Appl. Set-Valued Anal. Optim. 7 (2025), No. 2, pp. 209-221
Available online at http://asvao.biemdas.com
https://doi.org/10.23952/asvao.7.2025.2.05

ON THE CONVERGENCE OF OPTIMIZATION PROBLEMS WITH KERNEL DENSITY ESTIMATED PROBABILISTIC CONSTRAINTS

MICHAEL SCHUSTER

Department of Mathematics, Friedrich-Alexander-Universität Erlangen-Nürnberg, Germany

Abstract. Uncertainty plays a significant role in applied mathematics and probabilistic constraints are widely used to model uncertainty in various fields even if probabilistic constraints often demand computational challenges. Kernel density estimation (KDE) provides a data-driven approach for properly estimating probability density functions and efficiently evaluating corresponding probabilities. In this paper, we investigate optimization problems with probabilistic constraints, where the probabilities are approximated using a KDE approach. We establish sufficient conditions under which the solution of the KDE approximated optimization problem converges to the solution of the original problem as the sample size goes to infinity. The main results of this paper include three theorems: (1) For sufficiently large sample sizes, the solution of the original problem is also a solution of the approximated problem, if the probabilistic constraint is passive; (2) The limit of a convergent sequence of solutions of the approximated problems is a solution of the original problem, if the KDE uniformly converges; (3) We provide sufficient conditions for the existence of a convergent sequence of solutions of the approximated problems.

Keywords. Chance constraints; Kernel density estimation; Convergence analysis; Probabilistic constrained optimization; Stochastic optimization.

2020 Mathematics Subject Classification. 90C15, 62G07.

1. Introduction and Motivation

Uncertainty is always present in real world applications. Insofar it is important for decision-makers to consider this uncertainty while making optimal choices. Mathematically, this necessitates a paradigm shift from deterministic optimization towards embracing the complexities of stochastic environments. Real-world parameters are usually not completely unknown, and they often follow a probability distribution. In this context, probabilistic constraints (also known as chance constraints) emerge as an excellent modelling tool to represent this uncertainty. Charnes, Cooper, and Symonds are often mentioned as the pioneers in this area with their works in the late 1950s on production planning; see [1, 2]. In [3] and [4] the authors provided a detailed overview and introduction on the broad area of stochastic programming and optimization with probabilistic constraints.

*Corresponding author.
E-mail address: michi.schuster@fau.de
Received 12 June 2024; Accepted 8 March 2025; Published online 25 March 2025.

Probabilistic constraints as modelling tools for uncertainty are now highly used in application driven problems. In [5], the authors gave an introduction to probabilistic constraints in energy management. Recent research results acknowledge the importance of probabilistic constraints in applications related to energy management; see [6, 7, 8] for an application to electricity networks; see [9, 10, 11, 12, 13] for an application to gas networks; see [14, 15, 16] for general hyperbolic problems. However, it is well-recognized that evaluating probabilistic constraints may be challenging due to high-dimensional probability distributions, or even unknown probability distributions, and the computational demands they impose can hinder efficient optimization. Kernel density estimation (see, e.g., [17, 18]) is a well-known and powerful statistical technique that first arose in the early 1950s in mathematics and statistics, where [19] is often cited as a foundational paper. Kernel density estimation provides a flexible and data-driven approach to approximate probability density functions and thus, to evaluate probabilistic constraints. Apart from the above mentioned books, [20] and [21] gave an excellent overview about kernel density estimation including discussions on kernel functions, bandwidth selection and convergence analysis.

In this paper, we consider optimization problems with probabilistic constraints, where the probabilistic constraint is approximated by a kernel density estimator (KDE). This approach was recently applied to evaluate probabilities in the context of gas network optimization; see [12, 13, 16]. The convergence of the KDE was highly discussed in the past; see, e.g., [19, 22] for the very early results. The convergence theory of the kernel density estimator and the corresponding probabilistic constraints was perfectly summarized in [17] and it was well explained in [20]. Convergence rates for the KDE were analyzed in, for example, [23, 24]. To our best knowledge, the convergence of optimization problems with KDE approximated probabilistic constraints was not analyzed before. In this paper, we show that the solution of optimization problems with probabilistic constraints can be approximated by the solution of optimization problems with KDE approximated probabilistic constraints in the sense that if the number of samples goes to infinity, then the solution of the approximated problem converges to the solution of the exact problem.

The paper is organized as follows: In Section 2, we first state our probabilistic constrained optimization problem. We then shortly introduce the concept of kernel density estimation and further, we state the main results of this work. The main results consist of three theorems. In the first theorem, we state that if the probabilistic constraint is passive, then we can always find a sufficiently large number of samples, s.t. the approximated optimization problem has the same solution as the original problem. In the second theorem, we show that if the kernel density estimator converges uniformly, then the limit of a convergent sequence of solutions of the approximated problem is a solution of the original problem. In the last theorem, we state sufficient conditions for the existence of a convergent sequence of solutions of the approximated problem. In Section 3.1, we first prove some auxiliary lemmas that are needed for the proofs of the main results. In Section 3.2, we state the proofs of the three main theorems.

2. Mathematical Modelling

In this section, we introduce and motivate our probabilistic constrained optimization problem, we introduce and discuss kernel density estimation, and we approximate the probabilistic

constraint by using a kernel density estimator (KDE) approach, in which the probability is evaluated by integrating the approximated probability density function.

2.1. **Optimization with probabilistic constraints.** We use the notation $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty, \infty\}$. Consider the constrained optimization problem

$$\begin{aligned} \min_{x \in \mathscr{V}} \quad & f(x) \\ \text{s.t.} \quad & \tilde{g}_i(x) \leq 0 \qquad i = 1, \cdots, m \in \mathbb{N}, \\ & x \in \mathscr{X} \end{aligned}$$

with a Banach space $\mathscr{V}$, a convex and monotonously increasing function $f : \mathscr{V} \to \bar{\mathbb{R}}$ and constraints $\tilde{g}_i : \mathscr{V} \to \mathbb{R}$ for $i = 1, \cdots, m \in \mathbb{N}$. Let $\emptyset \neq \mathscr{X} \subseteq \mathscr{V}$ be a closed set defined by further constraints, which does not depend on the uncertainty later. In various applications like finance, fluid transport, telecommunication, (chemical) engineering, and much more, the constraints may be perturbed or may depend on an uncertain parameter $\xi \in \mathbb{R}^d$, e.g., due to demographic, meteorological, environmental and other reasons. For $i = 1, \cdots, m$, we define the constraints

$$g_i : \mathscr{V} \times \mathbb{R}^d \to \mathbb{R}, \qquad g_i : (x, \xi) \mapsto g_i(x, \xi).$$

