

Multi-Scale Statistical Modeling of Random Telegraph Noise and Read-Noise Distribution in Sub-Micrometer CMOS Image Sensors

Jae Ho Kim*, Sungchul Kim, Yonghee Park, Jonghyun Go¹, and Jae Hoon Jung
 Computational Science and Engineering Team, ¹CIS Technology Design Team
 Semiconductor R&D Center, Samsung Electronics, Hwaseong, Gyeonggi, Republic of Korea
 *Email: icos.kim@samsung.com

Abstract—Random telegraph noise (RTN) is a critical limiter of image quality in highly scaled CMOS image sensors (CIS), where a small population of large-amplitude pixels forms a heavy tail in the read-noise (RN) distribution. Reconstructing this distribution over a full pixel array has been impractical with full Monte-Carlo transient simulation, whose cost grows prohibitively with the number of frames. This paper presents an ultrafast, multi-scale statistical framework that bridges atomistic density functional theory (DFT) with the macroscopic RN distribution: a TCAD-evaluated voltage deviation function (VDF) from the Green’s function method is combined with a fully analytical Markov-chain model of the correlated double sampling (CDS) readout, replacing per-frame time-stepping with a closed-form solution. Using this framework, we identify a dominant physical origin of the severe RTN tail in L-shaped source-follower (SF) transistors as the spatial-energetic coincidence of a corner trap with the CDS resonance window, and we show that the fluorine ion-implantation (F-IIP) passivation process suppresses the tail by lowering trap density and shifting the residual energetics out of resonance, in quantitative agreement with silicon measurements.

Index Terms—random telegraph noise, CMOS image sensor, density functional theory, source follower, Green’s function, correlated double sampling

I. INTRODUCTION

As the pixel pitch of CMOS image sensors (CIS) is aggressively scaled into the sub-micrometer regime, random telegraph noise (RTN) originating in the source-follower (SF) transistor has become the dominant contributor to dark temporal noise and the principal limiter of low-light image quality. In the area-constrained unit pixel, an L-shaped SF channel is widely adopted to maximize the effective channel length within the available active region (Fig. 1(a)). While correlated double sampling (CDS) cancels reset (kTC) and offset noise, the low-frequency flicker and RTN components of the SF persist through the CDS operation (Fig. 1(b)). Consequently, a small population of pixels exhibiting large-amplitude RTN forms a heavy tail in the measured read-noise (RN) distribution, and these tail pixels dominate the visually objectionable white-spot defects [1].

Predicting such tail pixels requires a model that connects the atomistic nature of the responsible traps to the macroscopic, circuit-level noise that is measured. Previous statistical models employed the Green’s function method to evaluate RTN amplitudes [2], building on the classical picture of trap-induced

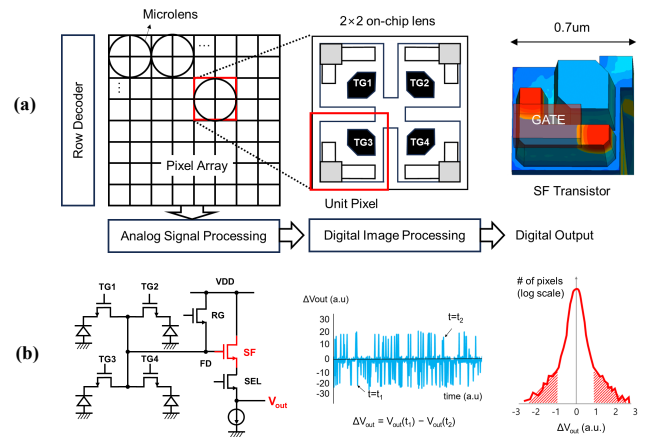


Fig. 1. (a) Schematic of the pixel array featuring a 2×2 on-chip lens layout for 0.7 μm pixels and the L-shaped source follower (SF) transistor structure. (b) Conceptual diagram of the CIS readout operation employing correlated double sampling (CDS) and the resulting noise output signals.

switching in MOS structures [5], but generally treated the trap signal in isolation and omitted the baseline analog read noise in the transient CDS domain [6]—which is essential for reconstructing the full pixel-variance histogram rather than only its amplitude. Building on our ab-initio study of fluorine-passivated defects [3], we propose a multi-scale simulator spanning atomistic DFT, device-level Green’s functions, and circuit-level continuous-time Markov chains.

