Comput Manag Sci (2017) 14:443–460
DOI 10.1007/s10287-017-0284-7

ORIGINAL PAPER

# A joint model of probabilistic/robust constraints for gas transport management in stationary networks

**T. González Grandón[1] · H. Heitsch[2] · R. Henrion[2]**

Received: 11 May 2017 / Accepted: 15 May 2017 / Published online: 5 June 2017


**Abstract** We present a novel mathematical algorithm to assist gas network operators in managing uncertainty, while increasing reliability of transmission and supply. As a result, we solve an optimization problem with a joint probabilistic constraint over an infinite system of random inequalities. Such models arise in the presence of uncertain parameters having partially stochastic and partially non-stochastic character. The application that drives this new approach is a stationary network with uncertain demand (which are stochastic due to the possibility of fitting statistical distributions based on historical measurements) and with uncertain roughness coefficients in the pipes (which are uncertain but non-stochastic due to a lack of attainable measurements). We study the sensitivity of local uncertainties in the roughness coefficients and their impact on a highly reliable network operation. In particular, we are going to answer the question, what is the maximum uncertainty that is allowed (shaping a 'maximal' uncertainty set) around nominal roughness coefficients, such that random demands in a stationary gas network can be satisfied at given high probability level for no matter which realization of true roughness coefficients within the uncertainty set. One ends up with a constraint, which is probabilistic with respect to the load of gas and robust with respect to the roughness coefficients. We demonstrate how such constraints

✉ R. Henrion
henrion@wias-berlin.de

T. González Grandón
gonzalet@hu-berlin.de

H. Heitsch
heitsch@wias-berlin.de

[1] Humboldt University Berlin, Rudower Chausse 25, 12489 Berlin, Germany

[2] Weierstrass Institute for Applied Analysis and Stochastics, Mohrenstrae 39, 10117 Berlin, Germany

can be dealt with in the framework of the so-called spheric-radial decomposition of multivariate Gaussian distributions. The numerical solution of a corresponding optimization problem is illustrated. The results might assist the network operator with the implementation of cost-intensive roughness measurements.

**Keywords** Chance constraint · Robust constraint · Uncertainty set · Spheric-radial decomposition

## 1 Introduction

### 1.1 Background

In the context of the liberalization paradigm, regulatory authorities have separated the natural gas transmission from production and services. Accordingly, the network operators are solely responsible for the transportation of gas, and gas traders only need to specify or nominate where they want to inject gas, at so-called entry points, or extract gas (loads), at so-called exit points. As a consequence, new mathematical challenges for the gas network operators have been introduced.

Presently, the reliability of the gas network operator depends on the accuracy of calculating the transport capacity and on the security of supply. This concern is called *nomination validation*, i.e., determine whether the given nominations of all entry and exit flows are technically and physical feasible under the available infrastructure (see Koch et al. 2015). This challenge is further complicated by the uncertainty of two parameters in the feasibility check: the coverage of future load and the pipes' roughness coefficients. When ensuring security of gas supply for end consumers, network operators have to quantify the coverage of uncertain future loads. The amount of gas that enters the network depends on volatile prices, and the amount of gas that exits is influenced by ambient temperature changes. Nevertheless, it is possible to model the amount of future load by means of a stochastic distribution based on historical data. Moreover, the roughness coefficient of a pipe influences the pressure loss of the gas flowing. However, the exact value of the roughness coefficients are only known at the time of commissioning of the network, but they change significantly through time due to the aging of pipe material and to deposition of particles. The traditional approach to estimate the transport capacity, is to set fixed values for the roughness coefficients based on the well-known Colebrook–White correlation (Colebrook 1939). Nonetheless, this approach underestimates the uncertainty of roughness coefficients, hence there is a miscalculation of the transport capacity.

In the research literature, there is a more in-depth study of nomination validation in Pfetsch et al. (2015). The robustness of natural gas flows is examined in Chertkov et al. (2015) and Gotzes et al. (2016) gives a explicit characterization of gas flow feasibility and considers the stochastic nature of exit loads. The present paper develops a novel algorithm to improve the nomination validation procedure, by taking into account these two types of uncertain parameters. The presence of non-stochastic (roughness coefficients) and stochastic (load) uncertainty motivates us to establish a joint robust/probabilistic constraint. As a result, we investigate the maximum of attain-

able uncertainty around the roughness coefficients for different pipes in the network, while keeping a high probability to satisfy the demands.

### 1.2 Proposed model: optimization under uncertainty

Data uncertainty prevails in many real world optimization models where it typically enters the inequality constraints describing the set of feasible decisions:

$$g_i(x, z) \geq 0 \quad (i = 1, \ldots, k)\,. \tag{1}$$

Here $x \in \mathbb{R}^n$ is a decision vector, $z \in \mathbb{R}^m$ is an uncertain parameter and $g : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^k$ refers to a constraint mapping. Overlooking the aspect of uncertainty would result in optimal decisions, which are notoriously non-robust with respect to deviations from the assumed deterministic data. When modeling uncertainty, two situations typically occur: in the first one, access to historical observations is given such that uncertainty can be modeled by means of a random vector $\xi$ obeying a certain estimated multivariate distribution. This allows one to turn (1) into a so-called (joint) probabilistic constraint:

$$\mathbb{P}\big(g(x, \xi) \geq 0\big) \geq p \in (0, 1) \tag{2}$$

(note that the first '$\geq$' sign is to be understood component-wise). The meaning of (2) is as follows: a decision $x$ is declared to be feasible if and only if the original random inequality system (1) is satisfied with at least probability $p$, a level usually chosen close to but not identical to one in order to guarantee sufficient robustness without excessive costs. For a standard reference on probabilistic constraints we refer to the monograph by Prékopa (1995).

A second situation arises when some uncertain parameter $\Phi$ is, for instance, fixed but simply unknown or it is random but its distribution is unknown because it cannot be observed/measured. In such case, a model is introduced that ensures feasibility of its solution for all possible realizations (equivalently: for the worst case scenario) of the uncertain parameter within some uncertainty set $\mathscr{U} \subseteq \mathbb{R}^s$. Then, (1) turns into a constraint with infinitely many inequalities

$$g(x, \Phi) \geq 0 \quad \forall \Phi \in \mathscr{U} \tag{3}$$

which can be equivalently written as a single worst case constraint

$$\inf_{\Phi \in \mathscr{U}} g(x, \Phi) \geq 0. \tag{4}$$

Clearly, if $g$ happens to be concave in the first variable, then the set of feasible decisions $x$ in (3) or (4) is convex, which is a favorable property of optimization problems. Note that both, the probabilistic and the worst case constraints (2), (3), (4) just depend on the decision vector $x$ but unlike (1) no longer on the concrete outcome of uncertainty. Hence, they can figure as ordinary well-defined constraints in optimization problems with some additional objective given. One then speaks of *probabilistic programming*

(in case of (2)) or of *robust optimization* (in case of (3)). For a basic monograph on robust optimization, we refer to Ben-Tal et al. (2009).

