Shot-Noise-Induced Failure in Nanoscale Flip-Flops—Part I: Numerical Framework

Pooya Jannaty, Florian C. Sabou, Son T. Le, Marco Donato, R. Iris Bahar, *Member, IEEE*, William Patterson, *Member, IEEE*, Joseph Mundy, and Alexander Zaslavsky

***Abstract*—As CMOS technology continues the path of miniaturization, noise-induced fluctuations raise heightened reliability concerns. In previous work, an analytical framework based on Markov queueing theory and Poisson shot noise was presented to model the probabilistic behavior of a CMOS flip-flop operated in the subthreshold regime. In this paper, this model is extended to also account for the above-threshold shot noise, where the noise distribution is no longer Poissonian. The formulas for the time-dependent charging and discharging of node capacitors of a four-transistor flip-flop are derived for different regimes of operation characterized by distinct Fano factors. The statistics of electron arrival and departure at node capacitors is incorporated in an algebraic representation based on Markov queueing theory to map the effects of charge fluctuations on the logic stability of a flip-flop. This framework is used in Part II of this work to investigate failure in time for end-of-roadmap CMOS at the 10-nm gate-length scale.**

***Index Terms*—CMOS devices, Fano factor, Markov processes, nonequilibrium Green's function (NEGF), numerical analysis, reliability, shot noise, soft errors, SRAMs.**

I. Introduction

IN LIGHT OF advances in silicon technology over the past decades that have resulted both in continuous downscaling of CMOS transistors and a proliferation of different geometries and operating parameters for such transistors, there is a strong motivation for the development of reliability analysis that can account for the specificity of a given technology. As a direct result of the miniaturization process, present-day electronic circuits consist of billions of logic devices, increasing the complexity of estimating noise-induced errors. While the failure rates for a given logic device might be very small, the aggregate error rate for the overall electronic circuit can lead to reliability concerns.

The vast gap in scale between the individual logic device and the very large integrated systems of such devices rules out the use of Monte Carlo techniques to estimate the mean time to failure. While perfectly suited to parameterize a single logic unit with high error rates, Monte Carlo approaches become computationally prohibitive when the error rates per device are low [1].

In earlier papers, an analytical approach based on Markov queueing theory for predicting fault rates due to noise in low-power devices [2], [3] operated in the subthreshold regime is described [4]–[6], where the noise in transistors is expressed in terms of Poissonian sources representing the forward and reverse components of the current [7]. However, the vast majority of logic circuits are operated above threshold, where the noise in transistors can no longer be described using Poissonian statistics. To address this critical aspect, in this paper, the probabilistic framework for the analysis of variations in the logic stability of memory circuits due to shot noise and thermal noise is extended to a regime of operation characterized by driving voltages higher than the threshold voltage V_T.

Conceptually, the proposed framework consists of two parts: an analysis of shot noise in nanoscale MOSFETs using a nonequilibrium Green's function (NEGF) formalism to calculate the statistics of electron arrival and departure at the node capacitors in a flip-flop and a probabilistic formulation based on queue theory and Markov chains that associates the noise-induced charge fluctuations with discrete transitions between corresponding states in a 2-D queue. Once reduced to this algebraic abstraction, a noise-induced error, a bit flip from logic "0" to "1" or *vice versa*, corresponds to a spontaneous transition between the two most stable states in the rectangular queue. This framework is exploited in Part II of this work [8] to examine error rates in ultimate-CMOS flip-flops at the $L_G = 10$ nm gate-length scale.

This paper is organized as follows. Section II presents the NEGF formalism for the analysis of shot noise under different regimes of transistor operation characterized by distinct Fano factors, defined as the ratio of noise variance to average current, $F \equiv S/2e\langle I\rangle$. In Section III, the charge fluctuation statistics are mapped via Markov chains to a 2-D queue, describing all possible states of a flip-flop in terms of the number of elementary charges on its node capacitors. Finally, Section IV presents the conclusions.

II. Shot Noise in Nanoscale MOSFETs

The Landauer formula for the average current and its analog for the low-frequency spectral density of noise in a ballistic two-terminal device is expressed as [9], [10]

$$\langle I\rangle = \frac{2e}{h}\int dE\,\{(f_S(E) - f_D(E))\,\mathrm{Tr}\,[\mathbf{T}(E)]\} \tag{1}$$

Manuscript received July 2, 2011; revised September 28, 2011 and November 16, 2011; accepted November 21, 2011. Date of publication January 4, 2012; date of current version February 23, 2012. This work was supported by the Defense Threat Reduction Agency under Basic Research Award HDTRA1-10-1-0013. The review of this paper was arranged by Editor J. S. Suehle.

The authors are with the School of Engineering and the Department of Physics, Brown University, Providence, RI 02912 USA (e-mail: Pooya@Brown.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TED.2011.2177983

$$S = \frac{4e^2}{h}\int dE\,\{[f_S(1-f_S)+f_D(1-f_D)]\,\mathrm{Tr}[\mathbf{T}] + (f_S-f_D)^2\mathrm{Tr}\,[\mathbf{T}(\mathbf{I}-\mathbf{T})]\} \quad (2)$$

where $\mathrm{Tr}[\mathbf{T}(E)]$ is the trace of the transmission function that is equal to the sum of all eigenchannel transmission values at energy E, and f_S and f_D are the Fermi–Dirac functions at source and drain terminals. The first two terms in (2) represent the thermal noise at the two semi-infinite contacts in equilibrium, and the third term represents the so-called partition noise, which arises due to the probabilistic nature of the injected carrier passing through the channel or not. The transmission function $\mathbf{T}(E)$ can be expressed based on the retarded and advanced Green's functions, i.e.,

$$\mathbf{T}(E) = \mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_D\mathbf{G}^a. \quad (3)$$

