Stochastic Ground-Delay-Program Planning in a Metroplex

Jun Chen* and Dengfeng Sun†
Purdue University, West Lafayette, Indiana 47906

DOI: 10.2514/1.G002964

This paper introduces a novel stochastic approach for the ground-delay-program planning under uncertainty, using chance-constrained optimization. The major advantage of the chance-constrained model is the ability to provide robust solutions with user-defined service level. The approach is compared with the Ball et al. ("A Stochastic Integer Program with Dual Network Structure and Its Application to the Ground-Holding Problem," Operations Research, Vol. 51, No. 1, Feb. 2003, pp. 167–171) model for selecting planned airport acceptance rates for airports in a metroplex, which is an interdependent system in close geographic proximity. The approaches were evaluated using real flight schedules and landing-capacity data from the New York City metroplex airports. Although the Ball et al. model was found to be more efficient, the chance-constrained model shows the ability to provide a quantized way to balance the solution's robustness and potential cost by choosing a proper service level. Moreover, the parallel-computing framework was demonstrated to be helpful in improving the computing efficiency, which suggests that deploying more computing resources would help solve a large-scale planning problem under uncertainty in the same framework.

I. Introduction

THE goal of air-traffic-management system is to balance traffic demand with system capacity through a variety of traffic management initiatives (TMIs). The ground-delay program (GDP) is one of the most effective strategic TMIs used to alleviate congestion costs, and it ensures safe and efficient air traffic [1]. In a GDP, flights are held on the ground at their origin airports when there is an expected reduction of landing capacity at the destination airport. The landing capacity is also referred to as airport acceptance rate (AAR), which describes the number of arrivals an airport is capable of accepting per hour. The assigned ground delay helps absorb airborne delay, such that the traffic supply–demand balance is maintained with cheaper and safer delay cost.

With rapid growth of air traffic, the airports in a metropolitan area cannot be considered as separated entities, but rather as an interdependent system, known as a metroplex [2]. A metroplex phenomenon is an interaction between two or more airports in close geographic proximity [3]. Adverse weather usually affects multiple airports in a metroplex simultaneously, such that the joint AARs of a metroplex are reduced, because adverse weather, such as fog, snow, wind, and reduced visibility, may require greater spacing between flights [4,5]. The imperfect weather forecast brings uncertainty into the GDP planning. Decisions made under uncertainty can cause airborne delays for multiple airports simultaneously, which greatly lower the efficiency in those busy metroplex airspace. This highlights the importance of addressing weather uncertainty in the GDP planning in a metroplex to mitigate congestions.

There has been considerable research on how to allocate ground delays efficiently and equitably. The first effort to tackle GDP dates back to 1987, when Odoni was among the first to describe the problem [6]. Following this, several deterministic models for the single-airport ground-holding problem (SAGHP) were developed [7,8]. The first stochastic models for SAGHP were introduced by Richetta and Odoni, which consider the uncertainty in AAR using stochastic integer programming under static [9] and partially dynamic [10] settings. Later, the Federal Aviation Administration implemented a new GDP paradigm, known as collaborative decision making (CDM), in which the airlines have more autonomy about their schedules. Under the CDM paradigm, the arrival slots are first allocated to individual flights based on the planned AARs (PAARs) [11]. Then, the airlines are allowed to exchange the arrival slots among themselves, which is the key feature of CDM [12].

Many models were proposed to assist the implementation of GDP under CDM. Ball et al. formulated an aggregative static stochastic model, which solves for optimal PAARs during different time intervals [13]. However, once the ground-holding strategies were decided "once and for all" at the beginning of planning time horizon, they could not be revised even for flights that have yet to depart. Mukherjee and Hansen improved this dynamic model by allowing for ground-holding revisions contingent on updated scenario realizations [14]. In all of the aforementioned models, the uncertainty in AAR was represented through a finite number of scenarios arranged in a probabilistic decision tree. As time progressed, branches of the tree were realized, resulting in better information about future capacities [15]. Unfortunately, the probabilistic-scenario-tree approach suffers significantly from the practical difficulty of not knowing the exact scenarios. Furthermore, it generally becomes intractable quickly as the number of scenarios increases, thereby posing substantial computational challenges.

This paper proposes an alternative method to incorporate probabilistic information instead of a scenario tree, called chance constraints. The idea is to constrain the chance of a constraint violation, given probabilistic information about future state disturbances. The major advantage of a chance-constrained model is the ability to provide robust solutions with a user-defined service level, in which the service level represents the chance of the constraints not being violated. The service level can be defined by the air-traffic authority or airlines. First, a chance-constrained model is developed based on the Ball et al. [13] model to provide a robust optimal PAAR with a required service level. To efficiently solve the chance-constrained optimization problem, a convex-approximation-based approach is proposed based on the Bernstein polynomial. To evaluate the chance-constrained model, both the Ball et al. model and the chance-constrained model are applied to a metroplex ground-delay problem (MGDP). The evaluation used real flight schedules from the New York City (NYC) metroplex airports: John F. Kennedy International (JFK), Newark Liberty International (EWR), and LaGuardia (LGA) Airports.

The rest of this paper is organized as follows. Section II introduces the chance-constrained model. Section III introduces a polynomial-approximation-based approach to overcome the limitation of the sampling method. Section IV sets up the distribution and flight plan for the experiment. Section V presents the experiment results. Section VI concludes this paper.

Received 10 April 2017; revision received 5 July 2017; accepted for publication 7 July 2017; published online 25 August 2017.

*Ph.D. Candidate, School of Aeronautics and Astronautics; chen1241@purdue.edu.

†Associate Professor, School of Aeronautics and Astronautics; dsun@purdue.edu. Senior Member AIAA.

II. Chance-Constrained Modeling

The proposed chance-constrained model is derived based on a previous static model for the SAGHP [13]. First, the static model is modified to schedule PAARs for all airports in a metroplex simultaneously. Then, the same problem is solved by the newly proposed chance-constrained model for comparison.

A. Static Model

The static model was introduced by Ball et al. in 2003 [13], which is a static approach to choosing PAARs under CDM procedures. It is static because decisions are made based only on the current state and do not take into account updated information [16]. We choose to modify the Ball et al. model to consider multi-airports in a metropolitan area (e.g., NYC metroplex). The formulation is

$$\min \sum_{t=1}^{T}\sum_{i=1}^{M}\left(\boldsymbol{c}_g Y_t^i + \sum_{q=1}^{Q} p_q \boldsymbol{c}_a Z_{t,q}^i\right)$$

subject to

$$\begin{aligned}
&X_t^i + Y_t^i - Y_{t-1}^i = S_t^i \quad t=1,\ldots,T+1, \quad i=1,\ldots M\\
&Z_{t-1,q}^i + X_t^i - Z_{t,q}^i \le D_{t,q}^i \quad t=1,\ldots,T+1,\\
&\quad i=1,\ldots,M, \quad q=1,\ldots,Q\\
&Z_{0,q}^i = Z_{T+1,q}^i = 0\\
&Y_0^i = Y_{T+1}^i = 0\\
&X_t^i, Y_t^i, Z_{t,q}^i \ge 0
\end{aligned} \tag{1}$$

The difference is that multiple SAGHPs are combined together by sampling the landing capacity of each airport $D_{t,q}^i$ simultaneously from a joint distribution. To present the joint landing-capacity distribution, the Ball et al. [13] model chooses to sample a finite set of landing-capacity scenarios with associated probabilities p_q, in which each scenario q represents one possible evolution of landing capacity over time. The parameter S_t^i is the number of scheduled arrival flight for interval t, airport i. The ground and air delays are represented by Y_t^i and $Z_{t,q}^i$, respectively. All flights are enforced to arrive within the time horizon by the first constraints in Eq. (1), because all the scheduled flights are absorbed by either PAARs or ground holdings, and the ground holdings are emptied by the fourth constraint. The second constraints ensure that the actual number of arrivals should not exceed the landing capacity, because the extra flights will be held in the air. Solving model described by Eq. (1) will provide the optimal PAARs X_t^i for each airport.

