 13. Cross-validation of the PCE approximation for the minimum cost $Q(\boldsymbol{x}, \boldsymbol{\xi})$. Symbols show the mean L_1 error for each PCE order, while the error bars represent two standard deviations below and above the mean.

$Q(\boldsymbol{x}, \boldsymbol{\xi})$ is smooth in $\boldsymbol{\xi}$, which makes it amenable to a PCE representation in terms of polynomials.

We next proceed to test the accuracy of PCE representations, from 1st to 4th order, for $Q(\boldsymbol{x}, \boldsymbol{\xi})$ with respect to actual model evaluations. Given the 16-dimensional stochastic space corresponding to the three wind sites, the number of sparse quadrature sample points necessary for the estimation of the PCE coefficients increases from about 50 for a 1st order expansion to about 2×10^5 for a 4th order expansion. We cross-validate our PCE representations relative to model solutions at 5×10^5 randomly chosen $\boldsymbol{\xi}$ samples. Fig. 13 shows the magnitude of the L_1 errors between actual model evaluations and the PCE-approximated cost. This error is relative to the expected minimum cost computed using all model evaluations available and is converted to percentages. The errorbars represent two standard deviations based on sample values below and above the mean L_1 error, respectively.

These results indicate a drop of about one order of magnitude, per polynomial degree, for the magnitude of the error between full model evaluations and the PCE results.

The above analysis explores the dependence of minimum cost for individual samples on the stochastic space corresponding to the uncertain wind at select sites, and the accuracy of the PCE approach in capturing this dependence. However, in the context of the SED problem we are only interested the accuracy of the expected minimum cost $\langle Q(\boldsymbol{x}, \boldsymbol{\xi}) \rangle$.[2] Hence,

[2]Our approach is not restricted to expected costs; other moments of Q can be similarly considered.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TPWRS.2016.2615334, IEEE Transactions on Power Systems

our goal here is to demonstrate the efficiency of a quadrature approach to estimate the expectation.

Toward this goal, we show the convergence of $\langle Q(\boldsymbol{x},\boldsymbol{\xi})\rangle$ as a function of the number of samples considered. The results shown with blue in Fig. 14 correspond to the relative error, in Eq. (7), between successive sparse quadrature evaluations for the first PCE coefficient. Here subscript i denotes the quadrature level, e.g., 33 model evaluations for level 1, 513 model evaluations for level 2, and so on.

In this figure we compare the PCE approach with results obtained with MC-based techniques. In the first MC approach labeled "MC" in the figure, scenarios are generated using random samples drawn from the underlying distribution of the RVs that control the wind power uncertainties. We augment the MC approach with scenario reduction techniques as follows. For each set of samples, we employ the forward selection algorithm [25], [26] to reduce the number of samples by half. The scenario reduction approach, labeled "MC-sr" in the figure, removes the MC samples that are "clustered" together, thus improving the coverage of the stochastic space with a smaller set of samples. For example, starting with a set of $n = 2\times 10^2$ scenarios, we downselect to $n = 10^2$ and then redistribute the weights from discarded samples to the ones that are retained. Statistics based on the reduced sample set are shown at position corresponding to 10^2 samples in Fig. 14. The algorithm is repeated for several realizations and sample sizes as shown in the figure.

We also employ quasi MC (QMC) sequences [43] to sample the stochastic space. The QMC approach employs low-discrepancy deterministic sequences as opposed to pseudo-random samples employed by Monte Carlo algorithms. The QMC technique generates samples that enhance samples uniformity (somewhat similar to scenario reduction methods) but are, in general, challenged by dimensionality.

The errors for the MC, MC-sr, and QMC results are computed using a similar approach as for the PCE results. However, because results for MC and MC-sr depend on randomly drawn sets of samples, the mean values are averaged over several realizations. For QMC we randomize the stochastic space ordering to generate several realizations. The specific formula is

$$E^j_{*,i} = |\overline{Q}^j_i - \overline{\overline{Q}}_{i+1}|/\overline{\overline{Q}}_{i+1} \tag{17}$$

where the subscript i denotes the number of samples used to compute the average cost $\overline{Q}^j_i$, e.g., 10 samples for $i=1$, 100 samples for $i=2$, etc. Here "*" stands for "MC", "MC-sr", and "QMC". We employed 10 realizations, indexed by superscript j, for each set of samples; the average cost over all realizations is denoted using a double overline. The dashed lines show power-law curve fits, $a\times N^{-b}$, for the data, which we use to analyze convergence rates as a function of the number of samples.

