MATHEMATICS OF OPERATIONS RESEARCH
Vol. 34, No. 1, February 2009, pp. 1–25
ISSN 0364-765X | EISSN 1526-5471 | 09 | 3401 | 0001

DOI 10.1287/moor.1080.0352

On Safe Tractable Approximations of Chance-Constrained Linear Matrix Inequalities

Aharon Ben-Tal
Faculty of Industrial Engineering and Management, MINERVA Optimization Center, Technion–Israel Institute of Technology, Technion City, Haifa 32000, Israel, abental@ie.technion.ac.il

Arkadi Nemirovski
School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia 30332, nemirovs@isye.gatech.edu

In the paper we consider the chance-constrained version of an affinely perturbed linear matrix inequality (LMI) constraint, assuming the primitive perturbations to be independent with light-tail distributions (e.g., bounded or Gaussian). Constraints of this type, playing a central role in chance-constrained linear/conic quadratic/semidefinite programming, are typically computationally intractable. The goal of this paper is to develop a tractable approximation to these chance constraints. Our approximation is based on measure concentration results and is given by an explicit system of LMIs. Thus, the approximation is computationally tractable; moreover, it is also safe, meaning that a feasible solution of the approximation is feasible for the chance constraint.

Key words: chance constraints; linear matrix inequalities; convex programming; measure concentration
MSC2000 subject classification: Primary: 90C15, 90C22, 90C25; secondary: 60F10
OR/MS subject classification: Primary: programming/stochastic; secondary: mathematics/matrices
History: Received November 20, 2006; revised October 19, 2007 and August 13, 2008.

1. Introduction. In this paper we study uncertain linear matrix inequalities (LMIs)

$$\mathcal{A}(x,\zeta)\succeq 0, \tag{1}$$

where $x\in\mathbf{R}^m$ is the decision vector, $\zeta\in\mathbf{R}^d$ is *data perturbation*, the body $\mathcal{A}(x,\zeta)$ of the inequality is bi-affine mapping in x and ζ taking values in the space $\mathbf{S}^n$ of symmetric $n\times n$ matrices, specifically:

$$\mathcal{A}(x,\zeta)=\mathcal{A}_0[x]+\sum_{l=1}^{d}\zeta_l\mathcal{A}_l[x], \tag{2}$$

where the matrices $\mathcal{A}_0[x],\ldots,\mathcal{A}_d[x]\in\mathbf{S}^n$ are affine in x. Here $A\succeq B$ means that A, B are symmetric matrices such that the matrix $A-B$ is positive semidefinite. We are interested in the case when (1) is a constraint in an optimization problem we wish to solve, and our goal is to process such an uncertain constraint. Given the basic role played by LMI constraints in modern convex optimization and the fact that the data in real-life optimization problems in many cases are uncertain (not known exactly the time the problem is to be solved), the question of how to process an uncertain LMI constraint is of major interest.

For the time being, there are two main approaches to treating uncertain constraints. The more traditional one, offered by stochastic programming, utilizes a stochastic uncertainty model: ζ is assumed to be a random vector with known (perhaps only partially) distribution. Here a natural way is to pass from the uncertain constraint (1) to its *chance-constrained* version—the usual—“certain”—constraint

$$p(x):=\inf_{P\in\mathcal{P}}\mathrm{Prob}_{\zeta\sim P}\{\mathcal{A}(x,\zeta)\succeq 0\}\geq 1-\epsilon, \tag{3}$$

where $\mathcal{P}$ is the family of all probability distributions of ζ compatible with our a priori information, and $\epsilon\in(0,1)$ is a given tolerance. An alternative to this approach, offered by robust optimization, is based on an “uncertain-but-bounded” model of data perturbations where all our a priori knowledge of ζ is that it belongs to a given *uncertainty set* $\mathcal{Z}$. In this case, a natural way is to replace the uncertain constraint with its *robust counterpart*

$$\mathcal{A}(x,\zeta)\succeq 0\quad\forall\,\zeta\in\mathcal{Z}. \tag{4}$$

Note that both outlined approaches “as they are” usually lead to computationally intractable constraints. As far as the chance-constrained LMI (3) is concerned, typically the only way to check whether a given point belongs to its feasible set is to use Monte Carlo simulation with sample size of order ϵ^{-1}, and this is computationally too demanding when ϵ is small. Another difficulty comes from the fact that the feasible set of (3) is usually nonconvex. The latter complication does not arise with the robust optimization approach—the feasible set of (4) is always convex; unfortunately, the first difficulty—impossible to check efficiently whether this *semi-infinite* convex constraint is satisfied at a given point—may become even more severe than in the case of chance-constrained LMI. These tractability difficulties of processing the LMI (1) make it natural to replace such a constraint with a *safe tractable approximation*—a system $\mathcal{S}$ of efficiently computable convex constraints in variables x and, perhaps, additional variables u such that whenever (x, u) is feasible for $\mathcal{S}$, x is feasible for the constraint (1). For the time being, "tight" (in a certain precise sense) approximations of this type are known only for the robust counterpart type constraints (4), and only under specific restrictions on the structure of $\mathcal{A}(x,\zeta)$; see Ben-Tal et al. [3, 4, 5]. In this paper, we focus solely on chance-constrained LMIs (3). In this case, seemingly the only safe tractable approximation known in the literature is the one given by the general *scenario approach*. For a chance-constrained optimization program

$$\min_x \{f_0(x)\colon \operatorname{Prob}\{f_i(x,\zeta)\le 0\}\ge 1-\epsilon,\ i=1,\dots,I\},$$

its scenario approximation is the *random* optimization program

$$\min_x \{f_0(x)\colon f_i(x,\zeta^j)\le 0,\ i=1,\dots,I,\ j=1,\dots,J\},$$

where $\zeta^1,\dots,\zeta^J$ is a sample of independent realizations of ζ. Theoretical justification of this natural approximation scheme is presented in Calafiore and Campi [8] and de Farias and Van Roy [9]. In particular, it is shown in Calafiore and Campi [8] that if $f_0(x)$, $f_i(x,\zeta)$, $i=1,\dots,I$, are convex in $x\in\mathbf{R}^m$ and the sample size J is large enough:

$$J\ge J^* := \operatorname{Ceil}\left[2m\epsilon^{-1}\log(12/\epsilon)+2\epsilon^{-1}\log(2/\delta)+2m\right], \tag{5}$$

then an optimal solution to the approximation, up to probability $\le\delta$ of "bad sampling," is feasible for the chance-constrained problem. (For substantial extensions of this remarkable result to the case of *ambiguously* chance-constrained convex problems, see Erdogan and Iyengar [10].) Although pretty general (in particular, imposing no restrictions on how the random perturbations enter the constraints and how they are distributed) and tractable, the scenario approximation has an intrinsic drawback—it requires samples of order $1/\epsilon$, and thus becomes prohibitively computationally demanding when ϵ becomes small, like 10^{-5} or less. For affinely perturbed LMIs (2) with *independent of each other* "light-tail" perturbations ζ_l, $l=1,\dots,d$, this drawback can be circumvented by a kind of importance sampling; see Nemirovski and Shapiro [15]. In this paper, we work under the same assumptions as in Nemirovski and Shapiro [15], i.e., focus on affinely perturbed LMIs with independent-of-each-other light-tail random perturbations ζ_l, and develop a novel, safe, tractable approximation of the chance-constrained versions (3) of these LMIs. In contrast to the purely simulation-based approximations of Calafiore and Campi [8], Erdogan and Iyengar [10], and Nemirovski and Shapiro [15], our new approximation is *nearly analytic*. Specifically, by itself our approximation is an explicit semidefinite program depending on a pair of real parameters and completely independent of any samples. In order for this approximation to be safe, the pair of parameters in question should be "properly guessed," that is, should ensure the validity of a specific large-deviation-type inequality. In principle, we can point out appropriate values of these parameters in advance. However, to reduce the conservatism of the approximation, we allow for an "optimistic" choice of the parameters and introduce a specific simulation-based postoptimization *validation procedure* that allows us either to justify our "optimistic guess" (and thus guarantees "up to probability $\le\delta$ of bad sampling" that the solution we end up with is feasible for the chance constraint of interest), or else demonstrates that our guess was "too optimistic," in which case we can pass to an approximation with better-chosen parameters. It should be stressed that in principle the size J of the sample used in this "validation procedure" is completely independent of how small the tolerance ϵ is; all we need is $J\ge O(1)\ln(1/\delta)$.

The rest of the paper is organized as follows. In §2 we make our standing assumptions and outline and motivate our approximation strategy. This strategy is fully developed in §§3.1 and 3.2. In §4 we consider two important special cases of (3). In the first of them, all matrices $\mathcal{A}_l[x]$, $l=0,1,\dots,d$, are diagonal. This is the case of randomly perturbed *scalar linear inequalities* or, which is the same, about *chance-constrained linear programming*. In the second special case, the matrices $\mathcal{A}_l[x]$, $l=1,\dots,d$, are of the form $\lambda_l(x)G(x)+e(x)f_l^T(x)+f_l(x)e^T(x)$, where $e(x)$ and $f_l(x)$ are vectors (and, as always in this paper, $\mathcal{A}_l[x]$ is affine in x). This situation covers the case when (1) is a randomly perturbed *conic quadratic inequality* $\|A[x]\zeta+b[x]\|_2\le c^T[x]\zeta+d[x]$ ($A[x]$, $b[x]$, $c[x]$, $d[x]$ are affine in x); indeed,

$$\|A[x]\zeta+b[x]\|_2\le c^T[x]\zeta+d[x] \;\Leftrightarrow\; \underbrace{\left[\begin{array}{c|c} d[x] & b^T[x]\\ \hline b[x] & d[x]I\end{array}\right]}_{\mathcal{A}_0[x]}+\sum_{l=1}^d \zeta_l \underbrace{\left[\begin{array}{c|c} c_l[x] & A_l^T[x]\\ \hline A_l[x] & c_l[x]I\end{array}\right]}_{\mathcal{A}_l[x]}\succeq 0, \tag{6}$$

where $A_l[x]$ are the columns of $A[x]$, and $c_l[x]$ are the entries of $c[x]$. Note that "fully analytic" safe, tractable approximations of chance-constrained LPs were recently proposed in Nemirovski and Shapiro [14]; §4 contains a comparison of approximations from Nemirovski and Shapiro [14] with the one developed in this paper. Section 5 presents techniques allowing us to reduce the task of building a safe approximation for the chance-constrained LMI (3) (under *partially known* "light-tail" distributions of independent perturbations ζ_l) to a similar task for an appropriately chosen reference distribution of ζ (most notably a Gaussian one). The concluding §6 presents numerical illustrations.

2. Goals, assumptions, strategy. Recall that our ultimate goal is to process a given chance-constrained optimization problem of the form

$$\min_x \left\{ c^T x \colon \begin{array}{l} F(x) \le 0 \\ \mathrm{Prob}\left\{ \mathcal{A}_0[x] + \sum_{l=1}^d \zeta_l \mathcal{A}_l[x] \succeq 0 \right\} \ge 1-\epsilon \end{array} \right\}, \tag{7}$$

where $F(x)$ is an efficiently computable vector function with convex components, $\mathcal{A}_0[x], \ldots, \mathcal{A}_d[x]$ are symmetric matrices affinely depending on the decision vector x, $\epsilon \in (0,1)$ is a given tolerance, and $\zeta_1, \ldots, \zeta_d$ are random perturbations. What we intend to do is to replace in (7) the "troublemaking" chance constraint with a *safe tractable approximation*, the latter notion being defined as follows:

Definition 2.1. We say that an explicit system $\mathcal{S}$ of efficiently computable convex constraints on variables x and additional variables u is a safe, tractable approximation of the chance-constrained LMI

$$p(x) := \mathrm{Prob}\left\{ \mathcal{A}_0[x] + \sum_{l=1}^d \zeta_i A_i[x] \succeq 0 \right\} \ge 1-\epsilon \tag{8}$$

if whenever a vector x can be extended to a feasible solution (x,u) of $\mathcal{S}$, x is feasible for the chance constraint (8) (or, which is the same, if the projection X of the solution set of $\mathcal{S}$ on the space of x-variables is contained in the feasible set of (8)).

Note that the requirement that X is contained in the feasible set of (8) means that $\mathcal{S}$ produces a *sufficient* condition for (8) to be satisfied ("safety" of the approximation). Similarly, the requirement that $\mathcal{S}$ is a system of efficiently computable convex constraints implies that we can minimize efficiently convex functions over X ("tractability" of the approximation).

Replacing the chance constraint (8) in the optimization problem (7) with a safe tractable approximation we get an optimization problem in variables x, u with efficiently computable convex constraints, that is, we get an efficiently solvable problem, and feasible solutions of this problem are feasible for the problem of actual interest (7).

We shall address the problem of building a safe, tractable approximation of (8) under the following assumption on the random perturbations:

Assumption A. *The scalar random variables $\zeta_1, \ldots, \zeta_d$ are mutually independent with zero means and either* (a) *all ζ_l have bounded ranges, or* (b) *all ζ_l are Gaussian.*

Note that applying deterministic scalings $\zeta_l \mapsto \zeta_l/s_l$, $\mathcal{A}_l[x] \mapsto s_l \mathcal{A}_l[x]$, in the case of (a) *we can convert the ranges of ζ_l into the segment $[-1,1]$, and in the case of* (b) *we can enforce $\zeta_l \sim \mathcal{N}(0,1)$ for all l. Therefore, from now on, if not stated otherwise, we assume that either*

(A.1.) *ζ_l is supported on $[-1,1]$*, or

(A.2.) *$\zeta_l \sim \mathcal{N}(0,1)$ for all l.*

2.1. The strategy. The idea of the construction we are about to develop is simple. Essentially, what we are looking for is a verifiable sufficient condition for the relation

$$A_0 + \sum_{l=1}^d \zeta_l A_l \succeq 0 \tag{9}$$

to be satisfied with probability at least $1-\epsilon$; here $A_0, \ldots, A_d$ are given $n \times n$ symmetric matrices. Assuming, for the sake of argument, that ζ_l are symmetrically distributed and ϵ is small, this is basically the same as to seek a sufficient condition for the relation

$$\mathrm{Prob}\left\{ -A_0 \preceq S := \sum_{l=1}^d \zeta_l A_l \preceq A_0 \right\} \ge 1-\epsilon. \tag{10}$$

An evident *necessary* condition here is $A_0 \succeq 0$. Assuming a bit more, namely, that $A_0 \succ 0$, the condition of interest becomes

$$\mathrm{Prob}\left\{-I \preceq \hat{S} := \sum_{l=1}^{d} \zeta_l \hat{A}_l \preceq I\right\} \geq 1-\epsilon, \quad \hat{A}_l = A_0^{-1/2} A_l A_0^{-1/2}. \tag{11}$$

Now, in case (A.2) it is intuitively clear (and can be easily proved) that (11) implies that

$$\mathbf{E}\{\hat{S}^2\} = \sum_{l=1}^{d} \hat{A}_l^2 \preceq O(1)I \tag{$*$}$$

with some positive absolute constant $O(1)$. Thus, the condition $(*)$ is a *necessary* condition for (11), provided that we want the latter condition to be satisfied for all distributions of ζ compatible with Assumption (A.1). Now assume for a moment that a condition of the type $(*)$, namely, the condition

$$\sum_{l=1}^{d} \hat{A}_l^2 \preceq \gamma^2 I, \tag{12}$$

with properly chosen γ, is *sufficient* for (11) to be valid. Then we are basically done: It is immediately seen that (12) can be equivalently reformulated as the LMI

$$\mathrm{Arrow}(\gamma A_0, A_1, \ldots, A_d) \equiv \begin{bmatrix} \gamma A_0 & A_1 & \cdots & A_d \\ A_1 & \gamma A_0 & & \\ \vdots & & \ddots & \\ A_d & & & \gamma A_0 \end{bmatrix} \succeq 0 \tag{13}$$

in variables $A_0, A_1, \ldots, A_d$. It follows that when $A_l = \mathcal{A}_l[x]$, $l = 0, 1, \ldots, d$, depend affinely on a decision vector x (the situation we are interested in), our sufficient condition (13) for the validity of (10) (and thus for the validity of (9) as well) becomes an LMI in variables x and thus provides us with safe tractable approximation of (8). The level of conservatism of this approximation can be quantified by γ—the less γ is, the larger is the "gap" between the sufficient condition (12) and the necessary condition $(*)$. It is shown in Nemirovski [13] that in order for (12) to be *always* sufficient for (9) (i.e., independently of the structure of $\hat{A}_l \in \mathbf{S}^n$ and the random perturbations ζ_l—as long as they satisfy Assumption A), then γ should be *at most* $O(1)[\sqrt{\ln n} + \sqrt{\ln(1/\epsilon)}]^{-1}$. In Nemirovski [13], it is also proved that with properly chosen $O(1)$ and with $\gamma = O(1)[n^{1/6} + \sqrt{\ln(1/\epsilon)}]^{-1}$, condition (12) is sufficient for the validity of (10), and is conjectured that this conclusion remains true when $n^{1/6}$ is replaced with "unimprovable" $\sqrt{\ln n}$. This conjecture was justified recently; see Man-Cho So [12] and Proposition A.1 in the appendix. Note, however, that the outlined value of γ that provably makes (12) sufficient for (10) is worst-case oriented and thus might typically lead to an overly conservative approximation (13). The main idea of this paper is that, given any *guess* of (γ, χ) that makes (12) sufficient for the validity of (10), we can use a cheap simulation-based procedure to validate the result yielded by this guess, or else to refine our guess. Numerical results presented in §3 demonstrate that this approach can result in significantly less-conservative approximations of (9) than those associated with the above "provably safe" values of (γ, χ).