B. Chance-Constrained Model

The chance-constrained model aims to incorporate the constantly changing landing capacities, which are caused by adverse weather conditions, into the MGDP. The current models are rather deterministic or based on predefined scenarios (like the Ball et al. [13] model). This paper proposes to impose a probabilistic constraint on landing capacities, as follows:

$$\mathbb{P}(Z_{t-1}^i + X_t^i - Z_t^i \le \xi_t^i, \quad i=1,\ldots,M) \ge \alpha \quad t=1,\ldots,T \tag{2}$$

in which $\mathbb{P}(\cdot)$ is the probability measure for the stochastic landing capacities, meaning that the landing capacity will only raise a feasibility issue with the probability of $\alpha \in (0,1)$, in which α is also called service level in this paper. The random components ξ_t^i are random parameters that represent the correlated, stochastic landing capacities, and only correlated random capacities are meaningful for the MGDP planning because adverse weather conditions will usually affect multiple airports of the metroplex simultaneously. Thus, the MGDP planning under the stochastic landing capacities can be written as

$$\min \sum_{t=1}^{T}\sum_{i=1}^{M}(\boldsymbol{c}_g Y_t^i + \boldsymbol{c}_a Z_t^i)$$

subject to

$$\begin{aligned}
&X_t^i + Y_t^i - Y_{t-1}^i = S_t^i \quad t=1,\ldots,T+1, \quad i=1,\ldots,M\\
&\mathbb{P}(Z_{t-1}^i + X_t^i - Z_t^i \le \xi_t^i, \quad i=1,\ldots,M) \ge \alpha \quad t=1,\ldots,T\\
&Z_0^i = Z_{T+1}^i = 0\\
&Y_0^i = Y_{T+1}^i = 0\\
&X_t^i, Y_t^i, Z_t^i \ge 0
\end{aligned} \tag{3}$$

The difference from the deterministic model is that the capacity constraints are replaced with the probabilistic capacity constraint (2). This problem is referred to as chance-constrained MGDP optimization. The chance-constrained model directly uses the joint landing-capacity distribution rather than generating a predefined scenario set from the distribution (like the Ball et al. [13] model).

III. Convex-Approximation Approach

The chance programming indicates that some of the constraints may be violated at a well-controlled, very low chance. In general, the chance-programming problem is not easy to solve [17]. The traditional solution approach to chance programming is the sample average approximation (SAA). However, the SAA approach becomes intractable quickly due to the exponential growth of state space with the number of sampled scenarios. Moreover, the SAA approach will only yield a feasible solution rather than an optimal solution. Therefore, this section will propose a convex-approximation method to efficiently solve the chance-constrained model, which could overcome the computational limitation of the SAA method when solving large-scale problems.

A. Problem Definition

The chance constraint would greatly complicate the computational perspective of the problem because of the loss of convexity, in both its feasible set and the constraint itself. Even though it is extremely difficult to solve a chance-constrained optimization for a global optimal solution, there are exceptions. In [18], the author showed that, based on the log concavity of the distribution [19], the chance-constrained model in Sec. II would have a convex feasible set. The constraint would be equivalently transformed into a convex program, which would be efficiently solved, as long as the function and its gradient (or subgradient) are available. The detailed definitions and theorems can be found in the Appendix.

Suppose the landing-capacity distribution is log concave with a probability distribution function $F_\xi(\boldsymbol{x})$. (The log-concave assumption will be justified in Sec. IV.) Then, we have $1 - F_\xi(R_t\boldsymbol{x}) \ge \alpha$, and by taking the log of both sides, the chance-constrained MGDP model can be written in a concise form:

$$\min \quad \boldsymbol{c}'\boldsymbol{x}$$

subject to

$$g_t(\boldsymbol{x}) = \log[F_\xi(R_t\boldsymbol{x})] - \log(1-\alpha) \le 0 \quad t=1,\ldots,T \tag{4}$$

$$A\boldsymbol{x} = \boldsymbol{b}$$

$$\boldsymbol{x} \ge 0$$

in which $\boldsymbol{c}$ and $\boldsymbol{x}$ represent the vector of the weight coefficients and decision variable, respectively; A, $\boldsymbol{b}$, and R_t are the coefficients corresponding to the original linear constraints.

With the exception of chance constraints, $g_t(\boldsymbol{x})$, the model described by Eq. (4) is a linear program. Although constraint $g_t(\boldsymbol{x})$ is

Table 1 Error bounds as the degree k of $p_k(y)$ increases

Degree k	Error bound
4	1
5	0.67
6	0.48
7	0.36
8	0.28
9	0.22
10	0.18
12	0.13
20	0.048
40	0.012
50	0.008

nonlinear, it is convex by Theorem 1 in the Appendix [20], which makes the model described by Eq. (4) a convex program with respect to $\boldsymbol{x}$. For any feasible point $\boldsymbol{x}_0$ of the model described by Eq. (4), as long as we have the function value at $\boldsymbol{x}_0$ [i.e., $g_t(\boldsymbol{x}_0)$] and its gradient $\nabla g_t(\boldsymbol{x}_0)$ [subgradient if $g_t(\boldsymbol{x})$ is nondifferentiable] at $\boldsymbol{x}_0$, then a first-order gradient algorithm can be adopted to obtain the optimal solution [21].

B. Polynomial Approximation

The key to solving the model described by Eq. (4) is to effectively evaluate the gradient (or subgradient) of $g_t(\boldsymbol{x})$, because the gradient could lead to the deepest feasible search direction. This subsection will build a polynomial-based approximation of $g_t(\boldsymbol{x})$ and use the gradient of the polynomial to approximate its original. Such an approximation has two advantages: first, thanks to the shape-preserving property of the Bernstein polynomial (see the definition in Definition 3 of the Appendix), we would effectively control the approximation errors for both the function values and their gradient at the same time. Second, we show that, under a large-enough sample size, the obtained optimal solution will converge to the true optimal solution.

Suppose a feasible $\boldsymbol{x} \in \mathbb{R}^n$, such that $A\boldsymbol{x} = \boldsymbol{b}$, $\boldsymbol{x} \geq 0$, $\boldsymbol{x} = [\boldsymbol{x}^1; \ldots; \boldsymbol{x}^n]$, in which $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^n \in \mathbb{R}^1$. We would impose an upper bound and a lower bound on each component of $\boldsymbol{x}$, as follows:

$$\ell^i \leq \boldsymbol{x}^i \leq u^i, \qquad i = 1, \ldots, n \tag{5}$$

We are interested in

$$\nabla g_t(\boldsymbol{x}) := \left[\frac{\partial g_t(\boldsymbol{x})}{\partial \boldsymbol{x}^1}; \ldots; \frac{\partial g_t(\boldsymbol{x})}{\partial \boldsymbol{x}^n}\right] \tag{6}$$

and each component $\partial g_t(\boldsymbol{x})/\partial \boldsymbol{x}^i : \mathbb{R} \to \mathbb{R}$ is a univariate function with respect to $\boldsymbol{x}^i \in [\ell^i, u^i]$.

Let us define the ith marginal function of $g_t(\boldsymbol{x})$ as $g_t^i(\boldsymbol{x}^i)$, which is the univariate function with respect to $\boldsymbol{x}^i \in [\ell^i, u^i]$; $g_t^i(\boldsymbol{x}^i)$ is essentially the function $g_t(\boldsymbol{x})$ with $\boldsymbol{x}^1, \ldots, \boldsymbol{x}^{i-1}, \boldsymbol{x}^{i+1}, \ldots, \boldsymbol{x}^n$ as a constant value. In other words, the univariate function $g_t^i(\boldsymbol{x}^i)$ is the orthogonal projection of $g_t(\boldsymbol{x})$ onto x^i. Since $g_t(\boldsymbol{x})$ is convex, all of its marginal functions $g_t^i(\boldsymbol{x}^i)$, $i = 1, \ldots, n$ are convex with respect to $\boldsymbol{x}^i$.

