On Safe Tractable Approximations of Chance Constrained Linear Matrix Inequalities*

Aharon Ben-Tal† Arkadi Nemirovski‡

Abstract

In the paper, we consider the chance constrained version

$$\mathrm{Prob}\{A_0[x] + \sum_{i=1}^{d} \zeta_i A_i[x] \succeq 0\} \geq 1 - \epsilon,$$

of an affinely perturbed Linear Matrix Inequality constraint; here $A_i[x]$ are symmetric matrices affinely depending on the decision vector x, and $\zeta_1, ..., \zeta_d$ are independent of each other random perturbations with light tail distributions (e.g., bounded or Gaussian). Constraints of this type, playing the central role in Chance Constrained Linear/Conic Quadratic/Semidefinite Programming, typically are computationally intractable, which makes natural to look for their tractable approximations. The goal of this paper is to develop such an approximation. Our approximation is based on measure concentration results and is given by an explicit system of LMIs and thus is computationally tractable; it is also safe, meaning that a feasible solution of the approximation is feasible for the chance constraint as well.

Key words: chance constraints, linear matrix inequalities, convex programming, measure concentration.

1 Introduction

In this paper, we focus on 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 in x and in ζ mapping taking values in the space $\mathbf{S}^n$ of symmetric $n \times n$ matrices:

$$\mathcal{A}(x, \zeta) = \mathcal{A}_0[x] + \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x] \tag{2}$$

*research was partly supported by the BSF grant # 2002038. Research of the second author was partly supported by the NSF grant # 0619977

†Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, abental@ie.technion.ac.il

‡School of Industrial and Systems Engineering, Georgia Institute of Technology, nemirovs@isye.gatech.edu

with affine in x matrices $\mathcal{A}_0[x], ..., \mathcal{A}_d[x] \in \mathbf{S}^n$, and $A \succeq B$ 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 are interested 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 when the problem is solved), the question of how to process an uncertain LMI constraint is of definite interest.

For the time being, there are two major 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 the most 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 "uncertain-but-bounded" model of data perturbations where all our a priori knowledge of ζ is that it runs through a given *uncertainty set* $\mathcal{Z}$. In this case, the most 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 of ϵ^{-1}, and this is too computationally demanding when ϵ is small. Another difficulty, arising independently of what ϵ is, comes from the fact that the feasible set of (3) usually is non-convex, which makes problematic optimization under this constraint. The latter complication does not arise with the Robust Optimization approach – the feasible set of (4) always is convex; unfortunately, the first difficulty – impossibility to check efficiently whether this *semi-infinite* convex constraint is satisfied at a given point – becomes here even more severe than in the case of chance constrained LMI. Severe tractability difficulties with (3) and (4) "as they are" make it natural to replace such a constraint with its *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 of interest. For the time being, "tight", in 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 $\mathcal{A}(x, \zeta)$ [3, 4, 5]. In this paper, we focus solely on chance constrained LMIs (3). In this case, seemingly the only known from literature safe tractable approximation of the constraint is the one given by the general *scenario approach*. For a chance constrained optimization program

$$\min_x \left\{ f_0(x) : \mathrm{Prob}\left\{ f_i(x, \zeta) \leq 0 \right\} \geq 1 - \epsilon, \, i = 1, ..., I \right\},$$

its scenario approximation is the *random* optimization program

$$\min_x \left\{ f_0(x) : f_i(x, \zeta^j) \leq 0, \, i = 1, ..., I, j = 1, ..., J \right\},$$

where $\zeta^1, ..., \zeta^J$ is a sample of independent realizations of ζ. Theoretical justification of this natural approximation scheme is presented in [6, 7]. In particular, it is shown in [6] that if $f_0(x)$, $f_i(x,\zeta)$, $i = 1, ..., I$, are convex in $x \in \mathbf{R}^m$ and the sample size J is large enough:

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

then an optimal solution to the approximation, up to probability $\leq \delta$ of "bad sampling", is feasible for the chance constrained problem of interest. (For substantial extensions of this remarkable result to the case of *ambiguously* chance constrained convex problems, see [8]). While being 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 of $1/\epsilon$, 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, $\ell = 1, ..., d$, this drawback can be circumvented by a kind of importance sampling [13]. In this paper, we work under the same assumptions as in [13], 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 chance constrained versions of these LMIs. In contrast to the purely simulation-based approximations of [6, 8, 13], 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 the parameters in advance, thus arriving at a provably safe and completely simulation-free tractable approximation of (3). However, in order to reduce the conservatism of the approximation, we allow for an "optimistic" choice of the parameters and introduce a specific simulation-based post-optimization *validation procedure* which allows either to justify our "optimistic guess" (and thus guarantees, "up to probability $\leq \delta$ of bad sampling", that the solution we end up with is feasible for the chance constraint of interest), or 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 the validation procedure is completely independent of how small is the tolerance ϵ; all we need is $J \geq O(1)\ln(1/\delta)$.

The rest of the paper is organized as follows. In Section 2 we make our standing assumptions and outline and motivate our approximation strategy. This strategy is fully developed in Sections 3.1, 3.2. In Section 4 we consider two important special cases of (3). In the first of them, all matrices $\mathcal{A}_\ell[x]$, $\ell = 0, 1, ..., d$, are diagonal, that is, we are speaking about a bunch of randomly perturbed scalar linear inequalities, or, which is the same, about chance constrained Linear Programming. In the second special case, the matrices $\mathcal{A}_\ell[x]$, $\ell = 1, ..., d$, are of the form $\lambda_\ell(x)G(x) + e(x)f_\ell^T(x) + f_\ell(x)e^T(x)$, where $e(x)$ and $f_\ell(x)$ are vectors (and, as always in this paper, $\mathcal{A}_\ell[x]$ is affine in x). This situation covers the case when (1) is in fact a randomly perturbed *Conic Quadratic Inequality* $\|A[x]\zeta + b[x]\|_2 \leq c^T[x]\zeta + d[x]$ ($A[x], b[x], c[x], d[x]$ are affine in x); indeed,

$$\|A[x]\zeta + b[x]\|_2 \leq c^T[x]\zeta + d[x] \Leftrightarrow \underbrace{\left[\begin{array}{c|c} d[x] & b^T[x] \\ \hline b[x] & d[x]I \end{array}\right]}_{\mathcal{A}_0[x]} + \sum_{\ell=1}^{d} \zeta_\ell \underbrace{\left[\begin{array}{c|c} c_\ell[x] & A_\ell^T[x] \\ \hline A_\ell[x] & c_\ell[x]I \end{array}\right]}_{\mathcal{A}_\ell[x]} \succeq 0, \tag{6}$$

where $A_\ell[x]$ are the columns of $A[x]$, and $c_\ell[x]$ are the entries of $c[x]$. Note that "fully analytic" safe tractable approximations of chance constrained LPs were recently proposed in [11]; Section 4 contains a comparison of approximations from [11] with the one developed in this paper. Section 5 presents techniques allowing to reduce building a safe approximation of the chance constrained LMI (3) with *partially known* "light tail" distributions of (independent of each other) perturbations ζ_ℓ to a similar task for an appropriately chosen reference distribution of ζ, most notably, a Gaussian one. The concluding Section 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

$$\min_x \left\{ c^T x : \begin{array}{l} F(x) \leq 0 \\ \mathrm{Prob}\{\mathcal{A}_0[x] + \sum_{\ell=1}^d \zeta_\ell \mathcal{A}_\ell[x] \succeq 0\} \geq 1-\epsilon \end{array} \right\}, \tag{7}$$

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

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

if whenever a vector x can be extended to a feasible solution 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 of the constraint, we get an optimization problem in variables x, u with efficiently computable convex constraints, that is, we get a 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 random perturbations:

> **Assumption A:** *The scalar random variables $\zeta_1, ..., \zeta_d$ are mutually independent and with zero means. Besides this, either* (a) *all ζ_ℓ have bounded ranges, or* (b) *all ζ_ℓ are Gaussian.*
> Note that applying deterministic scalings $\zeta_\ell \mapsto \zeta_\ell / s_\ell$, $\mathcal{A}_\ell[x] \mapsto s_\ell \mathcal{A}_\ell[x]$, 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 $\zeta_\ell \sim \mathcal{N}(0,1)$ for all ℓ. Therefore from now on, if otherwise is not stated explicitly, we assume that we are either in the case of

A.1. *ζ_ℓ is supported on* $[-1,1]$,

or in the case of

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

2.1 The strategy

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

$$A_0 + \sum_{\ell=1}^{d} \zeta_\ell A_\ell \succeq 0 \tag{9}$$

to be satisfied with probability at least $1-\epsilon$; here $A_0, ..., A_d$ 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 ask for a sufficient condition for the relation

$$\mathrm{Prob}\{-A_0 \preceq S := \sum_{\ell=1}^{d} \zeta_\ell A_\ell \preceq A_0\} \geq 1-\epsilon. \tag{10}$$

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

$$\mathrm{Prob}\{-I \preceq \widehat{S} := \sum_{\ell=1}^{d} \zeta_\ell \widehat{A}_\ell \preceq I\} \geq 1-\epsilon,\ \widehat{A}_\ell = A_0^{-1/2} A_\ell A_0^{-1/2}. \tag{11}$$

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

$$\mathbf{E}\{\widehat{S}^2\} = \sum_{\ell=1}^{d} \widehat{A}_\ell^2 \preceq O(1)(\ln(1/\epsilon))^{-1} I \tag{12}$$

with a moderate absolute constant $O(1)$. Thus, the condition (12) 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. Now assume for a moment that (12) is not only necessary, but a *sufficient* as well condition for (11) to be valid. Then we are basically done: it is immediately seen that the condition (12) can be equivalently reformulated as an LMI in variables $A_0, ..., A_d$, specifically, the LMI

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

with $\gamma = \sqrt{O(1)\ln(1/\epsilon)}$. It follows that when $A_\ell = \mathcal{A}_\ell[x]$, $\ell = 0, 1, ..., d$, affinely depend on a decision vector x, as it is the case in the situation we are interested in, our, hopefully sufficient for the validity of (10) (and thus – for the validity of (9) as well), condition (13) becomes an LMI in variables x and thus provides us with safe tractable approximation of (8).