For the MC approach we recover the theoretical convergence rate, $b=0.5$, while for the PCE approach we obtain a much stronger convergence rate, $b=1.9$. The mean "MC-sr" results are essentially the same as "MC". The scenario reduction approach acts to reduce the variance of results, and the standard deviation in the "MC-sr" is about $75-80\%$ for all sets of samples shown in the figure. The QMC-based results show lower accuracy for smaller number of scenarios, but its accuracy is better than the plain MC-based results for larger number of scenarios.

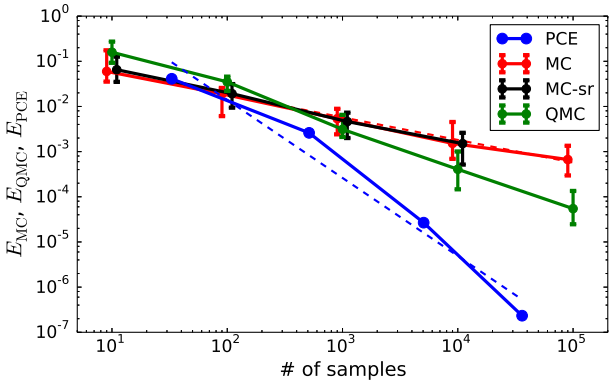

Fig. 14. Comparison of errors in the estimation of expected cost between PCE-based results, Monte Carlo (including with scenario reduction), and Quasi Monte Carlo estimates. Dashed lines show power-law fits, $a\times N^b$ to test the convergence rates of the MC and PCE approaches.

In this figure a 1% error level corresponds to about \$24K, while a 0.01% error level is in the hundreds of dollars. These results indicate that if a low degree of accuracy, *e.g.* 1%, is sufficient or desired, then the PCE approach exhibits a computational cost that is comparable to the MC approach. Nevertheless there is a considerable spread in the accuracy of these estimates, making MC estimates with a small number of samples somewhat unreliable. For situations where higher accuracies are required, for example of 0.01%, the PCE estimates require one to two orders of magnitude less model evaluations compared to the MC estimates.

V. Conclusion

In this manuscript we presented a proof-of-concept demonstration of efficient representation of uncertainty for the Stochastic Economic Dispatch problem. The proposed approach combines dimensionality reduction based on spatial and temporal dependencies followed by an efficient sampling of the reduced stochastic space. It is important to note that while we only look at three wind farms, the underlying dimensionality of the stochastic representation is much larger, proportional to the length of the time horizon employed in the study (currently 24h). For each wind site we cut down the dimensionality by a factor of 4 by considering a spectral representation Karhunen-Loève expansions (KLEs) for the wind speed. We were able to further cut the dimensionality by accounting for spatial correlations between wind sites that are geographically close.

In the second part, we reduce the computational costs of stochastic economic dispatch, by reducing the required number of samples. This approach is based on Polynomial Chaos expansion (PCE) models for the production cost that cover the uncertainty in renewable generation. The construction of the PCE terms is based on the projection of the model on increasingly higher basis modes. Consequently, the global mean square error between the surrogate model and the actual

simulations is easily controlled. Our experiments on the IEEE 118-bus system (augmented with renewables) showed that quadratic PCE models for the production cost showed pointwise relative errors less than approximately $1-2\%$ throughout the uncertain space for the renewable wind power. For relative accuracies of $o(10^{-4})$, the PCE approach is one to two orders more efficient than Monte Carlo-based estimates of the expected cost.

Nevertheless, the proposed algorithm faces the curse-of-dimensionality challenge, as the number of samples required by the PCE method exhibits a near exponential dependence on the number of stochastic dimensions. This is a universal challenge faced, to a varying degree, by all spectral non Monte Carlo methods for forward UQ, and is commonly resolved by judicious learning of the underlying intrinsic/effective low-dimensional structure of the model response. Based on the results presented in this manuscript, which suggest some stochastic directions are less important, we are now exploring algorithms based on global sensitivity analysis and optimized sparse quadrature to further decrease the computational expense associated with attaining a certain level of accuracy for the expected cost. These results will be part of a subsequent paper.