Upon inclusion of noncoherent scattering effects, (1) is generalized to [11]

$$\langle I\rangle = \frac{2e}{h}\int dE\mathrm{Tr}\left[\mathbf{\Sigma}_S^{<}\mathbf{G}^{>}-\mathbf{\Sigma}_S^{>}\mathbf{G}^{<}\right] \quad (4)$$

and (2) is generalized to [12], [13]

$$S = \frac{4e^2}{h}\int dE\mathrm{Tr}\left\{i\mathbf{\Gamma}_S\left[f_S\mathbf{G}^{>}-(1-f_S)\mathbf{G}^{<}\right]\times[\mathbf{I}-i\mathbf{\Gamma}_S(\mathbf{G}^r-\mathbf{G}^a)] + \mathbf{\Gamma}_S\mathbf{G}^{>}\mathbf{\Gamma}_S\mathbf{G}^{<}-f_S(1-f_S)\times(\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_S\mathbf{G}^r+\mathbf{\Gamma}_S\mathbf{G}^a\mathbf{\Gamma}_S\mathbf{G}^a)\}\right. \quad (5)$$

where $\mathbf{\Sigma}_S^{<}$ and $\mathbf{\Sigma}_S^{>}$ are the self energies for the contact S, and $\mathbf{G}^{<}$ and $\mathbf{G}^{>}$ have their usual definitions, i.e.,

$$\mathbf{G}^{<(>)} = \mathbf{G}^r\mathbf{\Sigma}^{<(>)}\mathbf{G}^a. \quad (6)$$

These results are valid for frequencies low enough so that transit time and scattering relaxation times can be ignored.

Even today's small L_G devices are believed to be operating in the partially ballistic regime [14]–[18], and the predicted end-of-roadmap $L_G = 10$ nm MOSFETs are expected to approach more closely ballistic behavior. Still, in order to examine the effects of scattering, we have chosen longitudinal acoustic (LA) phonon scattering as a representative scattering mechanism in the modeled devices. Although possible within the NEGF formalism, no attempt has been made to include a more complete set of scattering mechanisms to achieve more accurate results. Within the self-consistent Born approximation the self energies for the phonon scattering obey

$$\mathbf{\Sigma}_{\mathrm{scat.}}^{<(>)} = \gamma\mathrm{Diag}\left[\mathbf{G}^{<(>)}\right] \quad (7)$$

where γ is the interaction coupling factor [19]. Using (7) and the method described in [20] and assuming γ is small, (4) and (5) simplify to (see the Appendix for a proof)

$$\langle I\rangle = \frac{2e}{h}\int dE\,\{(f_S(E)-f_D(E))\,\mathrm{Tr}\,[\mathcal{T}(E)]\} \quad (8)$$

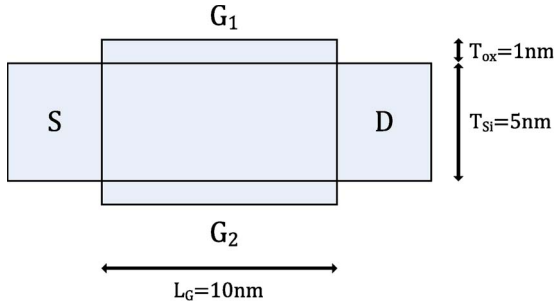

Fig. 1. Schematic of an ITRS-predicted LSTP DG-MOSFET with an effective gate length of $L_G = 10$ nm, Si channel thickness of $T_{\mathrm{Si}} = 5$ nm, and equivalent oxide thickness of $T_{\mathrm{ox}} = 1$ nm.

$$S = \frac{4e^2}{h}\int dE\,\{[f_S(1-f_S)+f_D(1-f_D)]\,\mathrm{Tr}[\mathcal{T}^2] + [f_S(1-f_D)+f_D(1-f_S)]\times\mathrm{Tr}\,[\mathcal{T}(\mathbf{I}-\mathcal{T})]\} \quad (9)$$

where

$$\mathcal{T}(E) = \mathbf{\Gamma}_S\mathbf{G}^r(\mathbf{\Gamma}_D+\mathbf{K}_D)\mathbf{G}^a \quad (10)$$

and

$$\mathbf{K}_D(E) = \gamma\mathrm{Diag}\,[\mathbf{G}^r(\mathbf{\Gamma}_D+\mathbf{K}_D)\mathbf{G}^a] \quad (11)$$

are the analogs of $\mathbf{T}(E)$ and $\mathbf{\Gamma}_D(E)$ for the noncoherent case. When scattering effects are off, $\gamma = 0$ and (8)–(10) reduce to (1)–(3). The reason S and D do not appear in a symmetric form in (10) is that this equation has been derived from (4), which describes $\langle I\rangle$ as the current entering the channel through contact S [20]. It should be noted that (2) could be rewritten in a form similar to that of (9) by expanding the quadratic term $(f_S-f_D)^2$.

Fig. 1 illustrates the schematic of an international technology roadmap for semiconductors (ITRS)-predicted [21] low standby power (LSTP) DG-MOSFET with an effective gate length of $L_G = 10$ nm, Si channel thickness of $T_{\mathrm{Si}} = 5$ nm, and equivalent oxide thickness of $T_{\mathrm{ox}} = 1$ nm. Fig. 2(a) and (b) shows, in linear and logarithmic scales, the I_{DS}–V_{DS}/V_{GS} characteristic of such a device, calculated using our 2-D Schrödinger–Poisson solver with LA scattering effects turned on and plotted over the V_{GS}–V_{DS} plane. Fig. 2(c) illustrates the Fano factor $F \equiv S/2e\langle I\rangle$ over the V_{GS}–V_{DS} plane for the same device, with scattering effects turned on and off. Evident from the figure is that, at low gate voltages, where the transistor is in the subthreshold regime, $F = 1$, indicating Poissonian noise. As V_{GS} increases, however, the noise becomes sub-Poissonian in both ballistic (shown in black) and scattering-dominated (shown in gray) cases with a Fano factor of $F < 1$. However, since we have chosen a relatively low V_{DD}, the Fano factor does not fall too far below unity ($F \sim 0.7$ for $V_{GS} = V_{DD} = 0.5$ V) as in diffusive ballistic conductors [22], [23].

At low drain voltages V_{DS}, the current vanishes along with the shot noise. However, since the thermal noise is still present, the overall noise is nonzero, causing the Fano factor to diverge.