Unfortunately, we do *not* know at present whether the condition (12) indeed is sufficient for the validity of (11). It is shown in [12] that the strongest result in this direction we can hope for could be the implication

$$\sum_{\ell=1}^{d} \widehat{A}_\ell^2 \preceq I \Rightarrow \forall t \geq O(1)\sqrt{\ln n} : \mathrm{Prob}\{\widehat{S} \preceq tI\} \geq 1 - O(1)\exp\{-O(1)t^2\}. \tag{14}$$

Were this implication true, the condition (13) would be sufficient for the validity of (11), provided that $\ln(1/\epsilon) \geq O(1)\ln n$. While it is *conjectured* in [12] that (14) is indeed true, the strongest for the time being *provably true* result in this direction is that *when replacing* $\sqrt{\ln n}$ *in* (14) *with* $n^{1/6}$, (14) *indeed becomes true, provided that* $\zeta_1, ..., \zeta_d$ *satisfy Assumption A and, in addition,* ζ_ℓ *have zero third moments*, see [12, Theorem 2].

The main idea of this paper is that we can, to some extent, circumvent the difficulty coming from the fact that (14) is a conjecture rather than a provable statement: we can act *as if* the conjecture were true and then use a cheap simulation-based procedure to validate the result (or to refine the conjecture).

3 Approximating chance constrained LMI's

3.1 Preliminaries on measure concentration

Our strategy heavily exploits the following fact:

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

$$\begin{array}{ll} (a) & \mathrm{Arrow}(B_0, B_1, ..., B_d) \succeq 0 \\ (b) & \mathrm{Prob}\left\{-\Upsilon B_0 \preceq \sum_{\ell=1}^{d} \zeta_\ell B_\ell \preceq \Upsilon B_0\right\} \geq 1 - \chi. \end{array} \tag{15}$$

Then

$$\begin{array}{l} \gamma \geq 1 \Rightarrow \mathrm{Prob}\left\{-\gamma\Upsilon B_0 \preceq \sum_{\ell=1}^{d} \zeta_\ell B_\ell \preceq \gamma\Upsilon B_0\right\} \geq 1 - \vartheta(\chi, \gamma), \\ \vartheta(\chi, \gamma) = \left\{ \begin{array}{ll} \frac{1}{1-\chi}\exp\{-\Upsilon^2(\gamma-1)^2/16\}, & \zeta \text{ satisfies A.1} \\ \exp\{-\frac{\phi^2(\chi)\gamma^2}{2}\}, & \zeta \text{ satisfies A.2} \end{array} \right. \end{array} \tag{16}$$

where $\phi(\cdot)$ *is the inverse error function:*

$$\int_{\phi(r)}^{\infty} \frac{1}{\sqrt{2\pi}} \exp\{-s^2/2\} ds = r, \; 0 < r < 1. \tag{17}$$

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 $B_0, B_1, ..., B_d$ to the matrices $I, B_0^{-1/2} B_1 B_0^{-1/2}, ..., B_0^{-1/2} B_d B_0^{-1/2}$, we immediately reduce the situation to the one with $B_0 = I$, which we assume from now on. In this case, (15.a) becomes simply $\sum_{\ell=1}^{d} B_\ell^2 \preceq I$. With this normalization, the validity of (16) in the case of A.1 is readily given by Lemma 1 in [12], and in the case of A.2 – by Theorem 1 in [11]. □

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

$$\vartheta = \begin{cases} \frac{1}{\Upsilon + 4\sqrt{\ln(\epsilon^{-1}(1-\chi)^{-1})}}, & \text{we are in the case of A.1} \\ \Upsilon^{-1} \min\left[\frac{\phi(\chi)}{\sqrt{2\ln(1/\epsilon)}}, 1\right], & \text{we are in the case of A.2} \end{cases} \tag{18}$$

where $\phi(\cdot)$ *is the inverse error function* (17). *Assume, further, that symmetric matrices* $A_0, ..., A_d$ *satisfy*

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

and, in addition, that

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

Then

$$\mathrm{Prob}\left\{-A_0 \preceq \sum_{\ell=1}^{d} \zeta_\ell A_\ell \preceq A_0\right\} \geq 1 - \epsilon. \tag{21}$$

Proof. Assume, first, that we are in the case of A.1, and let us set $\gamma = 1 + 4\Upsilon^{-1}\sqrt{\ln(\epsilon^{-1}(1-\chi)^{-1})}$, so that $\vartheta = \frac{1}{\gamma\Upsilon}$. Under the premise of Corollary, the matrices $B_0 = \vartheta A_0, B_1 = A_1, ..., B_d = A_d$ satisfy the relations (15) by (19), (20), whence, applying Theorem 3.1 to the matrices $B_0, ..., B_d$ and the just defined γ, we get

$$\begin{array}{l} \mathrm{Prob}\{-\underbrace{\gamma\Upsilon B_0}_{=A_0} \preceq \underbrace{\sum_{\ell=1}^{d} \zeta_\ell B_\ell}_{=\sum_{\ell=1}^{d}\zeta_\ell A_\ell} \preceq \gamma\Upsilon B_0\} \geq 1 - \frac{1}{1-\chi}\exp\{-(\gamma-1)^2\Upsilon^2/16\} \\ = 1 - \frac{1}{1-\chi}\exp\{\ln(\epsilon(1-\chi))\} = 1 - \epsilon, \end{array}$$

as required.

In the case of A.2 the reasoning is completely similar. □

3.2 The approximation

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

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

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

they satisfy the relation

$$\mathrm{Prob}\{-\Upsilon B_0 \preceq \sum_{\ell=1}^{d} \zeta_\ell B_\ell \preceq \Upsilon B_0\} \geq 1 - \chi \tag{23}$$

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

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

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

Let us set $B_0^* = \vartheta\mathcal{A}_0[x_*], B_1^* = \mathcal{A}_1[x_*], ..., B_d^* = \mathcal{A}_d[x_*]$; 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 of interest (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 *Validation procedure* as follows.

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

$$\pi = \min_{0\le p\le 1}\left\{p : \sum_{i=1}^{M}\left(\begin{array}{c} N \\ i \end{array}\right)(1-p)^i p^{N-i} \ge \delta\right\},$$

where $\delta \in (0,1)$ is a chosen in advance unreliability of our conclusions (say, $\delta = 10^{-12}$). We then check whether $\pi \ge 1-\chi$; 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 somehow the value of Υ and/or reducing 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 to validate the feasibility of x_* in the case when $p_* < 1-\chi$ is at most δ; since in the case of $p_* \ge 1-\chi$ x_* provably is feasible for (7), our randomized approximation scheme, while being tractable, indeed is safe up to probability of bad sampling $\le \delta$. □

The advantage of the outlined validation routine is that *when working with* χ *not too close to 0* (and we can afford ourselves to work with whatever $\chi \in (0, 1/2)$, say, with $\chi = 0.4$ or $\chi = 0.1$), *in the case of*

$$\mathrm{Prob}\left\{-\Upsilon B_0^* \preceq \sum_{\ell=1}^{d}\zeta_\ell B_\ell^* \preceq \Upsilon B_0^*\right\} \ge 1 - 0.8\chi, \tag{25}$$

(that is, in the case when an assumption slightly stronger than the one

$$\mathrm{Prob}\left\{-\Upsilon B_0^* \preceq \sum_{\ell=1}^{d}\zeta_\ell B_\ell^* \preceq \Upsilon B_0^*\right\} \ge 1 - \chi$$

we want to validate takes place) *the cardinality* N *of the sample which is sufficient to validate, with close to 1 probability* $1-\nu$, *the feasibility of* x_* *for* (7) *should not be too large.* Rough estimate shows that it suffices to take

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

With $\delta = \nu = 10^{-8}$, $\chi = 0.4$, this formula yields $N = 23026$; a more accurate computation shows that $N = 4541$ 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

$$\rho_*(\bar{c}) = \max_{x,\rho} \left\{ \rho : \begin{array}{l} F(x) \leq 0, c^T x \leq \bar{c}, \\ \text{Prob}\left\{ \mathcal{A}_0[x] + \rho \sum_{\ell=1}^d \zeta_\ell \mathcal{A}_\ell[x] \succeq 0 \right\} \geq 1 - \epsilon \end{array} \right\}, \tag{26}$$

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

$$\text{Opt} = \min_x \left\{ c^T x : F(x) \leq 0, \mathcal{A}_0[x] \succeq 0 \right\},$$

and then build the "tradeoff curve" $\tau(s) = \rho_*(\text{Opt} + s)$, $s > 0$ 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 exactly in the same fashion as above, but aiming at the problem (26) rather than at the problem (7), we would arrive at the approximation

$$\max_{x,\rho} \left\{ \rho : \begin{array}{l} F(x) \leq 0, c^T x \leq \bar{c}), \\ \text{Arrow}(\vartheta(\rho)\mathcal{A}_0[x], \mathcal{A}_1[x], ..., \mathcal{A}_d[x]) \succeq 0 \end{array} \right\} \tag{27}$$

where $\vartheta(\rho)$ is given by (18) with Υ replaced with $\rho\Upsilon$. Since $\vartheta(\rho)$ 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 quasiconvex* (and thus – efficiently solvable) *optimization problem*

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

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

$$\text{Prob}\left\{ -\mathcal{A}_0[x_*] \preceq \rho \sum_{\ell=1}^d \zeta_\ell \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*] \right\} \geq 1 - \epsilon. \tag{29}$$

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 $\mathrm{Arrow}(\vartheta\mathcal{A}_0[x_*], \mathcal{A}_1[x_*], ..., \mathcal{A}_d[x_*]) \succeq 0$ (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, since $\vartheta_*(x_*) = 0$ means that $\mathcal{A}_\ell[x_*] = 0$, $\ell = 1, ..., d$; since $\mathcal{A}_0[x_*] \succeq 0$ 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 $B_0 = \vartheta_*(x_*)\mathcal{A}_0[x_*]$, $B_\ell = \mathcal{A}_\ell[x_*]$, $\ell = 1, ..., d$, so that $\mathrm{Arrow}(B_0, B_1, ..., B_d) \succeq 0$. 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:

<u>*1. Building a grid of values of ρ.*</u> As we remember from Section 3.1, the implication (22)⇒(23) indeed holds true for "safe" values of Υ and χ, e.g., for $\chi = \chi_s = 0.4$ and $\Upsilon = \Upsilon_s = O(1)n^{1/6}$ 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

$$\mathrm{Prob}\left\{ -\mathcal{A}_0[x_*] \preceq \rho_s \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*] \right\} \geq 1 - \epsilon. \tag{30}$$

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

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

$$\mathrm{Prob}\left\{ -\mathcal{A}_0[x_*] \preceq \rho^+ \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*] \right\} \geq 1 - \epsilon$$

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

$$-\mathcal{A}_0[x_*] \preceq \rho^i \sum_{\ell=1}^{d} \zeta_\ell^i \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*]$$

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

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 $r_1 := \rho_s < r_2 < ... < r_K := \rho^+$ forms a geometric progression. This sequence forms a grid which we are about to use when building $\rho_*(x_*)$.