3. Approximating chance-constrained LMIs.

3.1. Preliminaries on measure concentration. Our strategy heavily exploits the following fact:

Theorem 3.1 ("Measure Concentration"). *Let $\zeta_1, \ldots, \zeta_d$ satisfy Assumption A, $\Upsilon > 0$, and $\chi \in (0, 1/2)$ be reals, and $B_0, \ldots, B_d$ be deterministic symmetric matrices such that*

$$\begin{aligned} &\text{(a)} \quad \mathrm{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0 \\ &\text{(b)} \quad \mathrm{Prob}\left\{-\Upsilon B_0 \preceq \sum_{l=1}^{d} \zeta_l B_l \preceq \Upsilon B_0\right\} \geq 1-\chi. \end{aligned} \tag{14}$$

Then

$$\gamma \geq 1 \;\Rightarrow\; \mathrm{Prob}\left\{-\gamma\Upsilon B_0 \preceq \sum_{l=1}^{d}\zeta_l B_l \preceq \gamma\Upsilon B_0\right\} \geq 1-\epsilon(\chi,\gamma),$$
$$\epsilon(\chi,\gamma) = \begin{cases} \dfrac{1}{1-\chi}\exp\{-\Upsilon^2(\gamma-1)^2/16\}, & \zeta \textit{ satisfies } (\mathrm{A.1}) \\ \mathrm{Erf}(\mathrm{ErfInv}(\chi)+(\gamma-1)\max[\Upsilon,\mathrm{ErfInv}(\chi)]), & \zeta \textit{ satisfies } (\mathrm{A.2}) \end{cases} \tag{15}$$

Here and in what follows Erf$(\cdot)$ *and* ErfInv$(\cdot)$ *are the error function and its inverse*:

$$\mathrm{Erf}(x) = \frac{1}{\sqrt{2\pi}}\int_x^\infty \exp\{-s^2/2\}\,ds, \qquad \mathrm{Erf}(\mathrm{ErfInv}(r)) \equiv r, \quad 0<r<1. \tag{16}$$

Proof. Under the premise of our theorem, we clearly have $B_0 \succeq 0$; by continuity reasons, it suffices to prove the theorem in the case of $B_0 \succ 0$. In this case, passing from the matrices $B_0, B_1, \ldots, B_d$ to the matrices $I, B_0^{-1/2}B_1B_0^{-1/2}, \ldots, B_0^{-1/2}B_dB_0^{-1/2}$, we immediately reduce the situation to the one with $B_0 = I$, which we assume from now on. In this case, (14.a) becomes simply $\sum_{l=1}^d B_l^2 \preceq I$. With this normalization, the validity of (15) in the case of (A.1) is readily given by Lemma 1 in Nemirovski [13], and in the case of (A.2) by the following refinement of Theorem 1 in Nemirovski and Shapiro [14]: □

Lemma 3.1. *Let $\zeta \sim \mathcal{N}(0, I_d)$, and let $Q \subset \mathbf{R}^d$ be a closed convex set such that* $\mathrm{Prob}\{\zeta \notin Q\} \leq \chi < 1/2$. *Then*
(i) *Q contains the centered at the origin $\|\cdot\|_2$-ball of radius* ErfInv(χ);
(ii) *If Q contains the centered at the origin $\|\cdot\|_2$-ball of a radius $r \geq \mathrm{ErfInv}(\chi)$, then for every $\gamma > 1$ one has*

$$\mathrm{Prob}\{\zeta \notin \gamma Q\} \leq \mathrm{Erf}((\gamma-1)r + \mathrm{ErfInv}(\chi)) \leq \mathrm{Erf}(\gamma\mathrm{ErfInv}(\chi)) \leq \exp\{-\gamma^2\mathrm{ErfInv}^2(\chi)/2\}. \tag{17}$$

Proof of Lemma 3.1. (i) is immediate. Indeed, assuming the opposite and invoking the separation theorem, Q is contained in a closed half-space $\Pi = \{x\colon e^Tx \leq r\}$ with a unit vector e and certain $r < \mathrm{ErfInv}(\chi)$, and therefore $\mathrm{Prob}\{\zeta \notin Q\} \geq \mathrm{Prob}\{\zeta \notin \Pi\} = \mathrm{Erf}(r) > \chi$, which is a contradiction.

(ii) is an immediate corollary of the following fact due to Borell [6]:

(!) *For every $\chi \in (0,1)$, $\epsilon \geq 0$ and every closed set $X \subset \mathbf{R}^d$ such that* $\mathrm{Prob}\{\zeta \notin X\} \leq \chi$ *one has* $\mathrm{Prob}\{\mathrm{dist}(\zeta, X) > \epsilon\} \leq \mathrm{Erf}(\mathrm{ErfInv}(\chi)+\epsilon)$, *where* $\mathrm{dist}(a,X) = \min_{x\in X}\|a-x\|_2$.

In the situation of (ii), Q contains the centered at the origin $\|\cdot\|_2$-ball B_r of the radius r, whence the set γQ, $\gamma \geq 1$, contains $Q + (\gamma-1)Q \supset Q + (\gamma-1)B_r$ and thus contains the set $\{x\colon \mathrm{dist}(x,Q) \leq \epsilon = (\gamma-1)r\}$. Invoking (!) with $X = Q$ and $\epsilon = (\gamma-1)r$, we get the first inequality in (17); the second inequality there is due to $r \geq \mathrm{ErfInv}(\chi)$, and the last inequality is well known. □

Lemma ⇒ case (A.2) of theorem: Let $Q = \{u \in \mathbf{R}^d\colon -\Upsilon I \preceq \sum_{l=1}^d u_l B_l \preceq \Upsilon I\}$, so that Q is a closed convex set in $\mathbf{R}^d$ such that $\mathrm{Prob}\{\zeta \notin Q\} \leq \chi < 1/2$ by (14.a). We claim that Q contains the centered at the origin $\|\cdot\|_2$-ball of the radius Υ (and thus, by Lemma 3.1.(i)), contains the centered at the origin $\|\cdot\|_2$-ball of the radius $\bar r = \max[\Upsilon, \mathrm{ErfInv}(\chi)]$). Indeed, when $\|u\|_2 \leq \Upsilon$, we have for every $e \in \mathbf{R}^n$:

$$\left\|\sum_{l=1}^d u_l B_l e\right\|_2 \leq \sum_{l=1}^d |u_l|\|B_l e\|_2 \leq \|u\|_2\left[\sum_{l=1}^d \|B_l e\|_2^2\right]^{1/2} \leq \Upsilon\left[\sum_{l=1}^d e^T B_l^2 e\right]^{1/2}$$
$$= \Upsilon\sqrt{e^T\left[\sum_{l=1}^d B_l^2\right]e} \leq \Upsilon\|e\|_2,$$

where the concluding inequality is due to $\sum_{l=1}^d B_l^2 \preceq I$. Thus, whenever $\|u\|_2 \leq \Upsilon$, we have $u \in Q$. We clearly have $\mathrm{Prob}\{-\gamma\Upsilon \preceq \sum_{l=1}^d \zeta_l B_l \preceq \gamma\Upsilon\} = 1 - \mathrm{Prob}\{\zeta \notin \gamma Q\}$; the latter quantity, by Lemma 3.1.(ii) applied with $r = \bar r$, is $\geq 1-\epsilon(\chi,\gamma)$. □

Corollary 3.1. *Given $\epsilon \in (0,1)$, $\Upsilon > 0$, $\chi \in (0,1/2)$, let us set*

$$\vartheta^{-1} = \begin{cases} \Upsilon + 4\sqrt{\ln(\epsilon^{-1}(1-\chi)^{-1})}, & \textit{we are in the case of } (\mathrm{A.1}) \\ \Upsilon + \max[\mathrm{ErfInv}(\epsilon)/\mathrm{ErfInv}(\chi)-1, 0]\min[\Upsilon, \mathrm{ErfInv}(\chi)], & \textit{we are in the case of } (\mathrm{A.2}). \end{cases} \tag{18}$$

Assume, further, that symmetric matrices $A_0, \ldots, A_d$ satisfy

$$\mathrm{Arrow}(\vartheta A_0, A_1, \ldots, A_d) \succeq 0 \tag{19}$$

and, in addition, that

$$\mathrm{Prob}\left\{-\Upsilon[\vartheta A_0] \preceq \sum_{l=1}^{d} \zeta_l A_l \preceq \Upsilon[\vartheta A_0]\right\} \geq 1-\chi. \tag{20}$$

Then

$$\mathrm{Prob}\left\{-A_0 \preceq \sum_{l=1}^{d} \zeta_l A_l \preceq A_0\right\} \geq 1-\epsilon. \tag{21}$$

Proof. Relations (19) and (20) imply that the matrices $B_0 = \vartheta A_0, B_1 = A_1, \ldots, B_d = A_d$ satisfy (14). It remains to apply Theorem 3.1 to the just-defined $B_0, B_1, \ldots, B_d$ and to $\gamma = (\vartheta\Upsilon)^{-1}$ and to note that with this γ one has $\epsilon(\chi,\gamma) \leq \epsilon$. □

3.2. The approximation. Our proposed way to process (7) is as follows.

1. **Building the approximation.** We start with somehow choosing parameters $\Upsilon > 0$, $\chi \in (0, 1/2)$ and act *as if* we were sure that whenever symmetric $n \times n$ matrices $B_0, \ldots, B_d$ satisfy

$$\mathrm{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0, \tag{22}$$

then they satisfy the relation

$$\mathrm{Prob}\left\{-\Upsilon B_0 \preceq \sum_{l=1}^{d} \zeta_l B_l \preceq \Upsilon B_0\right\} \geq 1-\chi. \tag{23}$$

Specifically, we replace the chance constraint (8) in (7) with the LMI

$$\mathrm{Arrow}(\vartheta\mathcal{A}_0[x], \mathcal{A}_1[x], \ldots, \mathcal{A}_d[x]) \succeq 0, \tag{24}$$

where ϑ is given by (18), and process the resulting optimization problem, arriving at its feasible solution x_*.

Let us set $B_0^* = \vartheta\mathcal{A}_0[x_*], B_1^* = \mathcal{A}_1[x_*], \ldots, B_d^* = \mathcal{A}_d[x_*]$; by construction, these matrices satisfy (22). *If these matrices satisfy* (23) *as well*, then by Corollary 3.1, x_* is a feasible solution to the chance-constrained problem (7). The difficulty, however, is that *unless we can prove that for* Υ, χ *in question, relation* (22) *always implies relation* (23)*, we cannot be sure in advance that the matrices* B_l^* *satisfy* (23) *and, consequently, cannot be sure that* x_* *is feasible for the chance-constrained problem* (7).

In order to overcome this difficulty, we use the following *validation procedure*.

2. **Validation procedure.** We generate a training sample of N independent realizations $\zeta^1, \ldots, \zeta^N$ of ζ and compute the number M of realizations for which the relation $-\Upsilon B_0^* \preceq \sum_{l=1}^{d} \zeta_l^i B_l^* \preceq \Upsilon B_0^*$ is *not* satisfied. We then use these statistics to get a $(1-\delta)$-reliable *lower bound* π on the probability $p_* = \mathrm{Prob}\{-\Upsilon B_0^* \preceq \sum_{l=1}^{d} \zeta_l^i B_l^* \preceq \Upsilon B_0^*\}$, specifically, set

$$\pi = \min_{0 \leq p \leq 1}\left\{p\colon \sum_{i=0}^{M} \binom{N}{i} (1-p)^i p^{N-i} \geq \delta\right\},$$

where $\delta \in (0,1)$ is a chosen in advance "unreliability level" (say, $\delta = 10^{-12}$). We then check whether $\pi \geq 1-\chi$; if it is the case, we claim that the feasibility of x_* for the problem of interest (7) is validated. Otherwise, we apply our approximation scheme anew, increasing the value of Υ and/or the value of χ.

Proposition 3.1. *For the outlined randomized approximation procedure, the probability of* x_* *being validated when in fact it is infeasible for* (7) *is at most* δ.

Proof. It is easily seen that the random quantity π is, with probability at least $1-\delta$, a lower bound on p_*. Thus, the probability of validating the feasibility of x_* in the case when $p_* < 1-\chi$ is at most δ; because x_* is provably feasible for (7), in the case of $p_* \geq 1-\chi$, is indeed safe up to probability of bad sampling $\leq \delta$. □

The advantage of the outlined validation routine is that *when working with* χ *not too close to* 0 (and we can afford to work with any $\chi \in (0, 1/2)$, say, $\chi = 0.25$ or $\chi = 0.1$), *in the case of*

$$\mathrm{Prob}\left\{-\Upsilon B_0^* \preceq \sum_{l=1}^{d} \zeta_l B_l^* \preceq \Upsilon B_0^*\right\} \geq 1-0.8\chi \tag{25}$$

(that is, validating an assumption

$$\mathrm{Prob}\left\{-\Upsilon B_0^* \preceq \sum_{l=1}^{d} \zeta_l B_l^* \preceq \Upsilon B_0^*\right\} \geq 1-\chi$$

slightly stronger than the one we wish to validate) *the cardinality* N *of the sample which is sufficient to validate, the feasibility of* x_* *for* (7) *with probability* $1-\nu$ *close to* 1 *should not be too large.* A rough estimate shows that it suffices to take

$$N \geq 100(\ln(1/\delta)+\ln(1/\nu))\chi^{-2}.$$

With $\delta=\nu=10^{-8}$, $\chi=0.25$, this formula yields $N=58{,}947$; a more accurate computation shows that $N=8{,}750$ also will do. It should be stressed that the sample size in question is completely independent of ϵ, which therefore can be arbitrarily small; this is in sharp contrast to what would happen if we were checking the fact that x_* is feasible for (8) by trying to estimate $p(x_*)$ (see (8)) by a straightforward Monte Carlo simulation in order to understand whether indeed $p(x_*)\geq 1-\epsilon$. Such a simulation would require a sample of cardinality $\geq O(1/\epsilon)$ and would therefore be completely impractical when ϵ is small, like 10^{-6} or less.

3.3. A modification. In many applications, it makes sense to pose problem (7) in a slightly different form, specifically, as the problem

$$\rho_*(\bar{c})=\max_{x,\rho}\left\{\rho\colon \begin{array}{l} F(x)\leq 0,\ c^Tx\leq\bar{c},\\ \mathrm{Prob}\left\{\mathcal{A}_0[x]+\rho\sum_{l=1}^d\zeta_l\mathcal{A}_l[x]\succeq 0\right\}\geq 1-\epsilon\end{array}\right\}, \tag{26}$$

Thus, instead of minimizing the value of the objective under the deterministic constraints and the chance constraint with the "reference" uncertainty level $\rho=1$, we are now maximizing the uncertainty level ρ for which the chance-constrained problem admits a feasible solution with the value of the objective $\leq\bar{c}$. In reality, we could, e.g., start with solving the "nominal" problem

$$\mathrm{Opt}=\min_x\{c^Tx\colon F(x)\leq 0,\ \mathcal{A}_0[x]\succeq 0\},$$

and then build the "trade-off curve" $\tau(s)=\rho_*(\mathrm{Opt}+s)$, $s>0$, which shows which uncertainty level could be tolerated given a "sacrifice" $s>0$ in the optimal value.

