



# Probabilistic constrained optimization on flow networks

Michael Schuster<sup>1</sup> · Elisa Strauch<sup>2</sup> · Martin Gugat<sup>1</sup> · Jens Lang<sup>2</sup>

Received: 28 July 2020 / Revised: 11 March 2021 / Accepted: 11 March 2021 / Published online: 21 April 2021  
© The Author(s) 2021

## Abstract

Uncertainty often plays an important role in dynamic flow problems. In this paper, we consider both, a stationary and a dynamic flow model with uncertain boundary data on networks. We introduce two different ways how to compute the probability for random boundary data to be feasible, discussing their advantages and disadvantages. In this context, feasible means, that the flow corresponding to the random boundary data meets some box constraints at the network junctions. The first method is the spheric radial decomposition and the second method is a kernel density estimation. In both settings, we consider certain optimization problems and we compute derivatives of the probabilistic constraint using the kernel density estimator. Moreover, we derive necessary optimality conditions for an approximated problem for the stationary and the dynamic case. Throughout the paper, we use numerical examples to illustrate our results by comparing them with a classical Monte Carlo approach to compute the desired probability.

**Keywords** Probabilistic constraints · Uncertain boundary data · Spheric radial decomposition · Kernel density estimation · Gas networks · Contamination of water

**Mathematics Subject Classification** 49K20 · 49K45

## 1 Introduction and motivation

In this paper, we present a method which describes how to deal with uncertain loads in the context of flow networks. The modeling and simulation of flow networks like gas flow, water flow and the diffusion of harmful substances inside flow networks become more and more important. So in this paper, we analyze the gas flow

---

✉ Michael Schuster  
michi.schuster@fau.de

<sup>1</sup> Department of Mathematics, Friedrich-Alexander University Erlangen-Nürnberg, Cauerstrasse 11, 91058 Erlangen, Germany

<sup>2</sup> Department of Mathematics, Technical University of Darmstadt, Dolivostrasse 15, 64293 Darmstadt, Germany

through a pipeline network in a stationary setting and the contamination of water in a dynamic setting. The aim of this paper is to solve probabilistic constrained optimization problems and to derive necessary optimality conditions for them in the context of flow networks.

Gas transport resp. general flow problems have been a goal of many studies. In general, such a flow problem is modeled as a system of hyperbolic balance laws based on e.g. the isothermal Euler equations (for gas flow, see e.g. Banda et al. 2006a, b; Gotzes et al. 2016; Gugat et al. 2015, 2018; Gugat and Schuster 2018) or the shallow water equations (for water flow, see e.g. Bastin et al. 2009; Coron 2002; Gugat and Leugering 2003; Gugat et al. 2004; Leugering and Schmidt 2002). The model in the stationary setting in this paper is based on the stationary isothermal Euler equations for modeling the gas flow through a pipeline network. In Domschke et al. (2017), Koch et al. (2015) one can find a great overview about the topic of gas transport, existing models, and network elements. The existence of a unique stationary state is shown in Gugat et al. (2015), the stationary states for real gas are analyzed in Gugat et al. (2018) and Gugat and Wintergerst (2018). The existence of solutions for the dynamic case have been analyzed in Gugat and Ulbrich (2017, 2018). Optimal control problems in gas networks have been studied e.g. in Bermúdez et al. (2015), Colombo et al. (2009) and Gugat and Herty (2009). For the problem of contamination of water by harmful substances, we use a linear scalar hyperbolic balance law. This has also been analyzed in Fügenschuh et al. (2007) and Gugat (2012).

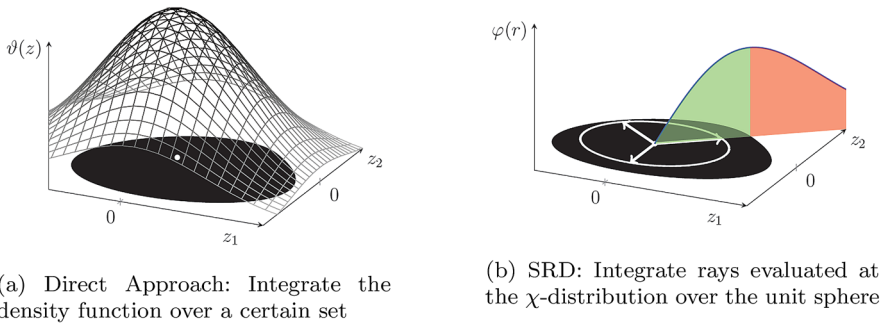
An important aspect of this paper is that we consider random boundary data. In the context of gas transport, that means that the loads (i.e., the gas demand) are random. In the context of water contamination, that means that the contaminant injection is random. This leads to optimization problems with probabilistic constraints (see e.g. Prékopa 1995; Shapiro et al. 2009). We also assume box constraints for the solution of the balance law at the network nodes and we define the set of feasible loads  $M$  as all loads, for which the solution of the balance law meets these box constraints. Our aim is to compute the probability for a random load vector to be feasible, i.e., we want to compute the probability for a random vector to be in a certain set  $M$ . So we identify the load vector with some random vector  $\xi$  on an appropriate probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  and we want to compute the probability

$$\mathbb{P}(\omega \in \Omega \mid \xi(\omega) \in M),$$

which we write as

$$\mathbb{P}(\xi \in M).$$

A direct approach to compute this probability is to integrate the probability density function of the balance law solution over the box constraints. However this density function may not be known. We use a *kernel density estimator* (see e.g. Nadaraya 1965; Parzen 1962; Scott and Terrell 1992) to obtain an approximation of this function. In Duller (2018) and Härdle et al. (2004) one can get a great overview about the area of nonparametric statistics. The authors in Gotzes et al. (2016) and Gugat and Schuster (2018) use the spheric radial decomposition (see e.g. van Ackooij et al.



**Fig. 1** Computing the probability for a random vector to be in a certain set: direct approach versus SRD

2018; van Ackooij and Henrion 2014; Farshbaf-Shaker et al. 2018; Gotzes et al. 2016; Gradón et al. 2016) for a similar gas flow problem to compute the desired probability.

**Theorem 1** (Spheric radial decomposition, see Gotzes et al. 2016, Theorem 2) *Let  $\xi \sim \mathcal{N}(0, R)$  be the  $n$ -dimensional standard Gaussian distribution with zero mean and positive definite correlation matrix  $R$ . Then, for any Borel measurable subset  $M \subseteq \mathbb{R}^n$  it holds that*

$$\mathbb{P}(\xi \in M) = \int_{\mathbb{S}^{n-1}} \mu_\chi \{ \hat{r} \geq 0 \mid \hat{r}Ls \in M \} d\mu_\eta(s), \tag{1}$$

where  $\mathbb{S}^{n-1}$  is the  $(n - 1)$ -dimensional sphere in  $\mathbb{R}^n$ ,  $\mu_\eta$  is the uniform distribution on  $\mathbb{S}^{n-1}$ ,  $\mu_\chi$  denotes the  $\chi$ -distribution with  $n$  degrees of freedom and  $L$  is such that  $R = LL^\top$  (e.g., Cholesky decomposition).

This result can be applied to general Gaussian distributions easily: For  $\xi \sim \mathcal{N}(\mu, \Sigma)$ , set  $\xi^* = D^{-1}(\xi - \mu) \sim \mathcal{N}(0, R)$  with  $D = \text{diag}(\sqrt{\Sigma_{ii}})$  and  $R = D^{-1}\Sigma D^{-1}$ . Then it follows  $\mathbb{P}(\xi \in M) = \mathbb{P}(\xi^* \in D^{-1}(M - \mu))$ . An algorithmic formulation of the spheric radial decomposition is given in Gotzes et al. (2016) and Gugat and Schuster (2018).

From now on we use SRD instead of spheric radial decomposition and KDE instead of kernel density estimator. The difference in both methods is shown in Fig. 1.

An advantage of the SRD is that we exploit almost all information we can get from the model. Thus the only numerical error occurs while approximating the spherical integral. The big disadvantage of the SRD is that we need to know an analytical solution of our model which we cannot always guarantee. Therefore we introduce a KDE, which estimates the probability density function of a random variable by using a sampling set of the variable.

**Definition 1** (*Kernel density estimator, see Gramacki (2018)*) Let  $y$  be a  $n$ -dimensional real-valued random variable with an absolutely continuous distribution and probability density function  $\rho$  with respect to the Lebesgue-measure. Moreover, let  $\mathcal{Y} = \{y^{S,1}, \dots, y^{S,N}\}$  be an independent and identically distributed sample of  $y$ . Then, the kernel density estimator  $\rho_N : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  is defined as

$$\rho_N(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K\left(H^{-\frac{1}{2}}(z - y^{S,i})\right),$$

with a symmetric positive definite bandwidth matrix  $H \in \mathbb{R}^{n \times n}$  and a  $n$ -variate density  $K : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  called kernel.

We apply the kernel density estimation to the balance law solution. If an analytical solution of the model is not known, we can compute the solution numerically and use a sampling set of approximated solutions. Then the desired probability can be computed by integrating the kernel density estimator over the box constraints. That means, we get an approximation error but we can analytically work with numerical solutions of our model. A KDE approach was used in Cailiau et al. (2018) for solving chance constrained optimal control problems with ODE constraints. But it was neither used in optimal control problems with PDE constraints and random boundary data nor in the context of continuous optimization with hyperbolic balance laws on networks. Throughout this paper, we illustrate the idea of the KDE in both, the stationary and the dynamic case, such that we can state necessary optimality conditions for optimization problems with probabilistic constraints. This paper is structured as follows:

In the next section, we consider stationary gas networks, similar to Gotzes et al. (2016) and Gugat and Schuster (2018). We first consider a simple model on a graph with only one edge to explain the ideas of the SRD and the KDE. We compare both results in a numerical computation with a classical Monte Carlo method (All numerical tests have been done with MATLAB<sup>®</sup>, version 2015a). Next we use both methods, the SRD and the KDE, to solve a model on a general tree-structured graph and again we compare both methods with a classical Monte Carlo method. Finally, we state necessary optimality conditions for probabilistic constrained optimization problems related to our stationary model. Last in this section we solve a probabilistic constrained optimization problem on a realistic gas network setting.

In Sect. 3, we consider a dynamic water network, in which the contaminant injection occurs at the boundaries. We consider a general linear hyperbolic balance law, which models the diffusion of harmful substances on a linear graph in order to discuss probabilistic constraints in the time dependent case and, whether the SRD can be expanded to this case. We also use the KDE for this model. Finally, we state necessary optimality conditions for probabilistic constrained optimization problems related to our dynamic model and solve a probabilistic constrained optimization problem for a realistic water contamination network setting.

## 2 Gas networks in a stationary setting

In this section, we consider stationary states in gas networks. As mentioned before, the model here is similar to the model in Gotzes et al. (2016) and Gugat and Schuster (2018). The main difference is that we fix an inlet pressure for our model, which the authors in Gotzes et al. (2016) and Gugat and Schuster (2018) did not. The network is described by a connected, directed, tree-structured graph  $G = (\mathcal{V}, \mathcal{E})$  (i.e., the graph does not contain cycles) with the vertex set  $\mathcal{V} = \{v_0, \dots, v_n\}$  and the set of edges  $\mathcal{E} = \{e_1, \dots, e_n\} \subseteq \mathcal{V} \times \mathcal{V}$ . We assume that the graph has only one inflow node  $v_0$  and that the other nodes are outflow nodes. Let node  $v_0$  be the root of the graph orientated away from the root. The graph is numbered from the root  $v_0$  using breadth-first search or depth-first search. Every edge  $e \in \mathcal{E}$  represents a pipe with positive length  $L^e$ . For  $x \in [0, L^e]$  we consider the stationary semi-linear isothermal Euler equations for horizontal pipes and ideal gases

$$\begin{cases} q_x^e(x) = 0, \\ (c^e)^2 p_x^e(x) = -\frac{\lambda^e}{2D^e} (R_S T)^2 \frac{q^e(x) |q^e(x)|}{p^e(x)}. \end{cases} \tag{2}$$

With  $p^e = p|_e$  we represent the restriction of the pressure defined over the network to a single edge  $e$  and  $p_x^e$  resp.  $q_x^e$  is the derivative of  $p^e$  resp.  $q^e$  w.r.t.  $x$ . Here,  $q^e$  is the flow along edge  $e$  and  $c^e, \lambda^e, D^e \in \mathbb{R}_{>0}$  denote the sound speed, the friction coefficient and the pipe diameter. The parameters  $R_S$  and  $T$  denote the specific gas constant of natural gas and the (constant) temperature. Note that  $q^e$  is constant on every edge. With  $q^e \geq 0$  we denote that gas flows along the orientation of edge  $e$  and with  $q^e \leq 0$  we denote that gas flows against the orientation of edge  $e$ .

We consider conservation of mass for the flow at the nodes (cf. Kirchhoff's first law). Let  $\mathcal{E}_+(v)$  resp.  $\mathcal{E}_-(v)$  be the set of all outgoing resp. ingoing edges at node  $v \in \mathcal{V}$ . Let  $b^v \in \mathbb{R}$  be the load at node  $v \in \mathcal{V}$ . With  $b^v \geq 0$  we denote that gas leaves the network at node  $v$  (exit node) and with  $b^v \leq 0$  that gas enters the network at node  $v$  (entry node). The equation for mass conservation for every node  $v \in \mathcal{V}$  is given by

$$\sum_{e \in \mathcal{E}_-(v)} q^e(L^e) = b^v + \sum_{e \in \mathcal{E}_+(v)} q^e(0). \tag{3}$$

Let  $p_i$  denote the pressure at the node  $v_i$  for  $i = 0, \dots, n$ . We assume continuity in pressure at every node, i.e., for all  $v_i \in \mathcal{V}$  it holds

$$\begin{aligned} p^{e_1}(L^{e_1}) &= p_i & \forall e_1 \in \mathcal{E}_-(v_i), \\ p^{e_2}(0) &= p_i & \forall e_2 \in \mathcal{E}_+(v_i). \end{aligned} \tag{4}$$

Therefore the pressure  $p_i$  is defined by the pressure  $p^e(L^e)$  resp.  $p^e(0)$  with ingoing resp. outgoing edge  $e$  at node  $v_i$ . We consider (positive) box constraints for the pressures at all outflow nodes  $v_1, \dots, v_n$ , s.t.

$$p_i \in [p_i^{\min}, p_i^{\max}] \quad \forall i \in \{1, \dots, n\}. \tag{5}$$

In addition, we impose pressure  $p_0$  at node  $v_0$ . So for the full graph, we consider the following model:

$$\left\{ \begin{array}{ll} q_x^e(x) = 0 & \forall e \in \mathcal{E}, \\ (c^e)^2 p_x^e(x) = -\frac{\lambda^e}{2D^e} (R_s T)^2 \frac{q^e(x)|q^e(x)|}{p^e(x)} & \forall e \in \mathcal{E}, \\ p^e(0) = p_0 & \forall e \in \mathcal{E}_+(v_0), \\ \sum_{e \in \mathcal{E}_-(v)} q^e(L^e) = b^v + \sum_{e \in \mathcal{E}_+(v)} q^e(0) & \forall v \in \mathcal{V} \\ p^{e_1}(L^{e_1}) = p_i & \forall v_i \in \mathcal{V}, e_1 \in \mathcal{E}_-(v_i) \\ p^{e_2}(0) = p_i & \forall v_i \in \mathcal{V}, e_2 \in \mathcal{E}_+(v_i) \\ p_i \in [p_i^{\min}, p_i^{\max}] & \forall i \in \{1, \dots, n\}. \end{array} \right. \tag{6}$$

Our aim is now to find loads  $b = (b^v)_{v \in \mathcal{V} \setminus \{v_0\}} \in \mathbb{R}_{\geq 0}^n$  (corresponding to the outflow nodes  $v_1, \dots, v_n$ ), s.t. the model (6) has a solution. We do not consider the load  $b^{v_0}$  at the inflow node  $v_0$ , because Eq. (3) provides the relation  $b^{v_0} = -\sum_{v \in \mathcal{V} \setminus \{v_0\}} b^v$ . We define

$$M := \{ b \in \mathbb{R}_{\geq 0}^n \mid \text{There exists a solution of (6)} \}$$

as the set of feasible loads. To go one step further we assume that  $b$  is random. This is motivated by reality. Because of the liberalization of the gas market,<sup>1</sup> the gas network is independent of the consumers and the gas companies. That means a gas network company must guarantee that the required gas can be transported through the network, but the company does not know the exact amount of gas a priori, so it can be seen as random. We assume

$$b \sim \mathcal{N}(\mu, \Sigma)$$

with mean value  $\mu \in \mathbb{R}_+^n$  and positive definite covariance matrix  $\Sigma \in \mathbb{R}^{n \times n}$  on an appropriate probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Motivated by the application, we assume  $\mu$  and  $\Sigma$  are chosen s.t. the probability that  $b$  takes positive values is almost 1. This assumption would not be needed if the gas demand would be modeled e.g. by a truncated Gaussian distribution but we focus on the non truncated case in this work. In Gotzes et al. (2016) and in Koch et al. (2015), Chapter 13, the authors explain why a multivariate Gaussian distribution is a good choice for the random load vector. So we want to know the probability that for a Gaussian distributed load vector  $b$ , the model (6) has a solution, i.e.,

$$\mathbb{P}(b \in M).$$

In the next subsections we will use two different approaches for computing this probability. First, to explain the ideas of our approaches we consider a graph with only one edge and later we use the introduced tree-structured graph with one input.

<sup>1</sup> [http://www.gesetze-im-internet.de/enwg\\_2005/index.html](http://www.gesetze-im-internet.de/enwg_2005/index.html).

### 2.1 Uncertain load on a single edge

As mentioned before, we will give two different ways to compute the probability for a random load value to be feasible. Here, we consider a single edge as graph. For the boundary conditions  $p(0) = p_0$  and  $q(L) = b$  model (2) has an analytical solution. Therefore our problem simplifies to verifying the following inequalities:

$$\begin{cases} q(x) = b, \\ p(x) = \sqrt{p_0^2 - \frac{\lambda}{c^2 D} (R_S T)^2 q(x) |q(x)| x}, \\ p_1 = p(L) \in [p^{\min}, p^{\max}]. \end{cases} \tag{7}$$

In fact, that means a random value  $b \in \mathbb{R}_{\geq 0}$  is feasible, if the pressure at the end satisfies the box constraints, i.e.

$$b \in M \iff b \geq 0 \text{ and } p_1 = \sqrt{p_0^2 - \phi b |b|} \in [p^{\min}, p^{\max}]$$

with  $\phi = \frac{\lambda}{c^2 D} (R_S T)^2 L$ . We can rewrite the box constraints for the pressure as inequalities. That means  $p_1 \in [p^{\min}, p^{\max}]$  iff

$$\begin{aligned} (p^{\min})^2 &\leq p_0^2 - \phi b |b|, \\ (p^{\max})^2 &\geq p_0^2 - \phi b |b|. \end{aligned} \tag{8}$$

Now we can use the SRD (introduced in Sect. 1) in an algorithmic way. We consider  $b \sim \mathcal{N}(\mu, \sigma^2)$  with mean value  $\mu \in \mathbb{R}_+$  and standard deviation  $\sigma \in \mathbb{R}_+$ . Because the unit sphere in this example is given by  $\{-1, 1\}$ , we need not to sample  $N$  points, we can use the unit sphere itself as sampling. For  $s \in \mathbb{S}^0$ , we set

$$b_s(\hat{r}) = \hat{r} \sigma s + \mu.$$

Note that in the multivariate Gaussian distribution, the covariance  $\Sigma$  is given, which contains the variances of the random variables on its diagonal. In the one dimensional case, the standard deviation is given, which is a Cholesky decomposition of the variance.

