Statistical Device Modeling with Arbitrary Model-Parameter Distribution via Markov Chain Monte Carlo

Hiroki Tsukamoto, Song Bian, and Takashi Sato
Graduate School of Informatics, Kyoto University Yoshida-hon-machi, Sakyo, Kyoto 606-8501, Japan
tsuka@easter.kuee.kyoto-u.ac.jp

***Abstract*—We propose a novel statistical device modeling methodology that can represent model-parameters of arbitrary distribution and correlation. The proposed modeling is based on Markov chain Monte Carlo method in which random samples are drawn from the target probability distribution. The proposed method is also independent of the device models, allowing us to apply the method for any device models. Through the validation, the proposed method successfully reproduced the two peaks of the model parameter distribution that generated the current distribution. In addition, the experiments on the measured current variations following a non-Gaussian distribution demonstrate that the proposed method reduced the modeling error significantly as compared to the conventional method that can only use normal distribution.**

***Index Terms*—statistical device modeling, device process variation, power MOSFET**

I. Introduction

Statistical circuit simulation, which considers the variation of device characteristics, is critically important for robust circuit design. Various statistical device modeling methods have been proposed thus far. Most of them assume that the model parameters as well as the device characteristics follow a normal distribution [1]. However, in reality, it is frequently reported that the model parameters and device characteristics follow a non-Gaussian distribution [2]–[5]. Moreover, the devices that are fabricated in multiple fabs can collectively exhibit a multimodal parameter distribution. Nevertheless, existing modeling methods can only approximate them using Gaussian distributions.

In this paper, we propose a new statistical device modeling technique that assumes no distributions and correlations. The contributions of this work are summarized as follows:

- **Statistical parameter modeling for arbitrary distributions with no underlying assumption**: We redefine statistical circuit simulations by introducing the concept of Markov chain Monte Carlo method for generating statistical model parameters. The proposed method is defined as a general procedure so that the model parameters represent arbitrary distributions including their correlations.
- **Model independence:** The proposed method uses the device models as a *translator* of the variation of model parameters to the target performance variation, such as measured drain currents. Hence, the proposed method is applicable for any device models with no change.
- **Solid use model**: We present a concrete use model of the proposed statistical modeling. Users generate random model parameters based on the probability density function (PDF) of measurement data provided by a manufacturer.

While we focus on the modeling of power MOSFETs as an example in this work, our modeling framework is general and applicable for any device models.

Algorithm 1 Metropolis sampling

```
Set the initial value x^(1)
for i = 1 to N − 1 do
    x′ is generated from a proposal distribution q(x′|x^(i))
    r ← P(x′)/P(x^(i))
    Draw R from uniform distribution with range 0 ≤ R < 1
    if r > R then
        x^(i+1) ← x′ (accept)
    else
        x^(i+1) ← x^(i) (reject)
    end if
end for
```

II. Statistical device modeling for arbitrary distribution

A. Markov chain Monte Carlo

We first review the Markov chain Monte Carlo (MCMC) method applied to the proposed modeling. MCMC is a method for generating random samples that follow a given multivariate probability distribution [6]. As a representative example, we explain the Metropolis method [7] which is sketched in Algorithm 1. Here, $\boldsymbol{x}$ is the random sample and $P(\boldsymbol{x})$ is the PDF that $\boldsymbol{x}$ should follow. Through Algorithm 1, we obtain N random samples $\{\boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)}, \ldots, \boldsymbol{x}^{(N)}\}$ that follow the PDF $P(\boldsymbol{x})$.

Though Algorithm 1 outlines the key idea of generating samples using MCMC, various modifications are proposed for improving the quality of the generated samples. In practice, between any two consecutive samples obtained according to Algorithm 1 suffers from correlation [6]. This correlation depends on the choice of the proposal distribution q. Choosing q that yields smaller distances between sample

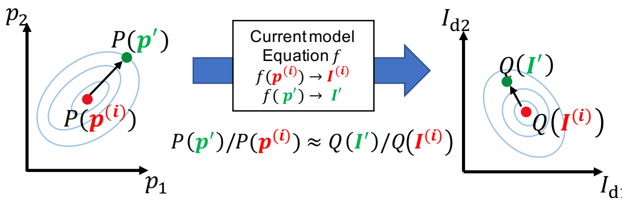


Fig. 1. Approximation of the probability density ratio.

candidates obviously leads to larger correlation. The samples hence should be subsampled to remove such correlations. For example, every one out of n_s samples can be used to improve independence. In addition, when the initial sample $\boldsymbol{x}^{(1)}$ is located away from a high probability region, which is usually the case when the PDF is unknown, the samples obtained at the beginning of this process may not correctly reflect the probability density of $P(\boldsymbol{x})$ [6]. Hence, the samples obtained until the process reaches the steady state are discarded as *burn-in*.