A crucial step in robust optimization is the appropriate choice of the uncertainty set $\mathscr{U}$. Simple-shaped sets like polyhedra or ellipsoids lead to problems with favorable complexity which allow one to deal with much larger dimensions than in the case of probabilistic constraints. That is why robust optimization is not only employed in the absence of statistical information but also as a nicer-behaved substitute of probabilistic constraints in a probabilistic setting. This is usually done by choosing an uncertainty set of desired probability $p$ in (2): $\mathbb{P}(\xi \in \mathscr{U}) = p$. This can be easily arranged, for instance, in the case of multivariate Gaussian distributions and ellipsoids of appropriate size. In such case the set of feasible decisions $x$ defined by (3) or (4) (with $\Phi$ replaced by $\xi$) is always contained in the set of feasible decisions defined by (2), such that robust optimization provides feasible yet conservative solutions to probabilistic programming. The conservatism, however, may be considerable up to the point of ending up at very small or even empty feasible sets possibly coming at much higher costs than under a probabilistic constraints. This effect motivates the consideration of probabilistic constraints in the presence of statistical information at least in moderate dimension.

Traditionally both approaches, chance constraint and robust optimization have been dealt with separately. Very often, however, one is faced with uncertain variables of both mentioned types. An example is the transportation management of gas networks. Here, the uncertain demands (loads) at the nodes of the network are usually provided with a rich historical data record allowing one to approximate their joint (multivariate) distribution and, thus, to model them by a random vector $\xi$. In contrast, the transport is influenced by roughness coefficients along the pipes for which there usually just exists some imprecise estimation but no concrete measurement. This leads us naturally to the consideration of uncertain inequalities (2) in which the uncertain variable has a stochastic and a non-stochastic part, i.e., $z = (\xi, \Phi)$. As a consequence, the originally separate models (3) and (4) have to be combined appropriately. There are essentially two different ways of doing so:

On the one hand, we formulate a probabilistic constraint (w.r.t. $\xi$) involving a robustified (w.r.t. $\Phi$) uncertain inequality system:

$$\mathbb{P}\big(g(x, \xi, \Phi) \geq 0 \quad \forall \Phi \in \mathscr{U}\big) \geq p. \tag{5}$$

A second possibility consists in modeling a family of probabilistic constraints first and then establishing a robustified version of these:

$$\mathbb{P}\big(g(x, \xi, \Phi) \geq 0\big) \geq p \quad \forall \Phi \in \mathscr{U}. \tag{6}$$

In analogy with finite random inequality systems, one may refer to (5) as a joint probabilistic constraint (over an infinite number of random inequalities) and to (6) as a system of individual probabilistic constraints (over a single random inequality each). Models (5) and (6) have been introduced and considered before in Van Ackooij et al. (2016) under the name of *hybrid robust/chance-constraint*. It has been observed there (and is easy to see) that (6) is a weaker constraint than (5). In this paper, we shall

focus on model (5) which takes the outer perspective of probabilistic programming. The alternative model (6) appears, for instance, in the context of so-called stochastic dominance constraints (see Dentcheva and Ruszczyński 2003). Both models have been considered in Van Ackooij et al. (2016) in the context of linear probabilistic constraints under discrete distributions. Our perspective is different in allowing for nonlinear probabilistic constraints under the (continuous) Gaussian distributions using the so-called spheric-radial decomposition. Moreover, the uncertainty set $\mathscr{U}$ will not be fixed in our problem but subject to optimization. The model will be illustrated for the example of demand satisfaction in a stationary gas network.

## 2 Description of the optimization problem

In this section, we present an optimization problem, from the viewpoint of the network operator, who needs to validate nominations under stochastic and non-stochastic uncertainty. We shall assume that the network is in steady state and is passive (i.e., does not contain active elements such as compressors, valves etc.). These simplifications allow us to maintain a purely algebraic model without combinatorial aspects. Moreover, we will focus our attention to the special case of a tree structured gas network $G = (\mathscr{V}, \mathscr{E})$, where $\mathscr{V}$ is the set of nodes representing interconnection points, and $\mathscr{E}$ is the set of arcs representing the pipelines. For including cycles into the consideration we refer to the recent work (Gotzes et al. 2017). For simplicity, we shall also assume that there is just one single injection node (entry), labeled zero, whereas there are $m$ additional nodes at which gas is withdrawn for consumption (exits). The unique entry will be declared as the root of the tree while the arcs in $\mathscr{E}$ are directed away from the root.

We will refer to $p, \hat{b}$ as the vectors of pressures and loads, respectively, at the nodes, and to $q$ as the vector of flows through the arcs of the network. It is well known that the following relations have to be met in order to satisfy the load $\hat{b}$ (being positive in the case of withdrawal and negative in the case of injection):

$$Aq = \hat{b} \tag{7}$$

$$A^T p^2 = (\operatorname{diag} \Phi)|q|q \tag{8}$$

$$p^{\min} \leq p \leq p^{\max} \tag{9}$$

Here, $A$ refers to the incidence matrix of the network, $\operatorname{diag} \Phi$ is a diagonal matrix whose diagonal is given by a vector $\Phi = (\Phi_e)_{e \in \mathscr{E}}$ of roughness coefficients for the pipes (arcs) and $p^{\min}$, $p^{\max}$ are vectors of lower and upper bounds, respectively, for the pressure at the given nodes. Functions of vectors, such as $(\cdot)^2$ or $|\cdot|(\cdot)$ are to be understood as vectors with entries defined componentwise. Equations (7), (8) refer to the first and second Kirchhoff laws (mass flow conservation and pressure drop equations), whereas the bound constraints (9) result from technological and contractual restrictions.

