SIAM J. OPTIM.
Vol. 19, No. 3, pp. 1211–1230

THE EXACT FEASIBILITY OF RANDOMIZED SOLUTIONS OF UNCERTAIN CONVEX PROGRAMS*

M. C. CAMPI† AND S. GARATTI‡

Abstract. Many optimization problems are naturally delivered in an uncertain framework, and one would like to exercise prudence against the uncertainty elements present in the problem. In previous contributions, it has been shown that solutions to uncertain convex programs that bear a high probability to satisfy uncertain constraints can be obtained at low computational cost through constraint randomization. In this paper, we establish new feasibility results for randomized algorithms. Specifically, the *exact* feasibility for the class of the so-called *fully-supported* problems is obtained. It turns out that all fully-supported problems share the same feasibility properties, revealing a deep kinship among problems of this class. It is further proven that the feasibility of the randomized solutions for all other convex programs can be bounded based on the feasibility for the prototype class of fully-supported problems. The feasibility result of this paper outperforms previous bounds and is not improvable because it is exact for fully-supported problems.

Key words. uncertain optimization, randomized methods, convex optimization, semi-infinite programming, robust optimization, chance-constrained

AMS subject classifications. 90C25, 90C15, 90C34, 68W20

DOI. 10.1137/07069821X

1. Introduction.

Uncertain convex optimization [21, 24, 25] deals with convex optimization in which the constraints are imprecisely known:

$$\text{UP}: \min_{x\in\mathcal{X}\subseteq\mathbb{R}^d} c^T x \qquad \text{subject to: } x\in\mathcal{X}_\delta,\quad \delta\in\Delta, \tag{1}$$

where UP stands for uncertain program, $\delta\in\Delta$ is an uncertain parameter, and $\mathcal{X}$ and $\mathcal{X}_\delta$ are convex and closed sets. Oftentimes, Δ is a set of infinite cardinality. The fact that the optimization objective is linear and does not carry any dependence on δ, that is, it is certain, is without loss of generality.

UP encompasses as special cases uncertain linear programs (LP), uncertain quadratic programs (QP), uncertain second-order cone programs (SOCP), and uncertain semi-definite programs (SDP) and plays a central role in many design endeavors such as [1, 15, 17, 14, 9, 24, 11, 6].

Dealing with uncertainty can be done along two distinct approaches. The first one consists in enforcing satisfaction of *all* constraints; that is, one optimizes the cost $c^T x$ over the set $\bigcap_{\delta\in\Delta}\mathcal{X}_\delta$ (see [2, 16, 3, 4]). Alternatively, one may want to satisfy the constraints with "high probability." Along this second approach one sees the uncertainty parameter δ as a random element with a probability $\mathbb{P}$ and seeks a solution that violates at most a fraction of constraints having small $\mathbb{P}$-probability (chance-constrained solution). Depending on the optimization problem at hand, $\mathbb{P}$

*Received by the editors July 24, 2007; accepted for publication (in revised form) May 13, 2008; published electronically November 19, 2008. This paper was supported by the MIUR national project "Identification and adaptive control of industrial systems."
http://www.siam.org/journals/siopt/19-3/69821.html
†Dipartimento di Elettronica per l'Automazione, Università di Brescia, via Branze 38, 25123 Brescia, Italia (marco.campi@ing.unibs.it, http://bsing.ing.unibs.it/∼campi/).
‡Dipartimento di Elettronica ed Informazione, Politecnico di Milano, piazza L. da Vinci 32, 20133 Milano, Italia (sgaratti@elet.polimi.it, http://home.dei.polimi.it/sgaratti/).

can have different interpretations. Sometimes, it is the actual probability with which the uncertainty parameter δ takes on value in Δ. Other times, it simply describes the relative importance attributed to different instances of δ. The use of a probabilistic description of uncertainty has a long history in optimization theory and dates back to the work [10] of Charnes and Cooper in the 1950s that in effect gave birth to the chance-constrained approach. See also [21, 22, 12, 25] for more information and [5] for a more in-depth comparison between deterministic and probabilistic uncertain optimization.

It is a fact that finding a solution carrying a high probability of constraint satisfaction is in general a very difficult task to achieve [21]. To circumvent this computational issue, recently, methodologies relying on the randomization over the set of constraints have been introduced [11, 5, 20, 6, 13]. Specifically, in [5, 6], the following randomized program RP_N is introduced, where N constraints $\delta^{(1)}, \ldots, \delta^{(N)}$ randomly extracted according to $\mathbb{P}$ in an independent fashion are simultaneously enforced:

$$\mathrm{RP}_N : \min_{x \in \mathcal{X} \subseteq \mathbb{R}^d} c^T x$$
$$\text{subject to: } x \in \bigcap_{i \in \{1,\ldots,N\}} \mathcal{X}_{\delta^{(i)}}.$$

RP_N is also known as "scenario program."

The distinctive feature of RP_N is that it is a program with a *finite* number of constraints, and, as such, it can be solved at low computational cost provided that N is not too large;[1] it is indeed a fact that RP_N has opened up new resolution avenues in uncertain optimization. On the other hand, the obvious question to ask is how feasible the solution of RP_N is; that is, how large the fraction of original constraints in Δ that are possibly violated by the solution x_N^* of RP_N is. Papers [5, 6] have pioneered a feasibility theory showing that x_N^* is feasible for the vast majority of the other unseen constraints—those that have not been used when optimizing according to RP_N—and this result holds in full generality, independently of the structure of the set of constraints Δ and the probability $\mathbb{P}$. So the vast majority of constraints take care of themselves, without explicitly accounting for them.

To allow for a sharper comparison with the results presented in this paper, we feel advisable to first recall in some detail the results in [5, 6]. The following notion of violation probability from [5] is central.

DEFINITION 1 (violation probability). *The* violation probability *of a given* $x \in \mathcal{X}$ *is defined as* $V(x) = \mathbb{P}\{\delta \in \Delta : x \notin \mathcal{X}_\delta\}$.

The problem addressed in [5, 6] is to evaluate if and when the violation probability of x_N^*, namely, $V(x_N^*)$, is below a satisfying level ϵ. To state the result precisely, note that $V(x_N^*)$ is a random variable since the solution x_N^* of RP_N is, due to the fact that it depends on the random extractions $\delta^{(1)}, \ldots, \delta^{(N)}$. Thus, $V(x_N^*) \leq \epsilon$ may hold for certain extractions $\delta^{(1)}, \ldots, \delta^{(N)}$, while $V(x_N^*) > \epsilon$ may be true for others. The following quantification of the "bad" extractions where $V(x_N^*) > \epsilon$ is the key result of [6]:

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \binom{N}{d}(1-\epsilon)^{N-d}. \tag{2}$$

[1] Depending on Δ and $\mathbb{P}$, the generation of N randomly extracted scenarios $\delta^{(1)}, \ldots, \delta^{(N)}$ from Δ can in itself be a nontrivial problem, and the reader is referred to [27, 8, 7] for further discussion on this issue.

Moving a fundamental step forward with respect to [6], in this paper we establish the validity of relation

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \tag{3}$$

(note that (3) holds with "="; that is, it is not a bound) for the prototype class of *fully-supported* problems according to Definition 3 in section 2. This result sheds new light on a deep kinship among all fully-supported problems, proving that their randomized solutions share the same violation properties, and writes a final word on the violation assessment for this type of problems.

It is further proven in this paper that the right-hand side of (3) is an upper bound for all convex problems; that is,

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \tag{4}$$

holds for all convex problems. Thus, in a real optimization problem one does not need to verify whether the problem is fully-supported or not, and $\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ can always be used as an upper bound for $\mathbb{P}^N\{V(x_N^*) > \epsilon\}$. This result (4) (i) cannot be improved (being tight for the prototype class of fully-supported problems) and (ii) outperforms the previous bound from [6] at times by a huge extent (note that when $\epsilon \to 0$, the previous bound (2) tends to $\binom{N}{d}$, while the new bound (4) goes naturally to 1!).

2. Main result.

The technical result of this paper is precisely stated in this section, followed by a discussion on the significance of the result.

For a fixed integer m and fixed given constraints $\delta^{(1)}, \dots, \delta^{(m)}$, program

$$\begin{aligned} \mathrm{P}_m : \min_{x \in \mathcal{X} \subseteq \mathbb{R}^d} & \ c^T x \\ \text{subject to: } & x \in \bigcap_{i \in \{1,\dots,m\}} \mathcal{X}_{\delta^{(i)}} \end{aligned} \tag{5}$$

is called a *finite instance with m constraints* of the uncertain optimization program UP in (1). For the time being, we make the following assumption.

Assumption 1. Every P_m is feasible, and its feasibility domain has a nonempty interior. Moreover, the solution x_m^* of P_m exists and is unique.

The existence and uniqueness of x_m^* are here assumed to streamline the presentation. The reader is referred to point 5 in the discussion in section 2.1 to see how these assumptions can be removed.

We recall the following fundamental definition and proposition. Definition 2 was introduced in [5], while Proposition 1 was originally stated in a slightly different but equivalent way in [18].