B. Key idea

In the proposed statistical device modeling, random model parameter sets $\{\boldsymbol{p}^{(1)}, \boldsymbol{p}^{(2)}, \ldots, \boldsymbol{p}^{(N)}\}$ that follow $P(\boldsymbol{p})$ are generated via MCMC to reproduce the distribution of measured device characteristics. Here, a model parameter set $\boldsymbol{p}$ consists of m model parameters, $\boldsymbol{p} = (p_1, p_2, \ldots, p_m)^T$, where p_j represents the j-th model parameter, such as threshold voltage **VTH**. In general, we do not know the PDF $P(\boldsymbol{p})$ that the model parameter sets follow. Instead, through measurements, we know the drain current distribution $Q(\boldsymbol{I})$. The essential idea of the proposed method is to indirectly approximate the model parameter distribution by randomly generating parameter sets through the process similar to MCMC so that simulated drain currents using the approximated device model parameters reproduce the measured current distribution.

Fig. 1 illustrates a key component of this idea in a two dimensional example. As we vary the model parameter set $\boldsymbol{p}$ in the parameter space p_1–p_2, the drain current $\boldsymbol{I}$ varies correspondingly in the current space I_{d1}–I_{d2} through the transformation of the current model equation f. Here, I_{d1} and I_{d2} represent the drain currents at different bias points. If the current density ratio $Q(\boldsymbol{I}')/Q(\boldsymbol{I}^{(i)})$ can be approximated by the ratio of the model parameters $P(\boldsymbol{p}')/P(\boldsymbol{p}^{(i)})$, we can sample model parameter sets that reproduce the measured current distribution via the process similar to MCMC.

C. The Proposed Statistical Modeling

The proposed statistical device modeling is shown in Algorithm 2. We first apply kernel density estimation (KDE) for the measured drain currents in an m-dimensional space to obtain PDF $Q(\boldsymbol{I})$. Starting from the initial model parameter set $\boldsymbol{p}^{(1)}$, lines 4 to 12 are repeated for $N-1$ times to generate a parameter set for each iteration. The next sample candidate $\boldsymbol{p}'$ is obtained by moving the current sample $\boldsymbol{p}^{(i)}$ in the model parameter space using a proposal distribution $q(\boldsymbol{p}'|\boldsymbol{p}^{(i)})$ in line 4. In line 5, $\boldsymbol{p}^{(i)}$ and $\boldsymbol{p}'$ are respectively converted to the drain currents $\boldsymbol{I}^{(i)}$ and $\boldsymbol{I}'$ for m bias points

Algorithm 2 Statistical Device Modeling via MCMC

Require: Measured current variation data of m bias points, $\boldsymbol{I}$
Ensure: N instances of model parameters $\boldsymbol{p}$ sampled from the unknown model parameters' PDF $P(\boldsymbol{p})$.

```
1:  Estimate the PDF Q(I) from the current variance data
2:  Set p^(1) as an initial instance of model parameter set
3:  for i = 1 to N − 1 do
4:      p' is generated from a proposal distribution q(p'|p^(i))
5:      I' ← f(p'), I^(i) ← f(p^(i))
6:      r ← Q(I')/Q(I^(i))
7:      Draw R from uniform distribution with range 0 ≤ R < 1
8:      if r > R then
9:          p^(i+1) ← p' (accept)
10:     else
11:         p^(i+1) ← p^(i) (reject)
12:     end if
13: end for
```

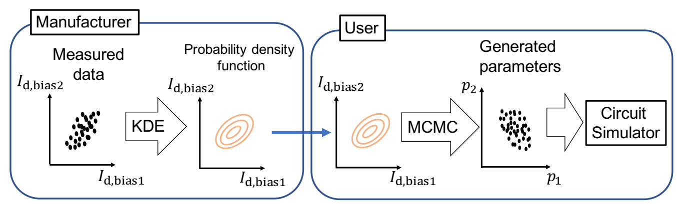


Fig. 2. Use case of the proposed statistical device modeling.

