Supplementary Information

A stochastic encoder using point defects in two-dimensional materials

Harikrishnan Ravichandran[1], Theresia Knobloch[2], Shiva Subbulakshmi Radhakrishnan[1], Christoph Wilhelmer[2], Sergei P Stepanoff[3], Bernhard Stampfer[2], Subir Ghosh[1], Aaryan Oberoi[1], Dominic Waldhoer[2], Chen Chen[4], Joan M Redwing[3,4,5], Douglas E Wolfe[1,3,5], Tibor Grasser[2], & Saptarshi Das[1,3,5,6]*

[1]*Engineering Science and Mechanics, Penn State University, University Park, PA 16802, USA*

[2]*Institute for Microelectronics (TU Wien), Gusshausstrasse 27–29, 1040 Vienna, Austria*

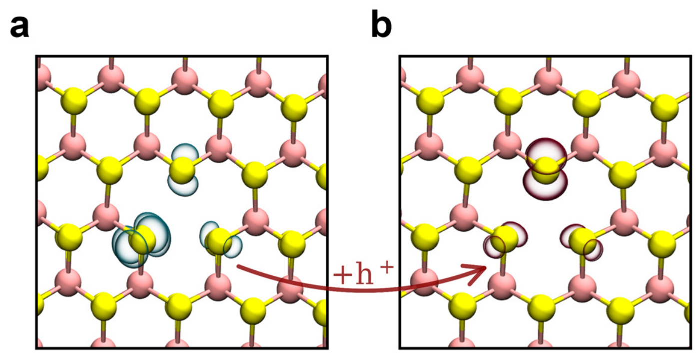
[3]*Materials Science and Engineering, Penn State University, University Park, PA 16802, USA*

[4]*2D Crystal Consortium Materials Innovation Platform, Penn State University, University Park, PA 16802, USA*

[5]*Materials Research Institute, Penn State University, University Park, PA 16802, USA*

[6]*Electrical Engineering, Penn State University, University Park, PA 16802, USA*

Supplementary Figure 1

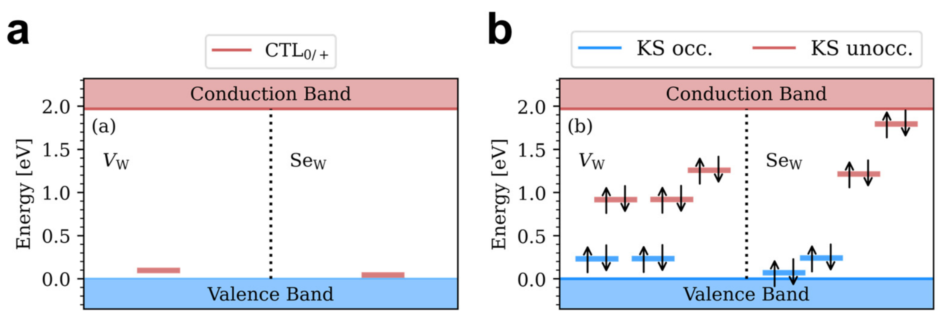


Supplementary Figure 1. DFT-optimized atomic structure of defects in monolayer WSe_2. *Atomic structure of a tungsten vacancy in a WSe_2 monolayer geometry optimized with DFT.* ***a)*** *Neutral charge state with the HOMO (blue bubbles) localized at the Se sites surrounding the vacancy*. ***b)*** *Positive charge state with the hole (LUMO, red bubbles) localized at Se sites.*

Supplementary Figure 2

Energy Levels

The energy levels of the investigated defects obtained from density functional theory (DFT) are discussed here. Both defect types, V_W and Se_W, introduce unoccupied and occupied Kohn-Sham (KS) states in the WSe_2 band gap which can host up to two electrons in opposite spin states. The occupied KS levels lie closely above the valence band maximum (VBM), indicating that these defects can potentially be classified as shallow hole traps and might therefore be responsible for generating the RTN signal measured in our devices. To confirm this assumption, the thermodynamic charge transition levels (CTLs) of both defect types are calculated for single hole trapping. Compared to the KS levels of a defect, which are directly obtained from DFT calculations of a fixed atomic configuration, the CTL also accounts for the



Supplementary Figure 2. Energy band diagrams. a) *Charge transition levels (CTLs) of the +/0 transition for a W vacancy (V_W, left) and a Se antisite (Se_W, right).* ***b)*** *Electronic Kohn-Sham (KS) levels from hybrid functional DFT in the WSe_2 band gap introduced by a V_W (left) and an Se_W (right). Occupied and empty KS levels are drawn as blue and red bars, respectively, with the two possible spin states depicted as upward and downward arrows. The CTLs and KS levels are given with respect to the valence band maximum of the pristine WSe_2 monolayer.*

energy change of the system upon charge trapping due to atomic relaxations and can thus be related to experimentally detected trap levels. The KS states and CTLs of both defect types are shown in the context of the electronic band gap of a WSe_2 monolayer in **Supplementary Figure 2**.

The CTL is calculated by comparing the formation energies of the defect in the positive and neutral charge states [1] and corresponds to the intersection point of two formation energy functions depending on the Fermi level. It is thus equal to the Fermi level at which the formation energies of the neutral and charged systems match and thus both charge states of the defect are equally stable. The CTL for a hole trap with respect to the valence band edge of the pristine bulk is given by:

$$CTL_{0/+1} = \frac{E_0^{tot} - E_{+1}^{tot} - E_{+1}^{corr}}{+1} - \epsilon_{VBM}^{bulk} + \Delta V_{0/p}$$