DEFINITION 2 (support constraint). *Constraint $\delta^{(r)}$, $r \in \{1, \dots, m\}$, is a* support constraint *for P_m if its removal changes the solution of P_m.*

PROPOSITION 1. *The number of support constraints for P_m is at most d, the size of x.*

Suppose now that Δ is endowed with a σ-algebra $\mathcal{D}$ and that a probability $\mathbb{P}$ over $\mathcal{D}$ is assigned. Further assume that m constraints $\delta^{(1)}, \dots, \delta^{(m)}$ are randomly extracted from Δ according to $\mathbb{P}$ in an independent fashion. Differently stated, the

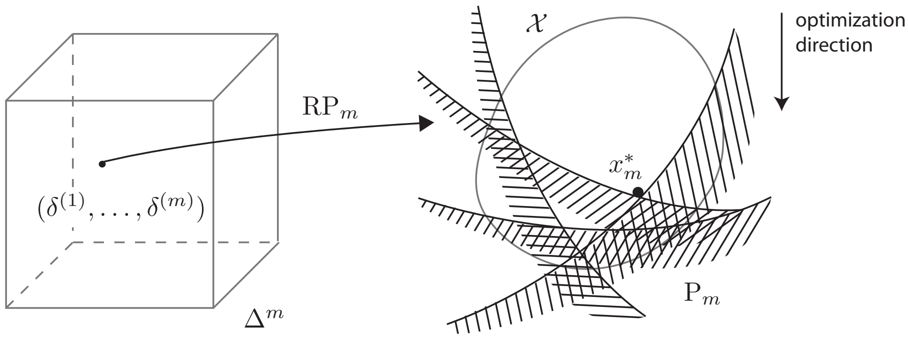

FIG. 1. *RP_m, a map from constraint multiextractions to finite instances P_m of the optimization problem.*

multiextraction $(\delta^{(1)}, \ldots, \delta^{(m)})$ is a random element from the probability space Δ^m equipped with the product probability $\mathbb{P}^m$. Each multiextraction $(\delta^{(1)}, \ldots, \delta^{(m)})$ generates a program P_m, and the map from Δ^m to P_m programs is a *randomized program* RP_m; see Figure 1. Note that this is the same as RP_N in section 1 with the only difference being that we have used here m to indicate the number of constraints, a choice justified by the fact that in this section m plays the role of a generic running argument taking on any integer value, while N represents in section 1 the fixed number of constraints picked by the user for the implementation of the randomized scheme.

We are now ready to introduce the notion of a *fully-supported* problem.

DEFINITION 3 (fully-supported problem). *A finite instance P_m, with $m \geq d$, is fully-supported if the number of support constraints of P_m is* exactly d. *Problem UP is fully-supported if, for any $m \geq d$, RP_m is fully-supported with probability* 1.

The main result of this paper is now stated in the following theorem.

THEOREM 1. *Under Assumption* 1,[2] *it holds that*

$$\mathbb{P}^N\{V(x^*_N) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}; \tag{6}$$

moreover, the bound is tight for all fully-supported *uncertain optimization problems; that is,*

$$\mathbb{P}^N\{V(x^*_N) > \epsilon\} = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}. \tag{7}$$

The proof is given in section 3. The measurability of $\{V(x^*_N) > \epsilon\}$, as well as the measurability of other sets, is assumed in this paper.

One interpretation of Theorem 1 is that the randomized solution is, with high probability, a feasible solution for a chance-constrained problem; see [21].

2.1. Discussion. The following comments are in order.

1. Equation (7) delivers the *exact* probability distribution of the violation $V(x^*_N)$ for all fully-supported problems. Since (7) holds independently of the nature and characteristics of the fully-supported problem, it establishes a fundamental kinship among problems of this prototype class.

[2]See point 5 in section 2.1 for relaxations of this assumption.

Table 1
β vs. β_{old} for different values of N ($\epsilon = 0.05$, $d = 10$).

N	150	300	450	600	750
β	0.78	0.06	$8.8\ 10^{-4}$	$4.8\ 10^{-6}$	$1.5\ 10^{-8}$
β_{old}	$8.8\ 10^{11}$	$4.8\ 10^{11}$	$1.3\ 10^{10}$	$1.1\ 10^{8}$	$4.8\ 10^{5}$

N	900	1050	1200	1350	1500
β	$3.5\ 10^{-11}$	$6.2\ 10^{-14}$	$9.2\ 10^{-17}$	$1.2\ 10^{-19}$	$1.4\ 10^{-22}$
β_{old}	$1.3\ 10^{3}$	2.9	$5.1\ 10^{-3}$	$7.5\ 10^{-6}$	$9.9\ 10^{-9}$

Table 2
N vs. N_{old} for different values of ϵ ($\beta = 10^{-5}$, $d = 10$).

ϵ	0.1	0.05	0.025	0.01	0.005	0.0025	0.001
N	285	581	1171	2942	5895	11749	29513
N_{old}	579	1344	3035	8675	18943	41008	112686

Bound (6) further asserts that all possible sources of non-fully-supportedness can only improve the feasibility properties of the problem.

2. The quantity $\beta := \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ in the right-hand side of (6) and (7) is the tail of a binomial distribution and goes rapidly (exponentially) to zero as N increases. Letting $\beta_{\text{old}} := \binom{N}{d}(1-\epsilon)^{N-d}$ (bound in (2) from [6]), Table 1 provides a comparison between β and β_{old}.

3. A typical use of Theorem 1 consists in selecting ϵ (violation parameter) and β (confidence parameter) in $(0,1)$ and then computing the smallest number N of constraints to be extracted in order to guarantee that $\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \beta$ by solving equation $\beta = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ for N. In Table 2, the values of N and of N_{old} obtained by using the bound in (2) are displayed for different values of ϵ, $\beta = 10^{-5}$ and $d = 10$.

4. A simple example illustrates Theorem 1.

$N = 1650$ points are independently extracted in $\mathbb{R}^2$ according to an unknown probability density $\mathbb{P}$, and the strip of smaller vertical width that contains all of the points is constructed; see Figure 2.

In mathematical terms—letting the points be $(u^{(i)}, y^{(i)})$, $i = 1, \ldots, N$, where u is the horizontal coordinate and y is the vertical coordinate—this amounts to solving the following program:

$$\begin{aligned} \mathrm{P}_N : \ & \min_{x_1, x_2, x_3 \in \mathbb{R}^3} x_1 \\ & \text{subject to: } \left| y^{(i)} - [x_2 u^{(i)} + x_3] \right| \leq x_1, \quad i = 1, \ldots, N, \end{aligned}$$

where $[x_2 u^{(i)} + x_3]$ is the median line of the strip and x_1 is the semiwidth of the strip.

What guarantee do we have that the strip contains at least 99% of the probability mass of $\mathbb{P}$?

One can easily recognize that this question is the same as asking for a guarantee, or a probability, that the violation is less than $\epsilon = 0.01$, and the answer can be found in Theorem 1: this probability is no less than $1 - \sum_{i=0}^{2} \binom{1650}{i} 0.01^i (1-0.01)^{1650-i} \approx 1 - 10^{-5}$. As a matter of fact, this probability is exact since, as it can be verified, this problem is fully-supported.

We can further ask for a different geometrical construction and look for the disk of smaller radius that contains all points; see Figure 3. Again, we are facing a finite

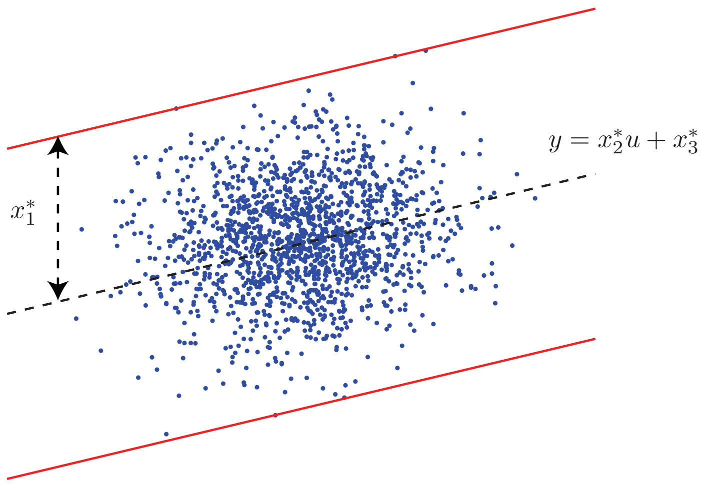

Fig. 2. *Strip of smaller vertical width.*

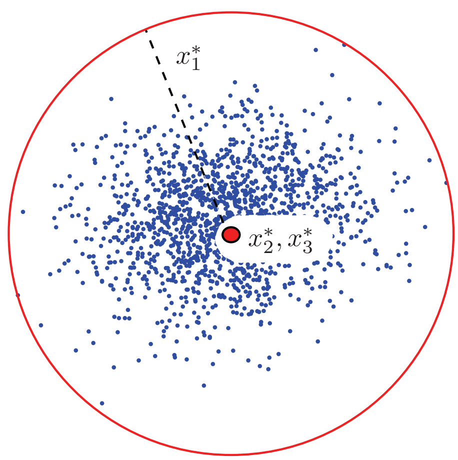

Fig. 3. *Disk of smaller radius.*

convex program

$$\begin{aligned} \mathrm{P}_N : \ & \min_{x_1, x_2, x_3 \in \mathbb{R}^3} x_1 \\ & \text{subject to: } \sqrt{(u^{(i)} - x_2)^2 + (y^{(i)} - x_3)^2} \leq x_1, \quad i = 1, \ldots, N, \end{aligned}$$

where (x_2, x_3) is the center of the disk and x_1 is its radius, and again we can claim with confidence $1 - 10^{-5}$ that the constructed disk will contain at least 99% of the probability mass. In this disk case, the figure $1 - 10^{-5}$ is a lower bound since the problem is not fully-supported, as it can be easily recognized by noting that a configuration with two points away from each other and all of the other points concentrated near the midposition of the first two points generates a disk where the segment joining the first two points is a diameter and only these two points are of support.