via a current model equation f. This conversion through the current model equation does not appear in the original MCMC. Through this conversion, current PDF Q of the corresponding model parameter set is evaluated, and compared with the previous one, i.e., the ratio $r = Q(\boldsymbol{I}')/Q(\boldsymbol{I}^{(i)})$ is calculated. Here, $Q(\boldsymbol{I}')/Q(\boldsymbol{I}^{(i)}) \approx P(\boldsymbol{p}')/P(\boldsymbol{p}^{(i)})$ holds as shown in Fig. 1. Then, we generate a uniform random number R, where $0 \le R < 1$. If r is greater than R, we accept the candidate $\boldsymbol{p}'$ as the next sample $\boldsymbol{p}^{(i+1)}$ (lines 8 and 9), otherwise $\boldsymbol{p}'$ is rejected and the current sample $\boldsymbol{p}^{(i)}$ becomes the next sample $\boldsymbol{p}^{(i+1)}$ (lines 10 and 11). The proposed algorithm yields random model parameter sets $\{\boldsymbol{p}^{(1)}, \boldsymbol{p}^{(2)}, \ldots, \boldsymbol{p}^{(N)}\}$ which reproduce the measured current distribution. In practice, the burn-in and the decorrelation by subsampling may be necessary but omitted in the algorithm description for the purpose of clarity. As we see in the above procedure, the proposed statistical device modeling method does not assume any underlying distributions for the model parameters nor correlations between them. This is a general method that is applicable not only for current variations but also for any other electrical characteristic variations, such as capacitance.

Fig. 2 shows an example use of the proposed method. A device manufacturer measures characteristic variation data, e.g., drain current, on a large number of devices. The manufacturer then calculates its PDF, $Q(\boldsymbol{I})$, through kernel density estimation (KDE). The manufacturer distributes $Q(\boldsymbol{I})$ to users. Using $Q(\boldsymbol{I})$, the users can generate model parameter sets by using their own models via the proposed method. Each generated model parameter set serves as a random model instance for Monte Carlo simulation.

TABLE I
MODEL PARAMETERS

Parameter	Description [Unit]
K	Current gain factor [A/V^2]
VTH	Threshold voltage [V]
CLM	Channel Length Modulation coefficient [V^{-1}]
MD	Mobility degradation coefficient [V^{-1}]
MDV	Channel voltage [V]
DELTA	Smoothing parameter [-]

III. EXPERIMENTS

A. Drain current model for power MOSFETs

On the basis of the simple threshold-voltage model, a drain current model for power MOSFETs is defined for use in the following experiments. This model equation serves as the function f in Fig. 1 and Algorithm 2. The model parameters of the model equation are summarized in Table I and are indicated in bold face. The drain current is expressed as:

$$I'_{\mathrm{d}} = \begin{cases} \mathbf{K}(V_{\mathrm{gs}} - \mathbf{VTH} - \frac{V_{\mathrm{ds,mod}}}{2}) \cdot V_{\mathrm{ds,mod}} & (V_{\mathrm{gs}} \geq \mathbf{VTH}) \\ 0 & (V_{\mathrm{gs}} < \mathbf{VTH}). \end{cases} \tag{1}$$

Here, the parameters such as channel length L, channel width W, carrier mobility μ, and oxide capacitance C_{ox} are collectively represented as $\mathbf{K}$ because each component in $\mathbf{K} = C_{\mathrm{ox}} \cdot \mu \cdot \frac{W}{L}$ is inseparable in the parameter fitting. In order to represent gradual transition between the linear and saturation regions of SiC power MOSFETs [8], $V_{\mathrm{ds,mod}}$ is introduced as in:

$$V_{\mathrm{ds,mod}} = \frac{V_{\mathrm{ds}}}{\left[1 + \left(\frac{V_{\mathrm{ds}}}{V_{\mathrm{gs}} - \mathbf{VTH}}\right)^{\mathbf{DELTA}}\right]^{1/\mathbf{DELTA}}}. \tag{2}$$

By considering the channel length modulation and mobility degradation, I_{d} becomes as follows:

$$I_{\mathrm{d}} = \begin{cases} \frac{1+\mathbf{CLM} \cdot V_{\mathrm{ds}}}{1+\mathbf{MD}(V_{\mathrm{gs}} - \mathbf{MDV})} \cdot I'_{\mathrm{d}} & (V_{\mathrm{gs}} \geq \mathbf{MDV}) \\ (1 + \mathbf{CLM} \cdot V_{\mathrm{ds}}) \cdot I'_{\mathrm{d}} & (V_{\mathrm{gs}} < \mathbf{MDV}). \end{cases} \tag{3}$$

B. Analysis on Simulation Data

In order to validate the proposed method, we first apply the proposed method to a problem where the underlying ground-truth is known.