Often, decisions x have to be taken before the uncertainty ξ that is sufficiently observed. In this paper, we focus on the case where the uncertain parameter is given by a probability distribution. Consider a random variable ξ with absolutely continuous probability distribution function $\mathscr{F}_\xi$ on an appropriate probability space $(\Omega, \mathscr{A}, \mathbb{P})$ with $\xi(\omega) \in \mathbb{R}^d$ for $\omega \in \Omega$. In the following, we write ξ instead of $\xi(\omega)$ for a realization of the random variable ξ. Note that $g(x, \cdot)$ as a vector of the g_i is also a m-dimensional random variable, which we assume to be absolutely continuous. Obviously, there is no guarantee that $g(x, \xi) \leq 0$ for all realizations of ξ. Due to the fact that the distribution of ξ is known, we can evaluate the probability $\mathbb{P}\big(g(x, \xi) \leq 0\big)$ and we agree on a decision x, if the probability larger than some probability level $\alpha \in (0, 1)$. This leads to the probabilistic constrained optimization problem

$$\left\{ \begin{aligned} \min_{x \in \mathscr{V}} \quad & f(x), \\ \text{s.t.} \quad & \mathbb{P}(\, g(x, \xi) \leq 0 \,) \;\geq\; \alpha, \\ & x \in \mathscr{X}. \end{aligned} \right. \tag{P_∞}$$

The main difficulty in analyzing and solving probabilistic constrained optimization problems like (P_∞) clearly emerges when evaluating the probabilistic constraint resp. when computing the probability. Due to the absolute continuity of ξ, a probability density function ρ_ξ of ξ exists and the probability depending on a decision x can be evaluated by integrating the probability density function over the set of feasible realizations, i.e.,

$$\mathbb{P}(\, g(x, \xi) \leq 0 \,) = \int_{M(x)} \rho_\xi(z)\, dz, \tag{2.1}$$

where the set of feasible realizations $M(x)$ is defined as ($i = 1, \cdots, m$)

$$M(x) := \big\{\, \omega \in \Omega \mid g_i(x, \xi(\omega)) \leq 0 \,\big\} \;\subseteq\; \mathbb{R}^d. \tag{2.2}$$

It is not difficult to see that evaluating this integral might causes problems, due to a high dimension d or due to an unknown structure of the set M. In [25], the authors gave an excellent

overview about how to efficiently integrate Gaussian distributions over polyhedral sets, but we neither focus on Gaussian distributions nor we restrict ourselves to polyhedral sets. In fact, since we have no restrictions on g, we can not even assume convexity of $M(x)$. If one is able to get sufficient information on M, then, for Gaussian distributions, in [26] the authors suggested to apply the *spheric radial decomposition* (see also [27, 28]) to efficiently evaluate the probability (2.1). Since this is usually not the case, we rewrite the probability (2.1) as

$$\mathbb{P}\left(g(x,\xi) \leq 0 \right) = \int\limits_{(-\infty,0]^m} \rho_{g(x)}(z)\, dz, \tag{2.3}$$

where $\rho_{g(x)}$ is the probability density function of the random variable $g(x,\cdot)$. So the set of integration now is quite convenient for evaluating the probability, but the probability density function is a priori unknown. This is where we apply kernel density estimation.

2.2. **Kernel density estimation.** In this section, we shortly introduce the concept of kernel density estimation (see [18, 20]), a non-parametric method used to estimate the probability density function of a random variable. Consider an independent and identically distributed sample

$$\mathscr{S}_\xi := \{\, y_1, \cdots, y_{N_{\mathrm{KDE}}} \,\} \subseteq \mathbb{R}^d,$$

of the random variable ξ. For every $y_i \in \mathscr{S}$ ($i \in \{1, \cdots, N_{\mathrm{KDE}}\}$), define the vector $Y_i(x) := g(x, y_i)$. Thus, for $x \in \mathscr{V}$, the sample

$$\mathscr{S}_{g(x)} := \{\, Y_i, \cdots, Y_{N_{\mathrm{KDE}}} \,\} \subseteq \mathbb{R}^m,$$

is an independent and identically distributed sample of the random variable $g(x,\cdot)$. For a multivariate kernel function $K : \mathbb{R}^m \to \mathbb{R}_{\geq 0}$ and a bandwidth matrix $H \in \mathbb{R}^{m\times m}$, the kernel density estimator (KDE) is defined by

$$\rho_{N_{\mathrm{KDE}},g(x)}(z) = \frac{1}{N_{\mathrm{KDE}}\sqrt{\det H}} \sum_{i=1}^{N_{\mathrm{KDE}}} K\left(H^{-\frac{1}{2}}(z - Y_i) \right), \tag{2.4}$$

which is an approximation of the probability density function $\rho_{g(x)}$. Note that H must be chosen suitable such that $H^{-1/2}$ exists. The corresponding probability is given by

$$\mathbb{P}_{N_{\mathrm{KDE}}}\left(g(x,\xi) \leq 0 \right) = \int\limits_{(-\infty,0]^m} \rho_{N_{\mathrm{KDE}},g(x)}(z)\, dz, \tag{2.5}$$

which is an approximation of the probability (2.3). In [12] and [16] the authors applyed a Gaussian product kernel and a diagonal bandwidth matrix, which leads to efficient numerical evaluation of the KDE. Different kernel functions were discussed in [29, 30] and the choice of bandwidth matrix was discussed in [31, 32].