Finally, let us compare the probability $1 - 10^{-5}$ with the probability that would have been obtained by applying the previous bound (2) from [6]. Applying the latter, we would find that this probability is no smaller than $1 - 48.4 = -47.4$, a negative number clearly devoid of any meaning and that does not allow us to draw any conclusion as far as the confidence is concerned.

5. Here we discuss the assumption of the existence and uniqueness of the solution of P_m. Suppose first that the solution exists but it may be nonunique. Then, the tie can be broken by selecting among the optimal solutions the one with the minimum Euclidian norm, and one can prove that Theorem 1 holds unchanged.

If we further relax the assumption that the solution exists (note that the solution may not exist even if P_m is feasible since the solution can drift away to infinity), extending Theorem 1 we can show that

$$\mathbb{P}^N\{x_N^* \text{ exists, and } V(x_N^*) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i},$$

where x_N^* is unique after applying the tie-break rule as above. In words, this result says that, when a solution is found, its violation exceeds ϵ with small probability only. In normal problems the nonexistence of the solution is a rare event whose probability exponentially vanishes with N.

3. Proof of Theorem 1.

We first prove that $\mathbb{P}^N\{V(x_N^*) > \epsilon\} = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ for *fully-supported* problems and then that $\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ for every problem.

PART 1: $\mathbb{P}^N\{V(x_N^*) > \epsilon\} = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ FOR FULLY-SUPPORTED PROBLEMS. Consider the solution x_d^* of RP_d (recall that d is the size of x), and let

$$F(\alpha) := \mathbb{P}^d\{V(x_d^*) \leq \alpha\} \tag{8}$$

be the probability distribution of the violation of x_d^*. It is a remarkable fact that this distribution is

$$F(\alpha) = \alpha^d, \tag{9}$$

independent of the problem type.

To prove (9), we have to consider multiextractions of m elements, where m is a generic integer greater than or equal to d. To each multiextraction $(\delta^{(1)}, \ldots, \delta^{(m)}) \in \Delta^m$, associate the indexes of the corresponding d support constraints (this is always possible except for a probability 0 set because the problem is *fully-supported*).[3] Further, group all multiextractions having the same indexes. In this way, $\binom{m}{d}$ sets $S_{\mathcal{I}}$ are constructed forming a partition (up to a probability 0 set) of Δ^m, where $\mathcal{I} \subset \{1, \ldots, m\}$ is a set of cardinality d containing the indexes of the support constraints. We claim that the probability of each of these sets is

$$\mathbb{P}^m\{S_{\mathcal{I}}\} = \int_0^1 (1-\alpha)^{m-d} F(\mathrm{d}\alpha), \tag{10}$$

where $F(\alpha)$ is defined in (8); using (10), later on in the proof, we shall show that $F(\alpha)$ must have the expression in (9).

[3]The fact that a fully-supported problem is one where the RP_m are fully-supported with probability 1, as opposed to *always* fully-supported, is a source of a bit of complication in the proof. On the other hand, requiring always fully-supportedness is too limitative since, e.g., extracting the same constraint m times results in a non-fully-supported P_m.

To establish (10) in a more concrete way, consider one of the sets $S_{\mathcal{I}}$, e.g., the set $S_{\bar{\mathcal{I}}}$ where the support constraints indexes are $1, \ldots, d$. Also let $\widetilde{S}_{\bar{\mathcal{I}}}$ be the set where $\delta^{(d+1)}, \ldots, \delta^{(m)}$ are not violated by the solution generated by $\delta^{(1)}, \ldots, \delta^{(d)}$. It is an intuitive fact that $S_{\bar{\mathcal{I}}}$ and $\widetilde{S}_{\bar{\mathcal{I}}}$ are the same up to a probability 0 set. To streamline the presentation, we accept in the following this fact for granted; however, the interested reader can find full details at the end of this Part 1 of the proof.

We next compute $\mathbb{P}^m\{\widetilde{S}_{\bar{\mathcal{I}}}\}$, which is the same as $\mathbb{P}^m\{S_{\bar{\mathcal{I}}}\}$.

Select fixed values for $\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(d)}$, and let α be the violation of the solution with these d constraints only. Then, the probability that $\delta^{(d+1)}, \ldots, \delta^{(m)}$ fall in the nonviolated set, that is, $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(d)}, \delta^{(d+1)}, \ldots, \delta^{(m)}) \in \widetilde{S}_{\bar{\mathcal{I}}}$, is $(1-\alpha)^{m-d}$.

Integrating over the domain Δ^d for $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(d)})$, we then have

$$
\begin{aligned}
&\mathbb{P}^m\left\{\widetilde{S}_{\bar{\mathcal{I}}}\right\}\\
&= [\text{letting } x^*_{\bar{\mathcal{I}}} \text{ be the solution with constraints } \bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(d)} \text{ only}]\\
&= \int_{\Delta^d} (1-\alpha(x^*_{\bar{\mathcal{I}}}))^{m-d} \mathbb{P}^d(\mathrm{d}\bar{\delta}^{(1)}, \ldots, \mathrm{d}\bar{\delta}^{(d)})\\
&= \int_0^1 (1-\alpha)^{m-d} F(\mathrm{d}\alpha),
\end{aligned}
$$

where the third equality is a change of variables from the domain $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(d)})$ to that of the violation of the corresponding solution.

Since $\mathbb{P}^m\{S_{\bar{\mathcal{I}}}\} = \mathbb{P}^m\{\widetilde{S}_{\bar{\mathcal{I}}}\}$ and this probability is the same for any other set $S_{\mathcal{I}}$, (10) remains proven.

Now turn back to (9). Recalling that the sets $S_{\mathcal{I}}$ form a partition of Δ^m up to a probability 0 set and that $\mathbb{P}^m\{\Delta^m\} = 1$, (10) yields

$$
\binom{m}{d} \int_0^1 (1-\alpha)^{m-d} F(\mathrm{d}\alpha) = 1 \quad \forall m \geq d. \tag{11}
$$

Expression $F(\alpha) = \alpha^d$ in (9) is indeed a solution of (11) (integration by parts); on the other hand, no other solutions exist since determining an F satisfying (11) is a moment problem for a distribution with finite support, and its solution is unique; see, e.g., Chapter II, section 12.9, Corollary 1 of [26]. Thus, it remains proven that $F(\alpha)$ must have the expression (9).

To conclude the proof of Part 1, consider now the problem with N constraints and partition set $\{(\delta^{(1)}, \ldots, \delta^{(N)}) : \; V(x^*_N) > \epsilon\}$ by intersecting it with the $\binom{N}{d}$ sets $S_{\mathcal{I}}$ grouping multiextractions such that the d support constraints have the same indexes. We then have

$$
\begin{aligned}
&\mathbb{P}^N\{V(x^*_N) > \epsilon\}\\
&= \mathbb{P}^N\Big\{ \cup_{\mathcal{I}} \{V(x^*_N) > \epsilon \text{ and } x^*_N \text{ is supported by the constraints}\\
&\qquad \text{with indexes in } \mathcal{I}\}\Big\}\\
&= [\mathbb{I}_A \text{ is the indicator function of set } A;\ \text{i.e., } \mathbb{I}_A = 1 \text{ over } A,\ \text{and } \mathbb{I}_A = 0 \text{ otherwise}]\\
&= \binom{N}{d} \int_{\Delta^d} (1-\alpha(x^*_{\bar{\mathcal{I}}}))^{N-d} \mathbb{I}_{\{\alpha(x^*_{\bar{\mathcal{I}}}) > \epsilon\}} \mathbb{P}^d(\mathrm{d}\bar{\delta}^{(1)}, \ldots, \mathrm{d}\bar{\delta}^{(d)})\\
&= \binom{N}{d} \int_\epsilon^1 (1-\alpha)^{N-d}\, F(\mathrm{d}\alpha)
\end{aligned}
$$

$$
\begin{aligned}
&= [\text{since } F(\mathrm{d}\alpha) = d\alpha^{d-1}\,\mathrm{d}\alpha] \\
&= \binom{N}{d} \int_{\epsilon}^{1} \left[(1-\alpha)^{N-d} d\alpha^{d-1}\right] \,\mathrm{d}\alpha \\
&= [\text{integrating by parts}] \\
&= \binom{N}{d} \left[-\frac{(1-\alpha)^{N-d+1}}{N-d+1} d\alpha^{d-1} \Big|_{\epsilon}^{1} + \int_{\epsilon}^{1} \frac{(1-\alpha)^{N-d+1}}{N-d+1} d(d-1)\alpha^{d-2}\,\mathrm{d}\alpha \right] \\
&= \binom{N}{d-1} \epsilon^{d-1}(1-\epsilon)^{N-d+1} + \binom{N}{d-1} \int_{\epsilon}^{1} (1-\alpha)^{N-d+1}(d-1)\alpha^{d-2}\,\mathrm{d}\alpha \\
&= \cdots \\
&= \binom{N}{d-1} \epsilon^{d-1}(1-\epsilon)^{N-d+1} + \ldots + \binom{N}{1}\epsilon(1-\epsilon)^{N-1} + \binom{N}{1} \int_{\epsilon}^{1} (1-\alpha)^{N-1}\,\mathrm{d}\alpha \\
&= \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^{i}(1-\epsilon)^{N-i}.
\end{aligned}
$$

Proof of the fact that $S_{\bar{\mathcal{I}}} = \widetilde{S}_{\bar{\mathcal{I}}}$ up to a probability zero set.