1) Simulation Setup: We generate correlated model parameters following a multi-modal distribution shown in Fig. 3. Here, the variation of two model parameters, **K** (current gain factor) and **VTH** (threshold voltage), are considered for ease of illustration. In total, 5000 parameter sets are generated using two mean values, variances, and correlation coefficients for **K** and **VTH**. These values are derived by the measured current distributions of the SiC MOSFETs [9] manufactured in two different fabrication lines. Within the 5000 devices, 3500 and 1500 devices use different mean values, variances and correlation coefficients. Then, we generate current samples using the model parameter sets through the current model equations in Eqs. (1)-(3). The obtained current variation and the estimated PDF are shown in Fig. 4.

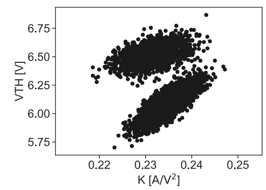


Fig. 3. Synthesized model parameter distribution based on the measured current distributions of the SiC MOSFETs manufactured in two different fabrication lines.

On the current distribution, we perform the proposed method to generate 5000 random model parameter sets of **K** and **VTH**. The initial parameter set $\boldsymbol{p}^{(1)} = (\mathbf{K}^{(1)}, \mathbf{VTH}^{(1)})^T$ for MCMC is randomly chosen from the intervals [0.1, 0.4) and [4, 8). The proposal distribution q during MCMC is an uncorrelated bivariate normal distribution with mean $\boldsymbol{p}^{(i)}$ and the variances, i.e., $(\mathbf{K}', \mathbf{VTH}')^T = (\mathbf{K}^{(i)} + \Delta_{\mathbf{K}}, \mathbf{VTH}^{(i)} + \Delta_{\mathbf{VTH}})^T$, where $\Delta_{\mathbf{K}} \sim N(0, 0.1^2)$, $\Delta_{\mathbf{VTH}} \sim N(0, 0.2^2)$. The total number of samples N is 101,000, in which the first 1,000 samples are discarded as burn-in. Then, every 20 samples of the remaining 100,000 samples are used as the generated sample, giving us a total number of 5,000 parameter sets.

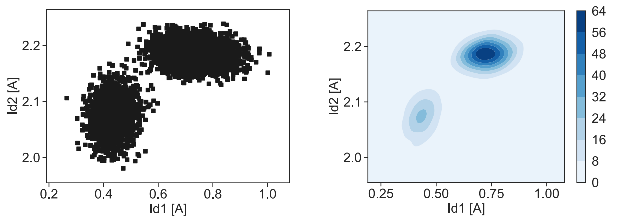
In comparison, we perform the conventional statistical device modeling [10]. Mean, variance, and correlation of the model parameters are derived, and 5,000 model parameter sets following the derived normal distribution are generated.

2) Results: The parameters generated by the proposed and the conventional methods are shown in Fig. 5. The solid black dots and histograms in (a) and (b) are identical and are the ground truth shown in Fig. 3. The diamonds in Fig. 5(a) are the generated model parameters using the existing method, while triangles in Fig. 5(b) are those using the proposed method. The corresponding histograms projected along **K** and **VTH** axes are also presented. The model parameters generated by the proposed method successfully reproduced the two clusters. There exist two high density locations, and the density is lower at about the center of this graph. On the other hand, in Fig. 5(a), those generated through the conventional method approximates the sample density with one peak. This difference is expected because the existing method assumes each model parameter to follow a normal distribution. Without assuming any distributions, the proposed method reproduced the distribution of model parameters more accurately than the existing method.