To guarantee, that  $b_s(\hat{r})$  is positive, we define the regular range

$$R_{s,\text{reg}} := \{\hat{r} \geq 0 \mid b_s(\hat{r}) \geq 0\}.$$

Thus we have

$$M_s = \{\hat{r} \in R_{s,\text{reg}} \mid b_s(\hat{r}) \in M\} = \{\hat{r} \in R_{s,\text{reg}} \mid b_s(\hat{r}) \text{ satisfies (8)}\}.$$

We insert  $b_s(\hat{r})$  in the inequalities in (8). Since  $b_s(\hat{r}) \geq 0$  on  $R_{s,\text{reg}}$ , we can write  $b_s^2$  instead of  $b_s |b_s|$ . Thus we have quadratic inequalities in  $r$ :

$$\begin{aligned} \hat{r}^2(\sigma^2 s^2 \phi) + \hat{r}(2\sigma s \mu \phi) + (\mu^2 \phi + (p^{\min})^2 - p_0^2) &\leq 0, \\ \hat{r}^2(\sigma^2 s^2 \phi) + \hat{r}(2\sigma s \mu \phi) + (\mu^2 \phi + (p^{\max})^2 - p_0^2) &\geq 0. \end{aligned}$$

So we can write the set  $M_s$  as a union of  $\kappa \in \mathbb{N}$  disjoint intervals, i.e.

$$M_s = \bigcup_{j=1}^{\kappa} I_{s,j}$$

with intervals  $I_{s,j} = [a_{s,j}, \bar{a}_{s,j}]$  and interval bounds  $a_{s,j}, \bar{a}_{s,j} \in \mathbb{R}$ ,  $a_{s,j} \leq \bar{a}_{s,j}$  ( $j = 1, \dots, \kappa$ ). Since the unit sphere contains only two values and

$$\begin{aligned} \mathbb{P}(b \in M) &= \frac{1}{2} \sum_{s \in \{-1,1\}} \mu_{\chi}(M_s) = \frac{1}{2} \sum_{s \in \{-1,1\}} \sum_{j=1}^{\kappa} \mu_{\chi}(I_{s,j}) \\ &= \frac{1}{2} \sum_{s \in \{-1,1\}} \sum_{j=1}^{\kappa} \mathcal{F}_{\chi}(\bar{a}_{s,j}) - \mathcal{F}_{\chi}(a_{s,j}), \end{aligned} \tag{9}$$

where  $\mathcal{F}_{\chi}(\cdot)$  is the cumulative distribution of the  $\chi$ -distribution, we can compute the probability for a random vector to be feasible. As mentioned before, the SRD gives us an efficient way to compute this probability, but we need to know the analytical solution of our system (6).

Another way to compute the probability for a random load vector to be feasible is the *kernel density estimator*. It is more general and does not require the analytical solution of our model. We consider the stochastic equation corresponding to (2) with random boundary condition  $b$  and coupling conditions (3), (4) which has also a solution  $\mathbb{P}$ -almost surely. Hence the pressure at node  $v_1$  is a random variable which we denote with  $p_1$ . The probability that a random load is feasible is equal to the probability that the pressure  $p_1$  is in the interval  $[p^{\min}, p^{\max}]$ . We assume that the variance of  $p_1$  is positive and that its distribution of the pressure  $p_1$  is absolutely continuous with probability density function  $\rho_p$ . Now, the probability  $\mathbb{P}(p_1 \in [p^{\min}, p^{\max}])$  can be computed by integrating the probability density function  $\rho_p$  over the pressure bound, so we get

$$\mathbb{P}(b \in M) = \mathbb{P}(p_1 \in [p^{\min}, p^{\max}]) = \int_{p^{\min}}^{p^{\max}} \rho_p(z) dz. \tag{10}$$

If the exact probability density function  $\rho_p$  is not known, we can approximate the function by a kernel density estimator. Then we integrate this estimator over the interval  $[p^{\min}, p^{\max}]$  to get an approximation of the probability  $\mathbb{P}(b \in M)$ .

As before, we consider

$$b \sim \mathcal{N}(\mu, \sigma^2)$$

with mean value  $\mu \in \mathbb{R}_+$  and standard deviation  $\sigma \in \mathbb{R}_+$ .

Let  $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N}\} \subseteq \mathbb{R}_{\geq 0}$  be independent and identically distributed samples of the random variable  $b$ . Let  $\mathcal{P}_{\mathcal{B}} = \{p_1(b^{S,1}), \dots, p_1(b^{S,N})\} \subseteq \mathbb{R}$  be the



pressures at the end of the edge for the different loads  $b^{S,i} \in \mathcal{B}$  ( $i = 1, \dots, N$ ), which are also independent and identically distributed. We use the KDE in one dimension with bandwidth  $h \in \mathbb{R}_+$  and with a Gaussian kernel (see e.g. Duller 2018; Scott and Terrell 1992) to get an approximation of the density function  $\rho_p$ :

$$\varrho_{p,N}(z) = \frac{1}{Nh} \sum_{i=1}^N \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z - p_1(b^{S,i})}{h}\right)^2\right). \tag{11}$$

**Remark 1** The choice of the bandwidth  $h$  is a separate topic. We refer to Duller (2018), Chapter 8 and Härdle et al. (2004), Chapter 3 for studies about optimal bandwidths for KDEs. Here we use the heuristic formula for the bandwidth given by

$$h = 1.06 \frac{\sigma_N}{\sqrt[5]{N}}, \tag{12}$$

where  $\sigma_N$  is the standard deviation of the sampling and  $N$  is the number of samples. The idea is, that we compute the bandwidth depending on the variance of the sampling. This is stated and explained e.g. in Gramacki (2018), Chapter 4.2 and in Turlach (1993).

For the previous bandwidth it holds  $(h + (Nh)^{-1}) \rightarrow 0$   $\mathbb{P}$ -almost surely for  $N \rightarrow \infty$ . Therefore, Devroye and Györfi (1985) (Chapter 6, Theorem 1) provides the  $L^1$ -convergence of the estimator:

$$\|\varrho_p - \varrho_{p,N}\|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P}\text{-almost surely.} \tag{13}$$

Thus, for an appropriate choice of the bandwidth  $h$ , we can use the KDE as an approximation for the exact probability density function of the pressure. From *Scheffé's lemma* (see Devroye and Györfi 1985), it follows

$$\left| \int_{p^{\min}}^{p^{\max}} \varrho_p(z) dz - \int_{p^{\min}}^{p^{\max}} \varrho_{p,N}(z) dz \right| \leq \frac{1}{2} \|\varrho_p - \varrho_{p,N}\|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P}\text{-almost surely.} \tag{14}$$

So with (11) we can approximate the probability for a random load vector to be feasible as follows:

$$\mathbb{P}(b \in M) = \mathbb{P}(p_1 \in [p^{\min}, p^{\max}]) \approx \int_{p^{\min}}^{p^{\max}} \varrho_{p,N}(z) dz =: \mathbb{P}_N(b \in M). \tag{15}$$

In this example, we can compute the sampling set  $\mathcal{P}_B$  analytically, because the analytical solution of model (2) is known. If this is not the case, e.g., for more complex systems like the stationary full Euler equations or Navier-Stokes equations (see Domschke et al. 2017), one can use numerical methods to solve the PDE and to get an approximated sampling set  $\mathcal{P}_B$ .

**Table 1** Values for the example with one edge

$p_0$	$p^{\min}$	$p^{\max}$	$\mu$	$\sigma$	$\phi$
60	40	60	4	0.5	100

**Table 2** Results for the example with one edge

	Test 1 (%)	Test 2 (%)	Test 3 (%)	Test 4 (%)	Test 5 (%)	Test 6 (%)	Test 7 (%)	Test 8 (%)
MC	82.99	82.88	82.83	82.95	83.09	82.77	82.83	82.58
KDE	82.83	82.75	82.69	82.74	82.91	82.58	82.70	82.43
SRD	82.75							

**Example 1** We give an example to illustrate that the results of the KDE and the SRD are similar. Therefore we use the values (without units) from Table 1. With the inequalities (8) and the values given in Table 1 we can see that  $b$  is feasible iff  $b \in [0, \sqrt{20}]$ . For comparing the probability we use a classical Monte Carlo (MC) method, in which we check the percentage number of points inside  $[0, \sqrt{20}]$ . For both, the MC method and the KDE approach, we use the same sampling of  $5 \cdot 10^4$  points. The bandwidth for the KDE is given by (12). The result for 8 tests is shown in Table 2.

The sphere of the SRD in this example is finite ( $\mathbb{S}^0 = \{-1, 1\}$ ), so the SRD gives always the exact probability (except numerical errors due to quadrature) of 82.75%. Also, the probabilities computed by MC and the KDE are quite similar. The mean probability in MC resp. KDE is 82.87% resp. 82.40% and the variances are 0.0238 resp. 0.0218, which is very close to the (exact) result of the SRD. Further we provide confidence intervals with confidence level 95% for the results stated in Table 2. The confidence level of 95% is quite common in statistics. An introduction to confidence intervals can be found in (Linde 2016, Chapter 8). The confidence interval for the MC probabilities is [82.74%, 82.99%] and the confidence interval for the KDE probability is [82.58%, 82.83%]. One can see that the (exact) probability (computed via SRD) is contained in both intervals. Thus one can see, that the KDE with the heuristic bandwidth (12) provides a suitable method to compute the desired probability.

A remaining question is how to choose the sample size s.t. the result is sufficiently accurate. In general it holds the larger the sample size the more accurate is the solution. A good sample size can be found by observing the sample error. In Roache (1998) and Schwer (2007) the authors suggest comparing the numerical result with two other numerical results for larger sample sizes. This mesh-to-mesh comparison is often used in numerics to determine the rate of convergence of the solution. With this approach an appropriate sample size can be chosen for the desired accuracy.

### 2.2 Uncertain loads on tree-structured graphs

For this subsection, we consider a tree-structured graph (i.e., the graph does not contain cycles)  $G = (\mathcal{V}, \mathcal{E})$  with one entry  $v_0$ , as introduced in the beginning of Sect. 2. Let the graph be numbered from the root  $v_0$  by breadth-first search or depth-first search and let the model (6) holds on every edge.

First we rewrite the system (6) using the solution of the isothermal Euler equations. We use the incidence matrix  $A^+ \in \mathbb{R}^{(n+1) \times n}$  of the graph. For an edge  $e_\ell \in \mathcal{E}$ , which connects the nodes  $v_i$  and  $v_j$  starting from node  $v_i$ , we have

$$A^+_{k,\ell} = \begin{cases} -1 & \text{if } k = i, \\ 1 & \text{if } k = j, \\ 0 & \text{else.} \end{cases}$$

A formal definition of the incidence matrix can be found e.g., in Gotzes et al. (2016), Gugat and Schuster (2018) and Gugat et al. (2020). We set  $A \in \mathbb{R}^{n \times n}$  as  $A^+$  without the first row (which corresponds to the root resp. the only entry node of the graph). The equation for mass conservation (3) can equally be written as

$$A q = b, \tag{16}$$

where  $q_j$  is the (constant) flow on the edge  $e_j$  and  $b_i$  is the load at node  $v_i$  ( $i, j = 1, \dots, n$ ). Due to the tree-structuredness of the graph,  $A$  is a square matrix with full rank and thus invertible. Numbering the graph by breadth-first search or depth-first search implies that the matrix  $A$  is upper triangular, just like its inverse. Further in (Gotzes et al. 2016, Section 3.2), the authors mention, that  $A^{-1}_{ij}$  is one if and only if the edge  $e_j$  is on the (unique) path from the root to node  $v_i$ , otherwise  $A^{-1}_{ij}$  is zero. Motivated by (7) we define the function

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^n, \quad g : b \mapsto (A^T)^{-1} \Phi \left( (A^{-1}b) \circ |A^{-1}b| \right),$$

where  $\Phi \in \mathbb{R}^{n \times n}$  is a diagonal matrix with the values  $\phi^e = \frac{\lambda^e}{(c^e)^2 D^e} (R_S T)^2 L^e$  ( $e \in \mathcal{E}$ ) at its diagonal. The product on the right has to be understood componentwise. The  $i$ -th component of this function states the pressure loss from the root  $v_0$  to node  $v_i$ . The term  $(A^{-1}b)$  comes from the equation for mass conservation (16) and contains the (constant) flows at the edges. With this function we get the following equivalence for feasible loads:

**Lemma 1** *A load vector  $b \in \mathbb{R}^n_{\geq 0}$  is feasible, i.e., there exists a solution of (6), iff the following system of inequalities holds for all  $k = 1, \dots, n$ :*

$$\begin{aligned} p_0^2 &\leq (p_k^{\max})^2 + g_k(b), \\ p_0^2 &\geq (p_k^{\min})^2 + g_k(b). \end{aligned}$$

**Proof** The result follows from (Gotzes et al. 2016, Corollary 1). In their setting, the inlet pressure  $p_0$  is not given explicitly, it is given inside a range  $[p_0^{\min}, p_0^{\max}]$ . Then

*Corollary 1* in Gotzes et al. (2016) states, that a load vector  $b \in \mathbb{R}_{\geq 0}^n$  is feasible, iff the following system of inequalities holds:

$$\begin{aligned} p_0^{\min} &\leq \min_{k=1,\dots,n} [p_k^{\max} + g_k(b)] \\ p_0^{\max} &\geq \max_{k=1,\dots,n} [p_k^{\min} + g_k(b)] \\ \max_{k=1,\dots,n} [p_k^{\min} + g_k(b)] &\leq \min_{k=1,\dots,n} [p_k^{\max} + g_k(b)]. \end{aligned}$$

In our setting the inlet pressure  $p_0$  is explicitly given, which means  $p_0^{\min} = p_0^{\max}$ . Then the third inequality directly follows from the first one and the second one and thus, our lemma is a special case of *Corollary 1* in Gotzes et al. (2016).  $\square$

**Lemma 2** *If  $p_i^{\min} \leq p_0 \leq p_j^{\max}$  for all  $i, j = 1, \dots, n$ , then the set of feasible loads  $M$  is convex.*

**Proof** The proof is equal to the proof of *Theorem 11* in Gugat et al. (2020) for linear graphs, i.e., graphs without junctions. Here, the proof also works for tree-structured networks, because the pressure at the inflow node is explicitly given. Thus, the inequality, where convexity for tree-structured networks breaks in Gugat et al. (2020), is redundant here (see Lemma 1).  $\square$

Now we use the SRD to compute the probability for a random load vector to be feasible. For this setting, the SRD approach is explained in detail in Gotzes et al. (2016) and Gugat and Schuster (2018). Let  $b \sim \mathcal{N}(\mu, \Sigma)$  with mean value  $\mu \in \mathbb{R}_+^n$  and positive definite covariance matrix  $\Sigma \in \mathbb{R}^{n \times n}$  for an appropriate probability space  $(\Omega, \mathcal{A}, \mathbb{P})$  be given. The next steps are similar to the case of one edge. For a point  $s$  of a sample  $\mathcal{S} := \{s_1, \dots, s_N\} \subseteq \mathbb{S}^{n-1}$  of  $N \in \mathbb{N}$  uniformly distributed points at the unit sphere  $\mathbb{S}^{n-1}$ , we set

$$b_s(\hat{r}) := \hat{r}\mathcal{L}s + \mu,$$

where  $\mathcal{L}$  is s.t.  $\mathcal{L}\mathcal{L}^T = \Sigma$ . Because we only consider outflows,  $b_s(\hat{r})$  must be positive. We define the regular range

$$R_{s,\text{reg}} := \{\hat{r} \geq 0 | b_s(\hat{r}) \geq 0\}.$$

From this, it follows that  $b_s(\hat{r}) \in R_{s,\text{reg}}$  is feasible, iff  $b_s(\hat{r})$  satisfies the inequalities in Lemma 1. The inequalities are quadratic in the variable  $r$  and we can write the sets  $M_s$  as unions of disjoint intervals. Thus, (9) gives us the probability for a random load vector to be feasible.

As in the previous subsection, we consider the stochastic equation corresponding to (2) with random load  $b$  and coupling conditions (3), (4). The  $n$ -dimensional random vector  $p$  denotes the pressure at the nodes  $v_1, \dots, v_n$ . The probability that  $b$  is feasible is equal to the probability that the pressure  $p$  at the nodes is in the pressure bounds, which we denote by  $\mathbb{P}(p \in P_{\min}^{\max})$  with  $P_{\min}^{\max} := \bigotimes_{i=1}^n [p_i^{\min}, p_i^{\max}]$ . We assume that the covariance matrix of the random vector  $p$  is positive definite

and that its distribution is absolutely continuous with probability density function  $\varrho_p$ . Now, it holds

$$\mathbb{P}(b \in M) = \mathbb{P}(p \in P_{\min}^{\max}) = \int_{P_{\min}^{\max}} \varrho_p(z) dz. \tag{17}$$

However the exact probability density function  $\varrho_p$  is not known. But we can approximate the function using a multidimensional kernel density estimation.

