nature electronics

Article https://doi.org/10.1038/s41928-023-01042-7

Reconfigurable mixed-kernel heterojunction transistors for personalized support vector machine classification

In the format provided by the authors and unedited

Supplementary Information

Reconfigurable Mixed-Kernel Heterojunction Transistors
for Personalized Support Vector Machine Classification

Table of Contents

Figure S1. SVM kernel classification

Figure S2. Gausssian standard deviation tuning

Table T1. Fit parameters extracted from Figure S2.

Figure S3. Effect of CNT density on Gaussian operation

Figure S4. Effect of CNT density on sigmoid operation

Figure S5. Effect of overlap on sigmoid operation

Figure S6. Parameter fitting

Figure S7. Sigmoid operation

Figure S8. Effect of Gaussian skew on SVM classification

Figure S9. MKH classification circuit

Table T2. Comparison of device performance with literature precedents.

Table T3. Performance of linear kernel SVM

Supplementary Note 1. Origin of sigmoid behavior

- Figure S10. MKH overlap diagram
- Figure S11. MKJ equivalent circuit
- Figure S12. Model device structure
- Figure S13. Effect of overlap on individual channel transfer characteristics
- Figure S14. Effect of overlap on heterojunction transfer characteristics

Supplementary Note 2. Electrical characterization parameters

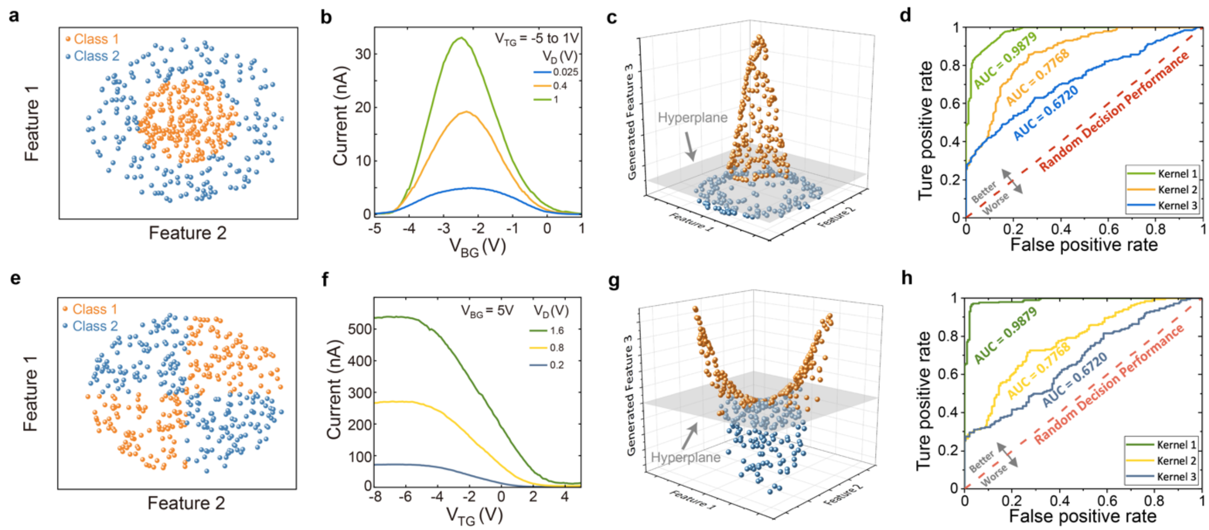

Figure S1. (a) Two classes of data are randomly generated with a concentric distribution shape in feature space. (b) Three Gaussian kernel functions generated by the mixed-kernel heterojunction (MKH) transistor. (c) A Gaussian kernel is suitable to classify the data from (a) by utilizing the geometric discrepancies around the center and tail areas in a Gaussian curve. Following application of an optimized Gaussian kernel, the data can be classified linearly with a hyperplane in the resulting three-dimensional Hilbert feature space. (d) A receiver operating characteristic (ROC) plot illustrates the diagnostic ability of the binary support vector machine (SVM) classifier upon application of the different Gaussian kernels in (b). A curve higher above the random decision performance line with a larger area-under-the-curve (AUC) value represents better classification performance. (e) Two classes of data are randomly distributed in a diagonal way, where one class mainly occupies the first and third quadrants and the other class mainly occupies the second and fourth quadrants. (f) Three sigmoid kernel functions generated by the MKH transistor. (g) A sigmoid kernel is suitable for classifying the data from (e). (h) ROC plot of SVM classification with the different sigmoid kernels from (f).