Appendix A
Formulation for Stochastic Economic Dispatch

The *expected* minimum production cost, denoted $\overline{Q}(\boldsymbol{x})$, is defined as

$$\overline{Q}(\boldsymbol{x}) = \mathrm{E}_{\boldsymbol{\xi}}\left[Q(\boldsymbol{x},\boldsymbol{\xi}(\omega))\right] \tag{18}$$

and the (multi-period) *stochastic economic dispatch problem* under a fixed unit commitment $\boldsymbol{x}$ is given by

$$Q(\boldsymbol{x},\boldsymbol{\xi}(\omega)) = \min_{\boldsymbol{f},\boldsymbol{p}\geq\boldsymbol{0},\boldsymbol{q}\geq\boldsymbol{0},\boldsymbol{\theta}} \quad \sum_{t\in T}\sum_{g\in G} c_g^P(p_g^t) + \sum_{t\in T}\sum_{i\in\mathcal{N}} M q_i^t \tag{19}$$

s.t.

$$\sum_{r\in R_i} p_r^t(\boldsymbol{\xi}(\omega)) + \sum_{g\in G_i} p_g^t + \sum_{\substack{e=(j,i)\\(j,i)\in\mathcal{E}}} f_e^t - \sum_{\substack{e=(i,j)\\(i,j)\in\mathcal{E}}} f_e^t = D_i^t - q_i^t,\ \forall i,t \tag{20a}$$

$$B_e(\theta_i^t - \theta_j^t) - f_e^t = 0, \quad \forall e=(i,j)\in\mathcal{E},t \tag{20b}$$

$$\underline{F}_e \leq f_e^t \leq \overline{F}_e, \quad \forall e,t \tag{20c}$$

$$\underline{P}_g x_g^t \leq p_g^t \leq \overline{P}_g x_g^t, \quad \forall g,t \tag{20d}$$

$$p_g^t - p_g^{t-1} \leq R_g^u x_g^{t-1} + S_g^u(x_g^t - x_g^{t-1}) + \overline{P}_g(1-x_g^t),\ \forall g,t \tag{20e}$$

$$p_g^{t-1} - p_g^t \leq R_g^d x_g^t + S_g^d(x_g^{t-1} - x_g^t) + \overline{P}_g(1-x_g^{t-1}),\ \forall g,t \tag{20f}$$

where R_g^u (R_g^d) and S_g^u (S_g^d) represent nominal ramp-up (ramp-down) and startup (shutdown) rates, respectively. Subscripts i and j denote bus indices defined over bus set $\mathcal{N}$, while the superscript t denotes specific time periods in the planning horizon $t=1,\cdots,T$. Subscript g denotes the generator index defined over generator set G, and subscript r denotes the renewable generator index defined over the renewable generator set R, with G_i and R_i respectively denoting the subsets of G and R associated with bus i. Subscript $e=(i,j)$ denotes the line index (and terminal buses (i,j)) defined over transmission line set $\mathcal{E}$, and f_e^t is the power flow though line e at time t. Renewable generation power (a parameter) is denoted by p_r^t while power from thermal generators (a variable) is denoted by p_g^t. The decision variables q_i^t denote the load shedding quantity at bus i at time period t. Because uncertain renewable generation is modeled as a random field, solution variables are necessarily random, and can be expressed as functions of the same set of random variables that are used to define the uncertain renewable generation. In subsequent sections we will present a method based on spectral representations of random variables leading to an efficient representation of uncertainties in solution variables.

The optimization objective in SED is to minimize the expected total production and loss-of-load costs. The first term in Eq. (19) represents total production cost, while the second term represents the loss-of-load penalty. The load-shedding penalty employs a large positive number M, typically \$5000 or \$6000 [5]. Eq. (20) specifies operational and physical constraints on grid components based on a *direct current (DC)* power flow model, and includes (in order): power balance at each time period and bus (20a); the power flow $B_e(\theta_i^t - \theta_j^t)$ through each line $e=(i,j)$ as a function of phase angle differences of the terminals and the reciprocal of the line reactance, B_e (20b); lower and upper bounds ($\underline{F}_e$ and $\overline{F}_e$, respectively) on line power flow (20c); lower and upper bounds ($\underline{P}_g$ and $\overline{P}_g$, respectively) for the standard generator power (20d); and generation ramp-up (20e) and ramp-down (20f) constraints for pairs of consecutive time periods. We employ a quadratic production cost function, $c_g^P(p_g^t)$, approximated using a set of piecewise linear segments. Carrión and Arroyo [1] show details concerning the linearization of the quadratic production cost functions.

Acknowledgment

This work was funded by the Laboratory Directed Research & Development (LDRD) program at Sandia National Laboratories. Sandia National Laboratories is a multiprogram laboratory operated by Sandia Corporation, a wholly owned subsidiary of Lockheed Martin Corporation, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

References

[1] M. Carrión and J. M. Arroyo, "A computationally efficient mixed-integer linear formulation for the thermal unit commitment problem," *IEEE T. Power Systems*, vol. 21, no. 3, pp. 1371–1378, 2006.