The contributions of this work are threefold. First, we develop an ultrafast multi-scale framework linking atomistic DFT trap properties, a TCAD-evaluated voltage deviation function (VDF), and an analytical continuous-time Markov-chain model of the CDS readout. Second, we show that the severe RTN tail arises not from trap amplitude alone but from the spatial-energetic coincidence of a trap at the corner of the L-shaped SF channel with the CDS resonance window. Third, we demonstrate, in agreement with silicon, that F-IIP passivation removes these resonant corner traps by lowering trap density and shifting the residual energetics out of resonance.

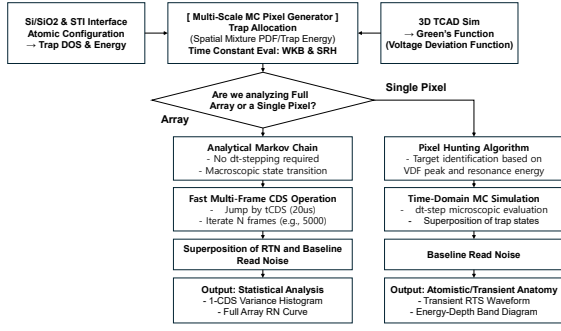


Fig. 2. The proposed multi-scale statistical simulation flow, integrating atomistic trap characteristics (DFT), the TCAD-evaluated voltage deviation function (VDF), and the analytical Markov-chain method for macroscopic CDS operations.

II. MULTI-SCALE SIMULATION FRAMEWORK

The overall simulation flow is illustrated in Fig. 2. The framework systematically integrates device physics, quantum mechanics, and circuit-level noise characteristics into a unified multi-scale methodology spanning three coupled scales. All key parameters—capture coefficients, trap densities, and trap energetics—are calibrated against silicon measurements, and a 2M-pixel ensemble is sampled to reconstruct the statistical noise distributions reported in Section III.

A. Atomistic Trap Characterization

The interface traps responsible for the SF RTN are amphoteric Si dangling bonds—the Pb centers at the Si/SiO₂ interface identified by electron spin resonance [8], [9]. Each Pb center is amphoteric, presenting a donor (+/0) level in the lower half of the gap (a hole trap) and an acceptor (0/−) level in the upper half (an electron trap), so the interface-trap density $D_{it}(E)$ follows the characteristic double-peak, U-shaped profile across the band gap. Our atomistic evaluation reproduces this distribution and indicates that, before passivation, the shallow lower-gap (hole) population dominates while the near-midgap deep states are the sparsest; F-IIP saturates these dangling bonds and lowers the population across the entire gap. This energy distribution sets the trap depths and barriers that enter the capture/emission kinetics below.

The capture (τ_c) and emission (τ_e) time constants of an individual trap are governed by Shockley–Read–Hall (SRH) statistics combined with Wentzel–Kramers–Brillouin (WKB) quantum tunneling, the latter depending exponentially on the trap depth and the local oxide field. The capture rate is expressed as

$$\tau_c^{-1} = c_n n_s f_t P_T, \quad (1)$$

where c_n is the capture coefficient, n_s the surface carrier density, f_t the trap occupancy factor, and P_T the tunneling probability,

$$P_T = \exp \left[-\frac{4\sqrt{2m^*}}{3\hbar q E_{ox}} \left(\phi_t^{3/2} - \phi_c^{3/2} \right) \right], \quad (2)$$

with m^* the effective mass, ϕ_t and ϕ_c the trap and conduction-band barrier heights, and E_{ox} the oxide field. The trap energy depths and reconstruction energetics are taken directly from our DFT calculations [3], performed within the generalized-gradient approximation [4]. To represent the realistic defect landscape, traps are generated from a Poisson–Lognormal compound distribution,

$$N_{total} = \mathcal{P}(\lambda_{cent}) + \mathcal{P}(\lambda_{edge}), \quad \lambda_{edge} \sim \text{LN}(\mu_{ln}, \sigma_{ln}^2), \quad (3)$$

which superposes a spatially uniform population of intrinsic center defects with a heavy-tailed, process-induced population clustered along the shallow-trench-isolation (STI) edges.

