HOKKAIDO UNIVERSITY

Title	Output probability density function estimation based on the weighted kernel density estimation method under variable distribution parameters
Author(s)	Guo, Yifan; Lu, Zhenzhou; Chen,Yizhou
Relation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026), June 28 - July 1, 2026, Sapporo, Japan
Citation	Proceedings of the 10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026)
Issue Date	2026-07-01
Doc URL	https://hdl.handle.net/2115/100479
Type	conference paper
File Information	C000272_1.pdf

Hokkaido University Collection of Scholarly and Academic Papers : HUSCAP

10th International Symposium on Reliability Engineering and Risk Management (ISRERM 2026),
28 June-01 July 2026, Sapporo, Japan
Kitahara, T., Miyamori, Y. & Beer, M. (editors)

Output probability density function estimation based on the weighted kernel density estimation method under variable distribution parameters

Yifan Guo[1,2,3,4], Zhenzhou Lu[2,3,5,*], and Yizhou Chen[2,3,6]

[1]*School of Mechanics and Transportation Engineering, Northwestern Polytechnical University, Xi'an, China.*
[2]*State Key Laboratory of Clean and Efficient Turbomachinery Power Equipment, School of Aeronautics, Northwestern Polytechnical University, Xi'an, China.*
[3]*National Key Laboratory of Aircraft Configuration Design, Northwestern Polytechnical University, Xi'an, China.*
[4]*yifanguo_nwpu@163.com*
[5]*zhenzhoulu@nwpu.edu.cn*
[6]*yizhouchen@mail.nwpu.edu.cn*
**Corresponding Author*

The probability density function (PDF) of structural output plays an important role in statistical characteristic analysis of structural performance. Studies have proposed effective algorithms for solving the output PDF under fixed distribution parameters (DP) of random input variables. However, when obtaining the output PDF for arbitrary DP realizations, these methods require repeated calculations. To avoid the enormous computational cost required by numerical simulations for obtaining the structural output PDF under arbitrary DP realizations, this paper proposes a weighted kernel density estimation (KDE) method based on a single set of samples in augmented space. First, the weights of augmented samples required for estimating the output PDF under given DP realization are derived based on the principle of probability conservation. Second, to address the tail estimation errors of KDE, an adaptive truncation strategy is further introduced. By setting the truncation point of the output PDF and obtaining corresponding important samples, the proposed weighted KDE can be applied to obtain the truncated output PDF, which is then assembled to the complete output PDF using the conditional probability formula. The adaptive strategy iteratively determines a series of truncation points that progressively leads the tail of the PDF, enabling more accurate estimation of the PDF tail. Examples demonstrate that the proposed method, requiring single simulation run, significantly outperforms traditional KDE methods in efficiency while ensuring accuracy.

Keywords: Output probability density function, kernel density estimation, importance sampling.

1 Introduction

In engineering practice, uncertainty in structural performance (model output) exists due to the propagation of structure parameter (model input) uncertainty, such as the randomness of structural dimension and external load. The probability density function (PDF) serves as an important tool for describing random uncertainty and analyzing the statistical characteristics of random variables. Estimating the PDF of structural output has become a core component of uncertainty quantification and risk assessment. Meanwhile, real engineering systems often exhibit complex non-normal characteristics, necessitating the use of flexible and computationally feasible PDF estimation methods to capture the distribution of structural output from large-scale sample data. However, as the complexity of modern engineering increases, the cost of obtaining structural output samples also rises, making the estimation of the PDF of structural output a fundamental challenge in scientific and engineering fields (Zhang 2026).

The commonly used PDF estimation methods can be roughly divided into two categories: parametric methods and non-parametric ones (Fryer 1977, Kham 2025). Parametric method assumes a predefined distribution form and estimates the required distribution parameter vector (DP) using samples. Gaussian mixture model (GMM) based methods are representative methods (Zhang 2025). Increasing the complexity of the preset model often helps capture complex distribution features of the data but simultaneously increases the number of parameters to be estimated, raises computational demands, and affects algorithm convergence. Non-parametric methods are data-driven and do not require a predefined distribution model, offering greater flexibility. Kernel density estimation (KDE) (Zhang 2023) is a theoretically rigorous non-parametric method capable of approximating complex distributions and has gained widespread recognition. The core idea of KDE is to construct a kernel function at each data point and then sum the contributions of all kernel functions.

