Performance Limit Prediction of Atomically Thin In_2O_3 Transistors

Zongmeng Yang,[1] Shibo Fang,[1] Linqiang Xu,[1] Qiuhui Li,[1] Jichao Dong,[1] Ying Li,[1] Baochun Wu[7], Ghafoor Mughira,[1] Peiqi Yang,[8] Ying Guo,[9,10] Shimin Hou,[8] Zhaochu Luo,[1,3] and Jing Lu[1,2,3,4,5,6*]

[1] State Key Laboratory for Mesoscopic Physics and School of Physics, Peking University, Beijing 100871, China

[2] Collaborative Innovation Center of Quantum Matter, Beijing 100871, China

[3] Beijing Key Laboratory for Magnetoelectric Materials and Devices (BKL-MEMD), Peking University, Beijing 100871, China

[4] Peking University Yangtze Delta Institute of Optoelectronics, Nantong 226010, China

[5] Key Laboratory for the Physics and Chemistry of Nanodevices, Peking University, Beijing 100871, China

[6] Beijing Key Laboratory of Quantum Devices, Peking University, Beijing 100871, China

[7] State Key Laboratory of Low Dimensional Quantum Physics and Department of Physics, Tsinghua University, Beijing 100871, China

[8] Center for Nanoscale Science and Technology, Key Laboratory for the Physics and Chemistry of Nanodevices, Department of Electronics, Peking University, Beijing 100871, China

[9] School of Materials Science and Engineering, Xi'an University of Technology, Xi'an, 710048, China

[10] School of Physics and Telecommunication Engineering, Shaanxi Key Laboratory of Catalysis, Shaanxi University of Technology, Hanzhong 723001, China

[*]Corresponding Author: jinglu@pku.edu.cn (Jing Lu)

Appendix A

Geometry Optimization The geometry is optimized by using plane-wave basis set implemented in the Vienna ab initio simulation package (VASP). The method is based on the density-functional theory (DFT) with projector augmented-wave pseudopotentials. The Perdew−Burke−Ernzerhof (PBE) functional on generalized gradient approximation (GGA) level is applied for the exchange-correlation interaction. The cutoff energy is 400 eV and the k-point mesh is set as 3×1×3. The convergence criteria for force and energy are set to be 0.05 eV/Å and 1×10^{-5} eV, respectively. A 20 Å vacuum layer is used to prevent periodic image interactions.

Elastic transport calculation The transport properties are calculated using DFT coupled to the nonequilibrium Green's Function (NEGF) method, implemented by the QuantumWise Atomistix ToolKit (ATK) 2022 package. While DFT-GGA method tends to underestimate the bandgaps of intrinsic semiconductors because it neglects the many-body effect, it is reliable in our transport simulation for two reasons. First, the dielectric layer screens the electron-electron interaction in the channel. For example, the bandgap of ML MoS_2 is 2.8 eV at GW level. But it is adjusted to 1.9 eV in FET with high-*k* dielectric which aligns with 1.76 eV calculated by DFT at GGA level. Second, the injection of numerous electrons into the channel contributes to the screening of electron-electron interaction. For instance, the bandgap of the heavily doped monolayer $MoSe_2$ at the GW level is adjusted to 1.59 eV, in agreement with the intrinsic value of 1.52 eV from DFT-GGA calculations and the measured 1.58 eV obtained through the angle-resolved photoemission spectroscopy. The simulation settings include the linear combination of atomic orbitals (LCAO) basis set, PseudoDojo pseudopotentials, density mesh cutoff of 125 Hartree, the *k*-point sampling of 3×1×92, and the electron temperature of 300 K. Dirichlet boundary condition is applied at the interface between the channel and the electrodes to maintain a fixed electrostatic potential. For the region connecting the channel to the vacuum, the Neumann boundary condition is used, as it fixes the derivative of the potential along the boundary direction. Periodic boundary condition is implemented in the direction perpendicular to the transport direction. Hartree potential $V_{\mathrm{H}}(\boldsymbol{r})$ is solved by the Poisson equation