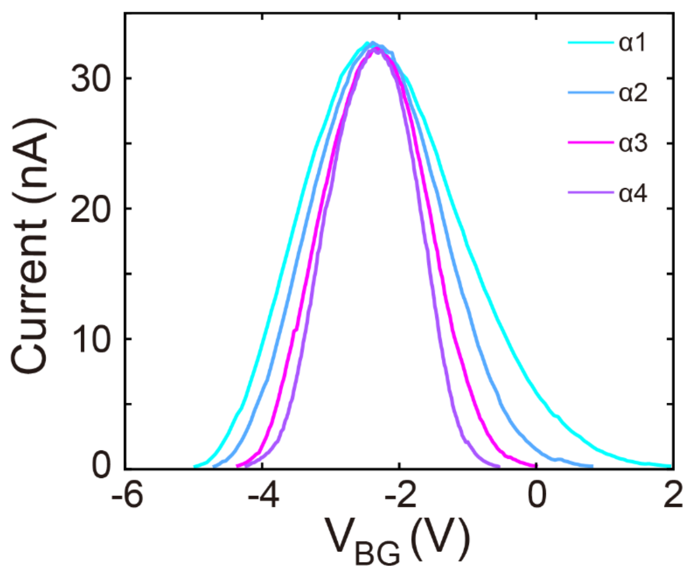

Figure S2. Gaussian curves showing standard deviation (σ) tuning with a fixed mean (μ). Here, α1, α2, α3, α4 correspond to specific biasing conditions where the differential (relative) gate sweep rates (dV_{BG}/dV_{TG}) are used to tune the curve width/height ratio (like in Fig. 1g), the top gate and bottom gate voltage offset (V_{TG}-V_{BG} (@V_{TG} = -5 V)) is used to shift the peak position (like in Fig. 1e), and the drain voltage is used to normalize the peak height (like in Fig. 1f). Additionally, a, b, and c correspond to the best-fit Gaussian parameters in the equation $f(x) = ae^{-(x-b)^2/2c^2}$, SSE is the sum of a squared estimate of errors, R-square is the coefficient of determination, DFE is the number of data points fit, and RMSE is the root-mean-square error.

Table T1. Fit parameters extracted from Figure S2.

	α1	**α2**	**α3**	**α4**
dV_{BG}/dV_{TG}	1	0.81	0.62	0.53
V_{TG}-V_{BG} (V) @V_{TG} = -5 V	0	0.28	0.62	0.75
V_D (V)	1	1.65	2.24	2.95
a (nA), A parameter	32.5709	32.9957	32.8986	33.264
b (V), μ parameter	-2.2757	-2.3346	-2.3481	-2.3772
c (V), σ parameter	1.618	1.3173	1.0788	0.9078
SSE	147.428	37.8757	48.5067	126.8309
R-square	0.9899	0.9977	0.997	0.9938
DFE	108	117	118	141
RMSE	1.1684	0.569	0.6412	0.9484

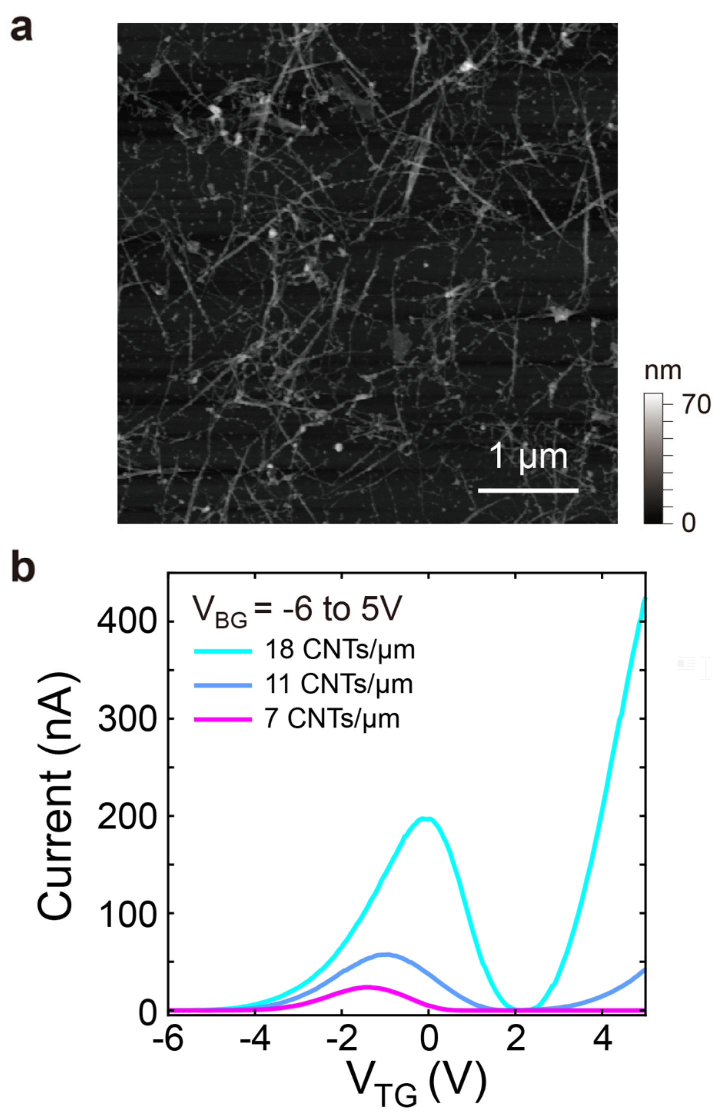

Figure S3. (a) Representative atomic force microscopy image that is used to quantify the linear carbon nanotube (CNT) density. (b) Drain current versus V_{TG} for different linear CNT densities in MKH transistors.