B. Device-Level VDF via Green's Function

To achieve numerical efficiency without sacrificing 3D physical accuracy, the RTN amplitude induced by each trapped charge is calculated using the Green's function method applied to the device's linearized system equations [2]:

$$\Delta V_i = \int G(\mathbf{r}, \mathbf{r}_i) \rho_{\text{trap}}(\mathbf{r}_i) d\mathbf{r}, \quad (4)$$

where $G(\mathbf{r}, \mathbf{r}_i)$ is the Green's function of the linearized device system and ρ_{trap} the trapped-charge distribution. Evaluating (4) over the full 3D L-shaped channel yields the TCAD-based VDF, a spatial map of the source-referred potential fluctuation caused by a defect at any coordinate. The VDF retains full 3D electrostatic accuracy while avoiding repeated transient device simulation.

C. From Langevin Dynamics to an Analytical Markov Chain

Physically, each trap is continually filled and emptied by carrier capture and emission, so its charge state $n(t) \in \{0, 1\}$, with $n = 1$ denoting the filled state, executes a random-telegraph process. Coarse-grained in time, this process obeys a Langevin equation,

$$\frac{dn}{dt} = -\frac{n - P_{\text{steady}}}{\tau_{\text{relax}}} + \Gamma(t), \quad (5)$$

whose first term relaxes the occupancy toward its steady-state filled probability $P_{\text{steady}} = \tau_e / (\tau_c + \tau_e)$, with relaxation time $\tau_{\text{relax}} = (\tau_c^{-1} + \tau_e^{-1})^{-1}$, while $\Gamma(t)$ is a zero-mean stochastic force representing the random capture and emission events. This balance of dissipation and random forcing is the microscopic origin of the RTN.

Because the two trap states and their constant rates τ_c^{-1} and τ_e^{-1} define a continuous-time Markov chain, the conditional probability of finding the trap filled after a CDS interval t_{CDS} is available in closed form—without integrating the intervening trajectory:

$$P_{\text{trans}} = P_{\text{steady}} + [S(t_1) - P_{\text{steady}}] \exp \left(-\frac{t_{\text{CDS}}}{\tau_{\text{relax}}} \right), \quad (6)$$

where $S(t_1)$ is the state sampled at the first read. Equation (6) is exactly the relaxation of the telegraph autocorrelation, which decays as $\exp(-t/\tau_{\text{relax}})$ (a Lorentzian spectrum) [7]. The CDS double sampling acts as a band filter on this spectrum,

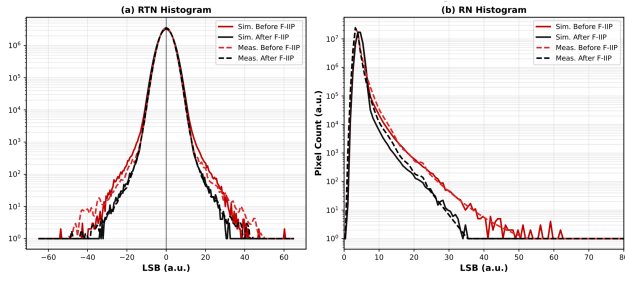


Fig. 3. Correlation between high-speed simulation and silicon measurement. (a) Simulated single-CDS histograms and (b) 20-frame variance (RN curve) distributions closely match the empirical data, accurately reflecting the noise characteristics before and after the F-IIP passivation process.

so that a trap contributes maximally when its τ_{relax} lies inside the CDS window—the “resonance” exploited in Section III.

The time-domain single-CDS output then superimposes the discrete RTN signal of all N_{traps} with the baseline analog read noise $n_{\text{read}}(k)$,

$$\Delta V_{\text{CDS}}(k) = \sum_{i=1}^{N_{\text{traps}}} \Delta V_i [S_i(k, t_2) - S_i(k, t_1)] + n_{\text{read}}(k), \quad (7)$$

and the total pixel variance is the sum of the RTN and read-noise variances,

$$\sigma_{\text{total}}^2 = \sigma_{\text{RTN}}^2 + \sigma_{\text{read}}^2. \quad (8)$$

Crucially, (6) replaces each trap’s entire transient trajectory with one closed-form evaluation per CDS frame. Reconstructing the noise statistics of a 2M-pixel array over thousands of frames—about four days by full Monte-Carlo time-stepping in our previous work [3]—is thereby cut to roughly three hours ($\sim 30\times$), making full-array statistical RTN prediction practical at the design stage.