Under slight assumptions on the bandwidth matrix (see, e.g., [17, Chapter 6, Theorem 1] and [20, Chapter 3.4]), the KDE L^1-converges $\mathbb{P}$-almost surely to the exact density function if the number of samples goes to infinity. That is,

$$\left\| \rho_{N_{\mathrm{KDE}},g(x)} - \rho_{g(x)} \right\|_{L^1(\mathbb{R}^m)} \xrightarrow{N_{\mathrm{KDE}}\to\infty} 0 \qquad \mathbb{P}-\text{almost surely.} \tag{2.6}$$

Further, due to *Scheffé's lemma* (see [17, Chapter 1, Theorem 1], the approximated probability (2.5) also converges to the exact probability (2.1). That is,

$$\left| \mathbb{P}_{N_{\mathrm{KDE}}}\left(g(x,\xi) \geq 0 \right) - \mathbb{P}\left(g(x,\xi) \geq 0 \right) \right| \xrightarrow{N_{\mathrm{KDE}}\to\infty} 0 \qquad \mathbb{P}-\text{almost surely.} \tag{2.7}$$

We now replace the probabilistic constraint in (P_∞) with the KDE approximated probability, and we define the approximated probabilistic constrained optimization problem

$$\begin{cases} \min\limits_{x\in\mathscr{V}} & f(x), \\ \text{s.t.} & \mathbb{P}_{N_{\mathrm{KDE}}}\left(g(x,\xi) \leq 0 \right) \geq \alpha, \\ & x \in \mathscr{X}. \end{cases} \tag{$\mathrm{P}_{N_{\mathrm{KDE}}}$}$$

Even if the approximated probability converges to the exact probability, it is not a priori clear that a sequence of solutions of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) also converges to a solution of (P_∞). In the next section, we give sufficient conditions for the convergence of optimal solutions of the approximated problem.

3. Main Results

In this section, we present the main results of this paper. We state sufficient conditions for the convergence of a sequence of solutions of the approximated probabilistic constrained optimization problem ($\mathrm{P}_{N_{\mathrm{KDE}}}$) to a solution of the exact probabilistic constrained optimization problem (P_∞). In fact, even if the convergence of the probability was stated in (2.7), it is not a priori clear whether the limit of a convergent sequence of solutions of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) is a solution to ($\mathrm{P}_\infty$) or if a convergent sequence of solutions to ($\mathrm{P}_{N_{\mathrm{KDE}}}$) even exists. In [12], the author presented a result. Under certain assumptions, for N_{KDE} sufficiently large, the solution of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) is close to a solution of (P_∞). In the following assumptions, under which a sequence of solutions of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) converges to a solution of (P_∞).

Throughout the last paragraphs, we already made a few assumptions that we summarize in the following paragraph. Consider a Banach space $\mathscr{V}$, a closed and nonempty set $\mathscr{X} \subseteq \mathscr{V}$, a continuous and monotonously increasing function $f : \mathscr{V} \to \bar{\mathbb{R}}$, a function $g : \mathbb{V} \times \mathbb{R}^d \to \mathbb{R}^m$, and a random variable ξ with probability density function ρ_ξ on an appropriate probability space $(\Omega, \mathscr{A}, \mathbb{P})$. Consider the probabilistic constrained optimization problem

$$\begin{cases} \min\limits_{x\in\mathscr{V}} & f(x), \\ \text{s.t.} & \mathbb{P}\left(g(x,\xi) \leq 0 \right) \geq \alpha, \\ & x \in \mathscr{X}, \end{cases} \tag{P_∞}$$

where the probability is evaluated by (2.3). Consider the approximated probabilistic constrained optimization problem

$$\begin{cases} \min\limits_{x\in\mathscr{V}} & f(x), \\ \text{s.t.} & \mathbb{P}_{N_{\mathrm{KDE}}}\left(g(x,\xi) \leq 0 \right) \geq \alpha, \\ & x \in \mathscr{X}, \end{cases} \tag{$\mathrm{P}_{N_{\mathrm{KDE}}}$}$$

where the probability is evaluated approximately by the KDE approach stated in (2.5).

Henceforth we assume that (P_∞) as well as ($P_{N_{KDE}}$) have at least one solution. The following theorems provide new results about optimization problems with kernel density estimator approximated probabilistic constraints:

Theorem 3.1. *Let $x^* \in \mathcal{V}$ be a solution of (P_∞) with $\mathbb{P}(\, g(x^*, \xi) \geq 0\,) > \alpha$. Then there exists N^* sufficiently large, s.t. x^* is a solution to ($P_{N_{KDE}}$) for all $N_{KDE} \geq N^*$ $\mathbb{P}$-almost surely.*

Theorem 3.2. *If the KDE $\mathbb{P}-$almost surely uniformly converges for every $x \in \mathcal{X}$, and if $(x^*_{N_{KDE}})_{N_{KDE} \geq 1}$ is a convergent sequence of solutions of ($P_{N_{KDE}}$), then its limit x^* $\mathbb{P}$-almost surely is a solution to (P_∞).*

Theorem 3.3. *If x^* is a unique solution to (P_∞) and if there exists a positive pair (ε, δ) such that, for all $x \in \mathcal{X} \setminus \mathcal{B}_\delta(x^*)$ with $\mathbb{P}(\, g(x, \xi) \leq 0\,) \geq \alpha$, $f(x) > f(x^*) + \varepsilon$, then there $\mathbb{P}$-almost surely exists a convergent sequence of solutions of ($P_{N_{KDE}}$).*

The first theorem can be interpreted as follows: If the probabilistic constraint in the original problem is passive, then the original problem can be seen as a deterministic problem locally around the optimal solution. Then we can $\mathbb{P}$-almost surely find a sufficiently large number of samples such that a solution of the original problem is also a solution to the approximated problem.

The second theorem states that uniform convergence of the KDE is sufficient such that the limit of a convergent sequence of solutions of the approximated problems is a solution to the original problem. Uniform convergence of the KDE is a rather weak assumption. In [33] the authors state that the uniform convergence only depends on the kernel functions and the bandwidth parameter if the original probability density function is uniformly continuous. Due to the existence of a probability density function, we already assumed the absolute continuity of the original probability distribution function. Since we have no additional restrictions on the kernel function and the bandwidth matrix apart from the KDE convergence stated in *Section 2.2*, both parameters can easily be chosen to guarantee uniform convergence, e.g., a Gaussian kernel function and a diagonal bandwidth matrix, as it was done in [16].

The third theorem gives a sufficient condition for the existence of a convergent sequence of solutions. The assumption in *Theorem 3.3* basically states that the objective value can only be infinitesimal close to the optimal objective value in an area close to the optimal solution. Note that the theory of convex optimization implies the assumption, i.e., the assumption is satisfied if the admissible set of (P_∞) is convex. Additionally, if the objective function is strictly convex, then there exists a unique solution of (P_∞) and thus also a convergent sequence of solutions of the approximated problems ($P_{N_{KDE}}$).