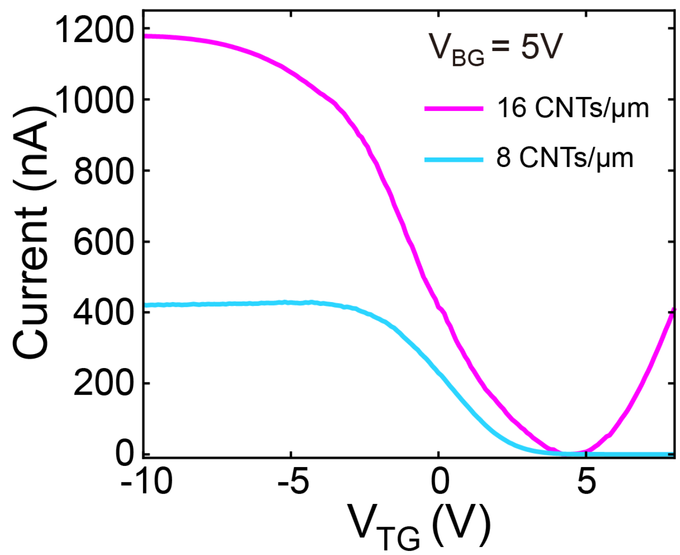

Figure S4. Drain current versus V_{TG} for different linear CNT densities in MKH transistors at a constant V_{BG}.

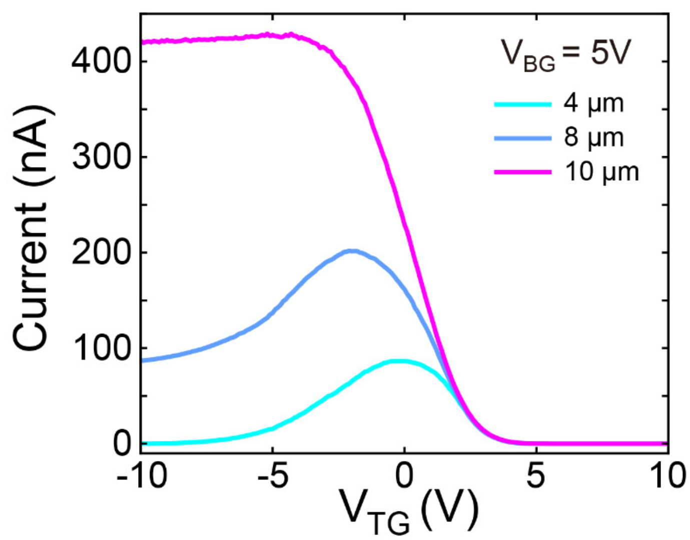

Figure S5. Drain current versus V_{TG} for different overlap region lengths in MKH transistors. The total length of the CNT and MoS_2 regions (Fig. 1c) is ~15 μm, and thus a 15 μm overlap would correspond to 100% of the channel area overlapping.