Let  $\mathcal{B} = \{b^{S,1}, \dots, b^{S,N}\}$  be independent and identically distributed non-negative samples of the random load vector  $b \sim \mathcal{N}(\mu, \Sigma)$ . Then, let  $\mathcal{P}_{\mathcal{B}} = \{p(b^{S,1}), \dots, p(b^{S,N})\} \subseteq \mathbb{R}^n$  be the pressures at the nodes  $v_1, \dots, v_n$  for the different loads  $b^{S,i} \in \mathcal{B}$  ( $i = 1, \dots, N$ ). These samples are also independent and identically distributed. Note that  $p_j(b^{S,i})$  ( $i \in \{1, \dots, N\}, j \in \{1, \dots, n\}$ ) is the pressure at node  $v_j$  and the  $j$ -th component of the pressure vector  $p(b^{S,i})$ . We introduce the general  $n$ -dimensional multivariate kernel density estimator (see e.g. Gramacki 2018):

$$\varrho_{p,N}(z) = \frac{1}{N \det(H)^{\frac{1}{2}}} \sum_{i=1}^N K\left(H^{-\frac{1}{2}}(z - p(b^{S,i}))\right)$$

with a symmetric positive definite bandwidth matrix  $H \in \mathbb{R}^{n \times n}$  and a  $n$ -variate density function  $K : \mathbb{R}^n \rightarrow \mathbb{R}_{\geq 0}$  called kernel. We choose the the standard multivariate normal density function as kernel. This kernel can be written as the product  $K(x) = \prod_{i=1}^n \mathcal{K}(x_i)$ , where the univariate kernel  $\mathcal{K} : \mathbb{R} \rightarrow \mathbb{R}$  is the standard univariate normal density function. Let  $\sigma_{N,i}^2$  denote the positive sample variance of the  $i$ th variable. Moreover, let  $V_N$  denote the diagonal matrix  $V_N = \text{diag}(\sigma_{N,1}^2, \dots, \sigma_{N,n}^2)$ . As suggested in Gramacki (2018), we use the bandwidth matrix

$$H = h_y^2 V_N \quad \text{with} \quad h_y = \left(\frac{4}{(n+2)N}\right)^{\frac{1}{n+4}}. \tag{18}$$

This choice simplifies the estimator to the following form

$$\begin{aligned} \varrho_{p,N}(z) &= \frac{1}{N \prod_{j=1}^n h_y \sigma_{N,j}} \sum_{i=1}^N \prod_{j=1}^n \mathcal{K}\left(\frac{z_j - p_j(b^{S,i})}{h_y \sigma_{N,j}}\right) \\ &= \frac{1}{N \prod_{j=1}^n h_y \sigma_{N,j}} \sum_{i=1}^N \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_y \sigma_{N,j}}\right)^2\right). \end{aligned} \tag{19}$$

As mentioned in Scott (2015) and Wand and Jones (1993) such product kernels are recommended and adequate in practice. But, in some situations using only diagonal bandwidth matrices could be insufficient and then general full bandwidth matrices are needed, e.g., depending on the sample covariance matrix.

In order to show the convergence of the multivariate KDE (19), we transform the random vector  $p$  via  $y = V_N^{-1/2} p$ . Thus we get the transformed sampling set

$$\mathcal{Y}_B = \{y_1^S, \dots, y_N^S\} := \left\{ V_N^{-1/2} p(b^{S,1}), \dots, V_N^{-1/2} p(b^{S,N}) \right\} = V_N^{-\frac{1}{2}} \mathcal{P}_B.$$

For the approximation of the probability density function of the transformed variable we use the multivariate KDE with the previous settings adapted to the data  $\mathcal{Y}_B$ . Thus we get the bandwidth matrix  $H = h_y^2 I_{n \times n}$ , because the transformed data have unit variance. This leads to a KDE with one single bandwidth  $h_y$  given by

$$\varrho_{y,N}(z) = \frac{1}{Nh_y^n} \sum_{i=1}^N \prod_{j=1}^n \mathcal{K} \left( \frac{z_j - y_{ij}^S}{h_y} \right). \tag{20}$$

Due to  $(h_y + (Nh_y^n)^{-1}) \xrightarrow{N \rightarrow \infty} 0$   $\mathbb{P}$ -almost surely, we can apply *Chapter 6, Theorem 1* in Devroye and Györfi (1985) to the estimator  $\varrho_{y,N}$ . Thus, it holds

$$\|\varrho_{y,N} - \varrho_y\|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P}\text{-almost surely,}$$

where  $\varrho_y$  is the exact probability density function of the random variable  $y$ . This density function is given by  $\varrho_y(z) = \varrho_p(V_N^{1/2}z) |\det(V_N)|^{1/2}$  according to the transformation. There is also a similar relation between the estimators:  $\varrho_{y,N}(z) = \varrho_{p,N}(V_N^{1/2}z) \det(V_N)^{1/2}$ . Using the previous relations the  $L^1$ -convergence for the KDE  $\varrho_{p,N}$  follows:

$$\|\varrho_{p,N} - \varrho_p\|_{L^1} = \|\varrho_{y,N} - \varrho_y\|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P}\text{-almost surely.} \tag{21}$$

Applying *Scheffé’s lemma* (see Devroye and Györfi 1985) yields

$$\left| \int_{P_{\min}^{\max}} \varrho_p(z) dz - \int_{P_{\min}^{\max}} \varrho_{p,N}(z) dz \right| \leq \frac{1}{2} \|\varrho_{p,N} - \varrho_p\|_{L^1} \xrightarrow{N \rightarrow \infty} 0 \quad \mathbb{P}\text{-almost surely.} \tag{22}$$

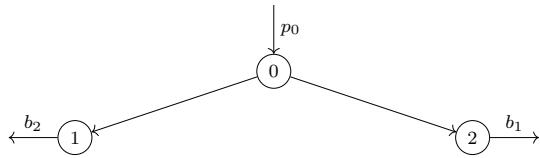
Thus, the integral of the estimator over the pressure bounds converges  $\mathbb{P}$ -almost surely to the probability  $\mathbb{P}(p \in P_{\min}^{\max})$ . Now, we can use this integral as an approximation for the probability  $\mathbb{P}(b \in M)$  in (17):

$$\mathbb{P}(b \in M) = \mathbb{P}(p \in P_{\min}^{\max}) \approx \int_{P_{\min}^{\max}} \varrho_{p,N}(z) dz =: \mathbb{P}_N(b \in M). \tag{23}$$

This multidimensional integral has some useful properties, which we will need when we derive the necessary optimality conditions. We want to mention again, that we can compute the sampling set  $\mathcal{P}_B$  analytically, because we know the solution of the stationary isothermal Euler equations.

**Example 2** This short example should illustrate the results of both approaches applied to the tree-structured graph with three nodes and two edges shown in Fig. 2. The values (without units) are given in Table 3.

**Fig. 2** Minimal tree-structured graph



**Table 3** Values for the example with two edges

$p_0$	$p^{\min}$	$p^{\max}$	$\mu$	$\Sigma$	$\phi$
60	$\begin{pmatrix} 40 \\ 30 \end{pmatrix}$	$\begin{pmatrix} 60 \\ 50 \end{pmatrix}$	$\begin{pmatrix} 4 \\ 4 \end{pmatrix}$	$\begin{pmatrix} 0.25 & 0 \\ 0 & 0.25 \end{pmatrix}$	$\begin{pmatrix} 100 \\ 100 \end{pmatrix}$

**Table 4** Results for the example with two edges

	Test 1 (%)	Test 2 (%)	Test 3 (%)	Test 4 (%)	Test 5 (%)	Test 6 (%)	Test 7 (%)	Test 8 (%)
MC	75.07	75.15	75.22	75.08	74.88	75.02	74.99	75.39
KDE	74.75	74.84	74.89	74.78	74.55	74.76	74.67	75.04
SRD	74.95							

From the inequalities in Lemma 1 we get

$$M = \left\{ b \in \mathbb{R}_{\geq 0}^2 \mid \begin{pmatrix} 0 \\ \sqrt{11} \end{pmatrix} \leq \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} \leq \begin{pmatrix} \sqrt{20} \\ \sqrt{27} \end{pmatrix} \right\}.$$

As in the last example, we compare both methods, the SRD and the KDE, with a classical Monte Carlo (MC) method. For the MC method and the KDE approach, we use the same sampling of  $1 \cdot 10^5$  points. The result for 8 tests is shown in Table 4.

The sampling of the SRD consists of  $1 \cdot 10^4$  points uniform distribution of the sphere  $S^1$ . Thus the SRD gives always the same (good) result of 74.95%, when rounded to 4 digits. Again, the MC method and the KDE approach are quite close. The mean probability in MC resp. KDE is 75.10% resp. 74.79% and the variance is 0.0235 resp. 0.0215. As in the example with one edge we provide the confidence intervals for confidence level 95% here. For the MC probability the confidence interval is [74.97%, 75.23%] and for the KDE probability it is [74.66%, 74.91%]. The (good) result of the SRD close to both intervals. Thus, also in the two dimensional case, the KDE approach is quite good for computing the desired probability. The computing time in every test is quite reasonable. The computation time of the MC method and the KDE approach needs less than one second, while the SRD needs almost two seconds, but the focus of the implementation was on correctness, not on efficiency. So the computing time of the implementation of course can be improved.

### 2.3 Stochastic optimization on stationary gas networks

In this subsection, we formulate necessary conditions for optimization problems with approximated probabilistic constraints. Both, MC and SRD, give algorithmic ways to compute the probability for a random load vector to be feasible. With a KDE approach, which provides a sufficiently good approximation of the probability (if the sample size is sufficiently large), we can get necessary optimality conditions for certain optimization problems with approximated probabilistic constraints. Define the set

$$\mathcal{P}_0 := \bigotimes_{i=1}^n [p_i^{\min}, \infty) \subseteq \mathbb{R}^n,$$

and let a function

$$f : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}, \quad (p^{\max}, p_0) \mapsto f(p^{\max}, p_0)$$

be given. For a probability level  $\alpha \in (0, 1)$  consider the optimization problems

$$\begin{cases} \min_{p^{\max} \in \mathcal{P}_0} & f(p_0, p^{\max}) \\ \text{s.t.} & \mathbb{P}(b \in M(p^{\max})) \geq \alpha, \end{cases} \tag{24}$$

and

$$\begin{cases} \min_{p_0 \in \mathbb{R}_{\geq 0}} & f(p_0, p^{\max}) \\ \text{s.t.} & \mathbb{P}(b \in M(p_0)) \geq \alpha. \end{cases} \tag{25}$$

Normally,  $\alpha$  is chosen large, s.t.  $\alpha$  is almost 1. As mentioned before, our aim here is to write down the necessary optimality conditions in an appropriate way. In Sect. 2.1 and Sect. 2.2 we stated the  $\mathbb{P}$ -almost surely convergence of the KDE to the exact probability density. Further the numerical examples provide good and accurate results. In fact, we formulate the optimality conditions for the approximated optimization problems

$$\begin{cases} \min_{p^{\max} \in \mathcal{P}_0} & f(p_0, p^{\max}) \\ \text{s.t.} & \mathbb{P}_N(b \in M(p^{\max})) \geq \alpha, \end{cases} \tag{26}$$

and

$$\begin{cases} \min_{p_0 \in \mathbb{R}_{\geq 0}} & f(p_0, p^{\max}) \\ \text{s.t.} & \mathbb{P}_N(b \in M(p_0)) \geq \alpha. \end{cases} \tag{27}$$

We mention again that due to the convergence results stated before, the approximated probabilistic constraint converges  $\mathbb{P}$ -almost surely to the exact probabilistic constraint for  $N \rightarrow \infty$ . We define

$$P_{\min}^{\max} := \bigotimes_{i=1}^n [p_i^{\min}, p_i^{\max}].$$



Our aim is now to integrate the kernel density estimator of the pressure at the nodes  $v_1, \dots, v_n$  over the pressure bounds. Since the following computations hold for the probability in (26) as well as in (27), we neglect the argument of  $M$  from here on. We have

$$\begin{aligned} \mathbb{P}_N(b \in M) &= \int_{p_{\min}^{\max}} \varrho_{p,N}(z) dz \\ &= \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \int_{p_{\min}^{\max}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz, \end{aligned}$$

with  $\varrho_{p,N}$  as in (19). Since  $P_{\min}^{\max}$  is a  $n$ -dimensional cuboid and  $\varrho_{p,N}(z)$  is continuous we can use *Fubini's Theorem*. Thus we have

$$\begin{aligned} \mathbb{P}_N(b \in M) &= \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \int_{p_1^{\min}}^{p_1^{\max}} \dots \int_{p_n^{\min}}^{p_n^{\max}} \prod_{j=1}^n \frac{1}{\sqrt{2\pi}} \\ &\quad \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz_n \dots dz_1, \end{aligned}$$

and as the density estimation of the pressure is a product of an exponential function in every dimension, we can exchange the integral and the product. It follows

$$\mathbb{P}_N(b \in M) = \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z_j - p_j(b^{S,i})}{h_j}\right)^2\right) dz_j.$$

We define

$$\varphi_{i,j} : \mathbb{R} \rightarrow \mathbb{R}, \quad \varphi_{i,j} : x \mapsto \left(\frac{x - p_j(b^{S,i})}{\sqrt{2}h_j}\right),$$

and we set  $\tau_{i,j} := \varphi_{i,j}(z_j)$  and use integration by substitution. Then with  $\varphi'_{i,j}(x) = (\sqrt{2} h_j)^{-1}$  we get

$$\begin{aligned} \mathbb{P}_N(b \in M) &= \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{i,j}^2(z_j)\right) dz_j \\ &= \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \prod_{j=1}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{2\pi}} \exp\left(-\tau_{i,j}^2\right) \sqrt{2} h_j d\tau_{i,j} \\ &= \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{\pi}} \exp\left(-\tau_{i,j}^2\right) d\tau_{i,j}. \end{aligned}$$

This formula contains the *Gauss error function* (see e.g. Andrews 1998):

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x \exp(-t^2) dt. \tag{28}$$

We insert the Gauss error function in the previous integral term and we obtain

$$\begin{aligned} \mathbb{P}_N(b \in M) &= \int_{p_{\min}^{\max}} \varrho_{p,N}(z) dz \\ &= \frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \prod_{j=1}^n \left[ \operatorname{erf}\left(\varphi_{i,j}(p_j^{\max})\right) - \operatorname{erf}\left(\varphi_{i,j}(p_j^{\min})\right) \right]. \end{aligned} \tag{29}$$

Now we consider problem (26). For  $\alpha \in (0, 1)$  we define

$$g_\alpha : \mathbb{R}^n \rightarrow \mathbb{R}, \quad p^{\max} \mapsto \alpha - \mathbb{P}_N(b \in M(p^{\max})).$$

Thus we have

$$g_\alpha(p^{\max}) = \alpha - \frac{1}{N} \sum_{i=1}^N \prod_{j=1}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{\pi}} \exp\left(-\tau_{i,j}^2\right) d\tau_{i,j}.$$

We compute the partial derivatives of  $g_\alpha$ . For  $k \in \{1, \dots, n\}$  we have

$$\begin{aligned} \frac{\partial}{\partial p_k^{\max}} g_\alpha(p^{\max}) &= -\frac{1}{N} \sum_{i=1}^N \left[ \prod_{j=1, j \neq k}^n \int_{\varphi_{i,j}(p_j^{\min})}^{\varphi_{i,j}(p_j^{\max})} \frac{1}{\sqrt{\pi}} \exp\left(-\tau_{i,j}^2\right) d\tau_{i,j} \right. \\ &\quad \left. \cdot \frac{1}{\sqrt{\pi}} \exp\left(-\varphi_{i,k}^2(p_k^{\max})\right) \frac{1}{\sqrt{2}h_k} \right], \end{aligned}$$

and with the Gauss error function (28) it follows

$$\begin{aligned} \frac{\partial}{\partial p_k^{\max}} g_\alpha(p^{\max}) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \left[ \prod_{j=1, j \neq k}^n \left[ \operatorname{erf}\left(\varphi_{i,j}(p_j^{\max})\right) - \operatorname{erf}\left(\varphi_{i,j}(p_j^{\min})\right) \right] \right. \\ &\quad \left. \cdot \frac{\sqrt{2}}{\sqrt{\pi}h_k} \exp\left(-\varphi_{i,k}^2(p_k^{\max})\right) \right]. \end{aligned} \tag{30}$$

Then, the  $k$ -th component of the gradient  $\nabla g_\alpha(p^{\max}) \in \mathbb{R}^n$  is given by (30). Note that for  $b^{S,1}, \dots, b^{S,N} \in \mathbb{R}^n$  ( $N > 1$ ) and  $p_i^{\max} > p_i^{\min}$  ( $i = 1, \dots, n$ ), the partial derivatives in (30) are negative for all  $p^{\max} \in \mathbb{R}^n$ .

**Remark 2** Since (26) has only one constraint, the *linear independent constraint qualification* (LICQ) holds for every  $\tilde{p}^{\max} > p^{\min}$  (componentwise) with  $g_\alpha(\tilde{p}^{\max}) = 0$ .

Now we can state necessary optimality conditions for the optimization problem (26):

**Corollary 1** *Let  $p^{*,\max} \in \mathbb{R}^n$  be a (local) optimal solution of (26). Since the LICQ holds in  $p^{*,\max}$ , there exists a multiplier  $\mu^* \geq 0$ , s.t.*

$$\begin{aligned} \nabla_{p^{\max}} f(p^{*,\max}, p_0) + \mu^* \nabla g_\alpha(p^{*,\max}) &= 0, \\ g_\alpha(p^{*,\max}) &\leq 0, \\ \mu^* g_\alpha(p^{*,\max}) &= 0. \end{aligned}$$

Thus,  $(p^{*,\max}, \mu^*) \in \mathbb{R}^{n+1}$  is a Karush–Kuhn–Tucker point.