<u>*2. Running simulations.*</u> At this step, we

1. Generate a training sample of N independent realizations $\zeta^1, ..., \zeta^N$ of ζ

2. For every $k = 1, ..., K$ compute the integers

$$M_k = \mathrm{Card}\{i \leq N : \neg(-\mathcal{A}_0[x_*] \preceq r_k \sum_{\ell=1}^{d} \zeta_\ell^i \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*])\}$$

and then – the reals

$$\widehat{\chi}_k = \max\left\{\chi \in [0,1] : \sum_{i=1}^{M_k} \left(\begin{array}{c} N \\ i \end{array}\right) \chi^i (1-\chi)^{N-i} \geq \delta/K\right\}.$$

Note that if

$$\chi_k = \mathrm{Prob}\left\{\neg(-\mathcal{A}_0[x_*] \preceq r_k \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*])\right\},$$

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

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

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

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

and in the case of A.2 we set

$$\rho_*(x_*) = \max_k\left\{r_k \min\left[\frac{\phi(\widehat{\chi}_k)}{\sqrt{2\ln(1/\epsilon)}}, 1\right] : 1 \leq k \leq K, \widehat{\chi}_k < 1/2\right\}. \tag{33}$$

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

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 $\widehat{\chi}_k \geq \chi_k$ for all $k = 1, ..., K$ (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 $\rho_* \equiv \rho_*(x_*) = \rho_{\rm s}$, 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 a well defined formula (33).

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

$$\Upsilon_k = \frac{1}{r_k\vartheta_*(x_*)},\ \vartheta_k = \frac{1}{\Upsilon_k + 4\sqrt{\ln(\epsilon^{-1}(1-\widehat{\chi}_k)^{-1})}},\ A_0 = \frac{\vartheta_*(x_*)}{\vartheta_k}\mathcal{A}_0[x_*],\ A_\ell = \mathcal{A}_\ell[x_*],\ \ell = 1, ..., d.$$

Then

$$\begin{array}{l}
\mathrm{Arrow}(\vartheta_k A_0, A_1, ..., A_d) = \mathrm{Arrow}(\vartheta_*(x_*)\mathcal{A}_0[x_*], \mathcal{A}_1[x_*], ..., \mathcal{A}_d[x_*]) \succeq 0,\\
\mathrm{Prob}\{-\underbrace{\Upsilon_k\vartheta_k A_0}_{r_k^{-1}\mathcal{A}_0[x_*]} \preceq \underbrace{\sum_{\ell=1}^{d} \zeta_\ell A_\ell}_{=\sum_{\ell=1}^d \zeta_\ell \mathcal{A}_\ell[x_*]} \preceq \Upsilon_k\vartheta_k A_0\} = \mathrm{Prob}\{-\mathcal{A}_0[x_*] \preceq r_k \sum_{\ell=1}^d \zeta_\ell \mathcal{A}_\ell[x_*] \preceq \mathcal{A}_0[x_*]\}\\
\geq 1 - \widehat{\chi}_k
\end{array}$$

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

The result for case A.2 can be proved in a completely similar way.. $\square$

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

$$\mathrm{Arrow}(B_0, B_1, ..., B_d) \succeq 0 \Rightarrow \mathrm{Prob}\{-\Upsilon B_0 \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell \preceq \Upsilon B_0\} \geq 1-\chi. \tag{34}$$

In particular, using in these Υ, χ in the approximation scheme of Section 3.2, we can avoid the necessity to use the Validating procedure.

4.1 Diagonal case

The first of the special cases we are about to consider is the one where $\mathcal{A}_0[x], \mathcal{A}_1[x], ..., \mathcal{A}_d[x]$ in (7) are diagonal matrices; we refer to this situation as to the *Diagonal case*. Note that in spite of its extreme simplicity, the Diagonal case is of definite interest: this 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 $B_\ell = \mathrm{Diag}\{B_\ell^1, ..., B_\ell^s\}$, $\ell = 0, 1, ..., d$, 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*

$$\forall \chi \in (0, 1/2) : \mathrm{Arrow}(B_0^j, ..., B_d^j) \succeq 0 \Rightarrow \mathrm{Prob}\{-\Upsilon(\chi) B_0^j \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell^j \preceq \Upsilon(\chi) B_0^j\} \geq 1-\chi.$$

Then one has

$$\forall \chi \in (0, 1/2) : \mathrm{Arrow}(B_0, ..., B_d) \succeq 0 \Rightarrow \mathrm{Prob}\{-\Upsilon(\chi/s) B_0 \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell \preceq \Upsilon(\chi/s) B_0\} \geq 1-\chi.$$

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

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

> **A.3.** *$\zeta_1, ..., \zeta_d$ are mutually independent and $\mathbf{E}\{\exp\{\zeta_\ell^2\}\} \leq \exp\{1\}$, $\ell = 1, ..., d$.*

(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*

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

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

$$\Upsilon = \Upsilon^{(n)}{\rm G}(\chi) = \phi(\chi/(2n)) \leq \sqrt{2\ln(n/\chi)}.$$

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

$$\|h\|_2 \leq 1 \Rightarrow \mathrm{Prob}\{|\sum_{\ell=1}^{d} \zeta_\ell h_\ell| > \Upsilon(\chi)\} \leq \chi,\ 0 < \chi < 1/2, \tag{35}$$

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

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

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 $\gamma \in \mathbf{R}$, $s_\ell = \sum_{r=1}^{\ell} \gamma h_r \zeta_r$, and $J = \{\ell : |h_\ell \gamma| > \sqrt{3/2}\}$. We have

$$\mathbf{E}\{\exp\{s_\ell\}\} = \mathbf{E}\{\exp\{s_{\ell-1}\}\exp\{\gamma h_\ell \zeta_\ell\}\} = \mathbf{E}\{\exp\{s_\ell\}\} \cdot \Theta_\ell, \quad \Theta_\ell = \mathbf{E}\{\gamma h_\ell \zeta_\ell\}. \tag{37}$$

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

$$\Theta_\ell \leq \left\{\begin{array}{ll} \exp\{2\gamma^2 h_\ell^2/3\}, & \ell \notin J \\ \exp\{7/12 + 2\gamma^2 h_\ell^2/3\}, & \ell \in J \end{array}\right. \tag{38}$$

Indeed, it is easily seen that

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

for all $t \in \mathbf{R}$, whence $\mathbf{E}\{\exp\{\gamma h_\ell \zeta_\ell\}\} \leq \mathbf{E}\{\exp\{2\gamma^2 h_\ell^2 \zeta_\ell^2/3\}\}$; when $\ell \notin J$, the latter expectation is at most $\left(\mathbf{E}\{\exp\{\zeta_\ell^2\}\}\right)^{2\gamma^2 h_\ell^2/3}$ by Hölder Inequality, as required in (38). Now let $\ell \in J$. We have $|\gamma h_\ell s| \leq s^2 + \gamma^2 h_\ell^2/4$ for all s, whence

$$\begin{array}{l} \mathbf{E}\{\exp\{\gamma h_\ell \zeta_\ell\}\} \leq \exp\{\gamma^2 h_\ell^2/4\}\mathbf{E}\{\exp\{\zeta_\ell^2\}\} \leq \exp\{1 + \gamma^2 h_\ell^2/4\} \\ \leq \exp\{7/12 + 2\gamma^2 h_\ell^2/3\} \end{array}$$

as required in (38).

Combining (37) and (38), we get

$$\begin{array}{l}\mathbf{E}\{\exp\{\gamma\sum_{\ell=1}^d h_\ell\zeta_\ell\}\}\leq\exp\{2\gamma^2[\sum_{\ell=1}^d h_\ell^2]/3\}\exp\{(7/12)\mathrm{Card}(J)\}\\ \leq\exp\{2\gamma^2/3\}\exp\{(7/12)\cdot(2/3)\cdot\gamma^2\},\end{array}$$

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

$$\mathbf{E}\{\exp\{\gamma\sum_{\ell=1}^d h_\ell\zeta_\ell\}\}\leq\exp\{19\gamma^2/18\},$$

whence, by Tschebyshev Inequality,

$$\mathrm{Prob}\{|\sum_{\ell=1}^d h_\ell\zeta_\ell|>\Upsilon\}\leq 2\min_{\gamma>0}\exp\{19\gamma^2/18-\gamma\Upsilon\}=2\exp\{-9\Upsilon^2/38\}.$$