4. Proof of the Main Results

In this section, we provide the proofs of *Theorem 3.1*, *Theorem 3.2* and *Theorem 3.3*. We further provide some auxiliary lemmas that are used for the proof. The proof of *Theorem 3.2* is based on a convergence result stated in [34]. First we state some definitions regarding the notation.

We define the admissible set $\mathcal{A}_\infty$ of optimization problem (P_∞) as

$$\mathcal{A}_\infty := \{\, x \in \mathcal{V} \mid \mathbb{P}(\, g(x^*, \xi) \leq 0\,) \geq \alpha,\, x \in \mathcal{X}\,\},$$

and we define the admissible set $\mathscr{A}_{N_{\mathrm{KDE}}}$ of optimization problem ($\mathrm{P}_{N_{\mathrm{KDE}}}$) as

$$\mathscr{A}_{N_{\mathrm{KDE}}} := \{\, x \in \mathscr{V} \mid \mathbb{P}_{N_{\mathrm{KDE}}}(\, g(x^*,\xi) \le 0\,) \;\ge\; \alpha,\, x \in \mathscr{X} \,\}.$$

4.1. **Auxiliary lemmas.**

Lemma 4.1. *Under the assumptions of Theorem 3.2, the probabilistic constraint in* ($\mathrm{P}_{N_{\mathrm{KDE}}}$) $\mathbb{P}-$*almost surely uniformly converges to the probabilistic constraint of* (P_∞) *on* $\mathscr{X}$ *as* N_{KDE} *goes to* ∞.

Proof of Lemma 4.1. Considering the probabilistic constraints $G_{N_{\mathrm{KDE}}}(x) \le 0$ of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) with

$$G_{N_{\mathrm{KDE}}}(x) := \alpha - \mathbb{P}_{N_{\mathrm{KDE}}}(\, g(x,\xi) \le 0\,),$$

and $G(x) \le 0$ of (P_∞) with $G(x) := \alpha - \mathbb{P}(\, g(x,\xi) \le 0\,)$, we have

$$\begin{aligned}
&\lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left|\, G_{N_{\mathrm{KDE}}}(x) - G(x) \,\right| \\
&= \lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left|\, \mathbb{P}(\, g(x^*,\xi) \le 0\,) - \mathbb{P}_{N_{\mathrm{KDE}}}(\, g(x^*,\xi) \le 0\,) \,\right| \\
&= \lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left| \int_{(-\infty,0]^m} \rho_{g(x)}(z)\, dz - \int_{(-\infty,0]^m} \rho_{g(x),N_{\mathrm{KDE}}}(z)\, dz \right| \\
&= \lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left| \int_{(-\infty,0]^m} \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z)\, dz \right|.
\end{aligned}$$

We use the triangular inequality to obtain

$$\begin{aligned}
&\lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left| \int_{(-\infty,0]^m} \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z)\, dz \right| \\
&\le \lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \int_{(-\infty,0]^m} \left|\, \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z) \,\right| dz.
\end{aligned}$$

Now we swap the supremum and the integral, which leads to

$$\begin{aligned}
&\lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \int_{(-\infty,0]^m} \left|\, \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z) \,\right| dz \\
&\le \lim_{N_{\mathrm{KDE}}\to\infty} \int_{(-\infty,0]^m} \sup_{x\in\mathscr{X}} \left|\, \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z) \,\right| dz.
\end{aligned}$$

Since the KDE $\mathbb{P}-$almost surely uniformly converges for every $x \in \mathscr{X}$, it also uniformly converges for the supremum of all $x \in \mathscr{X}$ and we can swap the limit and the integral, i.e.,

$$\begin{aligned}
&\lim_{N_{\mathrm{KDE}}\to\infty} \int_{(-\infty,0]^m} \sup_{x\in\mathscr{X}} \left|\, \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z) \,\right| dz \\
&= \int_{(-\infty,0]^m} \lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left|\, \rho_{g(x)}(z) - \rho_{g(x),N_{\mathrm{KDE}}}(z) \,\right| dz \qquad \mathbb{P}-\text{almost surely.}
\end{aligned}$$

Thus

$$\lim_{N_{\mathrm{KDE}}\to\infty} \sup_{x\in\mathscr{X}} \left| G_{N_{\mathrm{KDE}}}(x) - G(x) \right| = 0 \qquad \mathbb{P}-\text{almost surely},$$

and the lemma is proven. □

Lemma 4.2. *Under the assumptions of Theorem 3.2, if the number of samples N_{KDE} goes to infinity, then the admissible set $\mathscr{A}_{N_{KDE}}$ $\mathbb{P}$-almost surely converges to $\mathscr{A}_\infty$ in the sense that*

$$\lim_{N_{KDE}\to\infty} \mathscr{A}_{N_{KDE}} \subseteq \mathscr{A}_\infty \qquad \textit{and} \qquad \mathscr{A}_\infty \subseteq \lim_{N_{KDE}\to\infty} \mathscr{A}_{N_{KDE}}.$$

Proof of Lemma 4.2. Consider the probabilistic constraints $G_{N_{\mathrm{KDE}}}(x) \leq 0$ of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) with

$$G_{N_{\mathrm{KDE}}}(x) := \alpha - \mathbb{P}_{N_{\mathrm{KDE}}}(g(x,\xi) \leq 0),$$

and $G(x) \leq 0$ of (P_∞) with

$$G(x) := \alpha - \mathbb{P}(g(x,\xi) \leq 0).$$

Consider a convergent sequence $x_{N_{\mathrm{KDE}}} \to x^*$ with $x_{N_{\mathrm{KDE}}} \in \mathscr{A}_{N_{\mathrm{KDE}}}$. Due to *Lemma 4.1*, the corresponding sequence $G_{N_{\mathrm{KDE}}}(x_{N_{\mathrm{KDE}}})$ $\mathbb{P}$-almost surely also converges and

$$\lim_{N_{\mathrm{KDE}}\to\infty} G_{N_{\mathrm{KDE}}}(x_{N_{\mathrm{KDE}}}) \to G(x^*) \qquad \mathbb{P}\text{-almost surely}.$$