Now we consider problem (27). We slightly change the notation to add the explicit dependence on  $p_0$ , so we write  $p(b^{S,i}, p_0)$  ( $i = 1, \dots, N$ ) instead of  $p(b^{S,i})$  for the samples in the set  $\mathcal{P}_B$ . We redefine the function  $\varphi_{ij}$  ( $i = 1, \dots, N, j = 1, \dots, n$ ) as

$$\varphi_{ij} : \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R} \quad (x, y) \mapsto \left( \frac{x - p_j(b^{S,i}, y)}{\sqrt{2}h_j} \right),$$

and we define the constraint of (27) as

$$\gamma_\alpha : \mathbb{R} \rightarrow \mathbb{R} \quad p_0 \mapsto \alpha - \mathbb{P}_N(b \in M(p_0)).$$

Thus we have

$$\gamma_\alpha(p_0) = \alpha - \frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{ij}^2(z_j, p_0)\right) dz_j.$$

For the derivative with respect to  $p_0$ , it follows

$$\begin{aligned} \frac{d}{dp_0} \gamma_\alpha(p_0) &= -\frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \frac{d}{dp_0} \left( \prod_{j=1}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{ij}^2(z_j, p_0)\right) dz_j \right) \\ &= -\frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{ij}^2(z_j, p_0)\right) dz_j \\ &\quad \cdot \frac{d}{dp_0} \int_{p_k^{\min}}^{p_k^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{i,k}^2(z_k, p_0)\right) dz_k. \end{aligned}$$

Due to the *dominated convergence theorem* we can exchange the integral and the derivative, thus we have

$$\begin{aligned} \frac{d}{dp_0} \gamma_\alpha(p_0) &= -\frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{i,j}^2(z_j, p_0)\right) dz_j \\ &\quad \cdot \int_{p_k^{\min}}^{p_k^{\max}} \frac{d}{dp_0} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{i,k}^2(z_k, p_0)\right) dz_k \\ &= -\frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \int_{p_j^{\min}}^{p_j^{\max}} \frac{1}{\sqrt{2\pi}} \exp\left(-\varphi_{i,j}^2(z_j, p_0)\right) dz_j \\ &\quad \cdot \int_{p_k^{\min}}^{p_k^{\max}} \frac{1}{\sqrt{\pi} h_k} \exp\left(-\varphi_{i,k}^2(z_k, p_0)\right) \varphi_{i,k}(z_k, p_0) \frac{d}{dp_0} p_k(b_i, p_0) dz_k. \end{aligned}$$

We define  $\tau_{i,j} := \varphi_{i,j}(z_j, p_0)$  and since  $p_k(b_i, p_0)$  is independent of  $z_k$ , it follows

$$\begin{aligned} \frac{d}{dp_0} \gamma_\alpha(p_0) &= -\frac{1}{N \prod_{j=1}^n h_j} \sum_{i=1}^N \sum_{k=1}^n \prod_{j=1, j \neq k}^n \int_{\varphi_{i,j}(p_j^{\min}, p_0)}^{\varphi_{i,j}(p_j^{\max}, p_0)} \frac{h_j}{\sqrt{\pi}} \exp\left(-\tau_{i,j}^2\right) d\tau_{i,j} \\ &\quad \cdot \frac{d}{dp_0} p_k(b_i, p_0) \int_{\varphi_{i,k}(p_k^{\min}, p_0)}^{\varphi_{i,k}(p_k^{\max}, p_0)} \frac{\sqrt{2}}{\sqrt{\pi}} \exp\left(-\tau_{i,k}^2\right) \tau_{i,k} d\tau_{i,k}. \end{aligned}$$

The second integral can be solved analytically and yields:

$$\left[ \frac{d}{dx} \left( -\frac{1}{\sqrt{2\pi}} \exp(-x^2) \right) = \frac{\sqrt{2}}{\sqrt{\pi}} \exp(-x^2) x \right].$$

Hence, we have

$$\begin{aligned} \frac{d}{dp_0} \gamma_\alpha(p_0) &= -\frac{1}{N} \frac{1}{2^n} \sum_{i=1}^N \sum_{k=1}^n \left[ \prod_{j=1, j \neq k}^n \left[ \operatorname{erf} \left( \varphi_{i,j}(p_j^{\max}, p_0) \right) \right. \right. \\ &\quad \left. \left. - \operatorname{erf} \left( \varphi_{i,j}(p_j^{\min}, p_0) \right) \right] \right. \\ &\quad \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_k} \frac{d}{dp_0} p_k(b_i, p_0) \left[ -\exp \left( -\varphi_{i,k}^2(p_k^{\max}, p_0) \right) \right. \\ &\quad \left. \left. + \exp \left( -\varphi_{i,k}^2(p_k^{\min}, p_0) \right) \right] \right]. \end{aligned}$$

In the setting of the stationary gas networks it is true that

$$\frac{d}{dp_0} p_k(b^{S,i}, p_0) = \frac{p_0}{p_k(b^{S,i}, p_0)}.$$

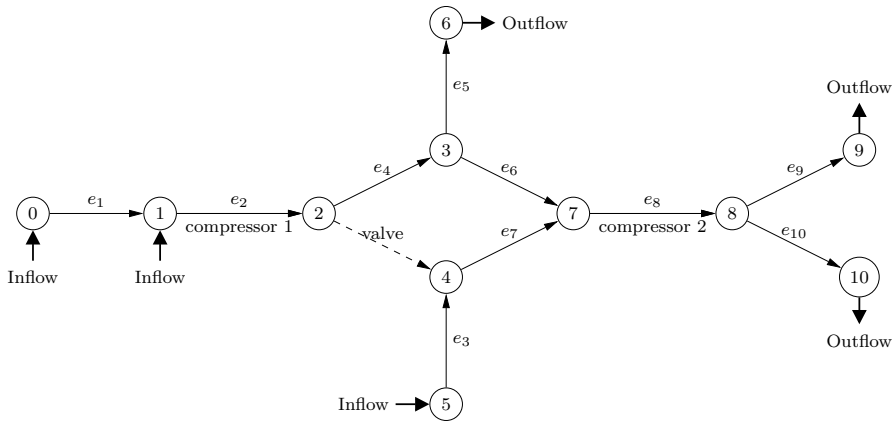


Fig. 3 A scheme of the GasLib-11

**Corollary 2** Let  $p_0^* \in \mathbb{R}$  be a (local) optimal solution of (27). Since the LICQ holds in  $p_0^*$  (cf. Remark 2), then there exists a multiplier  $\mu^* \geq 0$ , s.t.

$$\begin{aligned} \nabla_{p_0} f(p^{\max}, p_0^*) + \mu^* \nabla \gamma_\alpha(p_0^*) &= 0, \\ \gamma_\alpha(p_0^*) &\leq 0, \\ \mu^* \gamma_\alpha(p_0^*) &= 0. \end{aligned}$$

Thus the point  $(p_0^*, \mu^*) \in \mathbb{R}^2$  is a Karush–Kuhn–Tucker point.

If the objective function  $f$  is strictly convex and the feasible set is convex, then all necessary conditions stated here are sufficient. In this case, Corollary 1 and Corollary 2 give a characterization of the (unique) optimal solution of the approximated problems (26) and (27).

**Remark 3** The question whether the solutions of the approximated problems (26) and (27) converge to the solutions of (24) and (25), is out of scope of this work but all numerical results and tests hypothesize the convergence if the sample size goes to infinity.

### 2.4 Application to a realistic gas network

The *GasLib*<sup>2</sup> promotes research on gas networks by providing realistic benchmark instances. We use the *GasLib-11* as a meaningful example. A scheme of the GasLib-11 is shown in Fig. 3 and more information can be found at <http://gaslib.zib.de/testData.html>.

<sup>2</sup> <http://gaslib.zib.de/>.

**Table 5** Values for the GasLib-11

$p_0$	$p^{\min}$	$\mu$	$\Sigma$
$\begin{pmatrix} 60 \\ 58 \\ 60 \end{pmatrix}$	$\begin{pmatrix} 40 \\ 40 \\ 40 \end{pmatrix}$	$\begin{pmatrix} 20 \\ 15 \\ 18 \end{pmatrix}$	$\begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

The GasLib-11 consists in 11 nodes and 11 edges. Two of the edges represent compressor stations and one edge represents a valve. Compressor stations counteract the pressure loss caused by friction in the pipes. Here the compressor stations are modeled as frictionless pipes, that satisfy the equation

$$\frac{(p_{\text{in}})^2}{(p_{\text{out}})^2} = u,$$

as it is done in Gugat and Schuster (2018). This model for compressor stations is also suggested in Koch et al. (2015), where one gets an excellent overview about the details on how to model a compressor station. For our system we assume that the compressor at edge  $e_2$  is switched off, i.e.,  $u_{e_2} = 1$  (so this edge can be modeled as frictionless pipe) and that the compressor station at edge  $e_8$  increases the pressure by 20%, i.e.,  $u_{e_8} = 1.2$ . The valve is also modeled as a frictionless pipe in which gas can be transported if the valve is opened and which cannot be used for gas transportation if the valve is closed. We assume that the valve is closed, so this edge vanishes in our implementation. For the remaining edges ( $e_1, e_3, e_4, e_5, e_6, e_7, e_9, e_{10}$ ) we assume  $\phi_{e_i} = 1$ .

Further gas enters the network at the nodes  $v_0, v_1, v_5$  and is transported through the network to the nodes  $v_6, v_9$  and  $v_{10}$ . The values for the inlet pressure  $p_0 = [p_{v_0}, p_{v_1}, p_{v_5}]$ , the lower pressure bound  $p^{\min} = [p_{v_6}^{\min}, p_{v_9}^{\min}, p_{v_{10}}^{\min}]$  and the probability distribution at the exit nodes  $\mu = [\mu_{v_6}, \mu_{v_9}, \mu_{v_{10}}]$  and  $\Sigma$  with  $\text{diag}(\Sigma) = [\sigma_{v_6}^2, \sigma_{v_9}^2, \sigma_{v_{10}}^2]$  are given in Table 5.

Consider the linear function

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad f : p^{\max} \mapsto c^T p^{\max},$$

with  $c = \mathbb{1}_3$ . We first solve the deterministic problem

$$\begin{aligned} \min_{p^{\max} \in \mathcal{P}_0} & f(p^{\max}) \\ \text{s.t.} & b \in M(p^{\max}), \end{aligned} \tag{31}$$

where the load vector  $b$  is given by the mean value  $\mu$ . We use default setting of the MATLAB®-routine *fmincon* to solve (31), which is an interior-point algorithm. It returns

**Table 6** Probability  $\mathbb{P}(b \in M(p_{\text{deter}}^{\text{max}}))$  for the optimal deterministic upper pressure bounds

	Test 1 (%)	Test 2 (%)	Test 3 (%)	Test 4 (%)	Test 5 (%)	Test 6 (%)	Test 7 (%)	Test 8 (%)
MC	36.02	35.66	35.91	35.86	35.34	35.48	35.98	35.90
KDE	35.72	35.41	35.48	35.39	34.92	35.08	35.75	35.47

**Table 7** Stochastic optimal upper pressure bounds  $p_{\text{stoch}}^{\text{max}}$

Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8
$\begin{bmatrix} 47.51 \\ 53.33 \\ 52.44 \end{bmatrix}$	$\begin{bmatrix} 47.51 \\ 53.34 \\ 52.45 \end{bmatrix}$	$\begin{bmatrix} 47.52 \\ 53.33 \\ 52.46 \end{bmatrix}$	$\begin{bmatrix} 47.52 \\ 53.34 \\ 52.46 \end{bmatrix}$	$\begin{bmatrix} 47.51 \\ 53.35 \\ 52.46 \end{bmatrix}$	$\begin{bmatrix} 47.51 \\ 53.35 \\ 52.45 \end{bmatrix}$	$\begin{bmatrix} 47.53 \\ 53.33 \\ 52.44 \end{bmatrix}$	$\begin{bmatrix} 47.52 \\ 53.33 \\ 52.46 \end{bmatrix}$

$$p_{\text{det}}^{\text{max}} = \begin{bmatrix} 46.10 \\ 52.04 \\ 51.08 \end{bmatrix},$$

as optimal deterministic solution, i.e., as the lowest upper pressure bound for the nodes  $v_6, v_9$  and  $v_{10}$ . Now we consider the uncertain outflow at the nodes  $v_6, v_9$  and  $v_{10}$ . We compute the probability for a random load vector to be feasible with respect to the optimal deterministic pressure bounds by using (29). The probability  $\mathbb{P}(b \in M(p_{\text{det}}^{\text{max}}))$  for 8 tests (each with  $1 \cdot 10^5$  samples) is shown in Table 6. The probabilities for the deterministic optimal pressure bounds are unsatisfactory. The mean MC probability is 35.77% and the mean KDE probability is 35.40%. For a confidence level of 95% the confidence interval for the MC probability is [35.56%, 35.98%] and the confidence interval for the KDE probability is [35.16%, 35.64%]. So if the boundary data (i.e., the gas demand) is uncertain, the optimal deterministic pressure bounds are unserviceable in the sense that these bounds do not provide a good operating gas network for uncertain gas demand.

Next we consider the probabilistic constrained optimization problem (26). We set

$$\alpha := 0.75.$$

For arbitrary starting points the MATLAB<sup>®</sup>-routine *fmincon* sometimes struggles with finding a solution of (26) but the optimal deterministic solution appears to be a good choice for the starting point of the routine. The results of 8 Tests with  $1 \cdot 10^5$  sampling points, i.e., the optimal upper pressure bounds  $p^{\text{max}}$  at the nodes  $v_6, v_9$  and  $v_{10}$ , are shown in Table 7. In 8 more Tests we solve (26) by using Corollary 1. The points that satisfy the necessary optimality conditions are always good candidates for the optimal solution. To be more precise on that we solve the following optimization problem using again *fmincon*:

$$\begin{aligned}
 & \min_{p^{\max}, \mu} f(p^{\max}) \\
 \text{s.t.} \quad & \nabla_{p^{\max}} f(p^{\max}, p_0) + \mu \nabla g_{\alpha}(p^{\max}) = 0, \\
 & g_{\alpha}(p^{\max}) \leq 0, \\
 & \mu g_{\alpha}(p^{\max}) = 0, \\
 & \mu \geq 0.
 \end{aligned}$$

Here, we get values that are almost equal to the optimal solution stated in Table 7, they vary in a range of  $1 \cdot 10^{-6}$ . From this fact one could expect that the necessary optimality conditions stated in Corollary 1 are sufficient but we do not analyze this here.

One can see, that all results are almost equal. The optimal upper pressure bounds of the stochastic optimization problem (26) are slightly larger than the optimal upper pressure bounds of the deterministic optimization problem (31) but the probability for a random load vector to be feasible is 75%, as it is required in the probabilistic constraint. The computation time for a single test was about 20 min, where the optimization time was much less than a second. The 20 min were almost only needed to solve the GasLib-11  $1 \cdot 10^5$  times. The solution of the necessary optimality conditions needed a little bit more time than the direct solution using *fmincon* and (29), but solving the necessary optimality conditions leads to a solution more often even if the starting point of *fmincon* is badly chosen.

### 3 Dynamic flow networks

In this section, we extend the in Sect. 2 introduced methods to dynamic systems. We first discuss probabilistic constraints in a dynamic setting and time dependent random boundary data. Then we consider a model, which we can solve analytically to use the idea of the SRD for dynamic systems and compare it with the idea of the KDE. Last we also formulate necessary optimality conditions for optimization problems with probabilistic constraints in a dynamic setting.

#### 3.1 Time dependent probabilistic constraints and random boundary data

In Sect. 2, we computed the probability for a random vector to be feasible. So if we fix a point in time  $t^* \in [0, T]$ , we can use a similar procedure. But if we do not fix a point in time, we need an extension to the probabilistic constraint how it has to be understood for a time period. For a time dependent uncertain boundary function  $b(t)$  and a (time dependent) feasible set  $M(t)$ , a possible formulation (the one, that we will use later) for the probabilistic constraint is

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]) \geq \alpha. \tag{32}$$

That means, we want to guarantee, that a percentage  $\alpha$  of all possible random boundary functions (in an appropriate probability space  $(\Omega, \mathcal{A}, \mathcal{P})$ ) is feasible in every point in time  $t \in [0, T]$ . This is a very strong condition. In fact (32) is a so-called



probust constraint, which means it is a mix between a probabilistic constraint and a robust constraint. This class of constraints has been developed recently and is currently of big interest in research (see e.g. van Ackooij et al. 2020; Adelhütte et al. 2020). Another possibility is

$$\mathbb{P}(b \in M(t)) \geq \alpha \quad \forall t \in [0, T],$$

which means, that a random boundary function must be feasible with a percentage of  $\alpha$  in every time point  $t \in [0, T]$ . For our applications, this might not make sense, because we want to guarantee that problems for a gas consumer only occur in worst case scenarios. This probabilistic constraint only states, that even in these worst case scenarios, the problems for a consumer stay small, but these small problems can occur in every point in time. Probabilistic constraints of this type have been discussed in van Ackooij et al. (2016). A third possibility for the time dependent probabilistic constraint is an ergodic formulation:

$$\frac{1}{T} \int_0^T \mathbb{P}(b \in M(t)) dt \geq \alpha.$$

That means the ergodic probability during the time period  $[0, T]$  must be large enough. This formulation might make sense in other applications, but not for the flow problems which are considered here (with the same argument as before). Thus we use the formulation (32) for time dependent probabilistic constraints.

Next we discuss the uncertain boundary data. For a random boundary data, we use a representation as Fourier series as it is done in Farshbaf-Shaker et al. (2020). So for a deterministic boundary function  $b_D : [0, T] \rightarrow \mathbb{R}$  with  $b_D(0) = 0$  and for  $m = 0, 1, 2, \dots$ , we define the orthonormal series

$$\psi_m(t) := \frac{\sqrt{2}}{\sqrt{T}} \sin\left(\left(\frac{\pi}{2} + m\pi\right)\frac{t}{T}\right), \tag{33}$$

and the coefficients

$$a_m^0 := \int_0^T b_D(t)\psi_m(t)dt. \tag{34}$$

Then we can write the boundary function  $b_D(t)$  in a series representation

$$b_D(t) = \sum_{m=0}^{\infty} a_m^0 \psi_m(t). \tag{35}$$

Now for  $m \in \mathbb{N}_0$ , we consider the Gaussian distributed random variables  $a_m \sim \mathcal{N}(1, \sigma^2)$  for a mean value 1 and a standard deviation  $\sigma \in \mathbb{R}_+$  on an appropriate probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Then we consider the random boundary data

$$b(t, \omega) = \sum_{m=0}^{\infty} a_m(\omega)a_m^0 \psi_m(t). \tag{36}$$

Since the random variables  $a_m$  are all independent and identically distributed, we can use the fact that for  $b_D \in L^2(0, T)$ , we have also  $b \in L^2(0, T)$   $\mathbb{P}$ -almost surely. In Marcus and Pisier (1981), Hill (2012) and Farshbaf-Shaker et al. (2020) the authors state that this approach even guarantees better regularity and it also holds for a larger class of random variables.

**Remark 4** For the numerical tests, we truncate the Fourier series after  $N_F \in \mathbb{N}$  terms. Thus we use

$$b_D^{N_F}(t) = \sum_{m=0}^{N_F} a_m^0 \psi_m(t)$$

instead of (35) for the implementation of  $b_D$ . Because it holds

$$\lim_{N_F \rightarrow \infty} b_D^{N_F} = b_D,$$

this truncated Fourier series is a sufficient good expression for  $b_D(t)$  for  $N_F$  large enough. The question how to choose  $N_F$  strongly depends on the data  $b_D$  and on the desired accuracy of the Fourier series. In general one has to guarantee, that the truncation error is small. One criteria for finding a sufficient large number  $N_F$  is to state a bound for the  $L^2$ -truncation error. For  $\vartheta \in (0, 1)$  we require that  $N_F$  is chosen large enough, s.t.

$$\|b_D(t) - b_D^{N_F}(t)\|_{L^2}^2 \leq \vartheta \|b_D(t)\|_{L^2}^2.$$

Due to the convergence of the Fourier series it is always possible to find  $N_F$  large enough, s.t. the  $L^2$ -error bound is satisfied for all  $\vartheta \in (0, 1)$ . Another criteria for an sufficient large number  $N_F$  is a bound for the  $L^\infty$ -truncation error. For  $\vartheta \in (0, 1)$  we require that  $N_F$  is chosen large enough s.t.

$$\|b_D(t) - b_D^{N_F}(t)\|_{L^\infty} \leq \vartheta \left[ \sup_{\tau \in [0, T]} b_D(\tau) - \inf_{\tau \in [0, T]} b_D(\tau) \right].$$

The Gibbs phenomenon might cause problems regarding the  $L^\infty$ -error if  $b_D$  contains discontinuities (see e.g. Thompson 1992), so in this case the  $L^2$ -error is the better choice. For continuous functions  $b_D$  both estimates can be used to find a sufficient large number  $N_F$ . Usually  $\vartheta$  is chosen small, even close to zero, i.e.,  $\vartheta = 1\%$  or  $\vartheta = 0.1\%$ , but this choice depends on the operator. Similarly we use

$$b^{N_F}(t, \omega) := \sum_{m=0}^{N_F} a_m(\omega) a_m^0 \psi_m(t),$$

with  $\mathcal{N}(1, \sigma^2)$ -distributed random variables  $a_0, \dots, a_{N_F}$  instead of (36) as random boundary data for the implementation.



