Multivariate Probabilistic Collocation Method for Effective Uncertainty Evaluation With Application to Air Traffic Flow Management

Yi Zhou, *Student Member, IEEE,* Yan Wan, *Member, IEEE,* Sandip Roy, *Member, IEEE,* Christine Taylor, Craig Wanke, Dinesh Ramamurthy, and Junfei Xie, *Student Member, IEEE*

***Abstract*—Modern large-scale infrastructure systems have typical complicated structure and dynamics, and extensive simulations are required to evaluate their performance. The probabilistic collocation method (PCM) has been developed to effectively simulate a system's performance under parametric uncertainty. In particular, it allows reduced-order representation of the mapping between uncertain parameters and system performance measures/outputs, using only a limited number of simulations; the resultant representation of the original system is provably accurate over the likely range of parameter values. In this paper, we extend the formal analysis of single-variable PCM to the multivariate case, where multiple uncertain parameters may or may not be independent. Specifically, we provide conditions that permit multivariate PCM to precisely predict the mean of original system output. We also explore additional capabilities of the multivariate PCM, in terms of cross-statistics prediction, relation to the minimum mean-square estimator, computational feasibility for large dimensional parameter sets, and sample-based approximation of the solution. At the end of the paper, we demonstrate the application of multivariate PCM in evaluating air traffic system performance under weather uncertainties.**

***Index Terms*—Air traffic flow management, dynamical simulation, uncertainty evaluation.**

I. Introduction

THERE is a growing need to effectively and strategically manage large-scale infrastructure systems, such as air traffic systems, power grids, and environmental systems. These systems are typically subject to a wide range of uncertainties, which significantly complicate the evaluation and management of system performance. To give some examples, flow contingency management solutions are being developed for air traffic systems, which seek to automatically generate management plans over a 2–15 h lookahead time that are robust to weather uncertainties [26], [27]. In analogy, strategic resource scheduling and real-time surveillance/control algorithms are sought for the power grid, that are flexible to uncertainties in renewable generation and load, and robust to complex and uncertain fault events [12], [29]. As a step toward real-time management, techniques are needed for accurate yet computationally efficient evaluation/prediction of system performance over a range of parametric uncertainties. To address this need in broad infrastructure system applications, this paper develops a systematic method to effectively evaluate output statistics for systems with multiple uncertain input parameters.

System uncertainty evaluation problems are typically addressed from two angles: analytical solutions and simulations. However, because of the large scale and complicated nature of large-scale infrastructure systems, analytical solutions for system dynamics are typically unavailable; therefore, simulations using complicated computerized models are the primary approaches for uncertainty evaluation. Monte Carlo methods have been widely used to evaluate system performance under uncertainty by practitioners in many application domains [34]. Broadly, Monte Carlo methods have three main steps: 1) generation of a large number of samples covering the range of parameter uncertainties; 2) simulation of system performance for each parameter sample; and 3) summary of simulation outputs to obtain the system-performance statistics. Often, Monte Carlo methods may not be suitable for large-scale infrastructure applications, because of their inherent computational cost. Specifically, Monte Carlo methods typically require evaluation of a large number of simulations of the mapping of interest; since infrastructure-network simulations are rather computationally intensive, such exhaustive simulations are often impossible (especially when real-time decision-making is needed).

Driven by this limitation, our group (as well as others) have sought for alternative simulation methods to effectively evaluate parametric uncertainties. Our philosophy is that a limited number of simulations, if appropriately chosen, can provide adequate approximations of the mapping between uncertain input parameters and system performance over the range of likely parameter values. In particular, an adequate low-order mapping allows us to obtain statistical characterizations of system performance, and to evaluate system performance at any particular parameter value of interest. How to smartly

Manuscript received July 25, 2013; revised December 1, 2013; accepted February 18, 2014. Date of publication April 2, 2014; date of current version September 12, 2014. This work was supported in part by the National Science Foundation under Grant 1035386, in part by the contract from the MITRE Corporation, and in part by the U.S. Government under Grant DTFAWA-10-C-00080. This paper was recommended by Associate Editor S. Das.

Y. Zhou, Y. Wan, and J. Xie are with the Department of Electrical Engineering, University of North Texas, Denton, TX 76203 USA (e-mail: yan.wan@unt.edu).

S. Roy is with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA 99164 USA.

C. Taylor and C. Wanke are with the MITRE Corporation, McLean, VA 22102 USA.

D. Ramamurthy is with Enterprise Holdings, St. Louis, MO 63105 USA.

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMC.2014.2310712

choose the values of input parameters as simulation points and to construct the low-order mapping that allows best characterization of system-performance statistics is the key.

The probabilistic collocation method (PCM) is a method to evaluate the uncertainty of computationally expensive models at a low computational cost. It suggests a smart way of selecting simulation points to construct a low-order polynomial mapping between uncertain parameters and output or performance variable, that performs well over the likely range of parameter values [25], [32]. The idea behind PCM is as follows. Although system parameters are uncertain, we typically have some statistical knowledge about the parametric uncertainties; smartly utilizing this information allows us to find representative simulation points, from which a reduced-order mapping of high fidelity can be constructed. In particular, it was proved in [12] that the low-degree polynomial mapping generated by PCM is able to predict the mean output (performance) correctly, even if the actual mapping is a much higher degree polynomial. In [29], we further explored the statistical performance of PCM, and the practical use of it when data or empirical low-order statistics is available instead of the probabilistic density function (pdf).

Several studies have detailed applications of PCM in transportation- and power-network management [12], [16], [17], [23]. While these efforts are promising, a limitation is that only one uncertain input parameter is considered. In practice, large-scale infrastructure systems typically involve multiple (sometimes a large number of) spatially/temporally distributed uncertain parameters. In these cases, more than one uncertain parameter (which may be correlated) may exert significant impact on system performance or other output variables. Thus, we are motivated to develop smart simulation techniques analogous to PCM, and to understand their effectiveness and cost, when multiple correlated, uncertain parameters are present. Specially in this paper, we extend the formal analysis of the single-variable PCM to the multivariate case, where the uncertain parameters may or may not be independent.

The essence of our multivariate PCM approach lies in the smart selection of simulation points. Parsimonious selection of sampling points is needed for a range of applications, and has been widely studied (including in the specific context of uncertainty evaluation and mapping identification for complex systems). Here, we provide a brief and incomplete review of these methods, focusing primarily on differentiating our approach from related ones. For the purpose of polynomial interpolation, Chebyshev nodes are widely used to overcome the Runge's phenomenon observed in using equally-spaced sampling [5]. For the purpose of uncertainty evaluation, besides Monte Carlo methods (also called random samping), several other techniques have been proposed. For instance, the stratified sampling and Latin Hypercube Design (LHD) and their variants were developed to improve the coverage and projective properties [13], [15], [18]. Multiple steps are involved in these methods: first the space is subdivided to ensure full coverage, and then random sampling is used for each portion. These above methods are primarily designed to ensure appropriate coverage of the sampling space and to select sampling densities in the space, rather than to guarantee accurate estimation of output statistics. Gaussian Quadrature techniques also have been widely studied for uncertainty evaluation, both under the heading of PCM and using other terminology [14], [33], [35]. When multiple independent input parameters are involved, [14], [32], [33] suggest different procedures to select simulation points, among which [14] involves the most number of points. Gaussian Quadrature techniques have also been used to solve ODE/PDEs with uncertain parameters [6], [11], [16], [19]. These works mesh finite element decomposition in space and collocation methods on random variables, to approximate continuous high-dimensional solutions. All these works are relevant to our study, as they also consider the utilization of Gaussian Quadrature techniques when multiple uncertain variables are present. However, the purpose of our investigation in this paper is different from these studies in two aspects. First, these studies consider the approximation of system mapping using expansions of orthogonal polynomials (e.g., Hermite for Gaussian distribution, Laguerre for Gamma distribution, Jocobi for Beta distribution, and Legendra for Uniform distribution) with small error, instead of general polynomial system mapping and its low-order approximation that predicts precisely the same statistics as we do. Second, their studies are not focused on understanding how correlations in parametric uncertainties may affect expected performance while our study does consider correlated uncertainty.

Our major contribution here is a formal investigation of the properties of PCM for systems with multiple uncertain input parameters (see also [37] for the brief conference version). Specifically, we: 1) identify precise conditions on mapping functions and distributions to permit zero-error mean prediction, in both the independent and correlated cases; 2) in turn develop algorithms that obtain the best statistical performance; and 3) provide additional performance analyses, such as cross-statistics prediction, relation to minimum mean square estimation (MMSE), and computational feasibility analysis for large dimensional data. The multivariate PCM method that we develop overcomes the inadequacies of analytical uncertainty evaluation methods that do not work for complex infrastructure systems and Monte Carlo simulation methods that are computationally costly. In addition, our theoretical development provides a comprehensive understanding of the precise performance of PCM in multivariate settings and informs its practical use in performance evaluation and real-time decision making in large-scale infrastructure-type applications.

The remainder of the paper is organized as follows. In Section II, we describe and evaluate multivariate PCM when input parameters are independent. In Section III, we discuss further properties of the independent multivariate PCM, and computational issues. In Section IV, we provide results of the multivariate PCM when parameters are correlated. In Section V, we discuss the practical use of PCM when data or empirical low-order moments on low-order parameters, rather than explicit distributions of these parameters, are available. In Section VI, we present two examples to demonstrate the use of multivariate PCM to evaluate air traffic system performance under weather uncertainties. Section VII concludes the paper.

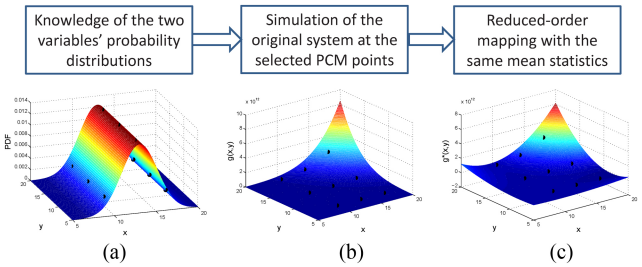


Fig. 1. (a) Joint probability density function $f_{X,Y}(x, y)$. (b) Original mapping $g(x, y)$. (c) Reduced-order multivariate PCM mapping $g^*(x, y)$. The black dots represent PCM points.

II. Independent Multivariate PCM

In this section, we consider a costly-to-simulate system with multiple input parameters subject to independent uncertainties. To effectively evaluate the dependence of system output on these input parameters and to obtain the statistics of system output under uncertainty, we study how to smartly select a limited number of simulations to construct a low-order polynomial mapping between the inputs and the output with good statistical performance.

A. Simple Two-Variable Case

To simplify the development of multivariate PCM, we first focus on the two-variable case. Specifically, let us consider a system of interest, for which we wish to identify the mapping between two variable input parameters x and y and an output of interest that is functionally dependent on the two parameters, say $g(x, y)$. Unfortunately, it is costly to compute the output for an input pair (e.g., because it requires a time-consuming simulation or a costly experiment), and hence we can evaluate $g(x, y)$ for only a limited number of pairs (x, y). With this limited ability to probe the system, we seek to accurately identify the mapping over a useful range of parameter values. More precisely, we model the two input parameters, for the evaluation task of interest, as independent random variables with known probability distribution $f_{X,Y}(x, y) = f_X(x) f_Y(y)$.

We propose a technique for selecting input pairs for simulation and in turn approximating the mapping of interest, which we call the two-variable PCM. The conceptual basis for the two-variable PCM (which is analogous to the one-variable case) is the following. A sparse number of points (pairs) are selected for evaluation, and these evaluations are used to obtain a low-degree polynomial mapping between the input parameters and the output; the points are specially selected, in such a way that the obtained low-order polynomial mapping can predict the mean output correctly even if the mapping in reality is of a much higher degree (see Fig. 1 for an illustration).

1) *Main Results:* In this section, we first illustrate the algorithm of the two-variable PCM in finding the PCM simulation points, the coefficients of the low-order PCM mapping, and then the predicted output mean. We then in Theorem 1 formalize that the low-degree approximation achieves the same mean as a higher-degree mapping of a general form, if simulation points are selected according to the algorithm. A simple example then follows to illustrate the use of two-variable PCM. The rest of the paper extends this theorem in multiple directions toward: 1) more than two input variables; 2) dependency among the input variables; 3) different general forms of original mappings; 4) further statistical performance analysis; and 5) input variables whose distributions are not explicitly known but specified by sample data.

Algorithm: The following algorithm shown in Table I, based on the distributions of random variables X and Y, $f_X(x)$ and $f_Y(y)$, constructs a low-order mapping of the form

$$g^*(x, y) = \sum_{i=0}^{n-1} \sum_{j=0}^{m-1} B_{i,j} x^i y^j \tag{1}$$

with the statistical property suggested by Theorem 1, where the coefficients $B_{i,j} \in R$.

In Theorem 1, we prove that a two-variable mapping $g(x, y)$ with the degrees of x up to $2n-1$ and y up to $2m-1$ can be approximated by a low-order mapping $g^*(x, y)$ with the degrees of x up to $n-1$ and y up to $m-1$, with the same expected values. We note that identifying the low-order PCM mapping $g^*(x, y)$ requires evaluating $g(x, y)$ at the simulation points (x_i, y_j), where x_i and y_j are the roots of the orthonormal polynomials $h_n(x)$ and $h'_m(y)$, respectively. Please refer to Appendix A for the proof.