Our approach is to approximate all of the marginal functions of $g_t^i(\boldsymbol{x}^i)$ with a convex, differentiable polynomial of degree k, $p_k(\boldsymbol{x}^i)$ at a fixed x. Then, we estimate $\partial g_t(\boldsymbol{x})/\partial \boldsymbol{x}^i$ by $p_k'(\boldsymbol{x}^i)$, such that the problem of approximating $g_t(\boldsymbol{x})$ is decomposed into n-independent univariate-approximation problems.

In this paper, the Bernstein polynomial is adopted to construct the approximation $p_k(\boldsymbol{x}^i)$. For the sake of simplifying the notation, we use $\phi(y)$ to represent one univariate function $g_t^i(\boldsymbol{x}^i)$. Without a loss of generality, we assume $y \in [0, 1]$, because we can make a linear change of variable, if necessary, to transform any finite interval $[\ell^i, u^i]$ onto [0,1].

Theorem 2 (Appendix) shows that the Bernstein polynomial can approximate any continuous univariate function on a closed interval [22]. However, for a convex function $\phi(y)$, its Bernstein-polynomial approximation may not be convex because the sampled data may not actually be convex due to an experimental numerical error. Besides, for the simple Bernstein polynomial, the degree of the polynomial needs to be doubled to halve the error [23]. Thus, we discard the idea of directly approximating $g_t^i(\boldsymbol{x}^i)$ by the Bernstein polynomial. Instead, Theorem 3 shows that, for any convex function $\phi(y)$, $y \in [0, 1]$, we can always approximate both $\phi(y)$ and its derivative $\phi'(y)$ within ϵ uniformly with the polynomial:

$$p_k(y) = \sum_{j=0}^{k} \boldsymbol{c}_j^* \psi_j(y) \tag{7}$$

of degree k, regardless of the differentiability of $\phi(y)$. As long as all of the coefficients in Eq. (7) are nonnegative, $\boldsymbol{c}_j^* \geq 0$, $j = 0, 1, \ldots, k$, $p_k(y)$ will be convex. We also need $k+1$ nonnegative coefficients $\boldsymbol{c}_0^*, \ldots, \boldsymbol{c}_k^*$ to construct $p_k(y)$. To obtain these coefficients, we need a set of points with coordinates $[y_i, \phi(y_i)]$, $i = 1, 2, \ldots, k+1$. We solve the following problem:

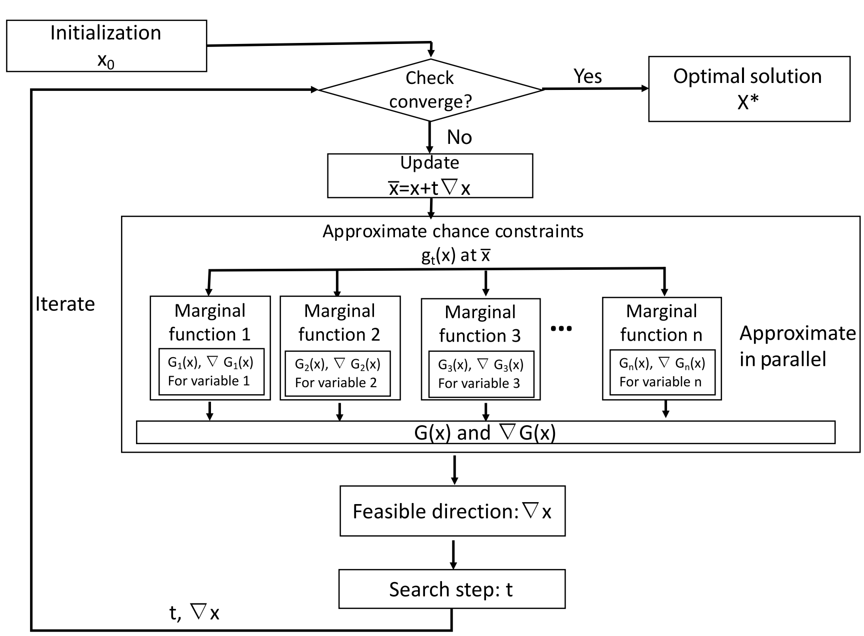
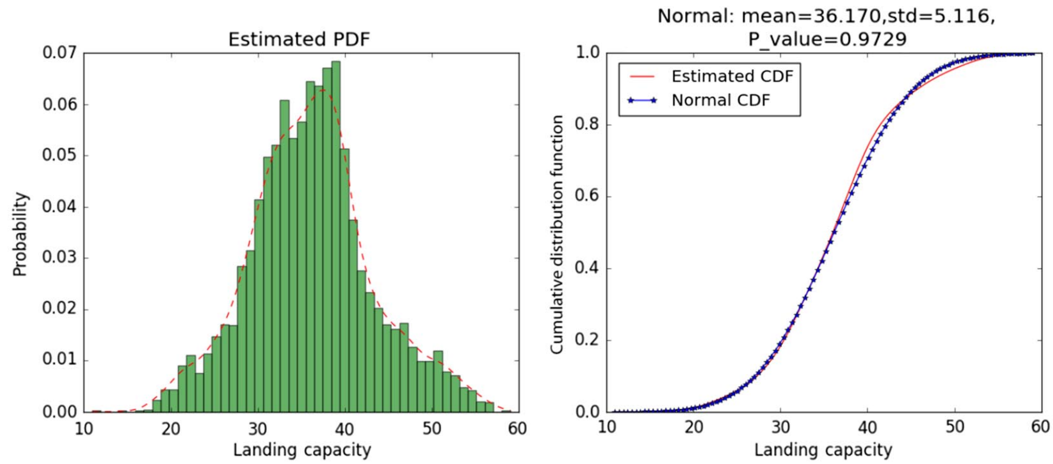
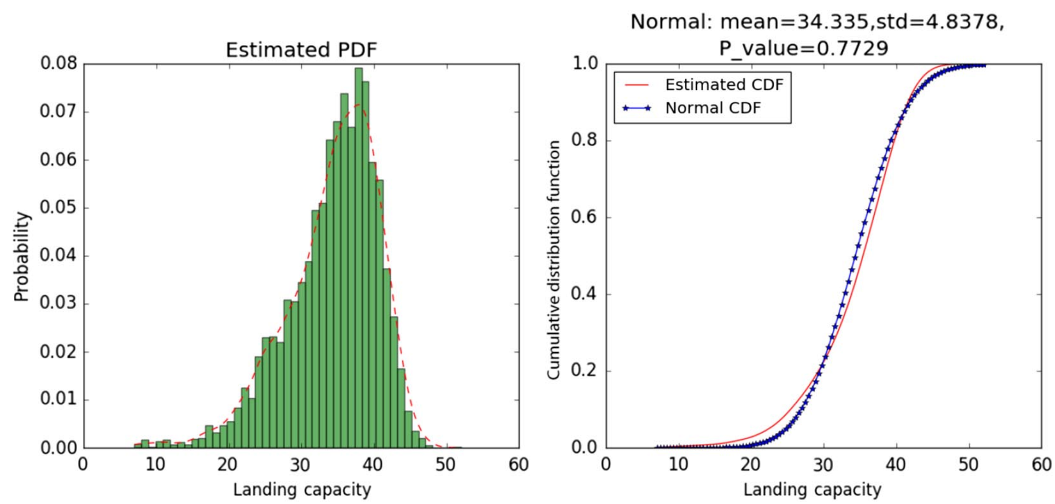