The advantage of (26) in our context is that here the safe, tractable approximation given by our approach does not require any a priori guess of Υ, χ. Indeed, assume that we start with certain Υ, χ which, we believe, ensure the validity of the implication "(22) $\Rightarrow$ (23)." Acting in exactly the same fashion as above, but aiming at the problem (26) rather than at the problem (7), we would arrive at the approximation

$$\max_{x,\rho}\left\{\rho\colon \begin{array}{l} F(x)\leq 0,\quad c^Tx\leq\bar{c},\\ \mathrm{Arrow}(\vartheta(\rho)\mathcal{A}_0[x],\mathcal{A}_1[x],\ldots,\mathcal{A}_d[x])\succeq 0\end{array}\right\} \tag{27}$$

where $\vartheta(\rho)$ is given by (18) with Υ replaced with $\rho\Upsilon$. Because $\vartheta(\rho)$ clearly decreases as ρ grows, we see that *as far as the x-component of an optimal solution to the resulting problem is concerned, this component is independent of our guesses* Υ, χ *and coincides with the x-component of the optimal solution to the quasi-convex* (and thus efficiently solvable) *optimization problem*

$$\min_{x,\vartheta}\left\{\vartheta\colon \begin{array}{l} F(x)\leq 0,\quad c^Tx\leq\bar{c},\quad \vartheta\geq 0,\quad \mathcal{A}_0[x]\succeq 0,\\ \mathrm{Arrow}(\vartheta\mathcal{A}_0[x],\mathcal{A}_1[x],\ldots,\mathcal{A}_d[x])\succeq 0\end{array}\right\}. \tag{28}$$

The fact that the resulting approximation is independent of any guess on Υ and χ does not resolve all of our difficulties—we still need to say what is the "feasibility radius" $\rho^*(x_*)$ of an optimal (or nearly so) solution x_* to (28), which we get when solving the latter problem, that is, what is the largest $\rho=\rho^*(x_*)$ such that

$$\mathrm{Prob}\left\{-\mathcal{A}_0[x_*]\preceq\rho\sum_{l=1}^d\zeta_l\mathcal{A}_l[x_*]\preceq\mathcal{A}_0[x_*]\right\}\geq 1-\epsilon. \tag{29}$$

Assume that x_* can be extended by certain ϑ to a feasible solution to (28). If the guess we started with were true, we could take as $\rho_+(x_*)$ the supremum of those $\rho>0$ for which $\vartheta(\rho)\geq\vartheta_*(x_*)$, where $\vartheta_*(x_*)$ is the smallest $\vartheta\geq 0$ such that $\mathrm{Arrow}(\vartheta\mathcal{A}_0[x_*],\mathcal{A}_1[x_*],\ldots,\mathcal{A}_d[x_*])\succeq 0$ (when x_* is an optimal solution to (28), $\vartheta_*(x_*)$ is exactly the optimal value in (28)). In the case when we are not sure that our guess is true, we can build a lower bound $\rho_*(x_*)$ on $\rho^*(x_*)$ via an appropriate modification of the validation procedure, specifically, as follows.

Assume that $\vartheta_*(x_*)>0$ (this is the only nontrivial case, because $\vartheta_*(x_*)=0$ means that $\mathcal{A}_l[x_*]=0$, $l=1,\ldots,d$; because $\mathcal{A}_0[x_*]\succeq 0$ due to the constraints in (28), in this case we clearly have $\rho^*(x_*)=+\infty$). Let us use the following.

Calibration procedure. Given x_*, $\vartheta_*(x_*) > 0$, let $B_0 = \vartheta_*(x_*)\mathcal{A}_0[x_*]$, $B_l = \mathcal{A}_l[x_*]$, $l = 1, \ldots, d$ satisfy $\text{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0$. Let, further, $\delta \in (0, 1)$ be a desired "unreliability level" of our conclusions (cf. the Validation procedure). We now carry out the following two steps:

1. *Building a grid of values of* ρ. As we remember from §3.1, the implication (22) ⇒ (23) indeed holds true for "safe" values of Υ and χ, e.g., for $\chi = \chi_s = 0.25$ and $\Upsilon = \Upsilon_s = O(1)\sqrt{\ln n}$ with appropriately chosen $O(1)$. From Corollary 3.1 it follows that if ϑ_s is given by (18) with $\chi = \chi_s$ and $\Upsilon = \Upsilon_s$, then, setting

$$\rho_s = \vartheta_s/\vartheta_*(x_*),$$

we have

$$\text{Prob}\left\{-\mathcal{A}_0[x_*] \preceq \rho_s \sum_{l=1}^{d} \zeta_l \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\right\} \geq 1 - \epsilon. \tag{30}$$

Indeed, the matrices $B_0, \ldots, B_d$ satisfy (22) and therefore satisfy (23) with $\chi = \chi_s$, $\Upsilon = \Upsilon_s$. Applying Corollary 3.1 to the matrices $A_0 = \vartheta_s^{-1} B_0 = \vartheta_s^{-1}\vartheta_*(x_*)\mathcal{A}_0[x_*] = \rho_s^{-1}\mathcal{A}_0[x_*]$, $A_l = B_l = \mathcal{A}_l[x_*]$, $l = 1, \ldots, d$, we conclude that (30) indeed holds true.

Now let us find $\rho^+ \geq \rho_s$ such that the relation

$$\text{Prob}\left\{-\mathcal{A}_0[x_*] \preceq \rho^+ \sum_{l=1}^{d} \zeta_l \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\right\} \geq 1 - \epsilon$$

is "highly unlikely" to be true. For example, assuming $\epsilon \ll 1/2$, we can generate a short (say, with $L = 100$ elements) pilot sample of realizations $\zeta^1, \ldots, \zeta^L$ of ζ; compute, for every $i \leq L$, the largest $\rho = \rho^i$ such that the relation

$$-\mathcal{A}_0[x_*] \preceq \rho^i \sum_{l=1}^{d} \zeta_l^i \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]$$

holds true; and take, as ρ^+, the maximum of ρ_s and of the median of $\{\rho^1, \ldots, \rho^L\}$.

Finally, we insert into the segment $[\rho_s, \rho^+]$ a moderate number $(K-2)$ of "intermediate" values of ρ, say, in such a way that the resulting sequence $r_1 := \rho_s < r_2 < \cdots < r_K := \rho^+$ forms a geometric progression. This sequence forms a grid that we are about to use when building $\rho_*(x_*)$.

2. *Running simulations.* At this step, we
   (i) Generate a training sample of N independent realizations $\zeta^1, \ldots, \zeta^N$ of ζ.
   (ii) For every $k = 1, \ldots, K$ compute the integers

$$M_k = \text{Card}\left\{i \leq N\colon \neg\left(-\mathcal{A}_0[x_*] \preceq r_k \sum_{l=1}^{d} \zeta_l^i \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\right)\right\}$$

and then the reals

$$\hat{\chi}_k = \max\left\{\chi \in [0,1]\colon \sum_{i=1}^{M_k} \binom{N}{i} \chi^i (1-\chi)^{N-i} \geq \delta/K\right\}.$$

Note that if

$$\chi_k = \text{Prob}\left\{\neg\left(-\mathcal{A}_0[x_*] \preceq r_k \sum_{l=1}^{d} \zeta_l \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\right)\right\},$$

then the probability for the random quantity $\hat{\chi}_k$ to be $< \chi_k$ is at most δ/K, so that

$$\text{Prob}\{\hat{\chi}_k \geq \chi_k,\ 1 \leq k \leq K\} \geq 1 - \delta. \tag{31}$$

3. *Specifying* $\rho_*(x_*)$. In the case of (A.1) we set

$$\rho_*(x_*) = \max_{1\leq k\leq K}\left\{\frac{r_k}{1 + 4r_k\vartheta_*(x_*)\sqrt{\ln(\epsilon^{-1}(1-\hat{\chi}_k)^{-1})}}\colon \hat{\chi}_k < 1/2\right\}, \tag{32}$$

and in the case of (A.2) we set

$$\rho_*(x_*) = \max_{1\leq k\leq K}\left\{\frac{r_k}{1 + \max[\text{ErfInv}(\epsilon)/\text{ErfInv}(\hat{\chi}_k) - 1, 0]\min[r_k\vartheta_*(x_*)\text{ErfInv}(\hat{\chi}_k), 1]}\colon \hat{\chi}_k < 1/2\right\}. \tag{33}$$

If these formulas are not well defined (e.g., there is no k such that $\hat{\chi}_k < 1/2$) or are well defined, but result in $\rho_*(x_*) < \rho_s$, we set $\rho_*(x_*)$ to the "safe" value ρ_s.

Note that the quantity $\rho_*(x_*)$ yielded by the calibration procedure is random.

Proposition 3.2. *Let $(x_*, \vartheta_*(x_*) > 0)$ be feasible for (28). Then, with the outlined calibration procedure, the probability for $(x_*, \rho_*(x_*))$ to be infeasible for (26) is $\leq \delta$.*

Proof. Assume that $\hat{\chi}_k \geq \chi_k$ for all $k = 1, \ldots, K$ (recall that this condition is valid with probability $\geq 1-\delta$), and let us prove that in this case $(x_*, \rho_*(x_*))$ is feasible for (26). We already know that this is the case when $\rho_* \equiv \rho_*(x_*) = \rho_s$, so that we can restrict ourselves with the case when $\rho_*(x_*)$ is given by a well-defined formula ((32) in the case of (A.1) or (33) in the case of (A.2)).

In the case of (A.1), let k be such that $\hat{\chi}_k < 1/2$ and $\rho_* = r_k/(1+4r_k\vartheta_*(x_*)\sqrt{\ln(\epsilon^{-1}(1-\hat{\chi}_k)^{-1})})$ (see (32)), and let

$$\Upsilon_k = \frac{1}{r_k \vartheta_*(x_*)}, \qquad \vartheta_k = \frac{1}{\Upsilon_k + 4\sqrt{\ln(\epsilon^{-1}(1-\hat{\chi}_k)^{-1})}}, \qquad A_0 = \frac{\vartheta_*(x_*)}{\vartheta_k}\mathcal{A}_0[x_*], \qquad A_l = \mathcal{A}_l[x_*], \quad l = 1, \ldots, d.$$

Then

$$\text{Arrow}(\vartheta_k A_0, A_1, \ldots, A_d) = \text{Arrow}(\vartheta_*(x_*)\mathcal{A}_0[x_*], \mathcal{A}_1[x_*], \ldots, \mathcal{A}_d[x_*]) \succeq 0,$$

$$\text{Prob}\Bigg\{-\underbrace{\Upsilon_k \vartheta_k A_0}_{r_k^{-1}\mathcal{A}_0[x_*]} \preceq \underbrace{\sum_{l=1}^{d}\zeta_l A_l}_{=\sum_{l=1}^{d}\zeta_l \mathcal{A}_l[x_*]} \preceq \Upsilon_k \vartheta_k A_0\Bigg\} = \text{Prob}\left\{-\mathcal{A}_0[x_*] \preceq r_k \sum_{l=1}^{d}\zeta_l \mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\right\}$$
$$\geq 1 - \hat{\chi}_k$$

where the concluding inequality is valid due to the fact that we are in the case of $\hat{\chi}_k \geq \chi_k$. Invoking Corollary 3.1, we conclude that $\text{Prob}\{-A_0 \preceq \sum_{l=1}^{d}\zeta_l A_l \preceq A_0\} \geq 1-\epsilon$, or, which is the same (due to $A_0 = (\vartheta_*(x_*)/\vartheta_k)\mathcal{A}_0[x_*] = (1/\rho_*)\mathcal{A}_0[x_*]$) as $\text{Prob}\{-\mathcal{A}_0[x_*] \preceq \rho_* \sum_{l=1}^{d}\mathcal{A}_l[x_*] \preceq \mathcal{A}_0[x_*]\} \geq 1-\epsilon$, as claimed.

The result for case (A.2) can be proved in a completely similar way. □

4. Special cases: Diagonal and arrow matrices. In this section, we consider two special cases where the chance-constrained LMI in (7) possesses a specific structure which, in principle, allows us to point out "moderate" Υ and χ which make valid the implication "(22) ⇒ (23)," that is, the implication

$$\text{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0 \;\Rightarrow\; \text{Prob}\left\{-\Upsilon B_0 \preceq \sum_{l=1}^{d}\zeta_l B_l \preceq \Upsilon B_0\right\} \geq 1-\chi. \tag{34}$$

In particular, using these Υ, χ in the approximation scheme of §3.2, we can avoid the necessity of using the validating procedure.

4.1. Diagonal case. The first special case we consider is where $\mathcal{A}_0[x], \mathcal{A}_1[x], \ldots, \mathcal{A}_d[x]$ in (7) are diagonal matrices. We refer to this situation as the *diagonal case*. Note that in spite of its simplicity, this case is of definite interest: It is the case of chance-constrained system of *linear inequalities*—the entity of primary interest for chance-constrained linear programming. We start with the following observation:

Lemma 4.1. *Let $\zeta \in \mathbf{R}^d$ be a random vector and $B_l = \text{Diag}\{B_l^1, \ldots, B_l^s\}$, $l = 0, 1, \ldots, d$, be block-diagonal matrices of common block-diagonal structure. Assume that for certain function $\Upsilon(\chi)$, $\chi \in (0, 1/2)$, and every $j \leq s$ the structure of the blocks B_l^j ensures the implication*

$$\forall \chi \in (0, 1/2): \quad \text{Arrow}(B_0^j, \ldots, B_d^j) \succeq 0 \;\Rightarrow\; \text{Prob}\left\{-\Upsilon(\chi)B_0^j \preceq \sum_{l=1}^{d}\zeta_l B_l^j \preceq \Upsilon(\chi)B_0^j\right\} \geq 1-\chi.$$

Then one has

$$\forall \chi \in (0, 1/2): \quad \text{Arrow}(B_0, \ldots, B_d) \succeq 0 \;\Rightarrow\; \text{Prob}\left\{-\Upsilon(\chi/s)B_0 \preceq \sum_{l=1}^{d}\zeta_l B_l \preceq \Upsilon(\chi/s)B_0\right\} \geq 1-\chi.$$

This statement is an immediate consequence of the fact that $\text{Arrow}(B_0, \ldots, B_d) \succeq 0$ if and only if $\text{Arrow}(B_0^j, \ldots, B_d^j) \succeq 0$ for every $j = 1, \ldots, s$.

Theorem 4.1. *Let* $B_0, B_1, \ldots, B_d$ *be diagonal* $n \times n$ *matrices satisfying* $\mathrm{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0$, *and* $\zeta_1, \ldots, \zeta_d$ *be random variables satisfying the assumption*:

(A.3) $\zeta_1, \ldots, \zeta_d$ *are independent, all with zero mean, and* $\mathbf{E}\{\exp\{\zeta_l^2\}\} \le \exp\{1\}$, $1 \le l \le d$

(*note that* (A.3) *is implied by* (A.1)). *Then the implication* (34) *holds true for every* $\chi \in (0, 1/2)$ *with*

$$\Upsilon = \Upsilon^{(n)}(\chi) = \frac{1}{3}\sqrt{38\ln(2n/\chi)}.$$

If, in addition to (A.3), *the entries in* ζ *are symmetrically distributed, then the above conclusion remains valid with*

$$\Upsilon = \Upsilon_{\mathrm{S}}^{(n)}(\chi) = \sqrt{3\ln(2n/\chi)}.$$

Finally, if ζ *satisfies* (A.2), *then the same conclusion remains valid with*

$$\Upsilon = \Upsilon_{\mathrm{G}}^{(n)}(\chi) = \mathrm{ErfInv}(\chi/(2n)) \le \sqrt{2\ln(n/\chi)}.$$

Proof. By Lemma 4.1, it suffices to prove the statement in the scalar case $n = 1$, where the relation $\mathrm{Arrow}(B_0, \ldots, B_d) \succeq 0$ means simply that $B_0 \ge \sqrt{\sum_{l=1}^d B_l^2}$. There is nothing to prove when $B_0 = 0$; assuming $B_0 > 0$ and setting $h_l = B_l/B_0$, all we need is to prove that whenever ζ satisfies (A.3) and $h \in \mathbf{R}^d$ is deterministic, then

$$\|h\|_2 \le 1 \;\Rightarrow\; \mathrm{Prob}\left\{\left|\sum_{l=1}^d \zeta_l h_l\right| > \Upsilon(\chi)\right\} \le \chi, \quad 0 < \chi < 1/2, \tag{35}$$

where $\Upsilon(\cdot)$, depending on the situation, is either $\Upsilon^{(1)}(\cdot)$, or $\Upsilon_{\mathrm{S}}^{(1)}(\cdot)$, or $\Upsilon_{\mathrm{G}}^{(1)}(\cdot)$. This result is readily given by standard facts on large deviations; to make the presentation self-contained, here is the demonstration. All we need is to prove that if $h \in \mathbf{R}^d$, $\|h\|_2 \le 1$, then

$$\forall\, \Upsilon > 0: \quad \mathrm{Prob}\left\{\left|\sum_{l=1}^d h_l \zeta_l\right| > \Upsilon\right\} \le \begin{cases} 2\exp\{-9\Upsilon^2/38\}, & \zeta \text{ satisfies (A.3)} \\ 2\exp\{-\Upsilon^2/3\}, & \zeta \text{ satisfies (A.3) and is symmetrically distributed} \\ 2\Phi(\Upsilon), & \zeta \sim \mathcal{N}(0, I_d), \end{cases} \tag{36}$$

where $\Phi(s) = \int_s^\infty (2\pi)^{-1/2}\exp\{-r^2/2\}\,dr$ is the error function.