Theorem 1: Consider a two-variable mapping $g(x, y)$ of the form

$$g(x, y) = \sum_{i=0}^{2n-1} \sum_{j=0}^{2m-1} A_{i,j} x^i y^j \tag{2}$$

where the coefficients $A_{i,j} \in R$, and n and m are integers greater than 1. With the assumption that the two variables x and y follow independent distributions $f_X(x)$ and $f_Y(y)$, respectively, the mapping $g(x, y)$ can be approximated by a low-order mapping $g^*(x, y)$ of the form shown in (1), such that $E[g(x, y)] = E[g^*(x, y)]$.

2) *Discussion and Example:* Now let us discuss the efficiency of the independent two-variable PCM. As the original mapping $g(x, y)$ is a polynomial with x up to the degree of $2n-1$ and y up to $2m-1$, a total of $(2n)(2m)$ simulations are required to uniquely identify the mapping. However, identifying $g^*(x, y)$ requires only nm simulations, as there are n roots for $h_n(x)$ and m roots for $h'_m(y)$. As such, the two-variable PCM can reduce the number of simulations by $3nm$.

Finally we use a simple example to illustrate the procedure and performance of the independent two-variable PCM as shown in Fig. 1. We consider a two-variable mapping $g(x, y) = x^5y^5 - 2x^5y^4 + 3x^5y^3 - 4x^5y^2 + 5x^5y - 6x^5 - 2x^4y^5 + 4x^4y^4 - 6x^4y^3 + 8x^4y^2 - 10x^4y + 12x^4 + 3x^3y^5 - 6x^3y^4 + 9x^3y^3 - 12x^3y^2 + 15x^3y - 18x^3 - 4x^2y^5 + 8x^2y^4 - 12x^2y^3 + 16x^2y^2 - 20x^2y + 24x^2 + 5xy^5 - 10xy^4 + 15xy^3 - 20xy^2 + 25xy - 30x - 6y^5 + 12y^4 - 18y^3 + 24y^2 - 30y + 36$ as shown in Fig. 1(b). Assuming that x is normally distributed with mean 13 and variance 4.6225 and y is uniformly distributed between 5 and 20, we show that the mapping can be well approximated by a polynomial with the degrees of x and y each up to 2.

To do that, we choose 9 simulation points based on the pdf of x and y, according to Step 1 in the algorithm. The locations of the simulation points

TABLE I
ALGORITHM FOR THE INDEPENDENT TWO-VARIABLE PCM

Step 1:	*Simulation point selection*
1.1	Compute the orthonormal polynomials of degree i for the random variable X, $h_i(x)$, for $i = 1, ..., n$ according to the following: *a)* Initialize $H_{-1}(x) = h_{-1}(x) = 0$ and $H_0(x) = h_0(x) = 1$ *b)* **For** $i = 1$ to n $H_i(x) = xh_{i-1}(x) - \langle xh_{i-1}(x), h_{i-1}(x)\rangle h_{i-1}(x) - \langle H_{i-1}(x), H_{i-1}(x)\rangle^{\frac{1}{2}} h_{i-2}(x)$, $h_i(x) = H_i(x)/\langle H_i(x), H_i(x)\rangle^{\frac{1}{2}}$. **End** *Note* that $H_i(x)$ represents the orthogonal polynomial of degree i for the random variable X. The second equation in the FOR loop normalizes $H_i(x)$. $\langle p(x), q(x)\rangle$ denotes the integration operation $\int p(x)q(x)f_X(x)dx$.
1.2	Find the roots of $h_n(x) = 0$ as the n PCM simulation points for X, denoted as x_1, x_2,, ..., x_n.
1.3	Repeat 1.1 and 1.2 for the random variable Y and obtain the orthonormal polynomials $h'_j(y)$, for $j = 1, ...m$, and the m PCM simulation points for Y, denoted as y_1, $y_2, ..., y_m$.
Step 2:	*Evaluation of system outputs at selected simulation points*
2.1	For each simulation points (x_i, y_j) where $i \in \{1, ...n\}$ and $j \in \{1...m\}$, run simulation and find the associated output $g(x_i, y_j)$.
Step 3:	*Identification of Mapping Coefficients*
3.1	Find the coefficients $b_{i,j}$ in the low-order PCM mapping: $g^*(x, y) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{i,j}h_i(x)h'_j(y)$ following: $\begin{bmatrix} b_{n-1,m-1} \\ \vdots \\ b_{n-1,0} \\ \vdots \\ b_{0,m-1} \\ \vdots \\ b_{0,0} \end{bmatrix} = K^{-1} \begin{bmatrix} g(x_1, y_1) \\ \vdots \\ g(x_1, y_m) \\ \vdots \\ g(x_n, y_1) \\ \vdots \\ g(x_n, y_m) \end{bmatrix}$, where $K = \begin{bmatrix} h_{n-1}(x_1)h'_{m-1}(y_1) & \cdots & h_{n-1}(x_1)h'_0(y_1) & \cdots & h_0(x_1)h'_0(y_1) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{n-1}(x_1)h'_{m-1}(y_m) & \cdots & h_{n-1}(x_1)h'_0(y_m) & \cdots & h_0(x_1)h'_0(y_m) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{n-1}(x_n)h'_{m-1}(y_1) & \cdots & h_{n-1}(x_n)h'_0(y_1) & \cdots & h_0(x_n)h'_0(y_1) \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ h_{n-1}(x_n)h'_{m-1}(y_m) & \cdots & h_{n-1}(x_n)h'_0(y_m) & \cdots & h_0(x_n)h'_0(y_m) \end{bmatrix}$.
3.2	The predicted output mean is $b_{0,0}$.
3.3	Reorganize terms in the low-order PCM mapping, and obtain coefficients $B_{i,j}$ in the low-order PCM mapping: $g^*(x, y) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} B_{i,j}x^i y^j$.

are (9.2761, 6.6905), (9.2761, 12.5000), (9.2761, 18.3095), (13, 6.6905), (13, 12.5000), (13, 18.3095), (16.7239, 6.6905), (16.7239, 12.5000), and (16.7239, 18.3095) as shown in Fig. 1(a). We then evaluate $g(x, y)$ at these 9 locations according to Step 2, and obtain the coefficients of the lower-oder PCM mapping according to the Step 3. The resulting PCM mapping up to the order of x^2y^2 is shown in Fig. 1(c).

From Fig. 1(a), we see that the 9 simulation points are located at the range of likely parameter values, in response to the probability distributions of x and y. Comparing Fig. 1(b) and (c) shows that the PCM mapping $g^*(x, y)$ approximates $g(x, y)$ very well over the likely domain of parameter values, despite the significantly reduced mapping orders and number of simulations required for the construction.

B. General Theorem and Procedure on Independent Multivariate PCM Mapping

Theorem 1 can be easily generalized to the multivariate PCM with more than two variables, as shown in Theorem 2. Please refer to Appendix B for the proof.

Theorem 2: Consider a multivariate mapping $g(x_1, x_2, ..., x_m)$ of the form

$$g(x_1, x_2, ..., x_m) = \sum_{k_1=0}^{2n_1-1}\sum_{k_2=0}^{2n_2-1}\cdots\sum_{k_m=0}^{2n_m-1} A_{k_1,...,k_m}\prod_{i=1}^{m} x_i^{k_i} \quad (3)$$

where the coefficients $A_{k_1,...,k_m} \in R$, and n_1, n_2, ...n_m are integers larger than 1. Assume that the variables x_1, x_2, ..., x_m follow independent distributions $f_{X_1}(x_1)$, $f_{X_2}(x_2)$, ..., and $f_{X_m}(x_m)$, respectively. Then, the mapping can be approximated by a low-order mapping $g^*(x_1, x_2, ..., x_m)$ of the form

$$g^*(x_1, x_2, ..., x_m) = \sum_{k_1=0}^{n_1-1}\sum_{k_2=0}^{n_2-1}\cdots\sum_{k_m=0}^{n_m-1} B_{k_1,...,k_m}\prod_{i=1}^{m} x_i^{k_i} \quad (4)$$

such that $E[g(x_1, x_2, ..., x_m)] = E[g^*(x_1, x_2, ..., x_m)]$, where the coefficients $B_{k_1,...,k_m} \in R$.

Theorem 2 shows that the independent multivariate PCM method can significantly reduce the number of simulations needed to construct a low-order mapping of the same mean. Similar to what we have argued in the two-variable case, constructing $g(x_1, x_2, ..., x_m)$ as shown in (3) requires at least $2^m\Pi_{i=1}^m n_i$ simulations. This is because $2n_i$ simulations are required to uniquely identify each variable up to the degree of $2n_i - 1$. However, the reduced-order mapping $g^*(x_1, x_2, ..., x_m)$ as shown in (4) requires only $\Pi_{i=1}^m n_i$ simulations evaluated at the roots of each $h_{n_i}^i(x_i)$. The independent multivariate PCM can thus save $(2^m - 1)\Pi_{i=1}^m n_i$ simulations without introducing any error to the mean prediction.

To complete the presentation and facilitate the use of the independent multivariate PCM, we briefly summarize the procedure to find the PCM mapping in its general form [as shown in (4)].

Step 1 (Simulation Point Selection): As a generalization to the two-variable case, $\Pi_{i=1}^m n_i$ simulation points are required to identify a m-variable PCM mapping shown in (4). Due

to the independence of variables, we only need to identify n_i points along the dimension of each x_i, for $i \in \{1, ..., m\}$. Combinations of the n_i points along each dimension constitute the set of $n_1 ... n_m$ simulation points. The n_i points associated with random variable x_i are the roots of the orthonormal polynomial $h^i_{n_i}(x_i)$, which can be found through the recursion illustrated in Step 1 in the algorithm in Section II-A.

Step 2 (Evaluation of System Outputs at Selected Simulation Points): For each simulation point identified in Step 1, we find the associated output through simulation. In total, $\Pi_{i=1}^{m} n_i$ simulations are needed. We stress that in many applications, this is the most time-consuming step. As the multivariate PCM can significantly reduce the number of simulations required to produce a mapping with correct mean prediction, this method is of significance to find high-dimensional mappings that are common to large-scale infrastructure applications.

Step 3 (Identification of Mapping Coefficients): Using the simulations in Step 2, we can explicitly express the low order mapping $g^*(x_1, ..., x_m)$ by identifying the coefficients in its orthonormal polynomial form. This result is a simple generalization of Step 3 of the algorithm in Section II-A for the two-variable case. In particular, the coefficients $a_{k_1,...,k_m}$ can be calculated using the matrix operation in (5). The vector a is arranged in a descending order of the subscripts starting from the last bit to the first bit, and the vector g is arranged in an increasing order of the root indices (i), again from the rightmost entry to the leftmost entry, where $1 \le i \le n_m$

$$L \begin{bmatrix} a_{n_1-1,...,n_m-1} \\ \vdots \\ a_{n_1-1,...,0} \\ a_{n_1-1,...n_{m-1}-2,n_m-1} \\ \vdots \\ a_{n_1-1,...n_{m-1}-2,0} \\ \vdots \\ a_{0,...,0} \end{bmatrix} = \begin{bmatrix} g(x_{1(1)}, x_{2(1)}, ..., x_{m(1)}) \\ \vdots \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m(n_m)}) \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m-1(2)}, x_{m(1)}) \\ \vdots \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m-1(2)}, x_{m(n_m)}) \\ \vdots \\ g(x_{1(n_1)}, x_{2(n_2)}, ..., x_{m(n_m)}) \end{bmatrix}. \tag{5}$$

Here, $L \in R^{n_1 n_2 ... n_m \times n_1 n_2 ... n_m}$ takes the following form:

$$\begin{bmatrix} h^1_{n_1-1}(x_{1(1)})...h^m_{n_m-1}(x_{m(1)}) & \cdots & h^1_0(x_{1(1)})...h^m_0(x_{m(1)}) \\ h^1_{n_1-1}(x_{1(1)})...h^m_{n_m-1}(x_{m(2)}) & \cdots & h^1_0(x_{1(1)})...h^m_0(x_{m(2)}) \\ \vdots & \ddots & \vdots \\ h^1_{n_1-1}(x_{1(n_1)})...h^m_{n_m-1}(x_{m(n_m)}) & \cdots & h^1_0(x_{1(n_1)})...h^m_0(x_{m(n_m)}) \end{bmatrix}.$$

In each row, the subscripts of h are arranged in a descending order from $n_1-1, n_2-1, ..., n_m-1$ to $0, 0, ...0$. In each column, the indices of roots are in an increasing order from $1, 1, ..., 1$ to $n_1, n_2, ...n_m$. The coefficients $B_{k_1,...,k_m}$ shown in (4) can then be easily derived from $a_{k_1,...,k_m}$.

III. Properties of Independent Multivariate PCM

As proved in Section II, the independent multivariate PCM can find a low-order polynomial mapping that correctly predicts the mean of the original higher-order mapping. In this section, we further characterize the performance of the independent multivariate PCM mapping. In particular, we show in Section III-A that the low-order mapping also predicts the correct cross-statistics. We then compare its performance with that of the MMSE. In Section III-B, we briefly discuss the numerical capability of the independent multivariate PCM in processing high-dimensional data.

A. Performance of Independent Multivariate PCM

In Theorem 3, we show that the low-order independent multivariate PCM can precisely predict the cross-statistics up to certain degree. Please refer to Appendix C for the proof.