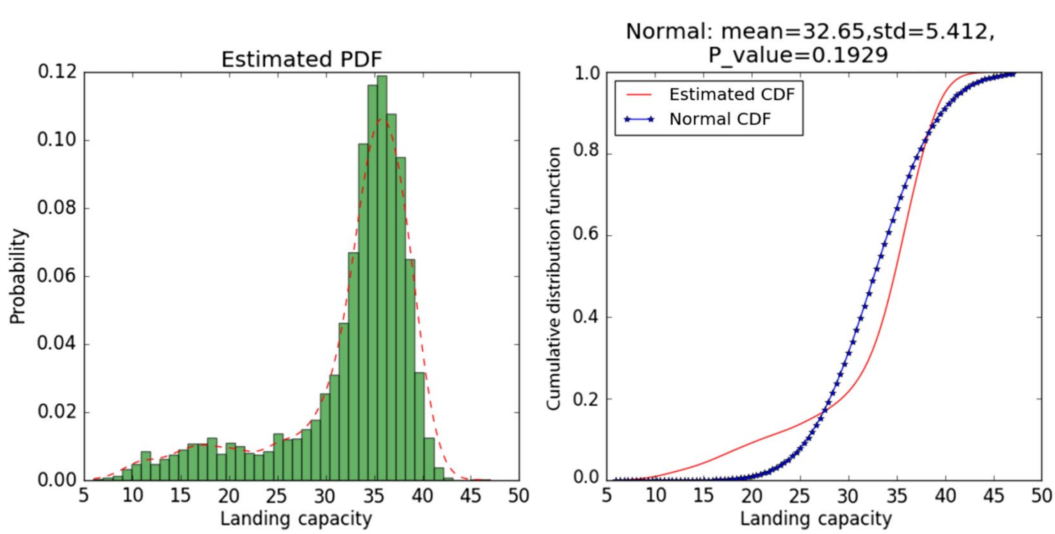
Fig. 1 Algorithm flowchart for the chance-constrained problem. (Color in online.)



a) Landing distribution at JFK



b) Landing distribution at EWR



c) Landing distribution at LGA

Fig. 2 Landing distributions for NYC metroplex. (Color in online.)

$$\min\left\{\max_{i=1,\ldots,k+1}|\phi(y_i)-\sum_{j=0}^{k}\boldsymbol{c}_j^*\psi_j(y_i)|,\boldsymbol{c}_j^*\geq 0\right\} \tag{8}$$

The $p_k(y)$ with coefficients $\boldsymbol{c}_0^*,\ldots,\boldsymbol{c}_k^*$ is called the best approximation of degree k.

We now need to determine the proper choice of the degree k. Theorem 4 shows that, if we approximate an r−differentiable function by $p_k(y)$, the error will be quickly reduced by increasing the order of the polynomial [24]. For example, when the degree increases from k to $k+1$, the rate of the error-bound reduction will be

$$\frac{(\pi/2)^r(|\phi^{(r)}(\omega)|/[(k-r+3)\ldots(k+1)(k+2)])}{(\pi/2)^r(|\phi^{(r)}(\omega)|/[(k-r+2)\ldots k(k+1)])}=\frac{k-r+2}{k+2}<1 \tag{9}$$

Table 2 Covariance matrix for the joint distribution of NYC metroplex landing capacities

	JFK	EWR	LGA
JFK	4.604	0.805	0.2592
EWR	0.805	3.652	1.633
LGA	0.2592	1.633	4.407

In the previous discussion (Proof 1 of Theorem 3), we assume that $\phi(y)$ is twice differentiable (i.e., $r = 2$). If we evaluate the error bound when $k = 4$, and $(\pi^2/2)[|\phi^{(2)}(\omega)|/4 \times 5]$ is the scale of the error base valued at 1, we present the results in the following Table 1. From Table 1, increasing k from 4 to 5 will reduce the error bound to 0.67 of its original value, while increasing k from 4 to 20 will reduce the error bound to 0.048 of its original value. When we increase k from 4 to 50, the new error bound will be reduced to 0.008 of its original value. Given the result of Theorem 3 and the good performance of the "best approximation" [i.e., $p_k(y)$], the error bound when $k = 4$ would already be a well-bounded value. Thus, when we use $k = 50$, the new error bound will be reduced to a fraction of 0.008, which should serve us adequately well.

At last, we determine the $k + 1$ coordinates {i.e., $[y_i, \phi(y_i)]$, $i = 1, \ldots, k + 1$} to solve for the coefficients $\mathbf{c}_0^*, \ldots, \mathbf{c}_k^*$ in Eq. (7). We choose the Chebyshev nodes [25] on [0,1], as follows:

$$y_i := \frac{1}{2} - \frac{1}{2}\cos\left(\frac{2i-1}{2k+2}\pi\right), \qquad i = 1, \ldots, k+1 \tag{10}$$

The technical detail regarding the minimization of the approximation error by adopting Chebyshev nodes is in Proposition 1 (Appendix).

The procedure to estimate $\phi'(y)$ by the polynomial $p_k(y)$ is summarized in the following steps:

1) Step 0: the polynomial $p_k(y)$ is defined by the formulation in Theorem 3.
2) Step 1: determine the overall error bound $\epsilon > 0$.
3) Step 2: choose the degree k based on Theorem 4.
4) Step 3: calculate $k + 1$ Chebyshev nodes y_i and coordinates $[y_i, \phi(y_i)]$, $i = 1, \ldots, k + 1$ (Proposition 1).
5) Step 4: solve the model (8) for the coefficients $\mathbf{c}_0^*, \ldots, \mathbf{c}_k^*$, and construct $p_k(y)$ in Eq. (7).
6) Step 5: use $p_k'(y)$ as an approximation of $\phi'(y)$.

C. Solving Framework

Based on the approximation approach in Sec. III.B, the function values and gradients for the chance constraint $g_t(\mathbf{x})$ can be estimated for any given point $\bar{\mathbf{x}}$. Therefore, a first-order algorithm can be adopted to solve the whole problem. In this paper, the feasible direction method [21] is adopted as the primary algorithm.

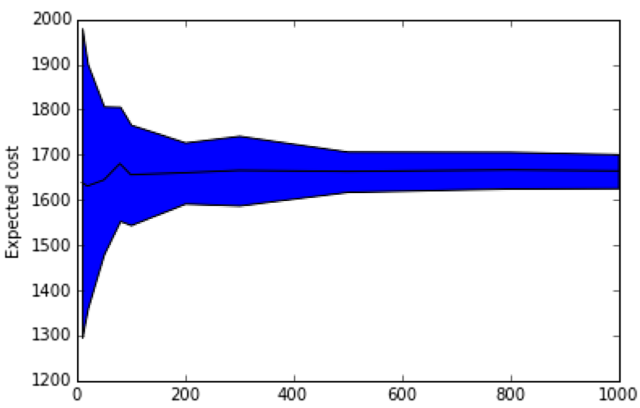


Fig. 3 Objective convergence along the number of sample scenarios. (Color in online.)

The flowchart of the solving algorithm for the chance-constrained problem is shown in Fig. 1. Recall that the construction of the polynomial approximation for each individual marginal function is independent. Therefore, at each step, the chance constraint can be approximated in parallel at the given point $\bar{\mathbf{x}}$. Then, the results are gathered to provide the first-order information, which is used to search for the feasible direction and optimal search step. Note that the algorithm needs to call the approximation process during every