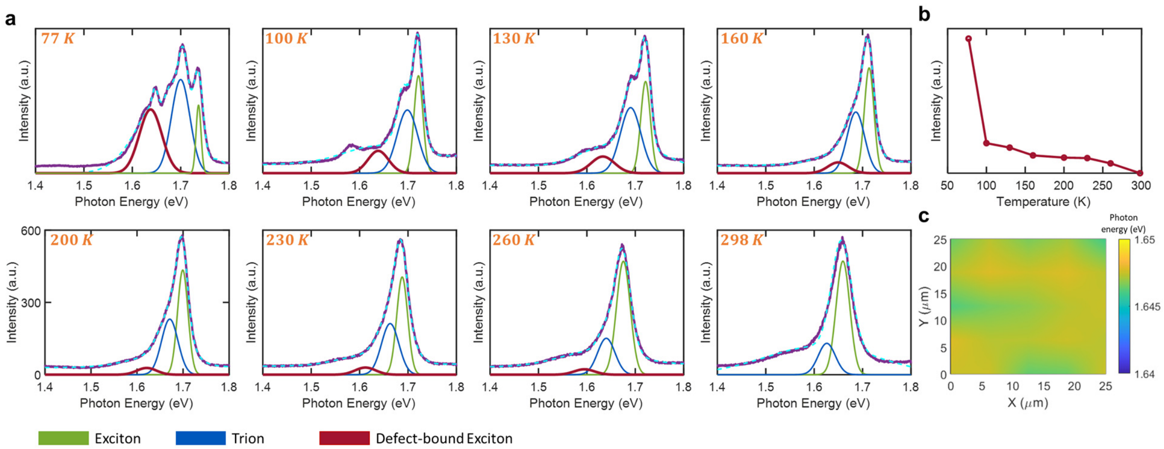
$Defect$	$CTL_{0/+1} - E_{VBM}$ (meV)	$E^{R}_{0/+}$ (meV)	$E^{R}_{+/0}$ (meV)
Se_W	45	23	23
V_W	97	196	117

Supplementary Table 1. Energy configurations of defects in monolayer WSe₂. *Energetic configurations of hole trapping* Se_W *and* V_W *defects obtained from hybrid functional DFT. Charge transition levels (CTL) are given with respect to the valence band maximum of the pristine* WSe_2 *monolayer.*

Here, E_q^{tot} is the total energy of the supercell containing the defect E in charge state q, $E_{\mathrm{q}}^{\mathrm{corr}}$is the electrostatic correction term for charged DFT calculations with periodic boundary conditions, ϵ_{VBM}^{Bulk} is the energy level of the VBM of the defect-free monolayer, and $\Delta V_{o/p}$ is a term to correct the electrostatic potential offset between the charged and uncharged systems far from the defect. Furthermore, the relaxation energies for hole capture ($E^{R}_{0/+}$) and release ($E^{R}_{+/0}$) at the defect sites are calculated. These energies correspond to the energy dissipated to the surrounding phonon bath due to structural relaxations after charge injection at the defect site. The calculated energies are presented in the **Supplementary Table 1**.

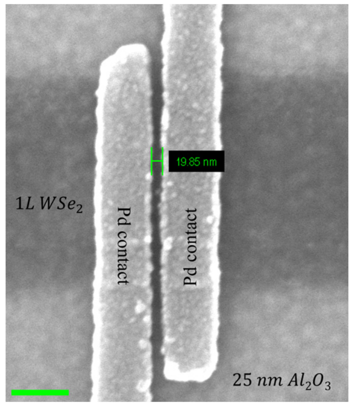
For both V_W and Se_W, the hole trap level lies closely above the VBM of pristine bulk WSe_2, with only small relaxation energies for both hole capture and release. The defect types can therefore be classified as hole traps and are probable defect candidates to explain the RTN signal measured in our 2D devices.

Supplementary Figure 3



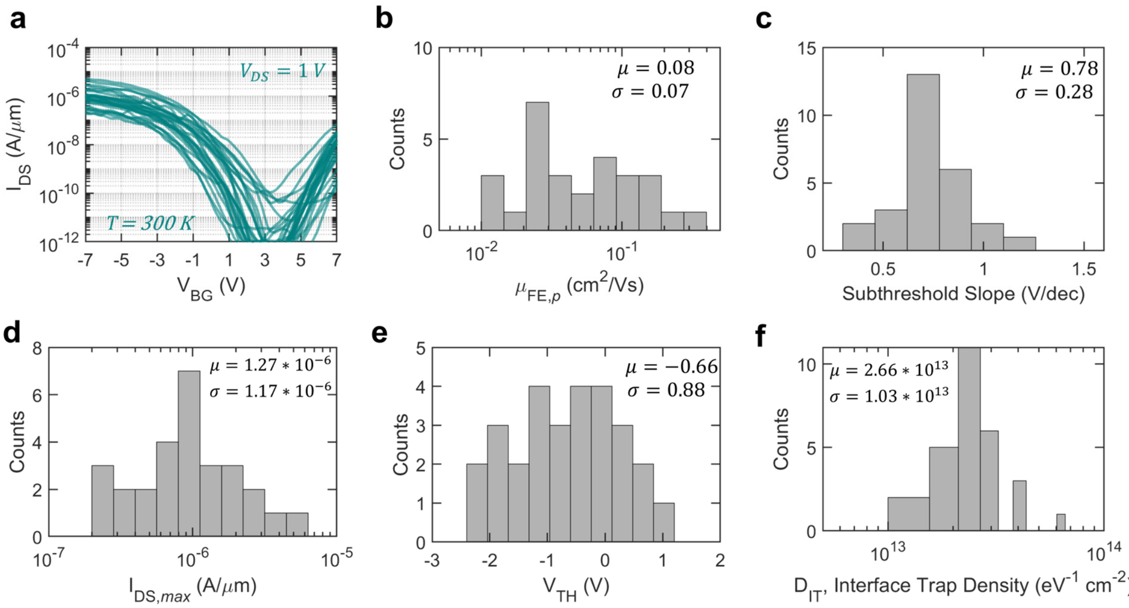
Supplementary Figure 3. Evolution of defect-bound exitonic peak in PL spectroscopy. a) *PL spectra measured at temperatures ranging from 77 K to 298 K showing the evolution of the defect-bound excitonic peak. At T =77 K, the intensity of defect-bound exciton peak is maximum. A sharp reduction in the defect-bound excitonic peak intensity is observed with increase in temperature.* ***b)*** *Plot of curve-fitted intensity of excitonic peak as a function of temperature.* ***c)*** *Spatial colormap for the defect-bound excitonic peak position, measured at 77 K over a 25 μm × 25 μm area. The mean peak position was found to be at ~ 1.647 eV with a standard deviation of ~0.001 eV.*