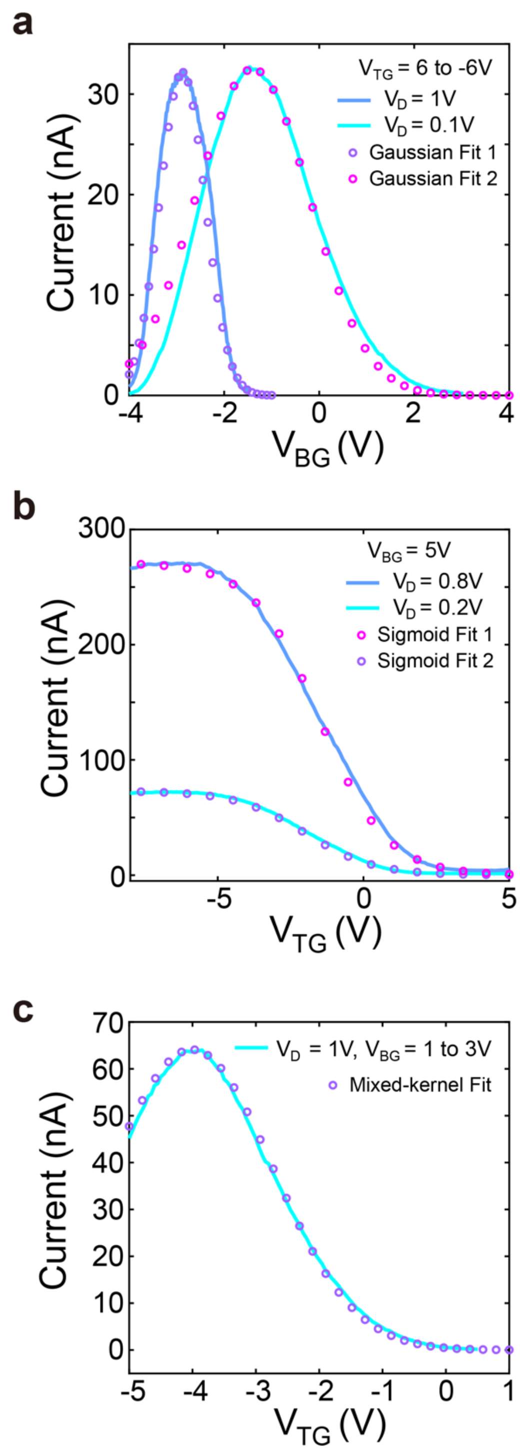

Figure S6. (a) Gaussian fitting of representative curves from Fig. 1g. (b) Sigmoid fitting of representative curves from Fig. 1h. (c) Mixed-kernel fitting of a representative curve from Fig. 1i.

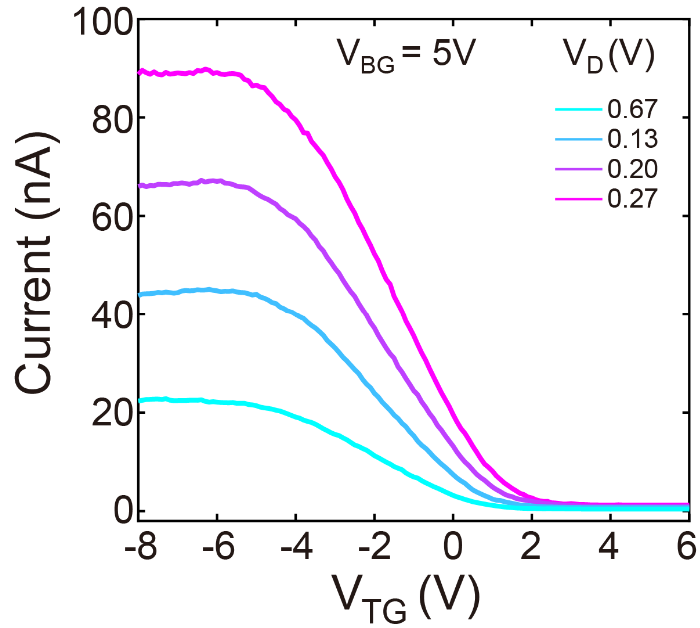

Figure S7. I_D-V_{TG} curves of the device in tunable sigmoid operation mode at lower currents. The bottom gate is held at a constant bias V_{BG} = 5 V.

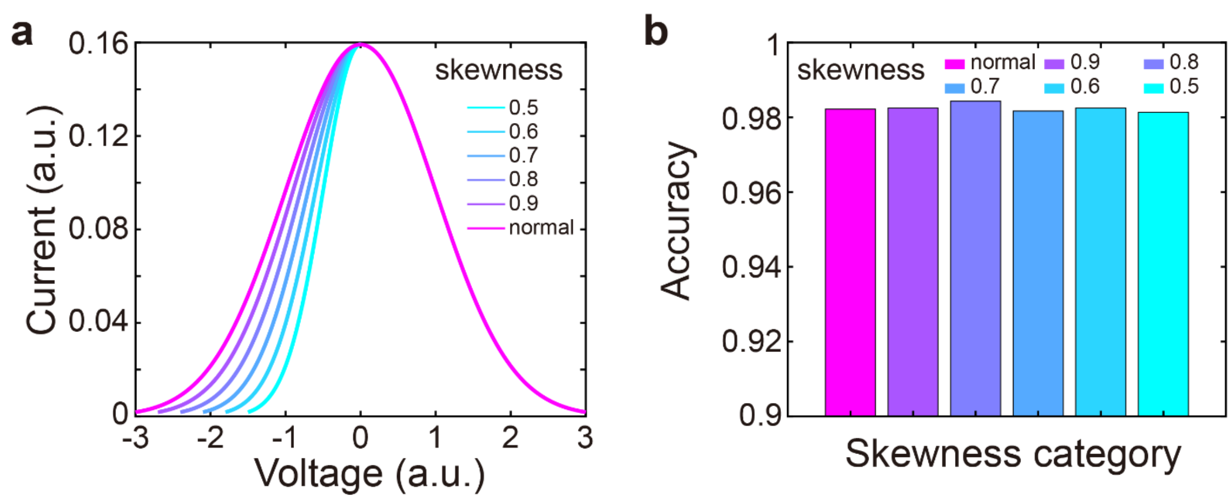

Figure S8. Effect of Gaussian skewness on SVM classification.