Theorem 3: Consider the use of a multivariable PCM mapping $g^*(x_1, x_2, ..., x_m)$ of the form $\sum_{k_1=0}^{n_1-1} \sum_{k_2=0}^{n_2-1} \cdots \sum_{k_m=0}^{n_m-1} B_{k_1,k_2,...,k_m} \prod_{i=1}^{m} x_i^{k_i}$ to approximate an original polynomial mapping $g(x_1, x_2, ..., x_m)$ of the form $\sum_{k_1=0}^{n_1+p_1} \sum_{k_2=0}^{n_2+p_2} \cdots \sum_{k_m=0}^{n_m+p_m} A_{k_1,k_2,...,k_m} \prod_{i=1}^{m} x_i^{k_i}$, for all $p_i \in \{0, ..., n_i - 1\}$. Assuming that all variables are independent, the low-order PCM can correctly predict the cross-statistics, that is

$$E\left[\left(\prod_{i=1}^{m} x_i^{l_i}\right) g^*(x_1, x_2, ..., x_m)\right] = E\left[\left(\prod_{i=1}^{m} x_i^{l_i}\right) g(x_1, x_2, ..., x_m)\right] \tag{6}$$

for all $l_i \in \{0, ..., n_i - 1 - p_i\}$.

In Theorem 4, we discuss the relationship between the independent multivariate PCM mapping and the MMSE estimator. In general, a low-order PCM mapping may not be the MMSE estimator among all polynomials of the same degree. However, the performance of PCM is attractive as reflected by the following two additional results: 1) the mean square error (MSE) performance of a PCM mapping may not be improved by adding any polynomial up to certain degree, and 2) a PCM mapping up to certain degree is the MMSE estimator among all polynomials of the same degree. Please refer to the Appendix D for the proof.

Theorem 4: Consider a multivariate mapping $\sum_{k_1=0}^{n_1+p_1} \sum_{k_2=0}^{n_2+p_2} \cdots \sum_{k_m=0}^{n_m+p_m} A_{k_1,k_2,...,k_m} \prod_{i=1}^{m} x_i^{k_i}$ for some $p_i \in \{0, ..., n_i - 1\}$. If a PCM mapping of the form

$$g^*(x_1, x_2, ..., x_m) = \sum_{k_1=0}^{n_1-1} \sum_{k_2=0}^{n_2-1} ... \sum_{k_m=0}^{n_m-1} a_{k_1,...,k_m} \prod_{i=1}^{m} h^i_{k_i}(x_i)$$

is used to fit the original mapping, the mean square error of the PCM fit cannot be improved by adding any polynomial with the degree of x_i up to $n_i - 1 - p_i$. Moreover, the lower order PCM mapping

$$g^*_r(x_1, x_2, ..., x_m) = \sum_{k_1=0}^{n_1-1-p_1} ... \sum_{k_m=0}^{n_m-1-p_m} a'_{k_1,...,k_m} \prod_{i=1}^{m} h^i_{k_i}(x_i)$$

is the MMSE mapping, among all polynomials with the degree of each x_i up to $n_i - 1 - p_i$.

B. Discussion on Numerical Issues

We note that the performance of PCM may be affected by numerical issues. As seen from (5), finding the coefficients $a_{n_1-1,...,n_m-1}$, and then $B_{n_1-1,...,n_m-1}$ in (4) (according to the procedure discussed in Section II-B) involves solving a large

equation array. Numerical issues (such as the loss of precision and even failure of getting a solution) may appear especially when the dimension of the array is large. Instead of providing a formal proof, we discuss the feasibility of this procedure through comparing it with an alternative procedure. We first describe the alternative procedure and then use the example presented in Section II-A to compare the performance of these two.

We note that once PCM points are selected and outputs at those points are simulated, we can find the mapping coefficients $B_{n_1-1,...,n_m-1}$ directly, instead of using the orthonormal bases. In particular

$$L' \begin{bmatrix} B_{n_1-1,...,n_m-1} \\ \vdots \\ B_{n_1-1,...,0} \\ B_{n_1-1,...n_{m-1}-2,n_m-1} \\ \vdots \\ B_{n_1-1,...n_{m-1}-2,0} \\ \vdots \\ B_{0,...,0} \end{bmatrix} = \begin{bmatrix} g(x_{1(1)}, x_{2(1)}, ..., x_{m(1)}) \\ \vdots \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m(n_m)}) \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m-1(2)}, x_{m(1)}) \\ \vdots \\ g(x_{1(1)}, x_{2(1)}, ..., x_{m-1(2)}, x_{m(n_m)}) \\ \vdots \\ g(x_{1(n_1)}, x_{2(n_2)}, ..., x_{m(n_m)}) \end{bmatrix} \tag{7}$$

where $L' \in R^{n_1 n_2 ... n_m \times n_1 n_2 ... n_m}$ takes the following form:

$$\begin{bmatrix} x_1^{n_1-1}(x_{1(1)})...x_m^{n_m-1}(x_{m(1)}) & \cdots & x_1^0(x_{1(1)})...x_m^0(x_{m(1)}) \\ x_1^{n_1-1}(x_{1(1)})...x_m^{n_m-1}(x_{m(2)}) & \cdots & x_1^0(x_{1(1)})...x_m^0(x_{m(2)}) \\ \vdots & \ddots & \vdots \\ x_1^{n_1-1}(x_{1(n_1)})...x_m^{n_m-1}(x_{m(n_m)}) & \cdots & x_1^0(x_{1(n_1)})...x_m^0(x_{m(n_m)}) \end{bmatrix}.$$

In the matrix L', $x_i^k(x_{i(j)})$ represents the kth power of the variable x_i evaluated at the jth simulation point of the variable x_i. The L' matrix is arranged as follows. In each row, the superscripts of x are arranged in a descending order from $n_1 - 1, n_2 - 1, ..., n_m - 1$ to $0, 0, ...0$. In each column, the indices of roots are in an increasing order from $1, 1, ..., 1$ to $n_1, n_2, ...n_m$.

The feasibility of finding the coefficients B in (7) or a in (5) is reflected by condition number, which describes the accuracy of numerical solutions to linear equation arrays [7]. Condition number is defined as the ratio between the largest and the smallest eigenvalues. We often consider the logarithm of condition number to be the loss of accuracy, although this estimation is not always precise. At the extreme, infinite condition number means that the matrix is singular and the equation array is unsolvable. We note that ill-conditioning typically leads to a drastic failure in solving equations rather than small numerical issues; hence, devising means for increasing solution precision rather than possible impacts of numerical errors may be the key concern. The possibility for ill-conditioning is dependent on a number of factors, including the number of variables, number of PCM points, and probability distribution; the possibility for numerical error arising from ill-conditioning may also depend on the nature of the system mapping, and of course the precision of the computing device. We here calculate and compare the condition numbers of the two matrices L' in (7) and L in (5), respectively, for the example in Section II-A. The condition number associated with L is 2.53, and associated with L' is 9.54×10^6 which is significantly larger. This simple comparison clearly shows the advantage of using orthogonal bases (5) over the x and y bases (7) to calculate PCM coefficients, especially for high-dimensional mappings.

IV. Correlated Multivariate PCM

The development in the previous sections is based on the assumption that uncertain variables are independent. In realistic modern infrastructure applications, system variables are often correlated. In general correlated settings, the PCM mapping obtained using the procedure described in Section II-B does not predict the correct mean. In this section, we explore conditions on the forms of original mapping and joint distribution to maintain the correct mean statistics.

A. Main Results

We first discuss the case when the original mapping does not contain any cross-terms among variables, and then the general case when cross-terms are present. In the first case, a low-order multivariate PCM mapping can be obtained in a way very similar to the independent case. The proof is omitted as it is a simple variation of the independent multivariate PCM [12].

Theorem 5: Consider a multivariate mapping $g_n(x_1, x_2, ..., x_m)$ of the form

$$g_n(x_1, x_2, ..., x_m) = \sum_{i=1}^{m} \sum_{k_i=0}^{2n_i-1} A'_{i,k_i} x_i^{k_i} \tag{8}$$

where the coefficients $A'_{i,k_i} \in R$ and n_i is an integer larger than 1, where $i \in 1, 2, ..., m$. Assume that the variables x_1, x_2, ..., x_m follow a joint distribution $f_{X_1 X_2 ... X_m}(x_1, x_2, ...x_m)$. The original mapping can be approximated by a lower-order mapping $g_n^*(x_1, x_2, ..., x_m)$ of the form

$$g_n^*(x_1, x_2, ..., x_m) = \sum_{i=1}^{m} \sum_{k_i=0}^{n_i-1} B'_{i,k_i} x_i^{k_i} \tag{9}$$

such that $E[g_n(x_1, x_2, ..., x_m)] = E[g_n^*(x_1, x_2, ..., x_m)]$, where the coefficients $B'_{i,k_i} \in R$ for $i \in 1, 2, ..., m$.

We now discuss the case when cross-terms are present. We show that under certain assumptions on conditional moments, a low-order PCM mapping can correctly predict the mean of an original mapping up to certain degree. Please refer to Appendix E for the proof.

Theorem 6: Consider a multivariate mapping $g_c(x_1, x_2, ..., x_m)$ of the form

$$g_c(x_1, x_2, ..., x_m) = \sum_{k_1=0}^{2n-1} \sum_{k_2=0}^{2n-1-k_1} \sum_{k_3=0}^{2n-1-(k_1+k_2)} ... \sum_{k_m=0}^{2n-1-(k_1+k_2+...+k_{m-1})} A'_{k_1,...,k_m} \prod_{i=1}^{m} x_i^{k_i} \tag{10}$$

where the coefficients $A'_{k_1,...,k_m} \in R$ and n is an integer larger than 1. Assume that the variables x_1, x_2,...,x_m follow a joint distribution $f_{X_1, X_2, ..., X_m}(x_1, x_2, ..., x_m)$, and the rth order conditional moment of x_i on x_{i+1}, ..., x_m is a polynomial of x_{i+1},...,x_m with the total degree not greater than r. The original

mapping can be approximated by a lower-order mapping $g_c^*(x_1, x_2, ..., x_m)$ with the same mean. The degree of each variable in $g_c^*(x_1, x_2, ..., x_m)$ is not greater than $n-1$.

Theorem 6 shows that in the two-variable case, if the rth conditional moment of one variable given the other is a rth order polynomial, the low-order PCM mapping with each variable up to the degree of $n-1$ predicts the correct mean when the total degree of the original mapping does not exceed $2n-1$. The procedure to pick PCM points here is different from that in the independent case, in the sense that the point locations are selected based on the marginal probability of one variable, and the conditional probability of the other. In particular, PCM points are selected through: 1) determining the y-coordinates (denoted as $y_1, ..., y_n$) of the PCM points according to the marginal probability of y and 2) finding the PCM points' x-coordinates according to the conditional probability $f_{X|Y}(x|y_i)$ for each selected y_i.

B. Discussion on Reduction of Computation Load

Here, let us also discuss the reduction of computational load. In the correlated case, the effectiveness of the multivariate PCM is generally not as good as that in the independent case. The original high-order mapping requires C_{2n-1+m}^m simulations, where C_p^q denotes the combination of p elements taken q of them at a time without repetition. However, the number of computations that can be reduced is generally not easy to write in the closed-form, due to the complicated conditional relationship. We here only discuss the two-variable and three-variable cases.

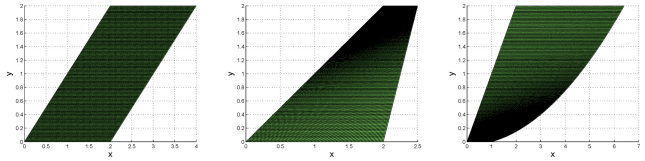
In the two-variable case, clearly, identifying the low-order PCM mapping $g_c^*(x, y)$ requires n^2 simulations to correctly predict the mean output. However, as the total degree of the original mapping is $2n-1$, $2n^2+n$ simulations are needed to identify all the coefficients in the original mapping. As such, more than half of the simulation time can be saved. Finally, we also note that to identify the low-order mapping, we need to directly find $B'_{i,j}$, instead of $e'_{l,k}$ at the orthogonal coordinates, as $e'_{l,k}$ is a varying function of y. In the three-variable case, the total number of simulations to identify $g_c(x_1, x_2, x_3)$ equals to $\frac{4}{3}n^3+n^2+\frac{2}{3}n$, because $g_c(x_1, x_2, x_3)$ has this number of coefficients. Similarly, we can check that the low-order PCM mapping requires $n^3 - C_n^3$ simulations. Hence, the number of simulations that can be saved is $\frac{1}{3}n^3 + n^2 + \frac{2}{3}n + C_n^3 = \frac{1}{2}n^3 + \frac{1}{2}n^2 + n$.

C. Discussion on Assumption

We note that the assumption on conditional moments is the key to ensure correct mean prediction in the correlated case (see Lemma 3 and Theorem 6). Here, we use the two-variable case to understand the generality of this condition, and then show examples to illustrate this condition.

In the two-variable case, the rth conditional moment of x can be written as

$$\int_{\alpha(y)}^{\beta(y)} x^r f_{X|Y} f(x|y)dx = \int_{\alpha(y)}^{\beta(y)} \frac{f_{X,Y}(x, y)}{\int_{\alpha(y)}^{\beta(y)} f_{X,Y}(x, y)dx} x^r dx = \frac{\int_{\alpha(y)}^{\beta(y)} f_{X,Y}(x, y)x^r dx}{\int_{\alpha(y)}^{\beta(y)} f_{X,Y}(x, y)dx}. \quad (11)$$



(a) (b) (c)

Fig. 2. Two random variables x and y are uniformly distributed in (a) parallelogram, (b) trapezium, and (c) irregular area. The conditional moment condition is satisfied for cases (a) and (b), but not for (c).