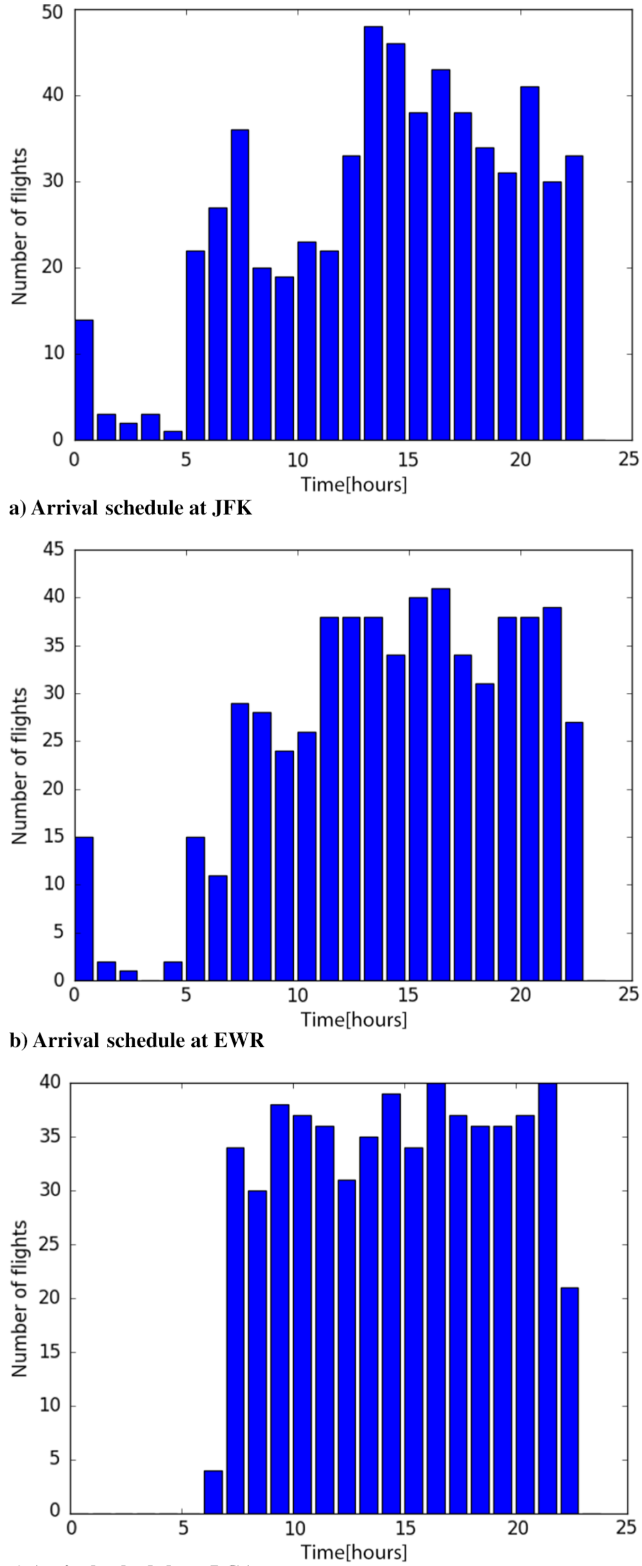


Fig. 4 Arrival schedules for NYC metroplex airports on 20 May 2016. (Color in online.)

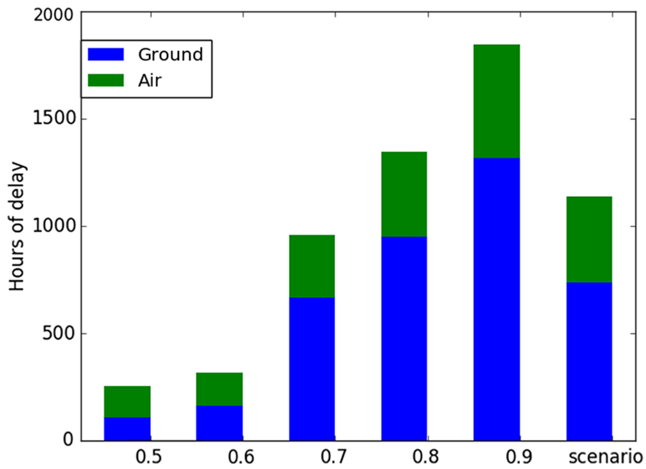


Fig. 5 Ground and air delays under various service levels. (Color in online.)

iteration until the final convergence. Therefore, the parallel-computing framework can greatly improve the computational efficiency by the fact that the approximation process has the most expensive computing cost of the whole process.

IV. Experimental Setup

The assumption that the landing-capacity distribution follows a log-concave distribution, but lacks closed-form distributional information, would be justified with two phases. First, in reality of air-traffic management, the historical data from the landing-capacity distribution are in the form of empirical distribution [26,27]. By using proper distribution-estimation methods (such as the kernel density estimation), the empirical distribution will be presented as a continuous distribution without the distributional information. By the Glivenko–Cantelli theorem (see [28]), the empirical distribution function estimates the cumulative distribution function (CDF) and converges with a probability of 1. That is, the empirical distribution can be presented as an underlying continuous distribution. Second, once the empirical distribution is in the format of a continuous distribution (but still lacks distribution information), this paper would further assume log concavity because so many commonly used distributions are, indeed, log concave. For example, the normal distribution, uniform distribution, gamma distribution (with a shape parameter greater than 1), beta distribution (with all parameters greater than 1), Weibull distribution, Laplace distribution, logistic distribution, exponential distribution, and extreme value distribution are log concave. There are only a few commonly used distributions that are not log concave, such as the log-normal distribution, *t* distribution, and chi-square distribution, which are often used to describe the distributions of various statics rather than random variables raised from real problems.

To evaluate the chance-constrained model for MGDP, a joint distribution of landing capacities is required. Moreover, to further support the log-concave distribution assumption, we would like to justify it with real empirical data. We analyzed the observed landing capacities of the NYC metroplex (JFK, EWR, and LGA) for 368 days from May 2015 through October 2015, and from May 2016 through October 2016. The landing capacities were computed based on the algorithm in [16,29] and the "arrivals for metric computations" data from the aviation system performance metrics database (ASPM) [30]. The individual distribution of landing capacities for each airport is shown in Fig. 2. The distributions are estimated by the kernel density estimation (KDE), which is a nonparametric way to estimate the probability density function (PDF) without assuming any distributional priori property [31]. In Fig. 2, the empirical data are shown in bar chart, and the KDE-based PDF is shown the dashed line. The individual distributions are demonstrated to be very close to the normal distribution. Figure 2 shows the CDFs of a normal distribution (dot line) and the CDF for the estimated distribution for each airport (solid line). A Kolmogorov–Smirnov test of normality was also applied, and the result, referred to as a *p* value, is indicated in the title of each subfigure. The *p* value generally indicates that the estimated distribution can be reasonably approximated as normal at the significant level 0.05. Therefore, the empirical landing capacities of all airports in the NYC metroplex are fitted by KDE to be a joint multinormal distribution, which is log concave. Please note that the convex-approximation method can work with any log-concave

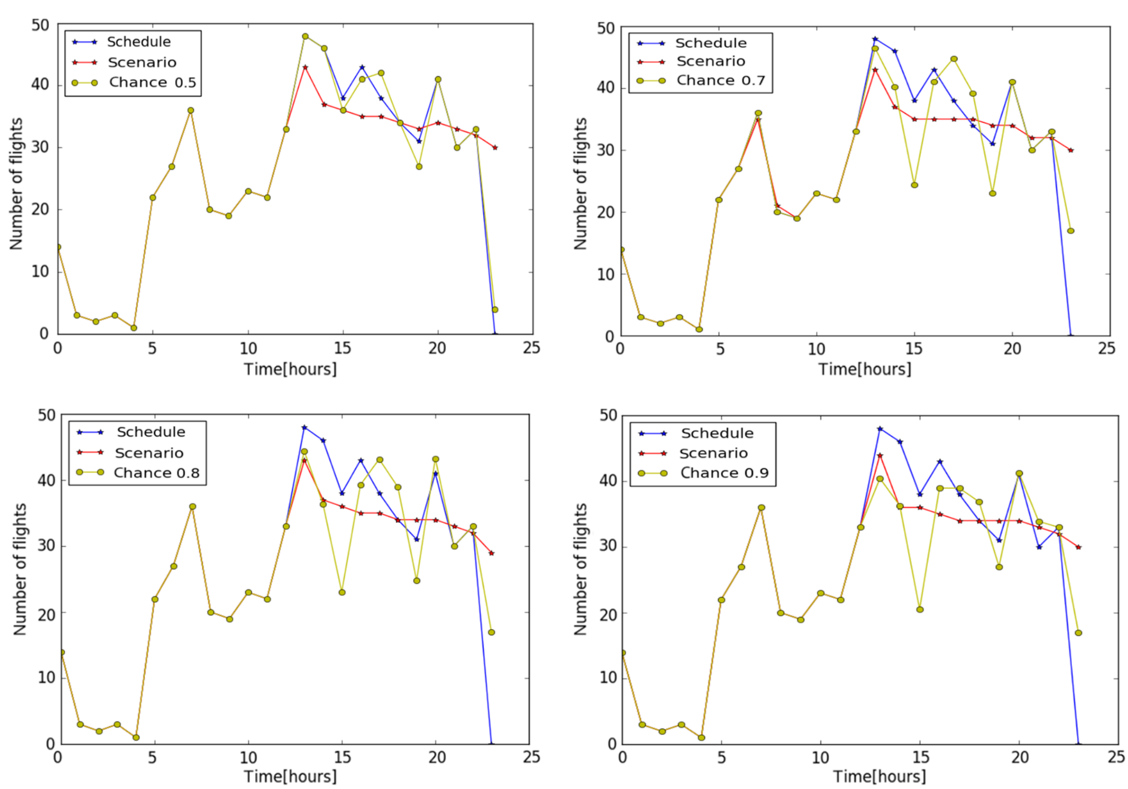


Fig. 6 Optimal PAARs of JFK under various service levels. (Color in online.)