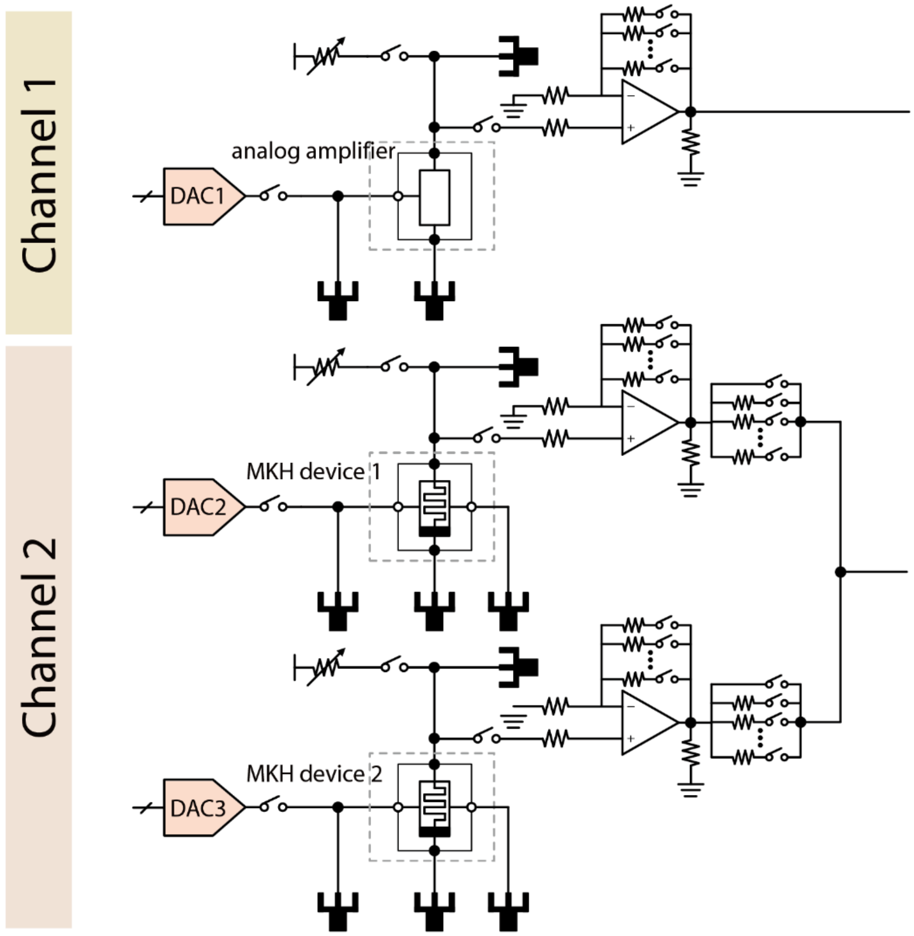

Figure S9. Schematic of the designed PCB board in Fig. 2d. Channel 1 contains an analog amplifier and an analog-to-digital converter to amplify the input signal from the ECG sensor and then send it to the SVM software. Channel 2 can contain either one or two MKH devices. Kernel functions being generated by one MKH device or each of two MKH devices are sent to the computer to be added into a mixed kernel.

Table T2. Comparison of device performance with literature precedents.

	Ref. 17	**Ref. 18**	**Current work**
Dynamic range of Gaussian mean (μ)	35 V (~3σ)	2.5 V (~2σ)	5.28 V (~5σ)
Deviation in Gaussian standard deviation (σ) while changing Gaussian mean (μ)	~50%	~25%	13.20%
Independent tunability of Gaussian mean (μ)	no	no	yes
Dynamic range of Gaussian standard deviation (σ)	10 V-15 V (50%)	1-1.25 V (25%)	0.585 V-1.22 V (109%)
Independent tunability of Gaussian standard deviation (σ)	no	no	yes
Dynamic range of peak amplitude (A)	0.3-1.5 μA	0.1-1.5 μA	0.05-0.2 μA
Independent tunability of peak amplitude	yes	yes	yes
Operating voltage (V)	50 V	6 V	6 V
Terminals	5	4	4

Table T3. Performance of linear kernel SVM

Type	**Accuracy**
Class1	0.859
Class2	0.873
Class3	0.716
Class4	0.868
Class5	0.719
average	0.807

Supplementary Note 1:

In the following section, we describe the structure and mechanisms of the MKH transistor. In the device, the heterojunction corresponds to an overlap region (Fig. S10). Because the MoS_2 in the overlap region is weakly screened by the porous CNT film from the top gate, there are two possible current paths: 'i_1' goes through the overlapping MoS_2, while 'i_2' goes through the non-overlapping MoS_2. An equivalent circuit model of the MKH transistor treats this as two MoS_2 transistors in parallel (Fig. S11), where the total conductance is given by:

$$G_{MKH} = \left(\frac{1}{G_P} + \frac{1}{G_{N,total}}\right)^{-1}$$
$$G_{N,total} = G_{N,overlap} + G_{N,non-overlap}$$

Changing the size of the overlap region changes the effective channel width of the screened (overlap) and unscreened (non-overlap) regions and their relative contribution to the conductance. To test the validity of this circuit model, we fabricated planar devices with a common terminal so that individual channels could be measured. As expected, for the purely unscreened (non-overlap) device (Fig. S12a), the MoS_2 is fully tunable, and the p-n device behavior is Gaussian (Fig. S13a). For the purely screened (overlap) device (Fig. S12b), the MoS_2 is only weakly tunable by the top gate, and the p-n device behavior shows only negative transconductance (Fig. S13b).