In fact, the DP of random input vector is not always fixed. In scenarios such as reliability-based design optimization (RBDO) and reliability updating (Guo 2025), DP also appears as vector variable. Therefore, it is necessary to analyze the structural output PDF under arbitrary realizations of DP. Obviously, estimating the output PDF through repeated sampling under different realizations of DP is a computationally expensive double-loop algorithm. This paper aims to overcome the computational inefficiency of double-loop algorithms and propose an efficient single-loop algorithm that does not require repeated sampling. In order to achieve single-loop simulation, this paper introduces a

weighted kernel density estimation algorithm (WKDE). The contribution of samples under arbitrary realizations of DP are derived which makes it possible to flexibly capture the characteristics of complex distributions using KDE method at arbitrary realizations of DP by re-weighting only one set of samples. Due to the high computational cost of complex performance function and the need for a large sample pool to accurately grasp the tail features of PDF, this paper designs a layered strategy on the basis of WKDE to further reduce the sample size. The basic idea of the layered strategy is to estimate the truncated PDF by obtaining a large number of samples at the tail of the distribution, thereby improving the estimation quality of the corresponding local regions of the PDF. The complete PDF can be efficiently obtained by repeatedly truncating smaller regions in the tail region of output distribution and extracting extra sample pool correspondingly. This paper uses a layered-strategy-based GMM importance sampling (IS) method to adaptively determine the tail truncation position, and combines the acceptance-rejection sampling (ARS) method to reorganize the samples of GMM corresponding to each layer to obtain distribution tail samples. In addition, this paper designs a sample mirror-data strategy to improve the estimation accuracy of WKDE near the truncation position of tail PDF where the step that starts from 0 occurs.

The remaining sections of this paper are arranged as follows. Section 2 proposes the single-loop framework of estimating output PDF under variable DP based on WKDE. Section 3 introduces the modified version of WKDE using layered IS strategy. Section 4 verifies the accuracy and efficiency of the proposed method with numerical examples. Section 5 presents the conclusion of this paper.

2 The Principle of the Weighted Kernel Density Estimation Method Under Variable Distribution Parameters

Denote $g(\boldsymbol{X})$ as the structure performance function involving an n-dimensional random input vector $\boldsymbol{X}=\{X_1,X_1,\cdots,X_n\}$ with PDF $f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta})$, and the DP $\boldsymbol{\theta}$ of $\boldsymbol{X}$ belongs to region $[\boldsymbol{\theta}^L,\boldsymbol{\theta}^U]$. The uncertainty of random input vector $\boldsymbol{X}$ propagates to the output $Y=g(\boldsymbol{X})$ through the deterministic law $g(\boldsymbol{X})$, thus the output Y is also a random variable with PDF $f_Y(y\,|\,\boldsymbol{\theta})$ related to the realization of DP $\boldsymbol{\theta}$. Once an $N_{\boldsymbol{X}}$-size sample pool $\boldsymbol{S}_{\boldsymbol{X}|\boldsymbol{\theta}^*}=\{\boldsymbol{x}_i^{\boldsymbol{\theta}^*}\}_{i=1}^{N_{\boldsymbol{X}}}$ of $\boldsymbol{X}$ under the realization $\boldsymbol{\theta}^*$ of DP $\boldsymbol{\theta}$ is generated from $f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta})$, the $N_{\boldsymbol{X}}$-size sample pool $\boldsymbol{S}_{Y|\boldsymbol{\theta}^*}=\{y_i^{\boldsymbol{\theta}^*}=g(\boldsymbol{x}_i^{\boldsymbol{\theta}^*})\}_{i=1}^{N_{\boldsymbol{X}}}$ of Y under the realization $\boldsymbol{\theta}^*$ of DP $\boldsymbol{\theta}$ can be obtained by calling the structure performance function $y=g(\boldsymbol{x})$. KDE can be used to estimate $f_Y(y\,|\,\boldsymbol{\theta}^*)$ by the kernel density function $k_{\boldsymbol{\theta}^*}(y)$ in Eq. (1).

$$f_Y(y\,|\,\boldsymbol{\theta}^*)\approx k_{\boldsymbol{\theta}^*}(y)=\frac{1}{N_{\boldsymbol{X}}}\sum_{i=1}^{N_{\boldsymbol{X}}}\frac{1}{\omega}K_{\boldsymbol{\theta}^*}\left(\frac{y-y_i^{\boldsymbol{\theta}^*}}{\omega}\right) \tag{1}$$

where ω is the bandwidth and $K_{\boldsymbol{\theta}^*}(t)$ is the normal kernel probability density function shown in Eq. (2) with sample mean $\overline{y}_{\boldsymbol{\theta}^*}$ in Eq. (3) and deviation $S_{\boldsymbol{\theta}^*}$ in Eq. (4).

