New *a priori* and *a posteriori* probabilistic bounds for robust counterpart optimization: III. Exact and near-exact *a posteriori* expressions for known probability distributions

Yannis A. Guzman^{c,a,b,*}, Logan R. Matthews^{c,a,b}, Christodoulos A. Floudas^{a,b}

^aArtie McFerrin Department of Chemical Engineering, Texas A&M University, College Station, TX, USA

^bTexas A&M Energy Institute, Texas A&M University, College Station, TX, USA

^cDepartment of Chemical and Biological Engineering, Princeton University, Princeton, NJ, USA

Abstract

The performance of robust optimization is closely connected with probabilistic bounds that determine the probability of constraint violation due to uncertain parameter realizations. In Part I of this work, new *a priori* and *a posteriori* probabilistic bounds were developed for cases when robust optimization is applied to uncertain optimization problems with parameters whose probability distributions were unknown. In Part II, the focus shifted to known probability distributions and *a priori* bounds. In this paper, new, tight *a posteriori* expressions are developed for constraints containing parameters with specific known distributions, that is, those attributed normal, uniform, discrete, gamma, chi-squared, Erlang, or exponential distributions. The nature of some of the expressions requires efficient implementations, and new algorithmic methods are discussed which greatly improve applicability. These new expressions are much tighter than existing bounds and greatly reduce the conservatism of robust solutions. The theoretical and algorithmic results of Parts I, II, and III allow for wider usage of robust optimization in process synthesis and operations research applications.

1. Introduction

One source of deviation between mathematical models and real-world performance is the presence of model parameters whose exact values are not yet known or are unable to be statically determined. These uncertain parameters can have drastic, detrimental effects on the objective values and feasibility of

*Corresponding author.

Email address: yguzman@alumni.princeton.edu (Yannis A. Guzman)

optimal solutions to mathematical optimization problems when they achieve realizations different from the static “nominal” values used to obtain the solutions. Robust optimization has received considerable attention as a method for incorporating parameter uncertainty into mathematical optimization problems due to (a) coverage of the continuous uncertain space, (b) low computational demands, and (c) high impact to the performance of the obtained solutions. Robust optimization addresses parameter uncertainty by providing solutions that are valid for either all possible parameter values, or for a subset of possible realizations described by an uncertainty set. The deterministic reformulations of models under uncertainty, known as robust counterparts, ensure the feasibility of solutions given an uncertainty set while also providing tractable optimization problems. Starting from the “worst-case” counterpart developed by Soyster (1973), the developments in robust optimization have been rapidly occurring in recent years.

After the work of Soyster, robust optimization was applied to problems such as uncertain least-squares (El Ghaoui and Lebret, 1997) or semi-definite optimization problems (El Ghaoui et al., 1998). Alternate formulations for uncertainty sets were also proposed; Ben-Tal and Nemirovski (2000) developed the interval + ellipsoidal uncertainty set for linear programs (LPs), while Bertsimas and Sim (2004) proposed an interval + polyhedral uncertainty set. With these new uncertainty sets arose concepts of nonzero probabilities of constraint violation, as these uncertainty sets did not always include the entire uncertain parameter space. Floudas and coworkers (Janak et al., 2007; Lin et al., 2004; Verderame and Floudas, 2009a,b) expanded robust counterpart optimization theory to include mixed-integer linear optimization problems (MILPs), and robust optimization became a comprehensive methodology for including uncertainty in optimization problems with linear or mixed-integer linear uncertain constraints. The applications are widespread, ranging from design under uncertainty (Floudas et al., 2001) to various scheduling problems (Janak et al., 2006; Lappas and Gounaris, 2016; Li et al., 2012a) to the capacitated vehicle routing problem (Gounaris et al., 2013), among others. Floudas et al. (2016) highlighted the power and importance of robust optimization for addressing uncertainty in other important applications in a recent perspective article on challenges and opportunities of multi-scale engineering for energy and the environment.

The theory of robust optimization has expanded into a variety of areas in order to improve its performance, as discussed in more detail elsewhere (Ben-Tal et al., 2009; Ben-Tal and Nemirovski, 2007; Bertsimas et al., 2011; Gabrel et al., 2014). As an example, adjustable-robust optimization has been developed for uncertain problems with recourse (Ben-Tal et al., 2003; Bertsimas and Goyal, 2012; Chen and Zhang, 2009). Theory regarding uncertainty sets and robust counterparts has also been expanded to increase robust optimization’s applicability (Averbakh and Zhao, 2008; Ben-Tal et al., 2006; Bertsimas and Brown, 2009). However, there is considerable potential for reductions in conservatism and increased applicability through traditional robust optimization with box, ellipsoidal, interval + ellipsoidal, polyhedral, and interval + polyhedral uncertainty sets by tightening the bounds on probabilities of constraint violation both

a priori and *a posteriori*, which this series of papers has and will continue to demonstrate.

A priori probability bounds determine the size of the uncertainty set necessary to guarantee an upper bound on the probability of constraint violation for all possible solutions to an uncertain model and are applied to formulate the robust counterpart. A tighter *a priori* bound will yield a smaller uncertainty set and a better robust objective function value at a given probability of constraint violation. Upper bounds on the probability of constraint violation, calculated *a priori*, have been proposed for a variety of situations and uncertainty sets, both when the parameter distribution is unknown (Ben-Tal and Nemirovski, 2000; Bertsimas and Sim, 2004) or known (Kang et al., 2013). These bounds were extended by Li et al. (2012b) to be broadly applicable across uncertainty sets. In Part I (Guzman et al., 2016a) and Part II (Guzman et al., 2016b) of this work, new *a priori* bounds were proposed for cases with unknown and known parameter distributions, respectively. For the first time, *a priori* bounds were provided for distributions that are unknown and asymmetric with limited information about the means. For cases where *a priori* bounds already existed, new bounds were provided that are tighter than the existing bounds. These new *a priori* bounds greatly improved the performance of traditional one-pass robust optimization techniques as demonstrated through LP and MILP example problems (Guzman et al., 2016a,b).

In general, *a posteriori* bounds are tighter than *a priori* bounds as they characterize one specific solution to an uncertain optimization problem, rather than all feasible solutions for a given uncertainty set. After solving a robust model which was formulated using robust counterpart constraints and *a priori* bounds, the obtained robust solution can be re-characterized using *a posteriori* bounds. An *a posteriori* bound will always be as tight or tighter than its corresponding *a priori* analogue, and may characterize the robust solution as still being conservative relative to the acceptable probability of constraint violation used to construct the robust counterpart *a priori*. *A posteriori* probabilistic bounds have been previously developed for a variety of parameter characteristics (Kang et al., 2013; Li et al., 2012b; Paschalidis et al., 2008). Obtaining a robust solution with a specific *a posteriori* probability of constraint violation is more challenging than ensuring a specific *a priori* probability. Li and Floudas (2014) proposed two methods for achieving a specific *a posteriori* probability. First, the *a posteriori* bounds can be placed directly into the model itself, thus treating *a posteriori* bounds as a sort of alternative to formulating robust counterpart constraints. This in general yields highly nonconvex models which in some cases can be optimized via state-of-the-art deterministic global optimization approaches such as GloMIQO or ANTIGONE (Gounaris et al., 2009; Misener and Floudas, 2010, 2012, 2013, 2014a,b; Misener et al., 2015, 2011). An iterative approach utilizing *a priori* and *a posteriori* bounds was also proposed that quickly produces high quality *a posteriori* solutions without including the bounds directly in the optimization problem. Thus, with the ability to approach solutions with specific *a posteriori* probabilities, increasing the tightness of *a posteriori* bounds can greatly reduce conservatism in robust opti-

mization. In Part I of this series (Guzman et al., 2016a), new *a posteriori* bounds were derived that are as tight or tighter than existing bounds for parameters with unknown distributions. Here, we derive new *a posteriori* expressions for cases when uncertain parameters are subject to specific distributions, including normal, uniform, discrete, gamma, chi-squared, Erlang, or exponential distributions, and which either calculate the exact probability of constraint violation or bound the exact probability and approach it to arbitrary proximity.

This work is structured in the following manner. First, detailed background on robust optimization and previous advances in probabilistic bounds will be provided in Section 2. Then, the derivations for expressions calculating exact or near-exact probabilities of constraint violation will be shown in Section 3. Because of the new *a posteriori* bounds for gamma distributions, a discussion of the utilization of gamma distributions in *a priori* bounds is provided in the appendices. For two of the expressions, efficient implementation is crucial; different algorithms, including an advantageous pruning algorithm for addend enumeration, are constructed and computationally evaluated in Section 4. Finally, in Section 5, the performance of the new expressions will be compared to existing *a posteriori* bounds in an example multiperiod planning problem.

2. Background

The advancements provided in this work are applicable to models with linear or mixed-integer linear uncertain objective function and constraints. Given continuous variables x_k and integer variables y_ℓ , consider inequality constraint i which, possibly after reformulations, exhibits the form

$$f_i(x, y) + \sum_k a_{ik}x_k + \sum_\ell b_{i\ell}y_\ell + \sum_m p_{im} \leq 0, \quad (1)$$

where the exact value of some or all of the parameters a_{ik} , $b_{i\ell}$, and p_{im} are unknown, and $f_i(x, y)$ can be nonlinear. Constraint (1) can be equivalently reformulated as:

$$\begin{aligned} f_i(x, y) + t_i &\leq 0 \\ -t_i + \sum_k a_{ik}x_k + \sum_\ell b_{i\ell}y_\ell + \sum_m p_{im} &\leq 0. \end{aligned} \quad (2)$$

Note that the uncertain parameters of (1) now participate in a linear inequality constraint. An objective function under uncertainty can also be reformulated into an uncertain constraint. Thus, we will assume that all uncertain parameters lie within uncertain constraints that are linear with continuous or integer variables. The typical presentation of robust optimization goes further and assumes that the entire model is an LP or MILP. We present the background section under this assumption so as to maintain consistency.

The general form of an LP or MILP under uncertainty is as follows:

$$\begin{aligned}
& \max_{x,y} \quad \sum_k \tilde{c}_k x_k + \sum_\ell \tilde{d}_\ell y_\ell \\
& \text{s.t.} \quad \sum_k \tilde{a}_{ik} x_k + \sum_\ell \tilde{b}_{i\ell} y_\ell \leq \tilde{p}_i \quad \forall i \\
& \quad \quad y_\ell \in \{0, 1\} \quad \forall \ell
\end{aligned} \tag{3}$$

where any parameter denoted with a tilde, say $\tilde{p}$, denotes a parameter subject to uncertainty. Different realizations of the uncertain parameters can change the solution, objective function value, or feasibility of (3). Model (3) can be equivalently written (Li et al., 2011, 2012b) as:

$$\begin{aligned}
& \max_{x,y,z} \quad z \\
& \text{s.t.} \quad z - \sum_k \tilde{c}_k x_k - \sum_\ell \tilde{d}_\ell y_\ell \leq 0 \\
& \quad \quad \tilde{p}_i x_0 + \sum_k \tilde{a}_{ik} x_k + \sum_\ell \tilde{b}_{i\ell} y_\ell \leq 0 \quad \forall i \\
& \quad \quad x_0 = -1 \\
& \quad \quad y_\ell \in \{0, 1\} \quad \forall \ell.
\end{aligned} \tag{4}$$

An uncertain parameter $\tilde{a}$ can be represented by a constant nominal value a , which is typically set to $\mathbf{E}[\tilde{a}]$, and a random variable ξ which subjects a to perturbations weighted by a known positive constant $\hat{a}$:

$$\tilde{a} = a + \xi \hat{a}. \tag{5}$$

Models (3) and (4) can be equivalently represented as:

$$\begin{aligned}
& \max_x \quad \sum_j c_j x_j \\
& \text{s.t.} \quad \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j \leq b_i \quad \forall i
\end{aligned} \tag{6}$$

where x_j is a continuous, integer, or fixed variable, and $J_i := \{j : \hat{a}_{ij} > 0\}$ is the set of indices where the coefficients of x_j in constraint i are subject to uncertainty.

A nominal solution to (6) is one obtained when each uncertain parameter $\tilde{a}_{ij}$ is at its nominal value a_{ij} , that is, when $\xi_{ij} = 0$, $\forall j \in J_i$, $\forall i$, while a robust solution will guarantee feasibility for more realizations of $\tilde{a}_{ij}$. The parameter space for which a robust solution is forced to guarantee feasibility is called an uncertainty set. We denote the uncertainty set of constraint i as U_i , and define it in terms of random vector ξ_i with elements ξ_{ij} , $j \in J_i$. For constraint i , U_i is defined to contain some or all realizations of ξ_i , including the point $\xi_i = \mathbf{0}$ which corresponds to the nominal solution. Those vectors in U_i are then imposed upon

the feasible space of constraint i . The robust counterpart of model (6) can thus be written as

$$\begin{aligned} \max_x \quad & \sum_j c_j x_j \\ \text{s.t.} \quad & \sum_j a_{ij} x_j + \max_{\xi_i \in U_i} \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j \right\} \leq b_i \quad \forall i. \end{aligned} \quad (7)$$

One way to construct uncertainty sets is to parameter realizations within some scaled norm distance of their nominal value. The simplest set, originally used by Soyster (1973) includes all possible values for bounded parameters, forming the worst-case interval set with box geometry. This can be generalized using the ∞ -norm distance around the nominal value in order to change the size of the box, and is called the box uncertainty set (Li et al., 2011):

$$U_i^\infty = \{\xi_i : \|\xi_i\|_\infty = \max_{j \in J_i} |\xi_{ij}| \leq \Psi_i\}. \quad (\text{B})$$

Soyster's worst-case set corresponds to $\Psi_i = 1$ and is known as the interval uncertainty set.

Similarly, the 1-norm or 2-norm can be used to define the uncertainty sets, resulting in the polyhedral and ellipsoidal uncertainty sets, respectively (Ben-Tal and Nemirovski, 1998, 1999, 2000; Bertsimas and Sim, 2004; El Ghaoui and Lebret, 1997; El Ghaoui et al., 1998; Li et al., 2011):

$$U_i^1 = \{\xi_i : \|\xi_i\|_1 = \sum_{j \in J_i} |\xi_{ij}| \leq \Gamma_i\} \quad (\text{P})$$

$$U_i^2 = \{\xi_i : \|\xi_i\|_2 = \sqrt{\sum_{j \in J_i} \xi_{ij}^2} \leq \Omega_i\}. \quad (\text{E})$$

Note that when set parameters Ψ_i , Γ_i , and Ω_i equal zero, the corresponding uncertainty set corresponds to the nominal parameter point; increasing the set parameters increases the size of U_i . These set parameters are generally represented as Δ_i for an arbitrary uncertainty set U_i . If it is assumed that $\xi_{ij} \in [-1, 1]$, $\forall j \in J_i$, all possible parameter realizations are included in the box uncertainty set when $\Psi_i = 1$ (i.e., the interval uncertainty set). Similarly, this occurs in polyhedral or ellipsoidal uncertainty sets when $\Gamma_i = |J_i|$ and $\Omega_i = \sqrt{|J_i|}$, respectively (Li et al., 2011). However, due to the geometry of the uncertainty sets at these values of Γ_i and Ω_i , parameter realizations beyond those in the worst-case set are included. For bounded uncertainty, the interval + polyhedral (Bertsimas and Sim, 2004) and interval + ellipsoidal (Ben-Tal and Nemirovski, 2000) uncertainty sets address this through intersection of the interval and polyhedral sets and intersection of the interval and ellipsoidal sets, respectively:

$$U_i^{1 \cap \infty} = \{\xi_i : \|\xi_i\|_1 \leq \Gamma_i, \|\xi_i\|_\infty \leq 1\}. \quad (\text{IP})$$

$$U_i^{2 \cap \infty} = \{\xi_i : \|\xi_i\|_2 \leq \Omega_i, \|\xi_i\|_\infty \leq 1\} \quad (\text{IE})$$

Thus, the interval + polyhedral and interval + ellipsoidal sets are never larger than the worst-case set for bounded uncertainty. Table A.1 summarizes the five uncertainty sets considered in this work. The inclusion of uncertainty sets in robust counterpart formulations is summarized in Table A.2 and discussed in detail in (Li et al., 2011, 2012b; Li and Floudas, 2014).