$S_{\bar{\mathcal{I}}} \subseteq \widetilde{S}_{\bar{\mathcal{I}}}$: Take a $(\delta^{(1)}, \ldots, \delta^{(m)}) \in S_{\bar{\mathcal{I}}}$ and eliminate a constraint among $\delta^{(d+1)}, \ldots, \delta^{(m)}$. Since this constraint is not of support, the solution remains unchanged; moreover, it is easy to see that the first d constraints are still the support constraints for the problem with $m-1$ constraints. If we now remove another constraint among those which are not of support, the conclusion is similarly drawn that the solution remains unchanged and that the first d constraints are still the support constraints for the problem with $m-2$ constraints. Proceeding this way until all constraints but the first d are removed, we obtain that the solution with the sole d support constraints $\delta^{(1)}, \ldots, \delta^{(d)}$ in place is the same as the solution with all m constraints. Since no constraint among $\delta^{(d+1)}, \ldots, \delta^{(m)}$ can be violated by the solution with all m constraints and such a solution is the same as the one with only the first d constraints, it follows that $(\delta^{(1)}, \ldots, \delta^{(m)}) \in \widetilde{S}_{\bar{\mathcal{I}}}$.

$\widetilde{S}_{\bar{\mathcal{I}}} \subseteq S_{\bar{\mathcal{I}}}$ up to a probability 0 set: Suppose now that $\delta^{(d+1)}, \ldots, \delta^{(m)}$ are not violated by the solution generated by $\delta^{(1)}, \ldots, \delta^{(d)}$, i.e., $(\delta^{(1)}, \ldots, \delta^{(m)}) \in \widetilde{S}_{\bar{\mathcal{I}}}$. A simple reasoning reveals that $(\delta^{(1)}, \ldots, \delta^{(m)})$ does not belong to any one of sets $S_{\mathcal{I}}$, $\mathcal{I} \neq \bar{\mathcal{I}}$. In fact, adding nonviolated constraints to $\delta^{(1)}, \ldots, \delta^{(d)}$ does not change the solution, and each of the added constraints can be removed back without altering the solution. Therefore, none of the constraints $\delta^{(d+1)}, \ldots, \delta^{(m)}$ can be of support, and hence the multiextraction is not in $S_{\mathcal{I}}$, $\mathcal{I} \neq \bar{\mathcal{I}}$. It follows that $\widetilde{S}_{\bar{\mathcal{I}}}$ is a subset of the complement of $\cup_{\mathcal{I}, \mathcal{I} \neq \bar{\mathcal{I}}} S_{\mathcal{I}}$, which is $S_{\bar{\mathcal{I}}}$ up to a probability 0 set.

PART 2: $\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}$ FOR EVERY PROBLEM. A non-fully-supported problem admits with nonzero probability randomized instances where the number of support constraints is less than d. A support constraint has to be an active constraint, and the typical reason for a lack of support constraints is that at the optimum the active constraints are less than d; see Figure 4. To carry on a proof along lines akin to those for the fully-supported case, we are well-advised to generalize the notion of solution to that of ball-solution; a ball-solution has always at least d active constraints. For simplicity, we henceforth assume that constraints are not trivial, i.e., $\mathcal{X}_\delta \neq \mathbb{R}^d \ \forall \delta \in \Delta$.

Roughly speaking, given an optimization problem whose solution is x_m^*, its *ball-solution* is a ball centered in x_m^* and whose radius has been enlarged until the ball

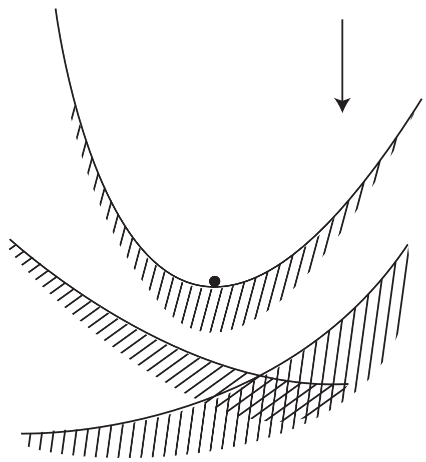

Fig. 4. *A two-dimensional problem with only one active constraint which is of support.*

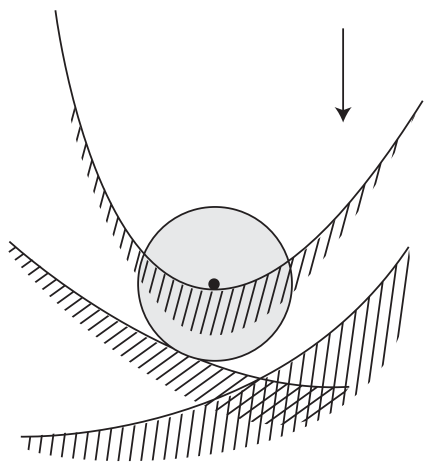

Fig. 5. *Ball-solution.*

touches the frontier of d constraints. See Figure 5 for an example of a ball-solution. The mathematical definition of a ball-solution is as follows.

Definition 4 (ball-solution). *Consider a finite instance P_m of UP with $m \geq d$, and let x_m^* be its solution. The* ball-solution $\mathcal{B}(x_m^*, r_m^*)$ *of P_m is the largest closed ball centered in x_m^* fully contained in the feasibility domain of all constraints, with the exception of at most $d-1$ of them; i.e., $\mathcal{X}_{\delta^{(i)}} \cap \mathcal{B}(x_m^*, r_m^*) = \mathcal{B}(x_m^*, r_m^*)$ for all i's, except at most $d-1$ of them.*

Note also that, when active constraints are d or more, $r_m^* = 0$ and $\mathcal{B}(x_m^*, r_m^*)$ reduces to the standard solution x_m^*. Moreover, a ball-solution $\mathcal{B}(x_m^*, r_m^*)$ need not be contained in $\mathcal{X}$, although its center x_m^* does.

The notion of active constraint can be generalized to balls by saying that a constraint is active for a ball if the ball touches the frontier of the constraint. If in addition the ball is fully contained in the constraint, then the constraint is said to be strictly active. See Figure 6 for a graphical illustration of active and strictly active constraints for a ball, while the precise definition is as follows.

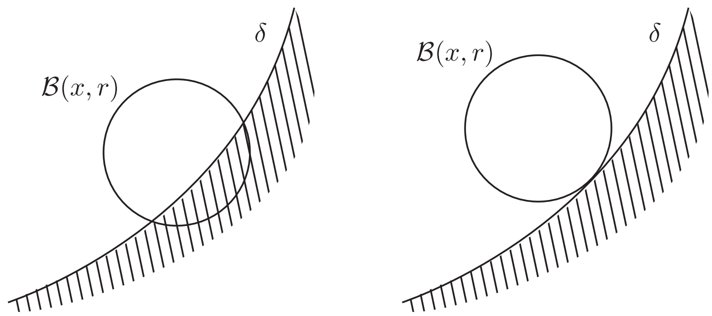

FIG. 6. *Active and strictly active constraints for a ball.*

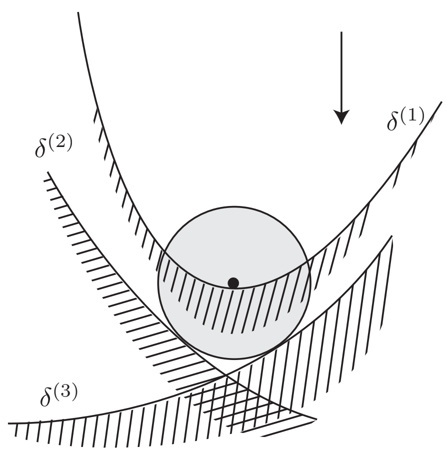

FIG. 7. *Only $\delta^{(1)}$ is a ball-support constraint.*

DEFINITION 5 (active constraint for a ball). *A constraint δ is* active *for a ball $\mathcal{B}(x,r)$ if $\mathcal{X}_\delta \cap \mathcal{B}(x,r) \neq \emptyset$ and $\mathcal{X}_\delta \cap \mathcal{B}(x,r+h) \neq \mathcal{B}(x,r+h)$ $\forall h > 0$. If in addition $\mathcal{X}_\delta \cap \mathcal{B}(x,r) = \mathcal{B}(x,r)$, $\mathcal{X}_\delta$ is said to be* strictly active.

If the ball is a single point, active and strictly active are the same and reduce to the standard notion of active.

By construction, a ball-solution has at least d active constraints. To go back to the track of the proof in Part 1, however, we need d support constraints, not just active constraints. The following definition naturally extends the notion of support constraint to the case of ball-solutions.