$$K_{\boldsymbol{\theta}^*}(t)=\frac{1}{\sqrt{2\pi S_{\boldsymbol{\theta}^*}}}\exp\left(-\frac{t^2}{2S_{\boldsymbol{\theta}^*}}\right) \tag{2}$$

$$\overline{y}_{\boldsymbol{\theta}^*}=\frac{1}{N_{\boldsymbol{X}}}\sum_{i=1}^{N_{\boldsymbol{X}}}y_i^{\boldsymbol{\theta}^*} \tag{3}$$

$$S_{\boldsymbol{\theta}^*}=\sum_{i=1}^{N_{\boldsymbol{X}}}(y_i^{\boldsymbol{\theta}^*}-\overline{y}_{\boldsymbol{\theta}^*})^2 \tag{4}$$

The following part introduces the principle of estimating $f_Y(y\,|\,\boldsymbol{\theta})$ under arbitrary realizations of DP in region $[\boldsymbol{\theta}^L,\boldsymbol{\theta}^U]$ with a single sample pool and WKDE. According to law of large numbers, $k_{\boldsymbol{\theta}^*}(y)$ can be seen as the estimation of an integral as shown in Eq. (5) since $\boldsymbol{S}_{Y|\boldsymbol{\theta}^*}$ is the sample pool of $f_Y(y\,|\,\boldsymbol{\theta}^*)$.

$$k_{\boldsymbol{\theta}^*}(y)=\frac{1}{N_{\boldsymbol{X}}}\sum_{i=1}^{N_{\boldsymbol{X}}}\frac{1}{\omega}K_{\boldsymbol{\theta}^*}\left(\frac{y-y_i^{\boldsymbol{\theta}^*}}{\omega}\right)\approx\frac{1}{\omega}\int_t K_{\boldsymbol{\theta}^*}\left(\frac{y-t}{\omega}\right)f_Y(t\,|\,\boldsymbol{\theta}^*)dt \tag{5}$$

Suppose an $N_{\boldsymbol{X}}$-size sample pool $\boldsymbol{S}_{\boldsymbol{X}}=\{\boldsymbol{x}_i\}_{i=1}^{N_{\boldsymbol{X}}}$ of $\boldsymbol{X}$ is generated from an assigned PDF $f_{\boldsymbol{X}}(\boldsymbol{x})$ that independent of $\boldsymbol{\theta}$, and corresponding $N_{\boldsymbol{X}}$-size sample pool $\boldsymbol{S}_Y=\{y_i=g(\boldsymbol{x}_i)\}_{i=1}^{N_{\boldsymbol{X}}}$ of Y can be obtained. Eq. (6) holds,

$$\int_{y\in\Omega}H(y)f_Y(y\,|\,\boldsymbol{\theta}^*)dy=\int_{g(\boldsymbol{x})\in\Omega}H(g(\boldsymbol{x}))f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)d\boldsymbol{x}=\int_{g(\boldsymbol{x})\in\Omega}H(g(\boldsymbol{x}))\frac{f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)}{f_{\boldsymbol{X}}(\boldsymbol{x})}f_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x} \tag{6}$$

where $H(\cdot)$ is an arbitrary function, whose domain is the same as the range of y. Therefore, Eq. (1) can be approximated by Eq. (7) with sample pools $\boldsymbol{S}_{\boldsymbol{X}}$, $\boldsymbol{S}_Y$ and weight function $W(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)=f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)\big/f_{\boldsymbol{X}}(\boldsymbol{x})$.

$$f_Y(y\,|\,\boldsymbol{\theta}^*) \approx k_{\boldsymbol{\theta}^*}(y) = \frac{1}{N_X}\sum_{i=1}^{N_X}\frac{1}{\omega}K_{\boldsymbol{\theta}^*}\left(\frac{y-y_i}{\omega}\right)W(\boldsymbol{x}_i\,|\,\boldsymbol{\theta}^*) \tag{7}$$

Similarly, Eqs. (3) and (4) can be approximated by Eqs. (8) and (9) with sample pools $\boldsymbol{S}_X$ and $\boldsymbol{S}_Y$.

$$\bar{y}_{\boldsymbol{\theta}^*} = \frac{1}{N_X}\sum_{i=1}^{N_X} y_i W(\boldsymbol{x}_i\,|\,\boldsymbol{\theta}^*) \tag{8}$$

$$S_{\boldsymbol{\theta}^*} = \sum_{i=1}^{N_X}(y_i-\bar{y}_{\boldsymbol{\theta}^*})^2 W(\boldsymbol{x}_i\,|\,\boldsymbol{\theta}^*) \tag{9}$$

By calculating the $W(\boldsymbol{x}_i\,|\,\boldsymbol{\theta}^*)$ $(i=1,2,\cdots,N_X)$ and substituting these weights into Eqs. (7), (8) and (9), $f_Y(y\,|\,\boldsymbol{\theta})$ under arbitrary realizations of $\boldsymbol{\theta}$ can be estimated by the WKDE within a single simulation run. The heigh efficiency layered strategy of determining the assigned PDF will be elaborated in section 3.