As mentioned before, only LA scattering was included in the simulations. Had optical phonon scattering been also included,

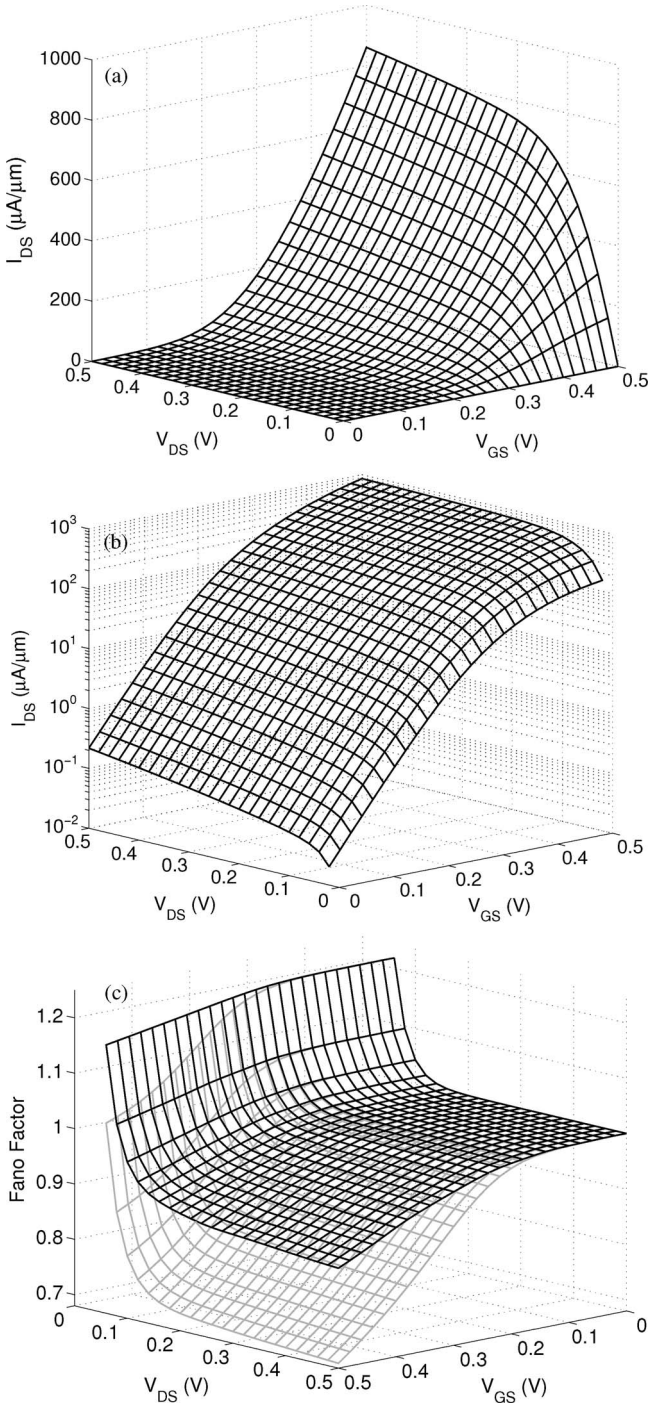

Fig. 2. I–V and noise characteristics for a DG-MOSFET with $L_G = 10$ nm and Si channel thickness of 5 nm, calculated using a 2-D Schrödinger–Poisson solver. The x and y axes represent the V_{GS} and V_{DS}, and the z-axis represents (a) the current in linear scale, (b) the current in logarithmic scale, and (c) the Fano factor in a fully ballistic case (shown in gray) and with LA scattering effects turned on (shown in solid black).

it would not have changed the near-unity value of the Fano factor below the threshold voltage. Instead, it would broaden the energy distribution of electrons in the drain, thereby suppressing shot noise in a wider gate-voltage range, i.e., it would suppress shot noise more in weak inversion and less in strong inversion compared with the ballistic transistor [24], [25].

It is well known that when the Fano factor $F = 1$, the probability distribution function (PDF) for the number of electrons passing through the device during a time interval T is a Poissonian with a mean value of $\mu = \langle I \rangle T/e$ [26]. This makes it possible to probabilistically model the noise-related behavior of CMOS devices operating in the subthreshold regime [5], [6]. However, for above-threshold operation, the PDF for noise becomes non-Poissonian, and different distribution functions based on the value of the Fano factor are needed, as discussed below.

A. Noise Distribution for $F > 1$ at Low V_{DS} Bias

As mentioned before, this situation only happens for low V_{DS} voltages when the mean current $\langle I \rangle$ approaches zero. Inspired by [7], this effect could be thought of as the two constituent Poisson sources representing the forward and the reverse currents in a MOSFET becoming close in value, resulting in the overall current vanishing. Assuming $P_f(x_f)$ and $P_r(x_r)$ represent the aforementioned Poisson sources, where x_f and x_r represent the forward- and reverse-going number of electrons, respectively, the PDF for the net current would be the convolution of the two. Thus

$$P_{\text{net}}(x) = \sum_{x=x_f-x_r} P_f(x_f) P_r(x_r) = P_f(x) * P_r(-x). \quad (12)$$

The characteristic function corresponding to $P_{\text{net}}(x)$ is expressed as

$$\begin{aligned} \phi_P(\omega) &= \mathrm{E}[e^{i\omega x}] = \sum_{x=-\infty}^{x=+\infty} \sum_{x=n-m} e^{i\omega x} P_f(n) P_r(m) \\ &= \exp\left[(\mu_f + \mu_r)\left(\cos(\omega) - 1\right) + i(\mu_f - \mu_r)\sin(\omega)\right]. \end{aligned} \quad (13)$$

Using the relation $\mathrm{E}[x^k] = (\partial^k / \partial(i\omega)^k)\phi_P(\omega)|_{\omega=0}$, the mean and the variance of $P_{\text{net}}(x)$ turn out to be

$$\langle I \rangle / e = \mu_f - \mu_r \quad (14)$$

$$S/2e^2 = \mu_f + \mu_r. \quad (15)$$

Inverting (14) and (15), the forward and reverse means μ_f and μ_r for the two constituent Poissonians can be written based on the average and variance of the total current. Thus

$$\mu_f = \left(S/2e^2 + \langle I \rangle / e\right)/2 \quad (16)$$

$$\mu_r = \left(S/2e^2 - \langle I \rangle / e\right)/2. \quad (17)$$

In the case of small V_{DS} and large V_{GS}, the noise PDF might slightly deviate from (12). However, this deviation has virtually no impact on failure rate calculation, since, as shown in Fig. 3 of Part II and also discussed in [5], the transitions between the two logic states of a flip-flop, which are represented by the two opposite corners on the 2-D queue, predominantly occur through a diagonal ribbon between the two corners (approximately marked by $V_{GS} + V_{DS} = V_{DD}$) where V_{DS} and V_{GS} adopt mid-range values, hence $V_{DS} \sim 0$ does not occur.