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

$$\mathbf{E}\{s_\ell\}=\mathbf{E}\{s_{\ell-1}\cosh(\gamma h_\ell\zeta_\ell)+\sinh(\gamma\sum_{r=1}^{\ell-1}h_r\zeta_r)\sinh(\gamma h_\ell\zeta_\ell\}=\mathbf{E}\{s_{\ell-1}\}\underbrace{\mathbf{E}\{\cosh(\gamma h_\ell\zeta_\ell)}_{\Theta_\ell}\},$$

whence

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

Setting $J=\{\ell:\gamma^2h_\ell^2\leq 2\}$ and taking into account that $\cosh(t)\leq\exp\{t^2/2\}$ for all t, for $\ell\notin J$ we have

$$\Theta_\ell=\mathbf{E}\{\cosh(\gamma h_\ell\zeta_\ell)\}\leq\mathbf{E}\{\exp\{\gamma^2h_\ell^2\zeta_\ell^2/2\}\}\leq\exp\{\gamma^2h_\ell^2/2\},$$

where the concluding inequality is given by the facts that $\gamma^2h_\ell^2/2\leq 1$ and $\mathbf{E}\{\exp\{\zeta_\ell^2\}\}\leq\exp\{1\}$ in view of the Hölder Inequality. When $\ell\in J$, we, same as above, have

$$\cosh(\gamma h_\ell\zeta_\ell)\leq\exp\{|\gamma h_\ell\zeta_\ell|\}\leq\exp\{\zeta^2+\gamma^2h_\ell^2/4\},$$

whence $\Theta_\ell\leq\exp\{1+\gamma^2h_\ell^2/4\}\leq\exp\{1/2+\gamma^2h_\ell^2/2\}$. We therefore get

$$\mathbf{E}\{\cosh(\gamma\sum_{\ell=1}^d h_\ell\zeta_\ell)\}\leq\exp\{\gamma^2[\sum_{\ell=1}^d h_\ell^2]/2\}\exp\{\mathrm{Card}(J)/2\},$$

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

$$\mathbf{E}\{\cosh(\gamma\sum_{\ell=1}^d h_\ell\zeta_\ell)\}\leq\exp\{3\gamma^2/4\}.$$

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

$$\mathrm{Prob}\{|\sum_{\ell=1}^d h_\ell\zeta_\ell|>\Upsilon\}\leq 2\inf_{\gamma>0}\exp\{3\gamma^2/4-\gamma\Upsilon\}=2\exp\{-\Upsilon^2/3\},$$

as required in (36). □

Comparison with other approximations of a chance constrained LP. As it was already mentioned, the diagonal case arises when solving chance constrained Linear Programming problems which we prefer to pose in the form of (26):

$$\max_{x,\rho}\left\{\rho: \begin{array}{l} Fx-f\geq 0, c^Tx\leq \bar{c} \\ \mathrm{Prob}\{A_\zeta x - b_\zeta \geq 0\}\geq 1-\epsilon \end{array}\right\},\quad [A_\zeta,b_\zeta]=[A^0,b^0]+\sum_{\ell=1}^d \zeta_\ell[A^\ell,b^\ell]$$
$$\Updownarrow$$
$$\max_{x,\rho}\left\{\rho: \begin{array}{l} Fx-f\geq 0, c^Tx\leq \bar{c} \\ \mathrm{Prob}\{\mathcal{A}_0[x]+\sum_{\ell=1}^d \zeta_\ell\mathcal{A}_\ell[x]\succeq 0\}\geq 1-\epsilon \end{array}\right\},\quad \mathcal{A}_\ell[x]=\mathrm{Diag}\{A^\ell x-b^\ell\},\ 0\leq\ell\leq d. \tag{39}$$

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

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

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

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

This problem involves chance constrained *scalar* linear inequalities which 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 quasiconvex problem

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

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 $\epsilon_i=\epsilon/I$, $i=1,...,I$. Note that *with this choice of* ϵ_i, *problem* (42) *is equivalent to* (40), up to rescaling $\rho\mapsto\rho/\phi(\epsilon/I)$. This, however, does not mean that the approximations are identical; while 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, while for (40) ρ_* is given by the Calibration routine. Experiments show that which of these two lower bounds is less conservative, it depends on 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, $\ell=1,...,d$, are of the form

$$B_\ell=[ef_\ell^T+f_\ell e^T]+\lambda_\ell G, \tag{43}$$

where $e, f_\ell \in \mathbf{R}^n$, $\lambda_\ell \in \mathbf{R}$ and $G \in \mathbf{S}^n$. Observe that we meet this case in the Chance Constrained Conic Quadratic Optimization, see (6). Indeed, the matrices $\mathcal{A}_\ell[x]$, $1 \leq \ell \leq d$, arising in (6) are, for every x, matrices of the form (43), and therefore all we need when building and processing the safe tractable approximation, as developed in Section 3.2, of the chance constrained LMI in (6) is the validity of (34) for matrices B_ℓ of the form (43).

Theorem 4.2 *Let $n\times n$ matrices $B_1, ..., B_d$ of the form (43) along with a matrix $B_0 \in \mathbf{S}^n$ satisfy the premise in (34). Let, further, $\zeta_1, ..., \zeta_d$ be independent random variables with zero means and such that $\mathbf{E}\{\zeta_\ell^2\} \leq \sigma^2$, $\ell = 1, ..., d$ (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*

$$\begin{array}{lll} (a) & 2\sigma\sqrt{2/\chi} & \textit{[general case]} \\ (b) & \min[2\sqrt{2/\chi}, 4 + 4\sqrt{\ln(2/\chi)}] & \textit{[case of A.1]} \\ (c) & \min[2\sigma\sqrt{2/\chi}, 10\sqrt{\ln(1/\chi)}] & \textit{[case of A.2]} \end{array} \tag{44}$$

one has

$$\Upsilon \geq \Upsilon(\chi)\frac{2\sqrt{2}\sigma}{\sqrt{\chi}} \Rightarrow \text{Prob}\{-\Upsilon B_0 \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell \preceq \Upsilon B_0\} \geq 1 - \chi, \tag{45}$$

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

Proof. First of all, when ζ_ℓ, $\ell = 1, ..., d$, satisfy A.3, we indeed have $\mathbf{E}\{\zeta_\ell^2\} \leq \exp\{1\} - 1$ due to $t^2 \leq \exp\{t^2\} - 1$ for all t. Further, by continuity argument, it suffices to consider the case where $\text{Arrow}(B_0, B_1, ..., B_d) \succeq 0$ and $B_0 \succ 0$. In this case, setting $A_\ell = B_0^{-1/2} B_\ell B_0^{-1/2}$, the relation $\text{Arrow}(B_0, ..., B_d) \succeq 0$ is equivalent to $\sum_{\ell=1}^d A_\ell^2 \preceq I$, and the target relation (45) is equivalent to

$$\Upsilon \geq \Upsilon(\chi) \Rightarrow \text{Prob}\{-\Upsilon I_n \preceq \sum_{\ell=1}^d \zeta_\ell A_\ell \preceq \Upsilon I_n\} \geq 1 - \chi$$

with $\Upsilon(\chi)$ announced in Theorem 4.2.

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

$$\text{Prob}\{\|\sum_{\ell=1}^d \zeta_\ell B_\ell\| \leq \Upsilon(\chi)\} \geq 1 - \chi, \tag{46}$$

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

Observe that A_ℓ, $1 \leq \ell \leq d$, also are of the form (43):

$$A_\ell = [gh_\ell^T + h_\ell g^T] + \lambda_\ell H \qquad\qquad [g = B_0^{-1/2}e, h_\ell = B_0^{-1/2}f_\ell, H = B_0^{-1/2}GB_0^{-1/2}]$$

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

$$A_\ell = \left[\begin{array}{c|c} q_\ell & r_\ell^T \\ \hline r_\ell & \lambda_\ell Q \end{array}\right]. \tag{47}$$

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

$$A_\ell^2 = \left[\begin{array}{c|c} q_\ell^2 + r_\ell^T r_\ell & q_\ell r_\ell^T + \lambda_\ell r_\ell^T Q \\ \hline q_\ell r_\ell + \lambda_\ell Q r_\ell & r_\ell r_\ell^T + \lambda_\ell^2 Q^2 \end{array}\right].$$

We conclude that $\sum_{\ell=1}^d A_\ell^2 \preceq I_n$ implies that $\sum_{\ell=1}^d (q_\ell^2 + r_\ell^T r_\ell) \le 1$ and $[\sum_{\ell=1}^d \lambda_\ell^2] Q^2 \preceq I_{n-1}$; since $\|Q^2\| = 1$, we arrive at the relations

$$(a)\quad \sum_{\ell=1}^d \lambda_\ell^2 \le 1 \qquad\qquad (b)\quad \sum_{\ell=1}^d (q_\ell^2 + r_\ell^T r_\ell) \le 1 \tag{48}$$

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

$$\begin{array}{rl} & S \equiv \sum_{\ell=1}^d \zeta_\ell A_\ell = [g(\underbrace{\sum_{\ell=1}^d \zeta_\ell p_\ell}_{\xi})^T + \xi^T g] + \mathrm{Diag}\{\underbrace{\sum_{\ell=1}^d \zeta_\ell q_\ell}_{\theta}, (\underbrace{\sum_{\ell=1}^d \zeta_\ell \lambda_\ell}_{\eta}) Q\} \\ \Rightarrow & \|S\| \le \|g\xi^T + \xi g^T\| + \max[|\theta|, |\eta| \|Q\|] = \|\xi\|_2 + \max[|\theta|, |\eta|]. \end{array}$$

Setting

$$\alpha = \sum_{\ell=1}^d r_\ell^T r_\ell, \ \ \beta = \sum_{\ell=1}^d q_\ell^2,$$

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

$$\begin{array}{ll} \mathbf{E}\{\xi^T \xi\} = \sum_{\ell,\ell'} \mathbf{E}\{\zeta_\ell \zeta_{\ell'}\} p_\ell^T p_{\ell'} = \sum_{\ell=1}^d \mathbf{E}\{\zeta_\ell^2\} r_\ell^T r_\ell & [\zeta_\ell \text{ are independent, } \mathbf{E}\{\zeta_\ell\} = 0] \\ \le \sigma^2 \sum_{\ell=1}^d r_\ell^T r_\ell \le \sigma^2 \alpha & \\ \Rightarrow \quad \mathrm{Prob}\{\|\xi\|_2 > t\} \le \frac{\sigma^2 \alpha}{t^2} \ \forall t > 0 & [\text{Tschebyshev Inequality}] \\ \mathbf{E}\{\eta^2\} = \sum_{\ell=1}^d \mathbf{E}\{\zeta_\ell^2\} \lambda_\ell^2 \le \sigma^2 \sum_{\ell=1}^d \lambda_\ell^2 \le \sigma^2 & [\text{see } (48.a)] \\ \Rightarrow \quad \mathrm{Prob}\{|\eta| > t\} \le \frac{\sigma^2}{t^2} \ \forall t > 0 & [\text{Tschebyshev Inequality}] \\ \mathbf{E}\{\theta^2\} = \sum_{\ell=1}^d \mathbf{E}\{\zeta_\ell^2\} q_\ell^2 \le \sigma^2 \beta & \\ \Rightarrow \quad \mathrm{Prob}\{|\theta| > t\} \le \frac{\sigma^2 \beta}{t^2} \ \forall t > 0 & [\text{Tschebyshev Inequality}] \end{array}$$