[2] "ISO New England: Forecast and scheduling reserve adequacy analysis ," www.iso-ne.com/support/training/courses/wem101/10_forecast_scheduling_callan.pdf, accessed: 2014-05-19.

[3] P. A. Ruiz, R. C. Philbrick, E. Zack, K. W. Cheung, and P. W. Sauer, "Uncertainty management in the unit commitment problem," *IEEE T. Power Systems*, vol. 24, no. 2, pp. 642–651, 2009.

[4] S. Takriti, J. Birge, and E. Long, "A stochastic model for the unit commitment problem," *IEEE T. Power Systems*, vol. 11, no. 3, pp. 1497–1508, 1996.

[5] A. Papavasiliou and S. S. Oren, "Multiarea Stochastic Unit Commitment for High Wind Penetration in a Transmission Constrained Network," *Operations Research*, vol. 61, no. 3, pp. 578–592, 2013.

[6] Q. Wang, J. Wang, and Y. Guan, "Stochastic unit commitment with uncertain demand response," *IEEE T. Power Systems*, vol. 28, no. 1, pp. 562–563, 2013.

[7] R. L.-Y. Chen, N. Fan, A. Pinar, and J.-P. Watson, "Contingency-constrained unit commitment with post-contingency corrective recourse," *Annals of Operations Research*, pp. 1–27, 2014.

[8] D. Bertsimas, E. Litvinov, X. A. Sun, J. Zhao, and T. Zheng, "Adaptive robust optimization for the security constrained unit commitment problem," *IEEE T. Power Systems*, vol. 28, no. 1, pp. 52–63, 2013.

[9] F. B. Thiam and C. L. DeMarco, "Optimal transmission expansion via intrinsic properties of power flow conditioning," in *2010 North American Power Symposium (NAPS)*, 2010, pp. 1–8.

[10] H. N. Najm, "Uncertainty Quantification and Polynomial Chaos Techniques in Computational Fluid Dynamics," *Annual Review of Fluid Mechanics*, vol. 41, no. 1, pp. 35–52, 2009.

[11] R. G. Ghanem and P. D. Spanos, *Stochastic Finite Elements: A Spectral Approach*. Springer Verlag, New York, 1991.

[12] B. Debusschere, K. Sargsyan, C. Safta, and K. Chowdhary, "UQ Toolkit," http://www.sandia.gov/UQToolkit, 2015.

[13] C. Safta, R. L.-Y. Chen, H. Najm, A. Pinar, and J.-P. Watson, "Toward using surrogates to accelerate solution of stochastic electricity grid operations problems," in *NAPS 2014*, Sept 2014, pp. 1–6.

[14] J. S. Hong, "Evaluation of the high-resolution model forecasts over the taiwan area during gimex," *Weather Forecast*, vol. 18, no. 5, pp. 836–846, 2003.

[15] E.-F. T.H.M., E.-S. E.F., and A. S. M.M., "One day ahead predictin of wind speed using annual trends," in *Proceedings of the Power Engineering Society General Meeting*, 2006, pp. 1–7.

[16] A. Papavasilious, S. Oren, and R. O'Neil, "Reserve requirements for wind power integration: A scenario-based stochastic programming framework," *IEEE T. Power Systems*, vol. 26, no. 4, pp. 2197–2206, 2011.

[17] I. Damousis, M. Alexiadis, J. Theocharis, and P. Dokopoulos, "A fuzzy model for wind speed prediction and power generation in wind parks using spatial correlation," *IEEE T. Energy Conversion*, vol. 19, no. 2, pp. 352–361, 2004.

[18] B. T.G. and T. J.B., "Locally recurrent neural networks for wind speed prediction using spatial correlation," *Information Sciences*, vol. 177, no. 24, pp. 5775–5797, 2007.

[19] P. Pinson, H. Madsen, H. A. Nielsen, G. Papaefthymiou, and B. Klöckl, "From probabilistic forecasts to statistical scenarios of short-term wind power production," *Wind Energy*, vol. 12, no. 1, pp. 51–62, 2009.

[20] A. M. Foley, P. G. Leahy, A. Maryuglia, and E. J. McKeogh, "Current methods and advances in forecasting of wind power generation," *Renewable Energy*, vol. 37, no. 1, pp. 1–8, 2012.

[21] Y. Zhang, J. Wang, and X. Wang, "Review on probabilistic forecasting of wind power," *Renewable and Sustainable Energy Reviews*, vol. 32, pp. 255–270, 2014.