The case of $\zeta \sim \mathcal{N}(0, I_d)$ is evident. Now assume that ζ satisfies (A.3). Let $\gamma \in \mathbf{R}$, $s_l = \sum_{r=1}^l \gamma h_r \zeta_r$, and $J = \{l\colon |h_l\gamma| > \sqrt{3/2}\}$. We have

$$\mathbf{E}\{\exp\{s_l\}\} = \mathbf{E}\{\exp\{s_{l-1}\}\exp\{\gamma h_l \zeta_l\}\} = \mathbf{E}\{\exp\{s_l\}\}\cdot\Theta_l, \quad \Theta_l = \mathbf{E}\{\gamma h_l \zeta_l\} \tag{37}$$

(we have taken into account that ζ_l is independent of s_{l-1}). We claim that

$$\Theta_l \le \begin{cases} \exp\{2\gamma^2 h_l^2/3\}, & l \notin J \\ \exp\{7/12 + 2\gamma^2 h_l^2/3\}, & l \in J. \end{cases} \tag{38}$$

Indeed, it is easily seen that

$$\exp\{t\} \le t + \exp\{2t^2/3\}$$

for all $t \in \mathbf{R}$, whence $\mathbf{E}\{\exp\{\gamma h_l \zeta_l\}\} \le \mathbf{E}\{\exp\{2\gamma^2 h_l^2 \zeta_l^2/3\}\}$; when $l \notin J$, the latter expectation is at most $(\mathbf{E}\{\exp\{\zeta_l^2\}\})^{2\gamma^2 h_l^2/3}$ by Hölder inequality, as required in (38). Now let $l \in J$. We have $|\gamma h_l s| \le s^2 + \gamma^2 h_l^2/4$ for all s, whence

$$\begin{aligned}\mathbf{E}\{\exp\{\gamma h_l \zeta_l\}\} &\le \exp\{\gamma^2 h_l^2/4\}\mathbf{E}\{\exp\{\zeta_l^2\}\} \le \exp\{1 + \gamma^2 h_l^2/4\} \\ &\le \exp\{7/12 + 2\gamma^2 h_l^2/3\}\end{aligned}$$

as required in (38).

Combining (37) and (38), we get

$$\begin{aligned}\mathbf{E}\left\{\exp\left\{\gamma\sum_{l=1}^d h_l\zeta_l\right\}\right\} &\le \exp\left\{2\gamma^2\left[\sum_{l=1}^d h_l^2\right]\Big/3\right\}\exp\{(7/12)\mathrm{Card}(J)\} \\ &\le \exp\{2\gamma^2/3\}\exp\{(7/12)\cdot(2/3)\cdot\gamma^2\},\end{aligned}$$

where the concluding inequality follows from the facts that $\|h\|_2 \le 1$ and that $h_l^2 > 3/(2\gamma^2)$ when $l \in J$, which combines with $\|h\|_2 \le 1$ to imply that $\mathrm{Card}(J) \le 2\gamma^2/3$. Thus,

$$\mathbf{E}\left\{\exp\left\{\gamma\sum_{l=1}^d h_l\zeta_l\right\}\right\} \le \exp\{19\gamma^2/18\},$$

whence, by Tschebyshev inequality,

$$\mathrm{Prob}\left\{\left|\sum_{l=1}^d h_l\zeta_l\right| > \Upsilon\right\} \le 2\min_{\gamma>0}\exp\{19\gamma^2/18 - \gamma\Upsilon\} = 2\exp\{-9\Upsilon^2/38\}.$$

Now let ζ satisfy (A.3) and be symmetrically distributed. For $\gamma > 0$, let us set $s_l = \cosh(\gamma\sum_{r=1}^l h_r\zeta_r)$. Then

$$\mathbf{E}\{s_l\} = \mathbf{E}\left\{s_{l-1}\cosh(\gamma h_l\zeta_l) + \sinh\left(\gamma\sum_{r=1}^{l-1}h_r\zeta_r\right)\sinh(\gamma h_l\zeta_l)\right\} = \mathbf{E}\{s_{l-1}\}\underbrace{\mathbf{E}\{\cosh(\gamma h_l\zeta_l)\}}_{\Theta_l},$$

whence

$$\mathbf{E}\{s_d\} = \Theta_1 \cdot \ldots \cdot \Theta_d.$$

Setting $J = \{l\colon\ \gamma^2 h_l^2 \le 2\}$ and taking into account that $\cosh(t) \le \exp\{t^2/2\}$ for all t, for $l \notin J$ we have

$$\Theta_l = \mathbf{E}\{\cosh(\gamma h_l\zeta_l)\} \le \mathbf{E}\{\exp\{\gamma^2h_l^2\zeta_l^2/2\}\} \le \exp\{\gamma^2h_l^2/2\},$$

where the concluding inequality is given by the facts that $\gamma^2h_l^2/2 \le 1$ and $\mathbf{E}\{\exp\{\zeta_l^2\}\} \le \exp\{1\}$ in view of the Hölder inequality. When $l \in J$, we, the same as above, have

$$\cosh(\gamma h_l\zeta_l) \le \exp\{|\gamma h_l\zeta_l|\} \le \exp\{\zeta^2 + \gamma^2h_l^2/4\},$$

whence $\Theta_l \le \exp\{1 + \gamma^2h_l^2/4\} \le \exp\{1/2 + \gamma^2h_l^2/2\}$. We therefore get

$$\mathbf{E}\left\{\cosh\left(\gamma\sum_{l=1}^d h_l\zeta_l\right)\right\} \le \exp\left\{\gamma^2\left[\sum_{l=1}^d h_l^2\right]\Big/2\right\}\exp\{\mathrm{Card}(J)/2\},$$

and, similarly to the previous case, $\mathrm{Card}(J) \le \gamma^2/2$, whence

$$\mathbf{E}\left\{\cosh\left(\gamma\sum_{l=1}^d h_l\zeta_l\right)\right\} \le \exp\{3\gamma^2/4\}.$$

When $|\sum_{l=1}^d h_l\zeta_l| > \Upsilon$, we have $\cosh(\gamma\sum_{l=1}^d h_l\zeta_l) > \exp\{\gamma\Upsilon\}/2$, so that

$$\mathrm{Prob}\left\{\left|\sum_{l=1}^d h_l\zeta_l\right| > \Upsilon\right\} \le 2\inf_{\gamma>0}\exp\{3\gamma^2/4 - \gamma\Upsilon\} = 2\exp\{-\Upsilon^2/3\},$$

as required in (36). □

Comparison with other approximations of a chance-constrained LP. As mentioned earlier, the diagonal case arises when solving chance-constrained linear programming problems that we prefer to pose in the form of (26):

$$\max_{x,\rho}\left\{\rho\colon\ \begin{array}{l} Fx - f \ge 0, \quad c^Tx \le \bar{c} \\ \mathrm{Prob}\{A_\zeta x - b_\zeta \ge 0\} \ge 1-\epsilon \end{array}\right\}, \quad [A_\zeta, b_\zeta] = [A^0, b^0] + \rho\sum_{l=1}^d \zeta_l[A^l, b^l]$$
$$\Updownarrow$$
$$\max_{x,\rho}\left\{\rho\colon\ \begin{array}{l} Fx - f \ge 0, \quad c^Tx \le \bar{c} \\ \mathrm{Prob}\left\{\mathcal{A}_0[x] + \rho\sum_{l=1}^d \zeta_l\mathcal{A}_l[x] \succeq 0\right\} \ge 1-\epsilon \end{array}\right\}, \quad \mathcal{A}_l[x] = \mathrm{Diag}\{A^lx - b^l\}, \quad 0 \le l \le d. \tag{39}$$

With our approximation scheme, the safe, tractable approximation of the resulting chance-constrained problem is, as it is immediately seen, the quasi-convex program

$$\max_{x,\rho}\left\{\rho:\quad \begin{array}{l} Fx-f\geq 0,\quad c^Tx\leq \bar c\\ \rho\sqrt{\sum_{l=1}^d\left[b_i^l+\sum_{j=1}^J A_{ij}^l x_j\right]^2}\leq \sum_j A_{ij}^0x_j-b_i^0,\quad 1\leq i\leq I\end{array}\right\} \tag{40}$$

where I, J are the row and the column sizes of A^l. There also exists a more traditional "constraint-by-constraint" way to process a chance constrained LP; specifically, we somehow choose positive ϵ_i, $\sum_i\epsilon_i=\epsilon$, and safely approximate (39) with the chance-constrained problem

$$\max_{x,\rho}\left\{\rho:\quad \begin{array}{l} Fx-f\geq 0,\quad c^Tx\leq \bar c\\ \mathrm{Prob}\left\{\sum_j A_{ij}^0x_j-b_i^0+\rho\sum_{l=1}^d\zeta_l\left[\sum_j A_{ij}^lx_j-b_i^l\right]\geq 0\right\}\geq 1-\epsilon_i,\quad 1\leq i\leq I\end{array}\right\}. \tag{41}$$

This problem involves chance-constrained *scalar* linear inequalities that are much easier to approximate than the original chance-constrained vector inequality appearing in (39). For the sake of simplicity, consider the case when $\zeta\sim\mathcal{N}(0,I)$ and $\epsilon<1/2$. In this case, (41) is *exactly equivalent* to the explicit quasi-convex problem

$$\max_{x,\rho}\left\{\rho:\quad \begin{array}{l} Fx-f\geq 0,\quad c^Tx\leq \bar c\\ \mathrm{ErfInv}(\epsilon_i)\rho\sqrt{\sum_{l=1}^d\left[b_i^l+\sum_{j=1}^J A_{ij}^l x_j\right]^2}\leq \sum_j A_{ij}^0x_j-b_i^0,\quad 1\leq i\leq I\end{array}\right\}. \tag{42}$$

Note that an attempt to treat the parameters ϵ_i of our construction as decision variables in (42) fails—the resulting problem loses convexity; this is why the parameters ϵ_i should be chosen in advance, and the most natural way to choose them is to set $\epsilon_i=\epsilon/I$, $i=1,\dots,I$. Note that *with this choice of* ϵ_i*, problem* (42) *is equivalent to* (40), up to rescaling $\rho\mapsto\rho/\mathrm{ErfInv}(\epsilon/I)$. This, however, does not mean that the approximations are identical; although both of them lead to the same optimal decision vector x_*, they differ in what is the resulting lower bound ρ_* on the true feasibility radius $\rho^*(x_*)$ of x_* (recall that this radius is the largest ρ for which (x_*,ρ) is feasible for the chance-constrained problem of interest (39)). Specifically, for approximation (42), ρ_* is exactly the optimal value of the approximation, whereas for (40) ρ_* is given by the calibration routine. Experiments show that which of these two lower bounds is less conservative depends on the problem's data, so that in practice it makes sense to build both these bounds and to use the larger of them.

4.2. Arrow case. We are about to justify the implication (34) in the *Arrow case*, where the matrices B_l, $l=1,\dots,d$, are of the form

$$B_l=[ef_l^T+f_le^T]+\lambda_lG, \tag{43}$$

where e, $f_l\in\mathbf{R}^n$, $\lambda_l\in\mathbf{R}$, and $G\in\mathbf{S}^n$. We meet this case in the chance-constrained conic quadratic optimization; see (6). Indeed, the matrices $\mathcal{A}_l[x]$, $1\leq l\leq d$, arising in (6) are, for every x, matrices of the form (43). Therefore, all we need when building and processing the safe tractable approximation, as developed in §3.2 for the chance-constrained LMI in (6), is the validity of (34) for matrices B_l of the form (43).

THEOREM 4.2. *Let the* $n\times n$ *matrices* $B_1,\dots,B_d$ *of the form* (43) *along with a matrix* $B_0\in\mathbf{S}^n$ *satisfy the premise in* (34)*. Let, further,* $\zeta_1,\dots,\zeta_d$ *be independent random variables with zero means and such that* $\mathbf{E}\{\zeta_l^2\}\leq\sigma^2$, $l=1,\dots,d$ *(note that in the cases of* (A.1) *and* (A.2)*, one can take* $\sigma=1$*, and in the case of* (A.3) *one can take* $\sigma=\sqrt{\exp\{1\}-1}$*). Then, for every* $\chi\in(0,1/2)$ *and with* $\Upsilon(\chi)$ *given by*

$$\begin{array}{lll}\text{(a)} & 2\sigma\sqrt{2/\chi} & [\textit{general case}]\\ \text{(b)} & \min[2\sqrt{2/\chi},\,4+4\sqrt{\ln(2/\chi)}] & [\textit{case } (A.1)]\\ \text{(c)} & 4+\mathrm{ErfInv}(\chi) & [\textit{case } (A.2)]\end{array} \tag{44}$$

one has

$$\Upsilon\geq\Upsilon(\chi)\ \Rightarrow\ \mathrm{Prob}\left\{-\Upsilon B_0\preceq\sum_{l=1}^d\zeta_lB_l\preceq\Upsilon B_0\right\}\geq 1-\chi, \tag{45}$$

that is, with our $\Upsilon(\chi)$*, the implication in* (34) *holds true.*

Proof. First of all, when ζ_l, $l=1,\ldots,d$, satisfy (A.3), we indeed have $\mathbf{E}\{\zeta_l^2\}\le\exp\{1\}-1$ due to $t^2\le\exp\{t^2\}-1$ for all t. Further, by continuity argument, it suffices to consider the case where

$$\mathrm{Arrow}(B_0,B_1,\ldots,B_d)\succeq 0 \quad\text{and}\quad B_0\succ 0.$$

In this case, setting $A_l=B_0^{-1/2}B_lB_0^{-1/2}$, the relation $\mathrm{Arrow}(B_0,\ldots,B_d)\succeq 0$ is equivalent to $\sum_{l=1}^d A_l^2\preceq I$, and the target relation (45) is equivalent to

$$\Upsilon\ge\Upsilon(\chi)\ \Rightarrow\ \mathrm{Prob}\left\{-\Upsilon I_n\preceq\sum_{l=1}^d\zeta_lA_l\preceq\Upsilon I_n\right\}\ge 1-\chi$$

with $\Upsilon(\chi)$ given by (44). Thus, all we need to prove is the following.

Lemma 4.2. *Let B_l, $l=1,\ldots,d$, be of the form of (43), let $B_0\succ 0$, and let the matrices $A_l=B_0^{-1/2}B_lB_0^{-1/2}$ satisfy $\sum_l A_l^2\preceq I$. Let ζ_l, further, satisfy the premise in Theorem 4.2. Then, for every $\chi\in(0,1/2)$, one has*

$$\mathrm{Prob}\left\{\left\|\sum_{l=1}^d\zeta_lB_l\right\|\le\Upsilon(\chi)\right\}\ge 1-\chi, \tag{46}$$

where $\|\cdot\|$ is the standard matrix norm (the largest singular value) and $\Upsilon(\chi)$ is given by (44).

Proof of Lemma 4.2. Observe that A_l, $1\le l\le d$ are also of the form (43):

$$A_l=[gh_l^T+h_lg^T]+\lambda_lH \qquad [g=B_0^{-1/2}e,\ h_l=B_0^{-1/2}f_l,\ H=B_0^{-1/2}GB_0^{-1/2}].$$

Note that by rescaling h_l we can ensure that $\|g\|_2=1$, and then rotate the coordinates to make g the first basic orth. In this situation, matrices A_l become matrices of the form

$$A_l=\left[\begin{array}{c|c} q_l & r_l^T\\ \hline r_l & \lambda_lQ\end{array}\right]. \tag{47}$$

Finally, by appropriate scaling of λ_l, we can ensure that $\|Q\|=1$. We have

$$A_l^2=\left[\begin{array}{c|c} q_l^2+r_l^Tr_l & q_lr_l^T+\lambda_lr_l^TQ\\ \hline q_lr_l+\lambda_lQr_l & r_lr_l^T+\lambda_l^2Q^2\end{array}\right].$$

We conclude that $\sum_{l=1}^d A_l^2\preceq I_n$ implies that $\sum_{l=1}^d(q_l^2+r_l^Tr_l)\le 1$ and $[\sum_{l=1}^d\lambda_l^2]Q^2\preceq I_{n-1}$; because $\|Q^2\|=1$, we arrive at the relations

$$\text{(a)}\ \sum_{l=1}^d\lambda_l^2\le 1 \qquad \text{(b)}\ \sum_{l=1}^d(q_l^2+r_l^Tr_l)\le 1. \tag{48}$$

Now let $p_l=(0,r_l^T)^T\in\mathbf{R}^n$. We have

$$S\equiv\sum_{l=1}^d\zeta_lA_l=\Bigg[g\Bigg(\underbrace{\sum_{l=1}^d\zeta_lp_l}_{\xi}\Bigg)^T+\xi g^T\Bigg]+\mathrm{Diag}\Bigg\{\underbrace{\sum_{l=1}^d\zeta_lq_l}_{\theta},\Bigg(\underbrace{\sum_{l=1}^d\zeta_l\lambda_l}_{\eta}\Bigg)Q\Bigg\}$$
$$\Rightarrow\ \|S\|\le\|g\xi^T+\xi g^T\|+\max[|\theta|,\,|\eta|\|Q\|]=\|\xi\|_2+\max[|\theta|,\,|\eta|].$$