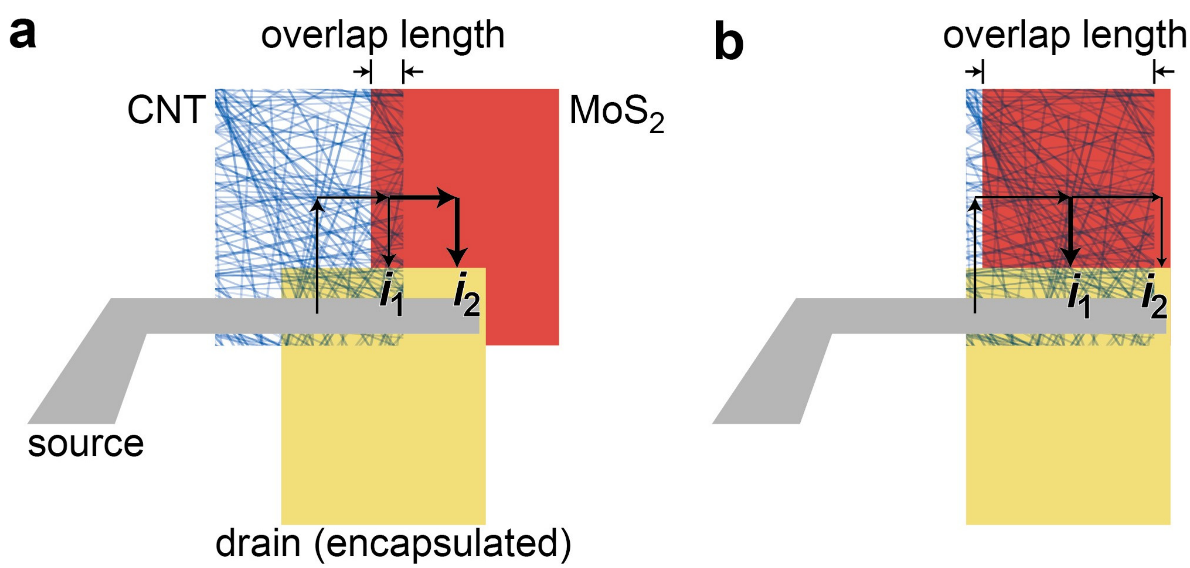

Figure S10. Top-down view of the device for (a) low overlap length and (b) high overlap length cases showing dual current paths.

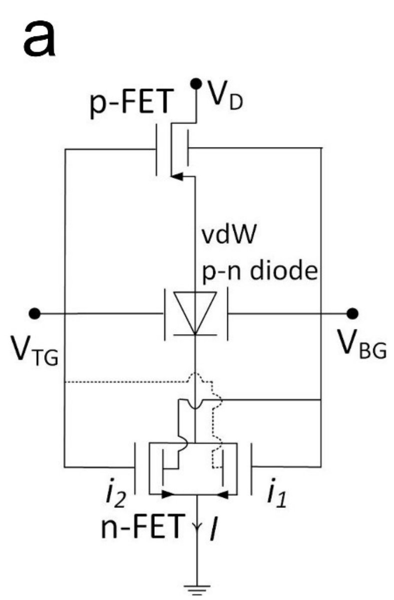

Figure S11. (a) Equivalent circuit diagram corresponding to a MKH device[1]. p-FET and n-FET correspond to the regions of individual semiconducting layers obtained by the offset between CNT and MoS_2 layers, respectively. The dashed wire shows the weakly screened gate line corresponding to the current path "i_1" in Fig. S10. (b) Increased conductance in the unscreened (non-overlap) region shown by the blue current path corresponding to Fig. S10a. (c) Increased conductance in the screened (overlap) region shown by the orange current path corresponding to Fig. S10b.

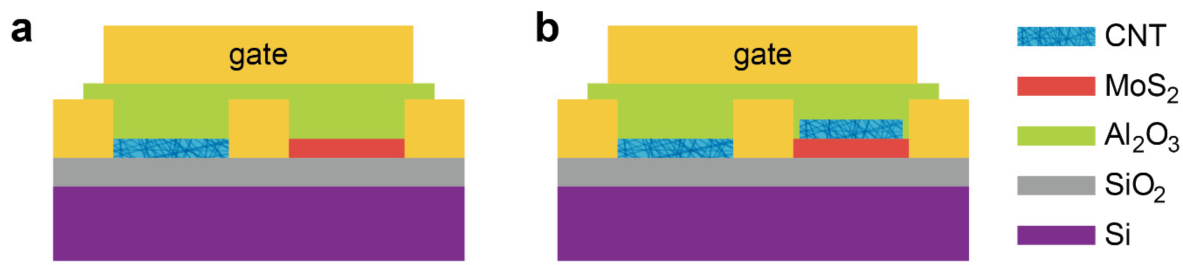

Figure S12. (a) Device structure used to model the unscreened current path. (b) Device structure used to model the screened current path. Here the CNTs were etched so they acted purely as a screening layer for the MoS_2.