Fig. 4 Linear graph with  $n + 1$  nodes

This representation of a random boundary function as Fourier series requires  $b_D(0) = 0$ . If this is not given, i.e., if  $b_D(0) \neq 0$ , one can shift  $b_D$  by  $b_D(0)$ , get the representation as Fourier series and shift this Fourier representation back by  $b_D(0)$ , as we do later in Example 3.

### 3.2 Deterministic loads for a scalar PDE

For  $(t, x) \in [0, T] \times [0, L]$  and constants  $d < 0, m \leq 0$ , we consider the deterministic scalar linear PDE with initial condition and boundary condition

$$\begin{cases} r_t(t, x) + dr_x(t, x) = mr(t, x), \\ r(0, x) = r_0(x), \\ r(t, L) = b(t). \end{cases} \tag{37}$$

Here,  $r$  is the concentration of the contamination. The term  $dr_x$  describes the transport of the contamination according to the water flow and the term  $mr$  describes the decay of the contamination. This equation models the flow of contamination in water along a pipe or in a network (see Gugat 2012; Fügenschuh et al. 2007). Assume  $C^0$ -compatibility between the initial and the boundary condition, which is  $r_0(L) = b(0)$ . We will specify the boundary condition later. We state  $b(t) \geq 0$ , if the water gets polluted and  $b(t) < 0$  if the water gets cleaned.

For initial data  $r_0 \in L^2(0, L)$  and boundary data  $b \in L^2(0, T)$ , a solution of (37) is in  $C([0, T], L^2(0, L))$  and it is analytically given by

$$r(t, x) = \begin{cases} \exp(mt) r_0(x - dt) & \text{if } x \leq L + dt, \\ \exp\left(m\frac{x-L}{d}\right)b\left(t - \frac{x-L}{d}\right) & \text{if } x > L + dt. \end{cases}$$

Now, we consider a linear graph  $G = (\mathcal{V}, \mathcal{E})$  with vertex set  $\mathcal{V} := \{v_0, \dots, v_n\}$  and the set of edges  $\mathcal{E} = \{e_1, \dots, e_n\} \subseteq \mathcal{V} \times \mathcal{V}$ . Every edge  $e_i \in \mathcal{E}$  has a positive length  $L_i$ . Linear means here, that every node has at most one outgoing edge (see Fig. 4). For a formal definition see Gugat et al. (2020).

Equation (37) holds on every edge. We assume conservation of the flow at the nodes, i.e.

$$r_i(t, L_i) = r_{i+1}(t, 0) + b_i(t) \quad \forall i = 1, \dots, n - 1 \quad \forall t \in [0, T],$$

where  $r_i$  denotes the contamination concentration on edge  $e_i$  and  $b_i$  denotes the boundary data at node  $v_i$ .

For constants  $d_k < 0, m_k \leq 0$ , the full model can be written as follows (with  $(t, x) \in [0, T] \times [0, L_k]$  on the  $k$ -th edge and  $k = 1, \dots, n$ ):

$$\begin{cases} r_k(0, x) = r_{k,0}(x), \\ (r_k)_t(t, x) + d_k(r_k)_x(t, x) = m_k r_k(t, x), \\ r_k(t, L_k) = \begin{cases} b_n(t) & \text{if } k = n, \\ r_{k+1}(t, 0) + b_k(t) & \text{else.} \end{cases} \end{cases} \tag{38}$$

The model can be interpreted as follows: The graph represents a water network, where the water is contaminated at the nodes  $v_i$  ( $i = 1, \dots, n$ ). This contamination is distributed in the graph in a negative way (due to  $d_k < 0$ ). We want to know the contamination rate at node  $v_0$ . Later, we assume the contamination rate at the nodes to be Gaussian distributed. Then for a time  $t^* \in [0, T]$ , we want to compute the probability for the contamination rate at node  $v_0$  to fulfill box constraints using both, the SRD and the KDE. In both cases, we also consider the general time dependent chance constraints discussed before. The next theorem states an analytical solution for the model (38).

**Theorem 2** *Let initial states  $r_{k,0} \in H^1(0, L_k)$  and boundary conditions  $b_k \in H^1(0, T)$  for  $k = 1, \dots, n$  be given. Then the solution of the  $k$ -th edge of (38) is in  $C^1([0, T], H^1(0, L_k))$  and it is analytically given for  $x \geq d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  by*

$$r_k(t, x) = \sum_{i=k}^n \exp\left(m_k \frac{x}{d_k} - \sum_{j=k}^i m_j \frac{L_j}{d_j}\right) b_i\left(t - \frac{x}{d_k} + \sum_{j=k}^i \frac{L_j}{d_j}\right).$$

For the case  $x < d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  the solution is given by

$$\begin{aligned} r_k(t, x) &= \sum_{i=k}^{\ell-1} \exp\left(m_k \frac{x}{d_k} - \sum_{j=k}^i m_j \frac{L_j}{d_j}\right) b_i\left(t - \frac{x}{d_k} + \sum_{j=k}^i \frac{L_j}{d_j}\right) \\ &\quad + \exp\left(m_\ell t - (m_\ell - m_k) \frac{x}{d_k} + \sum_{j=k}^{\ell-1} (m_\ell - m_j) \frac{L_j}{d_j}\right) r_{\ell,0} \\ &\quad \left(-d_\ell t + d_\ell \frac{x}{d_k} - d_\ell \sum_{j=k}^{\ell-1} \frac{L_j}{d_j}\right) \end{aligned}$$

for  $d_k t + d_k \sum_{j=k}^{\ell-1} \frac{L_j}{d_j} \leq x < d_k t + d_k \sum_{j=k}^{\ell} \frac{L_j}{d_j}$  and  $\ell \in \{k, \dots, n\}$  (with  $d_k < 0$  for  $k = 1, \dots, n$ ).

**Remark 5** We set  $t^* := \sum_{j=1}^n \frac{L_j}{|d_j|}$ . For points in time  $t \leq t^*$  the solution can depend explicitly on the initial condition. If we assume that  $|d_i|$  are absolute velocities and  $L_i$  are lengths, the information from the right boundary needs  $\frac{L_n}{|d_n|}$  seconds to travel along the  $n$ -th edge. Then after  $\frac{L_n}{|d_n|}$  seconds, the solution of the  $n$ -th edge only depends on the boundary data, but the solution of edge  $n - 1$  can still depend on the initial condition of the  $n$ -th edge. A scheme of characteristics for a graph with 4 edges is shown in Fig. 5.

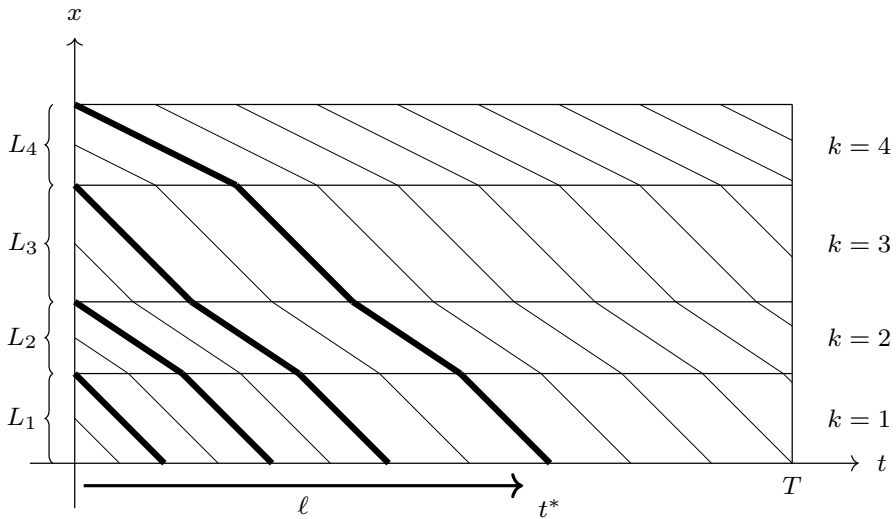


Fig. 5 Characteristics of (38) on a graph with 4 edges

**Remark 6** In the analytical solution of (38) we only distinguish if the solution depends on the initial or the boundary condition, depending on the time and the location in the pipe. So for the solution of edge 4 in Fig. 5, we distinguish between

$$x \geq L_4 + d_4t \quad \text{and} \quad x < L_4 + d_4t.$$

That means, at the beginning of edge 4 (for  $x = 0$ ), for  $0 \leq t \leq -\frac{L_4}{d_4}$ , the solution of edge 4 depends on the initial condition of edge 4. From this, it follows, that the solution of edge 3 in Fig. 5 depends only on the initial condition of edge 3 for

$$x < L_3 + d_3t.$$

It depends on the boundary conditions of edge 3 and the initial condition of edge 4 (due to the coupling condition) for

$$L_3 + d_3t \leq x < L_3 + d_3t + d_3 \frac{L_4}{d_4},$$

and it depends on the boundary conditions of edge 3 and edge 4 for

$$x \geq L_3 + d_3t + d_3 \frac{L_4}{d_4}.$$

This leads to the differentiation in Theorem 2 in the case  $x < d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  ( $k = 1, \dots, n$ ).

**Remark 7** With the result of Theorem 2 one can also derive analytical solutions of (38) for tree-structured graphs, but one has to take into account, that the flow at the

end of an edge (due to coupling conditions) can depend on more than one outgoing edges. That means the solution on a tree-structured graph is basically the sum over all paths of the solution stated in Theorem 2.

**Proof of Theorem 2** We define the following functions:

$$\begin{aligned} \alpha_{k,i}(x) &:= m_k \frac{x}{d_k} - \sum_{j=k}^i m_j \frac{L_j}{d_j}, \\ \beta_{k,i}(t, x) &:= t - \frac{x}{d_k} + \sum_{j=k}^i \frac{L_j}{d_j}, \\ \gamma_{k,\ell}(t, x) &:= m_\ell t - (m_\ell - m_k) \frac{x}{d_k} + \sum_{j=k}^{\ell-1} (m_\ell - m_j) \frac{L_j}{d_j}, \\ \delta_{k,\ell}(t, x) &:= -d_\ell t + d_\ell \frac{x}{d_k} - d_\ell \sum_{j=k}^{\ell-1} \frac{L_j}{d_j}. \end{aligned}$$

We consider the  $k$ -th edge in a linear graph with  $n$  edges ( $k \in \{1, \dots, n\}$ ).

*Step 1: The PDE in (38) holds*

For  $x \geq d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  we have

$$\frac{\partial}{\partial t} r_k(t, x) = \sum_{i=k}^n \exp(\alpha_{k,i}(x)) b'_i(\beta_{k,i}(t, x))$$

and

$$\begin{aligned} d_k \frac{\partial}{\partial x} r_k(t, x) &= d_k \sum_{i=k}^n \exp(\alpha_{k,i}(x)) \frac{m_k}{d_k} b_i(\beta_{k,i}(t, x)) \\ &\quad + d_k \sum_{i=k}^n \exp(\alpha_{k,i}(x)) b'_i(\beta_{k,i}(t, x)) \left(-\frac{1}{d_k}\right). \end{aligned}$$

Thus it follows

$$\frac{\partial}{\partial t} r_k(t, x) + d_k \frac{\partial}{\partial x} r_k(t, x) = m_k \sum_{i=k}^n \exp(\alpha_{k,i}(x)) b_i(\beta_{k,i}(t, x)) = m_k r_k(t, x).$$

So the PDE in the system (38) holds in the marked area in Fig. 6a. For  $x < d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  and  $\ell \in \{k, \dots, n\}$ , we have

$$\begin{aligned} \frac{\partial}{\partial t} r_k(t, x) &= \sum_{i=k}^{\ell-1} \exp(\alpha_{k,i}(x)) b'_i(\beta_{k,i}(t, x)) + \exp(\gamma_{k,\ell}(t, x)) m_\ell r_{\ell,0}(\delta_{k,\ell}(t, x)) \\ &\quad + \exp(\gamma_{k,\ell}(t, x)) r'_{\ell,0}(\delta_{k,\ell}(t, x)) (-d_\ell) \end{aligned}$$

and

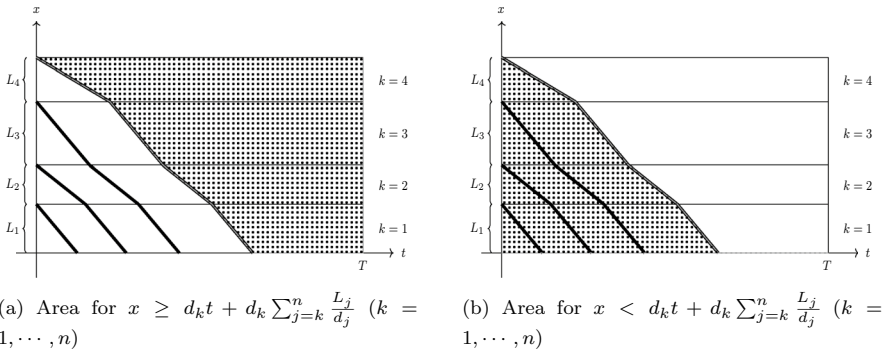


Fig. 6 Areas in which the PDE of system (38) holds

$$\begin{aligned}
 d_k \frac{\partial}{\partial x} &= d_k \sum_{i=k}^{\ell-1} \exp(\alpha_{k,i}(x)) \frac{m_k}{d_k} b_i(\beta_{k,i}(t, x)) \\
 &+ d_k \sum_{i=k}^{\ell-1} \exp(\alpha_{k,i}(x)) b'_i(\beta_{k,i}(t, x)) \left(-\frac{1}{d_k}\right) \\
 &+ d_k \exp(\gamma_{k,\ell}(t, x)) \left(\frac{-m_\ell + m_k}{d_k}\right) r_{\ell,0}(\delta_{k,\ell}(t, x)) \\
 &+ d_k \exp(\gamma_{k,\ell}(t, x)) r'_{\ell,0}(\delta_{k,\ell}(t, x)) \frac{d_\ell}{d_k}.
 \end{aligned}$$

It follows

$$\begin{aligned}
 \frac{\partial}{\partial t} r_k(t, x) + d_k \frac{\partial}{\partial x} r_k(t, x) &= m_k \sum_{i=k}^{\ell-1} \exp(\alpha_{k,i}(x)) b_i(\beta_{k,i}(t, x)) \\
 &+ m_k \exp(\gamma_{k,\ell}(t, x)) r_{\ell,0}(\delta_{k,\ell}(t, x)) \\
 &= m_k r_k(t, x),
 \end{aligned}$$

and the PDE in system (38) also holds in the marked area in Fig. 6b.

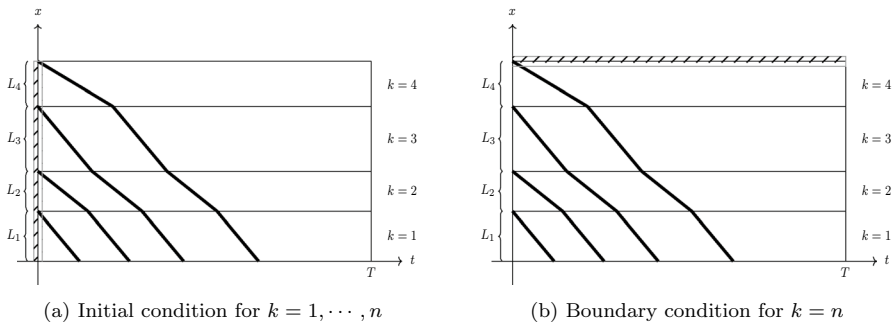
Step 2: The initial conditions in (38) hold

Next we show, that the initial conditions hold. For  $x < d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  and  $\ell = k$ , we have

$$r_k(0, x) = \sum_{i=k}^{k-1} \exp(\alpha_{k,i}(x)) b_i(\beta_{k,i}(0, x)) + \exp(\gamma_{k,k}(0, x)) r_{k,0}(\delta_{k,k}(0, x)).$$

Since sums from  $k$  to  $k - 1$  are equal to 0, this leads to  $\gamma_{k,k}(0, x) = 0$  and  $\delta_{k,k}(0, x) = x$ . Thus the initial conditions are fulfilled (see Fig. 7a).

Step 3: The boundary conditions in (38) hold



**Fig. 7** Areas in which the initial and the boundary conditions of (38) hold

For checking the boundary condition we consider (i.e.  $k = n$  and  $x \geq d_n t + L_n$ ), we have

$$\begin{aligned} r_n(t, L_n) &= \sum_{i=n}^n \exp(\alpha_{n,i}(L_n)) b_i(\beta_{n,i}(t, L_n)) \\ &= \exp(\alpha_{n,n}(L_n)) b_n(\beta_{n,n}(t, L_n)) = b_n(t), \end{aligned}$$

since  $\alpha_{n,n}(L_n) = 1$  and  $\beta_{n,n}(t, L_n) = t$  (see Fig. 7b).