If a model contains only bounded uncertain parameters, and if the solution must remain feasible for any realization of the uncertain parameters, only the interval uncertainty set should be used. However, it is often the case that a small, nonzero probability that constraint i is infeasible can be permitted, or that some of the parameters are unbounded; the objective function value can then be significantly improved when compared to the worst-case solution. To do this, a probability of constraint violation, ϵ_i^{prio} , is set before solving the model, representing an upper bound on the probability that constraint i is violated in the optimal solution due to uncertain parameter realizations:

$$\Pr \left\{ \sum_j a_{ij}x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij}x_j > b_i \right\} \leq \epsilon_i^{prio}. \quad (8)$$

Knowing the desired ϵ_i^{prio} , the size of set U_i is chosen by selecting Δ_i values that are less than the worst-case value. Methods that relate ϵ_i^{prio} to U_i are known as *a priori* probabilistic bounds. Once a robust counterpart is solved, the specific solution x^* can be used to determine an upper bound ϵ_i^{post} on the probability that the solution will be infeasible in constraint i with any parameter realization:

$$\Pr \left\{ \sum_j a_{ij}x_j^* + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij}x_j^* > b_i \right\} \leq \epsilon_i^{post}. \quad (9)$$

Methods relating specific solutions of the robust counterpart to probabilities of constraint violation are known as *a posteriori* probabilistic bounds. The tightness of *a priori* or *a posteriori* bounds is of great relevance to optimal solutions; a tighter *a priori* bound for a specified ϵ_i^{prio} will provide a better objective function value, while a tighter *a posteriori* bound provides a lower value of ϵ_i^{post} for a known optimal solution than looser bounds. Generally, when more distributional information on the uncertain parameters is available, tighter bounds can be applied.

2.1. Summary of best available *a priori* and *a posteriori* bounds

The remainder of this section will discuss the best available bounds given various situations (Guzman et al., 2016a,b). Uncertain parameters $\tilde{a}_{ij}$ are modeled as in (5); if $\tilde{a}_{ij}$ is subject to a bounded distribution, $\xi_{ij} \in [-1, 1]$. The parameter $\delta_i^{(1)}$ depends on U_i and is given in Table A.3. For *a posteriori* bounds, $h_i(x) := b_i - \sum_j a_{ij}x_j$.

In many cases, especially at low probabilities of constraint violation, knowing the probability distribution of the uncertain parameters can further decrease

the conservatism of *a priori* bounds. As discussed in Part II (Guzman et al., 2016b), these probability distributions can either be rigorously modeled or often conservatively approximated using the uniform distribution. For each probability distribution of random variable ξ_{ij} , the bounds for known distributions require the existence of moment generating function $M_{ij}(\theta)$ and the cumulant generating function $\Lambda_{ij}(\theta)$ corresponding to the distribution of ξ_{ij} :

$$\begin{aligned} M_{ij}(\theta) &:= \mathbf{E}[e^{\theta\xi_{ij}}] \\ \Lambda_{ij}(\theta) &:= \ln M_{ij}(\theta). \end{aligned}$$

The probability distributions can either be symmetric or asymmetric. The asymmetric bounds are dependent on parameters g_{ij} corresponding to each ξ_{ij} subject to an asymmetric probability distribution:

$$g_{ij} := \begin{cases} +1 & \text{if } M_{ij}(\theta) \geq M_{ij}(-\theta) \quad \forall \theta \geq 0 \\ -1 & \text{if } M_{ij}(\theta) \leq M_{ij}(-\theta) \quad \forall \theta \geq 0, \end{cases} \quad (10)$$

The full list of the best performing bounds is below.

- A. If there exist bounded uncertain parameters whose expected values are not known precisely, the only *a priori* bound available is (Guzman et al., 2016a):

$$\begin{aligned} \Pr \left\{ \sum_j a_{ij}x_j + \sum_{j \in J_i} \xi_{ij}\hat{a}_{ij}x_j > b_i \right\} \\ \leq \exp \left(\min_{\theta > 0} \left\{ -\theta\Delta_i + \sum_{k \in K_i} \ln G_{ik}(\theta) \right\} \right) \quad (\text{GMF1}) \end{aligned}$$

where $G_{ij}(\theta) := \mu_{ij} \sinh(\theta) + \cosh(\theta)$ and K_i contains the indices of the $\lceil \delta_i^{(1)} \rceil$ largest μ_{ij} . Constants μ_{ij} and nominal values a_{ij} are chosen such that $|\mathbf{E}[\xi_{ij}]| \leq \mu_{ij}$.

- B. If the uncertainty set U_i is an ellipsoidal or interval + ellipsoidal set, the following bound can improve the result of (GMF1) (Guzman et al., 2016a):

$$\begin{aligned} \Pr \left\{ \sum_j a_{ij}x_j + \sum_{j \in J_i} \xi_{ij}\hat{a}_{ij}x_j > b_i \right\} \\ \leq \exp \left(\min_{\theta > 0} \left\{ -\theta\Delta_i + |J_i| \ln \bar{G}_i \left(\theta / \sqrt{|J_i|} \right) \right\} \right) \quad (\text{GMF1}') \end{aligned}$$

where $\bar{\mu}_i = \max_{j \in J_i} \mu_{ij}$ and $\bar{G}_i(\theta) := \bar{\mu}_i \sinh(\theta) + \cosh(\theta)$.

- C. The *a posteriori* analogue of bounds (GMF1) and (GMF1') is (Guzman et al., 2016a):

$$\Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \ln G_{ij}(|\hat{a}_{ij} x_j^*| \theta) \right\} \right). \quad (\text{GMF4})$$

- D. When all the expected values of the uncertain parameters are known exactly, but the distributions remain completely unknown, the best available *a priori* bound for box, polyhedral, and interval + polyhedral sets is (Guzman et al., 2016a):

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq (\alpha^\alpha \beta^\beta)^{-\delta_i^{(1)}/2} \quad (\text{GMF2})$$

where $\alpha := 1 - \Delta_i / \delta_i^{(1)}$ and $\beta := 1 + \Delta_i / \delta_i^{(1)}$.

- E. For ellipsoidal and interval + ellipsoidal sets, the best available bound is (Guzman et al., 2016a):

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq (\alpha^\alpha \beta^\beta)^{-|J_i|/2}. \quad (\text{GMF2}')$$

- F. The *a posteriori* analogue of bounds (GMF2) and (GMF2') is (Paschalidis et al., 2008):

$$\Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \ln \cosh(\hat{a}_{ij} x_j^* \theta) \right\} \right). \quad (\text{PKL})$$

- G. Finally, if the uncertain parameters belong to symmetric distributions, the best available bound for box, polyhedral, and interval + polyhedral sets is (Guzman et al., 2016a):

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \frac{1}{2^{\tilde{n}}} \sum_{k=\lceil \tilde{v} \rceil}^{\tilde{n}} \binom{\tilde{n}}{k} \quad (\text{GMF3})$$

where $\tilde{\nu} := (\Delta_i + \tilde{n})/2$ and

$$\tilde{n} := \begin{cases} |J_i| + 1 - \operatorname{sgn} \Delta_i & \text{if } \lceil \Delta_i \rceil + |J_i| \text{ is even} \\ |J_i| - \operatorname{sgn} \Delta_i & \text{if } \lceil \Delta_i \rceil + |J_i| \text{ is odd.} \end{cases}$$

In certain situations, bound (GMF2') might still outperform bound (GMF3) with ellipsoidal and interval + ellipsoidal sets.

H. The *a posteriori* analogue of bound (GMF3) is (Guzman et al., 2016a):

$$\Pr \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^* > h_i(x^*) \right\} \leq \min \left\{ \frac{1}{2^{n_a}} \sum_{k=\lceil \nu_a \rceil}^{n_a} \binom{n_a}{k}, \frac{1}{2^{n_b}} \sum_{k=\lceil \nu_b \rceil}^{n_b} \binom{n_b}{k} \right\} \quad (\text{GMF5})$$

where

$$\begin{aligned} \nu_a &:= \frac{h_a + n_a}{2} \\ n_a &:= \begin{cases} |J_i^{\delta,*}| + 1 - \operatorname{sgn} h_a & \text{if } \lceil h_a \rceil + |J_i^{\delta,*}| \text{ is even} \\ |J_i^{\delta,*}| - \operatorname{sgn} h_a & \text{if } \lceil h_a \rceil + |J_i^{\delta,*}| \text{ is odd} \end{cases} \\ h_a &:= \frac{\Delta_i}{\max_{j \in J_i^{\delta,*}} |\delta_{ij}^*|} \\ J_i^{\delta,*} &:= \{j : |\delta_{ij}^*| > 0\} \\ \nu_b &:= \frac{h_b + n_b}{2} \\ n_b &:= \begin{cases} |J_i^*| + 1 - \operatorname{sgn} h_b & \text{if } \lceil h_b \rceil + |J_i^*| \text{ is even} \\ |J_i^*| - \operatorname{sgn} h_b & \text{if } \lceil h_b \rceil + |J_i^*| \text{ is odd} \end{cases} \\ h_b &:= \frac{h_i(x^*)}{\max_{j \in J_i^*} \hat{a}_{ij} |x_j^*|} \\ J_i^* &:= \{j : \hat{a}_{ij} |x_j^*| > 0\}, \end{aligned}$$

and constants δ_{ij}^* are given by their definitions and the solution of the robust counterpart.

I. The first bound for known, symmetric probability distributions was de-

rived by Kang et al. (2013) for the interval + polyhedral uncertainty sets:

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{j \in J_i} \Lambda_{ij}(\theta) \right\} \right). \quad (\text{KBP1})$$

This bound is tight for polyhedral-based uncertainty sets, but tighter bounds for the box and ellipsoidal-based uncertainty sets were derived in Part II (Guzman et al., 2016b). The derivation of these bounds require the definitions of Lemmas 1, 2, and 3 defined in Appendix D; Lemmas 2 and 3 are also transformed in Appendix D into conditions (C₁) and (C₂) which will be referenced subsequently.

- J. Bound (GMF6) is then defined which is equivalent to (KBP1) for the polyhedral and interval + polyhedral sets and tighter than the (KBP1) for the box, ellipsoidal, and interval + ellipsoidal sets:

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{k \in K_i} \Lambda_{ik}(\theta) \right\} \right) \quad (\text{GMF6})$$

In (GMF6), K_i is found using the ordering in (C₁) and contains the first $\lceil \delta_i^{(1)} \rceil$ indices. The parameter $\delta_i^{(1)}$ is uncertainty set dependent and is defined in Table A.3.

- K. Bound (GMF7) is similar to bound (GMF6) but utilizes the max operator rather than the ordering provided by (C₁):

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \delta_i^{(1)} \max_{j \in J_i} \Lambda_{ij}(\theta) \right\} \right). \quad (\text{GMF7})$$

This is equivalent in tightness to (GMF6) when each parameter has the same probability distribution, but can be conservative when the distributions differ.

- L. An ellipsoidal-specific adjustment to the (GMF7) can be made when condition (C₂) is satisfied for all $\theta > 0$, and can provide significant improvements:

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + |J_i| \max_{j \in J_i} \Lambda_{ij} \left(\theta / \sqrt{|J_i|} \right) \right\} \right). \quad (\text{GMF7'})$$

This bound is only specific to the ellipsoidal and interval + ellipsoidal sets.

- M. An asymmetric analogue for (KBP1) is defined by

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{j \in J_i} \Lambda_{ij}(g_{ij}\theta) \right\} \right). \quad (\text{GMF8})$$

- N. An asymmetric analogue for (GMF6) takes the form of

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{k \in K_i} \Lambda_{ik}(g_{ik}\theta) \right\} \right). \quad (\text{GMF9})$$

- O. Finally, bounds (GMF7) and (GMF7') have asymmetric analogues in (GMF10) and (GMF10').

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \delta_i^{(1)} \max_{j \in J_i} \Lambda_{ij}(g_{ij}\theta) \right\} \right) \quad (\text{GMF10})$$

$$\Pr \left\{ \sum_j a_{ij} x_j + \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j > b_i \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + |J_i| \max_{j \in J_i} \Lambda_{ij} \left(g_{ij}\theta / \sqrt{|J_i|} \right) \right\} \right) \quad (\text{GMF10'})$$

A summary of the *a priori* bounds for each case (unknown, known symmetric, known asymmetric) can be found in Tables 5 through 7. The best *a priori* bounds based on the available information for the uncertain parameters are shown in Table 8. These *a priori* bounds serve as important starting points when utilizing *a posteriori* bounds and improving the quality of robust solutions through iterative methods.

3. Exact expressions and near-exact bounds for *a posteriori* probabilities

In general, it is easier to derive an upper bound on the probability of constraint violation for one specific solution than it is to derive one for all possible solutions. Thus, *a posteriori* bounds are simpler to derive than *a priori* bounds. Furthermore, an *a posteriori* bound can be extended from the derivation of an *a priori* bound when specific solution values are used.

For uncertain constraint i and solution x^* , only those uncertain parameters $\tilde{a}_{ij}$ with nonzero coefficient x_j^* needs to be considered. With the active set of uncertain indices in constraint i defined as $J_i^* := \{j : \hat{a}_{ij} |x_j^*| > 0\}$, we can represent the *a posteriori* probability of constraint violation using random variables $\tilde{a}_{ij}$ or ξ_{ij} :

$$\Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} \quad (11)$$

$$= \Pr \left\{ \sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* > h_i(x^*) \right\} \quad (12)$$

$$= \Pr \left\{ \sum_{j \notin J_i^*} a_{ij} x_j^* + \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > b_i \right\} \quad (13)$$

$$= \Pr \left\{ \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > \tau_i(x^*) \right\} \quad (14)$$

$$(15)$$

where $h_i(x^*) := b_i - \sum_j a_{ij} x_j^*$ and $\tau_i(x^*) := b_i - \sum_{j \notin J_i^*} a_{ij} x_j^*$. Note that $\tau_i(x^*) = h_i(x^*) + \sum_{j \in J_i^*} a_{ij} x_j^*$. Kang et al. (2013) derived an analogue *a posteriori* bound to the distribution-dependent *a priori* bounds shown in Section 2:

$$\Pr \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^* > h_i(x^*) \right\} \leq \exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \Lambda_{ij}(\hat{a}_{ij} x_j^* \theta) \right\} \right). \quad (\text{KBP2})$$

Because x^* is known *a posteriori*, it is possible to represent the exact *a posteriori* probability, not probability bound, on constraint violation. For example,

using expression (12) and defining random variable $\Xi_i := \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^*$, the probability of constraint violation can be calculated as such:

$$\begin{aligned} \Pr \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij} x_j^* > h_i(x^*) \right\} &= \Pr \{ \Xi_i > h_i(x^*) \} \\ &= 1 - F_{\Xi_i}(h_i(x^*)), \end{aligned}$$

where $F_{\Xi_i}(x)$ is the cumulative distribution function of the distribution of Ξ_i . This requires the derivation of the distribution of a weighted sum of random variables, in this case ξ_{ij} with coefficients $\hat{a}_{ij} x_j^*$, which is in general not trivially constructed. In the sections that follow, we present cases for when this distribution can be constructed and then used to yield exact expressions or probabilistic bounds with arbitrary tightness in a robust optimization setting. Constraints with uncertain parameters which do not fall under one of these cases must default to the more conservative bound (KBP2); computational comparisons are made between (KBP2) and the expressions that follow in Section 5.

3.1. Normal distributions

Consider the case where uncertain parameters $\tilde{a}_{ij}$, $\forall j \in J_i^*$, are modeled with normal distributions:

$$\tilde{a}_{ij} \sim N(\mu_{ij}, \sigma_{ij}^2) \quad \forall j \in J_i^*$$

where μ_{ij} and σ_{ij}^2 are the expected value and variance of $\tilde{a}_{ij}$, respectively. To define ξ_{ij} , we can select the nominal and perturbation values of $\tilde{a}_{ij}$ to be its expected value and standard deviation:

$$\tilde{a}_{ij} = \mu_{ij} + \xi_{ij} \sigma_{ij} \quad \forall j \in J_i^*.$$

Then, ξ_{ij} are subject to the standard normal distribution, that is, $\xi_{ij} \sim N(0, 1)$, $\forall j \in J_i^*$. Random variables $X_{ij} := \hat{a}_{ij} x_j^* \xi_{ij}$ can be introduced, where $X_{ij} \sim N(0, \hat{a}_{ij}^2 x_j^{*2})$. The sum of normally distributed independent variables is normally distributed with an expected value and variance given by the sum of those of the component distributions; thus, $\Xi_i := \sum_{j \in J_i^*} X_{ij} \sim N(0, \sum_{j \in J_i^*} \hat{a}_{ij}^2 x_j^{*2})$. Therefore, if $\xi_{ij} \sim N(0, 1)$, $\forall j \in J_i^*$,

$$\begin{aligned} \Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} &= 1 - F_{\Xi_i}(h_i(x^*)) \\ &= 1 - \frac{1}{2} \left(1 + \operatorname{erf} \left(\frac{h_i(x^*)}{\sigma \sqrt{2}} \right) \right) \\ &= \frac{1}{2} \operatorname{erfc} \left(\frac{h_i(x^*)}{\sigma \sqrt{2}} \right), \quad (\text{GMF11}) \end{aligned}$$

where $\sigma^2 := \sum_{j \in J_i^*} \hat{a}_{ij}^2 x_j^{*2}$.