3 The Combination of the Weighted Kernel Density Estimation and Layered Importance Sampling

The expected PDF $\tilde{f}_X(\boldsymbol{x})$ in Eq. (10) contains the information of $\boldsymbol{\theta}$ in $[\boldsymbol{\theta}^L,\boldsymbol{\theta}^U]$,

$$\tilde{f}_X(\boldsymbol{x}) = \int_{\boldsymbol{\theta}} f_X(\boldsymbol{x}\,|\,\boldsymbol{\theta}) f_{\boldsymbol{\theta}}(\boldsymbol{\theta})d\boldsymbol{\theta} \tag{10}$$

where $\boldsymbol{\theta}$ is treated as a random vector on $[\boldsymbol{\theta}^L,\boldsymbol{\theta}^U]$ with uniform distribution PDF $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$. Suppose $\boldsymbol{S}_X^{\tilde{f}} = \{\boldsymbol{x}_i^{\tilde{f}}\}_{i=1}^{N_X}$ is the N_X-size sample pool of $\tilde{f}_X(\boldsymbol{x})$, WKDE can be used combining $\boldsymbol{S}_X^{\tilde{f}}$ and $\boldsymbol{S}_Y^{\tilde{f}} = \{y_i^{\tilde{f}} = g(\boldsymbol{x}_i^{\tilde{f}})\}_{i=1}^{N_X}$ to obtain the complete output PDF $f_Y(y\,|\,\boldsymbol{\theta}^*)$ $(\boldsymbol{\theta}\in[\boldsymbol{\theta}^L,\boldsymbol{\theta}^U])$. Match $\boldsymbol{S}_X^{\tilde{f}}$ and $\boldsymbol{S}_Y^{\tilde{f}}$ in sample pair pool $\boldsymbol{S}_{X,Y}^{\tilde{f}} = \{\boldsymbol{x}_i^{\tilde{f}}, y_i^{\tilde{f}}\}_{i=1}^{N_X}$

The KDE-based methods require a large number of tail samples to accurately capture PDF tail features. Directly expanding the size of $\boldsymbol{S}_{X,Y}^{\tilde{f}}$ can increase the tail samples of the distribution, but at the same time, it will increase a large number of samples in the middle of the distribution, resulting in a great increase in computational cost. Therefore, this section proposes a layered IS strategy to increase the sample size locally at a series of adaptively determined tail of the output distribution, to efficiently improving the estimation quality of the output PDF tail by WKDE.

The proposed layered IS strategy in this section is modified from that proposed by Kurtz and Song (2013). The case that the tail truncation position moves towards smaller y is elaborated as the following part. Suppose $b_1^- > b_2^- > \cdots$ is series of decreasing thresholds, and the corresponding IS PDF is defined in Eq. (11).

$$h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x}) = \frac{I_{y\le b_j^-}(\boldsymbol{x})\tilde{f}_X(\boldsymbol{x})}{\int_{\boldsymbol{x}} I_{y\le b_j^-}(\boldsymbol{x})\tilde{f}_X(\boldsymbol{x})d\boldsymbol{x}},\quad \text{where } I_{y\le b_j^-}(\boldsymbol{x}) = \begin{cases}1 & g(\boldsymbol{x})\le b_j^-\\ 0 & g(\boldsymbol{x})> b_j^-\end{cases} \tag{11}$$

where $P_f^-(b_j^-) = \int_{\boldsymbol{x}} I_{y\le b_j^-}(\boldsymbol{x})\tilde{f}_X(\boldsymbol{x})d\boldsymbol{x}$ is the normalization coefficient.