*Step 4: The coupling conditions in (38) hold*

Finally, we have to check the coupling conditions. For  $k = 1, \dots, n - 1$  and  $x \geq d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  it is  $\alpha_{k,i}(L_k) = \alpha_{k+1,i}(0)$  and  $\beta_{k,i}(t, L_k) = \beta_{k+1,i}(t, 0)$ . Thus we have

$$\begin{aligned} r_k(t, L_k) &= \sum_{i=k}^n \exp(\alpha_{k,i}(L_k)) b_i(\beta_{k,i}(t, L_k)) \\ &= \sum_{i=k}^n \exp(\alpha_{k+1,i}(0)) b_i(\beta_{k+1,i}(t, 0)) \\ &= \exp(\alpha_{k+1,k}(0)) b_k(\beta_{k+1,k}(t, 0)) \\ &\quad + \sum_{i=k+1}^n \exp(\alpha_{k+1,i}(0)) b_i(\beta_{k+1,i}(t, 0)) \\ &= b(t) + r_{k+1}(t, 0), \end{aligned}$$

and the coupling conditions are fulfilled (see Fig. 8a). For  $k = 1, \dots, n - 1$ ,  $x < d_k t + d_k \sum_{j=k}^n \frac{L_j}{d_j}$  and  $\ell \in \{k + 1, \dots, n\}$  we have  $\gamma_{k,\ell}(t, L_k) = \gamma_{k+1,\ell}(t, 0)$  and  $\delta_{k,\ell}(t, L_k) = \delta_{k+1,\ell}(t, 0)$ . It follows



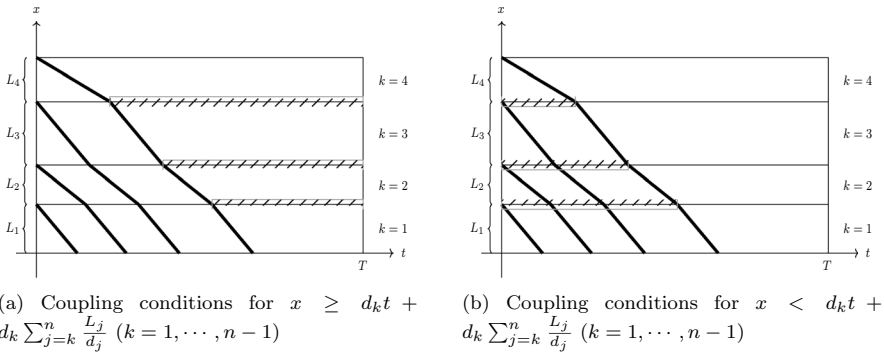


Fig. 8 Areas in which the coupling conditions of (38) hold

$$\begin{aligned}
 r_k(t, L_k) &= \sum_{i=k}^{\ell-1} \exp(\alpha_{k,i}(L_k)) b_i(\beta_{k,i}(t, L_k)) + \exp(\gamma_{k,\ell}(t, L_k)) r_{\ell,0}(\delta_{k,\ell}(t, L_k)) \\
 &= b_k(t) \sum_{i=k+1}^{\ell-1} \exp(\alpha_{k+1,i}(0)) b_i(\beta_{k+1,i}(t, 0)) \\
 &\quad + \exp(\gamma_{k+1,\ell}(t, 0)) r_{\ell,0}(\delta_{k+1,\ell}(t, 0)) \\
 &= b_k(t) + r_{k+1}(t, 0).
 \end{aligned}$$

So the coupling conditions also hold in this case (see Fig. 8b) and the theorem is proven. □

In the next section, we consider model (38) with uncertain boundary data.

### 3.3 Stochastic loads for a scalar PDE

The model (38) describes the distribution of the water contamination in a network which is contaminated by the consumers at the nodes except  $v_0$ . At the end of the network (at node  $v_0$ ), there are restrictions on the contamination rate. But because the contamination rate at the nodes cannot be known a priori, it can be seen as random. Of course one can expect a certain value from statistics or measurements, but this value is never exact. Therefore we use the random boundary data in a Fourier series representation, which was introduced before. In this section we first fix a point in time  $t^* \in [0, T]$  and compute the probability, that the solution of (38) with random boundary data satisfies box constraints at single point in time  $t^* \in [0, T]$ . Then we generalize this approach and compute the probability, that the box constraints are satisfied for all times  $t \in [0, T]$  (cf. (32)).

For  $m \in \mathbb{N}_0$ , we consider the Gaussian distributed random variables  $a_m \sim \mathcal{N}(\mathbb{1}_n, \Sigma)$  with mean value  $\mathbb{1}_n \in \mathbb{R}_+^n$  and positive definite covariance matrix  $\Sigma \in \mathbb{R}^{n \times n}$  on an appropriate probability space  $(\Omega, \mathcal{A}, \mathbb{P})$ . Then the random boundary data at node  $v_k$  (for  $k \in \{1, \dots, n\}$ ) is given by

$$b_k(t, \omega) = \sum_{m=0}^{\infty} a_{m,k}(\omega) a_{m,k}^0 \psi_m(t),$$

with coefficients

$$a_{m,k}^0 := \int_0^T (b_D)_k(t) \psi_m(t) dt,$$

and  $\psi_m$  defined in (33). Note again, that for the implementation, we cut the series after  $N_F \in \mathbb{N}$  terms, which is a good approximation of  $b$  for  $N_F$  large enough (see Remark 4) and as mentioned before, if  $(b_D)_k \in L^2(0, T)$ , then  $b_k \in L^2(0, T)$   $\mathbb{P}$ -almost surely. Because water cannot get cleaned at the nodes, we are only interested in positive boundary values (cf. Sect. 2). Therefore, we assume that  $b_D \in L^2(0, T)$  with  $b_D \geq 0$  and that the parameter  $\Sigma$  of the distribution of  $a_m$  are chosen s.t. the probability that  $b \geq 0$  is almost 1. In practice this can be done as follows: Let  $\gamma_k^* := \operatorname{argmin}_{t \in [0, T]} b_k^D(t)$  and let  $\mathfrak{S}_{k,+}$  resp  $\mathfrak{S}_{k,-}$  be the set of indices where  $a_{m,k}^0 \psi_m(\gamma_k^*) \geq 0$  resp. where  $a_{m,k}^0 \psi_m(\gamma_k^*) < 0$ . We split the Fourier series in positive and negative terms, it follows

$$b_k^D(\gamma_k^*) = \sum_{m \in \mathfrak{S}_{k,+}} a_{m,k}^0 \psi_m(\gamma_k^*) + \sum_{m \in \mathfrak{S}_{k,-}} a_{m,k}^0 \psi_m(\gamma_k^*).$$

Mention that the  $a_m$  all are identically distributed and  $a_{m,k}$  has the variance  $\sigma_k^2$ . We use that fact, that a random Gaussian number  $a_{m,k}(\omega)$  is in  $[1 - 3\sigma_k, 1 + 3\sigma_k]$  with probability 99.73%. The worst case for a random scenario with random numbers  $a_{k,m}(\omega) \in [1 - 3\sigma_k, 1 + 3\sigma_k]$  would be, if the positive terms get smaller and the negative terms get larger, i.e.,

$$\begin{aligned} b_k(\gamma_k^*) &= \sum_{m \in \mathfrak{S}_{k,+}} (1 - 3\sigma_k) a_{m,k}^0 \psi_m(\gamma_k^*) + \sum_{m \in \mathfrak{S}_{k,-}} (1 + 3\sigma_k) a_{m,k}^0 \psi_m(\gamma_k^*) \\ &= b_k^D(\gamma_k^*) - 3\sigma_k \left( \sum_{m \in \mathfrak{S}_{k,+}} a_{m,k}^0 \psi_m(\gamma_k^*) - \sum_{m \in \mathfrak{S}_{k,-}} a_{m,k}^0 \psi_m(\gamma_k^*) \right). \end{aligned}$$

From this it follows, that  $b_k(\gamma_k^*) \geq 0$ , if

$$\sigma_k \leq \frac{b_k^D(\gamma_k^*)}{3 \left( \sum_{m \in \mathfrak{S}_{k,+}} a_{m,k}^0 \psi_m(\gamma_k^*) - \sum_{m \in \mathfrak{S}_{k,-}} a_{m,k}^0 \psi_m(\gamma_k^*) \right)}.$$

For the implementation this is a quite cheap task since the terms of the Fourier series have to be computed anyway. When we would use a truncated Gaussian distribution for the  $a_{m,k}$  bounded from below by  $1 - 3\sigma_k$  and bounded from above by  $1 + 3\sigma_k$ , then we could guarantee that  $b_k(t, \omega)$  is non negative on  $[0, T]$ . As it is mentioned before, we want the solution at a time  $t^* \in [0, T]$  at the node  $v_0$  to satisfy box constraints, s.t.

$$r_1(t^*, 0) \in [r_0^{\min}, r_0^{\max}]. \tag{39}$$

So the full model in this subsection is given in (38). For this model, we define the set of feasible loads as

$$M(t^*) := \left\{ b \in L^2([0, T]; \mathbb{R}_{\geq 0}^n) \mid \begin{array}{l} r_k(t, x) \text{ is a solution of (38) (for } k = 1, \dots, n) \\ \text{such that } r_1(t^*, 0) \in [r_0^{\min}, r_0^{\max}] \end{array} \right\}. \tag{40}$$

Our aim in this subsection is, for a time  $t^* \in [0, T]$ , to compute the probability

$$\mathbb{P}(b \in M(t^*)),$$

which is the probability, that for a random boundary function  $b \in L^2(0, T)$ , the solution of the linear system (38) satisfies the box constraints (39) at a point in time  $t^* \in [0, T]$ . From Theorem 2 we know that

$$r_1(t, 0) = \begin{cases} \sum_{i=1}^n \exp\left(-\sum_{j=1}^i m_j \frac{L_j}{d_j}\right) b_i^\omega\left(t + \sum_{j=1}^i \frac{L_j}{d_j}\right) & t \geq -\sum_{j=1}^n \frac{L_j}{d_j}, \\ \sum_{i=1}^{\ell-1} \exp\left(-\sum_{j=1}^i m_j \frac{L_j}{d_j}\right) b_i^\omega\left(t + \sum_{j=1}^i \frac{L_j}{d_j}\right) & t < -\sum_{j=1}^n \frac{L_j}{d_j} \\ + \exp\left(m_\ell t + \sum_{j=1}^{\ell-1} (m_\ell - m_j) \frac{L_j}{d_j}\right) r_{\ell,0}\left(-d_\ell t - d_\ell \sum_{j=1}^{\ell-1} \frac{L_j}{d_j}\right) & (\ell = 1, \dots, n), \end{cases}$$

where  $b^\omega$  denotes the realization  $b(\omega)$  of the random boundary data for  $\omega \in \Omega$ . For  $i = 1, \dots, n$ , we define the (time dependent) values

$$C_i := \exp\left(-\sum_{j=1}^i m_j \frac{L_j}{d_j}\right),$$

and

$$C_i^0(t) := \exp\left(m_i t + \sum_{j=1}^{i-1} (m_i - m_j) \frac{L_j}{d_j}\right) r_{i,0}\left(-d_i t - d_i \sum_{j=1}^{i-1} \frac{L_j}{d_j}\right).$$

Then,  $b$  is feasible at time  $t^* \in [0, T]$ , iff

$$r_0^{\min} \leq \sum_{i=1}^n C_i b_i^\omega\left(t^* + \sum_{j=1}^i \frac{L_j}{d_j}\right) \leq r_0^{\max}, \tag{41}$$

for  $t^* \geq -\sum_{j=1}^n \frac{L_j}{d_j}$  and it is feasible, iff

$$r_0^{\min} \leq \sum_{j=1}^{\ell-1} C_i b_i^\omega\left(t^* + \sum_{j=1}^i \frac{L_j}{d_j}\right) + C_\ell^0(t^*) \leq r_0^{\max}, \tag{42}$$

for  $-\sum_{j=k}^{\ell-1} \frac{L_j}{d_j} \leq t^* < -\sum_{j=1}^n \frac{L_j}{d_j}$  ( $\ell \in \{1, \dots, n\}$ ). Due to the distribution of the random values  $a_m$  ( $m = 0, 1, \dots$ ), we have

$$b \sim \mathcal{N}(\mu_b(t), \Sigma_b(t)),$$

with  $\mu_b(\cdot) \in \mathbb{R}_+^n$  and  $\Sigma_b(\cdot) \in \mathbb{R}^{n \times n}$  positive definite. To compute the desired probability for a point in time  $t^* \in [0, T]$ , we use the idea of the SRD. For a point  $s \in \mathbb{S}^{n-1}$  at the unit sphere, we set

$$b_s(\hat{r}, t) = \hat{r} \mathcal{L}_b(t) s + \mu_b(t) = \hat{r} \pi_b(t) + \mu_b(t),$$

with  $\pi_b(t) = \mathcal{L}_b(t) v$  and  $\mathcal{L}$ , s.t.  $\mathcal{L}_b(t) \mathcal{L}_b^\top(t) = \Sigma_b(t)$ . Because we are only interested in positive boundary values, we define the regular range as

$$R_{s,\text{reg}} := \{\hat{r} \geq 0 \mid b_s(\hat{r}, t) \geq 0 \quad \forall t \in [0, T]\}.$$

Thus, similar to the stationary case, the (time dependent) one-dimensional sets

$$M_s(t^*) = \{\hat{r} \in R_{s,\text{reg}} \mid b_s(\hat{r}, \cdot) \in M(t^*)\}$$

at a point in time  $t^*$  can be computed by intersecting the regular range with the inequality (41) resp. (42). If  $t^* \geq -\sum_{j=1}^n \frac{L_j}{d_j}$ , then from (41) it follows

$$r_0^{\min} \leq \sum_{i=1}^n C_i \left[ \hat{r} \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) + \mu_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) \right] \leq r_0^{\max}. \quad (43)$$

Define the values

$$a_1 := \frac{r_0^{\min} - \sum_{i=1}^n C_i \mu_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}{\sum_{i=1}^n C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)},$$

and

$$a_2 := \frac{r_0^{\max} - \sum_{i=1}^n C_i \mu_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}{\sum_{i=1}^n C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}.$$

Then we have

$$M_s(t^*) = R_{s,\text{reg}} \cap \begin{cases} [a_1, a_2] & \text{if } \sum_{i=1}^n C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) \geq 0, \\ [a_2, a_1] & \text{else.} \end{cases}$$

If  $t^* < -\sum_{j=1}^n \frac{L_j}{d_j}$  ( $\ell = 1, \dots, n$ ), then from (42) it follows

$$r_0^{\min} \leq \sum_{j=1}^{\ell-1} C_j \left[ \hat{r} \pi_{b,j} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) + \mu_{b,j} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) \right] + C_\ell^0(t^*) \leq r_0^{\max}. \quad (44)$$

Define the values

$$a_1 := \frac{r_0^{\min} - C_\ell^0(t^*) - \sum_{i=1}^{\ell-1} C_i \mu_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}{\sum_{i=1}^{\ell-1} C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)},$$

and

$$a_2 := \frac{r_0^{\max} - C_\ell^0(t^*) - \sum_{i=1}^{\ell-1} C_i \mu_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}{\sum_{i=1}^{\ell-1} C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right)}.$$

Then we have

$$M_s(t^*) = R_{s,\text{reg}} \cap \begin{cases} [a_1, a_2] & \text{if } \sum_{i=1}^{\ell-1} C_i \pi_{b,i} \left( t^* + \sum_{j=1}^i \frac{L_j}{d_j} \right) \geq 0, \\ [a_2, a_1] & \text{else.} \end{cases}$$

Thus, for every point in time  $t^* \in [0, T]$ , the set  $M_s(t^*)$  can be represented as a union of disjoint intervals, which we can use to compute the probability for the random boundary function  $b$  to be feasible, like in (9).

Next, we approximate the probability  $\mathbb{P}(b \in M(t^*))$  by using the KDE approach introduced in Sect. 2.1. We consider the stochastic equation corresponding to (38) with random boundary data. Note that this equation has also a solution  $\mathbb{P}$ -almost surely. We assume that the distribution of the random variable  $r_1(t^*, 0)$  is absolutely continuous with probability density function  $\varrho_{r,t^*}$  for  $t^* \in [0, T]$ . Thus the point in time  $t^*$  has to be large enough, so that  $r_1(t^*, 0)$  depends on the random boundary data as it is explained in Remark 5. Similar to Sect. 2.1, it holds

$$\mathbb{P}(b \in M(t^*)) = \mathbb{P}(r_1(t^*, 0) \in [r_0^{\min}, r_0^{\max}]) = \int_{r_0^{\min}}^{r_0^{\max}} \varrho_{r,t^*}(z) dz.$$

In order to approximate the unknown probability density function by the KDE, we need a sampling set for the random variable  $r_1(t^*, 0)$ . Therefore, let

$$\mathcal{A}_m := \{a_m^{S,1}, \dots, a_m^{S,N}\}$$

for  $m = 1, \dots, N_F$  be independent and identically distributed sampling sets of  $a_m \sim \mathcal{N}(\mu, \Sigma)$  (with mean value  $\mu \in \mathbb{R}_+^n$  and  $\Sigma \in \mathbb{R}^{n \times n}$  positive definite). Let

$$\mathcal{B}_A := \{b^{S,1}, \dots, b^{S,N}\}$$

with  $b^{S,i} = \sum_{m=0}^{N_F} a_m^{S,i} a_m^0 \phi_m$  be the corresponding sampling of the random boundary function, where  $b^{S,i} \in L^2(0, T)$  with  $b^{S,i} \geq 0$  ( $i = 1, \dots, N$ )  $\mathbb{P}$ -almost surely. With this and Theorem 2 (or an appropriate numerical method for solving the system (38)), we define the sample

$$\mathcal{R}_* := \{r_1(t^*, 0, b^{S,1}), \dots, r_1(t^*, 0, b^{S,N})\},$$

where  $r_1(t^*, 0, b^{S,i})$  ( $i = 1, \dots, N$ ) is the solution of (38) at node  $v_0$  at time  $t^* \in [0, T]$  with boundary function  $b^{S,i} \in \mathcal{B}_A$ . The samplings  $\mathcal{B}_A$  and  $\mathcal{R}_*$  are also independent and identically distributed. Then for a bandwidth  $h \in \mathbb{R}_+$ , the probability density function  $\varrho_{r,t^*,N}$  is approximately given by

$$\varrho_{r,t^*,N}(z) = \frac{1}{N h} \sum_{i=1}^N \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2} \left(\frac{z - r_1(t^*, 0, b^{S,i})}{h}\right)^2\right).$$

We choose the bandwidth according to (12). Therefore, we get the same convergence results for the KDE resp. the approximated probability as in (13) resp. (14). So we can approximate the probability for a random boundary function  $b$  to be feasible at time  $t^* \in [0, T]$  by

$$\mathbb{P}(b \in M(t^*)) \approx \int_{r_0^{\min}}^{r_0^{\max}} \varrho_{r,t^*,N}(z) dz =: \mathbb{P}_N(b \in M(t^*)) \tag{45}$$

So far, the computation of the desired probability in this subsection was only for box constraints at a certain point in time  $t^* \in [0, T]$ , e.g., the end time  $t^* = T$ . As mentioned before, we are interested in box constraints for the full time period, which leads to a probabilistic constraint given in (32). The idea is, that  $r_1(t, 0)$  satisfies the box constraints for all  $t \in [0, T]$ , iff the maximum and the minimum value of  $r_1(t, 0)$  in  $[0, T]$  satisfies the box constraints:

$$\begin{aligned} r_1(t, 0) &\in [r_0^{\min}, r_0^{\max}] \quad \forall t \in [0, T] \\ &\Downarrow \\ \begin{bmatrix} r_0^{\min} \\ r_0^{\min} \end{bmatrix} &\leq \begin{bmatrix} \max_{t \in [0, T]} r_1(t, 0) \\ \min_{t \in [0, T]} r_1(t, 0) \end{bmatrix} \leq \begin{bmatrix} r_0^{\max} \\ r_0^{\max} \end{bmatrix}. \end{aligned}$$

We define the values

$$\underline{t} := \operatorname{argmin}_{t \in [0, T]} r_1(t, 0),$$

and

$$\bar{t} := \operatorname{argmax}_{t \in [0, T]} r_1(t, 0).$$

For the SRD we use a similar procedure as above. We have

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]) \Leftrightarrow \mathbb{P}(b \in M(\underline{t}) \text{ and } b \in M(\bar{t})).$$

That means, we need to intersect the regular range with two inequalities of the form (43) resp. (44) to get the set  $M_s(\underline{t}) \cap M_s(\bar{t})$  and to compute the desired probability. This is only possible, if one compute  $\underline{t}$  and  $\bar{t}$  for the deterministic boundary function  $b_D(t)$ . Otherwise, due to the randomness of the boundary functions  $b(t)$ , the *argmin* and the *argmax* can be shifted and thus, the  $\underline{t}$  and  $\bar{t}$  depend on this uncertainty. So a general  $\underline{t}$  and  $\bar{t}$  does not exist and it is not clear, where to evaluate the mean and the variance in (43) resp. (44). But if we use the deterministic boundary function, the  $\underline{t}$

and the  $\bar{t}$  does not meet the minimal and maximal values of the random boundary functions and thus, the result may not be significant.