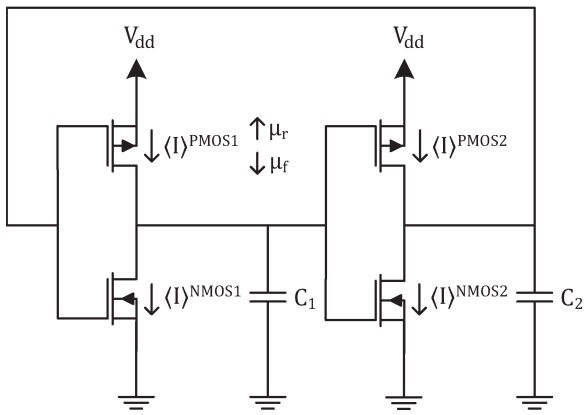

Fig. 3. Modeled flip-flop circuit. Capacitors C_1 and C_2 represent the node capacitance values associated with each inverter. For each transistor, the average currents $\langle I\rangle$ and noise values S determine the electron populations on the node capacitors. The current $\langle I\rangle$ can be decomposed into two Poisson sources of opposite directions with means μ_f and μ_r for low V_{DD} when the Fano factor $F > 1$. When $F < 1$, the statistics for the number of electron arrivals is given by (25), where $q = 1/F$.

B. Noise Distribution for $F < 1$ at High V_{GS} Bias

The interarrival-time distribution associated to a full shot noise $(F = 1)$ with an average count of μ in unit time is expressed as

$$G(\mu; t) = \mu e^{-\mu t} \tag{18}$$

which has a mean interarrival time of $\tau = 1/\mu$ and a variance of $\sigma_\tau^2 = 1/\mu^2$. Equation (18) is the unique solution of a maximum-entropy distribution problem with a fixed mean and describes the most disordered single-source shot noise. A generalization of (18), describing a more ordered arrival of particles, is based on a family of gamma distributions with a shape parameter q [27], [28]. Thus

$$G^{(1)}(q, \mu; t) = \frac{q^q \mu^q t^{q-1}}{\Gamma(q)} e^{-q\mu t}. \tag{19}$$

The characteristic function corresponding to (19) can be expressed as

$$\phi_G(\omega) = \mathrm{E}[e^{i\omega t}] = \left(1 - \frac{i\omega}{q\mu}\right)^{-q} \tag{20}$$

from which the mean and the variance for interarrival times can be found to be

$$\tau_q = 1/\mu \tag{21}$$

$$\sigma_{\tau_q}^2 = 1/q\mu^2 \tag{22}$$

where the mean is equal to that of the Poissonian in (18), whereas the variance is lowered by a factor of $1/q$. For $q = 1$, (19) simplifies to (18), describing independent processes. However, for $q > 1$, $G^{(1)}(q, \mu; t)$ approaches zero when $t \to 0$, suggesting that the process is somewhat biased toward not letting two consecutive electrons arrive too close to each other, thus making the arrivals more ordered.

This ordered arrival of electrons is observed, for example, when Coulombic correlations in the transistor channel are taken into account, where Monte Carlo simulations show that long-range Coulomb interactions cause a motional squeezing among electrons, which regularizes their arrival at the drain [29], [30]. However, as will be discussed later in this paper, the results of our computation are not particularly sensitive to the form of the interarrival-time distribution used, and (19) could be replaced by other distributions such as the Weibull distribution [27], as long as they exhibit the same characteristics up to the second moment and continuously reduce to (18) when $F \to 1$.

To calculate $P(q, \mu, T; n)$, the discrete probability distribution for the number of events n in a time interval $[0, T]$ associated with (19), we first formulate the probability for the sum of two consecutive interarrival times $t = t_1 + t_2$ as

$$G^{(2)}(q, \mu; t) = \int_0^t dz G^{(1)}(q, \mu; z) G^{(1)}(q, \mu; t - z) \tag{23}$$

which, using induction, could be easily generalized to that of the sum of n consecutive interarrival times. Thus

$$G^{(n)}(q, \mu; t) = \int_0^t dz G^{(1)}(q, \mu; z) G^{(n-1)}(q, \mu; t - z). \tag{24}$$

Now using (24), $P(q, \mu, T; n)$ can be rewritten as

$$P(q, \mu, T; n) = \int_0^T dt \left(G^{(n)}(q, \mu; t) - G^{(n+1)}(q, \mu; t)\right) \tag{25}$$

where the difference between probability of at least n events and that of at least $n + 1$ events occurring determines the probability of exactly n events happening during time T.

Unlike full shot noise, where the Fano factor is independent of time T, in the case of a sub-Poisson process characterized by (25), the Fano factor depends on T, ranging from $F \to 1$ for very small T to $F \to 1/q$ for large values of T. For frequencies low enough so that transit and scattering relaxation times can be ignored [26], the current noise shape parameter q can be set to its long-time value

$$q = 1/F. \tag{26}$$

It should be noted that the case of a full Poisson noise with $F = 1$ can be identically described as either a limiting case of $F > 1$, where $\mu_f = \langle I\rangle/e$ and $\mu_r = 0$ in (16) and (17) or as a limiting case of $F < 1$ with $q = 1$.