Supplementary Figure 4



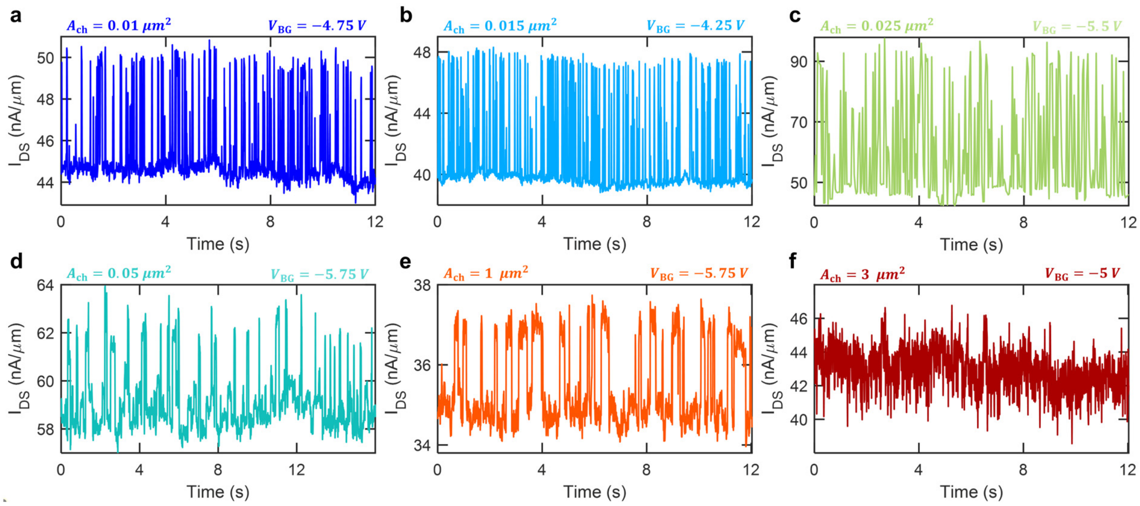
Supplementary Figure 4. Scanning electron microscopy analysis. *Zoomed-in SEM image confirming* L_{ch} *= 19.8* $\pm$ *0.05 nm. Scale bar, 100nm.*

Supplementary Figure 5



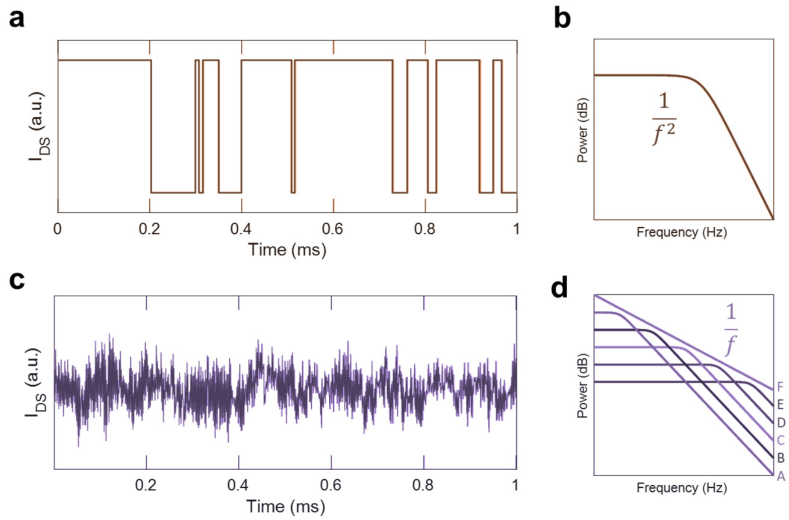
***Supplementary Figure 5. Electrical characterization of WSe_2 FETs**. Device statistics obtained from 30 WSe_2 FETs with L_{ch}= 20 nm and W_{ch}= 500 nm. showing **a)** the transfer characteristics and corresponding distributions for **b)** field-effect mobility ($\mu_{FE,p}$), **c)** subthreshold slope, **d)** maximum ON-current ($I_{DS,max}$), **e)** threshold voltage (V_{TH}), and **f)** interface trap density, respectively.*

Supplementary Figure 6



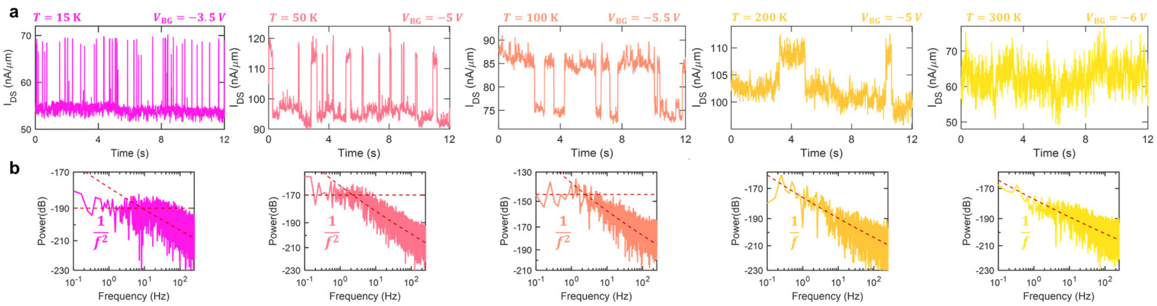
Supplementary Figure 6. Impact of channel area on RTNs. *RTN traces obtained at T =15 K from WSe_2 FETs with channel areas of* ***a)*** $A_{\text{ch}} = 0.01\ \mu\text{m}^2$, ***b)*** $A_{\text{ch}} = 0.015\ \mu\text{m}^2$, ***c)*** $A_{\text{ch}} = 0.025\ \mu\text{m}^2$, ***d)*** $A_{\text{ch}} = 0.05\ \mu\text{m}^2$, ***e)*** $A_{\text{ch}} = 1\ \mu\text{m}^2$, ***f)*** $A_{\text{ch}} = 3\ \mu\text{m}^2$. *With increasing channel area, the discrete nature of the RTN decreases due to the superimposition of RTNs stemming from multiple defects.*