Setting

$$\alpha=\sum_{l=1}^d r_l^Tr_l,\qquad \beta=\sum_{l=1}^d q_l^2,$$

we have $\alpha+\beta\le 1$ by (48.b). Besides this,

$$\begin{aligned}\mathbf{E}\{\xi^T\xi\} &= \sum_{l,l'}\mathbf{E}\{\zeta_l\zeta_{l'}\}p_l^Tp_{l'}=\sum_{l=1}^d\mathbf{E}\{\zeta_l^2\}r_l^Tr_l \quad [\zeta_l \text{ are independent, } \mathbf{E}\{\zeta_l\}=0]\\ &\le \sigma^2\sum_{l=1}^d r_l^Tr_l\le\sigma^2\alpha\end{aligned}$$
$$\Rightarrow\ \mathrm{Prob}\{\|\xi\|_2>t\}\le\frac{\sigma^2\alpha}{t^2}\quad\forall\, t>0\quad[\text{Tschebyshev inequality}]$$

$$\begin{aligned}
\mathbf{E}\{\eta^2\} &= \sum_{l=1}^d \mathbf{E}\{\zeta_l^2\}\lambda_l^2 \le \sigma^2 \sum_{l=1}^d \lambda_l^2 \le \sigma^2 \quad \text{[see (48.a)]}\\
&\Rightarrow \text{Prob}\{|\eta| > t\} \le \frac{\sigma^2}{t^2} \quad \forall\, t > 0 \quad \text{[Tschebyshev inequality]}\\
\mathbf{E}\{\theta^2\} &= \sum_{l=1}^d \mathbf{E}\{\zeta_l^2\} q_l^2 \le \sigma^2\beta\\
&\Rightarrow \text{Prob}\{|\theta| > t\} \le \frac{\sigma^2\beta}{t^2} \quad \forall\, t > 0 \quad \text{[Tschebyshev inequality]}
\end{aligned}$$

Thus, for every $\Upsilon > 0$ and all $\lambda \in (0, 1)$ we have

$$\begin{aligned}
\text{Prob}\{\|S\| > \Upsilon\} &\le \text{Prob}\{\|\xi\|_2 + \max[|\theta|, |\eta|] > \Upsilon\} \le \text{Prob}\{\|\xi\|_2 > \lambda\Upsilon\}\\
&\quad + \text{Prob}\{|\theta| > (1-\lambda)\Upsilon\} + \text{Prob}\{|\eta| > (1-\lambda)\Upsilon\} \le \frac{\sigma^2}{\Upsilon^2}\left[\frac{\alpha}{\lambda^2} + \frac{\beta+1}{(1-\lambda)^2}\right],
\end{aligned}$$

whence, due to $\alpha + \beta \le 1$, one has

$$\text{Prob}\{\|S\| > \Upsilon\} \le \frac{\sigma^2}{\Upsilon^2} \max_{\alpha\in[0,1]} \min_{\lambda\in(0,1)} \left[\frac{\alpha}{\lambda^2} + \frac{2-\alpha}{(1-\lambda)^2}\right] = \frac{8\sigma^2}{\Upsilon^2},$$

so that

$$\Upsilon \ge 2\sigma\sqrt{2/\chi} \;\Rightarrow\; \text{Prob}\{\|S\| > \Upsilon\} \le \chi, \tag{49}$$

which is the "general case" of our lemma (cf. (44.a)). It remains to justify the refinements in the cases of (A.1) and (A.2). In the case of (A.1), we have $\sigma \le 1$, so that whenever $\widetilde{\Upsilon} > 4$, we have $\text{Prob}\{\|S\| \ge \widetilde{\Upsilon}\} < 1/2$ by (49). Invoking Theorem 3.1, we conclude that for all $\gamma \ge 1$ we have $\text{Prob}\{\|S\| \ge \gamma\widetilde{\Upsilon}\} \le 2\exp\{-\widetilde{\Upsilon}^2(\gamma-1)^2/16\}$. Given $\chi \in (0, 1/2)$ and setting $\gamma = 1 + 4\widetilde{\Upsilon}^{-1}\sqrt{\ln(2/\chi)}$, we get $\text{Prob}\{\|S\| \ge \widetilde{\Upsilon} + 4\sqrt{\ln(2/\chi)}\} \le \chi$; because this relation holds true for every $\widetilde{\Upsilon} > 4$, we see that, in addition to (49), $\text{Prob}\{\|S\| \ge 4 + 4\sqrt{\ln(2/\chi)}\} \le \chi$, $0 < \chi < 1/2$, which proves the "(A.1)-version" of the lemma. Now let (A.2) be the case. Here (49) is satisfied with $\sigma = 1$, meaning that whenever $s \in (0, 1/2)$, we have $\text{Prob}\{\|S\| \ge 2\sqrt{2/s}\} \le s$. Applying Theorem 3.1 with s in the role of χ, we conclude that whenever $s \in (0, 1/2)$ and $\gamma \ge 1$, we have

$$\text{Prob}\left\{\|S\| \ge 2\gamma\sqrt{2/s}\right\} \le \text{Erf}\left(\text{ErfInv}(s) + (\gamma-1)\max[2\sqrt{2/s}, \text{ErfInv}(s)]\right).$$

It follows that setting

$$\Upsilon_*(\chi) = \inf_{s,\gamma}\left\{2\gamma\sqrt{2/s}: \begin{array}{l} s \in (0, 1/2), \quad \gamma \ge 1,\\ \text{ErfInv}(s) + (\gamma-1)\max[2\sqrt{2/s}, \quad \text{ErfInv}(s)] \ge \text{ErfInv}(\chi)\end{array}\right\},$$

we ensure the relation $\text{Prob}\{\|S\| \ge \Upsilon_*(\chi)\} \le \chi$ for all $\chi \in (0, 1/2)$. It is immediately seen that $\Upsilon(\chi)$, given in (44) for case (A.2), is an upper bound on $\Upsilon_*(\chi)$, so that (46) holds true in the case of (A.2). □

4.3. Simulation-free safe, tractable approximations of chance-constrained LMIs. Assume that the structure of LMI (8) ensures that the collections of matrices $\theta\mathcal{A}_0[x], \mathcal{A}_1[x], \ldots, \mathcal{A}_d[x]$, for all x and all $\theta \ge 0$, belong to a set $\mathcal{B}$ with the following property:

(P) *We can point out functions* $\Upsilon_1(\chi)$, $\Upsilon_2(\chi)$, $0 < \chi < 1/2$, *such that whenever a collection of matrices* $B_0, B_1, \ldots, B_d$ *belongs to* $\mathcal{B}$ *and satisfies the condition* $\text{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0$, *we have*

$$\begin{aligned}
&\forall(0 < \chi < 1/2): \quad \text{Prob}\left\{-\Upsilon_1(\chi)B_0 \preceq \sum_{l=1}^d \zeta_l B_l \preceq \Upsilon_1(\chi)B_0\right\} \ge 1-\chi \quad \textit{whenever } \zeta \textit{ satisfies } \text{(A.1)};\\
&\forall(0 < \chi < 1/2): \quad \text{Prob}\left\{-\Upsilon_2(\chi)B_0 \preceq \sum_{l=1}^d \zeta_l B_l \preceq \Upsilon_2(\chi)B_0\right\} \ge 1-\chi \quad \textit{whenever } \zeta \textit{ satisfies } \text{(A.2)}.
\end{aligned} \tag{50}$$

For example,

- using some deep results from functional analysis—the "noncommutative Khintchine inequality" (Buchholz [7]), it can be easily verified that (P) is true *for all matrices* $A, A_1, \ldots, A_d$, provided that $\Upsilon_{1,2}(\chi) = O(1)\sqrt{\ln(n/\chi)}$; see Proposition A.1 in the appendix or Man-Cho So [12]. The same is true when $\mathcal{B}$ is comprised of all collections of diagonal $n \times n$ matrices, see Theorem 4.1, and it is easily seen that in the latter case the outlined value of $\Upsilon(\chi)$ is, up to an $O(1)$ factor, the smallest possible;

- restricting $\mathscr{B}$ to be all collections $B_0, B_1, \ldots, B_d$ of symmetric $n \times n$ matrices with $B_1, \ldots, B_d$ of the form $e^T f_l + f_l^T e + \lambda_l G$, (P), it was shown in Theorem 4.2 that (50) is satisfied with $\Upsilon_{1,2}(\chi) = O(1)\sqrt{\ln(1/\chi)}$.

In the case of (P), we can build safe, tractable approximations of our problems of interest (7) and (26), avoiding the necessity to use simulations. Specifically, combining (50) with Corollary 3.1, we see that the problem

$$\min_x \left\{ c^T x\colon \begin{array}{l} F(x) \le 0 \\ \operatorname{Arrow}(\Theta \mathscr{A}_0[x], \mathscr{A}_1[x], \ldots, \mathscr{A}_d[x]) \succeq 0 \end{array} \right\},$$

$$\Theta^{-1} = \begin{cases} \inf\limits_{0<\chi<1/2} \left[\Upsilon_1(\chi) + 4\sqrt{\ln(\epsilon^{-1}(1-\chi)^{-1})}\right], & \text{case of (A.1)} \\ \inf\limits_{0<\chi<1/2} [\Upsilon_2(\chi) + \max[\operatorname{ErfInv}(\epsilon)/\operatorname{ErfInv}(\chi) - 1, 0]\min[\Upsilon_2(\chi), \operatorname{ErfInv}(\chi)]], & \text{case of (A.2)} \end{cases} \tag{51}$$

is a safe, tractable approximation of (7).

By exactly the same reasons, given a feasible solution $(x_*, \vartheta_* > 0)$ to (28) and setting $\rho_* = \Theta/\vartheta_*$, with Θ given by (51), we ensure that (x_*, ρ_*) is a feasible solution to (26).

It is not difficult to see that in the cases of chance-constrained linear and conic quadratic programming (covered by Theorems 4.1 and 4.2, respectively), the corresponding "simulation-free" safe, tractable approximations are not too conservative. For example, in the case of (A.2) there exists an absolute constant $C > 0$ such that a vector x that does *not* satisfy the constraint $\operatorname{Arrow}(C^{-1}\Theta\mathscr{A}_0[x], \mathscr{A}_1[x], \ldots, \mathscr{A}_d[x]) \succeq 0$ does not necessarily satisfy the chance constraint of interest (8), provided that $\epsilon n \le 1$. However, we shall see in §6 that in practice simulation-based approximations can be significantly less conservative than the simulation-free ones.

5. Majorization.

One way to bound from above the probability of violating a randomly perturbed LMI:

$$q(x) := \operatorname{Prob}\left\{\mathscr{A}_0[x] + \sum_{l=1}^d \zeta_l \mathscr{A}_l[x] \not\succeq 0\right\},$$

is to replace the random perturbations ζ with easier-to-handle perturbations $\hat{\zeta}$—to which we know how to bound from above the quantity

$$\hat{q}(x) := \operatorname{Prob}\left\{\mathscr{A}_0[x] + \sum_{l=1}^d \hat{\zeta}_l \mathscr{A}_l[x] \not\succeq 0\right\}.$$

If in addition $\hat{\zeta}$ is "more diffuse" than ζ, meaning that $\hat{q}(x) \ge q(x)$ for all x, we indeed end up with a bounding scheme for $q(\cdot)$. For example, let the entries in ζ be independent with zero means and *un*bounded ranges. With our present results, we cannot handle this situation unless ζ_l are Gaussian. In order to overcome this difficulty, we could replace ζ_l with "more diffuse" *Gaussian* random variables $\hat{\zeta}_l$, which we do know how to handle.

For the above idea to be meaningful we should properly specify the notion of "being more diffuse." We are about to present two specifications of this type, known as *monotone* and *convex* stochastic dominances, respectively.

5.1. Monotone dominance and comparison theorem.

For our purposes, it suffices to restrict ourselves with monotone dominance on the space $\mathscr{SU}$ of all symmetric w.r.t. 0 and unimodal probability distributions on the axis. The latter notion is defined as follows:

Definition 5.1. A probability distribution P on the axis is called unimodal and symmetric if P possesses a density $p(\cdot)$ that is an even function nonincreasing on $[0, \infty)$.[1]

A probability distribution $P \in \mathscr{SU}$ is said to be monotonically dominating another distribution $Q \in \mathscr{SU}$ (notation: $P \succeq_{\mathrm{m}} Q$, or, equivalently, $Q \preceq_{\mathrm{m}} P$), if $\int_t^\infty dP(s) \ge \int_t^\infty dQ(s)$ for every $t \ge 0$, or, equivalently,[2] $\int f(s)\,dP(s) \ge \int f(s)\,dQ(s)$ for every even and bounded function $f(s)$ that is nondecreasing on the nonnegative ray $\mathbf{R}_+$.

[1] In literature, a unimodal symmetric distribution is defined as a convex combination of the unit mass sitting at the origin and of what is called unimodal and symmetric in Definition 5.1. For the sake of simplicity, we forbid a mass at the origin; note that all results to follow remain valid when such a mass is allowed.

[2] This equivalence is well known; to be self-contained, we present the proof in the appendix.

With a slight abuse of notation, if ξ is a random variable with distribution P and probability density $p(\cdot)$, then every one of the relations $\xi \in \mathcal{SU}$, $p(\cdot) \in \mathcal{SU}$ is interpreted as the inclusion $P \in \mathcal{SU}$. Similarly, if ξ, η are random variables with distributions P, respectively, Q, and probability densities $p(\cdot)$, respectively, $q(\cdot)$, then every one of the relations $\eta \succeq_{\mathrm{m}} \xi$, $q(\cdot) \succeq_{\mathrm{m}} p(\cdot)$ means that $P, Q \in \mathcal{SU}$ and $P \succeq_{\mathrm{m}} Q$. Relation $\preceq_{\mathrm{m}}$ is the natural "counterpart" of the relation $\succeq_{\mathrm{m}}$.

Important facts on the monotone dominance that we need later are summarized in:

Proposition 5.1. (i) $\succeq_{\mathrm{m}}$ *is a partial order on* $\mathcal{SU}$.

(ii) *If* $p_i(\cdot) \preceq_{\mathrm{m}} q_i(\cdot)$, $i = 1, \ldots, I$, *and* $\alpha_i \geq 0$ *are such that* $\sum_i \alpha_i = 1$, *then* $\sum_i \alpha_i p_i(\cdot) \preceq_{\mathrm{m}} \sum_i \alpha_i q_i(\cdot)$.

(iii) *If* $\xi \in \mathcal{SU}$ *is a random variable, and* λ, $|\lambda| \geq 1$, *is a deterministic real, then* $\xi \preceq_{\mathrm{m}} \lambda\xi$.

(iv) *If* $p_i(\cdot) \in \mathcal{SU}$ *weakly converge as* $i \to \infty$ *to a probability density* $p(\cdot)$ (*meaning that* $\int g(s)p_i(s)\,ds \to \int g(s)p(s)\,ds$ *for every continuous* g *with compact support*), $q_i(\cdot) \in \mathcal{SU}$ *weakly converge as* $i \to \infty$ *to a probability density* $q(\cdot)$ *and* $p_i(\cdot) \preceq_{\mathrm{m}} q_i(\cdot)$ *for every* i, *then* $p(\cdot) \in \mathcal{SU}$, $q(\cdot) \in \mathcal{SU}$, *and* $p(\cdot) \preceq_{\mathrm{m}} q(\cdot)$.

(v) *If* $\{\xi_l \in \mathcal{SU}\}_{l=1}^n$, $\{\eta_l \in \mathcal{SU}\}_{l=1}^n$ *are collections of independent random variables such that* $\xi_l \succeq_{\mathrm{m}} \eta_l$, $l = 1, \ldots, n$, *and* λ_l, $l = 1, \ldots, n$, *are deterministic reals, then* $\sum_{l=1}^n \lambda_l \xi_l \succeq_{\mathrm{m}} \sum_{l=1}^m \lambda_l \eta_l$.

(vi) *Let* $\xi \in \mathcal{P}$ *be supported on* $[-1, 1]$, ζ *be uniformly distributed on* $[-1, 1]$, *and* $\eta \sim \mathcal{N}(0, 2/\pi)$. *Then* $\xi \preceq_{\mathrm{m}} \zeta \preceq_{\mathrm{m}} \eta$.

(vii) [*Comparison Theorem*] *Let* $\{\zeta_l \in \mathcal{SU}\}_{l=1}^d$, $\{\hat{\zeta}_{l=1}^d\}$ *be two collections of independent random variables such that* $\zeta_l \preceq_{\mathrm{m}} \hat{\zeta}_l$ *for all* l. *Then for every closed convex and symmetric w.r.t. the origin set* $Q \subset \mathbf{R}^d$, *one has*

$$\mathrm{Prob}\{\zeta := [\zeta_1; \ldots; \zeta_d] \in Q\} \geq \mathrm{Prob}\{\hat{\zeta} := [\hat{\zeta}_1; \ldots; \hat{\zeta}_d] \in Q\}.$$

To the best of our knowledge, some of the facts presented in Proposition 5.1, most notably the comparison theorem, are new; to be on the safe side, we provide full proofs of all these facts in the appendix.

5.2. Convex dominance and the majorization theorem. To conclude this section, we present another "Gaussian majorization" result. Its advantage is that the random variables ζ_l are not required to be symmetrically or unimodally distributed; what is needed, essentially, is just independence plus zero means. We start with recalling the definition of *convex dominance*. Let $\mathcal{R}_n$ be the space of Borel probability distributions on $\mathbf{R}^n$ with zero mean. For a random variable η taking values in $\mathbf{R}^n$, we denote by P_η the corresponding distribution, and we write $\eta \in \mathcal{R}_n$ to express that $P_\eta \in \mathcal{R}_n$. Let $\mathcal{CF}_n$ be the set of all *convex* function f on $\mathbf{R}^n$ with linear growth, meaning that there exists $c_f < \infty$ such that $|f(u)| \leq c_f(1 + \|u\|_2)$ for all u.