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

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

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

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

so that

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

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 $\mathrm{Prob}\{\|S\| \ge \widetilde{\Upsilon}\} < 1/2$ by (49). Invoking Theorem 3.1, we conclude that for all $\gamma \ge 1$ we have $\mathrm{Prob}\{\|S\| \ge \gamma\widetilde{\Upsilon}\} \le 2\exp\{-\widetilde{\Upsilon}^2(\gamma-1)^2/16\}$. Given $\chi \in (0, 1/2)$ and setting $\gamma = 1 + 4\widetilde{\Upsilon}^{-1}\sqrt{\ln(2/\chi)}$, we get $\mathrm{Prob}\{\|S\| \ge \widetilde{\Upsilon} + 4\sqrt{\ln(2/\chi)}\} \le \chi$; since this relation holds true for every $\widetilde{\Upsilon} > 4$, we see that in the case of A.1, in addition to (49), it holds $\mathrm{Prob}\{\|S\| \geq 4 + 4\sqrt{\ln(2/\chi)}\} \leq \chi$, $0 < \chi < 1/2$, which proves the "A.1-version" of 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 $\mathrm{Prob}\{\|S\| \geq 2\sqrt{2/s}\} \leq s$. Applying Theorem 3.1 with s in the role of χ, we conclude that whenever $s \in (0, 1/2)$ and $\gamma \geq 1$, we have $\mathrm{Prob}\{\|S\| \geq 2\gamma\sqrt{2/s}\} \leq \exp\{-\phi^2(s)\gamma^2/2\}$. It follows that setting

$$\Upsilon_*(\chi) = \inf_{s,\gamma} \left\{ 2\gamma\sqrt{2/s} : s \in (0, 1/2), \gamma \geq 1, \exp\{-\phi^2(s)\gamma^2/2\} \leq \chi \right\},$$

we ensure the relation $\mathrm{Prob}\{\|S\| \geq \Upsilon_*(\chi)\} \leq \chi$ for all $\chi \in (0, 1/2)$. It is immediately seen that $\Upsilon(\chi)$ given in the case A.2 by (44) 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 $\theta \geq 0$, 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* $\mathrm{Arrow}(B_0, B_1, \ldots, B_d) \succeq 0$, *we have*

$$\begin{array}{rr} \forall (0 < \chi < 1/2) : \mathrm{Prob}\{-\Upsilon_1(\chi)B_0 \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell \preceq \Upsilon_1(\chi)B_0\} \geq 1 - \chi \\ \textit{whenever } \zeta \textit{ satisfies A.1;} \\ \forall (0 < \chi < 1/2) : \mathrm{Prob}\{-\Upsilon_2(\chi)B_0 \preceq \sum_{\ell=1}^d \zeta_\ell B_\ell \preceq \Upsilon_2(\chi)B_0\} \geq 1 - \chi \\ \textit{whenever } \zeta \textit{ satisfies A.2.} \end{array} \qquad (50)$$

E.g.,

- when $\mathcal{B}$ is the set of collections of all $n \times n$ symmetric matrices, (P) is satisfied with $\Upsilon_{1,2}(\chi) = O(1)n^{1/6}\chi^{-1/2}$ (this is immediately given by the already cited results of [12]);
- when $\mathcal{B}$ is comprised of all collections of diagonal $n \times n$ matrices, (P) is satisfied with $\Upsilon_{1,2}(\chi) = O(1)\sqrt{\ln(n/\chi)}$, see Theorem 4.1;
- when $\mathcal{B}$ is comprised of all collections $B_0, B_1, \ldots, B_d$ of symmetric $n \times n$ matrices with $B_1, \ldots, B_d$ of the form $e^T f_\ell + f_\ell^T e + \lambda_\ell G$, (P) is satisfied with $\Upsilon_{1,2}(\chi) = O(1)\sqrt{\ln(1/\chi)}$, see Theorem 4.2.

In the case of (P), we can build safe tractable approximations of problems of interest (7), (26) avoiding any necessity in simulations. Specifically, a safe tractable approximation of (7) can be chosen as the problem

$$\begin{array}{c} \min\limits_x \left\{ c^T x : \begin{array}{l} F(x) \leq 0 \\ \mathrm{Arrow}(\Theta\mathcal{A}_0[x], \mathcal{A}_1[x], \ldots, \mathcal{A}_d[x]) \succeq 0 \end{array} \right\}, \\ \Theta = \left\{ \begin{array}{ll} \sup\limits_{0<\chi<1/2} \left[\Upsilon_1(\chi) + 4\sqrt{\ln(\epsilon^{-1}(1-\chi)^{-1})} \right]^{-1}, & \text{case of A.1} \\ \sup\limits_{0<\chi<1/2} \Upsilon_2^{-1}(\chi) \min\left[\frac{\phi(\chi)}{\sqrt{2\ln(\epsilon^{-1})}}, 1 \right], & \text{case of A.2} \end{array} \right. \end{array} \qquad (51)$$

Indeed, assume that x_* is a feasible solution of the approximation; we should prove that

$$\mathrm{Prob}\left\{\mathcal{A}_0[x_*] + \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x_*] \succeq 0\right\} \geq 1 - \epsilon. \tag{52}$$

Assuming that we are in the case of A.1, for every $\nu > 0$ we can point out $\chi \in (0, 1/2)$ such that $(1+\nu)\Upsilon_1(\chi)\Theta \geq 1$. Setting $A_0 = (1+\nu)\mathcal{A}_0[x_*]$, $A_\ell = \mathcal{A}_\ell[x_*]$ and invoking Corollary 3.1, we conclude that $\mathrm{Prob}\{(1+\nu)\mathcal{A}_0[x_*] + \sum_{\ell=1}^{d} \mathcal{A}_\ell[x_*] \succeq 0\} \geq 1 - \epsilon$; since $\nu > 0$ is arbitrary, (52) holds true. The reasoning in the case of A.2 is completely similar.

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, 4.2, respectively), the corresponding "simulation-free" safe tractable approximations of chance constrained problems of interest are not too conservative. E.g., in the case of A.2 there exists an absolute constant $C > 0$ such that a vector x which does *not* satisfy the constraint $\mathrm{Arrow}(C^{-1}\Theta\mathcal{A}_0[x], \mathcal{A}_1[x], ..., \mathcal{A}_d[x]) \succeq 0$ not necessarily satisfies the chance constraint of interest (8), provided that $\epsilon n \leq 1$. However, we shall see in Section 6 that in reality simulation-based approximations can be significantly less conservative than the simulation-free ones.

5 Majorization

One way to bound from above the probability

$$q(x) := \mathrm{Prob}\left\{\mathcal{A}_0[x] + \sum_{\ell=1}^{d} \zeta_\ell \mathcal{A}_\ell[x] \not\succeq 0\right\}$$

for a randomly perturbed LMI to be violated is to replace the random perturbations ζ with an easier-to-handle perturbations $\widehat{\zeta}$ – such that we know how to bound from above the quantity

$$\widehat{q}(x) := \mathrm{Prob}\left\{\mathcal{A}_0[x] + \sum_{\ell=1}^{d} \widehat{\zeta}_\ell \mathcal{A}_\ell[x] \not\succeq 0\right\}.$$

If, in addition, $\widehat{\zeta}$ is "more diffuse" than ζ, meaning that $\widehat{q}(x) \geq q(x)$ 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 $\widehat{\zeta}_\ell$ which we do know how to handle.

For the outlined idea to be meaningful, we should specify properly 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 $\mathcal{SU}$ of all symmetric w.r.t. 0 and unimodal probability distributions on the axis, the latter notion being defined as follows:

Definition 5.1 *A probability distribution P on the axis is called unimodal and symmetric, if P possess a density $p(\cdot)$ which is an even function non-increasing on $[0,\infty)$.*[1)]

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

With 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, resp., Q, and probability densities $p(\cdot)$, resp., $q(\cdot)$, then every one of the relations $\eta \succeq_{\rm m} \xi$, $q(\cdot) \succeq_{\rm m} p(\cdot)$ means that $P, Q \in \mathcal{SU}$ and $P \succeq_{\rm m} Q$. Relation $\preceq_{\rm m}$ is the natural "counterpart" of the relation $\succeq_{\rm m}$.

The important for us facts on the monotone dominance are as follows:

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

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

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

(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 $p_i(\cdot) \preceq_{\rm m} q_i(\cdot)$ for every i, then $p(\cdot) \in \mathcal{SU}$, $q(\cdot) \in \mathcal{SU}$ and $p(\cdot) \preceq_{\rm m} q(\cdot)$.*

(v) *If $\{\xi_\ell \in \mathcal{SU}\}_{\ell=1}^n$, $\{\eta_\ell \in \mathcal{SU}\}_{\ell=1}^n$ are collections of independent random variables such that $\xi_\ell \succeq_{\rm m} \eta_\ell$, $\ell = 1, ..., n$, and λ_ℓ, $\ell = 1, ..., n$, are deterministic reals, then $\sum_{\ell=1}^n \lambda_\ell \xi_\ell \succeq_{\rm m} \sum_{\ell=1}^m \lambda_\ell \eta_\ell$.*

(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 $\xi \preceq_{\rm m} \zeta \preceq_{\rm m} \eta$.*

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

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

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 Appendix.

5.2 Convex dominance and Majorization Theorem

To conclude this Section, we present another "Gaussian Majorization" result. Its advantage is that it does not require the random variables ζ_ℓ to be symmetrically or unimodally distributed; what we need, 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*

[1)]In literature, an 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 Appendix.

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 $\xi, \eta \in \mathcal{R}_n$. We say that η convexly dominates ξ (notation: $\xi \preceq_c \eta$, or $P_\xi \preceq_c P_\eta$, or $\eta \succeq_c \xi$, or $P_\eta \succeq_c P_\xi$) if*

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

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

A summary of relevant in our context facts on convex dominance is as follows.