The accuracy of the proposed and the conventional methods is quantitatively evaluated based on KL divergence [11]. We calculate probability densities of the model parameter samples generated using both methods by KDE. We calculate KL divergence between ground truth's PDF $P(\boldsymbol{p})$ and generated samples' PDF $P^*(\boldsymbol{p})$ by the following equation:

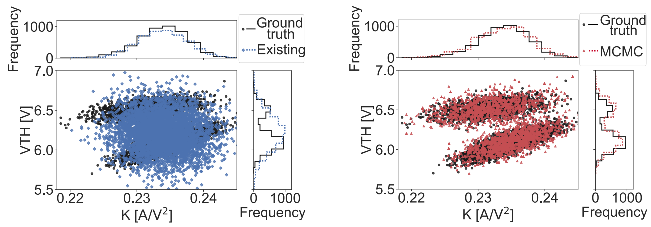
$$D_{\mathrm{KL}}(P||P^*) = \int P(\boldsymbol{p}) \log \frac{P(\boldsymbol{p})}{P^*(\boldsymbol{p})} \mathrm{d}\boldsymbol{p}. \tag{4}$$

The KL divergence of the proposed method is 0.065 and that



(a) Simulated current variation. (b) Estimated probability density.

Fig. 4. Generated current variation and its probability density. The bias conditions, I_{d1} and I_{d2} are chosen at (V_{gs}, V_{ds}) =(8 V, 20 V), (18 V, 1 V). A normal distribution is used as the kernel function for the KDE. The bandwidth in KDE is obtained by using the Scott's method [12].



(a) Conventional method (Gaussian). (b) Proposed method (MCMC).

Fig. 5. Comparison of the original and generated (extracted) model parameter variation.

of the existing method is 0.46. The proposed method reduced the error by 87% compared to the existing method, which may lead to critical yield estimation error.

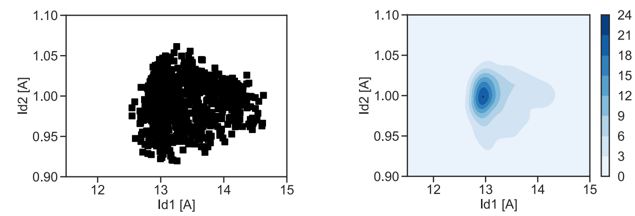
C. Analysis on Measured Data

We now apply the proposed method for modeling the measured current characteristic variations of 1020 SiC power MOSFETs [9].

We again demonstrate our method by using two parameters, **K** and **VTH** because of the simpplicity. The other parameters are treated as constant in this analysis. The bias conditions to define current distribution are (V_{gs}, V_{ds})=(20 V, 8 V) and (9 V, 9 V) for I_{d1} and I_{d2}, respectively, which is shown in Fig. 6(a). Then, we estimate the PDF of the current variation via KDE, using normal distribution as the kernel function. The estimated PDF of the measured current data is shown in Fig. 6(b).

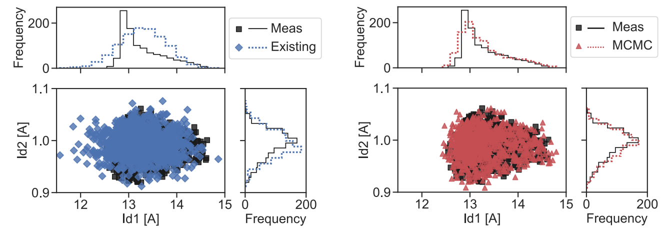
Then, we perform the proposed method to obtain random parameter sets. Here, the proposal distribution q and how to choose the initial parameter set $\boldsymbol{p}^{(1)}$ are the same as those in the previous simulation. The total number of samples N is 205,000, in which the first 1,000 samples are discarded as burn-in. Then, every 200 samples in the remaining 204,000 samples are used as the generated sample, giving us a total number of 1,020 sets of model parameters. In comparison, we perform the conventional statistical device modeling [10]. Mean, variance, and correlation of the model parameters are obtained, and then 1,020 model parameter sets following the normal distribution are generated.

In this experiment, since the true distribution of the model parameters is unknown, the accuracy is evaluated by comparing the current distributions reproduced by the generated model parameters. The obtained current distributions are shown in Fig. 7. The current characteristic variation generated by the proposed method overlaps closely to the measurement.



(a) Measured current variation. (b) Estimated probability density.

Fig. 6. Measured current variation and its probability density.



(a) Conventional method (Gaussian). (b) Proposed method (MCMC).

Fig. 7. Comparison of the measured and generated current variations.