Since $G_{N_{\mathrm{KDE}}}(x_{N_{\mathrm{KDE}}}) \to G(x^*)$ is a sequence of non positive numbers and the interval $(-\infty,0]$ is a closed set, the limit $G(x^*)$ is also a non positive number. Thus $G(x^*) \leq 0$, i.e.,

$$\mathbb{P}(g(x^*,\xi) \leq 0) \geq \alpha \qquad \mathbb{P}\text{-almost surely},$$

which implies $x^* \in \mathscr{A}_\infty$ $\mathbb{P}$-almost surely. The other way around, for $x \in \mathscr{A}_\infty$, we have

$$\alpha \leq \mathbb{P}(g(x,\xi) \leq 0) = \int_{(-\infty,0]^m} \rho_{g(x)}(z)\, dz.$$

From the convergence of the KDE, we have

$$\alpha \leq \int_{(-\infty,0]^m} \lim_{N_{\mathrm{KDE}}\to\infty} \rho_{g(x),N_{\mathrm{KDE}}}(z)\, dz \qquad \mathbb{P}\text{-almost surely},$$

and from the uniform convergence of the KDE for every $x \in \mathscr{X}$, we can swap the limit and the integral. This implies

$$\alpha \leq \lim_{N_{\mathrm{KDE}}\to\infty} \int_{(-\infty,0]^m} \rho_{g(x),N_{\mathrm{KDE}}}(z)\, dz \qquad \mathbb{P}\text{-almost surely},$$

which is equivalent to

$$\lim_{N_{\mathrm{KDE}}\to\infty} \mathbb{P}_{N_{\mathrm{KDE}}}(g(x,\xi) \leq 0) \geq \alpha \qquad \mathbb{P}\text{-almost surely}.$$

Thus we have $x \in \lim_{N_{\mathrm{KDE}}\to\infty} \mathscr{A}_{N_{\mathrm{KDE}}}$ $\mathbb{P}$-almost surely and the lemma is proven. □

Lemma 4.3. *Under the assumptions of Theorem 3.2, for every $x \in \mathscr{A}_\infty$, there $\mathbb{P}$-almost surely exists a convergent sequence $(x_{N_{KDE}})_{N_{KDE}\geq 1}$ with $x_{N_{KDE}} \in \mathscr{A}_{N_{KDE}}$ and $x_{N_{KDE}} \to x$ as $N_{KDE} \to \infty$.*

Proof of Lemma 4.3. Under the assumptions of *Theorem 3.2*, the admissible set $\mathscr{A}_\infty$ is non empty. For $x^* \in \mathscr{A}_\infty$, we define the sequence $(x_{N_{\mathrm{KDE}}})_{N_{\mathrm{KDE}}\geq 1}$ with $x_{N_{\mathrm{KDE}}} \in \mathscr{A}_{N_{\mathrm{KDE}}}$ such that the distance between x^* and $x_{N_{\mathrm{KDE}}}$ is minimal regarding the norm $\|\cdot\|$ of the Banach space, i.e.,

$$x_{N_{\mathrm{KDE}}} := \operatorname*{arg\,min}_{x\in\mathscr{A}_{N_{\mathrm{KDE}}}} \|x - x^*\|. \tag{4.1}$$

Note that, on the one hand, $x_{N_{\mathrm{KDE}}}$ is equal to x^* if and only if $x^* \in \mathscr{A}_{N_{\mathrm{KDE}}}$. On the other hand, it is not important whether the arg min is unique. Assume that $(x_{N_{\mathrm{KDE}}})_{N_{\mathrm{KDE}}\geq 1}$ does not converge to x^*. Then there exists $\varepsilon > 0$ such that

$$\|x_{N_{\mathrm{KDE}}} - x^*\| > \varepsilon \text{ for } N_{\mathrm{KDE}} \geq 1,$$

which is a contradiction to *Lemma 4.2*. Thus $x_{N_{\mathrm{KDE}}} \to x^*$ for $N_{\mathrm{KDE}} \to \infty$ and the lemma is proven. □

4.2. **Proof of the Main Results.**

Proof of Theorem 3.1. Let a solution $x^* \in \mathscr{X}$ with $\mathbb{P}(\, g(x^*,\xi) \leq 0\,) = \alpha + \varepsilon$ for $\varepsilon > 0$ be given. For $\varepsilon > 0$, we can find N^* sufficiently large, s.t. $\mathbb{P}$-almost surely. Then

$$\left|\, \mathbb{P}(\, g(x^*,\xi) \leq 0\,) - \mathbb{P}_{N_{\mathrm{KDE}}}(\, g(x^*,\xi) \leq 0\,) \,\right| \leq \varepsilon \qquad \forall N_{\mathrm{KDE}} \geq N^*. \tag{4.2}$$

This implies that $\mathbb{P}_{N_{\mathrm{KDE}}}(\, g(x^*,\xi) \leq 0\,) > \alpha$ for $N_{\mathrm{KDE}} \geq N^*$ (cf. (2.6), (2.7)). Thus x^* is admissible for ($\mathrm{P}_{N_{\mathrm{KDE}}}$). Assume that for $N_{\mathrm{KDE}} \geq N^*$ there $\mathbb{P}$-almost surely exists $\underline{x} \in \mathscr{A}_{N_{\mathrm{KDE}}}$ with $f(\underline{x}) < f(x^*)$. Then $\underline{x} \notin \mathscr{A}_\infty$ because otherwise this contradicts the optimality of x^*. Define

$$\bar{x} := \underline{x} + \lambda(x^* - \underline{x}),$$

with $\lambda \in (0,1)$, s.t. $\bar{x} \in \mathscr{A}_{N_{\mathrm{KDE}}} \cap \mathscr{A}_\infty$. Then $f(\underline{x}) < f(x^*) \leq f(\bar{x})$, due to the optimality of $\underline{x}$ and x^*. But this implies that the epigraph of f is not convex, which is a contradiction to the convexity of f. Thus, for $N_{\mathrm{KDE}} \geq N^*$, x^* is $\mathbb{P}$-almost surely a solution to ($\mathrm{P}_{N_{\mathrm{KDE}}}$). □

As mentioned before, for the proof of *Theorem 3.2*, we apply a result from [34], which requires additional notation. Let

$$\mathbf{IL}\{\mathscr{A}_{N_{\mathrm{KDE}}}\} := \{\, x \in \mathscr{V} \mid x = \lim_{N_{\mathrm{KDE}}\to\infty} x_{N_{\mathrm{KDE}}}, \text{ where } x_{N_{\mathrm{KDE}}} \in \mathscr{A}_{N_{\mathrm{KDE}}} \text{ for all but a finite number of } N_{\mathrm{KDE}}\text{'s } \},$$

be the set of inner limits of $\mathscr{A}_{N_{\mathrm{KDE}}}$. Further let $\mathscr{M}_{N_{\mathrm{KDE}}}$ and $\mathscr{M}_\infty$ be the set of optimal solutions of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) and ($\mathrm{P}_\infty$), respectively. For the reader's convenience, in the following, we state the theorem, which we apply in the proof of *Theorem 3.2*.