If $f_{X,Y}(x, y)$ does not depend on x [i.e., $f_{X,Y}(x, y)$ is a constant or a function of y], (11) can be further simplified to

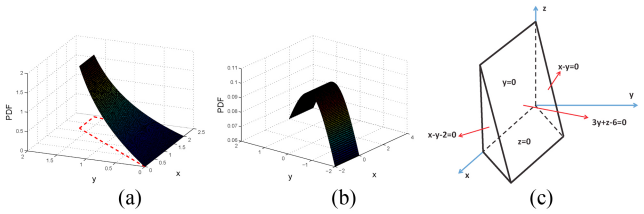
$$\int_{\alpha(y)}^{\beta(y)} x^r f_{X|Y}(x|y)dx = \frac{\frac{1}{r+1}[\beta^{r+1}(y) - \alpha^{r+1}(y)]}{\beta(y) - \alpha(y)} = \frac{1}{r+1}\left(\sum_{i=0}^{r} \beta^i(y)\alpha^{r-i}(y)\right) \quad (12)$$

for $0 \le i \le r$. Clearly, if $\beta(y)$ and $\alpha(y)$ are linear functions of y, the conditional moment is a rth order polynomial of y.

Some examples are shown in Fig. 2. Specifically, in Fig. 2(a), random variables x and y are uniformly distributed in the parallelogram defined by $0 \le y \le 2$ and $y \le x \le y+2$. As $\alpha(y) = y$ and $\beta(y) = y+2$ are linear functions of y, and moreover $f_{X,Y}(x, y) = 0.25$, the conditional moment requirement is clearly satisfied, and hence the low-order multivariate PCM predicts the correct mean. Similarly, in Fig. 2(b), a uniform distribution in the trapezium (e.g., specified by the boundaries $\alpha(y) = y$ and $\beta(y) = 0.25y + 2$) also satisfies the conditional moment requirement. In Fig. 2(c), the variables x and y are subject to a uniform distribution in the region $0 \le y \le 2$, $y \le x \le \sqrt{20y+1}$. As $\beta(y) = \sqrt{20y+1}$ is nonlinear, the conditional moment condition fails. Furthermore, we note that the joint distribution does not have to be uniformly distributed. As shown in Fig. 3(a), $f_{X,Y}(x, y)$ is $\frac{72y}{25(8-3y)}$ in the area defined by $0 \le y \le \frac{5}{3}$ and $y \le x \le 0.25y + 2$. As $f_{X,Y}(x, y)$ is only related to y, the conditional moment condition holds. Fig. 3(b) is another example. In this case, $f_{X,Y}(x, y)$ is $\frac{1}{2}\frac{1}{\sigma\sqrt{2\pi}}e^{-(y-\mu)^2/2\sigma^2}$ in an infinite area defined by two parallel lines (e.g., $y \le x \le y+2$, $-\infty \le y \le \infty$). The last example is concerned with three variables x, y, z. When these three variables are uniformly distributed in a volume [e.g., specified in Fig. 3(c)], the conditional moment condition also holds.

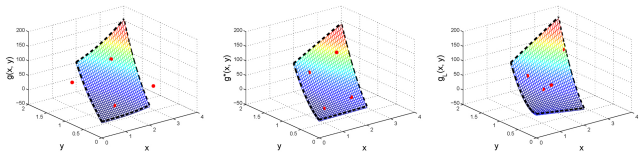
D. Comparative Example

We use a simple example to illustrate the advantage of the correlated multivariate PCM approach. In particular, we consider a two-variable mapping with the original form of $g(x, y) = x^3 + y^3 + xy^2 + x^2y + xy + y + x^2 + y^2 + 1$ as shown in Fig. 4(a). The two random variables x and y are uniformly distributed in the parallelogram area bounded by $0 \le y \le 2$ and $y \le x \le y + 2$, which is marked by the black dash lines in Fig. 4(a) with the joint pdf $f_{X,Y}(x, y) = 0.25$. The independent PCM method does not work, as it results in two simulation points outside the probability range



(a) (b) (c)

Fig. 3. First two are examples when the joint distribution is not uniform. (a) Trapezium. (b) Infinite area bounded by parallel lines. In both cases, the conditional moment requirements are satisfied. In the third example, x, y, and z are uniformly distributed in a (c) 3-D volume. Every surface of this volume is described by a linear function. The conditional moment requirement is satisfied in this case.



(a) (b) (c)

Fig. 4. (a) Original mapping and the four PCM points generated using the independent PCM method. (b) Four PCM points and low-order mapping generated using the correlated PCM approach. (c) Four simulation points and mapping generated by the LHD approach.

[see the red spots in Fig. 4(a)]. The correlated multivariate PCM method following Theorem 6 results in the four PCM points and the reduced-order mapping $1.33x-29.67y+23xy+7$ [shown in Fig. 4(b)] which predicts the correct mean of the original mapping 33.67. For illustration, we have also used the LHD method [35] to generate four simulation points. The produced low-order mapping $34xy-47y-11.96x+27$ and the four simulation points are shown in Fig. 4(c), with a predicted mean of 35.7, different from the original mean.

V. Empirical Data-Based Multivariate PCM

The full pdf of input parameters may not always be available in real applications. Instead, it is more common in practice that the distribution of input parameters is described in one of the following two forms: 1) sample data of input parameters based on historical records and 2) low-order moment description of the parameters. For the completeness of our presentation, we extend the results in [29] and discuss the use of sample-data-based and empirical low-order moment-based multivariate PCM.

A. Large Dataset

The procedure to use the sample-data-based multivariate PCM is summarized as follows: 1) finding from data the moments of x_i (and conditional moments in the correlated case); 2) calculating the (conditional) orthogonal polynomials according to (1) in [29]; 3) finding the roots of the orthogonal polynomials as PCM points; and 4) following Steps 2 and 3 in Section II-B to obtain the low-order PCM mapping in the independent case, or the brief discussion after Theorem 6 in the correlated case.

Theorem 7: Consider that each tuple in the data sets $\{(x_{11}, x_{21}, ..., x_{m1}), (x_{12}, x_{22}, ..., x_{m2}), ..., (x_{1p}, x_{2p}, ..., x_{mp})\}$ is drawn from a joint distribution $f_{X_1,X_2,...,X_m}(x_1, x_2, ..., x_m)$. With the assumption that x_1, x_2, ... and x_m have finite moments, the sample-data-based PCM mapping approaches to the pdf-based PCM mapping with probability 1 as $p \to \infty$.

Proof: To prove that the sample-data-based PCM matches with the pdf-based PCM with probability 1, we only need to show that the PCM points obtained directly using sample data converge to those obtained using the pdf, as $p \to \infty$. As PCM points are the roots of orthogonal polynomials or conditional orthogonal polynomials as shown in the proofs of Theorems 1–6, it suffices to show that the coefficients of the (conditional) orthogonal polynomials obtained using sample data converge to those obtained using the pdf. This problem is thus reduced to the convergence of sample (conditional) moments to real (conditional) moments with increasing size of data. The result follows naturally. Please refer to [29, Theorem 4] for the details of the proof in a single-variable case. ■

B. Small Dataset and Low-Order Moments

In the case that neither the full pdf nor large sample data sets is available, we may find PCM mapping based on very limited information such as small data sets, or low moments such as the mean and variance. Please refer to [29] for a detailed treatment of PCM with sparse data or low moments. We summarize the key procedure for the complete presentation here.

The best approach to find PCM mappings in these cases is to fit these low moments/small sample datasets with typical probability distributions (i.e., those are mathematically determined by very small number of parameters, such as the Gaussian, uniform, Gamma, and Beta distributions). Then, we can follow the same procedure as that of the pdf-based PCM to find low-order PCM mappings. Please refer to standard estimation approaches such as the maximum-likelihood estimation [20].

The probability distribution to choose is generally not easy due to the sparsity of information. Here, we briefly discuss the guidelines to select distribution functions. First, the selection should be application-dependent (see [20]). For instance, noise is typically captured by the Gaussian distribution, whereas waiting time is captured by the Gamma distribution. Second, hypothesis testing techniques can help to determine the best distribution. It also helps to understand the impact of choosing different distribution functions. One example suggests that choosing the Gaussian distribution instead of the uniform distribution will result in more-spread-out PCM points that better account for parameter values further away from the mean, but at a cost of worse estimation close to the mean. Rigorous analysis on the impact of higher moments on the movement of PCM points have been studied in [29].

VI. Applications to Air Traffic Flow Management

The development of multivariate PCM in this paper was motivated by practical needs in the field of air traffic flow management (ATFM), which is typically concerned with managing traffic flows at a long lookahead time frame

(e.g., 2–15 h). As traffic flow management plans are decided hours in advance, a wide range of weather possibilities exist. Therefore, it is important to effectively evaluate the performance of the air traffic system under uncertainty, so as to design the best management strategies that are robust to such uncertainty in real time. Here, we first provide a brief review of the literature on the evaluation and design of ATFM strategies at the long lookahead time, so as to make clear the contribution of our proposed approach.

Simple models of uncertainty as scenario trees and disturbances have been used to model uncertainty in ATFM; stochastic programming and model predictive control techniques have been applied on these models for decision-making [1]–[3], [8], [9]. More complicated stochastic models such as Bayesian networks and Markov chains have also been used [21], [38]. As pointed out in these papers, these methods become less effective when the scale and complexity of the traffic system increase. As the air traffic systems have complex nonlinear and intertwined dynamics [30], predicting their performance may require system-wide simulations. The Monte Carlo method has thus been studied [10], [22], [31]. The Monte Carlo approach does not enable real-time analysis under weather uncertainty, as the evaluation of system performance under multiple (and even spatiotemporally correlated) weather uncertainties requires time-consuming simulations for a large number of weather ensembles to obtain converging performance statistics. The multivariate PCM approach provides an alternative effective way for such simulations.

In this section, we demonstrate the use of multivariate PCM to evaluate the performance of air traffic systems under weather uncertainties using two examples. As seen from these examples, an important preliminary step for applying the multivariate PCM is to capture the range of weather uncertainties using a number of random variables. Next, the PCM points are selected along the dimension of each variable/parameter and then these selected points are simulated to provide a fast evaluation of traffic system performance. In the first simple example, we assume that the basic properties of an uncertain weather zone are captured by two independent variables (intensity and duration) with simple distributions, and then the total number of delayed aircraft (called backlog) for the flow entering this single zone is evaluated. The purpose of this example is not to precisely model uncertain weather, but to: 1) provide some background on air traffic flow management, including the performance metric and the simulation model and 2) show the practical use of PCM in the independent case. The second more realistic example is concerned with the evaluation of backlogs in two spatiotemporally-correlated weather zones.

A. Simple Single Region Example

In this simple example, we consider a single region (e.g., a sector) in the airspace subject to severe weather prediction. We aim to evaluate the impact of uncertain weather on the congestion in the region. Let us first describe the simple queuing model for our study. We note that realistic air traffic system is much more complicated. However, the abstracted model is justified based on the uncertainty present in the data at this time frame and the computational requirements of real-time systems [28], [30], [38].

In particular, severe weather events reduce the capacity $N[k]$ (i.e., the maximum number of aircraft allowed to enter at time k) of the region. The aircraft that are not allowed to pass due to the reduced capacity can be modeled as waiting in the queue at the boundary of the region [28], [30], [38]. The length of the queue, named backlog $B[k]$, is a natural performance metric to capture the congestion of the region. The dynamics of the queuing simulation model is shown below (see more details at [38])

$$B[k] = \max(B[k-1] + x[k] - N[k], 0) \tag{13}$$

where $x[k]$ is the incoming flow. The equation suggests that at each time step, a maximum number of $N[k]$ aircraft can enter the region; the remaining flow will wait to enter at the next time step, together with the new incoming flow. We note that the capacity $N[k]$ varies with time k; in particular, the presence of convective weather at time k reduces the capacity $N[k]$.

As weather duration and intensity may not be precisely predicted at a long lookahead time, weather-induced capacity reduction and the accumulated transient backlog become uncertain. It is of practical value to predict the backlog statistics under uncertain weather duration and intensity.

Assuming that no information about the two uncertain random variables is available other than the ranges, we model both variables as uniformly distributed for simplicity. In particular, we assume that the severe weather starts at the current time, but with uncertain intensity and end time. When the severe weather is present, it reduces the number of aircraft to cross the region in a unit time (20 min) $N[k]$ from 6 to a, where a is captured by a random variable uniformly distributed between 1 and 2.67. Moreover, we assume that the duration of severe weather is a random variable uniformly distributed between 2 and 6 h. When the severe weather completes, the capacity raises back to 6. For simplicity, we assume that the incoming flow (demand) $x[k]$ to the region is a deterministic sequence sampled from a Poisson process with the mean arrival rate of 3 per 20 min.