Referring to the method proposed by Kurtz and Song (2013), thresholds $b_1^- > b_2^- > \cdots$ can be adaptively determined and GMMs $G_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ $(j=1,2,\cdots)$ that approximate corresponding IS PDF $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ $(j=1,2,\cdots)$ can be obtained. Set $G_{\boldsymbol{X}|b_0^-}(\boldsymbol{x}) = \tilde{f}_X(\boldsymbol{x})$. The input sample pools $\boldsymbol{S}_{\boldsymbol{X}|b_j^-} = \{\boldsymbol{x}_i^{-(j)}\}_{i=1}^{N_X}$ and the output samples $\boldsymbol{S}_{Y|b_j^-} = \{y_i^{-(j)} = g(\boldsymbol{x}_i^{-(j)})\}_{i=1}^{N_X}$ corresponding $G_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ $(j=0,1,2,\cdots)$ are also obtained. The next step is obtaining the sample pools of $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ $(j=1,2,\cdots)$ by re-organizing $\boldsymbol{S}_{\boldsymbol{X}|b_j^-}$ and $\boldsymbol{S}_{Y|b_j^-}$ $(j=1,2,\cdots)$ with ARS. Introduce a standard uniformly distributed random variable U and generate samples $\boldsymbol{S}_U = \{u_i\}_{i=1}^{N_X}$ to form $\boldsymbol{S}_{U,\boldsymbol{X},Y|b_j^-} = \{u_i, \boldsymbol{x}_i^{-(j)}, y_i^{-(j)}\}_{i=1}^{N_X}$. According to ARS (Ridley and Forget 2021), $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^R = \{\boldsymbol{x}_i^{-(j)}, y_i^{-(j)}\}_{i=1}^{\cdots}$ can be obtained where $\boldsymbol{x}_i^{-(j)}$ $(i=1,2,\cdots)$ are the samples of $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$. Repeatedly generate $\boldsymbol{S}_U = \{u_i\}_{i=1}^{N_X}$, than form $\boldsymbol{S}_{U,\boldsymbol{X},Y|b_j^-}$ again and scan samples of $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ from $\boldsymbol{S}_{U,\boldsymbol{X},Y|b_j^-}$ to join into $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^R$ according to ARS until the size of $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^R$ becomes N_X. The tail truncation position b_j^- $(j=1,2,\cdots)$, IS sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^R$ $(j=0,1,2,\cdots)$ are ready and $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ of $\boldsymbol{x}_i^{-(j)} \in \boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^R$ $(j=0,1,2,\cdots)$ can be simulated by $\boldsymbol{S}_{\boldsymbol{X}|b_j^-}$ and $\boldsymbol{S}_{Y|b_j^-}$ of $G_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ $(j=0,1,2,\cdots)$ and an extra sample pool of $f_{\boldsymbol{\theta}}(\boldsymbol{\theta})$.

The original output PDF $f_Y(y\,|\,\boldsymbol{\theta}^*)$ under the DP realization of $\boldsymbol{\theta}^*$ can be derived as shown in Eq. (12),

$$f_Y(y\,|\,\boldsymbol{\theta}^*)=\frac{d}{dy}\Pr\{Y\le y\,|\,\boldsymbol{\theta}^*\}=\frac{d}{dy}\Big[\Pr\{Y\le y\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*,Y\le b_j^-\}\Pr\{Y\le b_j^-\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*\}\Big]$$
$$=\Pr\{Y\le b_j^-\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*\}\frac{d}{dy}\Big[\Pr\{Y\le y\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*,Y\le b_j^-\}\Big]=\Pr\{Y\le b_j^-\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*\}\cdot f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*) \tag{12}$$

where $\Pr\{Y\le b_j^-\,|\,\boldsymbol{\theta}=\boldsymbol{\theta}^*\}$ is the conditional probability and $f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*)$ is the tail PDF truncated by $y\le b_j^-$. In case of $y\le b_j^-$, $f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*)$ is of a step starts from 0 at b_j^-, i.e., $f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*)=0$ for $y>b_j^-$ and $f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*)\neq 0$ otherwise. Since WKDE behaves poorly near the step of PDF due to the limitations of direct KDE, WKDE is further modified. For $j>0$, mirror sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{RM}=\{\boldsymbol{x}_{i+N_X}^{-(j)}=\boldsymbol{x}_i^{-(j)},y_{i+N_X}^{-(j)}=2b_j^--y_i^{-(j)}\}_{i=1}^{N_X}\ (j=1,2,\cdots)$ are formed. The sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{\prime RM}=\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{R}\cup\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{RM}$ can be collected. Since, $2b_j^--y_i^{-(j)}$ and $y_i^{-(j)}$ are symmetric about b_j^-, the PDF $f_{Y|Y\le b_j^-}^{\prime RM}(y\,|\,\boldsymbol{\theta}^*)$ of $y_i^{-(j)}\in\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{\prime RM}$ is symmetric about b_j^- and $f_{Y|Y\le b_j^-}(y\,|\,\boldsymbol{\theta}^*)=2f_{Y|Y\le b_j^-}^{\prime RM}(y\,|\,\boldsymbol{\theta}^*)$ for $y\le b_j^-$. The weight function is $W(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)=P_f^-(b_j^-)\cdot f_{\boldsymbol{X}}(\boldsymbol{x}\,|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$.