DEFINITION 6 (ball-support constraint). *Constraint $\delta^{(r)}$, $r \in \{1,\dots,m\}$, is a* ball-support constraint *for P_m if its removal changes the ball-solution of P_m.*

An active constraint is not necessarily a ball-support constraint, nor does a P_m necessarily have to have d ball-support constraints (see Figure 7, where $\delta^{(2)}$ and $\delta^{(3)}$ are not of support). It is clear that the number of ball-support constraints is less than or equal to d. The case with less than d ball-support constraints is regarded as *degenerate* and needs to be treated separately. We thus split the remaining part of the proof in two sections: Part 2a (fully-ball-supported problems) and Part 2b (degenerate problems). Before proceeding, we are well-advised to give a formal definition of fully-ball-supported problems.

DEFINITION 7 (fully-ball-supported problem). *A finite instance P_m, with $m \geq d$, is fully-ball-supported if the number of ball-support constraints of P_m is d. Problem UP is fully-ball-supported if, for any $m \geq d$, RP_m is fully-ball-supported with probability* 1.

PART 2a: FULLY-BALL-SUPPORTED PROBLEMS. We start by introducing the notion of a constraint violated by a ball: a constraint δ is violated by $\mathcal{B}(x,r)$ if $\mathcal{X}_\delta \cap \mathcal{B}(x,r) \neq \mathcal{B}(x,r)$. The definition of probability of violation then generalizes naturally to the ball case.

DEFINITION 8 (violation probability of a ball). *The* violation probability *of a ball $\mathcal{B}(x,r)$, $x \in \mathcal{X}$, is defined as $V_\mathcal{B}(x,r) = \mathbb{P}\{\delta \in \Delta : \mathcal{X}_\delta \cap \mathcal{B}(x,r) \neq \mathcal{B}(x,r)\}$.*

Clearly, for any x, $V_\mathcal{B}(x,r) \geq V(x)$. Hence, if $\mathcal{B}(x_N^*, r_N^*)$ is the ball-solution of RP_N, we have

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \mathbb{P}^N\{V_\mathcal{B}(x_N^*, r_N^*) > \epsilon\}. \tag{12}$$

Below, we show that a result similar to (7) holds for fully-ball-supported problems, namely,

$$\mathbb{P}^N\{V_\mathcal{B}(x_N^*, r_N^*) > \epsilon\} = \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}, \tag{13}$$

and this result together with (12) leads to the thesis

$$\mathbb{P}^N\{V(x_N^*) > \epsilon\} \leq \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}.$$

The proof of (13) is verbatim the same as the proof of Part 1 provided that one substitutes

- *solution* with *ball-solution*,
- *support constraint* with *ball-support constraint*,
- *violation probability* V with *violation probability of a ball* $V_\mathcal{B}$,

with only one exception: the part where we proved that $\boldsymbol{S}_{\bar{\mathcal{I}}} \subseteq \widetilde{\boldsymbol{S}}_{\bar{\mathcal{I}}}$ has to be modified in a way that we spell out in the following.

The first rationale to conclude that "the solution with only the d support constraints $\delta^{(1)}, \ldots, \delta^{(d)}$ in place is the same as the solution with all m constraints" is still valid and leads in our present context to the fact that the ball-solution with only the d ball-support constraints $\delta^{(1)}, \ldots, \delta^{(d)}$ in place is the same as the ball-solution with all m constraints. Instead, the last argument with which we concluded that $\boldsymbol{S}_{\bar{\mathcal{I}}} \subseteq \widetilde{\boldsymbol{S}}_{\bar{\mathcal{I}}}$ is no longer valid since ball-solutions can violate constraints.

To amend it, suppose for the purpose of contradiction that a constraint among $\delta^{(d+1)}, \ldots, \delta^{(m)}$, say, $\delta^{(d+1)}$, is violated by the ball-solution with d constraints. Two cases can occur: (i) the ball-solution has only one strictly active constraint among $\delta^{(1)}, \ldots, \delta^{(d)}$; or (ii) it has more than one. In case (i), $d-1$ constraints among $\delta^{(1)}, \ldots, \delta^{(d)}$ are violated by the ball-solution so that, with the extra $\delta^{(d+1)}$ violated constraint, the number of violated constraints of the ball-solution with m constraints would add up to at least d, and this contradicts the definition of ball-solution. If instead (ii) is true, a simple thought reveals that, with one more constraint $\delta^{(d+1)}$ violated by the ball-solution, the strictly active constraints (which, in this case, are more than one) cannot be of ball-support for the problem with m constraints, and this contradicts the fact that $(\delta^{(1)}, \ldots, \delta^{(m)}) \in S_{\bar{\mathcal{I}}}$.

PART 2b: DEGENERATE PROBLEMS. For not being fully-ball-supported, a finite problem P_m needs to have more than one strictly active constraint, a circumstance which requires that constraints are not "generically" distributed. This observation is at the basis of the rather technical proof of Part 2b, which proceeds along the following steps:

Step 1. A constraint "heating" is introduced; heating scatters constraints around, and the resulting heated problem is shown to be fully-ball-supported; by resorting to the result in Part 2a, conclusions are derived about the violation properties of the heated problem.

Step 2. It is shown that the solution of the original problem is recovered by cooling the heated problem down.

Step 3. The violation properties of the original (nonheated) problem are determined from the violation properties of the heated problem by a limiting process.

Step 1 (*heating*). Let $\Delta' := \Delta \times \mathcal{B}_\rho$, where $\rho > 0$ is the *heating parameter* and $\mathcal{B}_\rho \subset \mathbb{R}^d$ is the closed ball centered in the origin with radius ρ, and let $\mathbb{P}' := \mathbb{P} \times \mathbb{U}$ be the probability in Δ' obtained as the product probability between $\mathbb{P}$ and the uniform probability $\mathbb{U}$ in $\mathcal{B}_\rho$. Each $z \in \mathcal{B}_\rho$ represents a constraint translation, and the heated uncertain program (HUP) is defined as

$$\begin{aligned}\text{HUP}: \min_{x \in \mathcal{X} \subseteq \mathbb{R}^d} & \; c^T x \\ & \text{subject to: } x \in [\mathcal{X}_\delta + z], \quad (\delta, z) \in \Delta',\end{aligned}$$

where $[\mathcal{X}_\delta + z]$ is set $\mathcal{X}_\delta$ translated by z, and the new uncertain parameter (δ, z) allows for different selections of $\mathcal{X}_\delta$ constraints as well as for any translation z in $\mathcal{B}_\rho$. We show that HUP is fully-ball-supported.

To start with, consider a given deterministic ball $\mathcal{B}(x, r)$. We first prove that the strictly active constraints $\delta' \in \Delta'$ for $\mathcal{B}(x, r)$ form a set of zero-probability $\mathbb{P}'$, and later on from this we shall conclude that HUP is fully-ball-supported.

Let $\delta' = (\delta, z)$, and $\mathbb{I}_A$ indicate the indicator function of set A, and write

$$\begin{aligned}
&\mathbb{P}'\{\delta' \text{ is strictly active for } \mathcal{B}(x,r)\} \\
&= \int_{\Delta'} \mathbb{I}_{\{\delta' \text{ is strictly active for } \mathcal{B}(x,r)\}} \mathbb{P}'(\mathrm{d}\delta') \\
&= [\text{by Fubini's theorem [23]}] \\
&= \int_{\Delta} \left[\int_{\mathcal{B}_\rho} \mathbb{I}_{\{(\delta,z) \text{ is strictly active for } \mathcal{B}(x,r)\}} \frac{\mathrm{d}z}{\mathrm{Vol}(\mathcal{B}_\rho)} \right] \mathbb{P}(\mathrm{d}\delta).
\end{aligned} \tag{14}$$

The result that

$$\mathbb{P}'\{\delta' \text{ is strictly active for } \mathcal{B}(x,r)\} = 0 \tag{15}$$

is established by showing that the term within square brackets in (14) is null for all δ's.

Fix a δ, and let $C = \{z \in \mathcal{B}_\rho : \mathcal{B}(x, r) \subseteq [\mathcal{X}_\delta + z]\}$ be the set of translations not violating $\mathcal{B}(x, r)$. We show that C is convex and that the set $\{z \in \mathcal{B}_\rho : (\delta, z) \text{ is strictly active for } \mathcal{B}(x, r)\}$ belongs to ∂C, the boundary of C. Since the boundary of a convex set has zero Lebesgue measure,[4] the desired result that the term within square brackets in (14) is null follows, viz.

$$\int_{\mathcal{B}_\rho} \mathbb{I}_{\{(\delta,z) \text{ is strictly active for } \mathcal{B}(x,r)\}} \frac{\mathrm{d}z}{\mathrm{Vol}(\mathcal{B}_\rho)} = 0. \tag{16}$$

The convexity of C is immediate: let $z_1, z_2 \in C$, that is, $\mathcal{B}(x,r) \subseteq [\mathcal{X}_\delta + z_1]$ and $\mathcal{B}(x,r) \subseteq [\mathcal{X}_\delta + z_2]$, or, equivalently, $\mathcal{B}(x,r) - z_1 \subseteq \mathcal{X}_\delta$ and $\mathcal{B}(x,r) - z_2 \subseteq \mathcal{X}_\delta$. From the convexity of $\mathcal{X}_\delta$, it follows that $\mathcal{B}(x,r) - \alpha z_1 - (1-\alpha)z_2 \subseteq \mathcal{X}_\delta$ $\forall \alpha \in [0,1]$; that is, $\alpha z_1 + (1-\alpha)z_2 \in C$ and C is convex.