In particular, the skew and the low probability region (outline) that determines circuit yield are very well represented. Meanwhile, the currents generated by the existing method lacks the skewness of the current distribution due to its normal assumption. The proposed method can simulate the statistical current variation more accurately as it does not assume any statistical distribution of the parameters.

The accuracy of the proposed and the conventional methods is quantitatively evaluated based on KL divergence. We calculate probability densities of the current variation generated using both methods by KDE. We calculate KL divergence between measured current's PDF $Q(\boldsymbol{I})$ and generated current's PDF $Q^*(\boldsymbol{I})$ by the following equation:

$$D_{\mathrm{KL}}(Q||Q^*) = \int Q(\boldsymbol{I}) \log \frac{Q(\boldsymbol{I})}{Q^*(\boldsymbol{I})} \mathrm{d}\boldsymbol{I}. \tag{5}$$

The KL distance of the proposed method is 0.036 and that of the existing method is 0.21. The proposed method reduced the error by 82% compared to the existing method, which may lead to critical yield estimation error.

IV. Conclusion

We proposed a novel and general statistical device modeling method in which no assumptions about the model parameter distribution are placed, and thus arbitrary distribution and correlations between the model parameters can be represented. In addition, the proposed modeling method is independent of a model equation. Through simulated and measured data, the accuracy of the proposed method is validated. The proposed method better approximated the distributions of model parameters and current variations than the existing method that relies on Gaussian assumption.

Ackowledgement

This work was supported by JST-OPERA Program Grant Number JPMJOP1841, Japan.

REFERENCES

[1] B. Cheng, D. Dideban, N. Moezi, C. Millar, G. Roy, X. Wang, S. Roy, and A. Asenov, "Statistical-variability compact-modeling strategies for BSIM4 and PSP," *IEEE Design & Test of Computers*, vol. 27, no. 2, pp. 26–35, 2010.

[2] U. Kovac, D. Dideban, B. Cheng, N. Moezi, G. Roy, and A. Asenov, "A novel approach to the statistical generation of non-normal distributed PSP compact model parameters using a nonlinear power method," in *Proc. International Conf. on Simulation of Semiconductor Processes and Devices*, 2010, pp. 125–128.

[3] A. Lange, C. Sohrmann, R. Jancke, J. Haase, B. Cheng, U. Kovac, and A. Asenov, "A general approach for multivariate statistical MOSFET compact modeling preserving correlations," in *Proc. European Solid-State Device Research Conference*, 2011, pp. 163–166.

[4] A. Lange, C. Sohrmann, R. Jancke, J. Haase, B. Cheng, A. Asenov, and U. Schlichtmann, "Multivariate modeling of variability supporting non-gaussian and correlated parameters," *IEEE Trans. Computer-Aided Design of Integrated Circuits and Systems*, vol. 35, no. 2, pp. 197–210, 2016.

[5] K. Revanth and V. Janakiraman, "Statistical compact model extraction for skew-normal distributions," *IET Circuits, Devices & Systems*, vol. 14, no. 5, pp. 576–585, 2020.

[6] W. Gilks, S. Richardson, and D. Spiegelhalter, *Markov Chain Monte Carlo in Practice*. CRC Press, 1995.

[7] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, "Equation of state calculations by fast computing machines," *J. Chemical Physics*, vol. 21, no. 6, pp. 1087–1092, 1953.

[8] G. Gildenblat, X. Li, W. Wu, H. Wang, A. Jha, R. van Langevelde, G. D. J. Smit, A. J. Scholten, and D. B. M. Klaassen, "PSP: An advanced surface-potential-based MOSFET model for circuit simulation," *IEEE Trans. Electron Devices*, vol. 53, no. 9, pp. 1979–1993, 2006.

[9] *SCT2450KE datasheet*, ROHM Co., Ltd.

[10] H. Jooypa and D. Dideban, "Statistical strategies to reproduce the propagation delay time variability using compact models," *IEEE Trans. Circuits and Systems II*, vol. 66, no. 11, pp. 1880–1884, 2019.

[11] S. Kullback and R. A. Leibler, "On information and sufficiency," *The Annals of Mathematical Statistics*, vol. 22, no. 1, pp. 79–86, 1951.

[12] D. W. Scott, *Multivariate Density Estimation: Theory, Practice, and Visualization*. John Wiley & Sons, 1992.