The KDE uses only the solution of the model for estimating an analytical probability density function, which we can use later for the optimization. To extend the KDE to probabilistic constraints like (32), we consider the minimal and maximal contamination concentration at node  $v_0$  in the time period  $[0, T]$ . We assume that the distribution of the random vector  $R := (\min_{t \in [0, T]} r_1(t, 0), \max_{t \in [0, T]} r_1(t, 0))^T$  is absolutely continuous with probability density function  $\varrho_R$ . Using this time independent variable, we get

$$\begin{aligned} \mathbb{P}(b \in M(t) \forall t \in [0, T]) &= \mathbb{P}(R \in [r_0^{\min}, r_0^{\max}] \times [r_0^{\min}, r_0^{\max}]) \\ &= \int_{[r_0^{\min}, r_0^{\max}] \times [r_0^{\min}, r_0^{\max}]} \varrho_R(z) dz. \end{aligned}$$

We use now a two dimensional KDE like in the stationary case for tree-structured graphs. For the set  $\mathcal{B}_A$  we define the sampling set of the random variable  $R$  by

$$\mathcal{R} = \left\{ \left( r_1^{S,1}, \bar{r}_1^{S,1} \right)^T, \dots, \left( r_1^{S,N}, \bar{r}_1^{S,N} \right)^T \right\}$$

with

$$r_1^{S,i} := \min_{t \in [0, T]} r_1(t, 0, b^{S,i}(t)) \quad (i = 1, \dots, N)$$

and

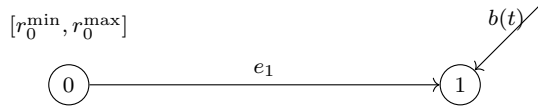
$$\bar{r}_1^{S,i} := \max_{t \in [0, T]} r_1(t, 0, b^{S,i}(t)) \quad (i = 1, \dots, N).$$

This sampling is also independent and identically distributed. Using the KDE given in (19) with bandwidth (18), the probability density function of the minimal and maximal contamination rate at node  $v_0$  in the time period  $[0, T]$  is approximately given by

$$\begin{aligned} \varrho_{R,N}(z) &= \frac{1}{Nh_y^2 \sigma_{N,1}^{\min} \sigma_{N,1}^{\max}} \sum_{i=1}^N \frac{1}{2\pi} \exp \left( -\frac{1}{2} \left( \frac{z_1 - r_1^{S,i}}{h_y \sigma_{N,1}^{\min}} \right)^2 \right) \\ &\quad \cdot \exp \left( -\frac{1}{2} \left( \frac{z_2 - \bar{r}_1^{S,i}}{h_y \sigma_{N,1}^{\max}} \right)^2 \right), \end{aligned}$$

where  $(\sigma_{N,1}^{\min})^2$  and  $(\sigma_{N,1}^{\max})^2$  are the variances of  $r_1^{S,i}$  and  $\bar{r}_1^{S,i}$  ( $i = 1, \dots, N$ ). For this estimator we get the same convergence results for the KDE resp. the approximated probability as in (21) and (22). Thus we approximate the desired probability as follows:

**Fig. 9** Example graph with 2 nodes



**Table 8** Values for the dynamic example

$r_0^{\min}$	$r_0^{\max}$	$\mu$	$\sigma$	$d$	$m$	$L$	$T$
2	6	1	0.25	-5	-1	1	4

$$\mathbb{P}(b \in M(t) \forall t \in [0, T]) \approx \int_{[r_0^{\min}, r_0^{\max}] \times [r_0^{\min}, r_0^{\max}]} \varrho_{R,N}(z) dz =: \mathbb{P}_N^*(b \in M(t) \forall t \in [0, T]).$$

**Remark 8** If  $r_1^{S,i}$  or  $\bar{r}_1^{S,i}$  is taken in the time period, in which the solution depends on the initial data, then the distribution function of the minimal resp. maximal contamination rate contains a discontinuity. Thus, a probability density function in the classical sense does not even exist. So one has to guarantee, that  $r_1^{S,i}$  and  $\bar{r}_1^{S,i}$  is not taken in the beginning of the time period, e.g., by excluding this part from the probabilistic constraint. As it is mentioned in Remark 5, for times  $t \geq t^*$  with

$$t^* = \sum_{j=1}^n \frac{L_j}{|d_j|}$$

the solution does not depend on the initial condition anymore. So instead of solving (32) one can solve

$$\mathbb{P}(b \in M(t) \forall t \in [t^*, T]) \geq \alpha.$$

Motivated by the application this makes sense since the initial state is either given a priori or can be chosen a priori s.t. all bounds are satisfied for small times.

**Example 3** Consider the graph with one edge shown in Fig. 9.

For the computation, we use  $r_0(x) = 5 \exp\left(\frac{m}{d}(x - L)\right)$  as initial function,  $b_D(t) = -2 \sin(2t) + 5$  and  $b(t, \omega) = \sum_{m=0}^{\infty} a_m(\omega) a_m^0 \psi_m(t) + 5$  as random boundary function. The coefficients  $a_m^0$  for the shifted function  $b_D(t) - 5$  are given by (34). The initial function is chosen s.t. the solution is constant at  $x = 0$  for small times, so we guarantee, that all minimal values are below this constant value and all maximal values are above this constant value (see Remark 8). The other values are given in Table 8.

Further, we use 101 points for the time discretization and we cut the Fourier series of the random boundary data after 30 terms. A sampling of 10 random boundary functions and the corresponding solutions are shown in Fig. 10.



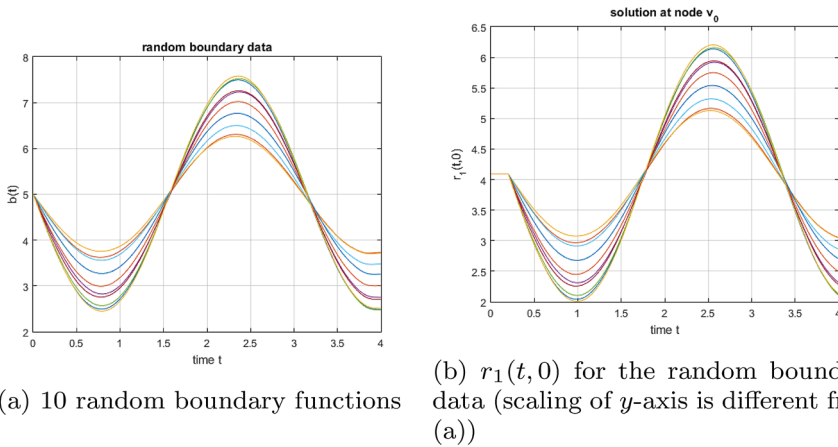


Fig. 10 Sampling of 10 random boundary functions and the corresponding solutions at node  $v_0$

Table 9 Results for the dynamic example with one edge

	Test 1 (%)	Test 2 (%)	Test 3 (%)	Test 4 (%)	Test 5 (%)	Test 6 (%)	Test 7 (%)	Test 8 (%)
MC	74.32	74.39	74.22	74.32	74.24	74.33	74.51	74.40
KDE	74.33	74.39	74.21	74.31	74.24	74.32	74.51	74.40

For the MATLAB<sup>®</sup> implementation, we use a sampling of  $1 \cdot 10^5$  boundary functions in terms of Fourier series. We compare the probabilities of the KDE again with a classical Monte Carlo method (MC). The MC method checks for a random boundary function, if the bounds are satisfied for every point of the time discretization. If this is not the case, this boundary function is not feasible. The results of the tests are shown in Table 9.

One can see, that the results of MC and the KDE are almost equal. The mean probability in MC resp. KDE is 74.38% resp. 74.37% and the variance is 0.0141 resp. 0.0142. For a confidence level of 95% the confidence interval for the MC probability is [74.26%, 74.42%], which is also the confidence interval of the KDE probability. Both methods are still quite fast. MATLAB<sup>®</sup> needs much more time ( $\sim 1$  minute) for the sampling and computing the random boundary data than for computing the probabilities, which is less than one second.

In this subsection, we have considered the SRD and the KDE in a dynamic setting based on the results from Sect. 2. First the box constraints only hold for a certain point in time  $t^* \in [0, T]$  and then, the box constraints hold for the full time period  $[0, T]$ .

### 3.4 Stochastic optimization on dynamic flow networks

In this subsection, we formulate necessary optimality conditions for the dynamic hyperbolic system introduced before. In the subsection before, we introduced different

ways to compute the probability for a random boundary function to be feasible. Using the KDE gives us a good approximation of this probability. Define the set

$$\mathcal{R}_0 := [r_0^{\min}, \infty),$$

and the function

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad (r_0^{\max}) \mapsto f(r_0^{\max}).$$

For a probability level  $\alpha \in (0, 1)$ , consider the optimization problem with the approximated probabilistic constraints

$$\begin{cases} \min_{r_0^{\max} \in \mathcal{R}_0} & f(r_0^{\max}) \\ \text{s.t.} & \mathbb{P}_N(b \in M(t) \forall t \in [0, T]) \geq \alpha \end{cases} \quad (46)$$

Similar to the stationary case, for  $k = 1, \dots, n$  and  $i = 1, \dots, N$ , we define

$$\varphi_{i,k}^{\min}(x) := \frac{x - \underline{r}_k^{S,i}}{\sqrt{2}h_y\sigma_{N,k}^{\min}} \quad \text{and} \quad \varphi_{i,k}^{\max}(x) := \frac{x - \bar{r}_k^{S,i}}{\sqrt{2}h_y\sigma_{N,k}^{\max}},$$

where

$$\underline{r}_k^{S,i} := \min_{t \in [0, T]} r_k(t, 0, b^{S,i}) \quad \text{and} \quad \bar{r}_k^{S,i} := \max_{t \in [0, T]} r_k(t, 0, b^{S,i}),$$

with samples  $b^{S,i} \in \mathcal{B}_{\mathcal{A}}$ . Then we can rewrite the computation of the desired probability using the error function as

$$\begin{aligned} \mathbb{P}_N(b \in M(t) \forall t \in [0, T]) &= \int_{[r_0^{\min}, r_0^{\max}] \times [r_0^{\min}, r_0^{\max}]} \varrho_{R,N}(z) dz \\ &= \frac{1}{4N} \sum_{i=1}^N \left[ \operatorname{erf} \left( \varphi_{i,1}^{\min}(r_0^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,1}^{\min}(r_0^{\min}) \right) \right] \\ &\quad \cdot \left[ \operatorname{erf} \left( \varphi_{i,1}^{\max}(r_0^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,1}^{\max}(r_0^{\min}) \right) \right]. \end{aligned}$$

We define the function

$$g_\alpha : \mathbb{R} \rightarrow \mathbb{R}, r^{\max} \mapsto \alpha - \mathbb{P}_N(b \in M(t), r^{\max} \forall t \in [0, T]).$$

Difference to the stationary case is, that the terms in both dimensions depend on the same upper bound, thus we have to use product rule to compute the derivative. It follows

$$\begin{aligned} \frac{d}{dr_0^{\max}} g_\alpha(r_0^{\max}) &= -\frac{1}{4N} \sum_{i=1}^N \left[ \operatorname{erf} \left( \varphi_{i,1}^{\min}(r_0^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,1}^{\min}(r_0^{\min}) \right) \right] \\ &\quad \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_y \sigma_{N,1}^{\max}} \exp \left( -(\varphi_{i,1}^{\max}(r_0^{\max}))^2 \right) \\ &\quad + \left[ \operatorname{erf} \left( \varphi_{i,1}^{\max}(r_0^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,1}^{\max}(r_0^{\min}) \right) \right] \\ &\quad \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_y \sigma_{N,1}^{\min}} \exp \left( -(\varphi_{i,1}^{\min}(r_0^{\max}))^2 \right). \end{aligned}$$

As it is mentioned in Remark 2, the LICQ is always fulfilled and we can state the necessary optimality conditions for the approximated problem (46):

**Corollary 3** *Let  $r_0^{*,\max} \in \mathbb{R}$  be a (local) optimal solution of (46). Since the LICQ holds in  $r_0^{*,\max}$ , there exists a multiplier  $\mu^* \geq 0$ , s.t.*

$$\begin{aligned} f'(r_0^{*,\max}) + \mu^* g'_\alpha(r_0^{*,\max}) &= 0, \\ g_\alpha(r_0^{*,\max}) &\leq 0, \\ \mu^* g_\alpha(r_0^{*,\max}) &= 0. \end{aligned}$$

Thus,  $(r_0^{*,\max}, \mu^*) \in \mathbb{R}^2$  is a Karush–Kuhn–Tucker point.

**Remark 9** In the  $n$ -dimensional case, in which we have bounds at  $n$  nodes, the computation of the desired probability is the following:

$$\begin{aligned} \mathbb{P}_N(b \in M(t) \forall t \in [0, T]) &= \frac{1}{4^n N} \sum_{i=1}^N \prod_{j=1}^n \left[ \operatorname{erf} \left( \varphi_{i,j}^{\min}(r_j^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,j}^{\min}(r_j^{\min}) \right) \right] \\ &\quad \cdot \left[ \operatorname{erf} \left( \varphi_{i,j}^{\max}(r_j^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,j}^{\max}(r_j^{\min}) \right) \right]. \end{aligned}$$

Mention that this is an  $2n$ -dimensional KDE since every boundary function provides two samples, one for the minimal values and one for the maximal values. The partial derivatives with respect to  $r_j^{\max}$  are given by

$$\begin{aligned} & \frac{\partial}{\partial r_j^{\max}} \mathbb{P}_N(b \in M(t) \forall t \in [0, T]) \\ &= \frac{1}{4^n N} \sum_{i=1}^N \prod_{j=1, j \neq k}^n \left[ \operatorname{erf} \left( \varphi_{ij}^{\min}(r_j^{\max}) \right) - \operatorname{erf} \left( \varphi_{ij}^{\min}(r_j^{\min}) \right) \right] \\ & \cdot \left[ \operatorname{erf} \left( \varphi_{ij}^{\max}(r_j^{\max}) \right) - \operatorname{erf} \left( \varphi_{ij}^{\max}(r_j^{\min}) \right) \right]. \\ & \cdot \left[ \operatorname{erf} \left( \varphi_{i,k}^{\min}(r_k^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,k}^{\min}(r_k^{\min}) \right) \right] \\ & \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_y \sigma_{N,k}^{\max}} \exp \left( -(\varphi_{i,k}^{\max}(r_k^{\max}))^2 \right) \\ & + \left[ \operatorname{erf} \left( \varphi_{i,k}^{\max}(r_k^{\max}) \right) - \operatorname{erf} \left( \varphi_{i,k}^{\max}(r_k^{\min}) \right) \right] \\ & \cdot \frac{\sqrt{2}}{\sqrt{\pi} h_y \sigma_{N,k}^{\min}} \exp \left( -(\varphi_{i,k}^{\min}(r_k^{\max}))^2 \right) \right], \end{aligned}$$

where  $(\sigma_{N,k}^{\max})^2$  resp.  $(\sigma_{N,k}^{\min})^2$  are the variances of the sampling of the maximal values resp. of the minimal values.

### 3.5 Application to a realistic network

For this section we basically use the graph of the GasLib-11 but since this graph was designed for gas transportation we slightly vary it. First we assume that the compressor edges are normal edges. As in section Sect. 2.4 we assume that the valve is closed, s.t. the edge between node  $v_2$  and  $v_4$  vanishes. The water is contaminated at the nodes  $v_6, v_9$  and  $v_{10}$  and the pollution distributes in the graph. We assume that the pollution equally distributes at node  $v_7$ , i.e., half of the pollution distributes in edge  $e_6$ , the other half in  $e_7$ . We define pollution bounds  $r_0^{\min}, r_0^{\max} \in \mathbb{R}_{\geq 0}$  for node  $v_0$ ,  $r_1^{\min}, r_1^{\max} \in \mathbb{R}_{\geq 0}$  for node  $v_1$  and  $r_5^{\min}, r_5^{\max} \in \mathbb{R}_{\geq 0}$  for node  $v_5$ . We want these bounds to be satisfied. A scheme of this network is shown in Fig. 11. For simplicity we assume that  $m = -0.1, d = -1$  and  $L = 1$  for every edge.