Consider now an interior point z of C (if any); i.e., there exists a ball centered in z all contained in C. This means that $[\mathcal{X}_\delta + z]$ can be moved around in all directions by a small quantity, and $\mathcal{B}(x,r)$ remains contained in it. It easily follows that (δ, z) cannot be strictly active, and, thus, $\{z \in \mathcal{B}_\rho : (\delta, z) \text{ is strictly active for } \mathcal{B}(x,r)\}$ has to belong to ∂C.

Wrapping up, (16) is established and, substituting in (14), (15) is obtained.

We next prove that (15) entails the fact that HUP is fully-ball-supported.

Consider a finite instance HP_m of HUP with $m \geq d$. One by one, eliminate $m-d$ constraints choosing at any time a constraint among those nonviolated by the ball-solution in such a way that the ball-solution does not change. This is certainly possible because the ball-support constraints are at most d. In the end, we are left with d constraints, say, the first d $\delta'^{(1)}, \ldots, \delta'^{(d)}$. A simple thought reveals that these d constraints are actually of ball-support for HP_m, provided that none of the other $m-d$ constraints that have been removed were strictly active.

Repeat the same above procedure for every m-ple of constraints (that is, for every HP_m generated by HUP), and group together all of the m-ples for which the procedure returns in the end the first d constraints $\delta'^{(1)}, \ldots, \delta'^{(d)}$. Call this group of m-ples G. We shall show that the probability of the m-ples in G such that HP_m is not fully-ball-supported is zero, and from this—by the observation that only a finite number $\binom{m}{d}$ of groups of m-ples can be similarly constructed—the final conclusion that HUP is fully-ball-supported will be secured.

Select fixed values $\bar{\delta}'^{(1)}, \ldots, \bar{\delta}'^{(d)}$ for the first d constraints, and consider the ball-solution $\mathcal{B}$ that these constraints generate. Let the other $m-d$ constraints vary in such a way that the m-ple $\bar{\delta}'^{(1)}, \ldots, \bar{\delta}'^{(d)}, \delta'^{(d+1)}, \ldots, \delta'^{(m)}$ belongs to G. For one such m-ple to correspond to a *non*-fully-ball-supported HP_m, at least one among the $m-d$ constraints $\delta'^{(d+1)}, \ldots, \delta'^{(m)}$ must be strictly active for $\mathcal{B}$, but we have proven in (15) that this happens with probability zero. Integrating over all possible values $\bar{\delta}'^{(1)}, \ldots, \bar{\delta}'^{(d)}$ for the first d constraints, the conclusion is drawn that the *non*-fully-ball-supported HP_m in G have zero probability.

Hence, by the above observation that there are only a finite number $\binom{m}{d}$ of groups and by the fact that $\binom{m}{d}$ times zero is zero, we obtain that HUP is fully-ball-supported.

To conclude Step 1, note that if we suppose to extract N constraints $\delta'^{(1)}, \ldots, \delta'^{(N)}$ from Δ', according to probability $\mathbb{P}'$ and in an independent fashion, and we denote by x'^*_N the corresponding solution, the result of Part 2a can be invoked to establish

[4]This simple fact follows from the observation that a convex set C in $\mathbb{R}^d$ either belongs to a flat of dimension $d-1$—and therefore C has zero $\mathbb{R}^d$ Lebesgue measure—or admits an interior point $\bar{z}$, and every half-line from $\bar{z}$ crosses the boundary of C in only one point (see, e.g., Propositions 1.1.13 and 1.1.14 in [19]).

that

$$(\mathbb{P}')^N\{V'(x'^*_N) > \epsilon\} \le \sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i}, \tag{17}$$

where $V'(x)$ is the probability of violation for the heated problem (i.e., $V'(x) = \mathbb{P}'\{(\delta, z) \in \Delta' : \ x \notin [\mathcal{X}_\delta + z]\}$). Equation (17) is the final result to which we wanted to arrive in this heating Step 1.

Step 2 (*cooling*). Fix a multiextraction $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(N)}) \in \Delta^N$, and consider x^*_N, the solution of the original optimization problem P_N with such constraints. We remark that in all of Step 2 the multiextraction $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(N)})$ is kept fixed and never changed throughout. Consider a closed ball $\mathcal{B}(x_f, r_f)$, $r_f > 0$, in the feasibility domain of P_N, which exists because the feasibility domain of P_N has a nonempty interior. Further, let $\rho_k \downarrow 0$ be a sequence of heating parameters monotonically decreasing to zero (*cooling of the heating parameter*) and such that $\rho_1 < \frac{r_f}{2}$. For all ρ_k, consider the heated versions of $(\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(N)})$, namely, $((\bar{\delta}^{(1)}, z_k^{(1)}), \ldots, (\bar{\delta}^{(N)}, z_k^{(N)}))$ where $z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}$, and let $x'^*_N(z_k^{(1)}, \ldots, z_k^{(N)})$ be the solution of the heated optimization problem HP_N with heated constraints $(\bar{\delta}^{(1)}, z_k^{(1)}), \ldots, (\bar{\delta}^{(N)}, z_k^{(N)})$. The goal of Step 2 is to prove that

$$\sup_{z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}} \left\| x'^*_N \left(z_k^{(1)}, \ldots, z_k^{(N)} \right) - x^*_N \right\| \longrightarrow 0 \quad \text{as } k \to \infty; \tag{18}$$

that is, the solution of the original problem is recovered by cooling the heated problem down.[5]

For brevity, from now on we omit the arguments $z_k^{(1)}, \ldots, z_k^{(N)}$ and write x'^*_N for $x'^*_N(z_k^{(1)}, \ldots, z_k^{(N)})$.

We first show that

$$\limsup_{k \to \infty} \sup_{z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}} c^T x'^*_N \le c^T x^*_N. \tag{19}$$

Following Figure 8, consider the convex hull $\mathrm{co}[\mathcal{B}(x_f, r_f) \cup x^*_N]$ generated by the feasibility ball $\mathcal{B}(x_f, r_f)$ and the solution x^*_N of the original problem with constraints $\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(N)}$. By convexity, $\mathrm{co}[\mathcal{B}(x_f, r_f) \cup x^*_N]$ is feasible for the original problem P_N. Construct the closed ball $\mathcal{B}(x_k, \rho_k) \subset \mathrm{co}[\mathcal{B}(x_f, r_f) \cup x^*_N]$ with radius ρ_k, whose center x_k is as close as possible to x^*_N and lies on the line segment connecting x_f with x^*_N (this ball exists since $\rho_1 < r_f$; the assumed stricter condition that $\rho_1 < \frac{r_f}{2}$ is required in the next construction). Clearly, $x_k \to x^*_N$ as $k \to \infty$. Since x_k is in the feasibility domain of P_N at a distance at least ρ_k from where $\bar{\delta}^{(1)}, \ldots, \bar{\delta}^{(N)}$ are violated, x_k is also in the feasibility domain of every heated problem HP_N with heating parameter ρ_k. Thus,

$$\limsup_{k \to \infty} \sup_{z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}} c^T x'^*_N \le \limsup_{k \to \infty} c^T x_k = c^T x^*_N;$$

that is, (19) holds.

[5]Although result (18) has an intuitive appeal, its proof is rather technical. The reader not interested in these technical details can jump to Step 3 from here without loss of continuity.

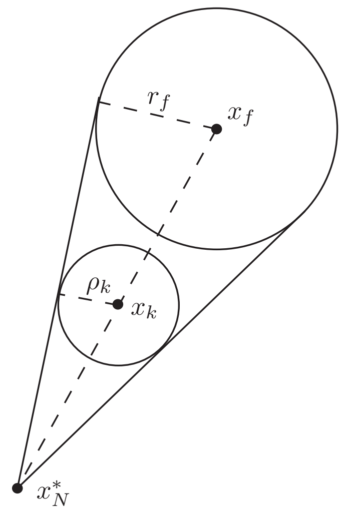

FIG. 8. *Convex hull of* $\mathcal{B}(x_f, r_f)$ *and* x_N^*, *and construction of* $\mathcal{B}(x_k, \rho_k)$.

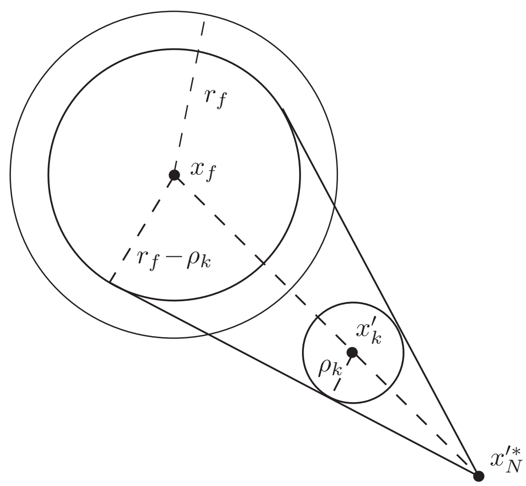

FIG. 9. *Convex hull of* $\mathcal{B}(x_f, r_f - \rho_k)$ *and* x'^*_N, *and construction of* $\mathcal{B}(x'_k, \rho_k)$.