Theorem 4.1. ***(cf.*** [34, Theorem 2.3]***)***
Let the following assumptions hold:

(1) The minimum sets $\mathscr{M}_{N_{KDE}}$ satisfies $\mathbf{IL}\{M_{N_{KDE}}\} \subseteq \mathscr{A}_\infty$.
(2) If x_{n_i} is in $\mathscr{M}_{n_i}$ for all i and $x_{n_i} \to x$ for $i \to \infty$, where $\{n_i\}$ is an increasing sequence, then $\liminf_{i\to\infty} f(x_{n_i}) \geq f(x)$.
(3) For each $x \in \mathscr{A}_\infty$, there exist $N \in \mathbb{N}$ and a sequence $\{x_{N_{KDE}}\}$ with $x_{N_{KDE}} \in \mathscr{A}_{N_{KDE}}$ for all $N_{KDE} \geq N$, and such that $x_{N_{KDE}} \to x$ and $\lim_{N_{KDE}\to\infty} f(x_{N_{KDE}}) = f(x)$.

Then $\mathbf{IL}\{M_{N_{KDE}}\} \subseteq \mathscr{M}_\infty$. In addition, if $\mathbf{IL}\{\mathscr{M}_{N_{KDE}}\}$ is nonempty, then the optimal values also converge.

Note that [34, Theorem 2.3] was adapted to our notation here. Now we can formulate the proof of *Theorem 3.2*.

Proof of Theorem 3.2. For this proof, we show that the assumptions of *Theorem 4.1* (cf. [34]) are satisfied. Consider a convergent sequence of solutions $x^*_{N_{\mathrm{KDE}}}$ with limit x^*_∞ and $x^*_{N_{\mathrm{KDE}}} \in \mathscr{M}_{N_{\mathrm{KDE}}}$. Since $\mathscr{X}$ is closed, we have $x^*_\infty \in \mathscr{X}$. The corresponding sequence $G_{N_{\mathrm{KDE}}}(x^*_{N_{\mathrm{KDE}}})$ is a sequence of non positive numbers. Due to the uniform convergence shown in *Lemma 4.1*, we have

$$\lim_{N_{\mathrm{KDE}}\to\infty} G_{N_{\mathrm{KDE}}}(x^*_{N_{\mathrm{KDE}}}) = G(x^*_\infty).$$

Assume that $G(x^*_\infty) > 0$. Then there exists $N^* \in \mathbb{N}$, s.t. $G_{N_{\mathrm{KDE}}}(x^*_{N_{\mathrm{KDE}}}) > 0$ for $N_{\mathrm{KDE}} \geq N^*$, which is a contradiction to $x_{N_{\mathrm{KDE}}} \in \mathscr{M}_{N_{\mathrm{KDE}}}$. Thus $G(x^*_\infty) \leq 0$ and $\mathbf{IL}\{\mathscr{M}_{N_{\mathrm{KDE}}}\} \subseteq \mathscr{A}_\infty$, that is, the first assumption holds of *Theorem 4.1* holds.

Consider the sequence (x_{n_i}) from *Theorem 4.1*, assumption 2. Since f is continuous and monotonously increasing, we have $\liminf_{i\to\infty} f(x_{n_i}) = f(x)$. Thus the second assumption of *Theorem 4.1* holds.

For each $x \in \mathscr{A}_\infty$, the existence of a convergent sequence $\{x_{N_{\mathrm{KDE}}}\}$ with limit x and $x_{N_{\mathrm{KDE}}} \in \mathscr{A}_{N_{\mathrm{KDE}}}$ for all $N_{\mathrm{KDE}} \geq N \in \mathbb{N}$ follows directly from *Lemma 4.3*. Due to the continuity of f, we have

$$\lim_{N_{\mathrm{KDE}}\to\infty} f(x_{N_{\mathrm{KDE}}}) = f\Big(\lim_{N_{\mathrm{KDE}}\to\infty} x_{N_{\mathrm{KDE}}}\Big) = f(x).$$

Thus the third assumption in *Theorem 4.1* also holds. This implies

$$\mathbf{IL}\{M_{N_{\mathrm{KDE}}}\} \subseteq \mathscr{M}_\infty,$$

i.e., the limits of all sequences of solutions of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) is a solution to ($\mathrm{P}_\infty$). □

Proof of Theorem 3.3. Consider a sequence of solutions $(x^*_{N_{\mathrm{KDE}}})_{N_{\mathrm{KDE}}\geq 1}$ of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) and recall the assumption that there exists a positive pair (ε, δ) such that

$$f(x) > f(x^*) + \varepsilon, \quad \forall x \in \mathscr{A}_\infty \backslash \mathscr{B}_\delta(x^*), \tag{4.3}$$

where x^* is the unique solution of (P_∞). We first show that we can find N^* sufficiently large, s.t. $x^*_{N_{\mathrm{KDE}}} \in \mathscr{B}_\delta(x^*)$ $\mathbb{P}$-almost surely. Assume that $x^*_{N_{\mathrm{KDE}}} \notin \mathscr{B}_\delta(x^*)$. Due to the continuity of f, for $\frac{\varepsilon}{2} > 0$ there exists $\delta_2 > 0$, s.t.

$$\|f(x) - f(x^*_{N_{\mathrm{KDE}}})\| < \frac{\varepsilon}{2}, \quad \forall x \in \mathscr{B}_{\delta_2}(x^*_{N_{\mathrm{KDE}}}). \tag{4.4}$$

Further, due to *Lemma 4.2*, for N_1 sufficiently large, we can find $0 < \delta_3 \leq \delta_2$, s.t.

$$\mathscr{B}_\delta(x^*) \cap \mathscr{B}_{\delta_3}(x^*_{N_{\mathrm{KDE}}}) = \emptyset, \quad \forall N_{\mathrm{KDE}} \geq N_1 \quad \mathbb{P}-\text{almost surely}. \tag{4.5}$$