$\nabla^2 V_{\mathrm{H}} = -e^2 n(\boldsymbol{r}) / 4\pi\varepsilon_r\varepsilon_0$ with the above boundary conditions, where $n(\boldsymbol{r})$ is the electron density, ε_r is the relative permittivity and *e* is the elementary charge. The TFT device is divided into the left (source)/right (drain) electrode and the central regions. The retarded Green's function of the central region is:

$$G_{k_{//}}(E) = \left[\left(E + i0^+\right)I - H_{k_{//}}(E) - \Sigma^L_{k_{//}}(E) - \Sigma^R_{k_{//}}(E)\right]^{-1} \quad (1)$$

where the subscript $k_{//}$ means the reciprocal lattice vector perpendicular to the transmission direction. 0^+, *I*, $H_{k_{//}}$ are positive infinitesimal, unit matrix, and the Hamiltonian of the central region with $k_{//}$, respectively. $\Sigma^{\mathrm{L(R)}}_{k_{//}}(E) = \tau^{\mathrm{L(R)}}_{k_{//}} G^{\mathrm{L(R)\,surf}}_{k_{//}}(E)\tau^{\mathrm{L(R)}\,?}_{k_{//}}$ is the self-energy that represents the interaction between the left (right) electrode and the central region. $\tau^{\mathrm{L(R)}}_{k_{//}}$ and $G^{\mathrm{L(R)\,surf}}_{k_{//}}(E) = \left[(E + i0^+)I - H^{\mathrm{L(R)}}_{k_{//}}(E)\right]^{-1}$ is the hopping matrix and the surface Green's function of the left (right) electrode, respectively. Then, the transmission coefficient $\mathsf{T}_{k_{//}}(E)$ is expressed as:

$$\mathsf{T}_{k_{//}}(E) = \mathrm{Tr}\left[\Gamma^{\mathrm{L}}_{k_{//}}(E) G_{k_{//}}(E) \Gamma^{\mathrm{R}}_{k_{//}}(E) \Gamma^{?}_{k_{//}}(E)\right] \quad (2)$$

where $\Gamma^{\mathrm{L(R)}}_{k_{//}}(E) = i\left[\Sigma^{\mathrm{L(R)}}_{k_{//}} - \left(\Sigma^{\mathrm{L(R)}}_{k_{//}}\right)^{\dagger}\right]$ is the broadening function. The wavefunction of the central region can be expressed as $\psi_{k_{//}} = G_{k_{//}}\left(\tau^{\mathrm{L}}_{k_{//}}\phi^{\mathrm{L}}_{k_{//}} + \tau^{\mathrm{R}}_{k_{//}}\phi^{\mathrm{R}}_{k_{//}}\right)$, where $\phi^{\mathrm{L(R)}}_{k_{//}}$ is the wavefunction of the left (right) electrode. Finally, we can get the source-drain current I_{ds} from the Landauer-Büttiker formula:

$$I_{\mathrm{ds}} = \frac{2e}{h}\int_{-\infty}^{+\infty} \mathsf{T}(E, V_{\mathrm{b}}, V_{\mathrm{g}})\left[f_{\mathrm{L}}(E - \mu_{\mathrm{L}}) - f_{\mathrm{R}}(E - \mu_{\mathrm{R}})\right]\mathrm{d}E \quad (3)$$

where the transmission function $\mathsf{T}(E, V_{\mathrm{b}}, V_{\mathrm{g}})$ is the average of $\mathsf{T}_{k_{//}}(E)$ over the irreducible Brillouin zone. $V_{\mathrm{b}} = \mu_{\mathrm{R}} - \mu_{\mathrm{L}}$ is the bias voltage, in which $\mu_{\mathrm{L(R)}}$ is the chemical potential of the left (right) electrode. f_{L} and f_{R} are the Fermi-Dirac distribution of the electrodes in the left and right electrodes, respectively.