Next, we construct a new convex hull which will allow us to reformulate goal (18) in a different, handier, way. Based on this reformulation, (18) will then be established in light of (19).

The new convex hull is $\mathrm{co}[\mathcal{B}(x_f, r_f - \rho_k) \cup x'^*_N]$; see Figure 9. Note that, for a given k, $\mathcal{B}(x_f, r_f - \rho_k)$ is a fixed ball, while instead x'^*_N depends on the specific choice of $z_k^{(1)}, \dots, z_k^{(N)} \in \mathcal{B}_{\rho_k}$; this means that there are actually as many convex hulls as choices of $z_k^{(1)}, \dots, z_k^{(N)}$. Moreover, $\mathrm{co}[\mathcal{B}(x_f, r_f - \rho_k) \cup x'^*_N]$ is feasible for problem HP_N with constraints translated by $z_k^{(1)}, \dots, z_k^{(N)}$ since $\mathcal{B}(x_f, r_f - \rho_k)$ and x'^*_N are. Construct then the closed ball $\mathcal{B}(x'_k, \rho_k) \subseteq \mathrm{co}[\mathcal{B}(x_f, r_f - \rho_k) \cup x'^*_N]$ with radius ρ_k, whose center x'_k is as close as possible to x'^*_N and lies on the line segment connecting x_f with x'^*_N (this ball exists since $\rho_1 < \frac{r_f}{2}$). Note that x'_k depends on $z_k^{(1)}, \dots, z_k^{(N)}$, too.

Since x'_k is in the feasibility domain of HP_N with constraints translated by $z_k^{(1)}, \ldots, z_k^{(N)}$ at a distance at least ρ_k from where these translated constraints are violated, x'_k is also in the feasibility domain of P_N.

What is different from the previous convex hull construction is that we cannot here easily conclude that $x'_k \to x'^*_N$ as $k \to \infty$ since x'^*_N is not a fixed point (it depends on $z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}$, a ball that changes with k). We can still, however, secure a result that goes along a similar line, namely, that

$$x'_k = \alpha_k x_f + (1-\alpha_k) x'^*_N, \qquad \text{where } \alpha_k = \frac{\rho_k}{r_f - \rho_k} \longrightarrow 0 \text{ as } k \to \infty, \tag{20}$$

as it results from Figure 9 by a simple proportion argument.[6] Reorganizing terms in this equation, we obtain $x'^*_N - x^*_N = -\frac{\alpha_k}{1-\alpha_k}(x_f - x^*_N) + \frac{1}{1-\alpha_k}(x'_k - x^*_N)$, from which

$$\|x'^*_N - x^*_N\| \le \frac{\alpha_k}{1-\alpha_k}\|x_f - x^*_N\| + \frac{1}{1-\alpha_k}\|x'_k - x^*_N\|.$$

We are now ready to reformulate goal (18) in a different way.

Note that the norm in (18) is the same as the left-hand side of the latter equation. On the right-hand side, $\|x_f - x^*_N\|$ is a fixed quantity multiplied by scalar $\frac{\alpha_k}{1-\alpha_k}$ which goes to zero. So, this first term vanishes. In the second term, scalar $\frac{1}{1-\alpha_k} \to 1$, and hence (18) is equivalent to

$$\sup_{z_k^{(1)}, \ldots, z_k^{(N)} \in \mathcal{B}_{\rho_k}} \|x'_k - x^*_N\| \longrightarrow 0 \quad \text{as } k \to \infty. \tag{21}$$

The goal of establishing (18) is finally achieved by proving (21) by contradiction.

Suppose that (21) is false; then, for a given $\mu > 0$, we can choose translations $\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)} \in \mathcal{B}_{\rho_k}$, $k = 1, 2, \ldots$, such that

$$\left\| x'_k\left(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)}\right) - x^*_N \right\| > \mu \quad \forall k,$$

where we have here preferred to explicitly indicate dependence of x'_k on $\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)}$.

Note that $x'_k(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)})$ is asymptotically superoptimal for problem P_N:

$$\begin{aligned}
&\limsup_{k\to\infty} c^T x'_k\left(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)}\right) \\
&\le [\text{using (20) and since } \alpha_k \to 0] \\
&\le \limsup_{k\to\infty} \sup_{z_k^{(1)}, \ldots, z_k^{(N)}} c^T x'^*_N \\
&\le [\text{using (19)}] \\
&\le c^T x^*_N.
\end{aligned} \tag{22}$$

The line segment connecting $x'_k(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)})$ with x^*_N intersects the surface of the ball with center x^*_N and radius μ in a point that we name x'^S_k. x'^S_k is still feasible for P_N being a convex combination of x^*_N and $x'_k(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)})$, both feasible points for P_N. In addition, since $x'_k(\bar{z}_k^{(1)}, \ldots, \bar{z}_k^{(N)})$ is asymptotically superoptimal for P_N (see (22)) and x^*_N is the solution of P_N, x'^S_k is asymptotically superoptimal for P_N, too, i.e., $\limsup_{k\to\infty} c^T x'^S_k \le c^T x^*_N$. Finally, since x'^S_k belongs to a compact, it admits a convergent subsequence to, say, x'^S_∞, a point which is still feasible for P_N due to

[6] Note that (20) does not imply that $x'_k \to x'^*_N$ since x'^*_N could in principle escape to infinity.

the fact that the feasibility domain of P_N is closed. x'^S_∞ would thus be feasible and superoptimal for P_N, so contradicting the uniqueness of the solution of P_N.

This concludes Step 2.

Step 3 (*drawing the conclusions*). The theorem statement that $\mathbb{P}^N\{V(x^*_N) > \epsilon\} \leq \sum_{i=0}^{d-1}\binom{N}{i}\epsilon^i(1-\epsilon)^{N-i}$ is established in this Step 3 along the following lines: by the convergence result (18) in Step 2, a bad multiextraction $(\bar{\delta}^{(1)},\dots,\bar{\delta}^{(N)})$ (i.e., one such that $V(x^*_N) > \epsilon$) is shown to generate bad heated multiextractions $((\bar{\delta}^{(1)},z^{(1)}_k),\dots,(\bar{\delta}^{(N)},z^{(N)}_k))$ for k large enough; we thus have that the probability of bad multiextractions can be bounded by the probability of bad heated multiextractions; by then using the bound for the probability of bad heated multiextractions derived in Step 1, the thesis follows.

Fix a *bad* multiextraction $(\bar{\delta}^{(1)},\dots,\bar{\delta}^{(N)}) \in \Delta^N$, and consider x^*_N, the solution of the optimization problem P_N with constraints $\bar{\delta}^{(1)},\dots,\bar{\delta}^{(N)}$. For an additional constraint $\delta \in \Delta$ to be violated by x^*_N, x^*_N must belong to the complement of $\mathcal{X}_\delta$, i.e., $\mathcal{X}^c_\delta$. Since $\mathcal{X}^c_\delta$ is open, we then have the fact that there exists a small enough ball centered in x^*_N fully contained in $\mathcal{X}^c_\delta$. Thus,

$$\{\delta \in \Delta :\ x^*_N \notin \mathcal{X}_\delta\} = \bigcup_{n=1,2,\dots}\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/n) \subseteq \mathcal{X}^c_\delta\}, \tag{23}$$

and

$$\begin{aligned}\epsilon &< [\text{since } \left(\bar{\delta}^{(1)},\dots,\bar{\delta}^{(N)}\right) \text{ is bad}]\\ &< V(x^*_N)\\ &= \mathbb{P}\{\delta \in \Delta :\ x^*_N \notin \mathcal{X}_\delta\}\\ &= [\text{using (23)}]\\ &= \mathbb{P}\Big\{\bigcup_{n=1,2,\dots}\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/n) \subseteq \mathcal{X}^c_\delta\}\Big\}\\ &= \lim_{n\to\infty}\mathbb{P}\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/n) \subseteq \mathcal{X}^c_\delta\},\end{aligned}$$

from which there exists a $\bar{n}$ such that

$$\mathbb{P}\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/\bar{n}) \subseteq \mathcal{X}^c_\delta\} > \epsilon. \tag{24}$$

Let us now heat the constraints $\bar{\delta}^{(1)},\dots,\bar{\delta}^{(N)}$ up by translation parameters $z^{(1)}_k,\dots,z^{(N)}_k \in \mathcal{B}_{\rho_k}$ and ask the following question: is it true that the heated multiextraction $((\bar{\delta}^{(1)},z^{(1)}_k),\dots,(\bar{\delta}^{(N)},z^{(N)}_k))$ is bad for HUP with heating parameter ρ_k? It turns out that the answer is positive for k large enough, a fact that is proven next.