Supplementary Figure 7



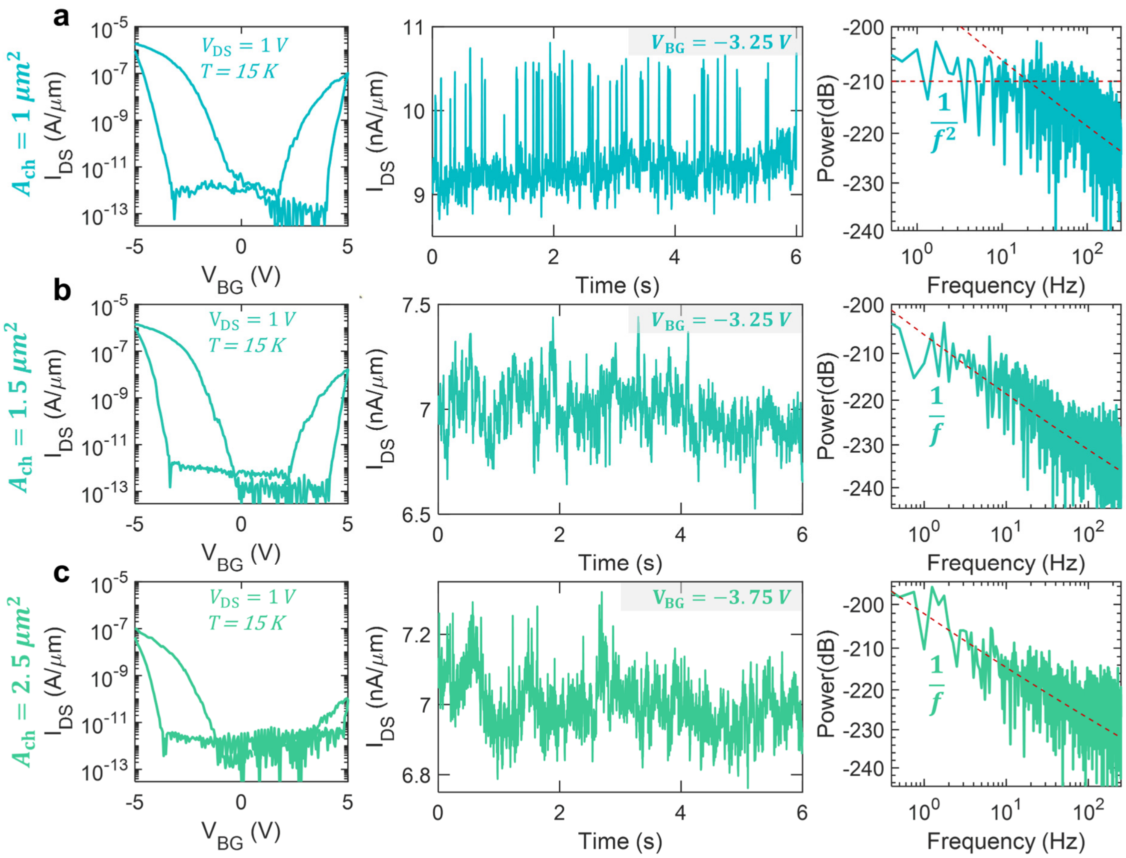
Supplementary Figure 7. Simulation of 1/f and 1/f² noise spectra. *Simulation results showing* ***a)*** *plot of time series data of I_{DS} with discrete "digital" fluctuations that can be observed as RTN traces.* ***b)*** *Corresponding power spectral density (PSD) of the time series data shown in* ***(a)****. The PSD characteristics follow a Lorentzian profile (slope =$1/f^2$) indicating the presence of one or few defects with a distinct characteristic time constant.* ***c)*** *Plot of time series data (I_{DS}) with analog fluctuations resulting from the superposition of several digital RTN.* ***d)*** *Corresponding PSD of the time series data shown in* ***(b)*** *following a $1/f$ noise profile as shown by curve F. The curves denoted A-E are indicative of individual RTN with a different characteristic time constant and curve F is obtained by the summation of the different RTN traces. Curve F represents the universally observed $1/f$ flicker noise spectra.*

Supplementary Figure 8



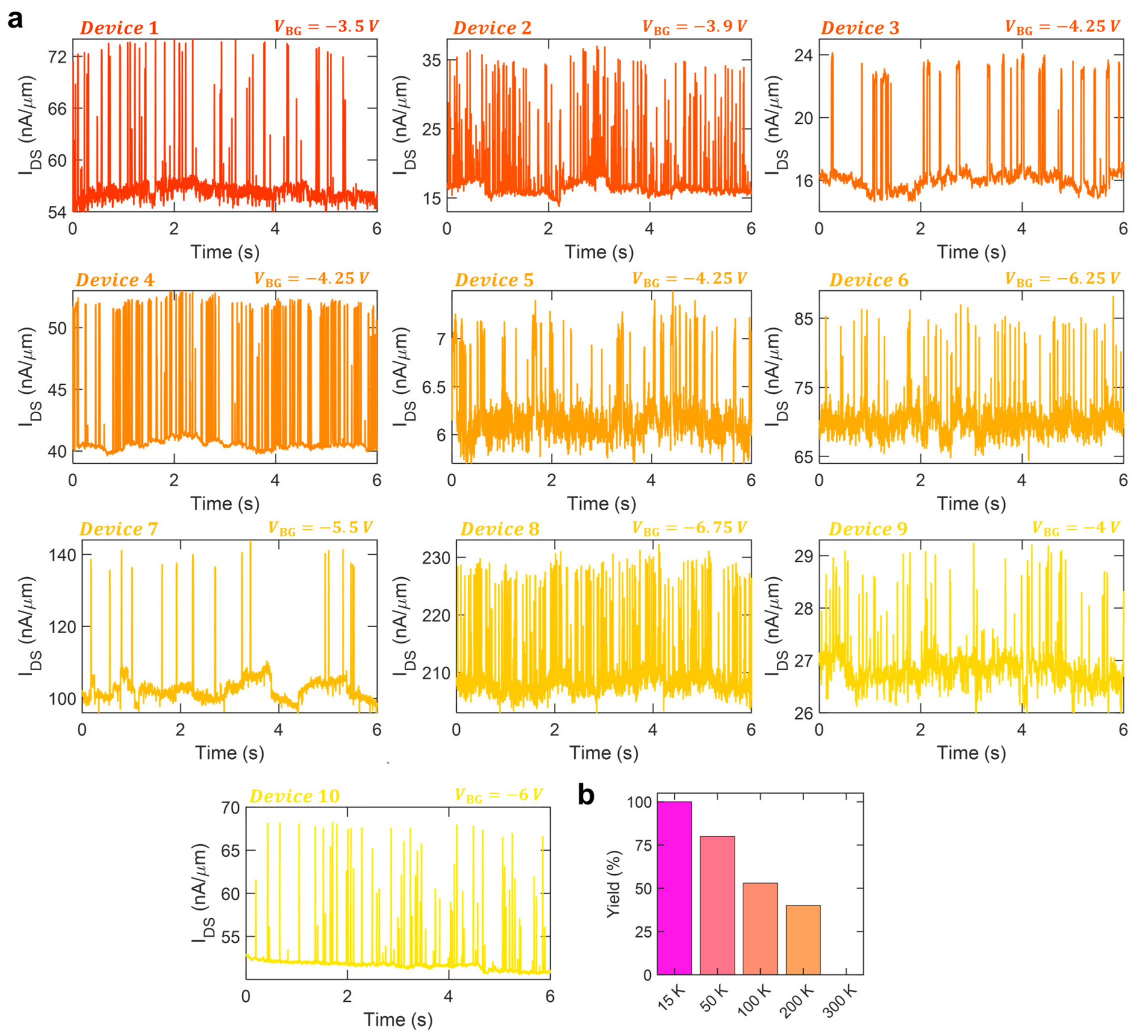
Supplementary Figure 8. Impact of temperature on RTN. a) *RTN traces obtained from an ultra-scaled WSe_2 FET with a channel area of $A_{ch} = 0.01\ \mu m^2$ using V_{DS} = 1 V at different temperatures, T = 15, 50, 100, 200, and 300 K.* ***b)*** *PSD plots of the corresponding RTN traces obtained in* ***(a)****. With increasing temperature, the short-lived, spike-like nature of the two state RTN transitions into a more rectangular one with increased lifetime at each state and the RTN traces completely vanish at $T > 200$ K.*