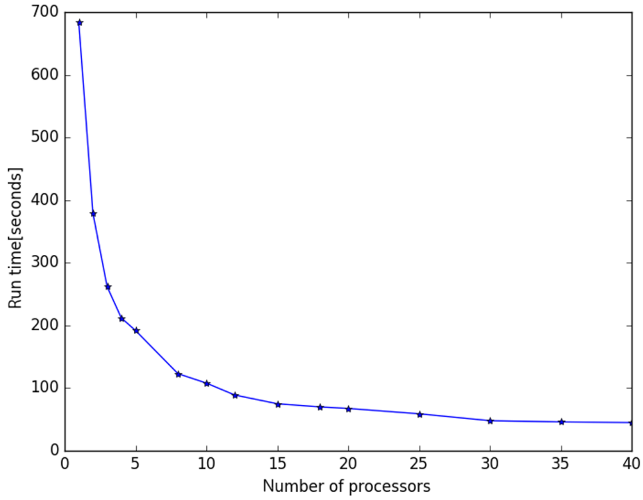


Fig. 7 Computation time decreasing as a function of the number of processors. (Color in online.)

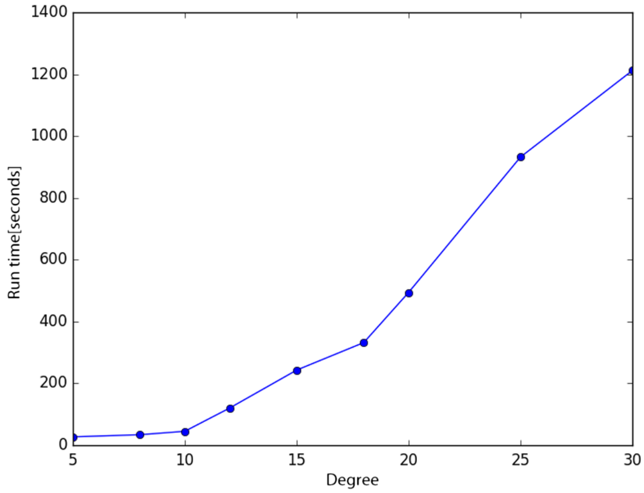


Fig. 8 Computation time associated with different polynomial degrees. (Color in online.)

distribution; the normality is not required for general cases. The covariance matrix is shown in Table 2, which confirms the correlation between the airports in the same metroplex.

Scenarios for the Ball et al. [13] model were generated by successively sampling landing capacities from the joint distributions. Different scenario samples will result in different minimum expected costs. Figure 3 shows the minimum expected cost along various sample sizes. The blue area is the 95% confidence interval for each sample size. We chose to sample 500 scenarios because the 95% confidence interval of the expected cost almost keeps the same with more than 500 scenarios. Each of the 500 scenarios is assigned with the same $1/500$ probability to calculate the expected cost of the static model in Sec. II.

The two approaches of MGDP were evaluated using the same flight schedules for the NYC metroplex on 20 May 2016. The number of arrivals per hour for this day for all airports, taken from the ASPM database, is shown in Fig. 4. We discretized the flight schedule and modified the last time interval with infinite capacity to ensure all flights could land within the time horizon. Solutions to the Ball et al. [13] model were found using the Gurobi mathematical programming solver [32]. The chance programming was implemented in the Python programming language.

V. Experimental Results

One of the key advantages of the chance-constrained model is the ability to provide robust solutions with a user-defined service level. Figure 5 shows the total delays of the NYC metroplex under various service levels from 0.5 to 0.9, and the delays with the 500-scenario static model. The relative cost ratio of air to ground delay is chosen to be 2: $c_a/c_g = 2$. In Fig. 5, the number of delays will increase with the service level, which is consistent with the intuition that the high service level will result in conservative solutions. This is also confirmed by the ratio of ground to air delays. The numbers above each bar in Fig. 5 represent the ratio of ground to air delays. The low service level, which represents aggressive planning, produces more air delays than ground delays. On the other hand, the high service level, which represents conservative planning, results in more ground delays than air delays. However, the scenario-based method (Ball et al. [13] model) will only produce one average result with respect to the minimum expected cost, which lacks the ability to adjust the planning strategy under different service levels.

Figure 6 shows the details of the optimal PAARs at JFK under various service levels. The red line shows the optimal PAARs for each hour with Ball et al. [13] model, referred to as scenario, and the yellow line represents the optimal PAARs with the chance-constrained model under a certain service level. By comparing both the results with the flight schedule (the blue line), we can find that both of the optimal PAARs are almost identical with the schedule before 1200 hrs, which represents the slack time. In the peak period, the scenario-based method slightly reduced the number of planned arrivals and compensated the schedule in the last time step. However, the chance-constrained model generates totally different solutions. In general, the low service level will provide aggressive planning, in which the PAAR almost follows the schedule. On the other hand, the conservative planning under the high service level will assign more ground delays in the beginning to avoid possible air delays, and then compensates the schedule in the latter time step. Meanwhile, similar results are observed for the other two airports (EWR and LGA). Therefore, the results demonstrate that the conservativeness level is positively correlated with the robustness level. The chance-constrained model, introduced in this paper, provides a quantized way to balance the solution's robustness and potential cost by choosing a proper service level.

As mentioned in Sec. III, the parallel-computing framework can greatly improve the computational efficiency. Figure 7 shows the computation time with different numbers of processors, in which the polynomial degree is fixed to be 10. The computation time decreases as more processors are launched under the parallel-computing framework. However, the speedup is not linear to the number of processors, and the run time is decreased from 684 to 45 s with 40 processors. The reason is that the parallel-programming model has inherent overheads, such as communication and synchronization between processors. In addition, the Ball et al. [13] model with 500 scenarios only takes 5 s, which is nine times faster than the chance-constrained model. The chance-constrained model is slower because it needs to fit the polynomial based on the distribution to approximate the first-order information, whereas the Ball et al. model saves time by solving a linear programming with sparse samplings of the distribution. Although the Ball et al. model is more efficient, it lacks the ability to provide robust solutions with service level. Moreover, the 45 s run time is still efficient in practice for selecting PAARs in each hour, because the planning is often performed several hours ahead. For a larger-size problem with more than three airports, the gap could be further reduced by deploying more computing processors.

As the other key factor of the computation time, the relationship between run time and the polynomial degree is demonstrated in Fig. 8, in which the number of processors is fixed to be 40. Even though a high polynomial degree could help reduce the approximation error, the increase of degree could also explode the computation time. Therefore, the choice of the polynomial degree is a balance between the computation time and the solution accuracy. In fact, the degree can be chosen between 10 and 15 to provide a good approximation, based on the experimental experience. Further increasing the degree level will provide little help for the quality of the solution.

VI. Conclusions

This paper introduces a novel chance-constrained approach for ground-delay program planning under uncertainty. The major advantage of the chance-constrained model is the ability to provide robust solutions with a user-defined service level. The approach is compared with a seminal model for selecting planned AARs for airports in the New York City metroplex. Although the seminal model was found to be more efficient, the chance-constrained model shows the ability to provide a quantized way to balance the solution's robustness and potential cost by choosing a proper service level. Moreover, the parallel-computing framework based on the convex-approximation method was demonstrated to be helpful in improving the computing efficiency, which suggests that simply increasing the number of processors would help solve a large-scale planning problem under uncertainty with the same method and framework.