The case that the tail truncation position moves towards bigger y is similarly. Increasing thresholds $b_1^+<b_2^+<\cdots$ are also adaptively determined and assigned as the tail truncation position. The IS PDFs corresponding to b_j^+ $(j=1,2,\cdots)$ are defined in Eq. (13),

$$h_{\boldsymbol{X}|b_j^+}(\boldsymbol{x})=\frac{I_{y\ge b_j^+}(\boldsymbol{x})\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})}{\int_{\boldsymbol{x}}I_{y\ge b_j^+}(\boldsymbol{x})\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}},\ \text{where } I_{y\ge b_j^+}(\boldsymbol{x})=\begin{cases}1 & g(\boldsymbol{x})\ge b_j^-\\ 0 & g(\boldsymbol{x})<b_j^-\end{cases} \tag{13}$$

where $P_f^+(b_j^+)=\int_{\boldsymbol{x}}I_{y\ge b_j^+}(\boldsymbol{x})\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})d\boldsymbol{x}$ is the normalization coefficient.

Set $G_{\boldsymbol{X}|b_0^+}(\boldsymbol{x})=G_{\boldsymbol{X}|b_0^-}(\boldsymbol{x})$. IS sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^+}^{R}$, $h_{\boldsymbol{X}|b_j^-}(\boldsymbol{x})$ of $\boldsymbol{x}_i^{-(j)}\in\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{R}$ mirror sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^+}^{RM}=\{\boldsymbol{x}_{i+N_X}^{+(j)}=\boldsymbol{x}_i^{+(j)},y_{i+N_X}^{+(j)}=2b_j^+-y_i^{+(j)}\}_{i=1}^{N_X}\ (j=1,2,\cdots)$ and the total sample pools $\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{\prime RM}=\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{R}\cup\boldsymbol{S}_{\boldsymbol{X},Y|b_j^-}^{RM}$ can be obtained refers to the aforementioned layered IS strategy.

With all these sample pools, the complete PDF can be estimated by applying WKDE in segments combining **Table 1** and Eqs. (7), (8) and (9). The order of determining tail truncation position can varies. In section 4, b_j^+ $(j=1,2,\cdots)$ are determined after the needed b_j^- $(j=1,2,\cdots)$ have been determined.

Table 1. Required information for estimating complete output probability density function by the proposed methods.

Method	Value of y	$W(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)$	Sample set	Estimation of $f_Y(y\,\|\,\boldsymbol{\theta}^*)$
WKDE	Arbitrary	$f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y}^{\tilde{f}}$	$k_{\boldsymbol{\theta}^*}(y)$
WKDEL*	$b_j^-\ge y>b_{j+1}^-\ (j>0)$	$P_f^-(b_j^-)\cdot f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^-}^{R}$	$k_{\boldsymbol{\theta}^*}(y)$
	$b_1^+\ge y>b_1^-$	$f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^-}^{R}$	
	$b_{j+1}^+>y\ge b_j^+\ (j>0)$	$P_f^+(b_j^+)\cdot f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^+}^{R}$	
WKDELM**	$b_j^-\ge y>b_{j+1}^-\ (j>0)$	$P_f^-(b_j^-)\cdot f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^-}^{\prime RM}$	$2k_{\boldsymbol{\theta}^*}(y)$
	$b_1^+\ge y>b_1^-$	$f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^-}^{R}$	$k_{\boldsymbol{\theta}^*}(y)$
	$b_{j+1}^+>y\ge b_j^+\ (j>0)$	$P_f^+(b_j^+)\cdot f_{\boldsymbol{X}}(\boldsymbol{x}\,\|\,\boldsymbol{\theta}^*)\big/\tilde{f}_{\boldsymbol{X}}(\boldsymbol{x})$	$\boldsymbol{S}_{\boldsymbol{X},Y\|b_j^+}^{\prime RM}$	$2k_{\boldsymbol{\theta}^*}(y)$

*: WKDE combined with layered IS strategy (L)
**: WKDE combined with layered IS strategy (L) and enhanced by mirror-data strategy (M)

4 Examples

In this section, two examples are presented. Double-loop KDE (DKDE), the proposed WKDE, WKDEL and WKDELM are compared.