Supplementary Figure 9



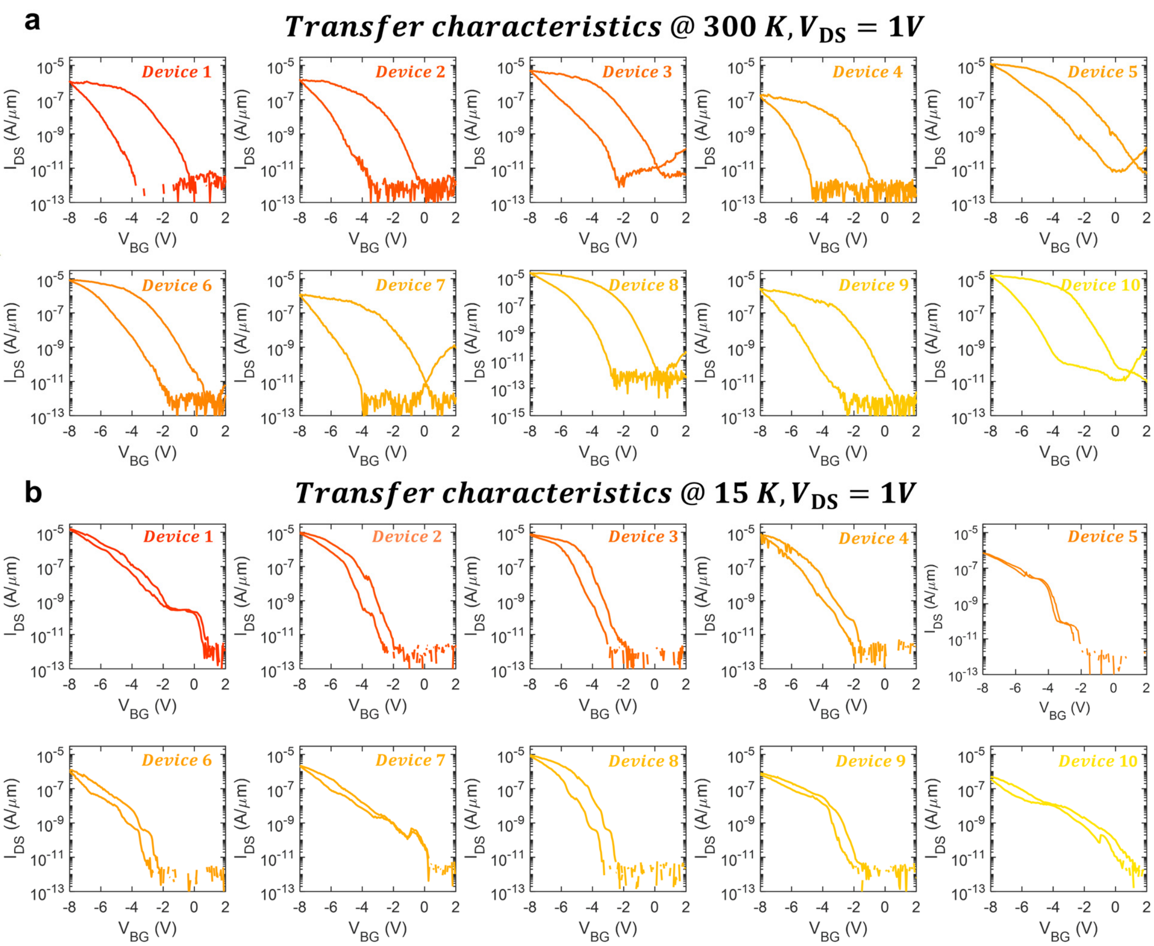
Supplementary Figure 9. Impact of change in dielectric on RTN. *Transfer characteristics, RTN traces, and PSD plots obtained from WSe_2 FETs fabricated on 20 nm ALD-grown hafnium zirconium oxide (HZO) with Pd/Au contacts and channel areas of* ***a)*** $A_{ch} = 1\ \mu m^2$, ***b)*** $A_{ch} = 1.5\ \mu m^2$, *and* ***c)*** $A_{ch} = 2.5\ \mu m^2$, *each measured at V_{DS} = 1 V and T = 15 K. RTN is present in devices where* $A_{ch} = 1\ \mu m^2$.