3.2. Uniform distributions

Consider the case where uncertain parameters $\tilde{a}_{ij}$, $\forall j \in J_i^*$, are modeled with continuous uniform distributions supported in $[a_{ij}^L, a_{ij}^U]$:

$$\tilde{a}_{ij} \sim U(a_{ij}^L, a_{ij}^U) \quad \forall j \in J_i^*.$$

To define ξ_{ij} such that $\xi_{ij} \in [-1, 1]$, we can select the nominal and perturbation values of $\tilde{a}_{ij}$ as such:

$$a_{ij} = \frac{a_{ij}^U + a_{ij}^L}{2}$$

$$\hat{a}_{ij} = \frac{a_{ij}^U - a_{ij}^L}{2}.$$

The distribution of the weighted sum of independent uniform random variables with arbitrary support was derived in (Kamgar-Parsi et al., 1995). We can then introduce $X_{ij} := \xi_{ij} \hat{a}_{ij} x_j^* \sim U(-\hat{a}_{ij} x_j^*, \hat{a}_{ij} x_j^*)$ and define $\Xi_i := \sum_{j \in J_i^*} X_{ij}$. Therefore, if $\xi_{ij} \sim U(-1, 1)$, $\forall j \in J_i^*$,

$$\begin{aligned} \Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} &= 1 - F_{\Xi_i}(h_i(x^*)) \\ &= 1 - \left(-\frac{1}{2}\right)^n \frac{1}{n! \chi_i} \sum_{s_1=\pm 1} \cdots \sum_{s_n=\pm 1} (\psi_i)(h_i(x^*) - \omega_i)^n H(h_i(x^*) - \omega_i) \\ &= \left(-\frac{1}{2}\right)^n \frac{1}{n! \chi_i} \sum_{s_1=\pm 1} \cdots \sum_{s_n=\pm 1} (\psi_i)(h_i(x^*) - \omega_i)^n H(\omega_i - h_i(x^*)) \\ &= \frac{1}{n! \chi_i} \sum_{s_1=\pm 1} \cdots \sum_{s_n=\pm 1} (\psi_i) \left(\frac{\omega_i - h_i(x^*)}{2} \right)^n H(\omega_i - h_i(x^*)) \quad (\text{GMF12}) \end{aligned}$$

where

$$\begin{aligned} j &= 1, 2, \dots, n \\ n &:= |J_i^*| \\ \chi_i &:= \prod_{j=1}^n \hat{a}_{ij} x_j^* \\ \psi_i &:= \prod_{j=1}^n s_j \\ \omega_i &:= \sum_{j=1}^n s_j \hat{a}_{ij} x_j^* \\ H(x) &:= \frac{1}{2} (1 + \operatorname{sgn} x). \end{aligned}$$

Note that $H(x)$ is the Heaviside function; between $H(x)$ and the previous term with exponent n , only addends with $\omega_i - h_i(x^*) > 0$ are nonzero, which include at most half of the possible addends since $h_i(x^*)$ is expected to be nonnegative and Ξ_i follows a symmetric distribution. These factors are taken into account in the construction of an efficient implementation of (GMF12) (see Section 4).

3.3. Discrete distributions

Consider the case where $\tilde{a}_{ij}$, $j \in J_i^*$, are subject to finite discrete distributions which are not necessarily identical. Parameter $\tilde{a}_{ij}$ can realize one of m_{ij} possible values, that is,

$$\tilde{a}_{ij} \in \{a_{ij1}, a_{ij2}, \dots, a_{ijk}, \dots, a_{ijm_{ij}}\}$$

The value p_{ijk} is the probability that $\tilde{a}_{ij}$ will realize the k th value in the set above, that is,

$$\Pr\{\tilde{a}_{ij} = a_{ijk}\} = p_{ijk}.$$

Converting from $\tilde{a}_{ij}$ to ξ_{ij} :

$$\begin{aligned} a_{ijk} &= a_{ij} + \xi_{ijk} \hat{a}_{ij} \\ \xi_{ij} &\in \{\xi_{ij1}, \xi_{ij2}, \dots, \xi_{ijk}, \dots, \xi_{ijm_{ij}}\} \\ \Pr\{\xi_{ij} = \xi_{ijk}\} &= \Pr\{\tilde{a}_{ij} = a_{ijk}\} = p_{ijk}. \end{aligned}$$

The convention of keeping $\xi \in [-1, 1]$ can be maintained by setting $\hat{a}_{ij}$ as follows:

$$\hat{a}_{ij} = \max_k |a_{ijk} - a_{ij}|,$$

where the nominal value a_{ij} can be set to the expected value of $\tilde{a}_{ij}$:

$$a_{ij} = \sum_k p_{ijk} a_{ijk}.$$

Using this notation, we can define the moment generating function $M_{ij}(\theta)$ of ξ_{ij} as

$$M_{ij}(\theta) = \sum_{k=1}^{m_{ij}} p_{ijk} \exp(\xi_{ijk} \theta) \quad (16)$$

and use (16) in bound (KBP2). However, the exact probability distribution for $\sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^*$ is available. For example, with $|J_i^*| = 3$ and $m_{ij} = 2$, $\forall j \in J_i^*$, and given that $\sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^*$ equals some value y only when, say, $\xi_i = (\xi_{i12}, \xi_{i21}, \xi_{i31})$, then

$$\Pr\left\{\sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* = y\right\} = p_{i12} p_{i21} p_{i31}$$

Thus, by defining the following relation:

$$S_{ij} := \{(\xi_{ijk}, p_{ijk}), k = 1, 2, \dots, m_{ij}\},$$

we can denote the exact *a posteriori* probability, given that constraint i only contains uncertain parameters subject to discrete distributions, as follows:

$$\Pr \left\{ \sum_j a_{ij} x_j^* + \sum_{j \in J_i^*} \xi_{ij} \hat{a}_{ij} x_j^* > b_i \right\} = \sum_{(\xi_{i1k}, p_{i1k}) \in S_{i1}} \cdots \sum_{(\xi_{in k}, p_{in k}) \in S_{in}} (\psi'_i) H'(\omega'_i - h_i(x^*)) \quad (\text{GMF13})$$

where

$$\begin{aligned} j &= 1, 2, \dots, n \\ n &:= |J_i^*| \\ \psi'_i &:= \prod_{j=1}^n p_{ijk} \\ \omega'_i &:= \sum_{j=1}^n \xi_{ijk} \hat{a}_{ij} x_j^* \\ H'(x) &:= \begin{cases} 1 & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

Since $h_i(x^*)$ is expected to be nonnegative, not every possible addend in (GMF13) needs to be considered in an efficient implementation (see Section 4).

3.4. Gamma, chi-squared, Erlang, and exponential distributions

Consider the case where uncertain parameters $\tilde{a}_{ij}$, $\forall j \in J_i^*$, are modeled with gamma distributions with shape parameters $\alpha_{ij} > 0$ and scale parameters $\beta_{ij} > 0$:

$$\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij}) \quad \forall j \in J_i^*.$$

Gamma distributions are asymmetric and supported on $(0, \infty)$. The probability density function and cumulative distribution function of $\tilde{a}_{ij}$ are

$$\begin{aligned} f(x) &= \frac{x^{\alpha_{ij}-1} e^{-x/\beta_{ij}}}{\beta_{ij}^{\alpha_{ij}} \Gamma(\alpha_{ij})} \\ F(x) &= \frac{\gamma(\alpha_{ij}, x/\beta_{ij})}{\Gamma(\alpha_{ij})} \end{aligned}$$

where $\Gamma(z)$ is the gamma function and $\gamma(z, x)$ is the lower incomplete gamma function:

$$\begin{aligned} \Gamma(z) &= \int_0^\infty t^{z-1} e^{-t} dt \\ \gamma(z, x) &= \int_0^x t^{z-1} e^{-t} dt. \end{aligned}$$

We also provide the upper incomplete gamma function $\Gamma(z, x)$, which will be used subsequently:

$$\Gamma(z, x) = \int_x^\infty t^{z-1} e^{-t} dt.$$

A gamma-distributed variable with parameters α_{ij} and β_{ij} as described above has an expected value and variance of $\alpha_{ij}\beta_{ij}$ and $\alpha_{ij}\beta_{ij}^2$, respectively. Alternatively, for a gamma-distributed variable with expected value μ_{ij} and standard deviation σ_{ij} , parameter $\alpha_{ij} = \mu_{ij}^2/\sigma_{ij}^2$ and parameter $\beta_{ij} = \sigma_{ij}^2/\mu_{ij}$. We caution the reader to pay special attention to the parameterization used to define the gamma distribution, as the notation throughout the literature is highly inconsistent. Here we adopt the notation of Moschopoulos (1985), whose work is central to this section.

The results of this section also apply to chi-squared, Erlang, and exponential distributions, as they are special cases of the gamma distribution. If $\tilde{a}_{ij}$ is subject to a chi-squared distribution with r_{ij} degrees of freedom, its probability density function is

$$f(x) = \frac{x^{r_{ij}/2-1} e^{-x/2}}{2^{r_{ij}/2} \Gamma(r_{ij}/2)}.$$

Thus, $\tilde{a}_{ij} \sim \text{Gamma}(r_{ij}/2, 2)$.

If $\tilde{a}_{ij}$ is subject to an Erlang distribution with (integer) shape parameter k_{ij} and rate parameter λ_{ij} , its probability density function is

$$f(x) = \frac{\lambda_{ij}^{k_{ij}} x^{k_{ij}-1} e^{-\lambda_{ij}x}}{(k_{ij}-1)!} = \frac{\lambda_{ij}^{k_{ij}} x^{k_{ij}-1} e^{-\lambda_{ij}x}}{\Gamma(k_{ij})}.$$

Thus, $\tilde{a}_{ij} \sim \text{Gamma}(k_{ij}, 1/\lambda_{ij})$.

If $\tilde{a}_{ij}$ is subject to an exponential distribution with rate parameter λ_{ij} , its probability density function is

$$f(x) = \lambda_{ij} e^{-\lambda_{ij}x}.$$

Thus, $\tilde{a}_{ij} \sim \text{Gamma}(1, 1/\lambda_{ij})$.

First, it should be noted that the use of gamma distributions is amenable to an *a priori* robust optimization setting by using the *a priori* bounds proposed in Part II (Guzman et al., 2016b); see Appendix B for a detailed treatment of this topic. To calculate *a posteriori* bounds on the probability of constraint violation with arbitrary proximity to the exact probability, the results of Moschopoulos (1985) are only able to be applied to a robust solution with either $x_j^* > 0$, $\forall j \in J_i^*$, or $x_j^* < 0$, $\forall j \in J_i^*$. It is likely that variables with gamma-distributed uncertain coefficients, supported on $(0, \infty)$, are constrained to be nonnegative and thus the former situation ($x_j^* > 0$) would apply. If this case occurs in the objective function of a maximization problem, the corresponding robust constraint after reformulation will have uncertain parameters with coefficients

$x'_j = -x_j < 0$ and thus the latter situation is likely as well. The following propositions derive *a posteriori* probabilistic bounds for these two cases and utilize the following definitions:

$$\begin{aligned}
\beta_i^L &:= \min_{j \in J_i^*} \beta_{ij} |x_j^*| \\
C_i &:= \prod_{j \in J_i^*} \left(\frac{\beta_i^L}{\beta_{ij} |x_j^*|} \right)^{\alpha_{ij}} \\
\nu_k &:= \sum_{j \in J_i^*} \alpha_{ij} \left(1 - \frac{\beta_i^L}{\beta_{ij} |x_j^*|} \right)^k \\
\delta_0 &:= 1, \quad \delta_n := \frac{1}{n} \sum_{k=1}^n \nu_k \delta_{n-k}, \quad n = 1, 2, \dots \\
\rho_i &:= \sum_{j \in J_i^*} \alpha_{ij}.
\end{aligned} \tag{17}$$

Proposition 1. *If $\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ and $x_j^* > 0, \forall J_i^*$, then for integer $N \geq 0$,*

$$\begin{aligned}
&\Pr \left\{ \sum_{j \notin J_i^*} a_{ij} x_j^* + \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > b_i \right\} \\
&\leq 1 - C_i \sum_{n=0}^N \delta_n \frac{\gamma(\rho_i + n, \tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \quad (\text{GMF14}^+)
\end{aligned}$$

where the inequality approaches an equality as $N \rightarrow \infty$.

Proposition 2. *If $\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ and $x_j^* < 0, \forall J_i^*$, then for integer $N \geq 0$,*

$$\begin{aligned}
&\Pr \left\{ \sum_{j \notin J_i^*} a_{ij} x_j^* + \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > b_i \right\} \\
&\leq 1 - C_i \sum_{n=0}^N \delta_n \frac{\Gamma(\rho_i + n, -\tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \quad (\text{GMF14}^-)
\end{aligned}$$

where the inequality approaches an equality as $N \rightarrow \infty$.

We recommend implementations of these bounds to include C_i with the summation addends when its value is low to pair with possibly large values of δ_n . Proofs for these bounds are provided in Appendix C.

4. Implementation

The ability to use (GMF12) or (GMF13) in an *a posteriori* setting can provide stark reductions in conservatism due to the fact that they calculate exact *a posteriori* probabilities. However, they pose significant implementation challenges. Each summand of (GMF12) is a function of vector s ,

$$s = (s_1, s_2, \dots, s_n)^\top,$$

with $s_j \in \{-1, 1\}$ and $n := |J_i^*|$, yielding 2^n summands. Each summand of (GMF13) is a function of some particular realization ξ'_i of vector ξ_i , that is,

$$\xi'_i = (\xi_{i1k_{i1}}, \xi_{i2k_{i2}}, \dots, \xi_{ink_{in}}),$$

with $k_{ij} \in \{1, \dots, m_{ij}\}$, yielding $\prod_{j=1}^n m_{ij}$ possible summands. For example, 10 uncertain parameters, each of which can realize 3 possible values, will yield 1000 possible summands with (GMF13).

Because of the very similar nature of (GMF12) and (GMF13), their implementations can follow the same algorithms. Inclusion of summands in the probability calculation depends on the comparison of $h_i(x^*)$ with a particular value (ω_i and ω'_i , respectively). This value will be generically referred to as ω_i in the remainder of this section. We define ragged matrix T_i , which is comprised of n row vectors T_{ij} , where

$$T_{ij} := (t_{ij1}, t_{ij2}, \dots, t_{ijk}, \dots, t_{ijm_{ij}}).$$

For (GMF12) (uniform distributions), t_{ijk} is equal to $\hat{a}_{ij}x_j^*$ or $-\hat{a}_{ij}x_j^*$; thus, m_{ij} always equals 2 and $T_i \in \mathbb{R}^{n \times 2}$. For (GMF13) (discrete distributions), $t_{ijk} = \xi_{ijk}\hat{a}_{ij}x_j^*$ and m_{ij} is the number of possible realizations ξ_{ijk} of ξ_{ij} ; if, say, $m_{ij} = \bar{m}_i$, $j = 1, \dots, n$, then we could denote $T_i \in \mathbb{R}^{n \times \bar{m}_i}$. Thus, each possible value of ω_i is a sum of n values, one drawn from each row of T_i . For each element t_{ijk} of every row T_{ij} in T_i , there is a corresponding element which also contributes to the probability calculation as a multiplicand of a product (ψ_i and ψ'_i , respectively). The product will be generically referred to as ψ_i in the remainder of this section, and we define ragged matrix V_i with row vectors V_{ij} and elements v_{ijk} which has a one-to-one correspondence to T_i . For (GMF12), v_{ijk} is equal to 1 or -1 (corresponding to the possible values of s_j), while for (GMF13), $v_{ijk} = p_{ijk}$. A brute-force implementation could traverse the enumeration tree via recursion and, at a tree leaf, (i) compare that leaf's ω_i with $h_i(x^*)$, and (ii) if required, add that leaf's summand (which is a function of ω_i and ψ_i) to the calculation of the probability. The brute-force approach is summarized in Algorithm 1.