### 2.1 Explicit representation of feasible loads

The operator of a gas network is obliged to ensure the feasibility of loads $b$ for a given network environment (technical equipment, pressure bounds, booked capacities for injecting or withdrawing gas). If he wishes to change this environment in a systematic way, e.g., by solving some optimization problem, then it is important to have access to an explicit description of feasibility in terms of the boundary data without further dependence on physical state variables $p, q$. Such description was provided in Gotzes et al. (2016). The following result provides a description of the feasibility of load vectors $b$ by means of an explicit inequality system:

**Proposition 2.1** (See Gotzes et al. 2016, Cor. 1) *Consider a tree $G = (\mathscr{V}, \mathscr{E})$ with the unique entry referring to the root node zero. Then, the load vector $\hat{b} := (b_0, b)$ with $b_0$ being the load at the entry and b being the vector of withdrawals at exits is feasible in the sense of Eqs. (7), (8), (9) if and only if the following system of inequalities is satisfied for $k, \ell = 0, \ldots, m,\ k \neq \ell$:*

$$h_k(b, \Phi) + \left(p_k^{\max}\right)^2 - h_\ell(b, \Phi) - \left(p_\ell^{\min}\right)^2 \geq 0, \tag{10}$$

*where, for $k = 0, \ldots, m$ we have put:*

$$h_k(b, \Phi) := \begin{cases} \sum\limits_{e \in \Pi(k)} \Phi_e \left( \sum\limits_{t \geq h(e)} b_t \right)^2 & if\ k \geq 1 \\ 0 & if\ k = 0 \end{cases}.$$

*Here, for $k, \ell \in \mathscr{V}$, denote $k \succeq l$ if, in $G$, the unique directed path from the root to $k$, denoted $\Pi(k)$, passes through $l$. Moreover, $h(e)$ refers to the head of arc $e \in \mathscr{E}$.*

The fact that the feasibility of the original enhanced vector $\hat{b}$ is already determined by the feasibility of its exit part $b$ in the inequality system above follows from the total load (sum of all components) always being zero. Hence, in the following we will only speak of the feasibility of the load vector at exits from the very beginning.

We emphasize that both variables of the function $h_k$ occurring in the inequality system (10) are uncertain, of different nature, however: the load vector $b$ is of stochastic nature because historical data on its realization are available. In contrast, the roughness coefficients $\Phi_e$ may be endowed only by some nominal values, from which they differ in an unobservable way, for instance, by the aging of pipes' material underground. Our aim is to consider both types of uncertainty in a joint probabilistic/robust model as detailed in the introduction.

To set up this model we start by emphasizing the variability of roughness coefficients in the definition of feasible loads via (10): for any given vector $\Phi$ of roughness coefficients, we denote the set of feasible loads $b$ (satisfying (10)) by $M_\Phi$. In order to take into account the uncertain (non-stochastic) nature of $\Phi$, we assume that it belongs to some uncertainty set $\mathscr{U}_\delta$ which is parameterized by a vector $\delta \in \mathbb{R}^{|\mathscr{E}|}$. For instance, $\mathscr{U}_\delta$ could be a rectangle with side lengts $\delta_e$ or an ellipsoid with principal axes $\delta_e$ around some nominal vector $\hat{\Phi}$ of roughness coefficients. We strengthen now the definition

of feasibility of a load vector $b$ w.r.t. uncertainty in $\Phi$ by requiring that it satisfies (10) **for all** $\Phi$ in the uncertainty set. This means, for each fixed shape parameter $\delta$, the set of feasible loads $b$ is given by

$$M_\delta := \bigcap_{\Phi \in \mathscr{U}_\delta} M_\Phi \quad (\delta \in \mathbb{R}^{|\mathscr{E}|}). \tag{11}$$

Often, the choice of the uncertainty set, in particular of its size, is somehow arbitrary. This motivates us, in this paper, not to fix it but to consider this set as variable (via the parameter $\delta$) and subject to optimization.

On the other hand, in order to address the stochastic uncertainty of the load vector $b$ we will understand it from now on as the outcome of an $m$-dimensional random vector $\xi$, where $m = \#\mathscr{V} - 1$ is the number of exits. In particular, we will assume $\xi$ to be a Gaussian random vector $\xi \sim \mathscr{N}(\mu, \Sigma)$ with mean $\mu$ and covariance matrix $\Sigma$. In a strict sense, exit loads cannot follow a Gaussian distribution because the latter allows negative values. This effect, however, is negligible if the relative standard deviations are small. Moreover, the methodology we are presenting here is easily adapted to Gaussian-like distributions (truncated Gaussian, Log-normal) which are good candidates for modeling stochastic exit loads (Koch et al. 2015, Chapter 13).

Accordingly, we turn the original inclusion $b \in M_\delta$ from (11) into the probabilistic constraint $\mathbb{P}(\xi \in M_\delta) \geq p$, where $p \in (0, 1)$ is a probability level (close to one) chosen in the modeling process. Observe that this inequality now is a constraint on the shape parameter $\delta$ for the uncertainty set $U_\delta$. Since we consider $\delta$ as a variable, it defines a set of feasible shape parameters as

$$M := \left\{\delta \in \mathbb{R}^{|\mathscr{E}|} \,\middle|\, \mathbb{P}(\xi \in M_\delta) \geq p\right\}. \tag{12}$$

The interpretation of this constraint is as follows: a shape parameter $\delta$ is feasible if and only if the probability of random load vectors $b$ being admissible in the sense of (10) **for all** roughness coefficients $\Phi \in U_\delta$ is at least $p$. Clearly, an increase of $\delta$ and, hence of the uncertainty set $U_\delta$ will result in a stronger 'for all' condition and, thus, in a decrease of the probability of the event $\xi \in M_\delta$.

With this setup, we are interested in a maximum amount of uncertainty for the roughness coefficients which still allows us to technically satisfy - under the given pressure bounds - the random loads for all uncertain roughness coefficients with a given probability $p$. This leads us to the optimization problem

$$\max\{f(\delta) \mid \delta \in M\}, \tag{13}$$

where the objective $f$ characterizes the amount of uncertainty, e.g., $f(\delta) := \operatorname{vol} U_\delta$ or $f(\delta) := \|\delta\|$. The solution of (13) may provide the network operator with an idea at what precision at least he needs to know the roughness coefficient in the context of safe network operation. Such information could be used, for instance, when trying to roughly estimate these coefficients by indirect measurements via the solution of an inverse problem (see Egger et al. 2017) or in order to identify critical parts of the network where it is more important to do so than in other parts.