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

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

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

(iv) *If ξ is symmetrically distributed w.r.t. 0 and $t \geq 1$ is deterministic, then $t\xi \succeq_c \xi$.*

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

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

(vii) *Assume that $\xi \in \mathcal{R}_n$ is supported in the unit cube $\{u : \|u\|_\infty \leq 1\}$ 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. Let also $\eta \sim \mathcal{N}(0, (\pi/2)I_n)$. Then $\xi \preceq_c \eta$.*

(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 $\xi \preceq_c \eta$.*

(ix) [Majorization Theorem [11]] *Let $\eta \sim \mathcal{N}(0, I_d)$, and let $\zeta \in \mathcal{R}_d$ be such that $\zeta \preceq_c \eta$. 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*

$$\mathrm{Prob}\{\zeta \notin \gamma Q\} \leq \inf_{1 \leq \beta < \gamma} \frac{1}{\gamma - \beta} \int_\beta^\infty \exp\{-r^2 \phi^2(\chi)/2\} dr, \tag{53}$$

where $\phi(\cdot)$ is the inverse error function (17).

All of the facts above, except for Majorization Theorem, are well-known; all proofs can be found in [11].

5.3 Calibration based on Gaussian majorization

We can utilize the outlined facts in the Calibration procedure, specifically, as follows.

Utilizing Comparison Theorem. Assume that the perturbations ζ_ℓ are independent and possess unimodal and symmetric distributions P_ℓ such that $P_\ell \preceq_m \mathcal{N}(0, \sigma^2)$ 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 Comparison Theorem, we conclude that for every deterministic symmetric matrices $A_0, A_1, ..., A_d$ and every $r > 0$ we have

$$\mathrm{Prob}\{-A_0 \preceq \frac{r}{\sigma} \sum_{\ell=1}^d \zeta_\ell A_\ell \preceq A_0\} \leq \mathrm{Prob}\{-A_0 \preceq r \sum_{\ell=1}^d \eta_\ell A_\ell \preceq A_0\}. \tag{54}$$

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

$$\rho^* = \max\left\{\rho : \mathrm{Prob}\{-A_0 \preceq \rho \sum_{\ell=1}^d \zeta_\ell A_\ell \preceq A_0\} \geq 1-\epsilon\right\}. \tag{55}$$

By (54), 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

$$r_* = \max\left\{r : \mathrm{Prob}\{-A_0 \preceq r \sum_{\ell=1}^d \eta_\ell A_\ell \preceq A_0\} \geq 1-\epsilon\right\}$$

and to set $\rho_* = r_*/\sigma$. This approach allows to extend the above constructions beyond the scope of Assumption A; moreover, we shall see in Section 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 $\zeta_1, ..., \zeta_d$ are independent with zero means, and that we can point out $\sigma > 0$ such that $P_{\zeta_\ell} \preceq_{\mathrm{c}} \mathcal{N}(0, \sigma^2)$. 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 $A_0, ..., A_d$, ϵ, δ to the Calibration procedure and applying Majorization Theorem to the closed convex set

$$Q = Q_s = \{u \in \mathbf{R}^d : -sA_0 \preceq \sum_{\ell=1}^d u_\ell A_\ell \preceq sA_0\},$$

we conclude that

$$\begin{array}{l} \forall (s > 0 : \chi(s) \equiv \mathrm{Prob}\{\eta \notin Q_s\} \equiv 1 - \mathrm{Prob}\{-A_0 \preceq s^{-1} \sum_{\ell=1}^d \eta_\ell A_\ell \preceq A_0\} < 1/2) : \\ \mathrm{Prob}\{-\gamma s\sigma A_0 \preceq \sum_{\ell=1}^d \zeta_\ell A_\ell \preceq \gamma\sigma s A_0\} = 1 - \mathrm{Prob}\{\sigma^{-1}\zeta \notin \gamma Q_s\} \geq 1 - \Psi(\gamma, \chi(s)), \\ \qquad \Psi(\gamma, \chi) = \inf\limits_{1 \leq \beta < \gamma} \frac{1}{\gamma-\beta} \int_\beta^\infty \exp\{-r^2\phi^2(\chi)/2\} dr. \end{array} \tag{56}$$

In order to bound from below ρ^* (see (55)), let us 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 $\{r^k > 0, \widehat{\chi}^k < 1/2\}_{k=1}^{\bar{K}}$ such that "up to probability of bad sampling $\leq \delta$" we have

$$\chi^k := \mathrm{Prob}\{\neg(-(r^k)^{-1}A_0 \preceq \sum_{\ell=1}^d \eta_\ell A_\ell \preceq (r^k)^{-1}A_0)\} = \mathrm{Prob}\{\eta \notin Q_{s_k}\},\ s_k = 1/r^k;$$

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

$$\rho_* = \max_{1 \leq k \leq \bar{K}} \frac{r^k}{\sigma\gamma_k},\ \gamma_k = \min\left\{\gamma \geq 1 : \Psi(\gamma, \widehat{\chi}^k) \leq \epsilon\right\},\ k = 1, ..., \bar{K}$$

and invoking (56), we see that ρ_* is a lower bound on ρ^*, provided that $\chi^k \leq \widehat{\chi}^k$, $1 \leq k \leq \bar{K}$, 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 illustrations to follow, 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 "standalone" Calibration procedure. Recall that this procedure is aimed at building $(1-\delta)$-reliable lower bound ρ_* on the quantity

$$\rho^* = \max\left\{\rho : p(\rho) := \mathrm{Prob}\{-A_0 \preceq \rho \sum_{\ell=1}^{d} \zeta_\ell A_\ell \preceq A_0\} \geq 1-\epsilon\right\}, \tag{57}$$

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

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

1. What is the best strategy to use the procedure? Should we use the plain Calibration procedure (PCP) whenever possible, or it is better to use the Gaussian majorization version (GCP) of this procedure?
2. As we have seen in Section 4.3, there are situations where not too conservative, at least from the theoretical viewpoint, 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?
3. How conservative, from practical perspective, is the Calibration procedure?

The answers to these questions suggested by our rather intensive numerical experimentation are as follows:

- Calibration procedure significantly outperforms simulation-free lower bounding;
- For small ϵ, GCP significantly outperforms PCP, while for "large" ϵ the situation is opposite, so that it makes sense to use both these versions of Calibration procedure, choosing the best (the largest) of the resulting lower bounds on ρ^*;
- Conservatism of the Calibration procedure is not that disastrous: the ratio r^*/r_* is usually well within one order of magnitude.

These observations are illustrated in Table 1 reporting on an experiment as follows. We randomly generate $d = 32$ matrices $A_1, ..., A_d$ of size 32×32 and of prescribed structure, specifically, full ("general case"), diagonal ("diagonal case") and of the form $\left[\begin{array}{c|c} & f^T \\ \hline f & \end{array}\right]$, f being a vector ("arrow case"), and scale the generated matrices to ensure that $\mathrm{Arrow}(\theta I_{32}, A_1, ..., A_d) \succeq 0$ if and only if $\theta \geq 1$; the input to the Calibration procedure is the collection $A_0 = I_{32}$, $A_1, ..., A_{32}$, $\vartheta_* = 1$.

Data in Table 1 correspond to 100,000-element training sample. Note that while the performance of Calibration procedure somehow improves when the sample size grows (see Table 2), this phenomenon is rather moderate.

		case								
		general			diagonal			arrow		
ϵ	P	ρ_*	$\widehat{\epsilon}$	ρ^*/ρ_*	ρ_*	$\widehat{\epsilon}$	ρ^*/ρ_*	ρ_*	$\widehat{\epsilon}$	ρ^*/ρ_*
1.e-2	G	1.67e-2	5.6e-3	1.1	4.04e-2 [2.77e-2]	3.3e-3	1.1	2.20e-2 [1.52e-3]	8.2e-3	1.1
	U	2.05e-2 (3.17e-2)	0.0 7.45e-3	1.6 1.05	5.06e-2 (7.12e-2) [9.75e-3]	1.0e-5 2.5e-3 0.0	1.7 1.2 8.8	2.75e-2 (3.43e-2) [2.78e-3]	0.0 1.2e-4 0.0	1.5 1.2 14.8
	B	1.09e-2 (1.94e-2)	0.0 3.3e-3	1.9 1.1	2.55e-2 (3.92e-2) [9.75e-3]	0.0 8.5e-4 0.0	2.0 1.3 5.2	1.41e-2 (2.12e-2) [2.43e-3]	0.0 1.0e-5 0.0	1.8 1.2 10.5
1.e-4	G	1.12e-2	0.0	1.3	2.54e-2 [1.99e-2]	0.0	1.4	1.42e-2 [1.08e-3]	0.0	1.3
	U	1.43e-2 (2.83e-3)	0.0 0.0	2.1 11.2	3.19e-2 (9.36e-3) [7.70e-3]	0.0 0.0 0.0	1.8 6.2 7.5	1.78e-2 (2.48e-3) [2.15e-3]	0.0 0.0 0.0	2.0 14.6 16.8
	B	8.63e-3 (2.68e-3)	0.0 0.0	2.1 6.6	1.93e-2 (8.84e-3) [7.70e-3]	0.0 0.0 0.0	2.0 4.3 4.9	1.08e-2 (2.39e-3) [1.93e-3]	0.0 0.0 0.0	2.0 8.7 11.2
1.e-6	G	9.17e-3	0.0	1.6	2.08e-2 [8.11e-3]	0.0	1.6	1.16e-2 [8.78e-4]	0.0	1.6
	U	1.16e-2 (2.35e-3)	0.0 0.0	2.5 12.3	2.60e-2 (7.79e-3) [6.62e-3]	0.0 0.0 0.0	2.3 7.7 9.1	1.45e-2 (2.48e-3) (1.83e-3)	0.0 0.0 0.0	2.2 13.1 17.7
	B	7.15e-3 (2.24e-3)	0.0 0.0	2.6 8.3	1.60e-2 (7.44e-3) [6.62e-3]	0.0 0.0 0.0	2.3 5.0 5.6	8.94e-3 (1.99e-3) [1.67e-3]	0.0 0.0 0.0	2.4 10.8 12.8

Table 1: Experiments with standalone Calibration procedure, $\delta = 1.e\text{-}6$. Column "P": identical to each other distributions of $\zeta_1, ..., \zeta_d$; 'G', 'U', 'B' stand for $\mathcal{N}(0,1)$, Uniform$[-1,1]$ and Uniform$\{-1;1\}$, respectively. Column "ρ_*": lower bounds on ρ^* as given by Calibration procedure with 100,000-element training sample. Data without parentheses: GCP; data in parentheses: PCP; data in brackets: simulation-free bounds, see Section 4.3. Column "$\widehat{\epsilon}$": empirical, over 100,000-element sample, values of $1 - p(\rho_*)$. Column "ρ^*/ρ_*": upper bounds on ρ^*/ρ_* given by empirical, over 100,000-element sample, estimates of ρ^*.