This implies that (4.3) holds for $x^*_{N_{\mathrm{KDE}}}$ with $N_{\mathrm{KDE}} \geq N_1$ $\mathbb{P}$-almost surely. On the other hand, for N_2 sufficiently large, we have

$$\mathscr{B}_{\delta_3}(x^*_{N_{\mathrm{KDE}}}) \cap \mathscr{A}_\infty \neq \emptyset \qquad \mathbb{P}-\text{almost surely},$$

i.e., there $\mathbb{P}-$almost surely exists $x_1 \in \mathscr{B}_{\delta_3}(x^*_{N_{\mathrm{KDE}}}) \cap \mathscr{A}_\infty$ with $x_1 \notin \mathscr{B}_\delta(x^*)$. Due to (4.4) and (4.5), we have

$$\|f(x_1) - f(x^*_{N_{\mathrm{KDE}}})\| < \frac{\varepsilon}{2}, \quad \forall N_{\mathrm{KDE}} \geq \max\{N_1, N_2\} \quad \mathbb{P}-\text{almost surely},$$

which implies

$$f(x^*_{N_{\mathrm{KDE}}}) > f(x^*) + \frac{\varepsilon}{2} \quad \mathbb{P}-\text{almost surely.} \tag{4.6}$$

Next, due to the continuity of f, for $\frac{\varepsilon}{2} > 0$, there also exists $\delta_4 > 0$, s.t.

$$\|f(x) - f(x^*)\| < \frac{\varepsilon}{2}, \quad \forall x \in \mathscr{B}_{\delta_4}(x^*). \tag{4.7}$$

Again, due to *Lemma 4.2*, for N_3 sufficiently large, we can find $0 < \delta_5 \leq \delta_4$, s.t.

$$\mathscr{A}_{N_{\mathrm{KDE}}} \cap \mathscr{B}_{\delta_5}(x^*) \neq \emptyset \qquad \mathbb{P}-\text{almost surely,}$$

i.e., there $\mathbb{P}-$almost surely exists $x_2 \in \mathscr{A}_{N_{\mathrm{KDE}}} \cap \mathscr{B}_{\delta_5}(x^*)$. Due to the continuity (4.7), we have

$$f(x_2) < f(x^*) + \frac{\varepsilon}{2} \qquad \mathbb{P}-\text{almost surely,}$$

and consequently (cf. (4.6))

$$f(x_2) < f(x^*_{N_{\mathrm{KDE}}}) \qquad \mathbb{P}-\text{almost surely,}$$

which is a contradiction to the optimality of $x^*_{N_{\mathrm{KDE}}}$. Thus, for $N^* := \max\{N_1, N_2, N_3\}$, we have

$$x^*_{N_{\mathrm{KDE}}} \in \mathscr{B}_{\delta}(x^*), \quad \forall N_{\mathrm{KDE}} \geq N^* \quad \mathbb{P}-\text{almost surely.}$$

This implies that, for all $\delta^* \geq \delta$,

$$\|x^*_n - x^*_m\| < 2\delta, \quad \forall n, m \geq N^* \quad \mathbb{P}-\text{almost surely.}$$

Due to the optimality of x^*, for all $0 < \delta^* \leq \delta$, we can find $\varepsilon^* > 0$, s.t. $f(x) > f(x^*) + \varepsilon^*$ for all $x \in \mathscr{A}_\infty \backslash \mathscr{B}_{\delta^*}(x^*)$. With the same argument as above, we can find N^* sufficiently large, s.t.

$$x^*_{N_{\mathrm{KDE}}} \in \mathscr{B}_{\delta^*}(x^*), \quad \forall N_{\mathrm{KDE}} > N^* \quad \mathbb{P}-\text{almost surely,}$$

and thus

$$\|x^*_n - x^*_m\| \leq 2\delta^*, \quad \forall n, m > N^* \quad \mathbb{P}-\text{almost surely.}$$

Hence, the sequence of solutions $(x^*_{N_{\mathrm{KDE}}})_{N_{\mathrm{KDE}} \geq 1}$ of ($\mathrm{P}_{N_{\mathrm{KDE}}}$) is a Cauchy sequence and the theorem is proven. □

5. Conclusions

In this paper, we analyzed the convergence of probabilistic constrained optimization problems, where the probabilities were approximated by a kernel density estimator approach. We shown that, under sufficient conditions, the solution of the KDE approximated optimization problems converges to the solution of the original optimization problem as the number of samples goes to infinity. Our three main results demonstrated that the KDE based approach can effectively approximate the probabilistic constraints. First we showed that, for sufficiently large sample sizes, the solution of the original probabilistic constrained optimization problem is also a solution to the KDE approximated probabilistic constrained optimization problem, if the probabilistic constraint is passive. Further we shown that the uniform convergence of the KDE is sufficient such that the limit of a convergent sequence of solutions of the approximated problems is a solution to the original problem. As third main result, we given sufficient conditions for the existence of a convergent sequence of solutions of the approximated problems.

Our theoretical framework is based on general assumptions for the optimization problems that are well-known in optimization theory. The assumptions on the KDE are rather weak and

flexible, allowing for practical applications in a wide range of stochastic environments. Specifically, the kernel functions and bandwidth parameters can be chosen to ensure smooth and continuously differentiable approximations of the probability, which, in contrast to methods like a Monte Carlo evaluation of the probability, allow the application of e.g. a gradient method. Thus combining the KDE and optimization problems with probabilistic constraints is a promising approach for solving these class of optimization problems numerically efficient.

Future research should focus on the convergence speed of optimization problems with KDE approximated probabilistic constraints. While considerable research exists on the convergence rates of the KDE itself, the connection between KDE convergence speed and convergence speed of KDE approximated optimization problems has not been analyzed yet. This relation could yield important insights into an optimal choice of kernel functions and bandwidth parameters, as well as improving the efficiency of KDE approximated probabilistic constraints in various practical applications.

Acknowledgments

This work was supported by Deutsche Forschungsgemeinschaft (DFG) in the Collaborative Research Centre CRC/Transregio 154, Mathematical Modelling, Simulation and Optimization Using the Example of Gas Networks, Project C03, Projektnummer 239904186.