## 3 Determination of the probability of $M_\delta$

The key for solving the optimization problem (13) is clearly the verification of the inclusion $\delta \in M$, which according to (12) amounts to the computation of the probability of the event $\xi \in M_\delta$. This task faces two difficulties which we will address in this section: first, the set $M_\delta$, given as an infinite intersection of sets $M_\Phi$ has to be made explicit and second, the Gaussian probability of such a set has to be determined efficiently.

### 3.1 Efficient description of the set $M_\delta$

By definition of $M_\Phi$ as consisting of all load vectors $b$ satisfying (10), we may write $M_\delta$ in (11) as

$$\begin{aligned} M_\delta = \big\{ & b \mid \\ & h_k(b,\Phi) + \left(p_k^{\max}\right)^2 - h_\ell(b,\Phi) - \left(p_\ell^{\min}\right)^2 \geq 0 \\ & \forall \Phi \in U_\delta \ \forall k,\ell = 0,\ldots,m \ (k \neq \ell) \big\}. \end{aligned}$$

The difficulty in working with the set $M_\delta$ consists in the fact that it is defined by infinitely many constraints, because the uncertainty set is infinite in general. Evidently, we have the following equivalent reformulation:

$$\begin{aligned} M_\delta = \big\{ & b \mid \\ & \inf_{\Phi \in U_\delta} h_k(b,\Phi) + \left(p_k^{\max}\right)^2 - h_\ell(b,\Phi) - \left(p_\ell^{\min}\right)^2 \geq 0 \\ & \forall k,\ell = 0,\ldots,m \ (k \neq \ell) \big\}. \end{aligned}$$

In the typical case of compact uncertainty sets $U_\delta$, we may finally represent $M_\delta$ as

$$\begin{aligned} M_\delta = \big\{ & b \mid \\ & h_k(b,\Phi^*_{k\ell}(\delta,b)) + \left(p_k^{\max}\right)^2 - h_\ell(b,\Phi^*_{k\ell}(\delta,b)) - \left(p_\ell^{\min}\right)^2 \geq 0 \\ & \forall k,\ell = 0,\ldots,m \ (k \neq \ell) \big\}, \end{aligned} \tag{14}$$

where

$$\Phi^*_{k\ell}(\delta,b) := \operatorname*{argmin}_{\Phi \in U_\delta} \left\{ h_k(b,\Phi) - h_\ell(b,\Phi) \right\}. \tag{15}$$

This last representation (14) has the advantage of involving only a finite number of inequalities in contrast to the original description. However, it comes at the price of solving the optimization problem (15) for each $b$ and $\delta$. Fortunately, this is easily done for specific simple enough uncertainty sets, for instance, rectangles or ellipsoids. This is mainly a consequence of the difference function $h_k - h_\ell$ appearing in (15) being linear in $\Phi$ by definition of $h_k$ in Proposition 2.1. Minimizing a linear function over a rectangle or ellipsoid can be done explicitly in terms of the coefficients of this linear function.

### 3.2 Explicit solutions for simple uncertainty sets

We will provide in the following explicit solutions for the functions $\Phi^*_{k\ell}(\delta, b)$ introduced in (15) in case of ellipsoidal and rectangular uncertainty sets. We will assume that our uncertainty sets are symmetric around some nominal (or guessed) value $\hat{\Phi}$ for the roughness coefficients. In the case of an ellipsoid, we define for $\delta \in \mathbb{R}^{|\mathscr{E}|}_+$:

$$U_\delta := \left\{\Phi \in \mathbb{R}^{|\mathscr{E}|} \middle| (\Phi - \hat{\Phi})^T \Sigma_\delta (\Phi - \hat{\Phi}) \leq 1\right\}, \tag{16}$$

where $\Sigma_\delta$ is a diagonal matrix with entries $\delta_1, \ldots, \delta_{|\mathscr{E}|}$.

In the following, referring back to the notation introduced in Proposition 2.1, we will use the assignments

$$\mathbf{1}_{e,k} := \begin{cases} 1 & \text{if } e \in \Pi(k) \\ 0 & \text{otherwise} \end{cases}$$

and

$$\gamma_e(b) := \sum_{t \in \mathscr{V},\, t \succeq h(e)} b_t \,. \tag{17}$$

This notation will allow us to rewrite the definition for $h_k(b, \Phi)$ in Proposition 2.1 as

$$h_k(b, \Phi) = \sum_{e \in \mathscr{E}} \Phi_e \gamma_e^2(b) \mathbf{1}_{e,k} \quad (k = 1, \ldots, m). \tag{18}$$

**Lemma 3.1** *For the uncertainty set* (16) *the functions introduced in* (15) *have the following explicit representation for* $k, \ell = 0, \ldots, m$ *with* $k \neq l$ *and* $e \in \mathscr{E}$*:*

$$[\Phi^*_{k\ell}(\delta, b)]_e = \hat{\Phi}_e + \frac{(\mathbf{1}_{e,\ell} - \mathbf{1}_{e,k})\gamma_e^2(b)/\delta_e}{\sqrt{\sum_{s \in (\Pi(k) \cup \Pi(l)) \setminus (\Pi(k) \cap \Pi(l))} \gamma_s^4(b)/\delta_s}}.$$

*Proof* It is well-known - and easy to show by writing down the necessary optimality conditions - that for a given cost vector $c \neq 0$, the optimization problem

$$\min_{\Phi} \left\{c^T \Phi | (\Phi - \hat{\Phi})^T \Sigma_\delta (\Phi - \hat{\Phi}) \leq 1\right\} \tag{19}$$

has the unique solution

$$\Phi^* = \hat{\Phi} - \frac{1}{\sqrt{c^T \Sigma_\delta^{-1} c}} \Sigma_\delta^{-1} c. \tag{20}$$

Clearly, the optimization problem defining $\Phi^*_{k\ell}(\delta, b)$ in (15) has the form of (19) with the coefficients of the cost vector given by (according to (18))

$$c_e := (\mathbf{1}_{e,k} - \mathbf{1}_{e,\ell})\gamma_e^2(b) \quad (e = 1, \ldots, |\mathscr{E}|).$$

Taking into account that $\Sigma_\delta$ is a diagonal matrix with diagonal entries $\delta_e$, the optimal solution $\Phi^*_{k\ell}(\delta, b)$ can be read off component-wise from (20) to yield the asserted formula. □