III. RESULTS AND DISCUSSION

A. Quantitative Validation against Silicon

Fig. 3 compares the high-speed simulation with the silicon measurement. The simulated single-CDS histograms in Fig. 3(a) and the 20-frame variance (RN curve) distributions in Fig. 3(b) closely match the measured data, both before and after F-IIP passivation. The simulated median, distribution width, and tail-bit ratio above 30 LSB all reproduce the measured values within the calibration tolerance, confirming that the framework captures both the bulk of the distribution and its heavy tail. The comparison further makes the effect of the passivation directly visible: the heavy tail is strongly suppressed, and the former tail pixels are redistributed into the center peak. The physical origin of this suppression is analyzed in Section III-B.

B. Spatial-Energetic Origin of the RTN Tail and the F-IIP Mechanism

By integrating the 3D VDF and the quantum tunneling probabilities, a 2M-pixel ensemble is rigorously filtered to isolate

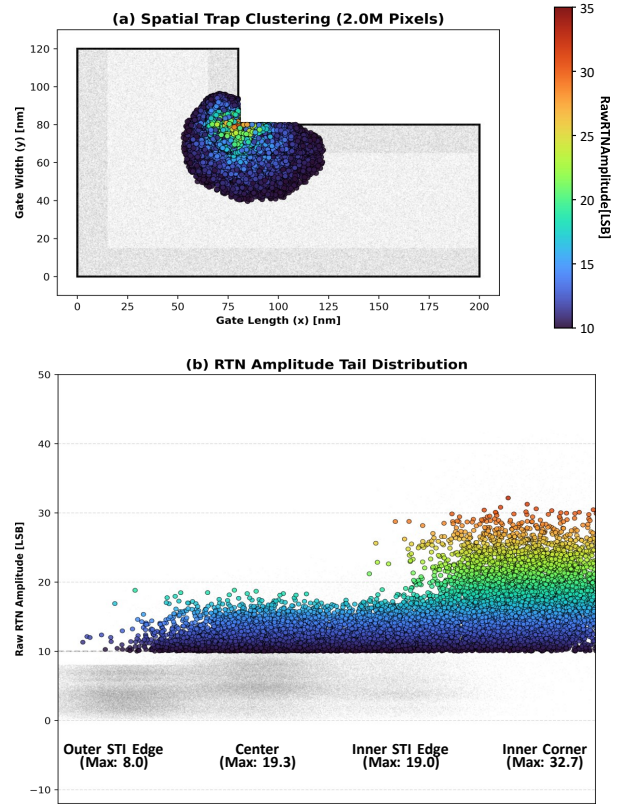


Fig. 4. Spatial and regional analysis of active RTN traps. (a) Spatial map of active single traps (>10 LSB) across the transistor geometry. (b) Amplitude scatter plot categorized by structural region, revealing that extreme RTN tails (> 30 LSB) are concentrated at the inner corner due to structural current crowding.

active single traps (>10 LSB). Their spatial map in Fig. 4(a) and their regional amplitude distribution in Fig. 4(b) reveal that extreme RTN fluctuations (>30 LSB) predominantly originate at the inner corner of the L-shaped channel, where structural current crowding maximizes the local VDF.

To resolve the origin of the noise tail at the level of individual pixels, we performed a statistical “pixel-hunting” simulation. Fig. 5 presents representative hunted pixels before passivation. A normal center-peak pixel (top row) lacks any high-sensitivity trap and exhibits stable noise. In contrast, a defective tail pixel (bottom row) generates extreme >30 LSB noise deterministically: such a pixel arises only when a trap is simultaneously located at the inner corner (large VDF, spatial condition) and possesses a relaxation time commensurate with the CDS window (resonance, temporal-energetic condition). The tail is therefore not a matter of trap amplitude alone, but of the coincidence of these two necessary conditions.

These two conditions also explain why the F-IIP passivation is so effective. Our atomistic calculations show that fluorine passivates the corner-related dangling-bond defects, thereby

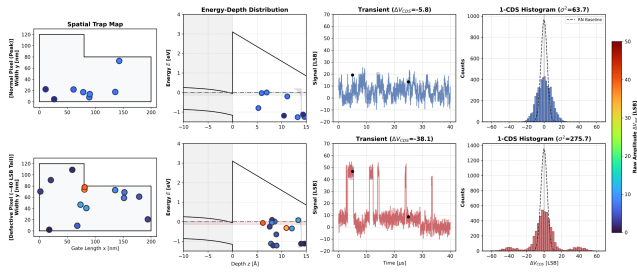


Fig. 5. Detailed anatomy of hunted representative pixels before the F-IIP passivation process. (Top row) A normal center-peak pixel exhibiting stable noise. (Bottom row) A defective tail pixel showing extreme RTN fluctuations (>30 LSB). Each case displays, from left to right, the spatial trap distribution, the atomistic energy-depth band diagram, the transient RTN waveform, and the resulting single-CDS histogram.