[22] Y. Dvorkin, Y. Wang, H. Padzic, and D. Kirschen, "Comparison of scenario reduction techniques for the stochastic unit commitment," in *Proceedings of the Power Engineering Society General Meeting*, 2014, pp. 1–5.

[23] J. Sumaili, H. Keko, V. Miranda, Z. Zhou, A. Botterud, and J. Wang, "Clustering-based wind power scenario reduction technique," in *Proceedings of 17th Power Systems Computation Conference*, 2011.

[24] D. Arthur and S. Vassilvitskii, "K-means++: the advantages of careful seeding," in *Proceedings of 18th Annual ACM-SIAM Symposium on Discrete Algorithms*, 2007, pp. 1027–1035.

[25] J. Dupacova, N. Growe-Kuska, and W. Romisch, "Scenario reduction in stochastic programming: an approach using probability metrics," *Math. Program.*, vol. 95, pp. 493–511, 2003.

[26] H. Heitsch and W. Romisch, "Scenario reduction algorithms in stochastic programming," *Comput. Optim. Appl.*, vol. 2-3, pp. 187–206, 24.

[27] K. Murphy, *Machine Learning: A Probabilistic Perspective*. The MIT Press, 2012, ch. Clustering, pp. 877–908.

[28] D. Xiu and G. E. Karniadakis, "The Wiener-Askey polynomial chaos for stochastic differential equations," *SIAM J. Scientific Computing*, vol. 24, no. 2, pp. 619–644, 2002.

[29] O. L. Maître and O. Knio, *Spectral Methods for Uncertainty Quantification: With Applications to Computational Fluid Dynamics (Scientific Computation)*, 1st ed. Springer, April 2010.

[30] O. Ernst, A. Mugler, H.-J. Starkloff, and E. Ullmann, "On The Convergence of Generalized Polynomial Chaos Expansions," *ESAIM: Mathematical Modelling and Numerical Analysis*, vol. 46, pp. 317–339, 2012.

[31] S. A. Smolyak, "Quadrature and interpolation formulas for tensor products of certain classes of functions," *Soviet Mathematics Dokl.*, vol. 4, pp. 240–243, 1963.

[32] F. Nobile, R. Tempone, and C. Webster, "A sparse grid stochastic collocation method for partial differential equations with random input data," *SIAM Journal on Numerical Analysis*, vol. 46, no. 5, pp. 2309–2345, 2008.

[33] ——, "An anisotropic sparse grid stochastic collocation method for partial differential equations with random input data," *SIAM J. Num. Anal.*, vol. 46, no. 5, pp. 2411–2442, 2008.

[34] C. W. Clenshaw and A. R. Curtis, "A method for numerical integration on an automatic computer," *Numerische Mathematik*, vol. 2, pp. 197–205, 1996.

[35] T. Patterson, "The optimum addition of points to quadrature formulae," *Mathematics of Computation*, vol. 22, no. 104, pp. 847–856, 1968.

[36] "NREL: Transmission Grid Integration - Western Wind Dataset," http://www.nrel.gov/electricity/transmission/wind_integration_dataset.html, accessed: 2015-02-28.

[37] G. J. Székely, M. L. Rizzo, and N. K. Bakirov, "Measuring and testing dependence by correlation of distances," *Annals of Statistics*, vol. 35, pp. 2769–2794, 2007.

[38] C. Safta, K. Sargsyan, H. N. Najm, K. Chowdhary, B. Debusschere, L. P. Swiler, and M. S. Eldred, "Probabilistic Methods for Sensitivity Analysis and Calibration of Computer Models in the NASA Challenge Problem," *J. Aerospace Information Systems*, 2015, in press.

[39] C. E. Rasmussen and C. K. I. Williams, *Gaussian Processes for Machine Learning*. MIT Press, 2006.

[40] "ELIA: Wind-Power Generation Data," http://www.elia.be/en/grid-data/power-generation/wind-power, accessed: 2015-07-26.

[41] T. Gerstner and M. Griebel, "Numerical integration using sparse grids," *Numerical Algorithms*, vol. 18, pp. 209–232, 1998.

[42] "Power Systems Test Case Archive ," http://www.ee.washington.edu/research/pstca/, accessed: 2014-05-01.

[43] L. Kocis and W. J. Whiten, "Computational investigations of low-discrepancy sequences," *ACM Trans. Math. Softw.*, vol. 23, no. 2, pp. 266–294, Jun. 1997.