Alternatively to (16), we now introduce a family of rectangular uncertainty sets, again centered around some nominal value $\hat{\Phi}$ by means of

$$U_\delta := [\hat{\Phi} - \delta, \hat{\Phi} + \delta] \quad (\delta \in \mathbb{R}_+^{|\mathscr{E}|}). \tag{21}$$

**Lemma 3.2** *For the uncertainty set* (21) *the functions introduced in* (15) *are actually independent of b and have the following explicit representation for* $k, \ell = 0, \ldots, m$ *with* $k \neq l$ *and* $e \in \mathscr{E}$*:*

$$[\Phi^*_{k\ell}(\delta, b)]_e = \begin{cases} \hat{\Phi}_e - \delta_e & \text{if } e \in \Pi(k) \backslash \Pi(l) \\ \hat{\Phi}_e + \delta_e & \text{if } e \in \Pi(l) \backslash \Pi(k) \\ \hat{\Phi}_e & \text{otherwise} \end{cases}.$$

*Proof* We observe from (18) and (17) that the following holds true for $k, \ell = 0, \ldots, m$ with $k \neq l$ and $e \in \mathscr{E}$:

$$\min_{\Phi \in \mathscr{U}_\delta} \{h_k(b, \Phi) - h_\ell(b, \Phi)\} = \min_{\Phi \in \mathscr{U}_\delta} \sum_{e \in \mathscr{E}} (\mathbf{1}_{e,k} - \mathbf{1}_{e,\ell}) \Phi_e \gamma_e^2(b).$$

With $\mathscr{U}_\delta$ being a rectangle and with the sum above being separable in the components $\Phi_e$, the minimization can be carried out component-wise. Accordingly, each component $\Phi_e$ is chosen in the interval $[\hat{\Phi}_e - \delta_e, \hat{\Phi}_e + \delta_e]$ as to minimize the expression

$$(\mathbf{1}_{e,k} - \mathbf{1}_{e,\ell}) \Phi_e \gamma_e^2(b).$$

Since the coefficients $\gamma_e^2(b)$ are non-negative, and $(\mathbf{1}_{e,k} - \mathbf{1}_{e,\ell})$ equals $\pm 1$ (or zero in which case the choice of $\Phi_e$ is arbitrary), we can choose the minimizing component as

$$[\Phi^*_{k\ell}]_e := \hat{\Phi}_e - (\mathbf{1}_{e,k} - \mathbf{1}_{e,\ell}) \delta_e.$$

By definition, this reduces to the formula asserted in the statement of our Lemma. □

### 3.3 An algorithm for computing the probability of feasible random exit loads

Now, that we are given an explicit description of the set $M_\delta$ in (14) for the special case of elliptical or rectangular uncertainty sets (upon substituting the functions $\Phi^*_{k\ell}(\delta, b)$ by the formulae obtained in Lemmas 3.1 and 3.2), we could use this finite inequality system in order to test the feasibility of simulated outcomes of the random load $b$ according to the given Gaussian distribution. The averaged number of feasible simulations would yield the Monte Carlo estimate for the desired probability $\mathbb{P}(\xi \in M_\delta)$.

Such Monte Carlo approach has two drawbacks: first it may come with a comparatively large variance for the obtained probability estimation and, second, it does not provide us with information about the sensitivity of this probability with respect to changes of $\delta$. This sensitivity (derivative) information is crucial, however, in order to set up any efficient algorithm of nonlinear optimization in order to solve problem (13). Therefore, we will alternatively make use of the so-called *spheric-radial decomposition of Gaussian random vectors* (see, e.g. Déak 2000; Genz and Bretz 2009):

**Theorem 3.1** *Let $\xi$ be an $m$-dimensional Gaussian random vector distributed according to $\xi \sim \mathcal{N}(\mu, \Sigma)$. Then for any Borel measurable subset $A \subseteq \mathbb{R}^m$ it holds that*

$$\mathbb{P}(\xi \in A) = \int_{v \in \mathbb{S}^{m-1}} \mu_\chi\{r \geq 0 \,|\, (rLv + \mu) \cap A \neq \emptyset\} d\mu_\eta$$

*with $\xi = \chi L\eta$, where $L$ is such that $\Sigma = LL^T$ (e.g., Cholesky decomposition), $\chi$ has a chi-distribution $\mu_\chi$ with $m$ degrees of freedom and $\eta$ has a uniform distribution $\mu_\eta$ over the Euclidean unit sphere $\mathbb{S}^{m-1}$.*

In order to evaluate the integrand in the spheric integral above, one has to be able to compute, for any fixed direction $v \in \mathbb{S}^{m-1}$, the $\chi$-probability of the one-dimensional set

$$\{r \geq 0 \,|\, (rLv + \mu) \cap A \neq \emptyset\}.$$

Since we are interested in the probability of the set $A := M_\delta$, this amounts by (14) to characterizing the set

$$\{r \geq 0 \,|\, g(\delta, rLv + \mu) \geq 0\} \quad (v \in \mathbb{S}^{m-1}), \tag{22}$$

where we set

$$g(\delta, b) := \min_{\substack{k,\ell=0,\ldots,m \\ k \neq l}} \left\{ h_k(b, \Phi^*_{k\ell}(\delta, b)) + (p_k^{\max})^2 - h_\ell(b, \Phi^*_{k\ell}(\delta, b)) - (p_\ell^{\min})^2 \right\}. \tag{23}$$

Using the idea of spheric-radial decomposition presented in Theorem 3.1, we propose the following algorithm for computing the probability $\mathbb{P}(\xi \in M_\delta)$ with a fixed value of $\delta$:

**Algorithm 3.1** *Let $\delta \in \mathbb{R}^{|\mathscr{E}|}$ be arbitrary, $\xi \sim \mathcal{N}(\mu, \Sigma)$ and $L$ such that $LL^T = \Sigma$.*

1. *Sample $N$ points $\{v_1, v_2, \ldots, v_N\}$ uniformly distributed on the sphere $\mathbb{S}^{m-1}$.*
2. *$i := 0$; $S := 0$*
3. *$i := i + 1$;*
   *Find the zero's of the one-dimensional function (in $r$ for $\delta$ fixed) $\theta_\delta(r) := g(\delta, rLv_i + \mu)$ with $g$ defined in (23) and represent the set $M^i_\delta := \{r \geq 0 \,|\, \theta_\delta(r) \geq 0\}$ corresponding to (22) as a disjoint union of intervals: $M^i_\delta = \cup_{j=1}^{\rho}[\alpha_j(\delta), \beta_j(\delta)]$, where $\alpha_j(\delta)$, $\beta_j(\delta)$ are the zero's obtained before and ordered appropriately.*