lowering their density and shifting the remaining trap levels away from the resonance energy [3]. Consequently, the two necessary ingredients of a tail pixel—a large-VDF corner trap *and* a resonance-matched relaxation time—are rarely satisfied together, so the spatial-energetic coincidence is statistically eliminated and former tail pixels revert to the center peak. This atomistic picture is fully consistent with the after-passivation distributions of Fig. 3.

A subtlety specific to the device is worth noting. Although F-IIP lowers the Pb-center density across the entire gap, in the present n-channel SF the surface Fermi level sits near the conduction band, so the dominant lower-gap (hole) traps remain filled and electrically static and do not contribute to the RTN; it is the upper-gap (electron) traps near the Fermi level that drive the switching. Because the Pb center is amphoteric, passivating a single dangling bond removes its donor and acceptor levels together, so the large reduction observed in the hole-side trap density and the suppression of the RTN-active electron-side traps are two manifestations of the same passivation event. The RTN impact of the lower-gap traps would instead be assessed in a p-channel device, where the surface Fermi level approaches the valence band.

IV. CONCLUSION

We developed an ultrafast multi-scale statistical framework linking atomistic trap properties to macroscopic pixel noise. We identified that the severe RTN tail in sub-micrometer CIS originates from the overlap of process-induced defect clustering at the L-shaped corner with the energetic CDS resonance, and verified that F-IIP passivation suppresses this noise by statistically removing the resonant corner traps, in agreement with silicon. The approach is a robust tool for optimizing future low-noise CIS architectures.

REFERENCES

[1] Y. Higashi, N. Momo, H. Sasaki, H. S. Momose, T. Ohguro, and K. Matsuzawa, “Comprehensive Understanding of Random Telegraph Noise with Physics Based Simulation,” *Symp. VLSI Technology Dig. Tech. Papers*, 2011, p. 200.

[2] Jun-Myung Woo, Hong-Hyun Park, Sung-Min Hong, In-Young Chung, Hong Shick Min, and Young June Park, “Statistical noise analysis of CMOS image sensors using the Green’s function method,” *IEEE Trans. Electron Devices*, vol. 56, no. 11, pp. 2481–2488, Nov. 2009.

[3] Jae Ho Kim, Sungchul Kim, Hong-Lae Park, Gijae Kang, Seonghoon Ko, Jongsu Yoon, Wook Lee, Yonghee Park, Jonghyun Go, HyunChul Kim, and Dae Sin Kim, “Atomistic analysis on random telegraph noise of SF transistors in sub-micrometer CIS pixels,” in *Proc. Int. Conf. Simulation of Semiconductor Processes and Devices (SISPAD)*, 2024.

[4] J. P. Perdew, K. Burke, and M. Ernzerhof, “Generalized gradient approximation made simple,” *Phys. Rev. Lett.*, vol. 77, no. 18, pp. 3865–3868, 1996.

[5] M. J. Kirtan and M. J. Uren, “Noise in solid-state microstructures: A new perspective on individual defects, interface states and low-frequency ($1/f$) noise,” *Adv. Phys.*, vol. 38, no. 4, pp. 367–468, 1989.

[6] H. Tian, B. Fowler, and A. El Gamal, “Analysis of temporal noise in CMOS photodiode active pixel sensor,” *IEEE J. Solid-State Circuits*, vol. 36, no. 1, pp. 92–101, Jan. 2001.

[7] S. Machlup, “Noise in semiconductors: Spectrum of a two-parameter random signal,” *J. Appl. Phys.*, vol. 25, no. 3, pp. 341–343, 1954.

[8] E. H. Poindexter, P. J. Caplan, B. E. Deal, and R. R. Razouk, “Interface states and electron spin resonance centers in thermally oxidized (111) and (100) silicon wafers,” *J. Appl. Phys.*, vol. 52, no. 2, pp. 879–884, 1981.

[9] P. M. Lenahan and J. F. Conley, Jr., “What can electron paramagnetic resonance tell us about the Si/SiO₂ system?” *J. Vac. Sci. Technol. B*, vol. 16, no. 4, p. 2134, 1998.