C. Use of Different Interarrival-Time Distributions

As mentioned earlier, the use of a gamma distribution for electron interarrival times is not unique, and as long as the conditions described in Section II-B are met (i.e., having predefined means and variances and continuously reducing to the Poisson distribution as the Fano factor approaches unity), the queue is not sensitive to the exact choice of an interarrival-time distribution. This has been verified in Part II [8] by testing the queue using different types of distributions as a replacement for (19) and obtaining the same results. Further investigation

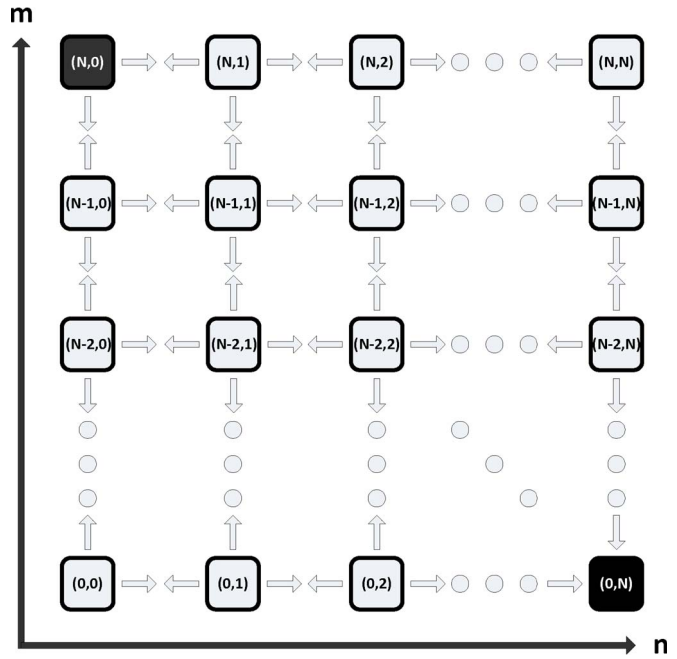

Fig. 4. Two-dimensional Markov chain for a flip-flop. Each axis labels different states corresponding to different voltages on capacitors C_1 and C_2.

showed that the independence of the final results from the exact shape of the interarrival-time distributions stems from the large number of states in the queue, rendering the relevance of higher moments minimal, as soon as the number of such states exceeds $\sim$10. This situation is very similar in practice to conditions of the central limit theorem [31].

III. Markov Chain for Above-Threshold Devices

In this section, an outline of the probabilistic formulation using Markov chains is presented. The approach presented here extends the one in [5] and [6] to model above-threshold devices by incorporating new statistics introduced in the previous section. The inputs for this method are the number and values of node capacitance, which determine the dimensionality of the problem, and full $I_{DS}-V_{DS}/V_{GS}$ curves as well as the noise values or the corresponding Fano factor at every bias point.

Assume the states are laid out as in Fig. 4, where the horizontal axis represents the voltage [or charge through $Q = Q(V)$] on capacitor 1 of the flip-flop, and the vertical axis represents the voltage on capacitor 2 of the same device. Depending on the relevant Fano factors of the pMOS and nMOS transistors of an inverter in Fig. 3, the time-dependent event generators corresponding to charging and discharging of capacitor C_i ($i = 1, 2$) can take on different forms.

- $F > 1$ for both pMOSi and nMOSi, in which case, based on (16), (17), and the directional currents defined in Fig. 3, the charging and discharging event generators for capacitor C_i take on the forms

$$P^{(i)}_{\text{charge}}(t) = \left(\mu_f^{\text{pMOS}i} + \mu_r^{\text{nMOS}i}\right) e^{-\left(\mu_f^{\text{pMOS}i}+\mu_r^{\text{nMOS}i}\right)t} \quad (27)$$

$$P^{(i)}_{\text{disch.}}(t) = \left(\mu_r^{\text{pMOS}i} + \mu_f^{\text{nMOS}i}\right) e^{-\left(\mu_r^{\text{pMOS}i}+\mu_f^{\text{nMOS}i}\right)t}. \quad (28)$$

- $F < 1$ for both pMOSi and nMOSi, in which case the transistor with a positive mean current going out of C_i is responsible for discharging it (normally the nMOS), and the transistor with a positive mean current going into C_i (normally the pMOS) is responsible for charging C_i where the event generator is given by (19).
- $F < 1$ for one of the MOSFETs (say pMOSi) and $F > 1$ for the other MOSFET (nMOSi). In this case, the only discharging random source is a Poissonian with a mean of $\mu_f^{\text{nMOS}i}$. However, there are two charging random sources of different types: a Poissonian with a mean of $\mu_r^{\text{nMOS}i}$ and a gamma random source due to the pMOS. Whichever of these sources triggers first will cause an event yielding

$$P^{(i)}_{\text{charge}}(t) = G^{(1)}\left(q, \langle I\rangle^{\text{pMOS}i}; t\right) \int_t^\infty dz G\left(\mu_r^{\text{nMOS}i}; z\right) + G\left(\mu_r^{\text{nMOS}i}; t\right) \int_t^\infty dz G^{(1)}\left(q, \langle I\rangle^{\text{pMOS}i}; z\right) \quad (29)$$

where $G(\mu_r^{\text{nMOS}i}; z)$ and $G^{(1)}(q, \langle I\rangle^{\text{pMOS}i}; t)$ are given by (18) and (19), respectively.

It should be noted that the four event generators corresponding to charging and discharging capacitors C_1 and C_2 are state-dependent and responsible for transitions between the states on the queue. For example, for the transition $(m, n) \to (m + 1, n)$ to happen at time t, capacitor C_1 should be charged by a single electron while other transitions (to the other three neighbor states) are being blocked until time t by formulating the conditional probability for a direct transition as

$$P_{\text{direct}}(m, n; m+1, n; t) = P^{(1)}_{\text{charge}}(m, n; t) \int_t^\infty dz P^{(1)}_{\text{disch.}}(m, n; z) \times \int_t^\infty dz P^{(2)}_{\text{charge}}(m, n; z) \int_t^\infty dz P^{(2)}_{\text{disch.}}(m, n; z) \quad (30)$$

where the state dependence of the event generators are explicitly shown for clarity. Similar relations hold for transitions to the other neighbor states.