Recall that x'^*_N is the solution with constraints $(\bar{\delta}^{(1)},z^{(1)}_k),\dots,(\bar{\delta}^{(N)},z^{(N)}_k)$, and define $d_k := \sup_{z^{(1)}_k,\dots,z^{(N)}_k \in \mathcal{B}_{\rho_k}}\|x'^*_N - x^*_N\|$ which, by (18), goes to 0 as $k \to \infty$. Pick a $\bar{k}$ such that

$$d_k + \rho_k < 1/\bar{n} \quad \forall k \geq \bar{k}.$$

All heated solutions x'^*_N are apart from x^*_N by at most d_k, and all heated constraints $(\delta,z) \in \Delta\times\mathcal{B}_{\rho_k}$ are apart from the corresponding unheated constraint δ by at most ρ_k. Thus, if $k \geq \bar{k}$, all heated versions of a constraint δ in the set $\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/\bar{n}) \subseteq \mathcal{X}^c_\delta\}$ on the left-hand side of (24) are violated by x'^*_N. That is,

$$\{\delta \in \Delta :\ \mathcal{B}(x^*_N,1/\bar{n}) \subseteq \mathcal{X}^c_\delta\}\times\mathcal{B}_{\rho_k} \subseteq \{(\delta,z) \in \Delta\times\mathcal{B}_{\rho_k} :\ x'^*_N \notin [\mathcal{X}_\delta+z]\} \quad \forall k \geq \bar{k}. \tag{25}$$

Then, for any $z_k^{(1)}, \dots, z_k^{(N)} \in \mathcal{B}_{\rho_k}$ and for any $k \geq \bar{k}$, we have

$$\begin{aligned} V'(x'^*_N) &= \mathbb{P}'\{(\delta, z) \in \Delta \times \mathcal{B}_{\rho_k} : \ x'^*_N \notin [\mathcal{X}_\delta + z]\} \\ &\geq [\text{using (25)}] \\ &\geq \mathbb{P}'\Big\{\{\delta \in \Delta : \ \mathcal{B}(x^*_N, 1/\bar{n}) \subseteq \mathcal{X}^c_\delta\} \times \mathcal{B}_{\rho_k}\Big\} \\ &= [\text{recalling that } \mathbb{P}' = \mathbb{P} \times \mathbb{U}] \\ &= \mathbb{P}\{\delta \in \Delta : \ \mathcal{B}(x^*_N, 1/\bar{n}) \subseteq \mathcal{X}^c_\delta\} \cdot \mathbb{U}\{\mathcal{B}_{\rho_k}\} \\ &> [\text{since } \mathbb{U}\{\mathcal{B}_{\rho_k}\} = 1 \text{ and using (24)}] \\ &> \epsilon, \end{aligned}$$

i.e., $((\bar{\delta}^{(1)}, z_k^{(1)}), \dots, (\bar{\delta}^{(N)}, z_k^{(N)}))$ is bad for HUP with heating parameter ρ_k for any $z_k^{(1)}, \dots, z_k^{(N)} \in \mathcal{B}_{\rho_k}$ when $k \geq \bar{k}$. In turn, this entails that

$$\int_{\mathcal{B}^N_{\rho_k}} \mathbb{I}_{\{V'(x'^*_N) > \epsilon\}} \frac{\mathrm{d}z^N}{\mathrm{Vol}(\mathcal{B}^N_{\rho_k})} = 1 \quad \forall k \geq \bar{k}. \tag{26}$$

Finally,

$$\begin{aligned} &\sum_{i=0}^{d-1} \binom{N}{i} \epsilon^i (1-\epsilon)^{N-i} \\ &\quad \geq \quad [\text{using (17)}] \\ &\quad \geq \quad (\mathbb{P}')^N \{V'(x'^*_N) > \epsilon\} \\ &\quad = \quad \int_{\Delta^N} \left[\int_{\mathcal{B}^N_{\rho_k}} \mathbb{I}_{\{V'(x'^*_N) > \epsilon\}} \frac{\mathrm{d}z^N}{\mathrm{Vol}(\mathcal{B}^N_{\rho_k})} \right] \mathbb{P}^N(\mathrm{d}\delta^N) \\ &\quad \geq \quad \int_{\{V(x^*_N) > \epsilon\}} \left[\int_{\mathcal{B}^N_{\rho_k}} \mathbb{I}_{\{V'(x'^*_N) > \epsilon\}} \frac{\mathrm{d}z^N}{\mathrm{Vol}(\mathcal{B}^N_{\rho_k})} \right] \mathbb{P}^N(\mathrm{d}\delta^N) \\ &\overrightarrow{k \to \infty} [\text{recalling (26) and by the dominated convergence theorem [26]}] \\ &\overrightarrow{k \to \infty} \int_{\{V(x^*_N) > \epsilon\}} \mathbb{P}^N(\mathrm{d}\delta^N) \\ &\quad = \quad \mathbb{P}^N\{V(x^*_N) > \epsilon\}. \end{aligned}$$

This concludes the proof. ∎

Acknowledgments. The authors would like to thank Marco Dalai for a careful reading of an earlier version of this paper and an anonymous referee for many suggestions that helped improve the paper.

REFERENCES

[1] A. Ben-Tal and A. Nemirovski, *Robust truss topology design via semidefinite programming*, SIAM J. Optim., 7 (1997), pp. 991–1016.
[2] A. Ben-Tal and A. Nemirovski, *Robust convex optimization*, Math. Oper. Res., 23 (1998), pp. 769–805.
[3] A. Ben-Tal and A. Nemirovski, *Robust solutions of uncertain linear programs*, Oper. Res. Lett., 25 (1999), pp. 1–13.
[4] A. Ben-Tal, A. Nemirovski, and C. Roos, *Robust solutions of uncertain quadratic and conic-quadratic problems*, SIAM J. Optim., 13 (2002), pp. 535–560.

[5] G. Calafiore and M. C. Campi, *Uncertain convex programs: Randomized solutions and confidence levels*, Math. Program., 102 (2005), pp. 25–46.
[6] G. Calafiore and M. C. Campi, *The scenario approach to robust control design*, IEEE Trans. Automat. Control, 51 (2006), pp. 742–753.
[7] G. Calafiore and F. Dabbene, *A probabilistic framework for problems with real structured uncertainty in systems and control*, Automatica, 38 (2002), pp. 1265–1276.
[8] G. Calafiore, F. Dabbene, and R. Tempo, *Randomized algorithms for probabilistic robustness with real and complex structured uncertainty*, IEEE Trans. Automat. Control, 45 (2000), pp. 2218–2235.
[9] G. Calafiore and B. Polyak, *Stochastic algorithms for exact and approximate feasibility of robust LMIs*, IEEE Trans. Automat. Control, 46 (2001), pp. 1755–1759.
[10] A. Charnes and W. W. Cooper, *Chance constrained programming*, Management Sci., 6 (1959), pp. 73–79.
[11] D. P. de Farias and B. Van Roy, *On constraint sampling in the linear programming approach to approximate dynamic programming*, Math. Oper. Res., 29 (2004), pp. 462–478.
[12] D. Dentcheva, *Optimization models with probabilistic constraints*, in Probabilistic and Randomized Methods for Design under Uncertainty, G. Calafiore and F. Dabbene, eds., Springer, London, 2005, pp. 47–96.
[13] E. Erdogan and G. Iyengar, *Ambiguous chance constrained problems and robust optimization*, Math. Program. Ser. B, 107 (2006), pp. 37–61.
[14] L. El Ghaoui and G. Calafiore, *Robust filtering for discrete-time systems with bounded noise and parametric uncertainty*, IEEE Trans. Automat. Control, 46 (2001), pp. 1084–1089.
[15] L. El Ghaoui and H. Lebret, *Robust solutions to least-squares problems with uncertain data*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 1035–1064.
[16] L. El Ghaoui, F. Oustry, and H. Lebret, *Robust solutions to uncertain semidefinite programs*, SIAM J. Optim., 9 (1998), pp. 33–52.
[17] L. El Ghaoui and S. I. Niculescu, *Robust decision problems in engineering: A linear matrix inequality approach*, in Advances in Linear Matrix Inequality Methods in Control, L. El Ghaoui and S. I. Niculescu, eds., SIAM, Philadelphia, 1999.
[18] V. L. Levin, *Application of E. Helly's theorem to convex programming, problems of best approximation and related questions*, Sb. Math., 8 (1969), pp. 235–247.
[19] R. Lucchetti, *Convexity and Well-posed Problems*, CMS Books Math., Springer, New York, 2006.
[20] A. Nemirovski and A. Shapiro, *Scenario approximations of chance constraints*, in Probabilistic and Randomized Methods for Design under Uncertainty, G. Calafiore and F. Dabbene, eds., Springer, London, 2005, pp. 3–48.
[21] A. Prékopa, *Stochastic Programming*, Kluwer, Boston, 1995.
[22] A. Prékopa, *Probabilistic programming*, in Stochastic Programming, A. Ruszczyǹski and A. Shapiro, eds., Handbooks Oper. Res. Managment Sci. 10, Elsevier, London, UK, 2003, pp. 267–352.
[23] W. Rudin, *Real and Complex Analysis*, 3rd ed., McGraw–Hill, New York, 1987.
[24] A. Ruszczyǹski and A. Shapiro, *Stochasting programming models*, in Stochastic Programming, A. Ruszczyǹski and A. Shapiro, eds., Handbooks Oper. Res. Managment Sci. 10, Elsevier, London, UK, 2003, pp. 1–64.
[25] A. Ruszczyǹski and A. Shapiro, *Lectures on Stochastic Programming*, preprint, 2008.
[26] A. N. Shiryaev, *Probability*, Springer, New York, 1996.
[27] R. L. Smith, *Efficients Monte-Carlo procedures for generating points uniformly distributed over bounded regions*, Oper. Res., 32 (1984), pp. 1296–1308.