It is worthwhile to notice that the actual relationship between the output (total backlog for a span of 30 h from the current time) and inputs (weather intensity and duration) is nonlinear [28], [38]. We show here that based on the distributions of weather intensity and duration, we can smartly choose only a limited number of simulations to construct a low-order multivariate PCM mapping that matches well with the original nonlinear mapping obtained using Monte Carlo simulations. Fig. 5(a) and (b) show the original mapping and the 2×2-order PCM mapping using the algorithm in Section II-A1. The two mappings match very well with a mean square error of 2.4 within the probability range. The mean total backlog obtained using the two mappings are 163.91 and 163.42, respectively. It is also interesting to notice that the performance of mean predictor improves significantly when the order of the PCM mapping increases from 1×1 (with the mean backlog of 165.52) to 2×2, but not much when

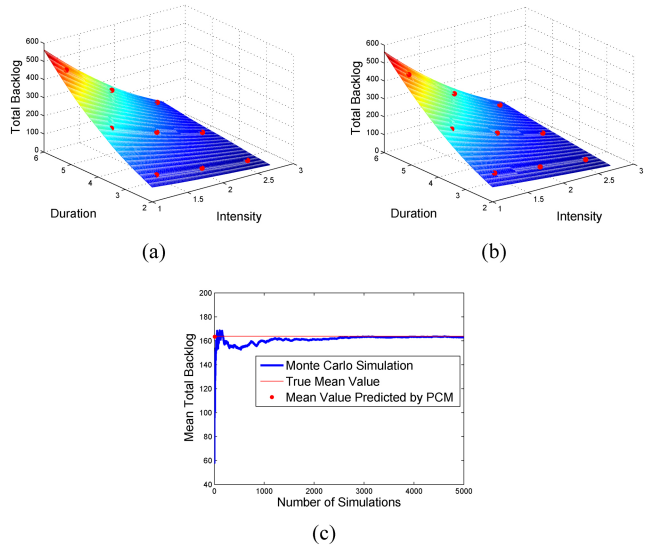


Fig. 5. (a) Original mapping between the input set including weather duration and weather intensity, and the output (30-h total backlog) obtained using the Monte Carlo simulation. (b) Reduced-order multivariate PCM mapping is based on the 9 sample points marked as red spots on the plot. (c) Comparison of the number of simulation runs needed to predict the correct mean total backlog over a 30-h span. The blue curve shows the means predicted by the accumulative Monte Carlo runs with the number specified on the x-axis. The red thin line shows the true mean estimated by the value that the Monte Carlo simulation (of two random variables) converges to. The red spot corresponds to the nine simulation runs required for the PCM method to predict the true mean.

the order is further increased to 3×3 (with the mean backlog of 163.64). Finally, Fig. 5(c) demonstrates the computational load that can be saved using the low-order 2×2 PCM. Monte Carlo simulation typically requires a large number of runs to converge (in this case 676 runs to estimate with an error less than 2%), while the 2×2-order PCM mapping only needs nine simulations to achieve a similar performance.

B. Spatiotemporally Correlated Two-Region Example

In the second example, we apply empirical data-based multivariate PCM to effectively evaluate the impact of spatiotemporally correlated uncertain weather events on transient traffic backlogs. In particular, we consider two scheduled streams of flows entering two neighboring regions A and B. Furthermore, a cold front (indicating of weather front associated with convective weather) passes Region A and then Region B, producing capacity reductions [see Fig. 6(a) and (b)]. Predicting transient backlog statistics caused by the uncertain progress of cold front is critical to design effective flow-management initiatives to reduce traffic delay.

Here, we used six random parameters to capture the stochastic weather propagation: S_A, N_A, D_A, S_B, N_B, and D_B. In particular, the cold front hits Region A from time S_A with a reduced capacity N_A. We assume that the capacity reduction remains constant for a span of D_A before the weather leaves the region. After that, its capacity returns to its regular value N_{RA}. Similarly, Region B undergoes weather-induced capacity reduction from time S_B, with capacity reduced from the normal value N_{RB} to N_B, for a span of D_B.

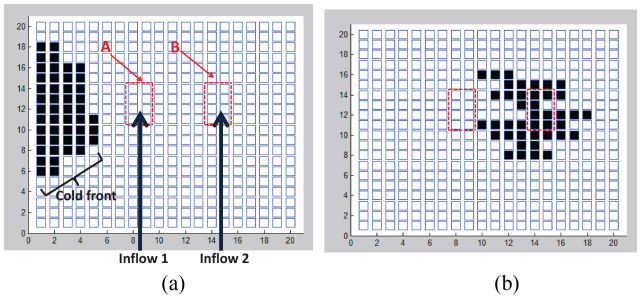


Fig. 6. (a) Illustration of a spatiotemporal correlated two-region example. (b) Development of cold font causes capacity reduction.

Different from the previous example, the full pdf of the aforementioned weather parameters is not directly available. We utilized the weather simulator [24] to generate a large set of weather ensembles covering the range of weather uncertainty. In particular, the simulated airspace is decomposed into small grids, with white color representing normal capacity and black color denoting capacity reduction caused by convective weather (see Fig. 6). The stochastic propagation of convective weather is governed by the influence model [4], with parameters estimated from hourly probabilistic weather forecasts.[1] The estimated weather simulator runs at a finer resolution of 15 min [24].

For each generated weather ensemble, we then found the six weather parameters (S_A, N_A, D_A, S_B, N_B, and D_B). Here, weather start times S_A and S_B are marked by the first time the region has at least two black grids. Durations D_A and D_B are similarly defined by the differences between the first time the region has at most two black grids afterward and S_A and S_B. Average capacities N_A and N_B during the span of weather are calculated as the normal capacity scaled by the average fraction of white grids during this span. Due to the spatiotemporal correlation of weather propagation, these parameters are subject to interdependency. In particular, as the weather front passes Region A and then region B, due to the spatial weather propagation delay, we expect the start time S_B is closely related to the end time of region A (which is expressed as $S_A + D_A$), and therefore both S_A and D_A. Similarly, due to the correlation of propagation speeds for weather in these two regions, we expect that the duration D_B is dependent on D_A. Moreover, due to the correlation of weather intensities across time, we assume that the capacity N_B is dependent upon N_A. In this example, the normal capacity N_{RA} and N_{RB} of the two regions are both 10 per unit time (defined as $\Delta t = 15$ min). The simulator generates a large number of ensembles (one million in this case to guarantee the convergence of probability distributions).

We then applied the empirical data-based multivariate PCM approach to find the PCM points. To do that, we first found the sample moments of D_A, S_A, and N_A, as they serve as the parameters that D_B, S_B, N_B are conditioned upon. From these sample moments, the PCM points along each of the

[1] A majority of weather forecasts are deterministic. The Very Short Range Ensemble Forecast System (VSREF) provides hourly probabilistic weather forecasts [36].

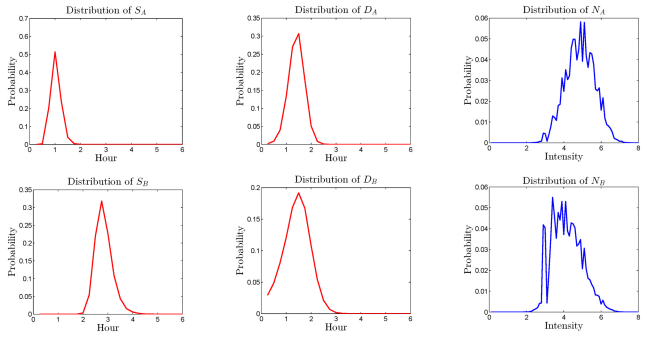


Fig. 7. Probability density functions of six weather parameters.

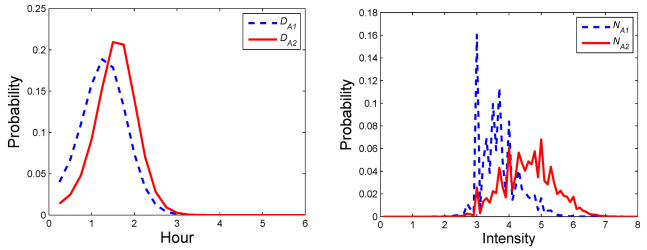


Fig. 8. (a) Distribution of D_B conditioned upon two selected PCM coordinates of D_A. (b) Distribution of N_B conditioned upon two selected PCM coordinates of N_A.

dimensions D_A, S_A, and N_A were obtained. We note that S_A plays a more important role in the uncertainty evaluation and needs to be sampled with more number of points for the following reasons: 1) due to the growth of uncertainty over time (which can be seen from the distributions plotted in Fig. 7), the information closest to the initial simulation time is the most trustworthy and 2) this trustworthy early information is important to be estimated correctly as its error will affect the estimation performance of other parameters capturing later time characteristics. We therefore chose 5 points along S_A as 0.75 h, 1 h, 1.25 h, 1.75 h, and 2.5 h. We also selected 2 PCM points along the dimension of D_A as 1 h and 1.75 h, and 2 points for N_A as 4.0602, and 5.6358, respectively. Next, we identified PCM points along the dimensions of D_B, S_B and N_B. Due to the spatiotemporal correlation of weather parameters in Region A and Region B, each 3-tuple PCM point (D_A, S_A, and N_A) resulted in different conditional distributions of weather parameters for Region B (see examples shown in Fig. 8), and thus different set of PCM points. Specifically, the PCM points along the dimensions of D_B, S_B were then selected based on the conditional sample moments of D_B and S_B for each combination of the PCM points for D_A and S_A. Similarly, the PCM points along the dimension of N_B were selected from the conditional sample moments of N_B on each of the PCM points for N_A. A total of $2^5 \times 5 = 160$ PCM points were selected.

We next evaluated the total backlog $\sum B[k]$ during a span of 24 h at each of the 6-tuple PCM points. We assumed that two streams of deterministic flows sampled from a Poisson process (with mean 5 per 15 min) enter the two regions. Using (13), the total backlogs at all PCM points were then used to obtain the low-order polynomial PCM mapping $f(D_A, S_A, N_A, D_B, S_B, N_B)$ (with 160 terms) between weather parameters and the total backlog, from which the mean backlog can be obtained. Furthermore, we noted that as the total backlog of the two regions is the summation of backlogs at each region, the mapping $f(D_A, S_A, N_A, D_B, S_B, N_B)$ can be expressed as $f_A(D_A, S_A, N_A)+f_B(D_B, S_B, N_B)$, where f_A and f_B are functions of weather parameters in individual regions. This expression significantly reduced the coefficients/terms in the polynomial mapping function from 160 to $5 \times 2^2+2^3 = 28$, by ignoring the cross terms that involve weather parameters of both regions. Least square mean estimation was then used to identify the 28 coefficients in the reduced order mapping.

TABLE II
PERFORMANCE COMPARISON AMONG THE MONTE CARLO AND PCM APPROACHES WITH TWO DIFFERENT EXPRESSIONS

	Backlog of Region A	Backlog of Region B	Total Backlog
Monte Carlo Simulation	40.7	72.09	112.8
PCM with 160 terms	40.5	74.2	114.7
PCM with 28 terms	40.5	74	114.5

We compared the performance of the PCM mappings with that of the Monte Carlo simulation. As shown in Table II, both PCM mappings predicted the total mean backlogs well with errors less than 2%. We also evaluated the efficiency of the PCM approaches. As seen in Fig. 9, the Monte Carlo method requires 11716 simulations (marked as the black spot) for the mean prediction to fall within a threshold marked by the black dashed lines. However, the PCM method only needs 160 simulations (marked as the red spot) to reach the same threshold. We also note that selecting PCM points based on conditional distributions improves the performance of the mean prediction. To display the results intuitively, we demonstrate the mappings generated by the aforementioned empirical data-based multivariate PCM approaches, when four out of the six weather parameters are fixed: $S_A = 1.25$ h, $D_A = 1$ h, $D_B = 1.75$ h, and $N_A = 5.6358$. As shown in Fig. 10, the 160-term PCM mapping and the 28-term PCM mapping generated by selecting four PCM points along the dimensions of S_B and N_B match well with the mapping generated by the Monte Carlo method.

VII. CONCLUSION

Motivated by effective uncertainty evaluation needs in large-scale infrastructure systems, we develop in this paper the multivariate PCM approach which allows using a few smartly selected simulation points to construct a low-order mapping between multiple uncertain input parameters and system output, which predicts the correct mean output of the original system of a higher-order. Besides describing the algorithm of the multivariate PCM, we develop mathematical conditions to permit correct mean output prediction in terms of the probability distributions of the input parameters and forms of original system mappings. Both independent and correlated cases are discussed. We also provide additional performance analysis of the multivariate PCM in terms of predicting other important statistics, and the practical use of the method when data or low-order moments are available instead of probability

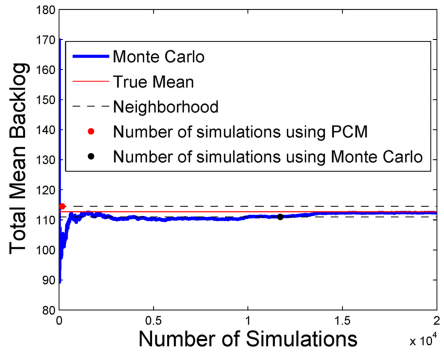


Fig. 9. Comparison of simulation time between the Monte Carlo and PCM approaches. The mean backlog estimated by the Monte Carlo method is shown in blue. The mean backlog estimated using the PCM approach is marked as the red spot. The dashed lines show a neighborhood around the true mean defined by the estimated mean using the PCM approach, marked by the red spot. The black spot denotes the number of simulations for the Monte Carlo method to reach and stay within the neighborhood.

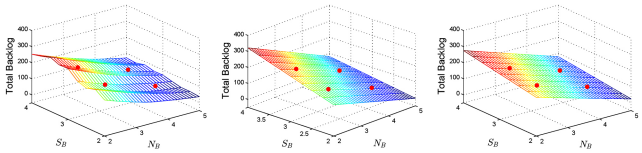


Fig. 10. Mappings with fixed S_A, D_A, D_B, and N_A generated by (a) Monte Carlo simulation (b) PCM mapping with 160 terms, and (c) PCM mapping with 28 terms.

distribution. In terms of numerical issues, we note that in the independent PCM case, using orthonormal bases to calculate PCM mapping coefficients reduces the ill-conditioning of the calculation; this advantage does not exist for the correlated PCM. The final example at the end of the paper demonstrates the use of this method to evaluate the impact of multiple dependent weather uncertain parameters on the statistics of air traffic system performance. Future works include applying this method for large-scale traffic examples, exploring capabilities of the reduced-order mapping such as optimization and sensitivity analysis, and selecting a subset of the PCM points when further knowledge of the original system mapping is available.