The mean time for an indirect transition from any state (m, n) to an arbitrary state (u, v) on the queue using multiple single-step transitions could be split into a single-step transition from (m, n) to its immediate neighbor (r, s) and an indirect transition from (r, s) to (u, v) summed over all neighbor states (r, s). Thus

$$T(m, n; u, v) = \sum_{(r,s)} \int_0^\infty \int_0^\infty dt_1 dt_2 (t_1 + t_2) \times P_{\text{direct}}(m, n; r, s; t_1) P_{\text{indirect}}(r, s; u, v; t_2). \quad (31)$$

Plugging (30) into (31), carrying out the integrals and noting that $\int_0^\infty t_2 P_{\text{indirect}}(r, s; u, v; t_2) dt_2 = T(r, s; u, v)$ and

$\int_0^\infty P_{\text{indirect}}(r,s;u,v;t_2)dt_2 = 1$, one eliminates the direct knowledge of $P_{\text{indirect}}(r,s;u,v;t_2)$ and comes to the equation represented below, which is analogous to the one in [5] for subthreshold-operated devices. Thus

$$T(m,n;u,v) = A(m,n) + \sum_{(r,s)} R(m,n;r,s)T(r,s;u,v) \tag{32}$$

where

$$R(m,n;r,s) = \int_0^\infty dt_1 P_{\text{direct}}(m,n;r,s;t_1) \tag{33}$$

$$A(m,n) = \sum_{(r,s)} \int_0^\infty dt_1 t_1 P_{\text{direct}}(m,n;r,s;t_1). \tag{34}$$

The integrals in (33) and (34) are not elementary and should be carried out numerically; we used the adaptive four-point Simpson integration for our calculations.

To gain some insight into (32), it should be noted that $A(m,n)$ represents the average residency time, i.e., the average time the system resides in state (m,n) upon arrival, the inverse of which represents the escape rate, and $R(m,n;r,s)$ denotes the probability of moving to the neighboring state (r,s) as a next step such that $\sum_{(r,s)} R(m,n;r,s) = 1$.

Equation (32) repeated for every state (m,n) on the queue yields a system of equations for mean transit times to the state (u,v) from anywhere on the queue, which could be solved using sparse matrix techniques combined with multiple precision arithmetic to take care of any ill-conditioning issues [5].

If (m,n) represents the logic "0" of the flip-flop with C_1 fully charged and C_2 empty of charge and (u,v) represents the logic "1" of the flip-flop with the roles of C_1 and C_2 interchanged, then $T(m,n;u,v)$, given by (32), will represent the mean time to spontaneous failure for the device due to shot noise, taking into account the exact $I_{DS}-V_{DS}/V_{GS}$ as well as $S_{DS}-V_{DS}/V_{GS}$ and other technology-dependent characteristics of the CMOS device.

The steady-state occupation probability for every state $P_{s.s.}(m,n)$ can be calculated using the fact that under steady-state conditions, the incoming probability flux is equal to the outgoing flux. Thus

$$A^{-1}(m,n)P_{s.s.}(m,n) = \sum_{(r,s)} \frac{R(r,s;m,n)}{A(r,s)} P_{s.s.}(r,s). \tag{35}$$

IV. Conclusions and Future Work

This paper presents a probabilistic framework for the investigation of the effect of noise on the logic stability of memory devices operated above the threshold voltage V_T. By using a Schrödinger–Poisson solver that includes only LA phonon scattering as a model scattering mechanism via the NEGF method (without any effort to reflect all scattering mechanisms relevant to the 10-nm-gate-length CMOS technology), the shot-noise-driven statistics of the charge fluctuations is decomposed into departure and arrival rates for the carriers at the node capacitors. Embedding the transient charge configuration of the flip-flop into a 2-D queue allows for a Markov-chain-based derivation of the error rates through an algebraic solution yielding the transit times between the two most stable logic-level states in the queue.

The proposed analytical formalism for the shot-noise effects has been applied to investigate different regimes of operations characterized by distinct Fano factors and benchmark values for the gate and drain voltages. Using sparse matrix algorithms and multiple precision arithmetic (MPA) to solve for the system of equations for the mean transit times allows for a detailed analysis of noise-induced errors in the ITRS-predicted 10-nm-gate-length technology, as carried out in Part II of this work.

While the main focus of this work was failure rates due to shot noise, the same Markov formalism will be applied in the future to other relevant noise sources such as low-frequency-dominant $1/f$ noise and RTS noise.

Appendix
Derivation of (9)

The low frequency noise spectral density is given by (5), which is repeated here

$$\begin{aligned} S = \frac{4e^2}{h} \int dE \operatorname{Tr} \{ & i\mathbf{\Gamma}_S \left[f_S \mathbf{G}^> - (1-f_S)\mathbf{G}^< \right] \\ & \times \left[\mathbf{I} - i\mathbf{\Gamma}_S(\mathbf{G}^r - \mathbf{G}^a) \right] \\ & + \mathbf{\Gamma}_S \mathbf{G}^> \mathbf{\Gamma}_S \mathbf{G}^< - f_S(1-f_S) \\ & \times (\mathbf{\Gamma}_S \mathbf{G}^r \mathbf{\Gamma}_S \mathbf{G}^r + \mathbf{\Gamma}_S \mathbf{G}^a \mathbf{\Gamma}_S \mathbf{G}^a) \}. \end{aligned} \tag{A.1}$$

The following relations are assumed:

$$\mathbf{G}^< = if_S \mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a + if_D \mathbf{G}^r(\mathbf{\Gamma}_D + \mathbf{K}_D)\mathbf{G}^a \tag{A.2}$$

$$\begin{aligned} \mathbf{G}^> = & -i(1-f_S)\mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a \\ & - i(1-f_D)\mathbf{G}^r(\mathbf{\Gamma}_D + \mathbf{K}_D)\mathbf{G}^a \end{aligned} \tag{A.3}$$

$$\begin{aligned} \mathbf{G}^r - \mathbf{G}^a &= \mathbf{G}^> - \mathbf{G}^< \\ &= -i\mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a - i\mathbf{G}^r(\mathbf{\Gamma}_D + \mathbf{K}_D)\mathbf{G}^a. \end{aligned} \tag{A.4}$$