REFERENCES

[1] A. Charnes, W.W. Cooper, Chance-constrained programming, Manag. Sci. 6 (1959), 73-79.
[2] A. Charnes, W.W. Cooper, G.H. Symonds, Cost horizons and certainty equivalents: An approach to stochastic programming of heating oil, Manag. Sci. 4 (1958), 235-263.
[3] A. Prékopa, Probabilistic programming, In: A. Ruszczynski, A. Shapiro (ed.) Stochastic Programming, vol. 10, pp. 267–351, Elsevier, 2003.
[4] A. Shapiro, D. Dentcheva, A. Ruszczynski, Lectures on Stochastic Programming: Modeling and Theory, MPS-SIAM, 2009.
[5] W. van Ackooij, R. Zorgati, R. Henrion, A. Möller, Chance constrained programming and its applications to energy management, In: I. Dritsas (ed.) Stochastic Optimization - Seeing the Optimal for the Uncertain, pp. 291–320, IntechOpen, 2011.
[6] S. Goettlich, R. Korn, K. Lux, Optimal control of electricity input given an uncertain demand, Math. Methods Oper. Res. 90 (2019), 1-28.
[7] L. Roald, G. Andersson, Chance-constrained AC optimal power flow: Reformulations and efficient algorithms, IEEE Trans. Power Syst. 33 (2018), 2906-2918.
[8] L. Roald, D. Pozo, A. Papavasiliou, D. Molzahn, J. Kazempour, A. Conejo, Power systems optimization under uncertainty: A review of methods and applications, Electr. Power Syst. Res. 214 (2023), 108725.
[9] D. Adelhütte, D. Aßmann, T. González Gradón, M. Gugat, H. Heitsch, R. Henrion, F. Liers, S. Nitsche, R. Schultz, M. Stingl, D. Wintergerst, Joint model of probabilistic-robust (probust) constraints applied to gas network optimization, Vietnam J. Math. 49 (2020), 1097-1130.
[10] T. González-Gradón, H. Heitsch, R. Henrion, A joint model of probabilistic/robust constraints for gas transport management in stationary networks, Comput. Manag. Sci. 14 (2017), 443-460.
[11] M. Gugat, M. Schuster, Stationary gas networks with compressor control and random loads: Optimization with probabilistic constraints, Math. Prob. Eng. 2018 (2018), 7984079.
[12] M. Schuster, Nodal control and probabilistic constrained optimization using the example of gas networks, Dissertation at FAU Erlangen-Nürnberg, Germany, 2021. https://opus4.kobv.de/opus4-trr154/frontdoor/index/index/docId/410.

[13] E. Strauch, Adaptive multi-level Monte Carlo and stochastic collocation methods for hyperbolic partial differential equations with random data on networks, Dissertation at TU Darmstadt, Germany, 2023. https://tuprints.ulb.tu-darmstadt.de/23310/.
[14] M. Hassan Farshbaf-Shaker, M. Gugat, H. Heitsch, R. Henrion, Optimal neumann boundary control of a vibrating string with uncertain initial data and probabilistic terminal constraints, SIAM J. Control. Optim. 58 (2020), 2288-2311.
[15] S. Göttlich, O. Kolb, K. Lux, Chance-constrained optimal inflow control in hyperbolic supply systems with uncertain demand, Optimal Control Appl. Methods 42 (2021), 566-589.
[16] M. Schuster, E. Strauch, M. Gugat, J. Lang, Probabilistic constrained optimization on flow networks, Optim. Eng. 23 (2022), 1-50.
[17] L. Devroye, L. Gyorfi, Nonparametric Density Estimation: The L1 View, Wiley, 1985.
[18] B.W. Silverman, Density Estimation for Statistics and Data Analysis, CRC Press, 2018.
[19] E. Parzen, On estimation of a probability density function and mode, Ann. Math. Statist. 33 (1962), 1065-1076.
[20] A. Gramacki, Nonparametric Kernel Density Estimation and Its Computational Aspects, Springer, 2018.
[21] D.W. Scott, Multivariate Density Estimation: Theory, Practice, and Visualization, Wiley, 2015.
[22] E. Nadaraya, On non-parametric estimates of density functions and regression curves, Theory Probab. Appl. 10 (1965), 186-190.
[23] T. Duong, M. Hazelton, Convergence rates for unconstrained bandwidth matrix selectors in multivariate kernel density estimation, Multivariate Anal. 93 (2005), 417-433.
[24] J. Kim, J. Shin, A. Rinaldo, L. Wasserman, Uniform convergence rate of the kernel density estimator adaptive to intrinsic volume dimension, In: K. Chaudhuri, R. Salakhutdinov (ed.) Proceedings of the 36th International Conference on Machine Learning, vol. 97, pp. 3398–3407, PMLR 2019.
[25] A. Genz, F. Bretz, Computation of Multivariate Normal and t Probabilities, Springer, 2009.
[26] C. Gotzes, H. Heitsch, R. Henrion, R. Schultz, On the quantification of nomination feasibility in stationary gas networks with random load, Math. Methods Oper. Res. 84 (2016), 427-457.
[27] I. Deák, Subroutines for computing normal probabilities of sets — computer experiences, Ann. Oper. Res. 100 (2000), 103-122.
[28] W. van Ackooij, R. Henrion, Gradient formulae for nonlinear probabilistic constraints with Gaussian and Gaussian-like distributions, SIAM J. Optim. 24 (2014), 1864-1889.
[29] W. Härdle, M. Müller, Multivariate and semiparametric kernel regression, SFB 373 Discussion Papers, Humboldt University of Berlin, 1997. https://ideas.repec.org/p/zbw/sfb373/199726.html.
[30] N. Langrené, X. Warin, Fast and stable multivariate kernel density estimation by fast sum updating, J. Comput. Graph. Statist. 28 (2019), 596-608.
[31] C. Duller, Einführung in die nichtparametrische Statistik mit SAS, R und SPSS, Springer, 2018.
[32] W. Härdle, A. Werwart, M. Müller, S. Sperlich, Nonparametric and Semiparametric Models, Springer (2004).
[33] D. Wied, R. Weißbach, Consistency of the kernel density estimator - A survey, Statist. Papers 52 (2012), 1-21.
[34] J. Alvarez-Mena, O. Hernández-Lerma, Convergence and approximation of optimization problems, SIAM J. Optim. 15 (2005), 527-539.