4.1 *Linear example*

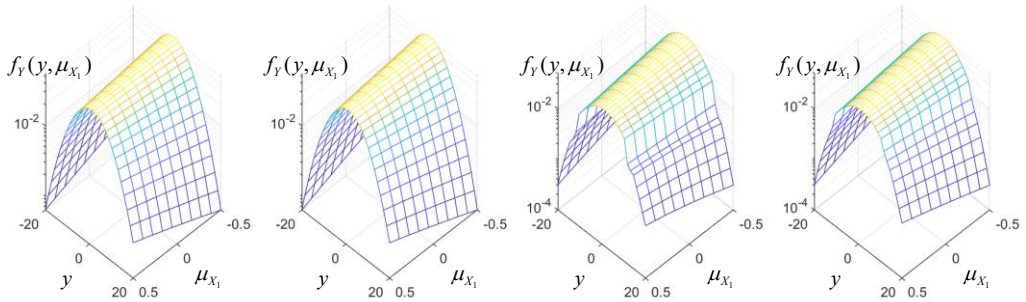
The performance function of example 1 is shown in Eq. (14), and it involves an 10-dimensional independently and normally distributed random input vector $(X_1,X_2,\cdots,X_n)$, i.e., $X_1\sim N(\mu_{X_1},1)$ and $X_i\sim N(0,1)$ $(i=2,3,\cdots,n)$. The random DP $\mu_{X_1}\in[-0.5,0.5]$.

$$g(\boldsymbol{x}) = \sum_{i=1}^{n}\left(3 - \frac{2(i-1)}{n-1}\right)x_i \tag{14}$$

The average performance of DKDE, WKDE, WKDEL and WKDELM are compared. Each method runs for 10 times. **Table 2**, **Figure 1** and **Figure 2** are the results of example 1. Denote total sample size as N, the mean of coefficient of variation as Cov_{mean}. Average of relative error (ARE) at each single discrete point represents the mean of relative error among 10 runs, and ARE_{mean} is the mean of ARE. It can be seen from the results the layered strategy save sample size compared with WKDE and the mirror-data strategy further improves the accuracy of output estimation.

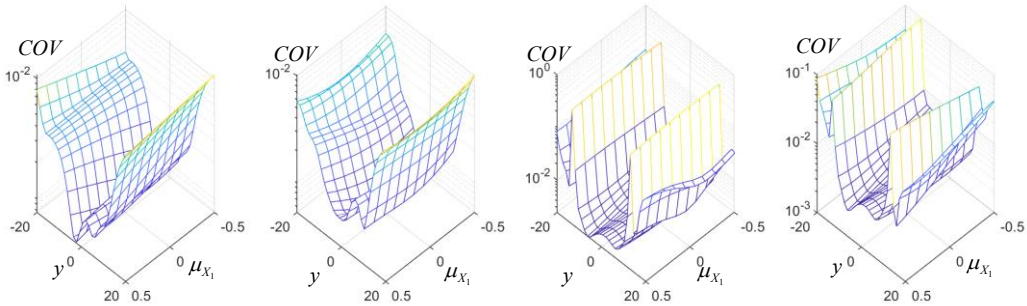
Table 2. The results of example 1.

Method	N	Cov_{mean}	ARE_{mean}
DKDE	1×10^{7}	3.25×10^{-3}	2.40×10^{-1}
WKDE	1×10^{6}	2.73×10^{-3}	2.41×10^{-1}
WKDEL	5×10^{5}	4.06×10^{-2}	1.30×10^{-1}
WKDELM	5×10^{5}	1.54×10^{-2}	7.58×10^{-2}



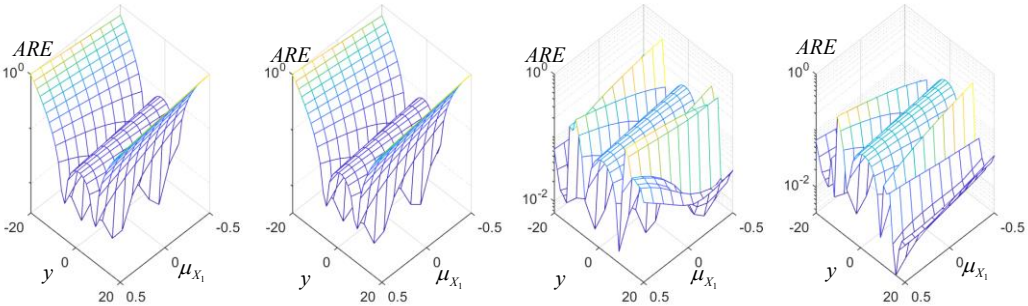
(*a*) DKDE (*b*) WKDE (*c*) WKDEL (*d*) WKDELM

Figure 1. This is the caption for the figure in 9 pt font. The caption must be centered.