Substituting (A.2)–(A.4) into (A.1), making use of (10) and a simplifying definition $\mathbf{A} \equiv \mathbf{\Gamma}_S \mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a$, one can simplify the terms in (A.1) as follows:

$$\begin{aligned} & i\mathbf{\Gamma}_S \left[f_S \mathbf{G}^> - (1-f_S)\mathbf{G}^< \right] \\ & \quad = 2f_S(1-f_S)\mathbf{\Gamma}_S \mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a \\ & \qquad + \left[f_S(1-f_D) + f_D(1-f_S) \right] \mathbf{\Gamma}_S \mathbf{G}^r(\mathbf{\Gamma}_D + \mathbf{K}_D)\mathbf{G}^a \\ & \quad = 2f_S(1-f_S)\mathbf{A} + \left[f_S(1-f_D) + f_D(1-f_S) \right] \mathcal{T} \end{aligned} \tag{A.5}$$

$$\begin{aligned} & \mathbf{I} - i\mathbf{\Gamma}_S(\mathbf{G}^r - \mathbf{G}^a) \\ & \quad = \mathbf{I} - \mathbf{\Gamma}_S \mathbf{G}^r(\mathbf{\Gamma}_S + \mathbf{K}_S)\mathbf{G}^a - \mathbf{\Gamma}_S \mathbf{G}^r(\mathbf{\Gamma}_D + \mathbf{K}_D)\mathbf{G}^a \\ & \quad = \mathbf{I} - \mathbf{A} - T \end{aligned} \tag{A.6}$$

$$\begin{aligned}
&\mathbf{\Gamma}_S\mathbf{G}^{>}\mathbf{\Gamma}_S\mathbf{G}^{<} \\
&\quad = \mathbf{\Gamma}_S\left[-i(1-f_S)\mathbf{G}^r(\mathbf{\Gamma}_S+\mathbf{K}_S)\mathbf{G}^a\right. \\
&\qquad \left. -i(1-f_D)\mathbf{G}^r(\mathbf{\Gamma}_D+\mathbf{K}_D)\mathbf{G}^a\right]\mathbf{\Gamma}_S \\
&\qquad \times\left[if_S\mathbf{G}^r(\mathbf{\Gamma}_S+\mathbf{K}_S)\mathbf{G}^a+if_D\mathbf{G}^r(\mathbf{\Gamma}_D+\mathbf{K}_D)\mathbf{G}^a\right] \\
&\quad = f_S(1-f_S)\mathbf{A}^2+f_D(1-f_D)\mathcal{T}^2 \\
&\qquad +\left[f_S(1-f_D)+f_D(1-f_S)\right]\mathbf{A}\mathcal{T} \qquad \text{(A.7)} \\
&\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_S\mathbf{G}^r+\mathbf{\Gamma}_S\mathbf{G}^a\mathbf{\Gamma}_S\mathbf{G}^a \\
&\quad = (\mathbf{\Gamma}_S\mathbf{G}^r)^2+(\mathbf{\Gamma}_S\mathbf{G}^a)^2 \\
&\quad = \left[\mathbf{\Gamma}_S(\mathbf{G}^r-\mathbf{G}^a)\right]^2+2\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_S\mathbf{G}^a \\
&\quad = \left[-i\mathbf{\Gamma}_S\mathbf{G}^r(\mathbf{\Gamma}_S+\mathbf{K}_S)\mathbf{G}^a-i\mathbf{\Gamma}_S\mathbf{G}^r(\mathbf{\Gamma}_D+\mathbf{K}_D)\mathbf{G}^a\right]^2 \\
&\qquad +2\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_S\mathbf{G}^a=-(\mathbf{A}+\mathcal{T})^2+2\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{\Gamma}_S\mathbf{G}^a \qquad \text{(A.8)}
\end{aligned}$$

where we used the cyclic property of the trace operator as needed. Putting (A.5)–(A.8) back into (A.1) yields

$$\begin{aligned}
S=\frac{4e^2}{h}\int dE\,\{&f_S(1-f_S)\mathrm{Tr}\left[\mathcal{T}^2+\mathbf{O}(\gamma)\right]+f_D(1-f_D) \\
&\times\ \mathrm{Tr}[\mathcal{T}^2]+\left[f_S(1-f_D)+f_D(1-f_S)\right] \\
&\times\mathrm{Tr}\left[\mathcal{T}(\mathbf{I}-\mathcal{T})\right]\} \qquad \text{(A.9)}
\end{aligned}$$

where $\mathbf{O}(\gamma)=2\mathbf{\Gamma}_S\mathbf{G}^r\mathbf{K}_S\mathbf{G}^a$, which is proportional to the coupling constant γ through $\mathbf{K}_S$, is very small in the case of a weak electron-phonon interaction. Moreover, assuming an applied bias of $\mu_S-\mu_D\gg kT$, the term containing $f_S(1-f_D)$ becomes dominant, and the effect of $\mathbf{O}(\gamma)$ becomes negligible.

ACKNOWLEDGMENT

The authors would like to thank Dr. A. Cresti for a helpful discussion.

REFERENCES

[1] M. Choudhury and K. Mohanram, "Reliability analysis of logic circuits," *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.*, vol. 28, no. 3, pp. 392–405, Mar. 2009.

[2] B. Calhoun and A. Chandrakasan, "A 256-kb 65-nm sub-threshold SRAM design for ultra-low-voltage operation," *IEEE J. Solid-State Circuits*, vol. 42, no. 3, pp. 680–688, Mar. 2007.

[3] H. Klauk, U. Zschieschang, J. Pflaum, and M. Halik, "Ultralow-power organic complementary circuits," *Nature*, vol. 445, no. 7129, pp. 745–748, Feb. 2007.

[4] F. C. Sabou, D. Kazazis, W. R. Bahar, R. I. Patterson, J. Mundy, and A. Zaslavsky, "Markov chain analysis of thermally induced soft errors in subthreshold nanoscale CMOS circuits," *IEEE Trans. Device Mater. Rel.*, vol. 9, no. 3, pp. 494–504, Sep. 2009.

[5] P. Jannaty, F. C. Sabou, R. I. Bahar, W. R. Patterson, J. Mundy, and A. Zaslavsky, "Full two-dimensional Markov chain analysis of thermal soft errors in subthreshold nanoscale CMOS devices," *IEEE Trans. Device Mater. Rel.*, vol. 11, no. 1, pp. 50–59, Mar. 2011.

[6] P. Jannaty, F. C. Sabou, M. Gadlage, R. I. Bahar, W. Patterson, J. Mundy, A. Reed, R. A. Weller, R. D. Schrimpf, and A. Zaslavsky, "Two-dimensional Markov chain analysis of radiation-induced soft errors in subthreshold nanoscale CMOS devices," *IEEE Trans. Nucl. Sci.*, vol. 57, no. 6, pp. 3768–3774, Dec. 2010.