Algorithm 1 Brute-force enumeration algorithm for (GMF12) or (GMF13)

```

1: function CALCULATEPROBABILITY_BRUTEFORCE( $d, T_i, V_i, h_i, \chi_i$ )
2:   input:
3:     (global var.) distribution case  $d \in \{\text{uniform, discrete}\}$ 
4:     (global var.) data  $T_i$  with  $n$  rows and elements  $t_{ijk}$ ; row  $j$  has length  $m_{ij}$ 
5:     (global var.) data  $V_i$  corresponding to and of equivalent size as  $T_i$ , with elements  $v_{ijk}$ 
6:     (global var.) value  $h_i$ 
7:     (global var.) value  $\chi_i$  (relevant only when  $d = \text{uniform}$ )
8:   output:
9:     probability prob of constraint violation
10:  if  $n = 0$  then                                      $\triangleright$  Special case for  $|J_i^*| = 0$ 
11:    if  $h_i \geq 0$  then return 0                        $\triangleright h_i \geq 0$  with any feasible  $x^*$ 
12:    else return 1
13:    end if
14:  else if  $h_i = 0$  and  $d = \text{uniform}$  then              $\triangleright$  Special case for  $h_i(x^*) = 0$ 
15:    return 0.5
16:  else
17:    return ENUMTREE_BRUTEFORCE(1,1,0)                  $\triangleright j \leftarrow 1, \psi_i \leftarrow 1, \omega_i \leftarrow 0$ 
18:  end if
19: end function

20: function ENUMTREE_BRUTEFORCE( $j, \psi_i, \omega_i$ )
21:   input:
22:     tree level  $j \in \{1, 2, \dots, n+1\}$ 
23:     running product  $\psi_i$ 
24:     running sum  $\omega_i$ 
25:   output:
26:     probability prob (from all leaves which descend from this node)
27:    $prob \leftarrow 0$ 
28:   if  $j = n+1$  then                                    $\triangleright$  This call is at a leaf
29:     if  $\omega_i > h_i$  then                                $\triangleright$  Include addend
30:       if  $d = \text{discrete}$  then  $prob \leftarrow \psi_i$ 
31:       else  $prob \leftarrow \psi_i ((\omega_i - h_i) / 2)^n / (n! \chi_i)$   $\triangleright d = \text{uniform}$ 
32:       end if
33:     end if
34:   else                                                  $\triangleright$  This call is at a node
35:     for  $k \leftarrow 1 \dots m_{ij}$  do
36:        $prob \leftarrow prob + \text{ENUMTREE\_BRUTEFORCE}(j+1, \psi_i \cdot v_{ijk}, \omega_i + t_{ijk})$ 
37:     end for
38:   end if
39:   return prob
40: end function

```

To improve the efficiency of the brute-force algorithm, branches of the enumeration tree can be pruned. It can be observed that the inclusion of addends in (GMF12) and (GMF13) hinges on whether or not the inequality $\omega_i > h_i(x^*)$ holds at the addend's corresponding leaf. For some branches, it may be apparent that the inequality will not hold at some intermediate node. One simple way to obtain and increase the applicability of a node-level pruning criteria is to first rearrange the elements of T_i . First, the elements of each row vector T_{ij} can be sorted by descending value order. Then, the row vectors T_{ij} should be sorted by descending lexicographic order, that is, by comparing the values of two vectors' first elements, followed by their second elements if the first elements are equal, etc. After these sorting processes, each row of T_i will be in descending value order and the rows of T_i will be in descending lexicographic order; the first column of T_i will then be in descending value order. Ragged matrix V_i should be also rearranged to retain the original element-wise correspondence with the elements of T_i . Thus, the largest value of ω_i can be calculated as the sum of all T_{ij1} , $j = 1, \dots, n$. Furthermore, at level j' in the enumeration tree, the maximum possible value of ω_i given the inherited running sum ω'_i could be calculated as $\omega'_i + \sum_{j \geq j'} T_{ij1}$; if this value is not greater than $h_i(x^*)$, then that node and all of its descendants can be ignored. Finally, if a given node is ignored in this manner, then all of its subsequent sibling nodes (with lower corresponding values of T_{ijk}) can be ignored. These ideas can be succinctly incorporated into the brute-force algorithm; the result is shown in Algorithm 2.

Algorithm 2 Pruning algorithm for (GMF12) or (GMF13)

```
1: function CALCULATEPROBABILITY_PRUNING( $d, T_i, V_i, h_i, \chi_i$ )
2:   input:
3:     (global var.) distribution case  $d \in \{\text{uniform, discrete}\}$ 
4:     (global var.) sorted data  $T_i$  with  $n$  rows and elements  $t_{ijk}$ ; row  $j$  has length  $m_{ij}$ 
5:     ► The rows of  $T_i$  are sorted in descending lexicographic order.
6:     ► The elements of each row are sorted in descending value order.
7:     (global var.) data  $V_i$  corresponding to and of equivalent size as  $T_i$ , with elements  $v_{ijk}$ 
8:     (global var.) value  $h_i$ 
9:     (global var.) value  $\chi_i$  (relevant only when  $d = \text{uniform}$ )
10:  output:
11:    probability prob of constraint violation
12:  global  $r \leftarrow 0 \in \mathbb{R}^n$  ▷  $r$  holds the maximum remaining partial sum to  $\omega_i$  at each level
13:  for  $j \leftarrow 1 \dots (n-1)$  do
14:     $r_{n-j} \leftarrow r_{n-j+1} + t_{i(n-j+1)1}$ 
15:  end for
16:  if  $n = 0$  then ▷ Special case for  $|J_i^*| = 0$ 
17:    if  $h_i \geq 0$  then return 0 ▷  $h_i \geq 0$  with any feasible  $x^*$ 
18:    else return 1
19:    end if
20:  else if  $h_i = 0$  and  $d = \text{uniform}$  then ▷ Special case for  $h_i(x^*) = 0$ 
21:    return 0.5
22:  else
23:    return ENUMTREE_PRUNING(1,1,0) ▷  $j \leftarrow 1, \psi_i \leftarrow 1, \omega_i \leftarrow 0$ 
24:  end if
25: end function

26: function ENUMTREE_PRUNING( $j, \psi_i, \omega_i$ )
27:  input:
28:    tree level  $j \in \{1, 2, \dots, n+1\}$ 
29:    running product  $\psi_i$ 
30:    running sum  $\omega_i$ 
31:  output:
32:    probability prob (from all leaves which descend from this node)
33:   $prob \leftarrow 0$ 
34:  if  $j = n+1$  then ▷ This call is at a leaf
35:    ► Here,  $\omega_i$  is always greater than  $h_i$ 
36:    if  $d = \text{discrete}$  then  $prob \leftarrow \psi_i$ 
37:    else  $prob \leftarrow \psi_i ((\omega_i - h_i) / 2)^n / (n! \chi_i)$  ▷  $d = \text{uniform}$ 
38:    end if
39:  else ▷ This call is at a node
40:    for  $k \leftarrow 1 \dots m_{ij}$  do
41:      if  $\omega_i + t_{ijk} + r_j > h_i$  then ▷ Check if child node should be entered
42:         $prob \leftarrow prob + \text{ENUMTREE\_PRUNING}(j+1, \psi_i \cdot v_{ijk}, \omega_i + t_{ijk})$ 
43:      else
44:        break ▷ Bypass proposed and subsequent children nodes
45:      end if
46:    end for
47:  end if
48:  return  $prob$ 
49: end function
```

We evaluated the efficiency of both the brute-force algorithm and the algorithm with node pruning by (i) implementing each algorithm in C++ and (ii) generating random test problems with different numbers of uncertain parameters n for (GMF12) and (GMF13). For (GMF12), 100 test problems were generated for values of n ranging from 2 to 40, resulting in 3900 test problems. For a given test problem, values of $\hat{a}_{ij}x_j^*$ were drawn from a uniform random number generator over the range $[-1.5, -0.5] \cup [0.5, 1.5]$ to avoid near-zero values. The value of $h_i(x^*)$ was then drawn from a uniform random number generator over the range $[0.01, \sum_j |\hat{a}_{ij}x_j^*| - 0.01]$. For (GMF13), it was assumed that in a given test problem, each discrete random variable ξ_{ij} has the same number of values it can realize, that is, $m_{ij} = \bar{m}_i$, $j = 1, \dots, n$. One hundred test problems were generated for each combination of n , ranging from 2 to 20, with $\bar{m}_i$, ranging over even integers from 2 to 20; this resulted in 19,000 test problems. Values of $\hat{a}_{ij}x_j^*$ were generated as previously described, while values of p_{ijk} were drawn from a uniform random number generator over the range $[0, 1]$ and then scaled to ensure $\sum_k p_{ijk} = 1$. Values of ξ_{ijk} , $j \neq n$, were drawn from a uniform random number generator over the range $[-1, 1]$, ξ_{ijn} was calculated to ensure $\mathbf{E}[\xi_{ij}] = \sum_k p_{ijk}\xi_{ijk} = 0$, and then all values of ξ_{ijk} were scaled to ensure $\max_k |\xi_{ijk}| = 1$. The value of $h_i(x^*)$ was then drawn from a uniform random number generator over the range $[0.01, \sum_j \max_k \xi_{ijk}\hat{a}_{ij}x_j^* - 0.01]$.

Algorithms 1 and 2 were applied to all test problems. For each test problem and algorithm, a CPU time limit of 5 minutes was enforced. For a given case, CPU time was averaged over all test problems which finished within the CPU time limit. Table 1 compares the performance of Algorithms 1 and 2 when used with (GMF12), while Tables 2 and 3 show the performance of Algorithms 1 and 2, respectively, when used with (GMF13). While the implementations of Algorithms 1 and 2 were both highly efficient, the data supports the assertion that the simple pruning mechanism introduced in Algorithm 2 significantly reduces the number of included tree leaves. Algorithm 2 was able to complete about 93% of all test cases with (GMF12) within the time limit, versus about 82% for Algorithm 1. Likewise, Algorithm 2 was able to complete about 72% of all test cases with (GMF13) within the time limit, versus about 55% for Algorithm 1. Drastically larger numbers of uncertain parameters may yield the bounds difficult to compute within a specified time limit; in these extreme cases, alternatives include utilizing bound (KBP2) or pursuing a sampling approach to obtain an estimate of the probability of constraint violation guaranteed to some sample-size-dependent confidence value (Johnson, 2013). The latter method has been recently studied in the contexts of chance constrained and stochastic programming, and computational cost (tied to the number of required samples) scales better with increasing n when compared to the worst-case complexity of the algorithms here (Nemirovski et al., 2009; Nemirovski and Shapiro, 2007). Conversely, sampling requirements scale poorly with decreasing violation probabilities (Au and Beck, 2001; Nemirovski and Shapiro, 2007), whereas a small *a priori* probability of constraint violation typically corresponds to a large value of $h_i(x^*)$ which in turn sharply increases the efficiency of Algorithm 2 by increasing

pruning at earlier stages. Overall, the computational results support the computational viability of both (GMF12) and (GMF13) in the average case, even with the brute-force algorithm; for example, average CPU time of the brute-force algorithm with $n = 8$ and $\bar{m}_i = 12$ (yielding 4.3E+8 addends) was 1 CPU second.

Table 1: Average CPU seconds of Algorithms 1 and 2 with (GMF12)^{a,b}

n	Avg. CPU seconds (% completed)		n	Average CPU seconds (% completed)	
	Alg. 1	Alg. 2		Alg. 1	Alg. 2
2	0 (100%)	0 (100%)	21	0 (100%)	0 (100%)
3	0 (100%)	0 (100%)	22	0 (100%)	0 (100%)
4	0 (100%)	0 (100%)	23	0 (100%)	0 (100%)
5	0 (100%)	0 (100%)	24	0 (100%)	0 (100%)
6	0 (100%)	0 (100%)	25	1 (100%)	0 (100%)
7	0 (100%)	0 (100%)	26	2 (100%)	1 (100%)
8	0 (100%)	0 (100%)	27	4 (100%)	1 (100%)
9	0 (100%)	0 (100%)	28	9 (100%)	4 (100%)
10	0 (100%)	0 (100%)	29	15 (100%)	6 (100%)
11	0 (100%)	0 (100%)	30	26 (100%)	10 (100%)
12	0 (100%)	0 (100%)	31	54 (100%)	21 (100%)
13	0 (100%)	0 (100%)	32	108 (98%)	49 (99%)
14	0 (100%)	0 (100%)	33	171 (85%)	45 (91%)
15	0 (100%)	0 (100%)	34	281 (6%)	45 (76%)
16	0 (100%)	0 (100%)	35	- (0%)	25 (74%)
17	0 (100%)	0 (100%)	36	- (0%)	32 (74%)
18	0 (100%)	0 (100%)	37	- (0%)	30 (59%)
19	0 (100%)	0 (100%)	38	- (0%)	17 (59%)
20	0 (100%)	0 (100%)	39	- (0%)	28 (57%)
			40	- (0%)	50 (56%)

^a n is the number of uncertain parameters.

^b CPU time was averaged over the number of test problems, indicated as a percent in parenthesis, which finished at or within 300 CPU seconds.

Table 2: Average CPU seconds of Algorithm 1 with (GMF13)^{a,b}

n	Average CPU seconds and (% completed)									
	$\overline{m}_i = 2$	$\overline{m}_i = 4$	$\overline{m}_i = 6$	$\overline{m}_i = 8$	$\overline{m}_i = 10$	$\overline{m}_i = 12$	$\overline{m}_i = 14$	$\overline{m}_i = 16$	$\overline{m}_i = 18$	$\overline{m}_i = 20$
2	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
3	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
4	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
5	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
6	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
7	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	1 (100%)	2 (100%)	5 (100%)
8	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	1 (100%)	4 (100%)	13 (100%)	32 (100%)	73 (100%)
9	0 (100%)	0 (100%)	0 (100%)	1 (100%)	5 (100%)	26 (100%)	101 (100%)	– (0%)	– (0%)	– (0%)
10	0 (100%)	0 (100%)	0 (100%)	5 (100%)	45 (100%)	290 (73%)	– (0%)	– (0%)	– (0%)	– (0%)
11	0 (100%)	0 (100%)	2 (100%)	38 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
12	0 (100%)	0 (100%)	10 (100%)	283 (71%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
13	0 (100%)	0 (100%)	61 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
14	0 (100%)	1 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
15	0 (100%)	6 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
16	0 (100%)	22 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
17	0 (100%)	66 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
18	0 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
19	0 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)
20	0 (100%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)	– (0%)

^a n is the number of uncertain parameters; $\overline{m}_i$ is the number of discrete realizations per parameter.^b CPU time was averaged over the number of test problems, indicated as a percent in parenthesis, which finished at or within 300 CPU seconds.