4. *Compute the $\chi$-probability of $M_\delta^i$ according to*

$$\mu_\chi(M_\delta^i) = \sum_j F_\chi(\beta_j(\delta)) - F_\chi(\alpha_j(\delta)),$$

*where $F_\chi$ refers to the cumulative distribution function of the one-dimensional $\chi$-distribution with $m$ degrees of freedom. Put $S := S + \mu_\chi(M_\delta^i)$*
5. *If $i < N$ then go to 3.*
6. *Set $\mathbb{P}(\xi \in M_\delta) := S/N$.*

A few words on this algorithm are in order at this place. The algorithm clearly provides an approximation to the spheric integral in Theorem 3.1 by means of a finite sum based on sampling of the sphere and then averaging the values of the integrand over all samples. Of course, this approximation will improve with the sampling size which may be large depending on the dimension $m$ of the problem (i.e., exit nodes in the network) and on the desired precision for the probability.

We recall that the uniform distribution on the sphere $\mathbb{S}^{m-1}$ can be represented as the distribution of $\eta/\|\eta\|$ (Euclidean norm), where $\eta$ has a standard Gaussian distribution in $\mathbb{R}^m$, i.e., $\eta \sim \mathcal{N}(0, I)$. Then, the simplest idea to sample a point $v^i$ on the sphere as in step 1. of the algorithm would be to independently sample $m$ values $w_j$ of a one-dimensional standard normal distribution by using standard random generators and then putting $v_i := w/\|w\|$ for $w := (w_1, \ldots, w_m)$. When replacing such Monte Carlo sampling of the normal distribution (based on random number generators) by Quasi-Monte Carlo sampling (based on deterministic low discrepancy sequences), one observes a dramatic improvement in the precision of the result. For our gas network problem (with fixed roughness coefficients), this was revealed in Gotzes et al. (2016) . A further improvement is to be expected for direct Quasi-Monte Carlo sampling on the sphere (not via normalization of Gaussian distributions) as discussed in Brauchart et al. (2014).

We illustrate the construction of the sets $M_i$ in step 3) of Algorithm 3.1 for the special case of uncertainty sets given by rectangles (21): thanks to Lemma 3.2, the optimal coefficients $\Phi_{k\ell}^*(\delta, b)$ do not depend on $b$. Hence, we may simply write them as $\Phi_{k\ell}^*(\delta)$ with values according to Lemma 3.2. Now, by (23) and by definition of $h_k$ below (10), we have that the function in $r$ whose zero's are looked for in step 3) takes the form

$$\theta_\delta(r) = \min_{\substack{k,\ell=0,\ldots,m \\ k\neq l}} \left\{ \sum_{e\in\Pi(k)} \left[\Phi_{k\ell}^*(\delta)\right]_e \left( \sum_{t\geq h(e)} r L_t v_i + \mu_t \right)^2 \right.$$
$$\left. - \sum_{e\in\Pi(l)} \left[\Phi_{k\ell}^*(\delta)\right]_e \left( \sum_{t\geq h(e)} r L_t v_i + \mu_t \right)^2 \right\},$$

where $L_t$ denotes row $t$ of the matrix $L$. Clearly, each of the expressions inside the minimum is quadratic in $r$, hence, $\theta$ may be written as a minimum of quadratic functions

$$\theta_\delta(r) = \min_{\substack{k,\ell=0,\ldots,m \\ k\neq l}} \left\{ c_{k\ell}(\delta) + d_{k\ell}(\delta) r + e_{k\ell}(\delta) r^2 \right\}, \tag{24}$$

with coefficients easily identified from the formula above. Since the zero's of $\theta_\delta$ must be contained in the zero's of all single quadratic functions inside the minimum, one may proceed as follows: determine first all zero's of the single quadratic functions above and order them as $x_1, \ldots x_K$. Second, delete from this list all zero's $x_n$ for which $x_n < 0$ or $\theta_\delta(x_n) < 0$. The remaining list, say $y_1, \ldots, y_{K'}$ will consist of the positive zero's of $\theta_\delta$. Third, identify neighboring zero's, between which the function $\theta_\delta$ remains positive, in order to represent the set $M_i$ in step 3) as a union of disjoint intervals. For instance, if $\theta_\delta(0) > 0$, then we'll have that

$$M_i = [0, y_1] \cup [y_2, y_3] \cup \cdots,$$

whereas for $\theta_\delta(0) < 0$ the representation will be

$$M_i = [y_1, y_2] \cup [y_3, y_4] \cup \cdots.$$

Observe, that the last of these intervals will be closed by the last zero $y_{K'}$ if $\theta_\delta(y_{K'}) < 0$ or by $\infty$ else.

Once a representation of the set $M_i$ in step 3. as a union of disjoint intervals has been obtained, step 4. is easily accomplished by applying efficient standard routines for high-precision approximations of the one-dimensional cumulative distribution function of the $\chi$-distribution.

## 4 Numerical solution of the optimization problem

In this section, we describe the numerical solution of the optimization problem (13) and illustrate the results for a concrete example. As a solution method for nonlinear optimization problems subject to inequality constraints we have chosen the projected gradient method (Gill et al. 1981) as it behaves rather robust with respect to the inevitable inaccuracy in the computation of probabilities by means of Algorithm 3.1. Note, that this inaccuracy can be reduced at the cost of increasing computation time by enhancing the sample size $N$. In order to apply the projected gradient method it is crucial not only to determine the ($\delta$-dependent) probabilities $\mathbb{P}(\xi \in M_\delta)$ in (12) but also their gradients with respect to $\delta$.