[7] R. Sarpeshkar, T. Delbruck, and C. Mead, "White noise in mos transistors and resistors," *IEEE Circuits Devices Mag.*, vol. 9, no. 6, pp. 23–29, Nov. 1993.

[8] P. Jannaty, F. C. Sabou, S. T. Le, M. Donato, R. I. Bahar, W. R. Patterson, J. Mundy, and A. Zaslavsky, "Shot-noise-induced failure in nanoscale flip-flops: II. Failure rates in 10nm ultimate CMOS," *IEEE Trans. Electron Devices*, vol. 59, no. 3, pp. 807–812, Mar. 2012.

[9] M. Büttiker, "Scattering theory of current and intensity noise correlations in conductors and wave guides," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 46, no. 19, pp. 12 485–12 507, Nov. 1992.

[10] T. Martin and R. Landauer, "Wave-packet approach to noise in multichannel mesoscopic systems," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 45, no. 4, pp. 1742–1755, Jan. 1992.

[11] Y. Meir and N. S. Wingreen, "Landauer formula for the current through an interacting electron region," *Phys. Rev. Lett.*, vol. 68, no. 16, pp. 2512–2515, Apr. 1992.

[12] J.-X. Zhu and A. V. Balatsky, "Theory of current and shot-noise spectroscopy in single-molecular quantum dots with a phonon mode," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 67, no. 16, p. 165326, Apr. 2003.

[13] F. Haupt, T. C. V. Novotný, and W. Belzig, "Current noise in molecular junctions: Effects of the electron-phonon interaction," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 82, no. 16, p. 165441, Oct. 2010.

[14] S. Martinie, D. Munteanu, G. Le Carval, and J.-L. Autran, "Physics-based analytical modeling of quasi-ballistic transport in double-gate MOSFETs: From device to circuit operation," *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2692–2702, Nov. 2009.

[15] M. Gilbert and S. Banerjee, "Ballistic to diffusive crossover in III-V nanowire transistors," *IEEE Trans. Electron Devices*, vol. 54, no. 4, pp. 645–653, Apr. 2007.

[16] V. Barral, T. Poiroux, S. Barraud, O. Bonno, F. Andrieu, C. Buj-Dufournet, L. Brevard, D. Lafond, O. Faynot, D. Munteanu, J. Autran, and S. Deleonibus, "Electron mean-free-path experimental extraction on ultra-thin and ultra-short strained and unstrained FDSOI n-MOSFETs," in *Proc. IEEE SNW*, 2008, pp. 1–2.

[17] F. Boeuf, X. Jehl, M. Sanquer, and T. Skotnicki, "Experimental study of transport in nanoscale planar MOSFETs near the ballistic limit," *IEEE Trans. Nanotechnol.*, vol. 3, no. 1, pp. 105–109, Mar. 2004.

[18] R. Wang, H. Liu, R. Huang, J. Zhuge, L. Zhang, D.-W. Kim, X. Zhang, D. Park, and Y. Wang, "Experimental investigations on carrier transport in Si nanowire transistors: Ballistic efficiency and apparent mobility," *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 2960–2967, Nov. 2008.

[19] R. Lake, G. Klimeck, R. C. Bowen, and D. Jovanovic, "Single and multiband modeling of quantum electron transport through layered semiconductor devices," *J. Appl. Phys.*, vol. 81, no. 12, pp. 7845–7869, Jun. 1997.

[20] A. Cresti and G. P. Parravicini, "Dephasing effects and shot noise in quantum Hall wires: Green's function formalism," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 78, no. 11, p. 115313, Sep. 2008.

[21] *International Technology Roadmap for Semiconductors*, 2011. [Online]. Available: http://public.itrs.net

[22] C. Beenakker and C. Schönenberger, "Quantum shot noise," *Phys. Today*, vol. 56, pp. 37–42, May 2003.

[23] C. W. J. Beenakker and M. Büttiker, "Suppression of shot noise in metallic diffusive conductors," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 46, no. 3, pp. 1889–1892, Jul. 1992.

[24] H.-H. Park, S. Jin, Y. J. Park, and H. S. Min, "Quantum simulation of noise in silicon nanowire transistors with electron-phonon interactions," *J. Appl. Phys.*, vol. 105, no. 2, pp. 023712–023712–6, Jan. 2009.

[25] H. Tsuchiya and S. Takagi, "Influence of elastic and inelastic phonon scattering on the drive current of quasi-ballistic MOSFETs," *IEEE Trans. Electron Devices*, vol. 55, no. 9, pp. 2397–2402, Sep. 2008.

[26] Y. M. Blanter and M. Buttiker, "Shot noise in mesoscopic conductors," *Phys. Rep.*, vol. 336, p. 1, 2000.

[27] J. Fade, N. Treps, C. Fabre, and P. Refregier, "Optimal precision of parameter estimation in images with local sub-Poissonian quantum fluctuations," *Eur. Phys. J. D—Atomic, Molecular, Optical Plasma Physics*, vol. 50, no. 2, pp. 215–227, Dec. 2008.

[28] M. C. Teich, B. E. A. Saleh, and J. Peina, "Role of primary excitation statistics in the generation of antibunched and sub-poisson light," *J Opt. Soc. Amer. B*, vol. 1, no. 3, pp. 366–389, Jun. 1984.

[29] O. M. Bulashenko, J. Mateos, D. Pardo, T. González, L. Reggiani, and M. Rubıacute;, "Electron-number statistics and shot-noise suppression by Coulomb correlation in nondegenerate ballistic transport," *Phys. Rev. B, Condens. Matter Mater. Phys.*, vol. 57, no. 3, pp. 1366–1369, Jan. 1998.

[30] A. Betti, G. Fiori, and G. Iannaccone, "Shot noise in quasi one-dimensional FETs," in *IEDM Tech. Dig.*, 2008, pp. 1–4.

[31] J. A. Rice, *Mathematical Statistics and Data Analysis*, 2nd ed. Pacific Grove, CA: Duxbury Press, 1994.

Author photographs and biographies not available at the time of publication.