Table 3: Average CPU seconds of Algorithm 2 with (GMF13)^{a,b}

n	Average CPU seconds and (% completed)									
	$\overline{m}_i = 2$	$\overline{m}_i = 4$	$\overline{m}_i = 6$	$\overline{m}_i = 8$	$\overline{m}_i = 10$	$\overline{m}_i = 12$	$\overline{m}_i = 14$	$\overline{m}_i = 16$	$\overline{m}_i = 18$	$\overline{m}_i = 20$
2	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
3	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
4	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
5	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
6	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)
7	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	1 (100%)
8	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	0 (100%)	1 (100%)	4 (100%)	5 (100%)	17 (100%)
9	0 (100%)	0 (100%)	0 (100%)	0 (100%)	1 (100%)	2 (100%)	11 (100%)	50 (100%)	44 (86%)	38 (70%)
10	0 (100%)	0 (100%)	0 (100%)	1 (100%)	5 (100%)	38 (100%)	40 (74%)	33 (74%)	54 (57%)	40 (49%)
11	0 (100%)	0 (100%)	0 (100%)	4 (100%)	64 (99%)	42 (72%)	45 (60%)	31 (60%)	26 (43%)	23 (46%)
12	0 (100%)	0 (100%)	2 (100%)	47 (100%)	35 (75%)	45 (42%)	28 (49%)	39 (33%)	23 (44%)	36 (29%)
13	0 (100%)	0 (100%)	8 (100%)	30 (69%)	34 (52%)	20 (45%)	25 (37%)	31 (45%)	6 (32%)	37 (31%)
14	0 (100%)	0 (100%)	45 (98%)	40 (58%)	27 (34%)	29 (28%)	21 (32%)	39 (24%)	22 (27%)	19 (25%)
15	0 (100%)	1 (100%)	41 (76%)	19 (49%)	30 (27%)	21 (28%)	14 (20%)	26 (13%)	22 (19%)	29 (18%)
16	0 (100%)	4 (100%)	35 (64%)	34 (40%)	26 (32%)	20 (22%)	19 (22%)	34 (21%)	41 (15%)	59 (5%)
17	0 (100%)	11 (100%)	25 (51%)	24 (40%)	30 (23%)	12 (20%)	23 (17%)	6 (14%)	25 (12%)	36 (12%)
18	0 (100%)	40 (100%)	32 (46%)	16 (27%)	33 (25%)	13 (17%)	43 (13%)	24 (18%)	17 (13%)	13 (8%)
19	0 (100%)	31 (77%)	32 (35%)	43 (27%)	13 (12%)	20 (18%)	17 (14%)	55 (12%)	24 (14%)	40 (14%)
20	0 (100%)	35 (65%)	24 (41%)	52 (21%)	13 (12%)	16 (19%)	15 (13%)	36 (5%)	18 (10%)	21 (9%)

^a n is the number of uncertain parameters; $\overline{m}_i$ is the number of discrete realizations per parameter.^b CPU time was averaged over the number of test problems, indicated as a percent in parenthesis, which finished at or within 300 CPU seconds.

5. Comparisons of *a posteriori* bounds

As in Parts I and II of this work (Guzman et al., 2016a,b), a multiperiod production planning problem expanded from Ravindran (2008) and Li et al. (2012b) will be used to demonstrate the improvements that the newly derived *a posteriori* expressions provide to the robust solutions. The LP formulation of the problem can be seen in model (18).

$$\begin{aligned}
& \max_{x,y,z} && \sum_j (P_j z_j - C_j x_j - V_j y_j) \\
& \text{s.t.} && \sum_j (C_j x_j + V_j y_j) \leq B \\
& && I + x_1 - (y_1 + z_1) = 0 \\
& && y_{j-1} + x_j - (y_j + z_j) = 0 && \forall j = 2, \dots, 12 \\
& && y_{12} = F \\
& && x_j \leq U_j && \forall j \\
& && z_j \leq D_j && \forall j \\
& && x_j, y_j, z_j \geq 0 && \forall j.
\end{aligned} \tag{18}$$

This model represents the maximization of profit from a specific product over 12 months. Certain parameters, such as the selling prices (P_j), production costs (C_j), storage costs (V_j), demands (D_j), and production capacities (U_j) are defined for each monthly period j . The production amounts (x_j), amounts of product in storage (y_j), and amounts of product sold in tons (z_j) are subject to a variety of constraints. The total cost of production for the entire year is limited by the budget $B = \$800,000$, and it is stipulated that the initial (I) and final amounts (F) of product in storage must equal 500 tons. The parameters that vary on a monthly basis are provided in Table 4.

Table 4: Monthly Planning Data for the LP Production Planning Problem

Period	Selling Price (\$/ton)	Production Cost (\$/ton)	Storage Cost (\$/ton)	Production Capacity (tons)	Demand (tons)
j	P_j	C_j	V_j	U_j	D_j
1	180	20	2	1500	1100
2	180	25	2	2000	1500
3	250	30	2	2200	1800
4	270	40	2	3000	1600
5	300	50	2	2700	2300
6	320	60	2	2500	2500
7	300	80	2	1900	2300
8	270	55	2	1500	1600
9	250	45	2	1300	1800
10	180	30	2	1100	1500
11	180	30	2	1200	1100
12	160	25	2	1400	1200

To demonstrate the power of robust optimization, uncertainty is considered in the monthly selling prices; this means that there are 12 uncertain parameters ($|J_i| = 12$). To incorporate uncertainty in the objective function, the following transformation occurs as in Guzman et al. (2016a,b) with the variable ζ maximized and the following constraint added to the model:

$$\zeta - \sum_j \tilde{P}_j z_j + \sum_j (C_j x_j + V_j y_j) \leq 0. \quad (19)$$

The uncertainty selling prices are represented by $\tilde{P}_j = P_j + \xi_j \hat{P}_j$, and the perturbations are set to 20% of the nominal values. That is, $\hat{P}_j = 0.20 \cdot P_j$. Solving the multiperiod planning problem without any uncertainty results in a profit of \$4,207,706 over the 12 month period. However, at the worst case values, the guaranteed profit decreases to a value of \$3,206,165.

5.1. Price parameters with uniform distributions

For the examples in this section, it is assumed that the parameters are subject to a uniform distribution. As shown in Guzman et al. (2016a,b), the use of *a priori* and *a posteriori* bounds can provide robust solutions that are much better than this worst case value. Using the best bounds from Guzman et al. (2016a,b) for comparison, the increased tightness of the proposed exact *a posteriori* expression (GMF12) for uniform distributions will be demonstrated.

To demonstrate the tightness of the proposed exact *a posteriori* bounds, the interval + ellipsoidal uncertainty set was used to solve the multiperiod planning problem at different probabilities of constraint violation, setting Ω with a *a priori* bound (GMF7'). Figure 1 shows a comparison of the probabilities obtained using *a posteriori* bounds with specific solutions gained at *a priori* probabilities with bound (GMF7'). First, it is seen that while any *a posteriori* bound will

provide probabilities that are as tight as or tighter than their specific *a priori* analogue, a *posteriori* bound (KBP2) is only slightly better than its *a priori* analogue (GMF7'). This reiterates the very strong performance of (GMF7') as an *a priori* bound. In fact, (GMF7') is even better than the distribution-independent *a posteriori* bound (GMF5) at low probabilities, and (PKL) at all probabilities of constraint violation. Finally, the exact probabilities of constraint violation of (GMF12) are dominant at all probabilities of constraint violation. Figure 2 demonstrates the impact of this improvement on the profit of the multiperiod planning problem. While the *a priori* (GMF7') essentially provides an *a posteriori* level of profit in the context of (KBP2), (GMF12) improves the profits even further.

The performance of the best *a priori* bound relative to *a posteriori* bounds is not as strong for the interval + polyhedral set, as seen in Figures 3 and 4. Bound (GMF6) is used as the *a priori* reference; note that when $\xi_{ij}, \forall j \in J_i$ are subject to identical distributions, bounds (GMF6), (GMF7), and (KBP1) provide the same Γ_i values for the interval + polyhedral set. Clearly, a *posteriori* bound (KBP2) is tighter than *a priori* bound (GMF6) for the interval + polyhedral set, while the exact (GMF12) is even tighter. The greater *a posteriori* increase in tightness for the interval + polyhedral set demonstrates the need for an algorithm, such as that provided by Li and Floudas (2014), which will greatly improve the quality of solutions using *a posteriori* bounds.

5.2. Price parameters with discrete distributions

The performance of bound (GMF13) is also demonstrated using the multiperiod planning problem. Rather than assuming a continuous uniform distribution for the price parameters, here it is assumed that the distribution is discretized with five values for the random variables ξ_{ij} . The discretization is assumed to be symmetric, and each discretization is assumed to have the same probability of occurrence. The random variable is defined such that $\xi_{ij} \in \{-1, -0.5, 0, 0.5, 1\}$, with each value given a 20% probability of realization. This distribution is then used with the same *a priori* and *a posteriori* bounds as in the previous section, albeit with (GMF13) instead of (GMF12) due to the attribution of discrete distributions.

An interval + ellipsoidal uncertainty set was again utilized, with the results shown in Figures 5 and 6. These figures are very similar in form to Figures 1 and 2, with (GMF13) outperforming each of the *a priori* and *a posteriori* bounds in terms of overall tightness and profit. Once again, even with a discrete distribution used rather than a uniform distribution, a *a priori* bound (GMF7') still performs nearly identically to a *posteriori* bound (KBP2); this demonstrates that the strong performance of (GMF7') comes from the nature of the bound itself, rather than the distribution used. However, the drastic improvement of (GMF13) relative to all of the other bounds reiterates the improvements which can be gained through the use of an iterative algorithm to produce high quality *a posteriori* results. Overall, if all parameters can be modeled with discrete distributions, large improvements in performance may be made over the existing *a posteriori* bounds.

Figure 1: *A posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + ellipsoidal uncertainty set versus *a priori* probability of constraint violation, when all 12 uncertain parameters are subject to a uniform distribution; ϵ_i^{prio} was set by selecting Δ_i with bound (GMF7').

Figure 2: Profit versus *a posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + ellipsoidal uncertainty set when all 12 uncertain parameters are subject to a uniform distribution; profit versus ϵ_t^{prio} as obtained by *a priori* bound (GMF7') included for comparison.

Figure 3: *A posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + polyhedral uncertainty set versus *a priori* probability of constraint violation, when all 12 uncertain parameters are subject to a uniform distribution; ϵ_i^{prio} was set by selecting Δ_i with bound (GMF6).

Figure 4: Profit versus *a posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + polyhedral uncertainty set when all 12 uncertain parameters are subject to a uniform distribution; profit versus ϵ_i^{prio} as obtained by *a priori* bound (GMF6) included for comparison.

Figure 5: *A posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + ellipsoidal uncertainty set versus *a priori* probability of constraint violation, when all 12 uncertain parameters are subject to a symmetric discrete distributions with five discretizations of equal probability of realization; ϵ_i^{prio} was set by selecting Δ_i with bound (GMF7').

Figure 6: Profit versus *a posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the interval + ellipsoidal uncertainty set when all 12 uncertain parameters are subject to symmetric discrete distributions with five discretizations of equal probability of realization; profit versus ϵ_i^{prio} as obtained by *a priori* bound (GMF7') included for comparison.

Figure 7: *A posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the ellipsoidal uncertainty set versus *a priori* probability of constraint violation, when all 12 uncertain parameters are subject to normal distributions; ϵ_i^{prio} was set by selecting Δ_i with bound (GMF7').

5.3. Price parameters with normal distributions

To demonstrate the effectiveness of using (GMF11) for normally distributed parameters, it is now assumed that all of the uncertain price parameters are subject to a normal distribution about their nominal price. This approximation has limitations — a parameter representing the selling price should be positive, and thus is not truly unbounded — but it clearly shows the improvement of (GMF11) relative to the existing methods for *a posteriori* probabilities of constraint violation. For these calculations, the same values of P_j and $\hat{P}_j$ are used as in the previous sections; because this is with a normal distribution, this indicates that P_j is the mean price and $\hat{P}_j$ is the standard deviation of the selling price. The results can be seen in Figures 7 and 8.

An ellipsoidal uncertainty was utilized for this case, as the normal distribution is unbounded and thus an interval + ellipsoidal uncertainty set is not appropriate for the uncertain price parameters. Figure 7 shows the improvement on the basis of probability of constraint violation; the values of the *a posteriori* (KBP2) with the normal distribution gives essentially the same values as the *a priori* bound (GMF7'). However, the exact *a posteriori* probability of

Figure 8: Profit versus *a posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the ellipsoidal uncertainty set when all 12 uncertain parameters are subject to normal distributions; profit versus ϵ_i^{prio} as obtained by *a priori* bound (GMF7') included for comparison.

constraint violation gained using (GMF11) is a substantial improvement over the *a priori* (GMF7'). This improvement is evident from the objective function values as well. As seen in Figure 8, a considerable improvement in profit is seen using the exact probability (GMF11) relative to (GMF7') and (KBP2). At a 10% probability of constraint violation, for instance, there is a 7.9% increase in profit when using the (GMF11) rather than (KBP2). This significant reduction in conservatism simply due to better probabilistic characterization of the optimal solutions, and thus bound (GMF11) should be used *a posteriori* for any constraints that contain only uncertain parameters subject to normal distributions.

5.4. Price parameters with gamma distributions

Finally, the exact *a posteriori* probability of constraint violation for the gamma distribution is also demonstrated through the multiperiod planning problem. In this example, the selling prices are assumed to be subject to a gamma distribution with P_j representing the average price and $\hat{P}_j$ representing the standard deviation for the j th uncertain parameter. Because the gamma distribution is unbounded on one side, the ellipsoidal uncertainty set was utilized rather than the interval + ellipsoidal uncertainty set. Reformulating the objective function into an uncertain constraint results in coefficients $(-z_j)$ of uncertain parameters $\tilde{P}_j$; thus, bound (GMF14⁻) applies. The results can be seen in Figures 9 and 10.

Figure 9 shows the relative improvement of exact *a posteriori* probability (GMF14⁻) relative to asymmetric *a priori* bound (GMF10) and *a posteriori* bound (KBP2). As expected, the exact probability of constraint violation found using (GMF14⁻) is much tighter the other two probabilistic bounds used for the asymmetric gamma distribution. This has a tremendous impact on the overall profit found, as seen in Figure 9. In fact, at a 10% probability of constraint violation, the profit improves by about 7.0% by using the exact probability (GMF14⁻) rather than the bound (KBP2). Thus, there is a significant advantage to using (GMF14⁺) or (GMF14⁻) if all uncertain parameters in a constraint are subject to gamma distributions.

6. Conclusions

Robust optimization is a powerful tool for optimization under uncertainty, whose impact and performance is greatly improved as bounds on the probability of constraint violation are tightened. This three-part series has provided new and tighter bounds on probabilities of constraint violation for constraints with parameters whose distributions are unknown or known. In this paper, exact *a posteriori* probabilistic guarantees on constraint violation are provided for parameters subject to normal, uniform, or discrete probability distributions. These guarantees outperform existing *a posteriori* bounds, and uniform distributions can be valid approximations for many cases where parameter distributions are

Figure 9: *A posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the ellipsoidal uncertainty set versus *a priori* probability of constraint violation, when all 12 uncertain parameters are subject to gamma distributions; ϵ_i^{prio} was set by selecting Δ_i with bound (GMF10).

Figure 10: Profit versus *a posteriori* probability of constraint violation of each solution obtained from the multiperiod planning problem with the ellipsoidal uncertainty set when all 12 uncertain parameters are subject to gamma distributions; profit versus ϵ_i^{prio} as obtained by *a priori* bound (GMF10) included for comparison.

unknown. The computational implementation of these new bounds was discussed for ease of future use with robust optimization, and through a computational example involving a multiperiod planning problem, the improved impact on the objective function was demonstrated. These fundamental improvements in probabilistic bounds, throughout all three parts of this series of papers, will provide increased performance and applicability in a variety of multi-scale systems engineering applications under uncertainty.

Acknowledgments

Yannis A. Guzman and Logan R. Matthews dedicate this article to the memory of Christodoulos A. Floudas. The authors gratefully acknowledge financial support from the National Science Foundation (NSF CBET-1158849). LRM gratefully acknowledges a National Defense Science and Engineering Graduate (NDSEG) fellowship.

Table 5: Best performing *a priori* probabilistic bounds on constraint violation: unknown distributions.