Definition 5.2. Let ξ, $\eta \in \mathcal{R}_n$. We say that η convexly dominates ξ (notation: $\xi \preceq_{\mathrm{c}} \eta$, or $P_\xi \preceq_{\mathrm{c}} P_\eta$, or $\eta \succeq_{\mathrm{c}} \xi$, or $P_\eta \succeq_{\mathrm{c}} P_\xi$) if

$$\int f(u)\,dP_\xi(u) \leq \int f(u)\,dP_\eta(u)$$

for every $f \in \mathcal{CF}_n$.

The relevant facts on convex dominance that we need are summarized in:

Proposition 5.2. (i) $\preceq_{\mathrm{c}}$ *is a partial order on* $\mathcal{R}_n$.

(ii) *If* $P_1, \ldots, P_k, Q_1, \ldots, Q_k \in \mathcal{R}_n$, *and* $P_i \preceq_{\mathrm{c}} Q_i$ *for every* i, *then* $\sum_i \lambda_i P_i \preceq_{\mathrm{c}} \sum_i \lambda_i Q_i$ *for all nonnegative weights* λ_i *with unit sum.*

(iii) *If* $\xi_1, \ldots, \xi_k, \eta_1, \ldots, \eta_k \in \mathcal{R}_n$ *are independent random variables such that* $\xi_i \preceq_{\mathrm{c}} \eta_i$ *for every* i, *and* s_i *are deterministic reals, then* $\sum_i s_i \xi_i \preceq_{\mathrm{c}} \sum_i s_i \eta_i$.

(iv) *If* ξ *is symmetrically distributed w.r.t.* 0 *and* $t \geq 1$ *is deterministic, then* $t\xi \succeq_{\mathrm{c}} \xi$.

(v) *Let* $P_1, Q_1 \in \mathcal{R}_r$, $P_2, Q_2 \in \mathcal{R}_s$ *be such that* $P_i \preceq_{\mathrm{c}} Q_i$, $i = 1, 2$. *Then* $P_1 \times P_2 \preceq_{\mathrm{c}} Q_1 \times Q_2$. *In particular, if* $\xi_1, \ldots, \xi_n, \eta_1, \ldots, \eta_n \in \mathcal{R}_1$ *are independent and such that* $\xi_i \preceq_{\mathrm{c}} \eta_i$ *for every* i, *then* $(\xi_1, \ldots, \xi_n)^T \preceq_{\mathrm{c}} (\eta_1, \ldots, \eta_n)^T$.

(vi) *Let* $\xi \in \mathcal{R}_1$ *be supported on* $[-1, 1]$ *and* $\eta \sim \mathcal{N}(0, \pi/2)$. *Then* $\xi \preceq_{\mathrm{c}} \eta$.

(vii) *Assume that* $\xi \in \mathcal{R}_n$ *is supported in the unit cube* $\{u\colon \|u\|_\infty \leq 1\}$ *and is "absolutely symmetrically distributed," meaning that if* J *is a diagonal matrix with diagonal entries* ± 1, *then* $J\xi$ *has the same distribution as* ξ. *Also let* $\eta \sim \mathcal{N}(0, (\pi/2)I_n)$. *Then* $\xi \preceq_{\mathrm{c}} \eta$.

(viii) *Let* $\xi, \eta \in \mathcal{R}_n$, $\xi \sim \mathcal{N}(0, \Sigma)$, $\eta \sim \mathcal{N}(0, \Theta)$ *with* $\Sigma \preceq \Theta$. *Then* $\xi \preceq_{\mathrm{c}} \eta$.

(ix) [*Majorization theorem*] *Let* $\eta \sim \mathcal{N}(0, I_d)$, *and let* $\zeta \in \mathcal{R}_d$ *be such that* $\zeta \preceq_{\mathrm{c}} \eta$. *Let, further,* $Q \subset \mathbf{R}^d$ *be a closed convex set such that*

$$\chi \equiv \mathrm{Prob}\{\eta \notin Q\} < 1/2.$$

Then for every $\gamma > 1$, *one has*

$$\operatorname{Prob}\{\zeta \notin \gamma Q\} \le \inf_{1\le\beta<\gamma} \frac{1}{\gamma-\beta}\int_\beta^\infty \operatorname{Erf}(s\operatorname{ErfInv}(\chi))\,ds \le \inf_{1\le\beta<\gamma}\frac{1}{2(\gamma-\beta)}\int_\beta^\infty \exp\left\{-\frac{s^2\operatorname{ErfInv}^2(\chi)}{2}\right\}ds. \tag{52}$$

All of the above facts, except for the Majorization Theorem, are well known; proofs can be found in Nemirovski and Shapiro [14]. The present Majorization Theorem is a slight refinement of what is called "Majorization Theorem" in Nemirovski and Shapiro [14]; the proof of this refinement is given in the appendix.

5.3. Calibration-based on Gaussian majorization. We can utilize the preceding facts in the calibration procedure as follows.

Utilizing comparison theorem. Assume that the perturbations ζ_l are independent and possess unimodal and symmetric distributions P_l such that $P_l \preceq_{\mathrm{m}} \mathcal{N}(0,\sigma^2)$ for certain σ and all l (the latter is, e.g., the case when ζ_l are supported on $[-1,1]$ and $\sigma=\sqrt{2/\pi}$; see Proposition 5.1.(vi)). Setting $\eta\sim\mathcal{N}(0,I_d)$ and invoking the comparison theorem, we conclude that for every deterministic symmetric matrices $A_0, A_1, \ldots, A_d$ and every $r>0$ we have

$$\operatorname{Prob}\left\{-A_0 \preceq \frac{r}{\sigma}\sum_{l=1}^d \zeta_l A_l \preceq A_0\right\} \ge \operatorname{Prob}\left\{-A_0 \preceq r\sum_{l=1}^d \eta_l A_l \preceq A_0\right\}. \tag{53}$$

Given matrices $A_0,\ldots,A_d$ and $\vartheta_*>0$ such that $\operatorname{Arrow}(\vartheta_* A_0, A_1,\ldots,A_d)\succeq 0$, along with $\epsilon,\delta\in(0,1)$, the purpose of the calibration procedure is to build a (random) $(1-\delta)$-reliable lower bound on the quantity

$$\rho^* = \max\left\{\rho\colon \operatorname{Prob}\left\{-A_0\preceq \rho\sum_{l=1}^d \zeta_l A_l \preceq A_0\right\}\ge 1-\epsilon\right\}. \tag{54}$$

By (53), in order to build such a bound, we can apply the plain calibration procedure to find a $(1-\delta)$-reliable lower bound r_* on the quantity

$$r^* = \max\left\{r\colon \operatorname{Prob}\left\{-A_0\preceq r\sum_{l=1}^d \eta_l A_l\preceq A_0\right\}\ge 1-\epsilon\right\}$$

and to set $\rho_* = r_*/\sigma$. This approach allows us to extend the above constructions beyond the scope of Assumption A; moreover, we shall see in §6 that this approach makes sense even in the case when ζ obeys (A.1) and thus can be processed "as it is." The reason is that the constant factors in the measure concentration inequalities of Theorem 3.1 in the case of (A.2) are better than in the case of (A.1).

Utilizing majorization theorem. Now assume that the random variables $\zeta_1,\ldots,\zeta_d$ are independent with zero means, and that we can point out $\sigma>0$ such that $P_{\zeta_l}\preceq_{\mathrm{c}}\mathcal{N}(0,\sigma^2)$. Introducing $\eta\sim\mathcal{N}(0,I_d)$ and applying Proposition 5.2.(v), we conclude that $\zeta\preceq_{\mathrm{c}}\sigma\eta$. Given the input $A_0,\ldots,A_d$, ϵ, δ to the calibration procedure and applying Majorization Theorem to the closed convex set

$$Q = Q_s = \left\{u\in\mathbf{R}^d\colon\ -sA_0\preceq\sum_{l=1}^d u_l A_l\preceq sA_0\right\},$$

we conclude that

$$\begin{gathered}\forall\left(s>0\colon \chi(s)\equiv\operatorname{Prob}\{\eta\notin Q_s\}\equiv 1-\operatorname{Prob}\left\{-A_0\preceq s^{-1}\sum_{l=1}^d\eta_l A_l\preceq A_0\right\}<1/2\right)\colon\\ \operatorname{Prob}\left\{-\gamma s\sigma A_0\preceq\sum_{l=1}^d\zeta_l A_l\preceq\gamma\sigma s A_0\right\} = 1-\operatorname{Prob}\{\sigma^{-1}\zeta\notin\gamma Q_s\}\ge 1-\Psi(\gamma,\chi(s)),\\ \Psi(\gamma,\chi)=\inf_{1\le\beta<\gamma}\frac{1}{\gamma-\beta}\int_\beta^\infty\operatorname{Erf}(s\operatorname{ErfInv}(\chi))\,ds.\end{gathered} \tag{55}$$

In order to bound from below ρ^* (see (54)), we apply the calibration procedure *with artificial random perturbation* η *in the role of actual perturbation* ζ. Carrying out the first two steps of this procedure, we end up with a collection $\{r^k>0, \hat\chi^k<1/2\}_{k=1}^{\bar K}$ such that "up to probability of bad sampling $\le\delta$" we have, for $1\le k\le\bar K$,

$$\hat\chi^k \ge \chi^k := \operatorname{Prob}\left\{\neg\left(-(r^k)^{-1}A_0\preceq\sum_{l=1}^d\eta_l A_l\preceq(r^k)^{-1}A_0\right)\right\} = \operatorname{Prob}\{\eta\notin Q_{s_k}\},\quad s_k = 1/r^k;$$

this collection is obtained from the collection $\{r_k, \hat{\chi}_k\}_{k=1}^{K}$ built at step 2 of the procedure by discarding the pairs with $\hat{\chi}_k \geq 1/2$. Setting

$$\rho_* = \max_{1\leq k\leq \bar{K}} \frac{r^k}{\sigma\gamma_k}, \qquad \gamma_k = \min\{\gamma\geq 1\colon \Psi(\gamma, \hat{\chi}^k) \leq \epsilon\}, \quad k=1,\ldots,\bar{K}$$

and invoking (55), we see that ρ_* is a lower bound on ρ^*, provided that $\chi^k \leq \hat{\chi}^k$, $1\leq k\leq \bar{K}$, which happens with probability at least $1-\delta$.

Note that with straightforward modifications, Gaussian majorization can be used in the validation procedure.

6. Numerical illustrations. In the following illustrations, we focus on problem (26) and on its safe tractable approximation given by (28) and the calibration procedure.

6.1. The calibration procedure. We start with illustrating the "stand-alone" calibration procedure. Recall that this procedure is aimed at building $(1-\delta)$-reliable lower bound ρ_* on the quantity

$$\rho^* = \max\left\{\rho\colon\, p(\rho) := \mathrm{Prob}\left\{-A_0 \preceq \rho\sum_{l=1}^{d}\zeta_l A_l \preceq A_0\right\} \geq 1-\epsilon\right\}, \tag{56}$$

where $A_0, A_1, \ldots, A_d$ are given symmetric $n\times n$ matrices such that $\mathrm{Arrow}(\vartheta_* A_0, A_1,\ldots,A_d)\succeq 0$ for a given $\vartheta_* > 0$.

The questions we tried to answer in our experiments were as follows:

(i) What is the better strategy to be used in the procedure—the plain calibration procedure (PCP) or the Gaussian majorization version (GCP) of this procedure?

(ii) As we have seen in §4.3, there are situations where not too conservative guaranteed lower bounds on ρ^* can be built without simulations at all. Are these "100% reliable" lower bounds more attractive than those given by calibration procedure?

(iii) From a practical perspective, how conservative is the calibration procedure?

Answers to these questions, based on our rather intensive numerical experimentation, are as follows:

- The calibration procedure, at least its GCP-version, significantly outperforms the simulation-free lower bounding;
- GCP significantly outperforms PCP;
- The conservatism of the calibration procedure is not very severe: the ratio ρ^*/ρ_* is usually well within one order of magnitude.

These observations are summarized in Table 1; they are based on experiments performed as follows: We randomly generate $d=32$ matrices $A_1,\ldots,A_d$ of size 32×32 and of prescribed structure, specifically, full ("general case"), diagonal ("diagonal case"), and of the form

$$\left[\begin{array}{c|c} & f^T \\ \hline f & \end{array}\right] + \lambda I_{32},$$

f being a vector ("arrow case"), and scale the generated matrices to ensure that $\mathrm{Arrow}(\theta I_{32}, A_1,\ldots,A_d)\succeq 0$ if and only if $\theta\geq 1$; the input to the calibration procedure is the collection $A_0 = I_{32}$, $A_1,\ldots,A_{32}$, $\vartheta_* = 1$. Data in Table 1 correspond to 100,000-element training sample. Note that although the performance of the calibration procedure somehow improves when the sample size grows (see Table 2), this phenomenon is rather moderate.

6.2. Illustration: Chance-constrained truss topology design. A *truss* is a mechanical construction comprised of thin elastic *bars* linked with each other at *nodes*. In the simplest Truss topology design (TTD) problem, one is given a finite 2D or 3D nodal set, a list of allowed pair connections of nodes by bars, and an external load—a collection of forces acting at the nodes. The goal is to assign the tentative bars weights, summing up to a given constant, in order to get a truss most rigid w.r.t. the load (for details, see, e.g., Ben-Tal and Nemirovski [2, Chapter 15]). Mathematically, the TTD problem is the semidefinite program

$$\min_{\tau,t}\left\{\tau\colon \left[\begin{array}{c|c} 2\tau & f^T \\ \hline f & \sum_{i=1}^{n} t_i b_i b_i^T\end{array}\right]\succeq 0,\ t\geq 0,\ \sum_i t_i = 1\right\}, \tag{57}$$

TABLE 1. Experiments with stand-alone calibration procedure, $\delta = 1.\text{e}-6$, $n = 32$, $d = 32$.

		Case								
		General			Diagonal			Arrow		
ϵ	P	ρ_*	$\hat{\epsilon}$	$\rho^*/\rho_* \leq$	ρ_*	$\hat{\epsilon}$	$\rho^*/\rho_* \leq$	ρ_*	$\hat{\epsilon}$	$\rho^*/\rho_* \leq$
1.0e−2	G	9.8e−2	8.2e−3	4.6	9.8e−2	2.8e−3	3.4	1.6e−1	2.7e−3	2.5
		4.5e−1		1.0	3.0e−1		1.1	3.5e−1		1.1
		4.5e−1		1.0	3.0e−1		1.1	3.5e−1		1.1
	U	1.2e−1	0.0	6.9	1.2e−1	1.0e−5	4.9	2.0e−1	1.0e−5	3.5
		9.5e−2		8.9	8.6e−2		7.0	8.8e−2		8.0
		5.6e−1		1.5	3.8e−1		1.6	4.4e−1		1.6
	R	1.3e−2	0.0	40	1.3e−2	0.0	27	1.1e−1	1.0e−5	3.6
		9.5e−2		5.5	8.5e−2		4.2	8.6e−2		4.8
		3.3e−1		1.6	2.3e−1		1.5	2.7e−1		1.5
1.0e−4	G	8.6e−2	0.0	4.4	8.6e−2	2.0e−5	2.9	1.3e−1	2.0e−5	2.2
		3.5e−1		1.1	2.3e−1		1.1	2.6e−1		1.1
		3.5e−1		1.1	2.3e−1		1.1	2.6e−1		1.1
	U	1.1e−1	0.0	6.8	1.1e−1	0.0	4.2	1.6e−1	0.0	3.1
		7.1e−2		11	6.6e−2		6.9	6.7e−2		7.5
		4.3e−1		1.7	2.9e−1		1.6	3.2e−1		1.6
	R	8.7e−3	0.0	55	8.7e−3	0.0	32	9.8e−2	0.0	3.1
		7.1e−2		6.8	6.5e−2		4.3	6.6e−2		4.6
		2.6e−1		1.9	1.7e−1		1.7	1.9e−1		1.6
1.0e−6	G	7.9e−2	0.0	4.4	7.9e−2	0.0	2.9	1.1e−1	0.0	2.2
		2.7e−1		1.3	1.9e−1		1.2	2.0e−1		1.2
		2.7e−1		1.3	1.9e−1		1.2	2.0e−1		1.2
	U	9.9e−2	0.0	7.1	9.9e−2	0.0	3.7	1.4e−1	0.0	2.9
		5.9e−2		12	5.6e−2		6.7	5.7e−2		7.4
		3.4e−1		2.1	2.4e−1		1.6	2.6e−1		1.6
	R	7.0e−3	0.0	67	7.0e−3	0.0	39	8.8e−2	0.0	3.0
		5.9e−2		7.9	5.5e−2		4.9	5.6e−2		4.8
		2.1e−1		2.2	1.4e−1		2.0	1.5e−1		1.7

Notes. *Column* "P": identical to each other distributions of $\zeta_1, \ldots, \zeta_d$; G, U, R stand for $\mathcal{N}(0, 1)$, Uniform$[-1, 1]$, and Uniform$\{-1; 1\}$, respectively.