		ρ_*		
ϵ	P	$N=1,000$	$N=10,000$	$N=100,000$
1.e-2	G	1.05e-2	1.55e-2	1.67e-2
	U	1.32e-2 (3.85e-3)	1.93e-2 (3.12e-2)	2.05e-2 (3.17e-2)
	B	7.65e-3 (3.59e-3)	1.02e-2 (1.89e-2)	1.09e-2 (1.94e-2)
1.e-4	G	7.45e-3	9.92e-3	1.12e-2
	U	9.35e-3 (2.83e-3)	1.24e-2 (2.82e-3)	1.43e-2 (2.83e-3)
	B	5.64e-3 (2.68e-3)	7.53e-3 (2.67e-3)	8.63e-3 (2.68e-3)
1.e-6	G	6.09e-3	8.10e-3	9.17e-3
	U	7.63e-3 (2.35e-3)	1.01e-2 (2.35e-3)	1.16e-2 (2.35e-3)
	B	4.69e-3 (2.25e-3)	6.24e-3 (2.35e-3)	7.15e-2 (2.24e-3)

Table 2: Performance of Calibration procedure vs. size N of training sample, $\delta = 1$.e-6. Column "P": see Table 1. Data without parentheses: GCP; data in parentheses: PCP.

6.2 Illustration: chance constrained Truss Topology Design

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, and 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., [2, Chapter 15]). Mathematically, the TTD problem is the semidefinite program

$$\min_{\tau,t}\left\{\tau : \left[\begin{array}{c|c} 2\tau & f^T \\ \hline f & \sum_{i=1}^n t_i b_i b_i^T \end{array}\right] \succeq 0, t\geq 0, \sum_i t_i = 1\right\}, \tag{58}$$

where τ is (an upper bound on) the *compliance* – a natural measure of truss's 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_i's and f is the total # of degrees of freedom of the nodes.

The "nominal design" shown on Fig. 1.a) is the optimal solution to a toy TTD problem with 9×9 planar nodal grid and the load of interest f comprised of a single force (see Fig. 1.c)); this design uses just 12 of the original 81 nodes and 24 of the original 2,039 bars. In reality, the truss, of course, will be subject not only to the load of interest, but also to occasional relatively small loads somehow distributed along the nodes actually used by the construction, and it should withstand these loads. This is *by far* not the case with the truss on Fig. 1.a) – it can be crushed by a pretty small occasional load. Indeed, a typical randomly oriented load $\tilde{f}$ acting at the 12 nodes of nominal design and very small as compared to the load of interest ($\|\tilde{f}\|_2 \leq 10^{-7}\|f\|_2$) results in compliances which are about 10 times larger than the compliance w.r.t. the load of interest – a phenomenon illustrated on Fig. 1.b). A natural way to "cure" the nominal design is to resolve the TTD problem, explicitly imposing the requirement that the would-be truss should carry well occasional random loads. Specifically, we

- replace our original 81-point nodal set with the 12-point set of nodes actually used by the nominal design (Fig. 1.c)). Note that among these nodes, the two most left 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 Fig. 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:

$$\max_{\rho,t} \left\{ \rho : \begin{array}{l} \overbrace{\left[\begin{array}{c|c} 2\widehat{\tau} & f^T \\ \hline f & \sum_{i=1}^{54} t_i b_i b_i^T \end{array}\right]}^{A(t)} \succeq 0, t \geq 0, \sum_i t_i = 1 \\ \mathrm{Prob}_{\zeta \sim \mathcal{N}(0, I_{20})} \{ \underbrace{\left[\begin{array}{c|c} 2\widehat{\tau} & \rho\zeta^T \\ \hline \rho\zeta & \sum_{i=1}^{54} t_i b_i b_i^T \end{array}\right]}_{\mathcal{A}_0[t] + \rho \sum_{\ell=1}^{M} \zeta_\ell \mathcal{A}_\ell[t]} \succeq 0 \} \geq 1 - \epsilon \end{array} \right\}, \tag{59}$$

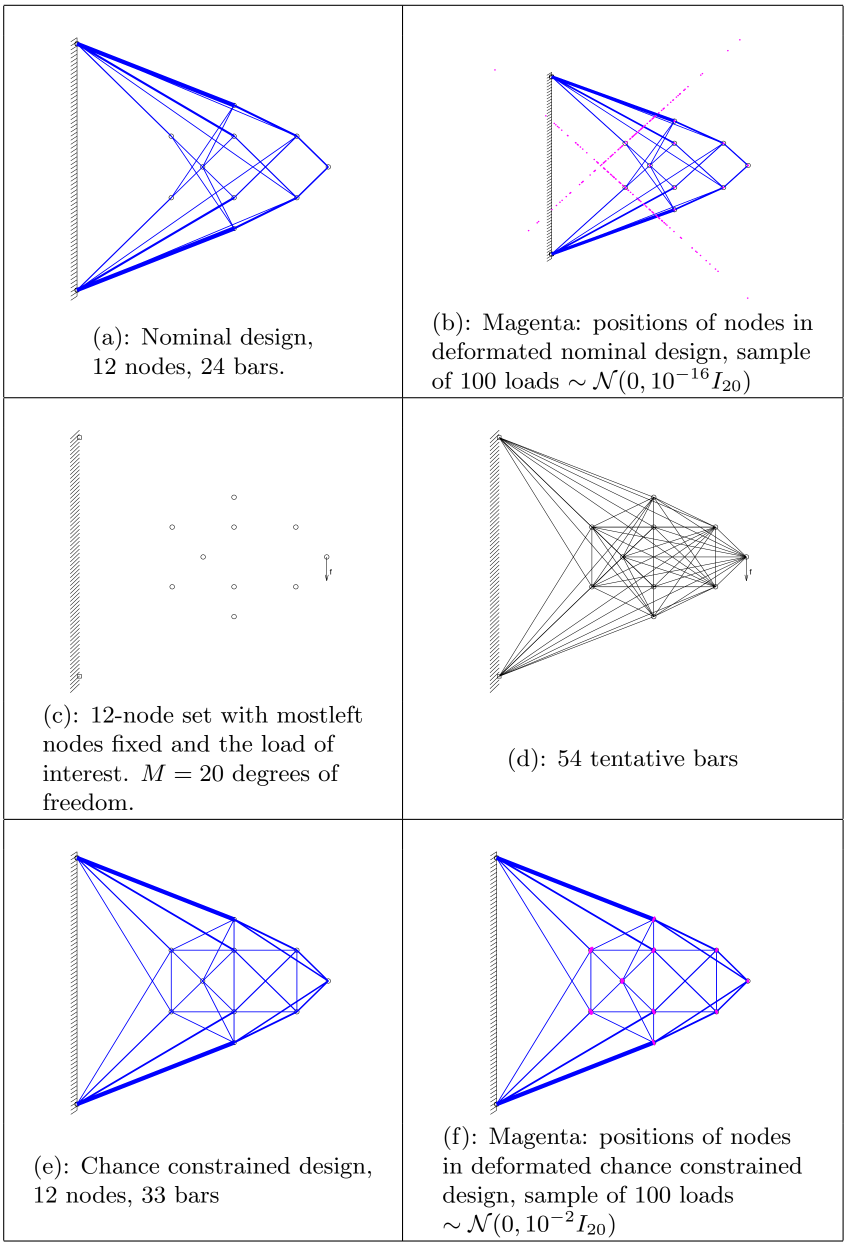
where $\widehat{\tau}$ is slightly greater than the optimal value τ_* in the original TTD problem (in our experiment, we set $\widehat{\tau} = 1.025\tau_*$). In other words, we are now looking for truss for which the compliance w.r.t. the load of interest is nearly optimal – is at most $\widehat{\tau}$, and which is capable to withstand 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 [1].

Implementing the outlined strategy, we built and solved the safe tractable approximation

$$\min_{\vartheta,t} \left\{ \vartheta : \begin{array}{l} A(t) \succeq 0, t \geq 0, \sum_i t_i = 1 \\ \mathrm{Arrow}(\vartheta \mathcal{A}_0[t], \mathcal{A}_1[t], ..., \mathcal{A}_M[t]) \succeq 0 \end{array} \right\} \tag{60}$$

(cf. (27)) of the chance constrained TTD problem (59). 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 (59). In our experiment, we worked with pretty high reliability requirements: $\epsilon = \delta =$ 1.e-10. The results are presented in Table 3 and are illustrated on Fig. 1. Note that we are in the arrow case, so that we can build a simulation-free lower bound on $\rho^*(t_*)$, see Section 4.3. With our data, this load is 4.01e-3 – nearly 10 times worse than the best simulation-based extremely reliable ($\delta =$ 1.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 Introduction); the latter, to the best of our knowledge, is the only existing alternative for processing chance constrained LMIs. The



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

(b): Magenta: positions of nodes in deformated nominal design, sample of 100 loads $\sim \mathcal{N}(0, 10^{-16} I_{20})$

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

(d): 54 tentative bars

(e): Chance constrained design, 12 nodes, 33 bars

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

Figure 1: Nominal and chance constrained designs.