Ballistic transport dominates when the channel length is below 10 nm. Recent research shows that the room-temperature ballistic ratio of the 2D InSe FET with 10-nm gate length can reach 83%. The

reliability of the DFT-NEGF approach without phonon scattering is demonstrated by the experiment. The transfer characteristic curve, I_{on}, delay time (τ), power delay product (PDP), and energy-delay product (EDP) of the *n*-type 5-nm-L_g carbon nanotube (CNT) MOSFET simulated by the DFT-NEGF method match well with the experimental results (e.g. I_{on} of 1775 and 1412 μA/μm for simulation and experiment, respectively). The experimental and simulated results of the transfer characteristic curve, I_{on}, τ, PDP, and EDP for the 10-nm-L_g InSe MOSFET are also consistent (e.g. I_{on} of 1497 and 1430 μA/μm for simulation and experiment, respectively).

Appendix B

The optimized In_2O_3 structure of 0.7-nm thickness cleaved from bulk is shown in Figure S1(a). Hydrogen atoms are added to the dangling bonds to stabilize the structure, and the lattice parameters are $a = b = 10.266$ Å. The band structure in Figure S1(b) shows that the 0.7-nm thickness hydrogenated In_2O_3 has a direct bandgap of 2.27 eV. The effective mass m^* of Γ point in the conduction band derived from the equation $\frac{1}{m^*} = \frac{1}{\hbar^2}\frac{d^2E}{dk^2}$ is $0.34m_0$, where E, k, and m_0 are energy, wave vector, and mass of the electron, respectively.

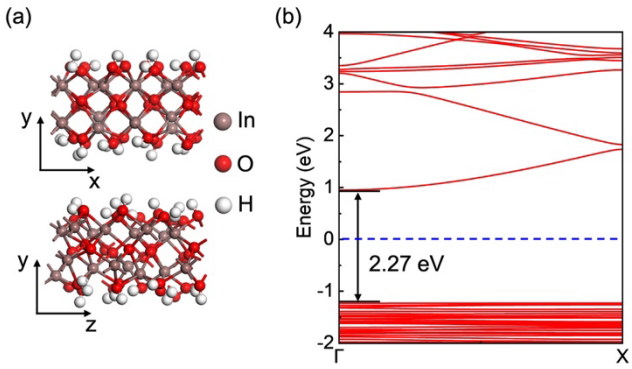
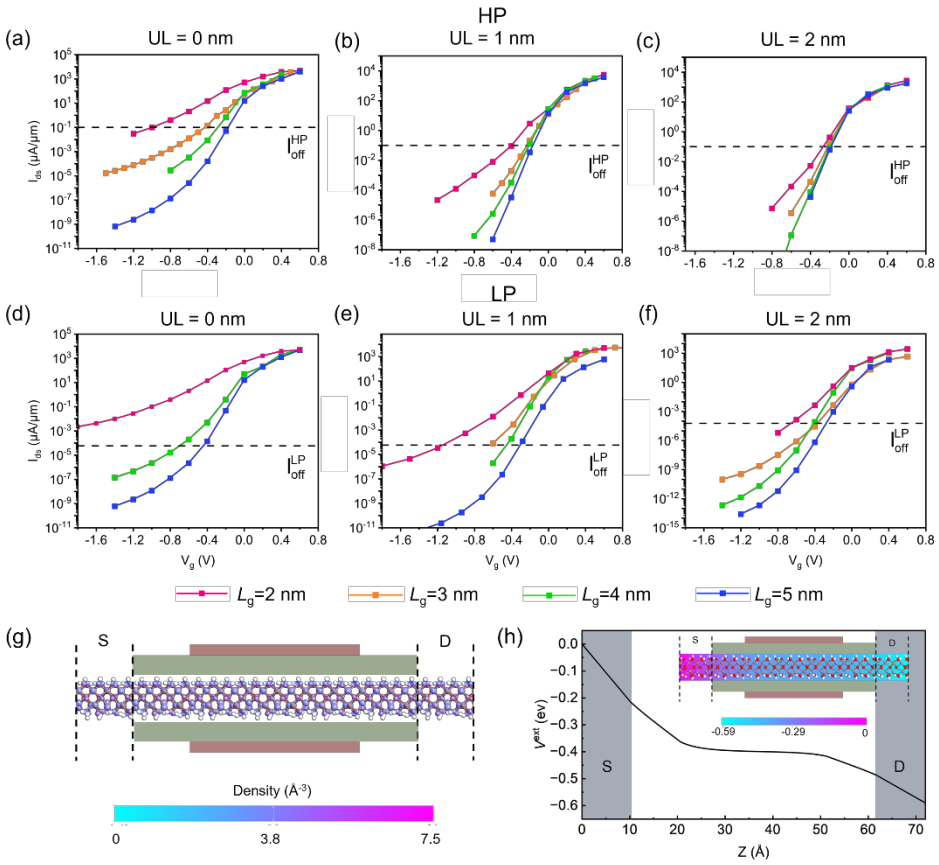