The boundary functions are given by

$$\begin{aligned} b_6(t) &= \sin(t) + 5, \\ b_9(t) &= \frac{1}{4}|t - 3| + 2, \\ \text{and } b_{10}(t) &= \frac{1}{(t - 1)^2 + \frac{1}{2}} + 3. \end{aligned} \tag{47}$$

The initial conditions are given by

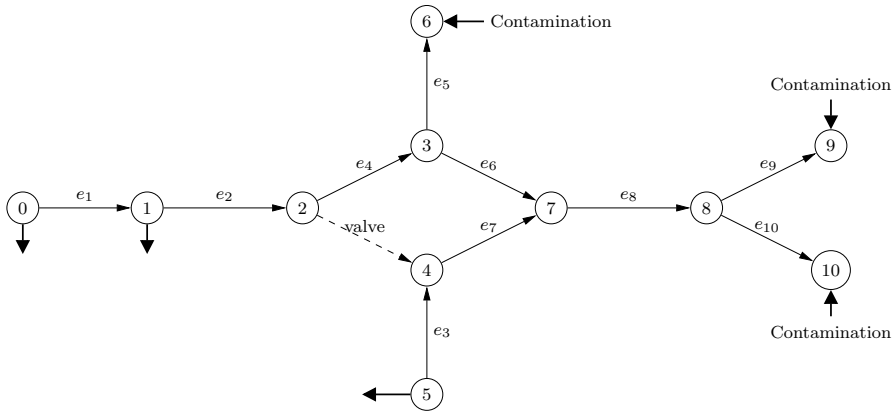


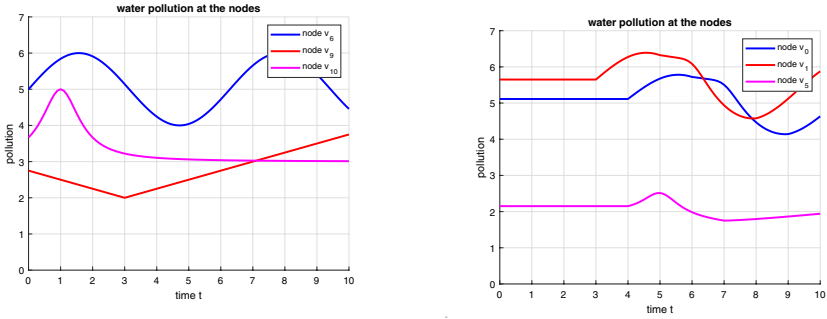
Fig. 11 A network scheme for water pollution

$$\begin{aligned}
 r_{e_{10},0}(x) &= \frac{11}{3} \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_9,0}(x) &= \frac{11}{4} \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_8,0}(x) &= \left(r_{e_9,0}(0) + r_{e_{10},0}(0)\right) \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_7,0}(x) &= \frac{1}{2} r_{e_8,0}(0) \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_6,0}(x) &= \frac{1}{2} r_{e_8,0}(0) \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_5,0}(x) &= 5 \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_4,0}(x) &= \left(r_{e_5,0}(0) + r_{e_6,0}(0)\right) \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_3,0}(x) &= r_{e_7,0}(0) \exp\left(\frac{m}{d}(x-L)\right), \\
 r_{e_2,0}(x) &= r_{e_4,0}(0) \exp\left(\frac{m}{d}(x-L)\right), \\
 \text{and } r_{e_1,0}(x) &= r_{e_2,0}(0) \exp\left(\frac{m}{d}(x-L)\right).
 \end{aligned}$$

The initial conditions are chosen, s.t. the solution at the nodes is constant as long as information from the boundary nodes needs to reach the nodes and the initial conditions satisfy the  $C^0$ -compatibility with the boundary conditions, which is

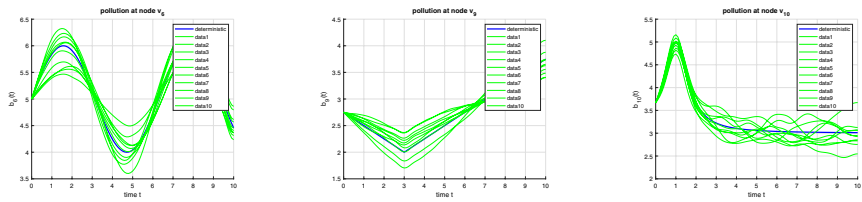
$$r_{e_5,0}(L) = b_6(0), \quad r_{e_9,0}(L) = b_9(0) \quad \text{and} \quad r_{e_{10},0}(L) = b_{10}(0).$$

The boundary functions and the solution at the nodes  $v_0, v_1$  and  $v_5$  are shown in Fig. 12. Since the boundary functions do not satisfy  $b_i(0) = 0$  ( $i \in \{6, 9, 10\}$ ), we compute the Fourier series for the functions  $(b_i(t) - b_i(0))$ , randomize them by multiplying a Gaussian distributed random number (with  $\mu = 1$  and  $\sigma^2 = 0.1$ ) to every summand of the series and then we add the constants  $b_i(0)$  to the random Fourier



(a) Deterministic Boundary Data at node  $v_6, v_9, v_{10}$  (b) Solution at node  $v_0, v_1$  and  $v_5$

**Fig. 12** Deterministic boundary data and deterministic solution



(a) Random scenarios for boundary data  $b_6(t)$  (b) Random scenarios for boundary data  $b_9(t)$  (c) Random scenarios for boundary data  $b_{10}(t)$

**Fig. 13** Random boundary scenarios at  $v_6, v_9$  and  $v_{10}$

series. For the implementation we use the first 30 terms of the Fourier series, i.e.,  $N_F = 30$ . Some random scenarios for the boundary functions are shown in Fig. 13.

The lower pollution bounds are given by

$$r^{\min} = \begin{bmatrix} r_0^{\min} \\ r_1^{\min} \\ r_5^{\min} \end{bmatrix} = \begin{bmatrix} 3.5 \\ 4 \\ 1 \end{bmatrix}.$$

Consider the linear function

$$f : \mathbb{R}^3 \rightarrow \mathbb{R}, \quad f : r^{\max} \mapsto c^T r^{\max},$$

with  $c = \mathbb{1}_3$ . We first solve the deterministic problem

$$\begin{aligned} \min_{r^{\max} \geq r^{\min}} & f(r^{\max}) \\ \text{s.t.} & b(t) \in M(r^{\max}) \forall t \in [0, T], \end{aligned} \tag{48}$$

**Table 10** Probability  $\mathbb{P}(b \in M(r_{\text{det}}^{\max}))$  for the optimal deterministic upper pollution bounds

	Test 1 (%)	Test 2 (%)	Test 3 (%)	Test 4 (%)	Test 5 (%)	Test 6 (%) (%)	Test 7 (%)	Test 8 (%)
MC	37.71	37.77	37.83	37.65	37.80	37.54	37.92	38.01
KDE	37.57	37.62	37.68	37.51	37.67	37.39	37.76	37.88

**Table 11** Stochastic optimal upper pollution bounds  $r_{\text{stoch}}^{\max}$

Test 1	Test 2	Test 3	Test 4	Test 5	Test 6	Test 7	Test 8
$\begin{bmatrix} 5.936 \\ 6.560 \\ 2.577 \end{bmatrix}$	$\begin{bmatrix} 5.938 \\ 6.562 \\ 2.574 \end{bmatrix}$	$\begin{bmatrix} 5.937 \\ 6.561 \\ 2.576 \end{bmatrix}$	$\begin{bmatrix} 5.935 \\ 6.559 \\ 2.578 \end{bmatrix}$	$\begin{bmatrix} 5.936 \\ 6.560 \\ 2.575 \end{bmatrix}$	$\begin{bmatrix} 5.935 \\ 6.559 \\ 2.576 \end{bmatrix}$	$\begin{bmatrix} 5.937 \\ 6.561 \\ 2.577 \end{bmatrix}$	$\begin{bmatrix} 5.936 \\ 6.560 \\ 2.575 \end{bmatrix}$

with  $T = 10$  and the boundary functions are given in (47). As in the stationary case we use the default setting of the MATLAB<sup>®</sup>-routine *fmincon* to solve (48), which is an interior-point algorithm. It returns

$$r_{\text{det}}^{\max} = \begin{bmatrix} r_0^{\max} \\ r_1^{\max} \\ r_5^{\max} \end{bmatrix} = \begin{bmatrix} 5.78 \\ 6.39 \\ 2.51 \end{bmatrix},$$

as optimal deterministic solution, i.e., as the lowest upper pollution bound for the nodes  $v_0, v_1$  and  $v_5$ . Now we consider uncertain water contamination at the nodes  $v_6, v_9$  and  $v_{10}$ . We compute the probability that the random contamination satisfies the bounds  $[r^{\min}, r_{\text{det}}^{\max}]$  at the nodes  $v_0, v_1$  and  $v_5$ , which is  $\mathbb{P}(b(t) \in M(r_{\text{det}}^{\max}) \forall t \in [0, T])$ . This is shown for 8 tests (each with  $1 \cdot 10^5$  samples) in Table 10.

The mean MC probability is 37.78% and the mean KDE probability is 37.63%. For a confidence level of 95% the confidence interval for the MC probability is [37.65%, 37.90%] and the confidence interval for the KDE probability is [37.51%, 37.76%]. Just like in the stationary case, the deterministic upper bound is unsatisfactory, so we consider the probabilistic constrained optimization problem (46) with  $\alpha := 0.75$ . The problems described in Sect. 2.4 also occur here but *fmincon* returns a solution if we choose the optimal deterministic solution as starting point. The results of 8 Tests with  $1 \cdot 10^5$  scenarios are shown in Table 11. This time we show 3 decimal places because otherwise the solutions would be equal. In 8 more Tests we solve (46) by using Corollary 3. The results vary from the optimal solutions computed by *fmincon* in a range of  $1 \cdot 10^{-7}$ . So here one could also expect that the necessary optimality conditions are sufficient but we do not analyze this here.

One can see, that all results are almost equal. The optimal upper pollution bounds of the probabilistic constrained optimization problem (46) are larger then the optimal upper pollution bounds of the deterministic optimization problem (48). The computation time is quite similar to the stationary case with the

difference that solving the necessary optimality conditions needed much more time here (about 2 min per test).

## 4 Conclusion

In this paper, we have shown two different ways to evaluate probabilistic constraints in the context of hyperbolic balance laws on graphs for both, a stationary and a dynamic setting with box constraints for the solution.

The spheric radial decomposition provides a good method for computing the probabilities for random boundary data to be feasible in the stationary case and in the dynamic case for box constraints for a certain point in time. Because the spheric radial decomposition is explicitly based on the analytical solution, it leads to good results. But as soon as the analytical solution is not given, the inequalities cannot be derived and so the spheric radial decomposition becomes an almost purely numerical method.

A kernel density estimator does not need the analytical solution, a numerically computed solution is sufficient, and it provides an estimated, but explicit representation of the probability density function, which can be used for computing the probabilistic constraint and for deriving necessary optimality conditions for probabilistic constrained optimization problems. In addition the kernel density estimator is smooth even if the exact probability density function is nonsmooth. The examples showed, that the kernel density estimator provides results almost as good as the results from the spheric radial decomposition and a classical Monte Carlo approach. Of course, the Monte Carlo approach is faster and easier to use, but it cannot be used to get any kind of analytical result like e.g. the necessary optimality conditions.

Another advantage of both methods is, that they do not depend on the topology of the graph. The topology of the graph influences the analytical solution, but the methods themselves work independently of the complexity of the graph. Further, the idea of the kernel density estimator can easily be used for any kind of partial differential equation with random boundary data, because it only requires a sufficiently accurate solution of the PDE, e.g., by using numerical methods, and this has been a goal of many papers in the last decades.

**Acknowledgements** The authors are supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) within the collaborative research center TRR154 “Mathematical modeling, simulation and optimisation using the example of gas networks” (Project-ID 239904186, TRR154/2-2018, TP B01 (Lang, Strauch), TP C03 (Gugat, Schuster) and TP C05 (Giesselmann, Gugat)).

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Declarations**

**Conflict of interest** The authors declare that there is no conflict of interest regarding the publication of this paper.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adelhütte D, Aßmann D, Gradón TG, Gugat M, Heitsch H, Henrion R, Liers F, Nitsche S, Schultz R, Stingl M, Wintergerst D (2020) Joint model of probabilistic-robust (proburst) constraints applied to gas network optimization. *Vietnam J Math*. <https://doi.org/10.1007/s10013-020-00434-y>
- Andrews LC (1998) *Special functions of mathematics for engineers*, 2nd edn. SPIE Press
- Banda MK, Herty M, Klar A (2006) Coupling conditions for gas networks governed by the isothermal Euler equations. *Netw Heterog Media* 1:295–314
- Banda MK, Herty M, Klar A (2006) Gas flow in pipeline networks. *Netw Heterog Media* 1:41–56
- Bastin G, Coron JM, d'Andréa Novel B (2009) On Lyapunov stability of linearised Saint–Venant equations for a sloping channel. *Netw Heterog Media* 4:177–187
- Bermúdez A, González-Díaz J, González-Diéguez FJ, González-Rueda AM, de Córdoba MPF (2015) Simulation and optimization models of steady-state gas transmission networks. *Energy Procedia* 64:130–139
- Caillaud JB, Cerf M, Sassi A, Trélat E, Zidani H (2018) Solving chance constrained optimal control problems in aerospace via kernel density estimation. *Optim Control Appl Methods* 39:1833–1858
- Colombo RM, Guerra G, Herty M, Schleper V (2009) Optimal control in networks of pipes and canals. *J Control Optim* 48:2032–2050
- Coron JM (2002) Local controllability of a 1-d tank containing a fluid modeled by the shallow water equations. *ESAIM: COCV* 8:513–554
- Devroye L, Györfi L (1985) *Nonparametric density estimation: the L1 view*. Wiley series in probability and mathematical statistics. Wiley
- Domschke P, Hiller B, Lang J, Tischendorf C (2017) Modellierung von Gasnetzwerken: Eine Übersicht. *Tech. Rep. 2717*, Technische Universität Darmstadt. <http://www3.mathematik.tu-darmstadt.de/fb/mathe/preprints.html>
- Duller C (2018) *Einführung in die nichtparametrische Statistik mit SAS, R und SPSS*, 2nd edn. Springer
- Farshbaf-Shaker MH, Henrion R, Hömberg D (2018) Properties of chance constraints in infinite dimensions with an application to pde constrained optimization. *Set-Valued Var Anal* 26:821–841
- Farshbaf-Shaker MH, Gugat M, Heitsch H, Henrion R (2020) Optimal Neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints. *SIAM J Control Optim* 58:2288–2311
- Fügenschuh A, Göttlich S, Herty M (2007) Water contamination detection. In: *Wirtschaftsinformatik proceedings*, 85
- Gotzes C, Heitsch H, Henrion R, Schultz R (2016) On the quantification of nomination feasibility in stationary gas networks with random load. *Math Methods Oper Res* 84:427–457
- Gradón TG, Heitsch H, Henrion R (2016) A joint model of probabilistic/robust constraints for gas transport management in stationary networks. *Comput Manag Sci* 14:427–457
- Gramacki A (2018) *Nonparametric kernel density estimation and its computational aspects*, 1st edn. Springer
- Gugat M (2012) Contamination source determination in water distribution networks. *SIAM J Appl Math* 72:1772–1791
- Gugat M, Herty M (2009) Existence of classical solutions and feedback stabilization for the flow in gas networks. *ESIAM Control Optim* 17:28–51
- Gugat M, Leugering G (2003) Global boundary controllability of the de St. Venant equations between steady states. *Ann Inst H Poincaré Anal Non Linéaire* 20(1):1–11

- Gugat M, Schuster M (2018) Stationary gas networks with compressor control and random loads: optimization with probabilistic constraints. *Math Prob Eng*. <https://doi.org/10.1155/2018/7984079>
- Gugat M, Ulbrich S (2017) The isothermal Euler equations for ideal gas with source term: product solutions, flow reversal and no blow up. *Aust J Math Anal Appl* 454:439–452
- Gugat M, Ulbrich S (2018) Lipschitz solutions of initial boundary value problems for balance laws. *Math Models Methods Appl Sci* 28(5):921–951
- Gugat M, Wintergerst D (2018) Transient flow in gas networks: traveling waves. *Int J Appl Math Comput Sci* 28:341–348
- Gugat M, Leugering G, Schmidt EJPG (2004) Global controllability between steady supercritical flows in channel networks. *Math Methods Appl Sci* 27:781–802
- Gugat M, Hante FM, Hirsch-Dick M, Leugering G (2015) Stationary states in gas networks. *Netw Heterog Media* 10(2):295–320
- Gugat M, Schultz R, Wintergerst D (2018) Networks of pipelines for gas with nonconstant compressibility factor: stationary states. *Comput Appl Math* 37:1066–1097
- Gugat M, Schultz R, Schuster M (2020) Convexity and starshapedness of feasible sets in stationary flow networks. *Netw Heterog Media* 15:171–195
- Härdle W, Werwatz A, Müller M, Sperlich S (2004) Nonparametric and semiparametric models. *Series in statistics*. Springer
- Hill M (2012) Convergence of random Fourier series. In: REU participant papers
- Koch T, Hiller B, Pfetsch ME, Schewe L (2015) Evaluating gas network capacities. *MOS-SIAM*
- Leugering G, Schmidt EJPG (2002) On the modelling and stabilization of flows in networks of open canals. *SICON* 41:164–180
- Linde W (2016) Probability theory. De Gruyter
- Marcus M, Pisier G (1981) Random Fourier series with applications to harmonic analysis. *Annals of mathematics studies*. Princeton University Press
- Nadaraya EA (1965) On non-parametric estimates of density functions and regression curves. *Theory Prob Appl* 10:186–190
- Parzen E (1962) On estimation of a probability density function and mode. *Ann Math Stat* 33:1065–1076
- Prékopa A (1995) Stochastic programming, 1st edn. Springer
- Roache P (1998) Verification and validation in computational science and engineering. Bertrams
- Schwer L (2007) An overview of the ptc 60/v&v 10: guide for verification and validation in computational solid mechanics: transmitted by L. E. Schwer, chair ptc 60v&v 10. *Eng Comput (Lond)* 23:245–252
- Scott DW (2015) Multivariate density estimation: theory, practice, and visualization, 2nd edn. Wiley
- Scott DW, Terrell GR (1992) Variable kernel density estimation. *Ann Math Stat* 20:1236–1265
- Shapiro A, Dentcheva D, Ruszczyński A (2009) Lectures on stochastic programming: modeling and theory. MPS-SIAM
- Thompson W (1992) Fourier series and the gibbs phenomenon. *Am J Phys* 60:425–429
- Turlach BA (1993) Bandwidth selection in kernel density estimation: a review. Discussion paper, Humboldt-Univ
- 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, Frangioni A, Oliveira W (2016) Inexact stabilized benders' decomposition approaches: with application to chance-constrained problems with finite support. *Comput Optim Appl* 65:637–669
- van Ackooij W, Aleksovska I, Zuniga MM (2018) (Sub-)differentiability of probabilistic functions with elliptical distributions. *Set-Valued Var Anal* 26:887–910
- van Ackooij W, Henrion R, Pérez-Aros P (2020) Generalized gradients for probabilistic/robust (probust) constraints. *Optimization* 69:1451–1479
- Wand MP, Jones MC (1993) Comparison of smoothing parametrizations in bivariate kernel density estimation. *J Am Stat Assoc* 88:520–528