### 4.1 Approximating the gradient of the probability function

The sensitivity information on the $\delta$-dependent probability function mentioned above can be gained directly from inside Algorithm 3.1: by step 6), the probability is given by $S/N$, where $S$ is updated in step 4). Accordingly, the derivative of the probability with respect to the parameter $\delta$ can be approximated by the expression $S'/N$, where $S'$ is updated by the derivatives with respect to $\delta$ of the updates of $S$:

$$\sum_j f_\chi(\beta_j(\delta))\nabla\beta_j(\delta) - f_\chi(\alpha_j(\delta))\nabla\alpha_j(\delta).$$

Here, $f_\chi$ refers to the density of the one-dimensional $\chi$-distribution with $m$ degrees of freedom, i.e., $f_\chi$ is the derivative of the distribution function $F_\chi$ from step 4). Since there exists an analytical expression for $f_\chi$, all one needs to know for evaluating the expression above is the gradients $\nabla\beta_j, \nabla\alpha_j$, i.e, the gradients with respect to $\delta$ of the appropriate zero's of the function $\theta_\delta$ defined in step 3). These are easily found by representing the appropriate zero's of the associated quadratic equation

$$c_{k\ell}(\delta) + d_{k\ell}(\delta)r + e_{k\ell}(\delta)r^2 = 0$$

in (24) as $r(\delta)$ by using the classical solution formula and then deriving analytically $r(\delta)$ with respect to $\delta$. We emphasize that the procedure sketched above corresponds to calculating the gradient of the approximated probability function. This does not automatically have to coincide with calculating an approximation of the gradient of the true probability function, i.e., both operations don't have to commute. It was found in Van Ackooij and Henrion (2014) and Van Ackooij and Henrion (2017) by deriving corresponding gradient formulae that both approaches coincide under convexity of the underlying random inequality system. Since this assumption is not satisfied in our case for the system (10), a rigorous justification of the procedure described above (differentiation 'inside the algorithm') is missing so far. On the other hand, our numerical experience suggests that the projected gradient method performs well in finding local solutions to problem (13).

### 4.2 Illustration of a toy problem

For the purpose of illustration we start by considering a toy example of problem (13). Here, a simple network consisting of one entry node, one passive node (Innode) without injection or consumption and two exit nodes with random load is given as in Fig. 1.

Observe, that the innode can be formally modeled as an exit with zero consumption. As there exist 3 arcs joining the nodes, we have three roughness coefficients $\Phi_e$. We will consider the maximization of rectangular uncertainty sets $U_\delta$ as introduced in (21). We assume the following data:

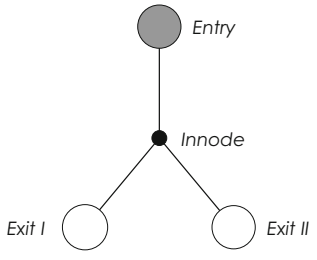


**Fig. 1** Structure of a simple network for a toy example containing one entry, one passive node and two exits

$$\begin{aligned}
&\hat{\Phi}_1 = \hat{\Phi}_2 = \hat{\Phi}_3 = 0.0015\\
&p^{\min} = (1, 1, 1, 1); \quad p^{\max} = (390, 200, 100, 120)\\
&\xi = (\xi_1, \xi_2) \sim \mathcal{N}\left((4100, 3900); \begin{pmatrix} 300 & 0 \\ 0 & 300 \end{pmatrix}\right)\\
&f(\delta) := (\delta_1)^{0.9} + (\delta_2)^{0.9} + (\delta_3)^{0.9}\\
&p = 0.8
\end{aligned}$$

Here, the components of the vector $\hat{\Phi}$ of average roughness coefficients correspond to the arcs (entry-innode, innode-exit 1, innode-exit 2), whereas the components of $p^{\min}$, $p^{\max}$ are labeled according to (entry, innode, exit 1, exit 2). The two exits are supposed to have Gaussian random load with means 4100 and 3900, respectively and standard deviation 300 both. We assume that the two loads are uncorrelated (which does not necessarily have to be the case in our approach). The objective function $f$ measuring the size of the uncertainty set $U_\delta$ has not been chosen on purpose as a norm. The reason is that, e.g., the one-norm (which would correspond to exponents one in the same expression) yields strongly 'concentrated' solutions, i.e. optimal rectangular uncertainty sets with many extremely short sides and a few long ones. In other words, in order to allow a larger uncertainty for a few roughness coefficients, one has to insist on extremely high accuracy for many other. Choosing exponents smaller than one makes it possible to control the contrast between different uncertainty ranges and to maintain a reasonable minimum amount. The probability level $p$ in (12) was set to 0.8. The numerical solution of problem (13) with the given data is:

$$\delta_1^* = 0.00014595; \quad \delta_2^* = 0.00006697; \quad \delta_3^* = 0.00020503.$$

Relating these values to the average roughness coefficients indicated above, the interpretation of this result is as follows: 80% of the random load scenarios at exits 1 and 2 are technically feasible in the sense of Prop. 2.1 **for all** roughness coefficients deviating relatively from their average values by at most 97.3, 44.6 and 136.7%, respectively. We can make a posterior check of this solution by simulating a set of exit load scenarios according to the Gaussian distribution given above and counting for how many of them feasibility holds true for all roughness coefficients in the calculated maximum uncertainty rectangle.

Figure 2 shows ten different simulated load patterns for the two exits (with values scattering around the averages indicated above indicated in the figure). For each simulation the optimal rectangular uncertainty with side lengths given above is shown along with all infeasible roughness coefficients colored inside. An empty rectangle indicates that the load pattern is feasible **for all** roughness coefficients in the rectangle. Then, by definition, this load scenario is counted as feasible. According to Fig. 2, eight out of ten load scenarios are feasible which corresponds (by chance exactly) to the chosen probability level $p = 0.8$. For the two infeasible scenarios the polyhedral set (actually, the complement of a polyhedron) of roughness coefficients in the rectangle violating feasibility is made visible.