Source	Label	Bound	Applicable U_i	Assumptions
(a)	(GMF1) ^b	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{k \in K_i} \ln G_{ik}(\theta) \right\} \right)$	B, E, IE, P, IP	$\xi_{ij} \in [-1, 1]$, known bounds on $\mathbf{E}[\xi_{ij}]$
(a)	(GMF1') ^c	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + J_i \ln \overline{G}_i \left(\theta / \sqrt{ J_i } \right) \right\} \right)$	E, IE	$\xi_{ij} \in [-1, 1]$, known bounds on $\mathbf{E}[\xi_{ij}]$
(a)	(GMF2) ^d	$(\alpha^\alpha \beta^\beta)^{-\delta_i^{(1)}/2}$	B, E, IE, P, IP	$\xi_{ij} \in [-1, 1]$, $\mathbf{E}[\xi_{ij}] = 0$
(a)	(GMF2') ^d	$(\alpha^\alpha \beta^\beta)^{- J_i /2}$	E, IE	$\xi_{ij} \in [-1, 1]$, $\mathbf{E}[\xi_{ij}] = 0$
(a)	(GMF3) ^e	$\frac{1}{2^{\tilde{n}}} \sum_{k=\lceil \tilde{\nu} \rceil}^{\tilde{n}} \binom{\tilde{n}}{k}$	B, E, IE, P, IP	$\xi_{ij} \in [-1, 1]$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric

^a Guzman et al., 2016a.^b $G_{ij}(\theta) := \mu_{ij} \sinh(\theta) + \cosh(\theta)$.^c $\overline{G}_i(\theta) := (\max_{j \in J_i} \mu_{ij}) \sinh(\theta) + \cosh(\theta)$.^d $\alpha := 1 - \Delta_i / \delta_i^{(1)}$, $\beta := 1 + \Delta_i / \delta_i^{(1)}$.^e $\tilde{\nu} := (\Delta_i + \tilde{n})/2$, $\tilde{n} := |J_i| + 1 - \text{sgn } \Delta_i$ or $|J_i| - \text{sgn } \Delta_i$.

Table 6: *A priori* probabilistic bounds on constraint violation: known, symmetric distributions.

Source	Label	Bound	Applicable ^a U_i	Assumptions
(b)	(KBP1)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{j \in J_i} \Lambda_{ij}(\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric
(c)	(GMF6)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{k \in K_i} \Lambda_{ik}(\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric, (C ₁)
(c)	(GMF7)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \delta_i^{(1)} \max_{j \in J_i} \Lambda_{ij}(\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric
(c)	(GMF7')	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + J_i \max_{j \in J_i} \Lambda_{ij} \left(\theta / \sqrt{ J_i } \right) \right\} \right)$	E, IE	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric, (C ₂)

^a Bound (KBP1) was formulated for IP uncertainty sets. Its usage was extended to other uncertainty sets in (Li et al., 2012b).^b Kang et al., 2013.^c Guzman et al., 2016b.

Table 7: *A priori* probabilistic bounds on constraint violation: known, asymmetric distributions.

Source	Label	Bound ^a	Applicable U_i	Assumptions
(^b)	(GMF8)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{j \in J_i} \Lambda_{ij}(g_{ij}\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, g_{ij} exists
(^b)	(GMF9)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \sum_{k \in K_i} \Lambda_{ik}(g_{ik}\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, g_{ij} exists, (C ₁)
(^b)	(GMF10)	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + \delta_i^{(1)} \max_{j \in J_i} \Lambda_{ij}(g_{ij}\theta) \right\} \right)$	B, E, IE, P, IP	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, g_{ij} exists
(^b)	(GMF10')	$\exp \left(\min_{\theta > 0} \left\{ -\theta \Delta_i + J_i \max_{j \in J_i} \Lambda_{ij} \left(g_{ij}\theta / \sqrt{ J_i } \right) \right\} \right)$	E, IE	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$, g_{ij} exists, (C ₂)

^a $g_{ij} = \pm 1$, see (10).^b Guzman et al., 2016b.

Table 8: Tightest *a priori* bounds by case.

Available Information	Uncertainty Set		
	B	E, IE	P, IP
ξ_{ij} bounded, known bounds on $\mathbf{E}[\xi_{ij}]$	(GMF1)	(GMF1) or (GMF1')	(GMF1)
ξ_{ij} bounded, $\mathbf{E}[\xi_{ij}]$ known	(GMF2)	(GMF2')	(GMF2)
ξ_{ij} bounded and symmetric	(GMF3)	(GMF2') or (GMF3)	(GMF3)
bell-shaped bounded distributions (conservatively applying uniform distributions)	(GMF3) or (GMF6) = (GMF7)	(GMF3) or (GMF7')	(GMF3) or (GMF6) = (GMF7) = (KBP1)
ξ_{ij} distribution known, symmetric	(GMF3) or (GMF6) ^a	(GMF3) or (GMF6) ^a or (GMF7') ^a	(GMF3) or (GMF6) ^a = (KBP1)
ξ_{ij} distribution known, asymmetric	(GMF9) ^b	(GMF9) ^b or (GMF10') ^b	(GMF8) = (GMF9) ^b

^a If auxiliary conditions do not apply, use (GMF7) or (KBP1) instead.

^b If auxiliary conditions do not apply, use (GMF8) or (GMF10) instead.

Table 9: Best performing *a posteriori* probabilistic bounds on constraint violation.

Source	Label	Bound ^a	Assumptions
(b)	(GMF4) ^c	$\exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \ln G_{ij}(\hat{a}_{ij} x_j^* \theta) \right\} \right)$	$\xi_{ij} \in [-1, 1]$, known bounds on $\mathbf{E}[\xi_{ij}]$
(b)	(GMF5) ^{e,f}	$\min \left\{ \frac{1}{2^{n_a}} \sum_{k=\lceil \nu_a \rceil}^{n_a} \binom{n_a}{k}, \frac{1}{2^{n_b}} \sum_{k=\lceil \nu_b \rceil}^{n_b} \binom{n_b}{k} \right\}$	$\xi_{ij} \in [-1, 1]$, $\mathbf{E}[\xi_{ij}] = 0$, symmetric
(d)	(PKL)	$\exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \ln \cosh(\hat{a}_{ij} x_j^* \theta) \right\} \right)$	$\xi_{ij} \in [-1, 1]$, $\mathbf{E}[\xi_{ij}] = 0$
(g)	(KBP2)	$\exp \left(\min_{\theta > 0} \left\{ -\theta h_i(x^*) + \sum_{j \in J_i} \Lambda_{ij}(\hat{a}_{ij} x_j^* \theta) \right\} \right)$	known $M_{ij}(\theta)$, $\mathbf{E}[\xi_{ij}] = 0$
this work	(GMF11)	$\frac{1}{2} \operatorname{erfc} \left(h_i(x^*) \left(2 \sum_{j \in J_i} \hat{a}_{ij}^2 x_j^{*2} \right)^{1/2} \right)$	$\xi_{ij} \sim N(0, 1)$
this work	(GMF12) ^h	$\left(-\frac{1}{2} \right)^n \frac{1}{n! \chi_i} \sum_{s_1=\pm 1} \cdots \sum_{s_n=\pm 1} (\psi)(h_i(x^*) - \omega_i)^n H(\omega_i - h_i(x^*))$	$\xi_{ij} \sim U(-1, 1)$
this work	(GMF13) ⁱ	$\sum_{(\xi_{i1k}, p_{i1k}) \in S_{i1}} \cdots \sum_{(\xi_{ink}, p_{ink}) \in S_{in}} (\psi'_i) H'(\omega'_i - h_i(x^*))$	ξ_{ij} is discrete
this work	(GMF14 ⁺) ^j	$1 - C_i \sum_{n=0}^N \delta_n \frac{\gamma(\rho_i + n, \tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)}$	$\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$, $x_j^* > 0$, $\forall j \in J_i^*$
this work	(GMF14 ⁻) ^j	$1 - C_i \sum_{n=0}^N \delta_n \frac{\Gamma(\rho_i + n, -\tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)}$	$\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$, $x_j^* > 0$, $\forall j \in J_i^*$

^a $h_i(x^*) := b_i - \sum_j a_{ij} x_j^*$, $\tau_i(x^*) := b_i - \sum_{j \notin J_i^*} a_{ij} x_j^*$.^b Guzman et al., 2016a.^c $G_{ij}(\theta) := \mu_{ij} \sinh(\theta) + \cosh(\theta)$.^d Paschalidis et al., 2008.^e $\nu_a := (h_a + n_a)/2$, $n_a := \lceil J_i^{\delta,*} \rceil + 1 - \operatorname{sgn} h_a$ or $\lceil J_i^{\delta,*} \rceil - \operatorname{sgn} h_a$, $h_a := \Delta_i / (\max_{j \in J_i^{\delta,*}} |\delta_{ij}^*|)$, $J_i^{\delta,*} := \{j : |\delta_{ij}^*| > 0\}$.^f $\nu_b := (h_b + n_b)/2$, $n_b := \lceil J_i^* \rceil + 1 - \operatorname{sgn} h_b$ or $\lceil J_i^* \rceil - \operatorname{sgn} h_b$, $h_b := h_i(x^*) / (\max_{j \in J_i^*} \hat{a}_{ij} |x_j^*|)$, $J_i^* := \{j : \hat{a}_{ij} |x_j^*| > 0\}$.^g Kang et al., 2013.^h $j = 1, 2, \dots, n$, $n := |J_i^*|$, $\psi := \prod_{j=1}^n s_j$, $\omega_i := \sum_{j=1}^n s_j \hat{a}_{ij} x_j^*$, $H(x) := \frac{1}{2} (1 + \operatorname{sgn} x)$, $\chi_i := \prod_{j=1}^n \hat{a}_{ij} x_j^*$.ⁱ $j = 1, 2, \dots, n$, $n := |J_i^*|$, $\psi'_i := \prod_{j=1}^n p_{ijk}$, $\omega'_i := \sum_{j=1}^n \xi_{ijk} \hat{a}_{ij} x_j^*$, $H'(x) := \{1 \text{ if } x > 0, 0 \text{ otherwise}\}$, $S_{ij} := \{(\xi_{ijk}, p_{ijk}), k = 1, \dots, m_{ij}\}$.^j $\beta_i^L := \min_{j \in J_i^*} \beta_{ij} |x_j^*|$, $C_i := \prod_{j \in J_i^*} \left(\frac{\beta_i^L}{\beta_{ij} |x_j^*|} \right)^{\alpha_{ij}}$, $\nu_k := \sum_{j \in J_i^*} \alpha_{ij} \left(1 - \frac{\beta_i^L}{\beta_{ij} |x_j^*|} \right)^k$, $\delta_0 := 1$, $\delta_{n \geq 2} := \frac{1}{n} \sum_{k=1}^n \nu_k \delta_{n-k}$, $\rho_i := \sum_{j \in J_i^*} \alpha_{ij}$.

Table A.1: Definition of uncertainty sets.

U^i	Parameter ^a	Definition	$\max \Delta_i^b$
Box	Ψ_i	$U_i^\infty = \{\xi_i : \ \xi_i\ _\infty \leq \Psi_i\}$	1
Ellipsoidal	Ω_i	$U_i^2 = \{\xi_i : \ \xi_i\ _2 \leq \Omega_i\}$	$\sqrt{ J_i }$
Interval + Ellipsoidal	Ω_i	$U_i^{2 \cap \infty} = \{\xi_i : \ \xi_i\ _2 \leq \Omega_i, \ \xi_i\ _\infty \leq 1\}$	$\sqrt{ J_i }$
Polyhedral	Γ_i	$U_i^1 = \{\xi_i : \ \xi_i\ _1 \leq \Gamma_i\}$	$ J_i $
Interval + Polyhedral	Γ_i	$U_i^{1 \cap \infty} = \{\xi_i : \ \xi_i\ _1 \leq \Gamma_i, \ \xi_i\ _\infty \leq 1\}$	$ J_i $

^a Original notation of parameter used here; Δ_i used for the unspecified case.

^b Applicable when $\xi_{ij} \in [-1, 1]$, $\forall j \in J_i$. The set J_i contains the indices of uncertain parameters in constraint i .

Appendix A. Reference Tables

Table A.1 summarizes the various uncertainty sets with respect to ξ_i . Table A.2 presents the robust counterparts of uncertain constraints. Table A.3 summarizes characteristics of δ_{ij} and defines parameters $\delta_i^{(1)}$ and $\delta_i^{(2)}$, the former of which is referenced in some of the probabilistic bounds. Table A.4 presents relevant information on common probability distributions, including the new information on standardized gamma distributions for *a priori* modeling.

Table A.2: Robust counterparts of constraint i by uncertainty set type.

U_i	$\sum_j a_{ij}x_j + \max_{\xi_i \in U_i} \left\{ \sum_{j \in J_i} \xi_{ij} \hat{a}_{ij}x_j \right\} \leq b_i$
B	$\sum_j a_{ij}x_j + \Psi_i \sum_{j \in J_i} \hat{a}_{ij} x_j \leq b_i$
E	$\sum_j a_{ij}x_j + \Omega_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 x_j^2} \leq b_i$
IE	$\sum_j a_{ij}x_j + \sum_{j \in J_i} \hat{a}_{ij} x_j - z_{ij} + \Omega_i \sqrt{\sum_{j \in J_i} \hat{a}_{ij}^2 z_{ij}^2} \leq b_i$
P	(1) $\sum_j a_{ij}x_j + \Gamma_i z_i \leq b_i$ (2) $z_i \geq \hat{a}_{ij} x_j , \forall j \in J_i$
IP	(1) $\sum_j a_{ij}x_j + \sum_{j \in J_i} p_{ij} + \Gamma_i z_i \leq b_i$ (2) $z_i \geq 0$ (3) $z_i + p_{ij} \geq \hat{a}_{ij} x_j , \forall j \in J_i$ (4) $p_{ij} \geq 0, \forall j \in J_i$

 Table A.3: Associated parameters of uncertainty sets.^a

U_i	Definition of δ_i	$\ \delta_i\ _1 \leq \delta_i^{(1)}$	$\ \delta_i\ _2 \leq \delta_i^{(2)}$
B	$\delta_{ij} := \frac{\hat{a}_{ij}x_j}{\sum_j \hat{a}_{ij} x_j }$	$\sum_j \delta_{ij} \leq 1$	$\sqrt{\sum_j \delta_{ij}^2} \leq 1$
E	$\delta_{ij} := \frac{\hat{a}_{ij}x_j}{\sqrt{\sum_j \hat{a}_{ij}^2 x_j^2}}$	$\sum_j \delta_{ij} \leq \sqrt{ J_i }^c$	$\sqrt{\sum_j \delta_{ij}^2} = 1$
IE ^b	$\delta_{ij} := \frac{\hat{a}_{ij}z_{ij}}{\sqrt{\sum_j \hat{a}_{ij}^2 z_{ij}^2}}$	$\sum_j \delta_{ij} \leq \sqrt{ J_i }^c$	$\sqrt{\sum_j \delta_{ij}^2} = 1$
P	$\delta_{ij} := \frac{\hat{a}_{ij}x_j}{\max_j \hat{a}_{ij} x_j }$	$\sum_j \delta_{ij} \leq J_i $	$\sqrt{\sum_j \delta_{ij}^2} \leq \sqrt{ J_i }$
IP ^{b,d}	$\delta_{ij} := \min \left\{ 1, \frac{\hat{a}_{ij} x_j }{\hat{a}_{ir} x_r } \right\}$	$\sum_j \delta_{ij} \leq J_i $	$\sqrt{\sum_j \delta_{ij}^2} \leq \sqrt{ J_i }$

^a Note that all $j \in J_i$, and that for all uncertainty sets, $\delta_{ij} \in [-1, 1]$. Notation uses vector δ_i with elements δ_{ij} .

^b Assumes that probability distributions are bounded.

^c Follows from the Cauchy-Schwarz inequality.

^d Index r is the index of the $(\max \{1, \lceil \Gamma_i \rceil\})$ th largest $\hat{a}_{ij} |x_j|$.

Table A.4: Relevant Information on Common Probability Distributions.