	$N = 1,000$	$N = 10,000$	$N = 100,000$
ρ_*	2.77e-2	3.37e-2	3.77e-2
$\rho^*(t_*)/\rho_* \leq$	2.5	2.0	1.8

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

Scenario approximation of the chance constrained problem of interest (59) is the semidefinite program

$$\max_{\rho,t}\left\{\rho : \begin{array}{l}\left[\begin{array}{c|c} 2\widehat{\tau} & f^T \\ \hline f & \sum_{i=1}^{54} t_i b_i b_i^T \end{array}\right] \succeq 0, t \geq 0, \sum_i t_i = 1 \\ \left[\begin{array}{c|c} 2\widehat{\tau} & \rho[\zeta^j]^T \\ \hline \rho\zeta^j & \sum_{i=1}^{54} t_i b_i b_i^T \end{array}\right] \succeq 0,\ 1 \leq j \leq J \end{array}\right\}, \tag{61}$$

where $\zeta^1, ..., \zeta^J$ 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 $\epsilon = \delta = 1.\text{e}{-}10$ requires a completely unrealistic sample size, this is why we ran the Scenario approximation with $\epsilon = 0.01$, $\delta = 0.001$. While 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 (61) turned out to be $\rho_{\rm SA} = 0.0797$. For comparison, our approximation with $\epsilon = 0.01$ and $\delta = 0.001$ results in $\rho_* = 0.105 \approx 1.31\rho_{\rm SA}$; keeping $\epsilon = 0.01$ and reducing δ to 1.e-6, we still get $\rho_* = 0.103 \approx 1.29\rho_{\rm SA}$. Note that the design given by (60) also seems to be better than the one given by (61): at uncertainty level $\rho = 0.105$, the empirical, over 100,000-element sample of random occasional loads, probabilities for the two designs in question to yield a compliance worse than the desired upper bound $\widehat{\tau}$ were 0.0077 and 0.0097, respectively. Thus, in the experiment we are reporting our approximation scheme is a clear winner.

7 Appendix: proofs of equivalence in Definition 5.1 and of Proposition 5.1

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

$$\left\{\forall f \in \mathcal{M} : \int f(s)p(s)ds \leq \int f(s)q(s)ds\right\} \Leftrightarrow \left\{\forall t \geq 0 : \int_t^\infty p(s)ds \leq \int_t^\infty q(s)ds\right\}. \tag{62}$$

By standard continuity arguments, the left condition in (62) 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

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

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 $\mathbf{R}^+$; since $P(\cdot)$, $Q(\cdot)$ are continuous, the latter is the case if and only if $Q(s) \geq P(s)$ for all $s \geq 0$. $\square$

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

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

1^0 Let us verify that $p*\widehat{p} \in \mathcal{SU}$. We should prove that the density $(p*\widehat{p})(s) = \int p(s-r)\widehat{p}(r)dr$ 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, $\widehat{p}$, in addition to being even and nonincreasing on $\mathbf{R}_+$, are smooth. In this case we have

$$(p * \widehat{p})'(s) = \int p'(s - r)\widehat{p}(r)dr = \int p(s - r)\widehat{p}'(r)dr = \int_{-\infty}^{0} (p(s - t) - p(s + t))\widehat{p}'(t)dt. \quad (63)$$

Let $s \geq 0$. Then for $t \leq 0$ we have $|s|+|t| = |s-t| \geq |s+t|$; and since 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 $s \geq 0, t \leq 0$. Since, in addition, $\widehat{p}'(t) \geq 0$ when $t \leq 0$, the concluding quantity in (63) is nonpositive, meaning that the density $p * \widehat{p}$ 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 $\mathbf{R}$ which 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, 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 $s \geq 0$, $t \leq 0$ and taking into account that f is even and is nondecreasing on $\mathbf{R}_+$, we have $f(s - t) = f(|s - t|) = f(|s| + |t|) \geq f(|s + t|) = f(s + t)$; since $p'(t) \geq 0$ when $t \leq 0$, we conclude that $\int_{-\infty}^{0}(f(s - t) - f(s + t))p'(t)dt \geq 0$ when $s \geq 0$.

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 $p, \widehat{p}, q, \widehat{q}$ belong to $\mathcal{SU}$. All we should prove is that when $p(\cdot) \preceq_{\rm m} q(\cdot)$ and $\widehat{p}(\cdot) \preceq_{\rm m} \widehat{q}(\cdot)$, then $(p * \widehat{p})(\cdot) \preceq_{\rm m} (q * \widehat{q})(\cdot)$.

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

$$\int f(s)(p * \widehat{p})(s)ds \leq \int f(s)(p * \widehat{q})(s)ds \quad (64)$$

for every even bounded function f which is nondecreasing on $\mathbf{R}_+$. By evident continuity reasons, it suffices to verify that (64) holds true for every $f \in \mathcal{M}_*$. Taking into account that p is even, we get

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

and by similar reasons

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

As we know from 2^0, $f * p \in \mathcal{M}_*$ whenever $f \in \mathcal{M}_*$, and (64) follows from the fact that $\widehat{p}(\cdot) \preceq_{\rm m} \widehat{q}(\cdot)$.

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

(vi): To prove that $\xi \preceq_{\rm m} \zeta$, observe that since $\xi \in \mathcal{SU}$ 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 $a \leq 1$. 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 $t \geq 0$ since $q(\cdot)$ is nonincreasing on $\mathbf{R}_+$. To prove that $\zeta \preceq_{\rm s} \eta$ is exactly the same as to verify that $\widetilde{P}(t) \geq \widetilde{Q}(t)$ for all $t \geq 0$. This is indeed the case when $0 \leq t \leq 1$, since $\widetilde{Q}(0) = 0$, $\widetilde{Q}'(0) = 1/2$ and $\widetilde{Q}$ is concave on $\mathbf{R}_+$, while $\widetilde{P}(t) = \frac{1}{2}t = \widetilde{Q}(0) + t\widetilde{Q}'(0)$ when $0 \leq t \leq 1$. And of course $\widetilde{P}(t) = 1/2 \geq \widetilde{Q}(t)$ when $t \geq 1$. (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

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

for every bounded piecewise continuous function f on the axis.

2^0. We have the following

Lemma 7.1 *Let $Q \subset \mathbf{R}^d$ be a nonempty convex compact set symmetric w.r.t. the origin, and let $p_1(\cdot), ..., p_d(\cdot), q(\cdot) \in \mathcal{SU}$ be such that $p_1(\cdot)$,...,$p_{d-1}(\cdot)$ are densities of uniform distributions and $p_d(\cdot) \preceq_{\rm m} q(\cdot)$. Then*

$$\int_Q p_1(x_1)p_2(x_2)...p_{d-1}(x_{d-1})p_d(x_d)dx \geq \int_Q p_1(x_1)p_2(x_2)...p_{d-1}(x_{d-1})q(x_d)dx. \tag{65}$$

Proof. Let Σ_ℓ, $1 \leq \ell < d$, 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 $\Sigma = \Sigma_1 \times ... \times \Sigma_{d-1} \times \mathbf{R}$, $\widehat{Q} = Q \cap \Sigma$, so that $\widehat{Q}$ is a convex compact set symmetric w.r.t. the origin, and let

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

The function $f(s)$ is even; denoting by Δ the projection of $\widehat{Q}$ 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

$$\int f(s)p_d(s)ds \geq \int f(s)q(s)ds \tag{66}$$

due to $p_d(\cdot) \preceq_{\rm m} q(\cdot)$. It remains to note that the left and the right hand sides in (65) are proportional, with a common positive coefficient, to the respective sides in (66). $\Box$

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

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

By continuity argument and in view of 1^0, it suffices to verify the same relation when $p_1(\cdot)$,...,$p_{d-1}(\cdot)$ are convex combinations of densities of uniform and symmetric w.r.t. the origin distributions. Since both sides in our target inequality are linear in every one of $p_1,...,p_{d-1}$, to prove the latter fact is the same as to prove it when every one of $p_1,...,p_{d-1}$ is a uniform distribution symmetric w.r.t. the origin. In the latter case, the required statement is given by Lemma 7.1. $\Box$

References

[1] Ben-Tal, A., Nemirovski, A. "Stable Truss Topology Design via Semidefinite Programming" – *SIAM Journal on Optimization* **7:4** (1997), 991-1016.

[2] Ben-Tal, A., and Nemirovski, A. *Lectures on Modern Convex Optimization: Analysis, Algorithms and Engineering Applications*, MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[3] Ben-Tal, A., and Nemirovski, A., "On tractable approximations of uncertain linear matrix inequalities affected by interval uncertainty" – *SIAM Journal on Optimization* **12** (2002), 811-833.

[4] Ben-Tal, A., Nemirovski, A., and Roos, C., "Robust solutions of uncertain quadratic and conic-quadratic problems", *SIAM Journal on Optimization* **13** (2002), 535-560.

[5] Ben-Tal, A., Nemirovski, A., Roos, C., "Extended Matrix Cube Theorems with applications to μ-Theory in Control" – *Mathematics of Operations Research* **28** (2003), 497–523.

[6] Calafiore, G., Campi, M.C., "Uncertain Convex Programs: Randomized Solutions and Confidence Levels" – *Mathematical Programming* **102** (2005), 25-46.

[7] de Farias, D.P., Van Roy, B., "On constraint sampling in the linear programming approach to approximate dynamic programming" – *Mathematics of Operations Research*, **29** (2004), 462-478.

[8] Erdogan, G., Iyengar, G., "Ambiguous chance constrained problems and robust optimization" – *Mathematical Programming* **107:1-2** (2006), 37-61.

[9] W.B. Johnson, G. Schechtman, "Remarks on Talagrand's deviation inequality for Rademacher functions", in: E. Odell, H. Rosenthal (Eds.), *Functional Analysis (Austin, TX 1987/1989)*, Lecture Notes in Mathematics 1470, Springer, 1991, 72-77.

[10] Nemirovski, A., On tractable approximations of randomly perturbed convex constraints, *Proceedings of the 42nd IEEE Conference on Decision and Control Maui, Hawaii USA, December 2003*, 2419-2422.

[11] A. Nemirovski, A. Shapiro, "Scenario Approximations of Chance Constraints", in: Calafiore, G., Dabbene, F, Eds., *Probabilistic Scenario Approximations of Chance Constraints*, Springer 2005.

[12] A. Nemirovski, "Sums of random symmetric matrices and quadratic optimization under orthogonality constraints" – *Mathematical Programming Ser. B* **109** (2007), 283-317.

[13] Nemirovski, A., Shapiro, A. "Convex Approximations of Chance Constrained Programs", *SIAM Journal on Optimization* **17:4** (2006), 969-996.