(*a*) DKDE (*b*) WKDE (*c*) WKDEL (*d*) WKDELM

Figure 2. This is the caption for the figure in 9 pt font. The caption must be centered.



(*a*) DKDE (*b*) WKDE (*c*) WKDEL (*d*) WKDELM

Figure 3. This is the caption for the figure in 9 pt font. The caption must be centered.

4.2 *Nonlinear example*

The performance function of example 2 is shown in Eq. (15), and it involves an 3-dimensional independently and normally distributed random input vector (X_1, X_2, X_3), i.e., $X_1 \sim N(\mu_{X_1}, 1)$ and $X_i \sim N(0,1)$ $(i = 2,3)$. The random DP $\mu_{X_1} \in [-0.5, 0.5]$.

$$g(\boldsymbol{x}) = 1.15^{\sum_{i=1}^{n}\left(3-\frac{2(i-1)}{n-1}\right)x_i} \tag{15}$$

Table 3 shows the average performance of four methods. The proposed WKDELM achieve similar overall accuracy of WKDE while greatly reducing sample size and ensuring acceptable overall convergency.

Table 3. The results of example 2.

Method	N	Cov_{mean}	ARE_{mean}	$ARE_{\max}$
DKDE	1×10^{7}	9.94×10^{-3}	1.15×10^{-1}	2.62×10^{-1}
WKDE	1×10^{6}	9.05×10^{-3}	1.14×10^{-1}	2.41×10^{-1}
WKDEL	3×10^{5}	4.48×10^{-2}	1.30×10^{-1}	4.51×10^{-1}
WKDELM	3×10^{5}	4.10×10^{-2}	1.24×10^{-1}	2.87×10^{-1}

5 Conclusion

Distribution parameter vector (DP) of random input vector $\boldsymbol{X}$ can be a variable vector in some scenarios and it is necessary to analyze the structural output probability density function (PDF) under arbitrary realizations of DP. This paper proposes a single-loop algorithm for estimating output PDF under variable DP based on the weighted kernel density estimation (WKDE) combined with the layered importance sampling (IS) strategy to reduce the heavy computational costs in double-loop algorithm. This paper introduces a WKDE framework in which a single sample pool is re-weighted at arbitrary realizations of DP to obtain the complete output PDF. This paper designs a layered IS strategy on the basis of WKDE to further reduce the sample size. The complete PDF can be efficiently assembled by repeatedly truncating smaller regions in the tail region of output distribution and extracting extra sample pool correspondingly for local PDF estimation. In addition, the WKDE combined with layered IS strategy is enhanced by the sample mirror-data strategy. In the listed examples, the proposed method significantly outperforms traditional double-loop KDE method in efficiency while ensuring accuracy.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. NSFC 12272300, 12572141).

References

Fryer, M. J., A review of some non-parametric methods of density estimation. *IMA J Appl Math* 20(3): 335–54, 1977.

Guo, Y. F., Lu, Z. Z., and Chen, Y. Z., Importance sampling for augmented failure probability updating under output observations, 25th QR2MSE, IET Conference Proceedings, Hohhot, Inner Mongolia, China, 2025.

Kurtz, N., and Song J. H., Cross-entropy-based adaptive importance sampling using Gaussian mixture. *Struct Saf*, 42: 35-44, 2013.

Kham, S. L., Chamnongthai, K., and Aunsri, N., A review of efficient techniques and applications of kernel density in particle filtering framework, *Inf Fusion*, 103169, 2025.

Ridley, G., and Forget, B., A simple method for rejection sampling efficiency improvement on SIMT architectures, Statistics and Computing, 31(3): 1-11, 2021.

Zhang, R., and Dai, H., Stochastic analysis of structures under limited observations using kernel density estimation and arbitrary polynomial chaos expansion. *Comput Methods Appl Mech Eng*, 403: 115689, 2023.

Zhang, F., Xu, J., and Wan, Z., An adaptive Gaussian mixture model for structural reliability analysis using convolution search technique, *Struct Saf*, 112:102548, 2025.

Zhang, F. T., Xu, J., and Wang, D., Frequency-domain approach to automated and efficient multivariate kernel density estimation for probabilistic modeling, *Reliab Eng Syst Saf*, 273: 112351, 2026.