Distribution ^a	$f_\xi(x)$	$M_\xi(\theta)$	$d\Lambda_\xi(\theta)/d\theta$	Rank for (C ₁)	Meets (C ₂)?
standardized gamma ^b	$\frac{(x + \sqrt{\alpha})^{\alpha-1} \alpha^{\alpha/2}}{\exp[\sqrt{\alpha}(x + \sqrt{\alpha})] \Gamma(\alpha)}$	$\exp(-\sqrt{\alpha}\theta) (1 - \theta/\sqrt{\alpha})^{-\alpha}$	$\frac{\theta}{1 - \theta/\sqrt{\alpha}}$	1	no
standard normal	$\frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right)$	$\exp\left(\frac{\theta^2}{2}\right)$	θ	2	yes
Rademacher (shifted Bernoulli)	$\begin{cases} 1/2 & \text{if } x = -1/2 \\ 1/2 & \text{if } x = 1/2 \\ 0 & \text{otherwise} \end{cases}$	$\cosh(\theta)$	$\tanh(\theta)$	3	yes
uniform ($\xi \in [-1, 1]$)	$\begin{cases} 1/2 & \text{if } x \in [-1, 1] \\ 0 & \text{otherwise} \end{cases}$	$\frac{\sinh(\theta)}{\theta}$	$\coth(\theta) - \frac{1}{\theta}$	4	yes
symmetric triangular ^c ($\xi \in [-1, 1]$, $c = 0$)	$\begin{cases} 1+x & \text{if } x \in [-1, 0] \\ 1-x & \text{if } x \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$	$\frac{2(\cosh(\theta) - 1)}{\theta^2}$	$\coth\left(\frac{\theta}{2}\right) - \frac{2}{\theta}$	5	yes
asymmetric triangular ^{c,d} ($\xi \in [a, b]$, $c \neq 0$, $\max\{-a, b\} = 1$, $a + b + c = 0$)	$\begin{cases} \frac{2(x-a)}{(b-a)(c-a)} & \text{if } x \in [a, c] \\ \frac{2(b-x)}{(b-a)(b-c)} & \text{if } x \in [c, b] \\ 0 & \text{otherwise} \end{cases}$	$\frac{2T(\theta)}{(b-a)(b-c)(c-a)\theta^2}$	$\frac{S(\theta)}{T(\theta)} + c - \frac{2}{\theta}$	6	yes
raised cosine ($\xi \in [-1, 1]$)	$\begin{cases} \frac{1}{2}(1 + \cos(\pi x)) & \text{if } x \in [-1, 1] \\ 0 & \text{otherwise} \end{cases}$	$\frac{\pi^2 \sinh(\theta)}{\theta(\pi^2 + \theta^2)}$	$\coth(\theta) - \frac{1}{\theta} - \frac{2\theta}{\pi^2 + \theta^2}$	7	yes

^a For random variable ξ , $f_\xi(x)$ is its probability density or probability mass function, $M_\xi(\theta) = \mathbf{E}[e^{\theta\xi}]$ is its moment generating function, and $\Lambda_\xi(\theta) = \ln M_\xi(\theta)$ is its cumulant generating function. Parameter g exists (see (10)) for the asymmetric distributions.

^b Transformed $\tilde{a} \sim \text{Gamma}(\alpha, \beta)$ so that expected value and variance of ξ are 0 and 1, respectively. Rank for (C₁) among standardized gamma distributions obtained by sorting by increasing α (given that $g = 1$ by Proposition 3). Domain for generating functions is $\theta \in (0, \alpha)$. See Sections 3.4 and Appendix B for details.

^c Parameter c is the mode of the triangular distribution of ξ .

^d Rank for (C₁) and satisfaction of (C₂) hold in the asymmetric case for $\Lambda_{ij}^\dagger(\theta) = \Lambda_{ij}(g\theta)$ given that $g = -\text{sgn } c$.
 $T(\theta) := (c-a)\exp(b\theta) - (b-a)\exp(c\theta) + (b-c)\exp(a\theta)$. $S(\theta) := (b-c)(c-a)(\exp(b\theta) - \exp(a\theta))$.

Appendix B. Robust optimization with gamma distributions and *a priori* bounds

Distribution-dependent *a priori* bounds (GMF8), (GMF9), (GMF10), and (GMF10') were derived in Part II (Guzman et al., 2016b) for a constraint with uncertain parameters subject to asymmetric distributions. The *a priori* bounds require that $\mathbf{E}[\xi_{ij}] = 0$ and depend on the existence of parameter g_{ij} , as shown in Section 2. The moment generating function of parameter $\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ is

$$M_{\tilde{a}_{ij}}(\theta) = (1 - \beta_{ij}\theta)^{-\alpha_{ij}} \quad \theta < 1/\beta_{ij}.$$

As noted in Section 3.4, $\tilde{a}_{ij}$ has an expected value of $\alpha_{ij}\beta_{ij}$ and a variance of $\alpha_{ij}\beta_{ij}^2$. Random variable ξ_{ij} can then be defined to have an expected value of zero and a variance of 1 by setting a_{ij} and $\hat{a}_{ij}$ to the expected value and standard deviation of $\tilde{a}_{ij}$, respectively:

$$a_{ij} := \alpha_{ij}\beta_{ij} \tag{B.1}$$

$$\hat{a}_{ij} := \sqrt{\alpha_{ij}}\beta_{ij} \tag{B.2}$$

$$\xi_{ij} = \frac{\tilde{a}_{ij} - a_{ij}}{\hat{a}_{ij}} = \frac{\tilde{a}_{ij}}{\sqrt{\alpha_{ij}}\beta_{ij}} - \sqrt{\alpha_{ij}}. \tag{B.3}$$

Given random variables X and Y , where $Y = aX + b$, their moment generating functions are related:

$$M_Y(\theta) = \exp(b\theta)M_X(a\theta).$$

Then, the moment generating function of ξ_{ij} , denoted as $M_{ij}(\theta)$, is defined as:

$$M_{ij}(\theta) = \exp(-\sqrt{\alpha_{ij}}\theta) (1 - \theta/\sqrt{\alpha_{ij}})^{-\alpha_{ij}} \quad \theta < \sqrt{\alpha_{ij}}. \tag{B.4}$$

The distribution of ξ_{ij} can be called the standardized gamma distribution (see Table A.4).

It only remains to identify g_{ij} so as to be able to use $M_{ij}(g_{ij}\theta)$ with one of the asymmetric *a priori* bounds.

Proposition 3. *If $\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$, and if ξ_{ij} and $M_{ij}(\theta)$ are defined as per equations (B.3) and (B.4), respectively, then $M_{ij}(\theta) \geq M_{ij}(-\theta)$ over $(0, \sqrt{\alpha_{ij}})$ and thus $g_{ij} = 1$.*

Proof. First, terms $\exp(-\sqrt{\alpha_{ij}}g_{ij}\theta)$ and $(1 - g_{ij}\theta/\sqrt{\alpha_{ij}})^{-\alpha_{ij}}$, $g_{ij} = \pm 1$, are noted to be positive over $|\theta| < \sqrt{\alpha_{ij}}$. Then, these terms can be written with

their series representation:

$$\begin{aligned}
\exp(-\sqrt{\alpha_{ij}}g_{ij}\theta) &= \left[\exp\left(-\frac{g_{ij}\theta}{\sqrt{\alpha_{ij}}}\right) \right]^{\alpha_{ij}} \\
&= \left[\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n g_{ij}^n \left(\frac{\theta}{\sqrt{\alpha_{ij}}}\right)^n \right]^{\alpha_{ij}} \\
(1 - g_{ij}\theta/\sqrt{\alpha_{ij}})^{-\alpha_{ij}} &= \left[\frac{1}{1 - \frac{g_{ij}\theta}{\sqrt{\alpha_{ij}}}} \right]^{\alpha_{ij}} \\
&= \left[\sum_{n=0}^{\infty} g_{ij}^n \left(\frac{\theta}{\sqrt{\alpha_{ij}}}\right)^n \right]^{\alpha_{ij}}.
\end{aligned}$$

Finally, it can be shown that $M(\theta) \geq M_{ij}(-\theta)$ over $(0, \sqrt{\alpha_{ij}})$. Defining $y := \theta/\sqrt{\alpha_{ij}}$ for convenience, and recalling that $\alpha_{ij} > 0$:

$$\begin{aligned}
M(y) &= \left[\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n (y)^n \right]^{\alpha_{ij}} \left[\sum_{n=0}^{\infty} (y)^n \right]^{\alpha_{ij}} & y \in (0, 1) \\
M(y)^{1/\alpha_{ij}} &= \left[\sum_{n=0}^{\infty} \frac{1}{n!} (-1)^n (y)^n \right] \left[\sum_{n=0}^{\infty} (y)^n \right] & y \in (0, 1) \\
&= \sum_{n=0}^{\infty} \left[\sum_{i=0}^n \frac{1}{i!} (-1)^i \right] y^n & y \in (0, 1) \\
M(-y)^{1/\alpha_{ij}} &= \left[\sum_{n=0}^{\infty} \frac{1}{n!} (y)^n \right] \left[\sum_{n=0}^{\infty} (-1)^n (y)^n \right] & y \in (0, 1) \\
&= \sum_{n=0}^{\infty} \left[\sum_{i=0}^n \frac{1}{i!} (-1)^{n-i} \right] y^n & y \in (0, 1) \\
&= \sum_{n=0}^{\infty} \left[\sum_{i=0}^n \frac{1}{i!} (-1)^i \right] (-1)^n y^n & y \in (0, 1)
\end{aligned}$$

The truncated series over index i is the exponential sum function $\exp_n(x)$ with

$x = -1$, and it is nonnegative at -1 for all n :

$$\begin{aligned}
\sum_{i=0}^0 \frac{1}{i!} (-1)^i &= 1 & \sum_{i=0}^4 \frac{1}{i!} (-1)^i &= \frac{1}{3} + \frac{1}{24} = \frac{3}{8} \\
\sum_{i=0}^1 \frac{1}{i!} (-1)^i &= 1 - 1 = 0 & \sum_{i=0}^5 \frac{1}{i!} (-1)^i &= \frac{3}{8} - \frac{1}{120} = \frac{11}{30} \approx 0.367 \\
\sum_{i=0}^2 \frac{1}{i!} (-1)^i &= \frac{1}{2} & \sum_{i=0}^n \frac{1}{i!} (-1)^i &= \frac{11}{30} + O(1/(6!)) \\
\sum_{i=0}^3 \frac{1}{i!} (-1)^i &= \frac{1}{2} - \frac{1}{6} = \frac{1}{3} & \sum_{i=0}^{\infty} \frac{1}{i!} (-1)^i &= e^{-1} \approx 0.368.
\end{aligned}$$

Thus, $M(y) \geq M(-y)$ over the domain because the nonnegative truncated series sums are being subtracted in $M(-y)$ when n is odd. $\square$

Gamma distributions are asymmetric, so distribution-dependent bounds (GMF8), (GMF9), and (GMF10), which apply to asymmetric distributions, must be used. Note that the cumulant generating function corresponding to (B.4) does not meet condition (C₂), and thus bound (GMF10') cannot be applied. With $g_{ij} = 1$,

$$\Lambda_{ij}(g_{ij}\theta) = \Lambda_{ij}(\theta) = -\sqrt{\alpha_{ij}}\theta - \alpha_{ij} \ln(1 - \theta/\sqrt{\alpha_{ij}}) \quad \theta \in (0, \sqrt{\alpha_{ij}}) \quad (\text{B.5})$$

$$\frac{d\Lambda_{ij}(\theta)}{d\theta} = \frac{\theta}{1 - \theta/\sqrt{\alpha_{ij}}} \quad \theta \in (0, \sqrt{\alpha_{ij}}). \quad (\text{B.6})$$

To use bound (GMF9), distributions must be ordered according to conditions (C₁). Note that for every $\theta \in (0, \sqrt{\alpha_{ij}})$, $d\Lambda_{ij}(\theta)/d\theta$ decreases with increasing α_{ij} . Thus, ordering gamma distributions of ξ_{ij} by condition (C₁) for bound (GMF9) is simply a matter of sorting ξ_{ij} by ascending α_{ij} :

$$\frac{\theta}{1 - \theta/\sqrt{\alpha_{ij}}} > \frac{\theta}{1 - \theta/\sqrt{\alpha_{ik}}} \quad \alpha_{ij} < \alpha_{ik}, \forall \theta \in (0, \sqrt{\alpha_{ij}}). \quad (\text{B.7})$$

When modeled as described in this section, gamma distributions will be ordered before the standard normal distribution (where $d\Lambda_{ij}(\theta)/d\theta = \theta$):

$$\frac{\theta}{1 - \theta/\sqrt{\alpha_{ij}}} > \theta \quad \forall \theta \in (0, \sqrt{\alpha_{ij}}). \quad (\text{B.8})$$

These findings are summarized in Table A.4.

Appendix C. Proofs for bounds (GMF14⁺) and (GMF14⁻)

Bounds (GMF14⁺) and (GMF14⁻), which approach exact probabilities with arbitrary proximity, rely on the cumulative distribution function of the sum of

gamma-distributed random variables, provided in (Moschopoulos, 1985). First, a key result in (Moschopoulos, 1985) will be presented and reformulated, and then the bounds will be substantiated.

Given random variables $X_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta'_{ij})$ with $\alpha_{ij}, \beta'_{ij} > 0, \forall j$, the cumulative distribution function of random variable Ξ_i , where $\Xi_i := \sum_j X_{ij}$, is (Moschopoulos, 1985)

$$F_{\Xi_i}(x) = C_i \sum_{n=0}^{\infty} \delta_n \int_0^x \frac{y^{\rho_i+n-1} e^{-y/\beta_i^L}}{\Gamma(\rho_i+n) (\beta_i^L)^{\rho_i+n}} dy \quad (\text{C.1})$$

where

$$\begin{aligned} \beta_i^L &:= \min_{j \in J_i^*} \beta'_{ij} \\ C_i &:= \prod_{j \in J_i^*} \left(\frac{\beta_i^L}{\beta'_{ij}} \right)^{\alpha_{ij}} \\ \nu_k &:= \sum_{j \in J_i^*} \alpha_{ij} \left(1 - \frac{\beta_i^L}{\beta'_{ij}} \right)^k \\ \delta_0 &:= 1, \quad \delta_n := \frac{1}{n} \sum_{k=1}^n \nu_k \delta_{n-k}, \quad n = 1, 2, \dots \\ \rho_i &:= \sum_{j \in J_i^*} \alpha_{ij}. \end{aligned} \quad (\text{C.2})$$

Note that at this point, the definitions in (C.2) somewhat differ from those in (17). Equation (C.1) can be simplified by a change of integration variable from y to z with $z := y/\beta_i^L$:

$$\begin{aligned} F_{\Xi_i}(x) &= C_i \sum_{n=0}^{\infty} \frac{\delta_n}{\Gamma(\rho_i+n) (\beta_i^L)^{\rho_i+n}} \int_{y=0}^{y=x} y^{\rho_i+n-1} e^{-y/\beta_i^L} dy \\ &= C_i \sum_{n=0}^{\infty} \frac{\delta_n}{\Gamma(\rho_i+n) (\beta_i^L)^{\rho_i+n}} \int_{z=0}^{z=x/\beta_i^L} (\beta_i^L z)^{\rho_i+n-1} e^{-z} \beta_i^L dz \\ &= C_i \sum_{n=0}^{\infty} \frac{\delta_n}{\Gamma(\rho_i+n)} \int_0^{x/\beta_i^L} z^{\rho_i+n-1} e^{-z} dz \\ &= C_i \sum_{n=0}^{\infty} \delta_n \frac{\gamma(\rho_i+n, x/\beta_i^L)}{\Gamma(\rho_i+n)}. \end{aligned} \quad (\text{C.3})$$

Gamma distributions exhibit the property that $cX_{ij} \sim \text{Gamma}(\alpha_{ij}, c\beta_{ij})$ for $c > 0$. Given $\tilde{\alpha}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ and $x_j^* > 0, \forall j \in J_i^*$, we can set $X_{ij} := \tilde{\alpha}_{ij} x_j^*$, where $X_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij} x_j^*)$. This is equivalent to stating $X_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij} |x_j^*|)$. Thus, replacing β'_{ij} with $\beta_{ij} |x_j^*|$ then equates the

definitions of (C.2) with (17). Bound (GMF14⁺) then follows:

$$\begin{aligned}
& \Pr \left\{ \sum_{j \notin J_i^*} a_{ij} x_j^* + \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > b_i \right\} \\
&= \Pr \left\{ \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > \tau_i(x^*) \right\} \\
&= 1 - F_{\Xi_i}(\tau_i(x^*)) \\
&= 1 - C_i \sum_{n=0}^{\infty} \delta_n \frac{\gamma(\rho_i + n, \tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \\
&\stackrel{1}{\leq} 1 - C_i \sum_{n=0}^N \delta_n \frac{\gamma(\rho_i + n, \tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \quad (\text{GMF14}^+)
\end{aligned}$$

where relation 1 holds because each term in the infinite series is nonnegative.