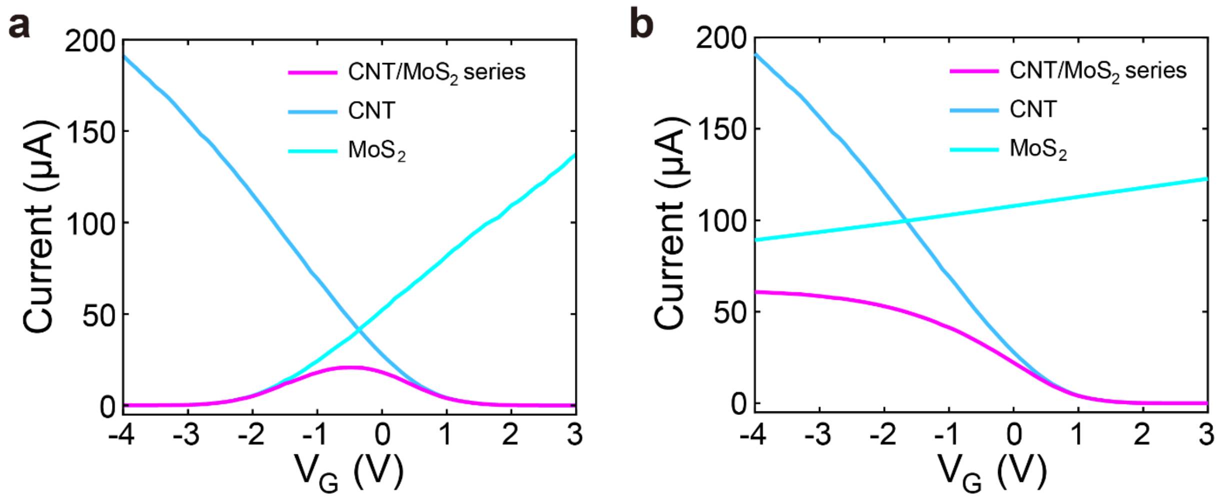

Figure S13. Measured transfer characteristics of CNT, MoS_2, and CNT/MoS_2 devices in (a) Fig. S12a and (b) Fig. S12b.

The total conductance of the MoS_2 region which is determined by the parallel unscreened and screened current paths, can be calculated from:

$$G_{N,total} = G_{N,overlap} + G_{N,non-overlap}$$

In Fig. S14, we plot calculated composite MoS_2 transfer curves, which correspond to 75% non-overlap/25% overlap (Fig. S14a) and 50% non-overlap/50% overlap cases (Fig. S14b). The total conductance of the circuit is plotted as calculated from:

$$G_{MKH} = \left(\frac{1}{G_P} + \frac{1}{G_{N,total}}\right)^{-1}$$

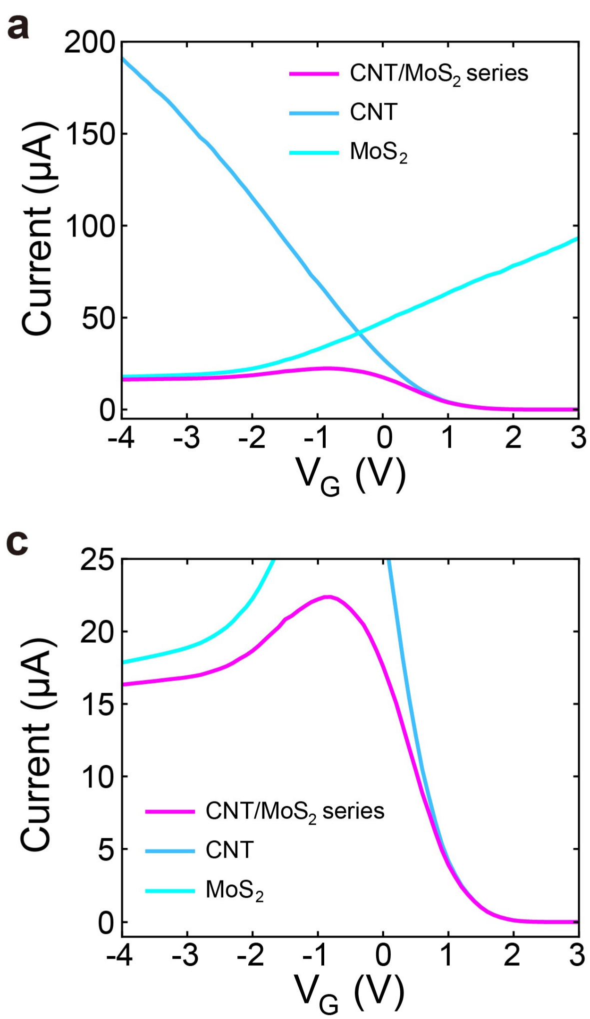

Figure S14. Transfer curves for (a) composite 75% non-overlap/25% overlap, (b) composite 50% non-overlap/50% overlap cases. (c,d) Zoomed-in plot from (a) and (b), respectively.