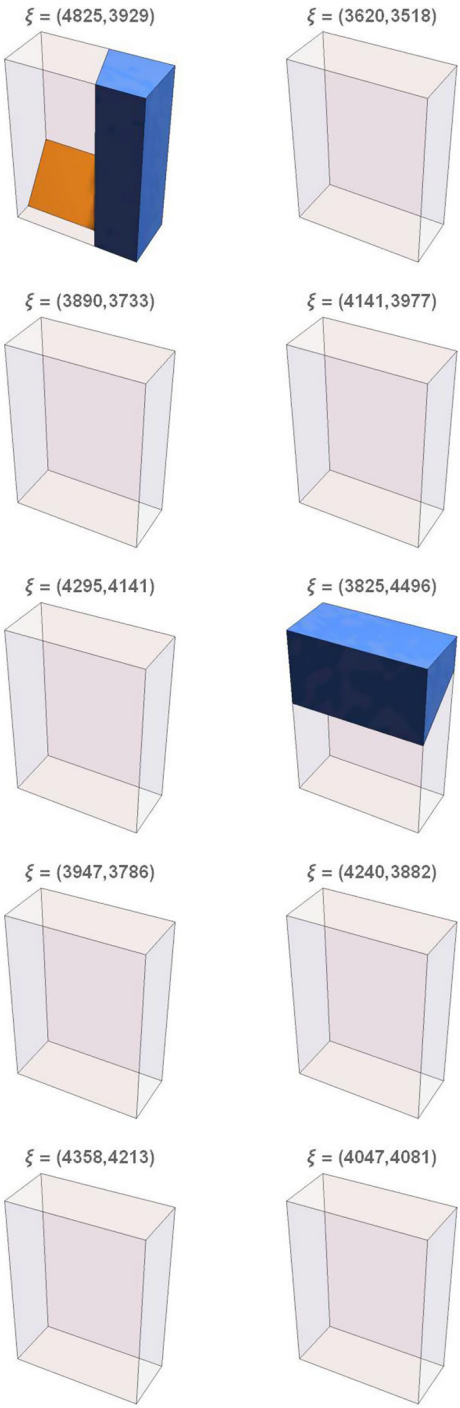


**Fig. 2** Simulation of ten exit load patterns along with infeasible roughness coefficients *colored* inside the optimal uncertainty set (color figure online)

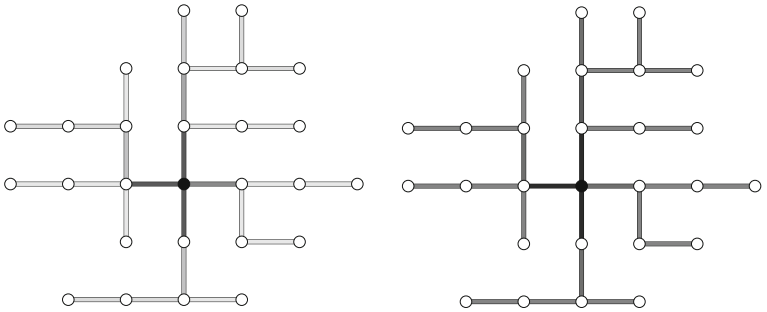

**Fig. 3** Computed sensitivity (displayed by *gray scale*) of the roughness coefficients for distinct pipes of a gas network with different probability levels: $p = 0.8$ (*left*) and $p = 0.9$ (*right*)

### 4.3 Numerical solution of a medium size problem

In this section, we provide the results of solving problem (13) for a more realistic setting with a network consisting of 27 nodes (1 entry and 26 exits). In other words, the random vector considered in the joint probabilistic/robust constraint follows a 26-dimensional Gaussian distribution. The network is illustrated in Fig. 3 with the entry located in the center. The pipes in the network are colored according to the relative uncertainty allowed for the corresponding roughness coefficients in an optimal (maximal) uncertainty set computed with the same objective as in the previous toy example. The colors follow a gray scale, where black means that only low uncertainty is allowed whereas white means high uncertainty can be tolerated. The solution in the left figure was obtained upon imposing a probability level $p = 0.8$ whereas the one in the right figure refers to a probability level $p = 0.9$. Not surprisingly, higher accuracy for roughness coefficients is required overall when increasing the probability level. More interestingly, higher accuracy is required at arcs close to the entry.

**Acknowledgements** The authors thank the Deutsche Forschungsgemeinschaft for support within project B04 in the Sonderforschungsbereich/Transregio 154 'Mathematical Modelling, Simulation and Optimization using the Example of Gas Networks'. Additionally, the first author expresses her gratitude to the SFB TRR 154 Integrated Graduate School.

## References

Ben-Tal A, El Ghaoui L, Nemirovski A (2009) Robust optimization. Princeton University Press, Princeton

Brauchart J, Saff E, Sloan I, Womersley R (2014) QMC designs: optimal order Quasi Monte Carlo integration schemes on the sphere. Math Comput 83:2821–2851

Chertkov M, Misra S, Vuffray M (2015) Monotonicity of dissipative flow networks renders robust maximum profit problem tractable: general analysis and application to natural gas flows. Math Optim Control. arXiv:1504.00910 [math.OC]

Colebrook C (1939) Turbulent flow in pipes with particular reference to the transition region between the smooth and rough pipe laws. J Inst Civ Eng I1:133–156

Déak I (2000) Subroutines for computing normal probabilities of sets-computer experiences. Ann Oper Res 100:103–122

Dentcheva D, Ruszczyński A (2003) Optimization with stochastic dominance constraints. SIAM J Optim 14(2):548–566

Egger H, Kugler T, Strogies N (2017) Parameter identification in a semilinear hyperbolic system. IOPsci Inverse Probl 33(5):055022

Genz A, Bretz F (2009) Computation of multivariate normal and t-probabilities, vol 195. Lecture notes in statistics. Springer, Heidelberg

Gill P, Murray W, Wright M (1981) Practical optimization. Academic Press, Cambridge

Gotzes C, Heitsch H, Henrion R, Schultz R (2016) Feasibility of nominations in stationary gas networks with random load. Math Methods Oper Res 84:427–457

Gotzes C, Nitsche S, Schultz R (2017) Probability of feasible loads in passive gas networks with up to three cycles. Preprint, available online at OPUS TRR154

Koch T, Hiller B, Pfetsch M, Schewe L (2015) Evaluating gas network capacities, MOS-SIAM series on optimization, vol 21. SIAM, Philadelphia

Pfetsch ME, Fügenschuh A, Geißler B, Geißler N, Gollmer R, Hiller B, Humpola J, Koch T, Lehmann T, Martin A, Morsi A, Rövekamp J, Schewe L, Schmidt M, Schultz R, Schwarz R, Schweiger J, Stangl C, Steinbach MC, Vigerske S, Willert BM (2015) Validation of nominations in gas network optimization: models, methods, and solutions. Optim Methods Softw 30(1):15–53. doi:10.1080/10556788.2014.888426

Prékopa A (1995) Stochastic programming. Kluwer, Dordrecht

Van Ackooij W, Henrion R (2014) Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions. SIAM J Optim 24:1864–1889

Van Ackooij W, Henrion R (2017) (Sub-) Gradient formulae for probability functions of random inequality systems under Gaussian distribution. SIAM/ASA J Uncertain Quantif 5:63–87

Van Ackooij W, Frangioni A, de Oliveira W (2016) Inexact stabilized Benders' decomposition approaches with application to chance-constrained problems withfinite support. Comput Math Appl 65:637–669