Supplementary Figure 10



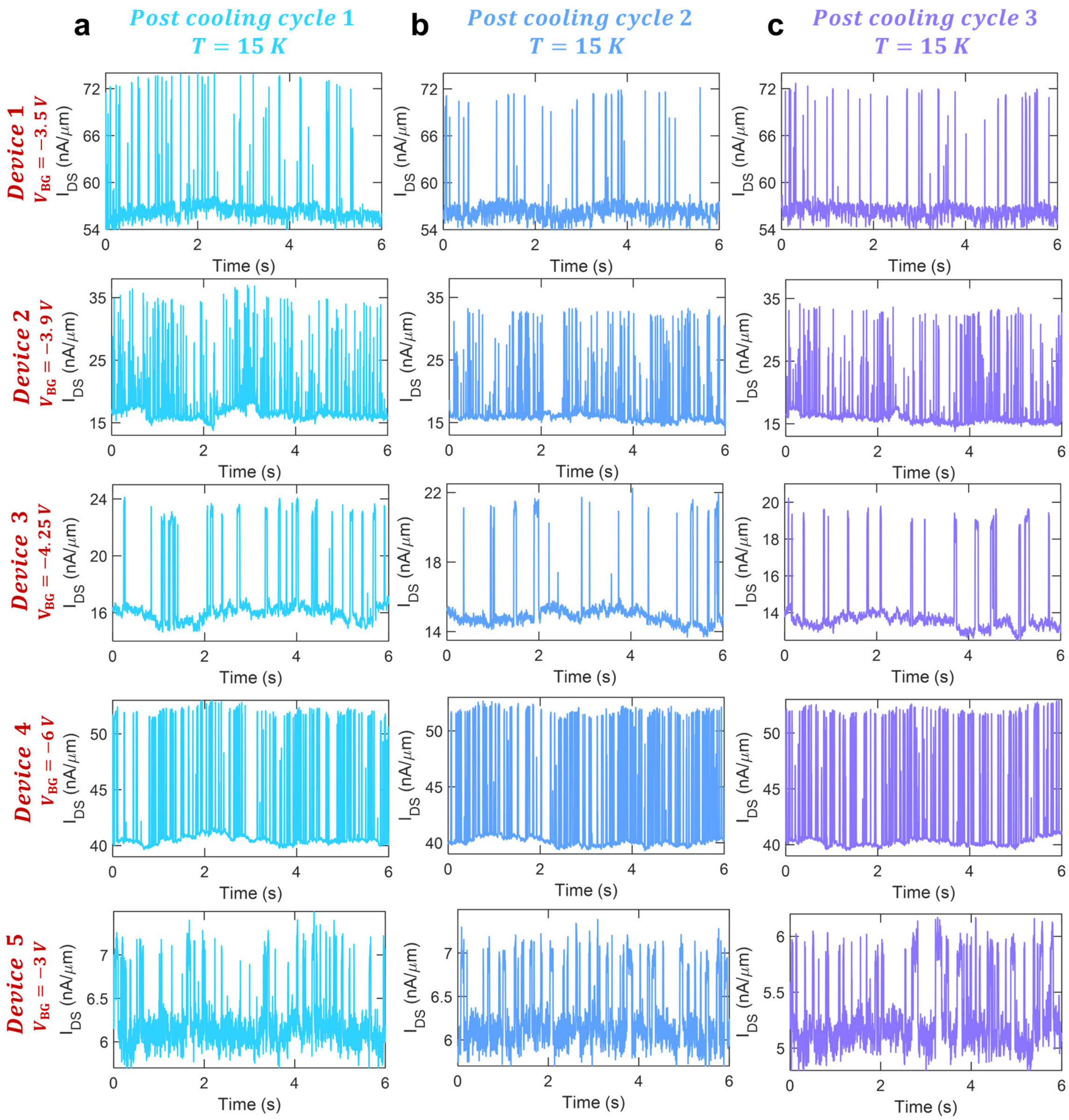
Supplementary Figure 10. RTN obtained from multiple devices. a) *RTN traces obtained from 10 ultra-scaled WSe_2 FETs with an A_{ch} of 0.01 μm^2 when measured at 15 K.* ***b)*** *Plot of % yield of RTN as a function of temperature for the same 10 ultra-scaled WSe_2 FETs measured at T = 15 K, 50 K, 100 K, 200 K and 300 K. As the temperature increases, the RTN yield steadily declines, transitioning from 100% at 15 K to 0% at 300 K.*

Supplementary Figure 11



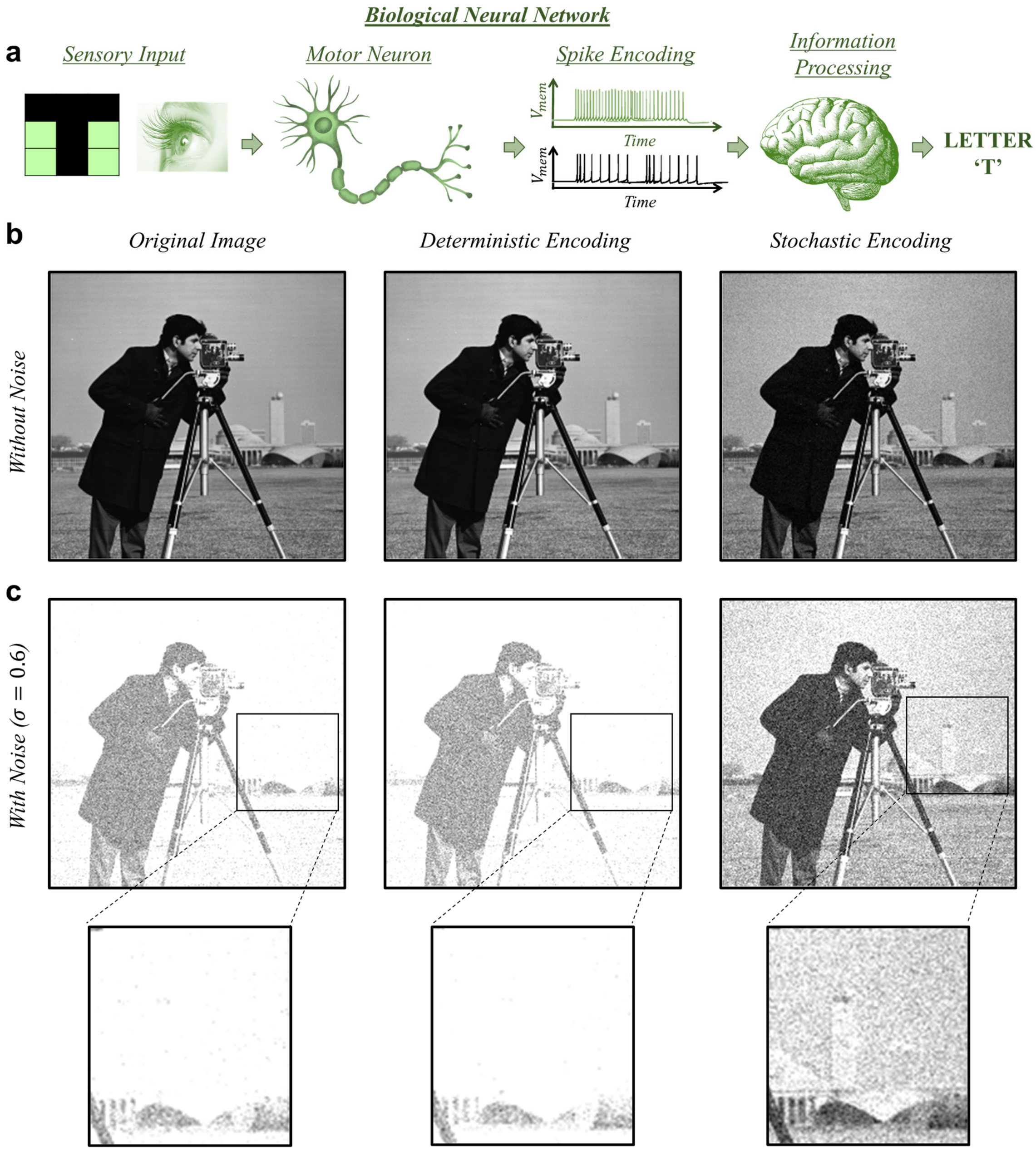
Supplementary Figure 11. Transfer characteristics of multiple devices yielding RTN. *Transfer characteristics of 10 ultra-scaled WSe_2 FETs having A_{ch} of 0.01* $\mu\mathrm{m}^2$ *measured at* ***a)*** *300 K and* ***b)*** *15 K with $V_{\mathrm{DS}} = 1V$.*