Appendix: Definitions and Theorems

Definition 1: A function $f(z) \geq 0,\ z \in \mathbb{R}^m$ is said to be logarithmically concave (in short form, log concave), if, for any z_1, z_2 and $0 < \lambda < 1$, we have the inequality:

$$f[\lambda z_1 + (1-\lambda)z_2] \geq [f(z_1)]^{\lambda}[f(z_2)]^{(1-\lambda)}$$

If $f(z) \geq 0$ for $z \in \mathbb{R}^m$, then this means that $\log f(z)$ is a concave function in $\mathbb{R}^m$.

Definition 2: A probability measure defined on the Borel sets of $\mathbb{R}^m$ is said to be logarithmically concave (log concave) if, for any convex subsets of $\mathbb{R}^m$: X, Y and $0 < \lambda < 1$, we have the inequality:

$$\mathbb{P}[\lambda X + (1-\lambda)Y] \geq [\mathbb{P}(X)]^{\lambda}[\mathbb{P}(Y)]^{(1-\lambda)}$$

in which $\lambda X + (1-\lambda)Y = \{z = \lambda x + (1-\lambda)y | x \in X, y \in Y\}$.

Based on these two definitions, we have

Theorem 1: If $\xi \in \mathbb{R}^m$ is a random variable, the probability distribution of which is log concave, then the probability distribution function $F_{\xi}(x) = \mathbb{P}(\xi \leq x)$ is a log-concave function in $\mathbb{R}^m$.

The proof of Theorem 1 and the rationale of Definitions 1 and 2 are presented in [20], and this paper omits them.

Definition 3: The Bernstein polynomial of a function $\phi(y)$, $y \in [0,1]$ is

$$B_k(\phi; y) := \sum_{j=0}^{k} \binom{k}{j} y^j (1-y)^{k-j} \phi(j/k) \tag{A1}$$

and

$$B_k(\phi; y) = \phi(0), \qquad B_k(\phi; 1) = \phi(1) \tag{A2}$$

Theorem 2 (Bernstein theorem): Let $\phi(y)$ be continuous on [0,1]. Then

$$\lim_{k\to\infty} B_k(\phi; y) = \phi(y) \tag{A3}$$

any point $y \in [0,1]$ and the limit (A3) hold uniformly in [0,1]. That is, given an $\epsilon > 0$, for all large-enough k, we have

$$|\phi(y) - B_k(\phi; y)| \leq \epsilon, \qquad y \in [0,1] \tag{A4}$$

The proof is in [22], and we omit it.

Theorem 3: There exists a sequence of component functions:

$$\psi_0(y), \psi_1(y), \psi_2(y), \ldots \tag{A5}$$

Each is convex on [0,1], such that any function $\phi(y)$ that is convex on [0,1] may be approximated with arbitrary accuracy on [0,1] by a sum of nonnegative multiples of the component functions.

The proof is in [23]. Because this result plays the central role of this paper, we will present the proof as a courtesy. We adopt our notations (not the original) as being consistent with our problem.

Proof: First, we assume that $\phi(y)$ is twice differentiable on [0,1] because, if otherwise, we can apply Theorem 2 to construct a (convex) Bernstein polynomial, which approximates $\phi(y)$ to within $\epsilon/2$ on [0,1] using a degree of $k > 2$. We then use the obtained Bernstein polynomial to replace $\phi(y)$. We use $\phi'(y)$ and $\phi''(y)$ to denote the first- and second-order derivatives of $\phi(y)$, respectively. Let

$$B_k(\phi''; y) = \sum_{j=0}^{k} \binom{k}{j} y^j (1-y)^{k-j} \phi''(j/k) \tag{A6}$$

represent the Bernstein polynomial of degree k for $\phi''(y)$. Let us observe that $y^j(1-y)^{k-j} \geq 0$ on [0,1] and that, in Eq. (A6), are being approximated by the sum of nonnegative multiples of the polynomials $y^j(1-y)^{k-j}$. For $k \geq 2$, define $p_k(y)$ by

$$p_k''(y) = B_{k-2}(\phi''; y), \qquad p_k'(0) = \phi'(0), \qquad p_k(0) = \phi(0) \tag{A7}$$

We see that $p_k(y)$ is a polynomial of degree at most k. We also define $\beta_{j,k}(y)$ for $2 \leq j \leq k$ by

$$\beta_{j,k}''(y) = y^{j-2}(1-y)^{k-j}, \qquad \beta_{j,k}'(0) = \beta_{j,k}(0) = 0 \tag{A8}$$

To complete the definition of polynomials $\beta_{j,k}(y)$, we define

$$\beta_{0,k}(y) = \mathrm{sign}[\phi(0)], \qquad \beta_{1,k}(y) = y\,\mathrm{sign}[\phi'(0)] \tag{A9}$$

The relevance of the choice of functions (A9) will be seen later. We then have

$$p_k(y) = \sum_{j=0}^{k} c_j^* \beta_{j,k}(y) \tag{A10}$$

in which $c_j^* \geq 0$ and $\beta_{j,k}''(y) \geq 0$ on [0,1]. Now, given any $\epsilon > 0$, applying Theorem 2, we have

$$|B_{k-2}(\phi''; y) - \phi''(y)| \leq \epsilon \tag{A11}$$

on [0,1]. That is

$$|p_k''(y) - \phi''(y)| \leq \epsilon \tag{A12}$$

on [0,1], and therefore, for $y \in [0,1]$

$$\left| \int_0^y [p_k''(t) - \phi''(t)]\,\mathrm{d}t \right| \leq \int_0^y |p_k''(t) - \phi''(t)|\,\mathrm{d}t \leq \epsilon y \leq \epsilon \tag{A13}$$

Using Eq. (A7), the inequality (A13) gives

$$|p_k'(y) - \phi'(y)| \leq \epsilon \tag{A14}$$

for $y \in [0,1]$. Similarly, another integration shows that

$$|p_k(y) - \phi(y)| \leq \epsilon \tag{A15}$$

for $y \in [0,1]$. Note that the polynomial $\beta_{j,k}(y)$ may be $\psi_0(y), \psi_1(y), \psi_2(y), \ldots$. We set

$$\psi_j(y) = \beta_{j,k}(y), \quad 2 \leq j \leq k, \quad \psi_0(y) = \mathrm{sign}[\phi(0)],$$
$$\psi_1(y) = y\,\mathrm{sign}[\phi'(0)] \tag{A16}$$

in which for $j \geq 2$

$$\beta_{j,k}''(y) = y^{j-2}(1-y)^{k-j} = y^{j-2} \sum_{i=1}^{k-j} (-1)^i \binom{k-j}{i} y^i \tag{A17}$$

and we have

$$\beta_{j,k}(y) = y^j \sum_{i=0}^{k-j} (-1)^i \left(\binom{k-j}{i} y^i / [(i+j)(i+j-1)] \right) \quad \text{(A18)}$$

Theorem 4 (Jackson's theorem V): If $\phi(y)$ is r differentiable on $y \in [0,1]$, and $\phi(y)$ is approximated by $p_k(y)$, which is defined in Theorem 3, then the approximation error of $\phi(y)$ on [0,1] satisfies

$$\max_{y\in[0,1]} |\phi(y) - p_k(y)| \le \left(\frac{\pi}{2}\right)^r \frac{|\phi^{(r)}(\omega)|}{[(k-r+2)\ldots k(k+1)]}, \quad k \ge r \quad \text{(A19)}$$

in which $\phi^{(r)}(\omega)$ represents the r-order derivative of $\phi(y)$ at some $\omega \in [0,1]$.

The proof of this theorem is in [24].

Proposition 1: Let $\text{poly}_k(y)$ be the polynomial constructed from $k+1$ coordinates $(y_i, \phi(y_i))$, $i = 1, \ldots, k+1$. Then

$$|\phi(y) - \text{poly}_k(y)| = \frac{\phi^{(k+1)}(\omega)}{(k+1)!} \Pi_{i=1}^{k+1} (y - y_i) \quad \text{(A20)}$$

in which ω lies in the smallest interval containing $y_1, \ldots, y_{k+1}$ and y.