Columns "ρ_*": lower bounds on ρ^*. Rows in a cell are as follows:

- *First row:* simulation-free bound (Gaussian majorization coupled with Proposition A.1, Theorem 4.1, Theorem 4.2, depending on whether A_l are general/diagonal/arrow)
- *Second row:* PCP calibration
- *Third row:* GCP calibration.

Gaussian majorization is based either on comparison, or on Majorization Theorem, depending on the type (U/R) of the distributions of ζ_l.

Columns "$\hat{\epsilon}$": empirical value, over 100,000-element sample, of $1 - p(\rho)$, see (56), ρ being set to the largest value in the corresponding cell of the ρ_*-column

Columns "$\rho_*/\rho^* \leq$": ratios of the empirical bound on ρ^* as yielded by 100,000 sample, to the corresponding lower bounds on ρ^* from the ρ_*-column.

where τ is (an upper bound on) the *compliance*—a natural measure of truss' rigidity (the less the compliance, the better), t_i are weights of the bars, f represents the external load, and b_i are readily given by the geometry of the nodal set. The dimension M of b_is and f is the total # of degrees of freedom of the nodes.

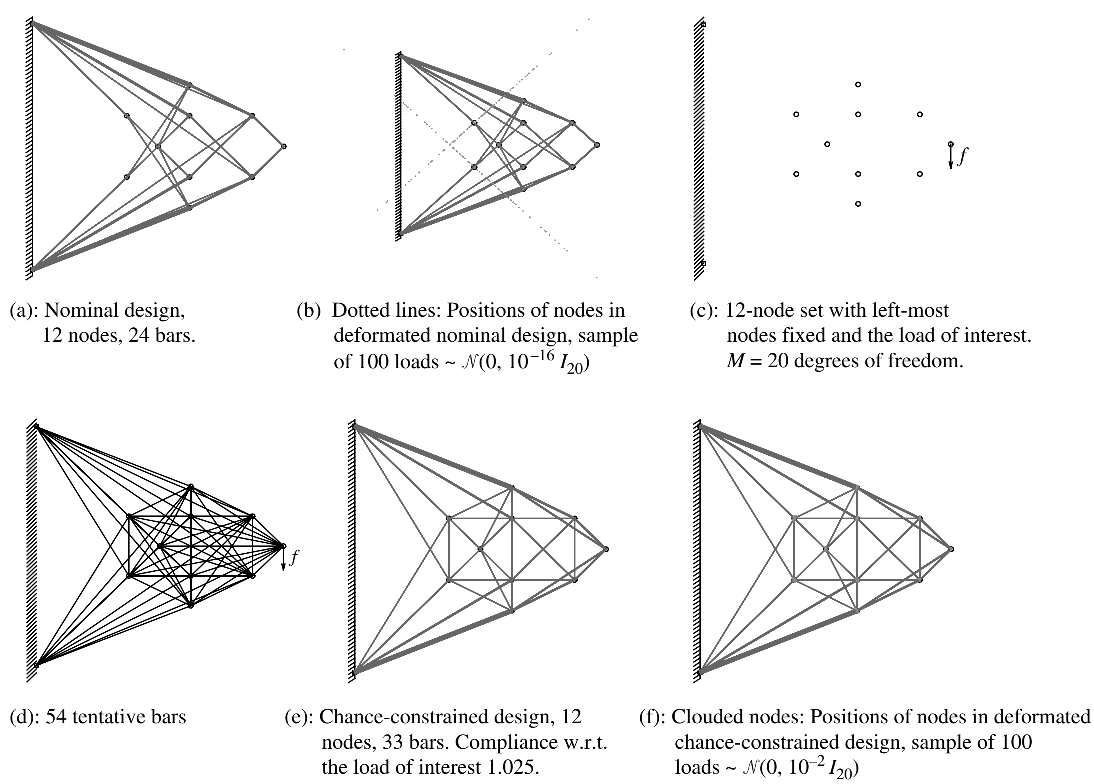
The "nominal design" shown in Figure 1(a) is the optimal solution to a small TTD problem with 9×9 planar nodal grid and where the load f is comprised of a single force (see Figure 1(c)). This design uses just 12 of the original 81 nodes and 24 of the potential 2,039 bars. In reality, the truss, of course, will be subject not only to the primary load f, but also to occasional secondary, relatively small, loads affecting the nodes used by the

TABLE 2. Performance of stand-alone calibration procedure vs. size N of training sample, $\delta = 1.\text{e} - 6$, $n = 32$, $d = 32$, general-type matrices A_l.

		ρ_*		
ϵ	P	$N = 1{,}000$	$N = 10{,}000$	$N = 100{,}000$
1.0e−2	G	3.7e−1	4.4e−1	4.5e−1
		3.7e−1	4.4e−1	4.5e−1
	U	9.4e−2	9.5e−2	9.5e−2
		4.7e−1	5.5e−1	5.6e−1
	R	9.4e−2	9.5e−2	9.5e−2
		2.5e−1	3.2e−1	3.3e−1
1.0e−4	G	2.5e−1	3.0e−1	3.5e−1
		2.5e−1	3.0e−1	3.5e−1
	U	7.0e−2	7.1e−2	7.1e−2
		3.1e−1	3.8e−1	4.4e−1
	R	7.1e−2	7.1e−2	7.1e−2
		1.7e−1	2.3e−1	2.6e−1
1.0e−6	G	2.0e−1	2.4e−1	2.7e−1
		2.0e−1	2.4e−1	2.7e−1
	U	5.9e−2	6.0e−2	5.9e−2
		2.5e−1	3.0e−1	3.4e−1
	R	5.9e−2	6.0e−2	5.9e−2
		1.4e−1	1.8e−1	2.1e−1

Notes. Column "P": see Table 1. The first number in "ρ_*"-cells corresponds to PCP, the second corresponds to GCP.

construction. The truss should, of course, withstand these loads as well. This is *by far* not the case with the truss on Figure 1(a)—it can be crushed by a very small occasional load. Indeed, a typical random load $\tilde{f}$ acting on the 12 nodes of the nominal design, and of very small size as compared to f, say ($\|\tilde{f}\|_2 \leq 10^{-7}\|f\|_2$), results in compliances that are about 10 times larger than the compliance caused by f—a phenomenon illustrated on Figure 1(b). A natural way to "cure" the nominal design is to reformulate the TTD problem, explicitly imposing



(a): Nominal design, 12 nodes, 24 bars.

(b) Dotted lines: Positions of nodes in deformated nominal design, sample of 100 loads ~ $\mathcal{N}(0, 10^{-16} I_{20})$

(c): 12-node set with left-most nodes fixed and the load of interest. $M = 20$ degrees of freedom.

(d): 54 tentative bars

(e): Chance-constrained design, 12 nodes, 33 bars. Compliance w.r.t. the load of interest 1.025.

(f): Clouded nodes: Positions of nodes in deformated chance-constrained design, sample of 100 loads ~ $\mathcal{N}(0, 10^{-2} I_{20})$

FIGURE 1. Nominal and chance-constrained designs.

the requirement that the would-be truss should carry occasional random loads well. Specifically, we

- replace the original 81-point nodal set with the 12-point set of nodes actually used by the nominal design (Figure 1(c)). Note that among these nodes, the two leftmost ones are fixed by boundary conditions ("are in the wall"), so that the total number M of degrees of freedom of this reduced nodal set is $2\times 10 = 20$;
- allow for all pair connections of the resulting 12 nodes by tentative bars (except for clearly redundant bar linking the two fixed nodes and the bars incident to more than 2 nodes); the resulting 54 tentative bars are shown on Figure 1(d);
- assume that the occasional loads are random $\sim \mathcal{N}(0, \rho^2 I_{20})$, where ρ is an uncertainty level, and take, as the "corrected" truss, the *chance-constrained design*—the optimal solution to the following chance-constrained semidefinite program:

$$\max_{\rho,\,t}\left\{\rho:\ \begin{array}{l}\overbrace{\left[\begin{array}{c|c}2\hat{\tau} & f^T\\ \hline f & \sum_{i=1}^{54} t_i b_i b_i^T\end{array}\right]}^{A(t)}\succeq 0,\quad t\geq 0,\quad \sum_i t_i = 1\\[2ex] \mathrm{Prob}_{\zeta\sim\mathcal{N}(0,I_{20})}\left\{\underbrace{\left[\begin{array}{c|c}2\hat{\tau} & \rho\zeta^T\\ \hline \rho\zeta & \sum_{i=1}^{54} t_i b_i b_i^T\end{array}\right]}_{\mathcal{A}_0[t]+\rho\sum_{l=1}^{M}\zeta_l\mathcal{A}_l[t]}\succeq 0\right\}\geq 1-\epsilon\end{array}\right\}, \tag{58}$$

where $\hat{\tau}$ is slightly greater than the optimal value τ_* in the original TTD problem (in our experiment, we set $\hat{\tau} = 1.025\tau_*$). In other words, we are now looking for truss for which the compliance w.r.t. the primary load f is nearly optimal—is at most $\hat{\tau}$, and which is capable of withstanding equally well to "nearly all" (up to probability ϵ) random occasional loads of the form $\rho\zeta$, $\zeta\sim\mathcal{N}(0, I_{20})$; under these restrictions, we intend to maximize ρ, i.e., to maximize (the $(1-\epsilon)$-quantile of) the rigidity of the truss w.r.t. occasional loads (cf. (26)). Note that the robust optimization version of the outlined strategy was proposed and discussed in full details in Ben-Tal and Nemirovski [1].
Implementing the outlined strategy, we built and solved the safe tractable approximation

$$\min_{\vartheta,\,t}\left\{\vartheta:\ \begin{array}{l}A(t)\succeq 0,\quad t\geq 0,\quad \sum_i t_i = 1\\ \mathrm{Arrow}(\vartheta\mathcal{A}_0[t], \mathcal{A}_1[t],\ldots,\mathcal{A}_M[t])\succeq 0\end{array}\right\} \tag{59}$$

(cf. (27)) of the chance-constrained TTD problem (58). After a feasible solution t_* to the approximation is found, we used the calibration procedure to build a $(1-\delta)$-reliable lower bound ρ_* on the largest $\rho = \rho^*(t_*)$ such that (t_*, ρ) is feasible for (58). In our experiment, we worked with pretty high reliability requirements: $\epsilon = \delta = 1.\mathrm{e}{-10}$. The results are presented in Table 3 and are illustrated on Figure 1. Note that we are in the arrow case, so that we can build a simulation-free lower bound on $\rho^*(t_*)$; see §4.3. With our data, this load is 4.01e−3—more than 10 times worse than the best simulation-based extremely reliable ($\delta = 1.\mathrm{e}{-10}$) bound presented in Table 3.

Comparison with the scenario approximation. We have used the TTD example to compare our approximation scheme with the scenario one (see the introduction). The latter, to the best of our knowledge, is the

TABLE 3. Lower bounds for $\rho^*(t_*)$ in the chance-constrained TTD problem vs. the size N of training sample, $\epsilon = \delta = 1.\mathrm{e}{-10}$.

	$N = 1{,}000$	$N = 10{,}000$	$N = 100{,}000$
ρ_*	0.035	0.041	0.043
$\rho^*(t_*)/\rho_* \leq$	1.99	1.64	1.56

only existing alternative for processing chance-constrained LMIs. The scenario approximation of the chance constrained problem of interest (58) is the semidefinite program

$$\max_{\rho,\,t}\left\{\rho:\ \begin{array}{l}\left[\begin{array}{c|c}2\hat\tau & f^T\\\hline f & \sum_{i=1}^{54} t_i b_i b_i^T\end{array}\right]\succeq 0,\quad t\geq 0,\quad \sum_i t_i=1\\[2ex] \left[\begin{array}{c|c}2\hat\tau & \rho[\zeta^j]^T\\\hline \rho\zeta^j & \sum_{i=1}^{54} t_i b_i b_i^T\end{array}\right]\succeq 0,\quad 1\leq j\leq J\end{array}\right\}, \tag{60}$$

where $\zeta^1,\ldots,\zeta^J$ is a sample drawn from $\mathcal{N}(0, I_{20})$; the sample size J is given by (5) where one should set $m=\dim t+\dim\rho=55$. Needless to say, the scenario approximation with the above $\epsilon=\delta=1.\mathrm{e}{-}10$ requires a completely unrealistic sample size; this is why we ran the scenario approximation with $\epsilon=0.01$, $\delta=0.001$. Although these levels of unreliability are by far too dangerous for actual truss design, they are acceptable in our current comparison context. With the outlined ϵ, δ, the sample size J as given by (5) is 42,701, and the optimal value in (60) turned out to be $\rho_{\mathrm{SA}}=0.0797$. For comparison, our approximation with $\epsilon=0.01$ and $\delta=0.001$ results in $\rho_*=0.105\approx 1.31\rho_{\mathrm{SA}}$; keeping $\epsilon=0.01$ and reducing δ to 1.e−6, we still get $\rho_*=0.103\approx 1.29\rho_{\mathrm{SA}}$. Note that the design given by (59) also seems to be better than the one given by (60): At uncertainty level $\rho=0.105$, the empirical probabilities (over 100,000-element sample of random occasional loads) for the two designs to yield a compliance worse than the desired upper bound $\hat\tau$ were 0.0077 and 0.0097, respectively. Thus, in the experiment we are reporting, our approximation scheme is a clear winner.

Appendix A. Some proofs.

Proof of Equivalence in Definition 5.1. We should prove that if $p(s)$, $q(s)$ are nonincreasing on $\mathbf{R}_+$ and such that $\int_{\mathbf{R}^+} p(s)\,ds=\int_{\mathbf{R}_+} q(s)\,ds$, and $\mathcal{M}$ is the family of all bounded nondecreasing functions on $\mathbf{R}_+$, then

$$\left\{\forall f\in\mathcal{M}:\ \int f(s)p(s)\,ds\leq\int f(s)q(s)\,ds\right\}\ \Leftrightarrow\ \left\{\forall t\geq 0:\ \int_t^\infty p(s)\,ds\leq\int_t^\infty q(s)\,ds\right\}. \tag{61}$$

By standard continuity arguments, the left condition in (61) is equivalent to the similar condition with $\mathcal{M}$ replaced with the space $\mathcal{CM}$ of all continuously differentiable bounded nondecreasing functions on $\mathbf{R}_+$.

Setting $P(s)=\int_s^\infty p(r)\,dr$, $Q(s)=\int_s^\infty q(r)\,dr$, for every $f\in\mathcal{CM}$ we have

$$\begin{aligned} I[f] &:= \int_0^\infty f(s)[q(s)-p(s)]\,ds = -\int_0^\infty f(s)\,dQ(s)+\int_0^\infty f(s)\,dP(s)\\ &= f(0)[Q(0)-P(0)]+\int_0^\infty f'(s)[Q(s)-P(s)]\,ds=\int_0^\infty f'(s)[Q(s)-P(s)]\,ds.\end{aligned}$$

We see that $I[f]\geq 0$ for every continuously differentiable nondecreasing and bounded f if and only if $\int_0^\infty g(s)[Q(s)-P(s)]\,ds\geq 0$ for every nonnegative summable function $g(\cdot)$ on $\mathbf{R}^+$; because $P(\cdot)$, $Q(\cdot)$ are continuous, the latter is the case if and only if $Q(s)\geq P(s)$ for all $s\geq 0$. □

Proof of Proposition 5.1. Relations (i)–(iv) are evident in view of the equivalence mentioned in Definition 5.1.

(v): Relation $\xi\succeq_{\mathrm{m}}\eta$ clearly implies that $\lambda\xi\succeq_{\mathrm{m}}\lambda\eta$ for every deterministic λ. In view of this fact, in order to prove (v) it suffices to prove that if the densities $p,\hat p,q,\hat q$ belong to $\mathcal{SU}$ and $p\preceq_{\mathrm{m}} q$, $\hat p\preceq_{\mathrm{m}}\hat q$, then $p*\hat p$ and $q*\hat q$ belong to $\mathcal{SU}$ and $p*\hat p\preceq_{\mathrm{m}} q*\hat q$.

1^0. Let us verify that $p*\hat p\in\mathcal{SU}$. We should prove that the density $(p*\hat p)(s)=\int p(s-r)\hat p(r)\,dr$ is even (which is evident) and is nonincreasing on $\mathbf{R}_+$. By standard approximation arguments, it suffices to verify the latter fact when the probability densities p, $\hat p$, in addition to being even and nonincreasing on $\mathbf{R}_+$, are smooth. In this case we have

$$(p*\hat p)'(s)=\int p'(s-r)\hat p(r)\,dr=\int p(s-r)\hat p'(r)\,dr=\int_{-\infty}^0 (p(s-t)-p(s+t))\hat p'(t)\,dt. \tag{62}$$

Let $s\geq 0$. Then for $t\leq 0$ we have $|s|+|t|=|s-t|\geq|s+t|$, and because p is even and nonincreasing on $\mathbf{R}_+$, we conclude that $p(s-t)=p(|s-t|)\leq p(|s+t|)=p(s+t)$, so that $p(s-t)-p(s+t)\leq 0$ when $s\geq 0$, $t\leq 0$. Because, in addition, $\hat p'(t)\geq 0$ when $t\leq 0$, the concluding quantity in (62) is nonpositive, meaning that the density $p*\hat p$ is even and is nonincreasing on $\mathbf{R}_+$.