Supplementary Figure 12



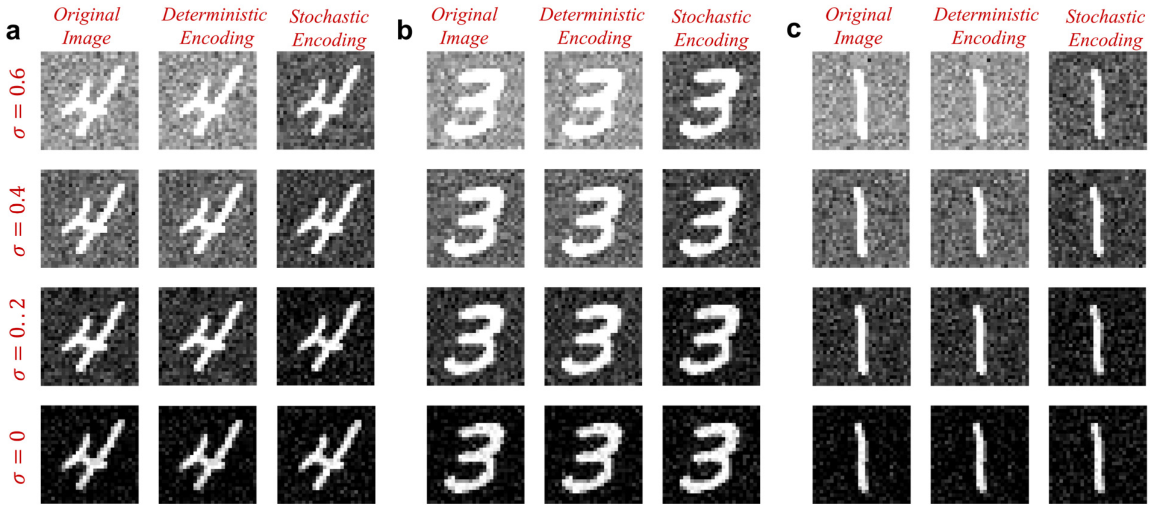
Supplementary Figure 12. Reproducibility of RTN. *RTN traces obtained from five WSe_2 FETs measured corresponding to **a)** post cooling cycle 1, **b)** post cooling cycle 2 and **c)** post cooling cycle 3. Each cooling cycle consists of measurements of RTN at 15 K followed by increasing the temperature to 300 K and again cooling down to 15 K. The RTN traces remain unaffected by the cooling cycles, indicating the reproducibility of RTN generation at 15 K.*

Supplementary Figure 13



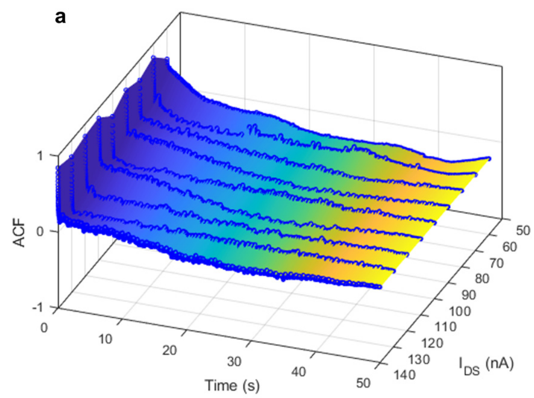
Supplementary Figure 13. Realization of biomimetic afferent neuron. a) *Schematic representation of input stimuli (images) being translated by visual afferent neurons into stochastic spike trains. The stochastic spike train with τ_{spike} represents the corresponding intensity of light in each pixel, which is subsequently processed by the BNN and inferred. Example of deterministic versus stochastic encoding of the "cameraman" image* ***b)*** *in the absence of noise and* ***c)*** *in presence of noise (σ = 0.6), respectively. Compared to deterministic encoding, stochastic encoding can resolve finer features in the image even in the presence of noise, as depicted in the inset plots for the figures in the lower panel. Image Credits: Eye - © Vladimir Voronin / Stock.adobe.com; Neuron - © Dyah / Stock.adobe.com; Brain - © jolygon / Stock.adobe.com; Cameraman image - "cameraman.tif" from MATLAB's Image Processing Toolbox, © Massachusetts Institute of Technology (MIT).*

Supplementary Figure 14



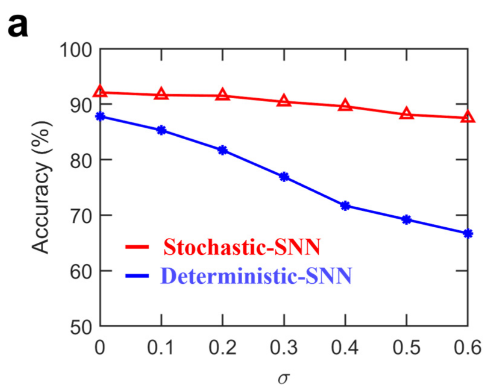
Supplementary Figure 14. Deterministic versus stochastic encoding. *Example of deterministic versus stochastic encoding of three representative MNIST handwritten digits, **a)** digit 4, **b)** digit 3, and **c)** digit 1, in the absence of noise and in the presence of varying standard deviation of noise (σ = 0.2, 0.4 and 0.6). Compared to deterministic encoding, stochastic encoding can resolve finer features in the image even in the presence of noise.*