Appendix A
Proof of Theorem 1

Proof: In order to prove $E[g(x, y)] = E[g^*(x, y)]$, we start with computing $E[g(x, y)]$, then construct $g^*(x, y)$ along this process, and finally show that the means of both $g(x, y)$ and $g^*(x, y)$ can be reduced to the same value.

As x and y are independent random variables, we find

$$E[g(x,y)] = \iint \sum_{i=0}^{2n-1}\sum_{j=0}^{2m-1} A_{i,j}x^i y^j f_X(x) f_Y(y)dxdy = \int \sum_{j=0}^{2m-1} y^j f_Y(y) \int \sum_{i=0}^{2n-1} A_{i,j}x^i f_X(x)dxdy. \quad (14)$$

The terms inside the internal integral $\sum_{i=0}^{2n-1} A_{i,j}x^i$ (for any j) can be expressed in terms of the series of orthonormal polynomials $h_n(x), ..., h_0(x)$ of degrees $n, ..., 0$ [12]

$$\sum_{i=0}^{2n-1} A_{i,j}x^i = h_n(x)\left(\sum_{i=0}^{n-1} a_{i+n,j}h_i(x)\right) + \sum_{i=0}^{n-1} a_{i,j}h_i(x) \quad (15)$$

where the coefficients $a \in R$, the first subscript of a represents the degree of x, and the second subscript stands for the degree of y. The orthonormal polynomials $h_i(x)$ satisfy

$$\langle h_i(x), h_j(x)\rangle = \int h_i(x)h_j(x)f_X(x)dx = \begin{cases}1, & if\ i=j\\ 0, & if\ i\neq j\end{cases} \quad h_0(x)=1. \quad (16)$$

A particular note is that the above definition leads to

$$\int h_i(x)h_0(x)f_X(x)dx = \int h_i(x)f_X(x)dx = 0 \quad (17)$$

for all $i \geq 1$, which we will frequently use later. We denote the roots of $h_n(x)$ as $x_1, x_2, ..., x_n$. In a single-variable PCM, these roots are the PCM points selected for computationally intensive simulations [29].

Due to the orthonormal properties of the variable x (16), $\int \sum_{i=0}^{2n-1} A_{i,j}x^i f_X(x)dx$ shown in the form of (15) can be reduced to $\int \sum_{i=0}^{n-1} a_{i,j}h_i(x)f_X(x)dx$. As such, (14) becomes

$$E[g(x,y)] = \int \sum_{j=0}^{2m-1} y^j f_Y(y) \int \left(\sum_{i=0}^{n-1} a_{i,j}h_i(x)\right) f_X(x)dxdy. \quad (18)$$

We now follow the same procedure to reduce the order of y. By rearranging the terms in (18) in a descending degree of $h_i(x)$ for all $i \in \{0, ..., n-1\}$, we obtain

$$E[g(x,y)] = \int \sum_{i=0}^{n-1} h_i(x)f_X(x) \int \left(\sum_{j=0}^{2m-1} a_{i,j}y^j\right) f_Y(y)dydx. \quad (19)$$

Introducing the jth-degree orthonormal polynomials $h'_j(y)$ with respect to the probability distribution $f_Y(y)$, and denoting the roots of $h'_m(y)$ as $y_1, y_2, ..., y_m$, we can express any polynomial of y up to the order of $2m-1$ in terms of $h'_0(y), ..., h'_m(y)$. In particular

$$\sum_{j=0}^{2m-1} a_{i,j}y^j = h'_m(y)\left(\sum_{j=0}^{m-1} b_{i,j+m}h'_j(y)\right) + \sum_{j=0}^{m-1} b_{i,j}h'_j(y) \quad (20)$$

where $b \in R$ are the coefficients. Again, the first subscript of b represents the degree of x, and the second stands for the total degree of y. Applying orthonormal properties again, (19) is further reduced to

$$E[g(x,y)] = \int \int \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{i,j}h_i(x)h'_j(y)f_X(x)f_Y(y)dxdy. \quad (21)$$

Note that the expression inside the above double integrals is $g^*(x, y)$. As $h_i(x)$ is an ith-order polynomial and $h'_j(y)$ is a

jth-order polynomial, we can easily find the parameters $B_{i,j}$ from $b_{i,j}$ and express (21) in terms of x and y as

$$g^*(x, y) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{i,j}h_i(x)h'_j(y) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} B_{i,j}x^i y^j. \quad (22)$$

Observation of (15) and (20) clearly suggests that $g(x, y)$ and $g^*(x, y)$ pass through the same set of points defined by (x_i, y_j) for all $i \in \{1, ..., n\}$ and $j \in \{1, ..., m\}$, as $h_n(x_i) = 0$ and $h'_m(y_i) = 0$. Finally, we also notice that both $E[g(x, y)]$ and $E[g^*(x, y)]$ can be reduced to $b_{0,0}$ through applying (17) to variable x first and then variable y. ■

Appendix B
Proof of Theorem 2

Proof: As this theorem is a straightforward generalization of the two-variable case, we only sketch the outline of the proof.

Through recursively expressing x_1, x_2,... x_m in terms of orthonormal polynomials, and applying the orthonormal properties (16), we obtain the reduced-order mapping $g^*(x_1, x_2, ..., x_m)$ of the form

$$g^*(x_1, x_2..., x_m) = \sum_{k_1=0}^{n_1-1}\sum_{k_2=0}^{n_2-1}...\sum_{k_m=0}^{n_m-1} a_{k_1,...,k_m}\prod_{i=1}^{m} h^i_{k_i}(x_i) \quad (23)$$

where the function $h^i_{k_i}(x_i)$ denotes the k_ith degree orthonormal polynomial with respect to the parameter x_i, and $a_{k_1,...,k_m}$ are the coefficients. Furthermore

$$E\left[g(x_1, x_2, ..., x_m)\right] = E\left[g^*(x_1, x_2, ..., x_m)\right] = a_{0,...,0}. \quad (24)$$

Rearranging the terms in 23 in terms of $x_1, x_2, ..., x_m$, we can easily find the coefficients $B_{k_1,...,k_m}$ such that $g^*(x_1, x_2, ..., x_m)$ is in the form of (4). ■

Appendix C
Proof of Theorem 3

We prove the two-variable case for the clarity of presentation. In this case, the theorem is reduced to the following Lemma. The general case follows naturally.

Lemma 1: Consider the use of a two-variable PCM mapping $g^*(x, y)$ of the form $\sum_{i=0}^{n-1}\sum_{j=0}^{m-1} B_{i,j}x^i y^j$ to approximate an original polynomial mapping $g(x, y)$ of the form $\sum_{i=0}^{n+n_1}\sum_{j=0}^{m+m_1} A_{i,j}x^i y^j$, for some $n_1 \in \{0, ..., n-1\}$ and $m_1 \in \{0, ..., m-1\}$. Assuming that the two variables are independent, the low-order PCM can correctly predict the cross-statistics up to certain degree. In particular

$$E\left[x^l y^k g^*(x, y)\right] = E\left[x^l y^k g(x, y)\right] \quad (25)$$

for all $l \in \{0, ..., n-1-n_1\}$ and $k \in \{0, ..., m-1-m_1\}$.

Proof: The cross-statistics can be expressed in the following due to the independence of variables:

$$\begin{aligned} E\left[x^l y^k g(x, y)\right] &= \iint x^l y^k \sum_{i=0}^{n+n_1}\sum_{j=0}^{m+m_1} A_{i,j}x^i y^j f_X(x) f_Y(y)dxdy \\ &= \int y^k \sum_{j=0}^{m+m_1} y^j f_Y(y) \int x^l \sum_{i=0}^{n+n_1} A_{i,j}x^i f_X(x)dxdy. \end{aligned} \quad (26)$$

For any fixed j, $\sum_{i=0}^{n+n_1} A_{i,j}x^i$ can be represented in terms of orthonormal polynomials [similar to (15)]

$$\sum_{i=0}^{n+n_1} A_{i,j}x^i = h_n(x)\left(\sum_{i_1=0}^{n_1} a_{n+i_1,j}h_{i_1}(x)\right) + \sum_{i=0}^{n-1} a_{i,j}h_i(x) \quad (27)$$

where the first subscript in the coefficient a represents the total degree of this term with respect to x, and the second represents the degree with respect to y inside the summation operator in (26). As $x^l\left(\sum_{i_1=0}^{n_1} a_{n+i_1,j}h_{i_1}(x)\right)$ is a polynomial with the degree of x less than or equal to $n-1$, where $j \in \{0, ..., m+m_1\}$, the orthogonality naturally leads to

$$\begin{aligned} E\left[x^l y^k g(x, y)\right] &= \int y^k \sum_{j=0}^{m+m_1} y^j f_Y(y) \int x^l \sum_{i=0}^{n-1} a_{i,j}h_i(x) f_X(x)dxdy \\ &= \int x^l \sum_{i=0}^{n-1} h_i(x) f_X(x) \int y^k a_{i,j} \sum_{j=0}^{m+m_1} y^j f_Y(y)dydx. \end{aligned} \quad (28)$$

Next, we reduce the order of y to $m-1$ similar to the above process. It is then not difficult to obtain

$$\begin{aligned} E\left[x^l y^k g(x, y)\right] &= \int x^l \sum_{i=0}^{n-1} h_i(x) f_X(x) \int y^k b_{i,j} \sum_{j=0}^{m-1} h'_j(y) f_Y(y)dydx \\ &= \iint x^l y^k \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{i,j}h_i(x)h'_j(y) f_X(x) f_Y(y)dxdy \\ &= E\left[x^l y^k g^*(x, y)\right]. \end{aligned} \quad (29)$$

■

Appendix D
Proof of Theorem 4

Again, we prove the two-variable case shown in the following lemma. The proof of the general case follows naturally.

Lemma 2: Consider a mapping $g(x, y) = \sum_{i=0}^{n+n_1}\sum_{j=0}^{m+m_1} A_{i,j}x^i y^j$ for some $n_1 \in \{0, ..., n-1\}$ and $m_1 \in \{0, ..., m-1\}$. If a PCM mapping of the form $g^*(x, y) = \sum_{i=0}^{n-1}\sum_{j=0}^{m-1} b_{i,j}h_i(x)h'_j(y)$ is used to fit the original mapping, the MSE of the PCM fit cannot be improved by adding any polynomial with the degree of x up to $n-1-n_1$, and the degree of y up to $m-1-m_1$. Moreover, the lower order mapping $g^*_r(x, y) = \sum_{i=0}^{n-1-n_1}\sum_{j=0}^{m-1-m_1} b'_{i,j}h_i(x)h'_j(y)$ is the MMSE mapping, among all polynomials with the degree of x up to $n-1-n_1$ and the degree of y up to $m-1-m_1$.

Proof: To prove the first part of the theorem, we construct $\bar{g}(x, y) = g^*(x, y) + \sum_{i=0}^{n-1-n_1}\sum_{j=0}^{m-1-m_1} C_{i,j}x^i y^j$, and show that

$E\left[\left(g(x,y)-g^*(x,y)\right)^2\right] \le E\left[\left(g(x,y)-\bar{g}(x,y)\right)^2\right]$. Here, the coefficients $C_{i,j} \in R$. The mean square error between $g(x,y)$ and $\bar{g}(x,y)$ can be expressed as

$$\begin{aligned} E\left[\left(g(x,y)-\bar{g}(x,y)\right)^2\right] &= E\left[\left(g(x,y)-g^*(x,y)\right)^2\right] \\ &+2E\Big[\left(g(x,y)-g^*(x,y)\right) \\ &\left(g^*(x,y)-\bar{g}(x,y)\right)\Big] \\ &+E\left[\left(g^*(x,y)-\bar{g}(x,y)\right)^2\right]. \end{aligned} \tag{30}$$

We note that $E\left[\left(g^*(x,y)-\bar{g}(x,y)\right)^2\right]$ is always nonnegative. Therefore, it is sufficient to show $E\Big[\left(g(x,y)-g^*(x,y)\right)\left(g^*(x,y)-\bar{g}(x,y)\right)\Big]=0$. To do that, we first notice

$$g^*(x,y)-\bar{g}(x,y) = -\left(\sum_{i=0}^{n-1-n_1}\sum_{j=0}^{m-1-m_1} C_{i,j}x^i y^j\right). \tag{31}$$

Moreover, $g(x,y)-g^*(x,y)$ is the sum of three terms: $\sum_{i=0}^{n_1}\sum_{j=0}^{m_1} b_{n+i,m+j}h_n(x)h_i(x)h'_m(y)h'_j(y)$, $\sum_{i=0}^{n_1}\sum_{j=0}^{m-1} b_{n+i,j}h_n(x)h_i(x)h'_j(y)$, and $\sum_{i=0}^{n-1}\sum_{j=0}^{m_1} b_{i,m+j}h_i(x)h'_m(y)h'_j(y)$. As each $h_i(x)$ in the first two terms is a polynomial of degree at most n_1, $-h_i(x)\left(\sum_{i=0}^{n-1-n_1}\sum_{j=0}^{m-1-m_1} C_{i,j}x^i y^j\right)$ is a polynomial of degree at most $n-1$ of variable x. Applying orthogonality with respect to the variable x, it is straightforward to find $E\Big[\left(g(x,y)-g^*(x,y)\right)\left(g^*(x,y)-\bar{g}(x,y)\right)\Big]$ equals to zero for the first two terms of $g(x,y)-g^*(x,y)$. Using the same argument for the third term and apply orthogonality to the variable y, we can find that $E\left[\left(g(x,y)-g^*(x,y)\right)\left(g^*(x,y)-\bar{g}(x,y)\right)\right]$ also equals to 0 for the third term.