2^0. Now let us verify that if $\mathcal{M}_*$ is the family of all even bounded and continuously differentiable functions on $\mathbf{R}$ that are nondecreasing on $\mathbf{R}_+$, then $f_+ = p * f \in \mathcal{M}_*$ whenever $f \in \mathcal{M}_*$. The only nontrivial claim is that f_+ is nondecreasing on $\mathbf{R}_+$, and when verifying it, we, the same as in 1^0, can assume that p is not only even and nonincreasing on $\mathbf{R}_+$, but is also smooth. In this case we have $f'_+(s) = \int f(s-r)p'(r)\,dr = \int_{-\infty}^0 (f(s-t) - f(s+t))p'(t)\,dt$. Assuming $s \ge 0$, $t \le 0$ and taking into account that f is even and is nondecreasing on $\mathbf{R}_+$, we have $f(s-t) = f(|s-t|) = f(|s|+|t|) \ge f(|s+t|) = f(s+t)$; because $p'(t) \ge 0$ when $t \le 0$, we conclude that $\int_{-\infty}^0 (f(s-t) - f(s+t))p'(t)\,dt \ge 0$ when $s \ge 0$.

3^0. Now we can conclude the proof of (v). We already know from 1^0 that the convolutions of every two of the four densities $p, \hat{p}, q, \hat{q}$ belong to $\mathcal{SU}$. All we should prove is that when $p(\cdot) \preceq_{\rm m} q(\cdot)$ and $\hat{p}(\cdot) \preceq_{\rm m} \hat{q}(\cdot)$, then $(p * \hat{p})(\cdot) \preceq_{\rm m} (q * \hat{q})(\cdot)$.

3^0.(a) Let us first verify that $(p * \hat{p})(\cdot) \preceq_{\rm m} (p * \hat{q})(\cdot)$, that is,

$$\int f(s)(p * \hat{p})(s)\,ds \le \int f(s)(p * \hat{q})(s)\,ds \tag{63}$$

for every even bounded function f that is nondecreasing on $\mathbf{R}_+$. By evident continuity reasons, it suffices to verify that (63) holds true for every $f \in \mathcal{M}_*$. Taking into account that p is even, we get

$$\int f(s)(p * \hat{p})(s)\,ds = \int f(s)p(s-t)\hat{p}(t)\,ds\,dt = \int (f * p)(t)\hat{p}(t)\,dt,$$

and by similar reasons

$$\int f(s)(p * \hat{q})(s)\,ds = \int (f * p)(t)\hat{q}(t)\,dt.$$

As we know from 2^0, $f * p \in \mathcal{M}_*$ whenever $f \in \mathcal{M}_*$, and (63) follows from the fact that $\hat{p}(\cdot) \preceq_{\rm m} \hat{q}(\cdot)$.

3^0.(b) The result of 3^0.(a) states that $p * \hat{p} \preceq_{\rm m} p * \hat{q}$. By the same result, but with swapped roles of "plain" and "$\widehat{\ }$" components, we further have $p * \hat{q} \preceq_{\rm m} q * \hat{q}$. As we know from (i), $\preceq_{\rm m}$ is a partial order, so that $p * \hat{p} \preceq_{\rm m} p * \hat{q}$ and $p * \hat{q} \preceq_{\rm m} q * \hat{q}$ imply the desired relation $p * \hat{p} \preceq_{\rm m} q * \hat{q}$. (v) is proved.

(vi): To prove that $\xi \preceq_{\rm m} \zeta$, observe that because $\xi \in \mathcal{SU}$ and ξ is supported on $[-1, 1]$, the density of ξ clearly is the weak limit of convex combinations of densities of uniform distributions on segments of the form $[-a, a]$ with $a \le 1$. Every one of these uniform distributions is $\preceq_{\rm m}$ the distribution of ζ by (iii), so that their convex combinations are $\preceq_{\rm m}$ the distribution of ζ by (ii). Applying (iv), we conclude that $\xi \preceq_{\rm m} \zeta$.

To prove that $\zeta \preceq_{\rm m} \eta$, let $p(\cdot)$ and $q(\cdot)$ be the respective densities (both of them belong to $\mathcal{SU}$), and let $\widetilde{P}(t) = \int_0^t p(s)\,ds = \frac{1}{2}\min[t, 1]$, $\widetilde{Q}(t) = \int_0^t q(s)\,ds$; this function is concave in $t \ge 0$ because $q(\cdot)$ is nonincreasing on $\mathbf{R}_+$. To prove that $\zeta \preceq_{\rm s} \eta$ is exactly the same as to verify that $\widetilde{P}(t) \ge \widetilde{Q}(t)$ for all $t \ge 0$. This is indeed the case when $0 \le t \le 1$, because $\widetilde{Q}(0) = 0$, $\widetilde{Q}'(0) = 1/2$, and $\widetilde{Q}$ is concave on $\mathbf{R}_+$, whereas $\widetilde{P}(t) = \frac{1}{2}t = \widetilde{Q}(0) + t\widetilde{Q}'(0)$ when $0 \le t \le 1$. And, of course, $\widetilde{P}(t) = 1/2 \ge \widetilde{Q}(t)$ when $t \ge 1$. (vi) is proved.

(vii): 1^0. Observe, first, that whenever $p(\cdot) \in \mathcal{SU}$, then there exists a sequence $\{p^t(\cdot) \in \mathcal{SU}\}_{t=1}^{\infty}$ such that

(a) every $p^t(\cdot)$ is a convex combination of densities of uniform symmetric w.r.t. 0 distributions;

(b) $p^t \to p$ as $t \to \infty$ in the sense that

$$\int f(s)p^t(s)\,ds \to \int f(s)p(s)\,ds \quad \text{as } t \to \infty$$

for every bounded piecewise-continuous function f on the axis.

2^0. We have the following:

Lemma A.1. *Let $Q \subset \mathbf{R}^d$ be a nonempty convex compact set symmetric w.r.t. the origin, and let $p_1(\cdot), \ldots, p_d(\cdot), q(\cdot) \in \mathcal{SU}$ be such that $p_1(\cdot), \ldots, p_{d-1}(\cdot)$ are densities of uniform distributions and $p_d(\cdot) \preceq_{\rm m} q(\cdot)$. Then,*

$$\int_Q p_1(x_1)p_2(x_2)\ldots p_{d-1}(x_{d-1})p_d(x_d)\,dx \ge \int_Q p_1(x_1)p_2(x_2)\ldots p_{d-1}(x_{d-1})q(x_d)\,dx. \tag{64}$$

Proof of Lemma A.1. Let Σ_l, $1 \le l < d$, be the support of the density p_l, so that Σ_l is a segment on the axis symmetric w.r.t. 0. Let us set $\Sigma = \Sigma_1 \times \cdots \times \Sigma_{d-1} \times \mathbf{R}$, $\hat{Q} = Q \cap \Sigma$, so that $\hat{Q}$ is a convex compact set symmetric w.r.t. the origin, and let

$$f(s) = \mathrm{mes}_{d-1}\{x \in \hat{Q}\colon x_d = s\}.$$

The function $f(s)$ is even; denoting by Δ the projection of $\hat{Q}$ onto the x_d-axis and applying the symmeterization principle of Brunn-Minkowski, we conclude that $f^{1/(d-1)}(s)$ is concave, even, and continuous on Δ, whence, of course, $f^{1/(d-1)}(s)$ is nonincreasing in $\Delta\cap\mathbf{R}_+$. We see that the function $f(s)$ is even, bounded, and nonnegative, and is nonincreasing on $\mathbf{R}_+$, whence

$$\int f(s)p_d(s)ds \geq \int f(s)q(s)\,ds \tag{65}$$

due to $p_d(\cdot)\preceq_{\mathrm{m}} q(\cdot)$. It remains to note that the left- and the right-hand sides in (64) are proportional, with a common positive coefficient, to the respective sides in (65). □

3^0. Now we can complete the proof of (vii). Clearly, all we need is to show that if $p_1(\cdot),\ldots,p_d(\cdot),q_d(\cdot)\in\mathcal{SU}$ and $p_d(\cdot)\preceq_{\mathrm{m}} q_d(\cdot)$, then

$$\int_Q p_1(x_1)p_2(x_2)\ldots p_{d-1}(x_{d-1})p_d(x_d)\,dx \geq \int_Q p_1(x_1)p_2(x_2)\ldots p_{d-1}(x_{d-1})q_d(x_d)\,dx.$$

By continuity argument and in view of 1^0, it suffices to verify the same relation when $p_1(\cdot),\ldots,p_{d-1}(\cdot)$ are convex combinations of densities of uniform and symmetric w.r.t. the origin distributions. Because both sides in our target inequality are linear in every one of $p_1,\ldots,p_{d-1}$, to prove the latter fact is the same as to prove it when every one of $p_1,\ldots,p_{d-1}$ is a uniform distribution symmetric w.r.t. the origin. In the latter case, the required statement is given by Lemma A.1. □

Proof of Proposition 5.2.(ix). Under the premise of the statement to be proved, Q contains the centered at the origin $\|\cdot\|_2$-ball of the radius $\mathrm{ErfInv}(\chi)$ (Lemma 3.1.(i)), so that the Minkowski function $\theta(x)=\inf\{t: t^{-1}x\in Q\}$ of Q belongs to $\mathcal{CF}_n$. Let $\beta\in[1,\gamma)$, and let $\delta(x)=\max[\theta(x)-\beta,0]$. We clearly have $\delta(\cdot)\in\mathcal{CF}_n$, so that

$$\int \delta(x)\,dP_\xi(x)\leq\int\delta(x)\,dP_\eta(x). \tag{66}$$

For $s\geq\beta$, let $p(s)=\mathrm{Prob}\{\eta\notin sQ\}=\mathrm{Prob}\{\delta(\eta)>s-\beta\}$. By Lemma 3.1.(ii) we have

$$s\geq\beta \;\Rightarrow\; p(s)\leq \mathrm{Erf}(s\mathrm{ErfInv}(\chi)). \tag{67}$$

We have $\int\delta(x)\,dP_\eta(x)=-\int_\beta^\infty(s-\beta)\,dp(s)=\int_\beta^\infty p(s)\,ds\leq\int_\beta^\infty\mathrm{Erf}(s\mathrm{ErfInv}(\chi))\,ds$, (the concluding inequality is due to (67)), whence $\int\delta(x)\,dP_\xi(x)\leq\int_\beta^\infty\mathrm{Erf}(s\mathrm{ErfInv}(\chi))\,ds$ by (66). Now, when $\xi\notin\gamma Q$, we have $\delta(\xi)\geq\gamma-\beta$. Invoking the Tschebyshev inequality, we arrive at

$$\mathrm{Prob}\{\xi\notin\gamma Q\}\leq\frac{\mathbf{E}\{\delta(\xi)\}}{\gamma-\beta}\leq\frac{1}{\gamma-\beta}\int_\beta^\infty\mathrm{Erf}(s\mathrm{ErfInv}(\chi))\,ds. \quad\square$$

Proof of Property (P) via Noncommutative Khintchine Inequality. We start with the following deep fact of functional analysis due to Lust-Piquard [11], Pisier [16], and Buchholz [7]; see Tropp [17, Proposition 10]:

Noncommutative Khintchine Inequality (NKI): *Let $\eta\sim\mathcal{N}(0,I_d)$, and let $Q_1,\ldots,Q_d$ be deterministic matrices. Then for every $p\in[2,\infty)$ one has*

$$\mathbf{E}\left\{\left|\sum_{l=1}^d\eta_lQ_l\right|_p^p\right\}\leq[2^{-1/4}\sqrt{p\pi/e}]^p\max\left[\left|\sum_{l=1}^dQ_lQ_l^T\right|_{p/2},\left|\sum_{l=1}^dQ_l^TQ_l\right|_{p/2}\right]^{p/2}, \tag{68}$$

where $|A|_p=\|\sigma(A)\|_p$, $\sigma(A)$ being the vector of singular values of a matrix A.

As an immediate corollary of NKI, we have the following

Proposition A.1. *Let $B_l\in\mathbf{S}^n$, $n\geq2$, be such that $\sum_{l=1}^dB_l^2\preceq I$, and let*

$$c_n=\inf_{2\leq p<\infty}[2^{-1/4}\sqrt{p\pi/e}n^{1/p}].$$

Then $c_n\leq O(1)\sqrt{\ln n}$ and for all $\chi\in(0,1/2)$ one has

$$\begin{gathered}\mathrm{Prob}\left\{-\Upsilon(\chi)I\preceq\sum_{l=1}^d\zeta_lB_l\preceq\Upsilon(\chi)I\right\}\geq1-\chi,\\ \Upsilon(\chi)=\begin{cases}16c_n\mathrm{ErfInv}(0.3\chi), & \textit{we are in the case of (A.1)}\\ c_n/\chi, & \textit{we are in the case of (A.2)}\end{cases}\end{gathered} \tag{69}$$

PROOF. When $p = \max[2, \ln n]$, we clearly have $2^{-1/4}\sqrt{p\pi/e}n^{1/p} \leq O(1)\sqrt{p} \leq O(1)\sqrt{\ln n}$, so that $c_n \leq O(1)\sqrt{\ln n}$. Now assume that (A.2) is the case. Applying NKI with $Q_l = B_l$ and taking into account that $|\sum_{l=1}^d \zeta_l B_l|_p \leq \|\sum_{l=1}^d \zeta_l B_l\|$ and taking into account that $|\sum_{l=1}^d B_l^2|_{p/2}^{p/2} \leq n$ due to $\sum_l B_l^2 \preceq I_n$, we get $\mathbf{E}\{\|\sum_{l=1}^d \zeta_l B_l\|^p\} \leq [2^{-1/4}\sqrt{p\pi/e}]^p n$, whence $\text{Prob}\{\|\sum_l \zeta_l B_l\| > \Upsilon\} \leq [2^{-1/4}\sqrt{p\pi/e}]n^{1/p}$ for every $\Upsilon > 0$ by Tschebyshev Inequality. The resulting bound is valid for every $p \in [2, \infty)$, and the (A.2)-version of (69) follows. The (A.1)-version of (69) follows from the (A.2)-version of this relation due to Gaussian majorization. □

Acknowledgments. This research was partly supported by BSF Grant 2002038; research of the second author was supported by NSF Grant 0619977. The authors are greatly indebted to A. Man-Cho So, who brought to their attention the noncommutative Khintchine inequality.

References

[1] Ben-Tal, A., A. Nemirovski. 1997. Stable truss topology design via semidefinite programming. *SIAM J. Optim.* **7**(4) 991–1016.
[2] Ben-Tal, A., A. Nemirovski. 2001. *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications*. MPS-SIAM Series on Optimization, SIAM, Philadelphia.
[3] Ben-Tal, A., A. Nemirovski. 2002. On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty. *SIAM J. Optim.* **12** 811–833.
[4] Ben-Tal, A., A. Nemirovski, C. Roos. 2002. Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM J. Optim.* **13** 535–560.
[5] Ben-Tal, A., A. Nemirovski, C. Roos. 2003. Extended matrix cube theorems with applications to μ-Theory in control. *Math. Oper. Res.* **28** 497–523.
[6] Borell, C. 1975. The Brunn-Minkowski inequality in Gauss space. *Inventiones Mathematicae* **30**(2) 207–216.
[7] Buchholz, A. 2001. Operator Khintchine inequality in the non-commutative probability. *Mathematische Annalen* **391** 1–16.
[8] Calafiore, G., M. C. Campi. 2005. Uncertain convex programs: Randomized solutions and confidence levels. *Math. Programming* **102** 25–46.
[9] de Farias, D. P., B. Van Roy. 2004. On constraint sampling in the linear programming approach to approximate dynamic programming. *Math. Oper. Res.* **29** 462–478.
[10] Erdogan, G., G. Iyengar. 2006. Ambiguous chance constrained problems and robust optimization. *Math. Programming B* **107**(1–2) 37–61.
[11] Lust-Piquard, F. 1986. Inégalités de khintchine dans C_p $(1 < p < \infty)$. *Comptes Rendus de l'Académie des Sciences de Paris, Série I* **393**(7) 289–292.
[12] Man-Cho So, A. 2008. Improved approximation bound for quadratic optimization problems with orthogonality constraints. Technical report, Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, N. T., Hong Kong.
[13] Nemirovski, A. 2007. Sums of random symmetric matrices and quadratic optimization under orthogonality constraints. *Math. Programming B* **109** 283–317.
[14] Nemirovski, A., A. Shapiro. 2006. Scenario approximations of chance constraints. G. Calafiore, F. Dabbene, eds. *Probabilistic and Randomized Methods for Design Under Uncertainty*. Springer-Verlag, New York, 3–48.
[15] Nemirovski, A., A. Shapiro. 2006. Convex approximations of chance constrained programs. *SIAM J. Optim.* **17**(4) 969–996.
[16] Pisier, G. 1998. Non-commutative vector valued L_p spaces and completely p-summing maps. *Astérisque* **247** 1–111.
[17] Tropp, J. A. 2008. The random paving property for uniformly bounded matrices. *Studia Mathematica* **185** 67–82.