This proposition is in [25]. Because we can apply Theorem 2 to approximate any continuous function by a Bernstein polynomial, which is differentiable, we can assume that the $(k+1)$-order derivative $\phi^{(k+1)}(y)$ exists, and it is a bounded value for $y \in [0,1]$. Thus, to reduce the error of approximation, we need to minimize $\Pi_{i=1}^{(k+1)}(y - y_i)$ by choosing the Chebyshev nodes on [0,1], as follows:

$$y_i := \frac{1}{2} - \frac{1}{2}\cos\left(\frac{2i-1}{2k+2}\pi\right), \quad i = 1, \ldots, k+1 \quad \text{(A21)}$$

Acknowledgments

The authors are very grateful to the editor and anonymous reviewers; in particular, the authors thank the anonymous reviewers' suggestion on justification for moving beyond the simple Bernstein approximation method and other related issues.

References

[1] Rios, J., "Aggregate Statistics of National Traffic Management Initiatives," *10th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference*, AIAA Paper 2010-9382, Sept. 2010.
[2] Wei, P., Chen, J.-T., Andrisani, D., and Sun, D., "Routing Flexible Traffic into Metroplex," *AIAA Guidance, Navigation, and Control Conference*, AIAA Paper 2011-6365, Aug. 2011.
[3] Atkins, S., "Observation and Measurement of Metroplex Phenomena," *2008 IEEE/AIAA 27th Digital Avionics Systems Conference*, IEEE Publ. Paper 3-E, Piscataway, NJ, Oct. 2008.
[4] Sridhar, B., Grabbe, S., and Mukherjee, A., "Modeling and Optimization in Traffic Flow Management," *Proceedings of the IEEE*, Vol. 96, No. 12, Dec. 2008, pp. 2060–2080. doi:10.1109/JPROC.2008.2006141
[5] Chen, J., DeLaurentis, D., and Sun, D., "Dynamic Stochastic Model for Converging Inbound Air Traffic," *Journal of Guidance, Control, and Dynamics*, Vol. 39, No. 10, Oct. 2016, pp. 2273–2283. doi:10.2514/1.G001379
[6] Odoni, A. R., "The Flow Management Problem in Air Traffic Control," *Flow Control of Congested Networks*, Springer, Berlin, 1987, pp. 269–288.
[7] Hoffman, R., and Ball, M. O., "A Comparison of Formulations for the Single-Airport Ground-Holding Problem with Banking Constraints," *Operations Research*, Vol. 48, No. 4, Aug. 2000, pp. 578–590. doi:10.1287/opre.48.4.578.12417
[8] Bertsimas, D., and Patterson, S. S., "The Air Traffic Flow Management Problem with Enroute Capacities," *Operations Research*, Vol. 46, No. 3, June 1998, pp. 406–422. doi:10.1287/opre.46.3.406
[9] Richetta, O., and Odoni, A. R., "Solving Optimally the Static Ground-Holding Policy Problem in Air Traffic Control," *Transportation Science*, Vol. 27, No. 3, Aug. 1993, pp. 228–238. doi:10.1287/trsc.27.3.228
[10] Richetta, O., and Odoni, A. R., "Dynamic Solution to the Ground-Holding Problem in Air Traffic Control," *Transportation Research Part A: Policy and Practice*, Vol. 28, No. 3, May 1994, pp. 167–185.
[11] Chang, K., Howard, K., Oiesen, R., Shisler, L., Tanino, M., and Wambsganss, M. C., "Enhancements to the FAA Ground-Delay Program Under Collaborative Decision Making," *Interfaces*, Vol. 31, No. 1, Feb. 2001, pp. 57–76. doi:10.1287/inte.31.1.57.9689
[12] Vossen, T. W. M., and Ball, M. O., "Slot Trading Opportunities in Collaborative Ground Delay Programs," *Transportation Science*, Vol. 40, No. 1, Feb. 2006, pp. 29–43. doi:10.1287/trsc.1050.0121
[13] Ball, M. O., Hoffman, R., Odoni, A. R., and Rifkin, R., "A Stochastic Integer Program with Dual Network Structure and Its Application to the Ground-Holding Problem," *Operations Research*, Vol. 51, No. 1, Feb. 2003, pp. 167–171. doi:10.1287/opre.51.1.167.12795
[14] Mukherjee, A., and Hansen, M., "A Dynamic Stochastic Model for the Single Airport Ground Holding Problem," *Transportation Science*, Vol. 41, No. 4, Nov. 2007, pp. 444–456. doi:10.1287/trsc.1070.0210
[15] Liu, P.-C. B., Hansen, M., and Mukherjee, A., "Scenario-Based Air Traffic Flow Management: From Theory to Practice," *Transportation Research Part B: Methodological*, Vol. 42, No. 7, 2008, pp. 685–702. doi:10.1016/j.trb.2008.01.002
[16] Cox, J., and Kochenderfer, M. J., "Optimization Approaches to the Single Airport Ground-Holding Problem," *Journal of Guidance, Control, and Dynamics*, Vol. 38, No. 12, 2015, pp. 2399–2406. doi:10.2514/1.G001081
[17] Nemirovski, A., and Shapiro, A., "Convex Approximations of Chance Constrained Programs," *SIAM Journal on Optimization*, Vol. 17, No. 4, 2006, pp. 969–996. doi:10.1137/050622328
[18] Prékopa, A., "Numerical Solution of Probabilistic Constrained Programming Problems," *Numerical Techniques for Stochastic Optimization*, Springer, Berlin, 1988, pp. 123–139.
[19] Saumard, A., and Wellner, J. A., "Log-Concavity and Strong Log-Concavity: A Review," *Statistics Surveys*, Vol. 8, Dec. 2014, pp. 45–114. doi:10.1214/14-SS107
[20] Shapiro, A., Dentcheva, D., and Ruszczyński, A., *Lectures on Stochastic Programming*, Soc. for Industrial and Applied Mathematics, Philadelphia, PA, Jan. 2009, pp. 50–58.
[21] Nesterov, Y., *Introductory Lectures on Convex Optimization: A Basic Course*, Vol. 87, Springer Science + Business Media, Berlin, 2013, pp. 33–50.
[22] Shisha, O., and Davis, P. J., "Interpolation and Approximation," *Mathematics of Computation*, Vol. 20, No. 93, Jan. 1966, p. 175. doi:10.2307/2004299
[23] Phillips, G., and Taylor, P., "Approximation of Convex Data," *BIT Numerical Mathematics*, Vol. 10, No. 3, 1970, pp. 324–332. doi:10.1007/BF01934201
[24] Cheney, E. W., *Introduction to Approximation Theory*, AMS Chelsea, Providence, RI, 1966, pp. 23–39.
[25] Stewart, G. W., *Afternotes on Numerical Analysis*, Soc. for Industrial and Applied Mathematics, Philadelphia, PA, Jan. 1996, pp. 53–59.
[26] Chen, J., Chen, L., and Sun, D., "Air Traffic Flow Management Under Uncertainty Using Chance-Constrained Optimization," *Transportation Research Part B: Methodological*, Vol. 102, Aug. 2017, pp. 124–141. doi:10.1016/j.trb.2017.05.014
[27] Provan, C. A., Cook, L., and Cunningham, J., "A Probabilistic Airport Capacity Model for Improved Ground Delay Program Planning," *2011 IEEE/AIAA 30th Digital Avionics Systems Conference (DASC)*, IEEE Publ., Piscataway, NJ, 2011, pp. 2B6–1.
[28] van der Vaart, A. W., *Asymptotic Statistics*, Vol. 3, Cambridge Univ. Press, Cambridge, United Kingdom, 2000, pp. 25–30.
[29] Cox, J., and Kochenderfer, M. J., "Probabilistic Airport Acceptance Rate Prediction," *AIAA Modeling and Simulation Technologies Conference*, AIAA Paper 2016-0165, Jan. 2016.
[30] Anon., "FAA Operations and Performance Data," Federal Aviation Administration, http://aspm.faa.gov [retrieved 24 March 2017].
[31] Silverman, B. W., *Density Estimation for Statistics and Data Analysis*, Vol. 26, CRC Press, Boca Raton, FL, 1986, pp. 345–350.
[32] Gurobi, O. I., "Gurobi Optimizer Reference Manual," 2017, http://www.gurobi.com/documentation/7.0/refman.pdf [retrieved 27 March 2017].