The 50% non-overlap/50% overlap case shows near-ideal sigmoid behavior with improved saturation in Fig. S14b,d, while decreasing the overlap to 25% (Fig. S14a,c) shows partial Gaussian behavior due to the mismatched saturation the MoS_2. This behavior reflects our experimental results for MKH devices with different levels of overlap (Fig. S4), although there are some quantitative deviations due to different device dimensions and processing variability.

Supplementary Note 2:

In the Gaussian operation mode, transfer plots (Fig. 1e-g) are plotted using the bottom gate voltage as a reference, while in the sigmoid/mixed operation mode, transfer plots (Fig. 1h-i) are plotted using the top gate voltage as a reference since the bottom gate is static. In the following discussion, we use the terms n-region and p-region to refer to the portions of the Gaussian and sigmoid curve where the conductance is dominated by n-type and p-type behavior, respectively.

Fig. 1d

Standard output curve measurement where the bottom gate and top gate are grounded, and V_D is swept from -1 V to 1 V. The p-n junction shows partial rectification, which can be tuned further using V_{TG} and V_{BG} if desired.

Fig. 1e

Transfer curves where V_{TG} is swept from -5 V to 1 V. V_{BG} is swept simultaneously at the same rate but with the fixed offset voltage specified in the legend, resulting in a shifting Gaussian mean (μ). Because the sweep rate for both gates is identical, the offset remains constant. The shift in μ occurs since changing the voltage offset changes the V_{BG} that is applied at a given V_{TG} and vice versa, which will shift the effective threshold voltage of the p-region and n-regions (Note: this effect is well-established in dual-gated thin film and nanomaterial devices[2]). The offset also changes the transconductance of the regions, and thus the drain voltage is changed so that the curves are approximately normalized. For example, for the rightmost curve:

Color	**dV_{BG}/ dV_{TG}**	**V_{TG}-V_{BG} (V) (@V_{TG} = -5 V)**	**V_D (V)**
Light blue	1	3	2

Fig. 1f

Transfer curves where V_{TG} is swept from -5 V to 1 V. V_{BG} is swept simultaneously at the same rate but with no offset. The drain voltage is changed, resulting in amplitude (A) modulation. For example, for the highest curve:

Color	**dV_{BG}/ dV_{TG}**	**V_{TG}-V_{BG} (V) (@V_{TG} = -5 V)**	**V_D (V)**
Light blue	1	0	1

Fig. 1g

Transfer curves where V_{TG} is swept from -5 V to 1 V. V_{BG} is swept with the offset fixed at zero (@V_{TG} = -5 V), but with different rates resulting in a changing Gaussian standard deviation σ. Because the offset at @V_{BG} = -5 V is 0 V, the threshold of the n-region is not significantly modified. However, the differential sweep rate significantly changes the transconductance (in the frame of

V_{BG}). At faster relative V_{TG} sweep rates, the effective transconductance of the n-region (in the frame of V_{BG}) will be more positive since the top gate partially gates the overlap n-type region, and fully gates the non-overlap n-type region. The transconductance of the p-region (in the frame of V_{BG}) will be more negative because of the faster V_{TG} sweep rate. Additionally, the effective threshold of the p-region will shift because there is a V_{BG}/V_{TG} offset at the p-region. The drain voltage is changed so that the curves are approximately normalized. The net effect is to significantly change σ. For example, for the leftmost curve:

Color	**dV_{BG}/dV_{TG}**	**V_{TG}-V_{BG} (V) (@V_{TG} = -5 V)**	**V_D (V)**
Light blue	0.4	0	2

Fig. 1h

Transfer curves where V_{TG} is swept from -8 V to 5 V with V_{BG} fixed at 5 V. The drain voltage is changed to demonstrate changing current levels.

Fig. 1i

Transfer curves where V_{TG} is swept from -5 V to 1 V. V_{BG} is swept simultaneously from β to 3V, which results in an offset that is fixed at -2 V (@V_{TG} = 1 V). The case where β = 3 V corresponds to static voltage bias rather than a sweep. Using the earlier model for sigmoid behavior, as β decreases and the relative V_{BG} sweep rate increases, the total transconductance (in the frame of V_{TG}) of the n-type channel becomes more positive, leading to larger transconductances left of the peak. Once β is sufficiently negative such that the n-type material is capable of being fully shut off within the specified V_{TG} range, the behavior is Gaussian. For example, for β = -2 V:

Color	**dV_{BG}/dV_{TG}**	**V_{TG}-V_{BG} (V) (@V_{TG} = 1 V)**	**V_D (V)**
Light blue	0.5	-2	1

References:

1 Sangwan, V. K. *et al.* Self-Aligned van der Waals Heterojunction Diodes and Transistors. *Nano Letters* **18**, 1421-1427, (2018).

2 Yi, J. *et al.* Double-Gate MoS_2 Field-Effect Transistors with Full-Range Tunable Threshold Voltage for Multifunctional Logic Circuits. *Advanced Materials* **33**, 2101036, (2021).