We next prove the second part of the theorem. This is sufficient to show that $E\left[x^l y^k\left(g(x,y)-g_r^*(x,y)\right)\right]$ equals to zero, for $l \in \{0,...,n-1-n_1\}$ and $k \in \{0,...,m-1-m_1\}$. Notice that

$$\begin{aligned} E\left[x^l y^k\left(g(x,y)-g_r^*(x,y)\right)\right] &= E\left[x^l y^k\left(g(x,y)-g^*(x,y)\right)\right] \\ &+E\left[x^l y^k\left(g^*(x,y)-g_r^*(x,y)\right)\right]. \end{aligned} \tag{32}$$

The first term is zero according to Theorem 3. The second term is also zero according to a proof similar to that of the first part of this theorem. In particular, each term in $g^*(x,y)-g_r^*(x,y)$ contains $h_i(x)h'_j(y)$ with degrees either $i \ge n-n_1$ or $j \ge m-m_1$. Invoking orthogonality, the second part is proved as well. ■

APPENDIX E
PROOF OF THEOREM 6

To ease understanding, we first investigate the correlated two-variable PCM in Lemma 3.

Lemma 3: Consider a two-variable mapping $g_c(x,y)$ of the form

$$g_c(x,y) = \sum_{j=0}^{2n-1}\sum_{i=0}^{2n-1-j} A'_{i,j}x^i y^j \tag{33}$$

where the coefficients $A'_{i,j} \in R$, and n is an integer greater than 1. Assume that the two variables x and y follow a joint distribution $f_{X,Y}(x,y)$ and the rth conditional moment of one variable given the other is at most a rth degree polynomial of the other variable. Then the mapping $g_c(x,y)$ can be approximated by a lower-order PCM mapping $g_c^*(x,y)$ of the form

$$g_c^*(x,y) = \sum_{j=0}^{n-1}\sum_{i=0}^{n-1} B'_{i,j}x^i y^j \tag{34}$$

such that $E\left[g_c(x,y)\right] = E\left[g_c^*(x,y)\right]$, where the coefficients $B'_{i,j} \in R$.

Proof: Without loss of generality, we assume that $\int x^r f_{X|Y}(x|y)dx$ is a rth degree polynomial of y. We first construct the expression of $g_c^*(x,y)$ from $g_c(x,y)$ without changing its mean. We then verify that $g_c(x,y)$ and the constructed $g_c^*(x,y)$ have the same values at the set of PCM points selected from the marginal and conditional probabilities. We can thus use these PCM points to uniquely identify $g_c^*(x,y)$.

To calculate the mean of the original mapping $g_c(x,y)$, we note that because $E\left[g_c(x,y)\right]$ can be expressed as $\sum_{j=0}^{2n-1}\sum_{i=0}^{2n-1-j} E\left[A'_{i,j}x^i y^j\right]$, we can focus on the calculation of each term $E\left[A'_{i,j}x^i y^j\right]$ first. To simplify $E\left[A'_{i,j}x^i y^j\right]$, we consider three cases with different ranges of i and j: Case 1 ($0 \le i < n$, $n-1 < j \le 2n-1$), Case 2 ($n-1 < i \le 2n-1$, $0 \le j < n$), and Case 3 ($0 \le i < n$, $0 \le j < n$). As in Case 3, the polynomial is already in the form of $g_c^*(x,y)$, we only focus on the first two cases. We show the calculation of $E\left[A'_{i,j}x^i y^j\right]$ for Case 1 only, as Case 2 follows a similar procedure.

In Case 1, when $n-1 < j \le 2n-1$, we can express y^j using the orthogonal polynomials of y. In particular, $y^j = h_n(y)\left(\sum_{k=0}^{j-n} b'_{j,k+n}h_k(y)\right) + \sum_{k=0}^{n-1} b'_{j,k}h_k(y)$, where the coefficients $b' \in R$, the first subscript of b' stands for the total degree of y, and the second represents the degree of y in each term. Moreover, as $0 \le i < n$, for each particular y, we can write x^i in terms of the conditional orthogonal polynomials of x given y, denoted as $h'_k(x|y)$, where $k \in \{0,...,i\}$ represents the degree of the orthogonal polynomial. In particular, $A'_{i,j}x^i = \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)$, where $(a'_{i,l}|y)$ is a parameter dependent upon y, and the format of the subscripts is the same as that of b', but for the variable x. The conditional orthogonal polynomial $h'_l(x|y)$ is associated with the conditional pdf $f_{X|Y}(x|y)$. As such, at different y, $h'_l(x|y)$ may have different expressions. Now, we are ready to calculate $E\left[A'_{i,j}x^i y^j\right]$ as

$$\begin{aligned} &E\left[A'_{i,j}x^i y^j\right] \\ &= \iint A'_{i,j}x^i y^j f_{X|Y}(x|y)f_Y(y)dxdy = \iint \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y) \\ &\left(h_n(y)\Big(\sum_{k=0}^{j-n} b'_{j,k+n}h_k(y)\Big) + \sum_{k=0}^{n-1} b'_{j,k}h_k(y)\right) f_{X|Y}(x|y)f_Y(y)dxdy \\ &= \iint \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)\left(h_n(y)\Big(\sum_{k=0}^{j-n} b'_{j,k+n}h_k(y)\Big)\right) \end{aligned}$$

$$f_{X|Y}(x|y)f_Y(y)dxdy + \iint \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)\left(\sum_{k=0}^{n-1} b'_{j,k}h_k(y)\right)$$

$$f_{X|Y}(x|y)f_Y(y)dxdy = \int h_n(y)\left(\sum_{k=0}^{j-n} b'_{j,k+n}h_k(y)\right)$$

$$\left(\int \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)f_{X|Y}(x|y)dx\right) f_Y(y)dy$$

$$+ \iint \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)\left(\sum_{k=0}^{n-1} b'_{j,k}h_k(y)\right) f_{X|Y}(x|y)f_Y(y)dxdy. \tag{35}$$

Now let us show that the first double integration in the last equation equals 0. Note that for each y, $\sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)$ is a polynomial of x with the degree up to i. Furthermore, according to the assumption that $\int x^i f_{X|Y}(x|y)dx$ is at most an ith degree polynomial of y, we find that $\left(\sum_{k=0}^{j-n} b'_{j,k+n}h_k(y)\right)$ $\left(\int \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)f_{X|Y}(x|y)dx\right)$ is a polynomial of y with the degree up to $i + j - n \le 2n - 1 - n = n - 1$. Applying orthogonality with respect to variable y, we can simplify (35) to

$$E[A'_{i,j}x^i y^j] = \iint \sum_{l=0}^{i}(a'_{i,l}|y)h'_l(x|y)\left(\sum_{k=0}^{n-1} b'_{j,k}h_k(y)\right) f_{X|Y}(x|y)f_Y(y)dxdy$$

$$= \iint \sum_{l=0}^{i}\sum_{k=0}^{n-1}(c'_{l,k}|y)h'_l(x|y)h_k(y)f_{X,Y}(x,y)dxdy \tag{36}$$

for $0 \le i < n$, $n - 1 < j \le 2n - 1$. Here, $(c'_{l,k}|y)$ is some parameter dependent upon y.

Following a similar procedure, we can show that in Case 2

$$E[A'_{i,j}x^i y^j] = \iint \left(h'_n(x|y)\left(\sum_{l=0}^{i-n}(a'_{i,l+n}|y)h'_l(x|y)\right)\right.$$

$$\left. + \sum_{l=0}^{n-1}(a'_{i,l}|y)h'_l(x|y)\right)\left(\sum_{k=0}^{j} b'_{j,k}h_k(y)\right) f_{X,Y}(x,y)dxdy$$

$$= \iint \sum_{l=0}^{n-1}\sum_{k=0}^{j}(d'_{l,k}|y)h'_l(x|y)h_k(y)f_{X,Y}(x,y)dxdy \tag{37}$$

where $(d'_{l,k}|y)$ is some parameter dependent upon y. As for both cases, $E[A'_{i,j}x^i y^j]$ can be expressed using orthogonal and conditional orthogonal polynomials with the degree up to $n-1$, we can write

$$E[g_c(x,y)] = \sum_{j=0}^{2n-1}\sum_{i=0}^{2n-1-j} E[A'_{i,j}x^i y^j]$$

$$= \iint \sum_{l=0}^{n-1}\sum_{k=0}^{n-1}(e'_{l,k}|y)h'_l(x|y)h_k(y)f_{X,Y}(x,y)dxdy \tag{38}$$

where $(e'_{l,k}|y)$ is some parameter dependent upon y. Noticing that $\sum_{l=0}^{n-1}\sum_{k=0}^{n-1}(e'_{l,k}|y)h'_l(x|y)h_k(y)$ is a polynomial with the degree of each variable up to $n - 1$, we can write

$$E[g_c(x,y)] = \iint \sum_{i=0}^{n-1}\sum_{j=0}^{n-1} B'_{i,j}x^i y^j f_{X,Y}(x,y)dxdy = E[g_c^*(x,y)]. \tag{39}$$

It is clear that $g_c^*(x,y)$ and $g_c(x,y)$ have the same mean.

Finally, let us verify that both $g_c^*(x,y)$ and $g_c(x,y)$ pass through the same set of selected PCM points. Let us denote the roots of $h_n(y)$ as $y_1, y_2, ..., y_n$ and the roots of $h'_n(x|y_j)$ as x_{ij} for $1 \le i \le n$, $1 \le j \le n$. The n^2 pairs of inputs (x_{ij}, y_j) are the selected PCM points. By expressing $g_c(x,y)$ and $g_c^*(x,y)$ in orthogonal forms, and from observing (35), (36) and (37), it is clear to see that $g_c(x,y)$ and $g_c^*(x,y)$ are identical at each pair of the PCM points. The proof is now complete. ■

We now generalize to more than two variables, and sketch the key steps to prove Theorem 6.

Proof: First of all, note that there is only one variable in the original mapping with degree greater than $n - 1$, as the total degree of $A'_{k_1,k_2,...,k_m}\prod_{i=1}^{m} x_i^{k_i}$ is at most $2n-1$. Without loss of generality, let us denote this variable as x_t, where $1 \le t \le m$. As such, we only need to find a low-order mapping with the degree of this variable x_t reduced to $n - 1$.

To do that, we express x_t in terms of orthogonal polynomials defined upon the conditional pdf $f_{X_t|X_{t+1},X_{t+2},...,X_m}(x_t|x_{t+1}, x_{t+2}, ..., x_m)$. Let us denote these orthonormal polynomials as $h_i^t(x_t|x_{t+1}, ..., x_m)$, where i is the degree of this polynomial. It is then easy to see that any term in $g_c(x_1, x_2, ..., x_m)$ that includes x_t, denoted as $A'_{k_1,k_2,...,k_t,...,k_m}\prod_{i=1}^{m} x_i^{k_i}$, can be expressed as

$$A'_{k_1,k_2,...,k_t,...,k_m}\prod_{i=1}^{m} x_i^{k_i} = \left(\prod_{i=1}^{t-1} x_i^{k_i}\right)\left(h_n^t(x_t|x_{t+1}, ..., x_m)\right.$$

$$\left(\sum_{l=0}^{k_t-n}\left(a'^{t}_{k_t,l+n}|x_{t+1}, ..., x_m\right)\right.$$

$$\left. h_l^t(x_t|x_{t+1}, ..., x_m)\right)$$

$$+ \sum_{l=0}^{n-1}\left(a'^{t}_{k_t,l}|x_{t+1}, ..., x_m\right)$$

$$\left. h_l^t(x_t|x_{t+1}, ..., x_m)\right)\left(\prod_{i=t+1}^{m} x_i^{k_i}\right) \tag{40}$$

where $a'^{t} \in R$ are the corresponding coefficients. Again, the first subscript of a'^{t} stands for the total degree of x_t, and the second represents the degree of x_t in each term.

In order to show that the degree of x_t can be reduced to a value smaller than n, we only need to show that in $E[g_c(x_1, x_2, ..., x_m)]$, any term involved with $h_l^t(x_t|x_{t+1}, ..., x_m)$

and with total degree greater than n equals 0. It is sufficient to show that

$$\iint \dots \int \left(\prod_{i=1}^{t-1} x_i^{k_i}\right) h_n^t(x_t|x_{t+1}, \dots, x_m) \left(\sum_{l=0}^{k_t-n} \left(a'^t_{k_t,l+n}|x_{t+1}, \dots, x_m\right) h_l^t(x_t|x_{t+1}, \dots, x_m)\right) \left(\prod_{i=1}^{t} f_{X_i|X_{i+1},\dots,X_m}(x_i|x_{i+1}, \dots, x_m)\right) dx_1 dx_2 \dots dx_t = 0. \quad (41)$$

Recall the assumption that each conditional moment of x_i (i.e., $\int x_i^{k_i} f_{X_i|X_{i+1},X_{i+2},\dots,X_m}(x_i|x_{i+1}, x_{i+2}, \dots, x_m)dx_i$) is a polynomial of $x_{i+1}, x_{i+2}, \dots, x_m$ with total degree k_i, $1 \le i \le t-1$. We could then find that the conditional moment of $\prod_{i=1}^{t-1} x_i^{k_i}$ expressed as

$$\iint \dots \int \left(\prod_{i=1}^{t-1} x_i^{k_i}\right) \left(\prod_{i=1}^{t-1} f_{X_i|X_{i+1},\dots,X_m}(x_i|x_{i+1}, \dots, x_m)\right) dx_1 \dots dx_{t-1}$$

is a polynomial of $x_t, x_{t+1}, \dots, x_m$ with the total degree not exceeding $\sum_{i=1}^{t-1} k_i$. When multiplying this polynomial with s $\sum_{l=0}^{k_t-n} \left(a'^t_{k_t,l+n}|x_{t+1}, \dots, x_m\right) h_l^t(x_t|x_{t+1}, \dots, x_m)$, we find that the maximum degree of x_t is $\sum_{i=1}^{t-1} k_i + k_t - n = \sum_{i=1}^{t} k_i - n \le n-1$. Equation (41) is proved according to the orthogonality.