Supplementary Figure 15



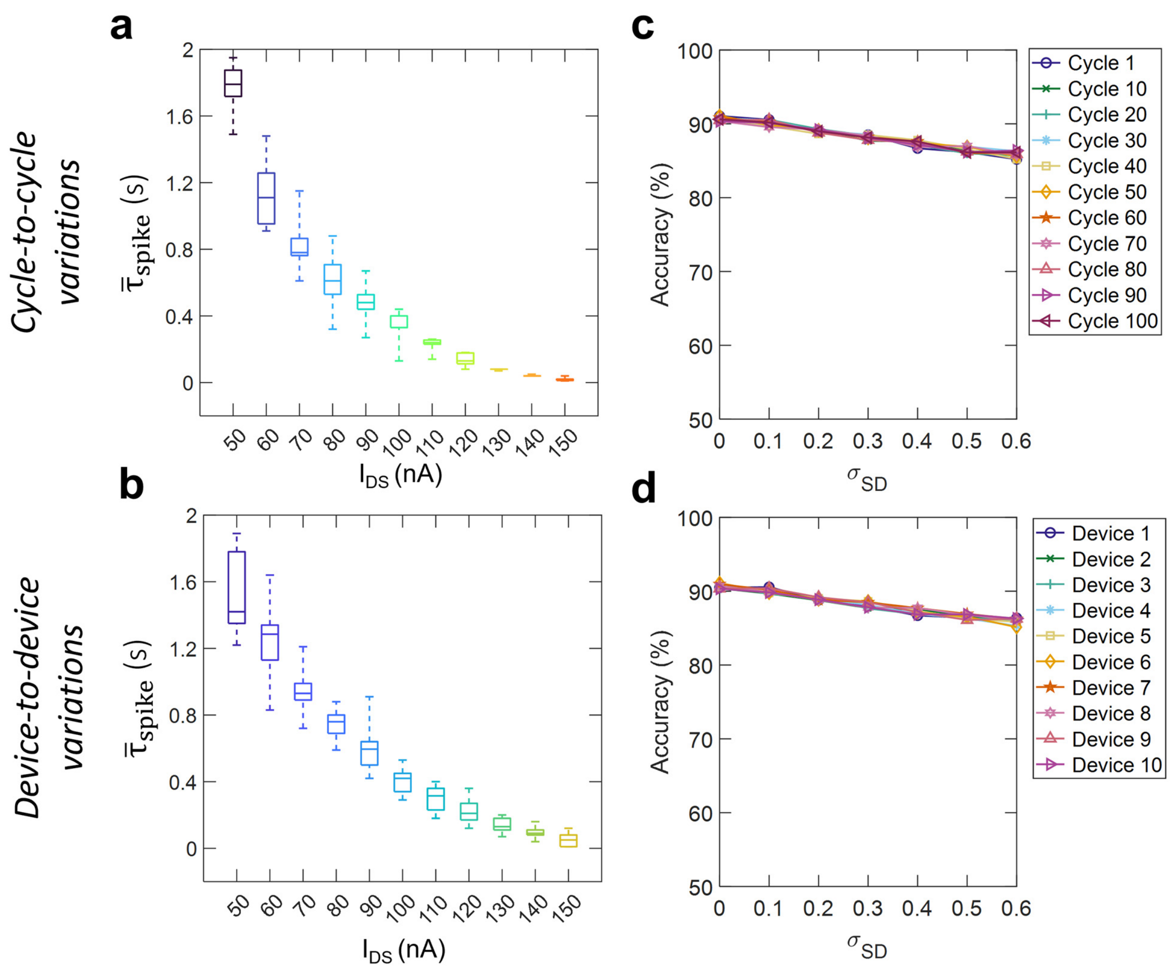
Supplementary Figure 15. Autocorrelation of stochastic voltage spikes. a) *Autocorrelation function (ACF) of the stochastic voltage spike trains as a function of* I_{DS} *and time-lag. The ACF plots show a near ideal value close to 0, confirming the truly random nature of the stochastic spike trains.*

Supplementary Figure 16



Supplementary Figure 16. Noise-resilient inference of handwritten digit MNIST dataset using stochastic SNN. *The MNIST handwritten digit dataset consists of 70,000 images, each 28×28-pixel, where 60,000 images correspond to training and the remaining 10,000 correspond to testing. A fully-connected artificial neural network (ANN) with 784, 100 and 10 neurons in the input, hidden and output layers, respectively, was trained on the MNIST dataset, achieving the total accuracy of 95 % for analog inputs over 150 epochs. The ANN was subsequently converted to a spiking neural network (SNN) to test inference accuracy using the 10,000 test images from the MNIST dataset. We obtained a testing accuracy of ~92%. We found that the inference accuracy only dropped from ~92% to 87% for stochastic spike encoding, whereas the inference accuracy dropped from 88% to 67% for deterministic encoding.*

Supplementary Figure 17



Supplementary Figure 17. Impact of cycle-to-cycle and device-to-device variation on inference accuracy of medical MNIST images. *Boxplots depicting the minimum, 25th percentile, median, 75th percentile and maximum values of* $\bar{\tau}_{\text{spike}}$ *as a function of the input* (I_{DS}) *for* ***a)*** *a single device over 100 cycles and* ***b)*** *ten devices over one cycle. The center line within the box represents the median value, with the bottom and top bounds of the box delineating the 25th and 75th percentiles, respectively. Whiskers represent the minimum and maximum values. The relationship between* $\bar{\tau}_{\text{spike}}$ *and* I_{DS} *is quantified using a fitting based on the empirical model* $\bar{\tau}_{\text{spike}} = \tau_0 \times exp\left(-\frac{I_0}{I_{DS}}\right)$. *Corresponding inference accuracy as a function of standard deviation* (σ_{SD}) *of gaussian noise added to medical MNIST images followed by stochastic encoding for* ***c)*** *cycle-to-cycle variations and* ***d)*** *device-to-device variations. The accuracy remains agnostic to cycle-to-cycle and device-to-device variation, indicating the robustness of the stochastic SNN for noise-tolerant inference applications.*

Supplementary References

[1] H.-P. Komsa, T. T. Rantala, and A. Pasquarello, "Finite-size supercell correction schemes for charged defect calculations," *Physical Review B,* vol. 86, no. 4, p. 045112, 07/13/2012, doi: 10.1103/PhysRevB.86.045112.