For bound (GMF14⁻), $F_{\Xi_i}(x)$ must be reformulated again. First, for any cumulative distribution function $F(x)$, it must hold that $F(\infty) \rightarrow 1$. This can be utilized with (C.3) to provide the following identity:

$$\begin{aligned}
F_{\Xi_i}(\infty) &= C_i \sum_{n=0}^{\infty} \delta_n \frac{\gamma(\rho_i + n, \infty)}{\Gamma(\rho_i + n)} \\
&= C_i \sum_{n=0}^{\infty} \delta_n = 1.
\end{aligned} \tag{C.4}$$

Then, using the additional identity

$$\Gamma(z) = \gamma(z, x) + \Gamma(z, x),$$

$F_{\Xi_i}(x)$ can be written as

$$\begin{aligned}
F_{\Xi_i}(x) &= C_i \sum_{n=0}^{\infty} \delta_n \frac{\gamma(\rho_i + n, x/\beta_i^L)}{\Gamma(\rho_i + n)} \\
&= C_i \sum_{n=0}^{\infty} \delta_n \frac{\Gamma(\rho_i + n) - \Gamma(\rho_i + n, x/\beta_i^L)}{\Gamma(\rho_i + n)} \\
&= C_i \sum_{n=0}^{\infty} \delta_n - C_i \sum_{n=0}^{\infty} \delta_n \frac{\Gamma(\rho_i + n, x/\beta_i^L)}{\Gamma(\rho_i + n)} \\
&\stackrel{1}{=} 1 - C_i \sum_{n=0}^{\infty} \delta_n \frac{\Gamma(\rho_i + n, x/\beta_i^L)}{\Gamma(\rho_i + n)} \tag{C.5}
\end{aligned}$$

where equality 1 is due to (C.4). Given $\tilde{a}_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij})$ and $x_j^* < 0$, $\forall j \in J_i^*$, we can set $X_{ij} := \tilde{a}_{ij}(-x_j^*)$. This is equivalent to setting $X_{ij} := \tilde{a}_{ij} |x_j^*|$,

where $X_{ij} \sim \text{Gamma}(\alpha_{ij}, \beta_{ij} | x_j^*)$, just as before. Finally, bound (GMF14⁻) follows from these definitions and (C.5):

$$\begin{aligned}
& \Pr \left\{ \sum_{j \notin J_i^*} a_{ij} x_j^* + \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > b_i \right\} \\
&= \Pr \left\{ \sum_{j \in J_i^*} \tilde{a}_{ij} x_j^* > \tau_i(x^*) \right\} \\
&= \Pr \left\{ \sum_{j \in J_i^*} \tilde{a}_{ij} (-x_j^*) < -\tau_i(x^*) \right\} \\
&= \Pr \{ \Xi_i < -\tau_i(x^*) \} \\
&\stackrel{1}{=} \Pr \{ \Xi_i \leq -\tau_i(x^*) \} \\
&= 1 - C_i \sum_{n=0}^{\infty} \delta_n \frac{\Gamma(\rho_i + n, -\tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \\
&\stackrel{2}{\leq} 1 - C_i \sum_{n=0}^N \delta_n \frac{\Gamma(\rho_i + n, -\tau_i(x^*)/\beta_i^L)}{\Gamma(\rho_i + n)} \quad (\text{GMF14}^-)
\end{aligned}$$

where equality 1 is due to the fact that Ξ_i is subject to a continuous distribution and thus $\Pr \{ \Xi_i = x \} = 0$ for any $x \in \mathbb{R}$, and relation 2 holds because each term in the infinite series is nonnegative.

Appendix D. Lemmas required for *a priori* bounds

Lemma 1. *Given the random variable ξ is subject to a probability distribution that is supported on the bounded interval $[a, b]$, and given that its moment generating function $M_\xi(\theta) = \mathbf{E}[e^{\theta\xi}]$ exists, then for any real number θ ,*

$$(i) \quad \mathbf{E}[e^{\theta\xi}] \leq \frac{b - \mathbf{E}[\xi]}{b - a} e^{\theta a} + \frac{\mathbf{E}[\xi] - a}{b - a} e^{\theta b}$$

(ii) *If ξ is subject to a symmetric distribution, $a = -1$, and $b = 1$, then*

$$\frac{d^n}{d\theta^n} \mathbf{E}[e^{\theta\xi}] \leq \begin{cases} \cosh(\theta) & \text{if } n \text{ is even} \\ \sinh(\theta) & \text{if } n \text{ is odd.} \end{cases}$$

Proof. See Lemma 1, Guzman et al. (2016b).

Lemma 2. *Let $f_j : X \rightarrow Y$, where $X = \{x_j \in \mathbb{R} : 0 \leq x_j^L \leq x_j \leq x_j^U\}$ and $Y = \{y \in \mathbb{R} : 0 \leq y < \infty\}$, be nondecreasing and log-convex over X , $j = 1, 2, \dots, n$, with*

$$\frac{f_1(x_i)}{f_1(x_j)} \geq \frac{f_2(x_i)}{f_2(x_j)} \geq \dots \geq \frac{f_n(x_i)}{f_n(x_j)} \quad \forall (i, j) : x_i, x_j \in X, x_i \geq x_j \quad (\text{D.1})$$

or

$$\frac{d \ln f_1(x)}{dx} \geq \dots \geq \frac{d \ln f_n(x)}{dx} \quad \forall x \in X. \quad (\text{D.2})$$

Given that the nonconvex program

$$\max_x \left\{ \prod_{j=1}^n f_j(x_j) : \sum_{j=1}^n x_j \leq x^\sigma; x_j \in X, \forall j \right\}, \quad (\text{P1})$$

is feasible (i.e., $x^L \leq x^\sigma/n$), then the point $x^* = (x_1^*, x_2^*, \dots, x_j^*, \dots, x_n^*)^\top$, where

$$x_j^* = \begin{cases} x^U & j \leq \bar{j} \\ x^m & j = \bar{j} + 1 \\ x^L & j \geq \bar{j} + 2, \end{cases}$$

$$\bar{j} := \left\lfloor \frac{x^\sigma - nx^L}{x^U - x^L} \right\rfloor,$$

$$x^m := x^\sigma - \bar{j}(x^U - x^L) - (n-1)x^L,$$

is a global maximizer to program (P1).

Proof. See Lemma 2, Guzman et al. (2016a).

Lemma 3. Let $f : X \rightarrow Y$, where $X = \{x \in \mathbb{R} : 0 \leq x \leq 1\}$ and $Y = \{y \in \mathbb{R} : 0 < y < \infty\}$, be nondecreasing over X , and let the composite function $(f \circ g)(x)$, where $g(x) = \sqrt{x}$, be concave over X . Then the program

$$\max_x \left\{ \sum_j f(x_j) : \sum_j x_j^2 \leq 1; x_j \in [0, 1], \forall j \right\} \quad (\text{P2})$$

with $j = 1, 2, \dots, n$, has an optimal solution $x_j^* = 1/\sqrt{n}$.

Proof. See Lemma 3, Guzman et al. (2016a).

The qualifying conditions of Lemmas 2 and 3 can be equivalently stated in terms of $\Lambda_{ij}(\theta)$ as:

$$\frac{d\Lambda_{ij}(\theta)}{d\theta} \geq \frac{d\Lambda_{ik}(\theta)}{d\theta} \quad \forall j, k \in J_i, j < k, \forall \theta \geq 0 \quad (\text{C}_1)$$

$$\Lambda_{ij}(\sqrt{\theta}) \text{ concave in } \theta \quad \forall \theta \geq 0. \quad (\text{C}_2)$$

The first condition (C₁) mandates that the indices in J_i are ordered such that the relationships in (C₁) hold, while the second condition (C₂) requires a specific property of cumulant generating function $\Lambda_{ij}(\theta)$.

References

- Au, S.-K., Beck, J. L., 2001. Estimation of small failure probabilities in high dimensions by subset simulation. *Probabilistic Engineering Mechanics* 16 (4), 263–277.
- Averbakh, I., Zhao, Y.-B., 2008. Explicit reformulations for robust optimization problems with general uncertainty sets. *SIAM Journal on Optimization* 18 (4), 1436–1466.
- Ben-Tal, A., Boyd, S., Nemirovski, A., 2006. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming* 107 (1), 63–89.
- Ben-Tal, A., El Ghaoui, L., Nemirovski, A., 2009. *Robust optimization*. Princeton University Press.
- Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A., 2003. Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99 (2), 351–376.
- Ben-Tal, A., Nemirovski, A., 1998. Robust convex optimization. *Mathematics of Operations Research* 23 (4), 769–805.
- Ben-Tal, A., Nemirovski, A., 1999. Robust solutions of uncertain linear programs. *Operations research letters* 25 (1), 1–13.
- Ben-Tal, A., Nemirovski, A., 2000. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical programming* 88 (3), 411–424.
- Ben-Tal, A., Nemirovski, A., 2007. Selected topics in robust convex optimization. *Mathematical Programming* 112 (1), 125–158.
- Bertsimas, D., Brown, D. B., 2009. Constructing uncertainty sets for robust linear optimization. *Operations Research* 57 (6), 1483–1495.
- Bertsimas, D., Brown, D. B., Caramanis, C., 2011. Theory and applications of robust optimization. *SIAM Review* 53 (3), 464–501.
- Bertsimas, D., Goyal, V., 2012. On the approximability of adjustable robust convex optimization under uncertainty. *Mathematical Methods of Operations Research* 77 (3), 323–343.
- Bertsimas, D., Sim, M., 2004. The price of robustness. *Operations research* 52 (1), 35–53.
- Chen, X., Zhang, Y., 2009. Uncertain linear programs: Extended affinely adjustable robust counterparts. *Operations Research* 57 (6), 1469–1482.

- El Ghaoui, L., Lebret, H., 1997. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications* 18 (4), 1035–1064.
- El Ghaoui, L., Oustry, F., Lebret, H., 1998. Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization* 9 (1), 33–52.
- Floudas, C. A., Gümüş, Z. H., Ierapetritou, M. G., 2001. Global optimization in design under uncertainty: feasibility test and flexibility index problems. *Industrial & Engineering Chemistry Research* 40 (20), 4267–4282.
- Floudas, C. A., Niziolek, A. M., Onel, O., Matthews, L. R., 2016. Multi-scale systems engineering for energy and the environment: Challenges and opportunities. *AIChE Journal* 62 (3), 602–623.
- Gabrel, V., Murat, C., Thiele, A., 2014. Recent advances in robust optimization: An overview. *European Journal of Operational Research* 235 (3), 471 – 483.
- Gounaris, C. E., Misener, R., Floudas, C. A., 2009. Computational comparison of piecewise- linear relaxations for pooling problems. *Industrial & Engineering Chemistry Research* 48 (12), 5742–5766.
- Gounaris, C. E., Wiesemann, W., Floudas, C. A., 2013. The robust capacitated vehicle routing problem under demand uncertainty. *Operations Research* 61 (3), 677–693.
- Guzman, Y. A., Matthews, L. R., Floudas, C. A., 2016a. New *a priori* and *a posteriori* probabilistic bounds for robust counterpart optimization: I. Unknown probability distributions. *Computers & Chemical Engineering* 84, 568–598.
- Guzman, Y. A., Matthews, L. R., Floudas, C. A., 2016b. New *a priori* and *a posteriori* probabilistic bounds for robust counterpart optimization: II. *A priori* bounds for known symmetric and asymmetric probability distributions. *Computers & Chemical Engineering*, In Press.
- Janak, S. L., , Floudas*, C. A., Kallrath, J., Vormbrock, N., 2006. Production scheduling of a large-scale industrial batch plant. ii. reactive scheduling. *Industrial & Engineering Chemistry Research* 45 (25), 8253–8269.
- Janak, S. L., Lin, X., Floudas, C. A., 2007. A new robust optimization approach for scheduling under uncertainty: II. Uncertainty with known probability distribution. *Computers & chemical engineering* 31 (3), 171–195.
- Johnson, M. E., 2013. Multivariate statistical simulation: A guide to selecting and generating continuous multivariate distributions. John Wiley & Sons.
- Kamgar-Parsi, B., Kamgar-Parsi, B., Brosh, M., 1995. Distribution and moments of the weighted sum of uniforms random variables, with applications in reducing monte carlo simulations. *Journal of Statistical Computation and Simulation* 52 (4), 399–414.

- Kang, S.-C., Brisimi, T. S., Paschalidis, I. C., 2013. Distribution-dependent robust linear optimization with applications to inventory control. *Annals of Operations Research*, 1–35.
- Lappas, N. H., Gounaris, C. E., 2016. Multi-stage adjustable robust optimization for process scheduling under uncertainty. *AIChE Journal* 62 (5), 1646–1667.
- Li, J., Misener, R., Floudas, C. A., 2012a. Scheduling of crude oil operations under demand uncertainty: A robust optimization framework coupled with global optimization. *AIChE Journal* 58 (8), 2373–2396.
- Li, Z., Ding, R., Floudas, C. A., 2011. A comparative theoretical and computational study on robust counterpart optimization: I. Robust linear optimization and robust mixed integer linear optimization. *Industrial & Engineering Chemistry Research* 50 (18), 10567–10603.
- Li, Z., Tang, Q., Floudas, C. A., 2012b. A comparative theoretical and computational study on robust counterpart optimization: II. Probabilistic guarantees on constraint satisfaction. *Industrial & Engineering Chemistry Research* 51 (19), 6769–6788.
- Li, Z., Floudas, C. A., 2014. A comparative theoretical and computational study on robust counterpart optimization: III. Improving the quality of robust solutions. *Industrial & Engineering Chemistry Research* 53 (33), 13112–13124.
- Lin, X., Janak, S. L., Floudas, C. A., 2004. A new robust optimization approach for scheduling under uncertainty: I. Bounded uncertainty. *Computers & chemical engineering* 28 (6), 1069–1085.
- Misener, R., Floudas, C. A., 2010. Global optimization of large-scale pooling problems: Quadratically constrained MINLP models. *Industrial & Engineering Chemistry Research* 49 (11), 5424–5438.
- Misener, R., Floudas, C. A., 2012. Global optimization of mixed-integer quadratically-constrained quadratic programs (MIQCQP) through piecewise-linear and edge-concave relaxations. *Mathematical Programming* 136 (1), 155–182.
- Misener, R., Floudas, C. A., 2013. GloMIQO: Global mixed-integer quadratic optimizer. *Journal of Global Optimization* 57 (1), 3–50.
- Misener, R., Floudas, C. A., 2014a. ANTIGONE: Algorithms for coNTinuous/Integer Global Optimization of Nonlinear Equations. *Journal of Global Optimization* 59 (2-3), 503–526.
- Misener, R., Floudas, C. A., 2014b. A framework for globally optimizing mixed-integer signomial programs. *Journal of Optimization Theory and Applications* 161 (3), 905–932.

- Misener, R., Smadbeck, J. B., Floudas, C. A., 2015. Dynamically generated cutting planes for mixed-integer quadratically constrained quadratic programs and their incorporation into GloMIQO 2. *Optimization Methods and Software* 30 (1), 215–249.
- Misener, R., Thompson, J. P., Floudas, C. A., 2011. APOGEE: Global optimization of standard, generalized, and extended pooling problems via linear and logarithmic partitioning schemes. *Computers & Chemical Engineering* 35 (5), 876–892.
- Moschopoulos, P. G., 1985. The distribution of the sum of independent gamma random variables. *Annals of the Institute of Statistical Mathematics* 37 (1), 541–544.
- Nemirovski, A., Juditsky, A., Lan, G., Shapiro, A., 2009. Robust stochastic approximation approach to stochastic programming. *SIAM Journal on optimization* 19 (4), 1574–1609.
- Nemirovski, A., Shapiro, A., 2007. Convex approximations of chance constrained programs. *SIAM Journal on Optimization* 17 (4), 969–996.
- Paschalidis, I., Kang, S., Li, K., 2008. Distribution-dependent robust linear optimization with asymmetric uncertainty and application to optimal control. In: *Proceedings of the 17th World Congress IFAC*. pp. 10069–10074.
- Ravindran, A. R., 2008. *Operations research methodologies*. Crc Press.
- Soyster, A. L., 1973. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research* 21 (5), 1154–1157.
- Verderame, P. M., Floudas, C. A., 2009a. Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty. i. robust optimization framework. *Industrial & engineering chemistry research* 48 (15), 7214–7231.
- Verderame, P. M., Floudas, C. A., 2009b. Operational planning of large-scale industrial batch plants under demand due date and amount uncertainty: II. Conditional value-at-risk framework. *Industrial & engineering chemistry research* 49 (1), 260–275.