The rest of the proof on selecting PCM points and constructing $g_c^*(x_1, x_2, \dots, x_m)$ follows directly from the proof of Theorem 6, and is thus omitted here. ■

ACKNOWLEDGMENT

Approved for Public Release; Distribution Unlimited. 14-0746

Notice: This work subjects to Federal Aviation Administration Acquisition Management System Clause 3.5–13, Rights In Data-General, Alt. III and Alt. IV (Oct. 1996). The contents of this material reflect the views of the authors and The MITRE Corporation and do not necessarily reflect the views of the FAA or the DOT. Neither the Federal Aviation Administration nor the Department of Transportation makes any warranty or guarantee, or promise, expressed or implied, concerning the content or accuracy of these views.

REFERENCES

[1] A. Agogino and J. Rois, "Robustness of two air traffic scheduling approaches to departure uncertainty," in *Proc. IEEE/AIAA 30th Digital Avionics Syst. Conf.*, Oct. 2011, pp. 2C6-1–2C6-8.

[2] A. Agustin, A. Alonso-Ayuso, L. F. Escudero, and C. Pizarro, "On air traffic flow management with rerouting. Part II: Stochastic case," *Eur. J. Oper. Res.*, vol. 219, no. 1, pp. 167–177, May 2012.

[3] G. Andreatta, P. Dell'Olmo, and G. Lulli, "An aggregate stochastic programming model for air traffic flow management," *Eur. J. Oper. Res.*, vol. 215, no. 3, pp. 697–704, Dec. 2011.

[4] C. Asavathiratham, S. Roy, B. Lesieutre, and G. Verghese, "The influence model," *IEEE Control Syst. Mag.*, vol. 21, no. 6, pp. 52–64, Dec. 2001.

[5] K. Atkinson, *An Introduction to Numerical Analysis*. Hoboken, NJ, USA: Wiley, 1988.

[6] I. Babuska, F. Nobile, and R. Tempone, "A stochastic collocation method for elliptic partial differential equations with random input data," *SIAM J. Numer. Anal.*, vol. 45, no. 3, pp. 1005–1034, May 2007.

[7] E. W. Cheney and D. R. Kaincaid, *Numerical Mathematics and Computing*. Boston, MA, USA: Cengage Learning, 2007.

[8] G. Clare and A. Richards, "Disturbance feedback for handling uncertainty in air traffic flow management," in *Proc. Eur. Control Conf.*, Jul. 2013, pp. 3246–3251.

[9] G. Clare, A. Richards, J. Escartin, and D. Martinez, "Air traffic flow management under uncertainty: Interactions between network manager and airline operations centre," in *Proc. 2nd SESAR*, Nov. 2012, pp. 1–8.

[10] J.-P. B. Clarke and S. Solak, "Air traffic flow management in the presence of uncertainty," in *Proc. 8th USA/Europe Air Traffic Manag. Res. Develop. Sem.*, Jun. 2009, pp. 1–10.

[11] J. Foo, X. Wan, and G. E. Karniadakis, "The multi-element probabilistic collocation method (ME-PCM): Error analysis and applications," *J. Comput. Phys.*, vol. 46, no. 5, pp. 2309–2345, May 2008.

[12] J. R. Hockenberry and B. C. Lesieutre, "Evaluation of uncertainty in dynamic simulations of power system models: The probabilistic collocation method," *IEEE Trans. Power Syst.*, vol. 19, no. 3, pp. 1483–1491, Aug. 2004.

[13] S. Hosder, R. W. Walters, and M. Balch, "Efficient samping for non-intrusive polynomial chaos applications with multiple uncertain input variables," in *Proc. AIAA/ASME/ASCE/AHS/ASC*, Apr. 2007, pp. 1–16.

[14] S. Isukapalli, *Uncertainty Analysis of Transport-Transformation Models*. New Brunswick, NJ, USA: Rutgers, New Jersey, 1999.

[15] R. Jin, W. Chen, and A. Sudjianto, "An efficient algorithm for constructing optimal design of computer experiments," *J. Stat. Plan. Inference*, vol. 134, no. 1, pp. 268–287, Sep. 2005.

[16] Q. Ma, Y. Wan, and D. Sun, "Probabilistic collocation method in solving euclidean model for air traffic flow management," in *Proc. AIAA Infotech@aerosp. Conf.*, Jun. 2012, pp. 1–6.

[17] Q. Ma, Y. Wan, and D. Sun, "Stochastic analysis of air-traffic system and its corresponding application in parameters prediction," in *Proc. Amer. Control Conf.*, Jun. 2012, pp. 1689–1694.

[18] M. D. Mckay, R. J. Beckman, and W. J. Conover, "A comparison of three methods for selecting values of input variables in the analysis of output from a computer code," *Technometrics*, vol. 21, no. 2, pp. 239–245, 1979.

[19] F. Nobile, R. Tempone, and C. Webster, "A sparse grid collocation method for elliptic partial differential equations with random input data," *SIAM J. Numer. Anal.*, vol. 229, no. 5, pp. 1536–1557, Mar. 2010.

[20] A. Papoulis and S. U. Pillai, *Probability, Random Variables and Stochastic Processes*. New York, NY, USA: McGraw-Hill, 2002.

[21] J. W. Pepper, K. R. Mills, and L. A. Wojcik, "Predictability and uncertainty in air traffic flow management," in *Proc. 5th USA/Eur. Air Traffic Manag. R&D Sem.*, Jun. 2003, pp. 1–10.

[22] J. Rios and A. Morando, "The value of reduced uncertainty in air traffic flow management," in *Proc. AIAA Guid., Navigation, Control Conf.*, Aug. 2011, pp. 1–14.

[23] S. Roy, D. Ramamurthy, and B. C. Lesieutre, "Studies on the probabilistic collocation method and its application to power system analysis," in *Proc. 36th North America Power Symp.*, Aug. 2004, pp. 1–8.

[24] S. Roy, Y. Wan, C. Taylor, and C. Wanke, "A stochastic network model for uncertain spatiotemporal weather impact at the strategic time horizon," in *Proc. 10th ATIO*, 2010, pp. 1–19.

[25] M. A. Tatang, "Direct incorporation of uncertainty in chemical and environmental engineering systems," Ph.D. dissertation, Dept. Chemical Eng., Massachusetts Inst. Technol., Cambridge, MA, USA, 1995.

[26] C. Taylor, C. Wanke, Y. Wan, and S. Roy, "A framework for flow contingency management," in *Proc. AIAA Aviation Technol., Integr., Oper. Conf.*, Sep. 2011, pp. 1–23.

[27] C. Taylor, C. Wanke, Y. Wan, and S. Roy, "A decision support tool for flow contingency management," in *Proc. AIAA Guid. Navigation Control Conf.*, Aug. 2012, pp. 1–21.

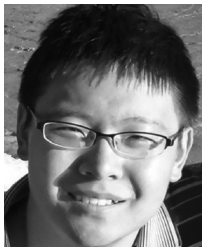
[28] Y. Wan and S. Roy, "A scalable methodology for evaluating and designing coordinated air traffic flow management strategies under uncertainty," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 4, pp. 644–656, Nov. 2008.

[29] Y. Wan and S. Roy, "Uncertainty evaluation through mapping identification in intensive dynamic simulations," *IEEE Trans. Syst., Man, Cybern. A, Syst. Humans*, vol. 40, no. 5, pp. 1094–1104, Sep. 2010.

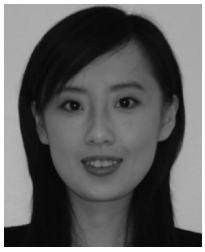
[30] Y. Wan, C. Taylor, S. Roy, C. Wanke, and Y. Zhou, "Dynamic queuing network model for flow contingency management," in *Proc. AIAA Guid., Navigation Control Conf. Exhibit*, 2011, pp. 1–22.

[31] C. Wanke, S. Mulgund, D. Greenbaum, and L. Song, "Modeling traffic prediction uncertainty for traffic management decision support," in *Proc. AIAA Guid., Navigation, Control Conf.*, Aug. 2004, pp. 1–13.

[32] M. Webster, M. A. Tang, and G. J. McRae, "Application of the probabilistic collocation method for an uncertainty analysis of a simple ocean model," Joint Program on the Science and Policy of Global Change, MIT, Cambridge, MA, USA, Jan. 1996.
[33] D. L. Wei, Z. S. Cui, and J. Chen, "Uncertainty quantification using polynomial chaos expansion with points of monomial cubature rules," *Comput. Structures*, vol. 86, nos. 23–24, pp. 2102–2108, Dec. 2008.
[34] M. M. Woolfson and G. J. Pert, *An Introduction to Computer Simulation*. New York, NY, USA: Oxford Univ. Press, 1999.
[35] F. Xiong, W. Chen, Y. Xiong, and S. Yang, "A new weighted stochastic response surface method for uncertainty propagation," in *Proc. AIAA/ISSMO Multidiscipl. Anal. Optimiz. Conf.*, Sep. 2010, pp. 1–13.
[36] M. Xue, S. Zobell, S. Roy, C. Taylor, Y. Wan, and C. Wanke, "Using stochastic, dynamic weather-impact models in strategic traffic flow management," in *Proc. 91st AMS Annu. Meet.*, Jan. 2011, pp. 1–23.
[37] Y. Zhou, D. Ramamurthy, Y. Wan, C. Taylor, and C. Wanke, "Multivariate probabilistic collocation method for effective uncertainty evaluation with application to air traffic management," in *Proc. Amer. Control Conf.*, Jun. 2013, pp. 6345–6350.
[38] Y. Zhou, Y. Wan, S. Roy, C. Taylor, and C. Wanke, "A stochastic modeling and analysis approach to strategic traffic flow management under weather uncertainty," in *Proc. AIAA Guid., Navigation Control Conf. Exhibit*, Aug. 2011, pp. 1–18.



Yi Zhou (S'11) was born in Nanjing, China, 1985. He received the B.S. degree from Jiang Su University, Zhenjiang, China, in 2007, and the M.S. degree from the University of North Texas, Denton, TX, USA, in 2011, where he is currently pursuing the Ph.D. degree. His current research interests include stochastic modeling, queuing theory, and air traffic flow management.



Yan Wan (S'08–M'09) received the B.S. degree from the Nanjing University of Aeronautics and Astronautics, Nanjing, China, in 2001, the M.S. degree from the University of Alabama, Tuscaloosa, AL, USA, in 2004, and the Ph.D. degree from Washington State University, Pullman, WA, USA, in 2009.

She was a Post-doctoral Scholar with the Control Systems Program, University of California at Santa Barbara, Santa Barbara, CA, USA. She is currently an Assistant Professor with the Department of Electrical Engineering, University of North Texas, Denton, TX, USA. Her current research interest include decision-making tasks in large-scale networks, with applications to air traffic management, airborne networks, sensor networking, and biological systems.

Dr. Wan was the recipient of the prestigious William E. Jackson Award (excellence in aviation electronics and communication) in 2009, presented by Radio Technical Commission for Aeronautics.

Sandip Roy (M'04) is currently an Associate Professor with the School of Electrical Engineering and Computer Science, Washington State University, Pullman, WA, USA, and an Affiliate Associate Professor with the School of Global Animal Health. His current research interests include developing tools for estimation and control of complex dynamical networks, and applying these tools to facilitate automated decision-making or management in infrastructure networks, particularly the air transportation system.

Christine Taylor received the B.S. degree in mechanical and aerospace engineering from Cornell University, Ithaca, NY, USA, and the M.S. and Ph.D. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA. Her dissertation research focused on integrating operations research methodologies into the design of aerospace systems.

She is currently a Lead Simulation and Modeling Engineer with the MITRE Corporations Center for Advanced Aviation Systems Development, McLean, VA, USA, specializing in decision support system development for traffic flow management applications.

Craig Wanke received the S.B, S.M, and Ph.D. degrees in aeronautical engineering from the Massachusetts Institute of Technology, Cambridge, MA, USA.

He is currently a Senior Principal Engineer with the MITRE Corporations Center for Advanced Aviation System Development, McLean, VA, USA. He has been with MITRE for 19 years, engaged in a wide range of decision support systems for pilots, air traffic controllers, and traffic managers.

Dinesh Ramamurthy received the B.S. degree from the University of Madras, Chennai, India, in 2003, and the M.S. degree from Washington State University, Pullman, WA, USA, in 2005.

Since 2006, he has been a Lead Software Engineer with the Enterprise Holdings, Inc., St. Louis, MO, USA.

Junfei Xie (S'13) received the B.S. degree from the University of Electronic Science and Technology of China, Chengdu, China, in 2012, and the M.S. degree from the University of North Texas, Denton, TX, USA, in 2013, where she is currently pursuing the Ph.D. degree with the Department of Computer Science and Engineering.

Her current research interests include airborne networks, complex information systems, and air traffic flow management.