Figure S1. (a) Atomic configuration of 0.7-nm thickness hydrogenated In_2O_3 structure; (b) Electronic band structure. The blue dot line represents the Fermi level;



Appendix C

Figure S2. Source-drain current (I_{ds})-gate voltage (V_g) characteristics of *n*-type DG ultrathin In_2O_3 TFTs with L_g = 2-5 nm, L_{UL} = 0-2 nm. (a-c) HP applications; (d-f) LP applications. The distribution of electron density (g) and external potential (h) of 3-nm-L_g 1-nm-UL In_2O_3 TFTs at on-state (V_g = 0.4 V, V_{ds} = 0.59 V), respectively.

Table S1. Benchmark of the *n*-type 0.7-nm thickness In_2O_3 TFTs with sub-5 nm L_g at ITRS 2013 standards.

	L_g (nm)	UL (nm)	I_{on} (μA μm^{-1})	I_{off} (μA μm^{-1})	SS (mV dec^{-1})	C_t (fF μm^{-1})	τ (ps)	PDP (fJ μm^{-1})
LP	5	0	155.5	5.9×10^{-5}	79	0.163	0.648	0.062
		1	289.8	5.9×10^{-5}	71	0.253	0.541	0.097
		2	306.8	5.9×10^{-5}	77	0.199	0.403	0.077
	4	0	7.0	6.9×10^{-5}	100	0.098	8.419	0.035
		1	248.0	6.9×10^{-5}	79	0.095	0.230	0.034

		2	185.5	6.9×10^{-5}	67	0.084	0.270	0.030
	3	1	7.0	8.0×10^{-5}	115	0.048	4.010	0.016
		2	26.1	8.0×10^{-5}	92	0.060	1.332	0.020
		3	38.3	8.0×10^{-5}	77	0.044	0.668	0.015
	2	1	0.03	9.1×10^{-5}	169	0.027	540.417	0.0086
		2	64.0	9.1×10^{-5}	103	0.031	0.169	0.0098
HP	5	0	1660.1	0.1	78	0.410	0.160	0.170
		1	2140.8	0.1	99	0.294	0.0879	0.120
		2	1095.7	0.1	108	0.172	0.101	0.0706
	4	0	1195.4	0.1	102	0.272	0.141	0.105
		1	2250.8	0.1	81	0.213	0.0586	0.082
		2	1203.6	0.1	111	0.140	0.0720	0.054
	3	0	203.0	0.1	151	0.141	0.411	0.049
		1	1021.0	0.1	112	0.23	0.13	0.08
		2	834.9	0.1	82	0.120	0.085	0.042
	2	0	11.0	0.1	227	0.030	1.582	0.0099
		1	414.0	0.1	162	0.090	0.124	0.029
		2	529.9	0.1	104	0.0847	0.091	0.028
	1	2	169.3	0.1	160	0.054	0.18	0.017
		4	55.1	0.1	165	0.023	0.231	0.007

Appendix D

The gate control and UL work mechanism are investigated by the local device density of states (LDDOS) and spectral currents of the 3-nm-L_g In_2O_3 TFTs with different UL. The channel part is located between the two orange dashed lines. μ_s and μ_d are the chemical potential of the left (source) and (drain) right electrode, respectively. The active energy height (Φ_B) is determined by the energy difference between the conduction band edge of the channel and μ_s. At off-state (Fig. 4 (a)-(c)), Φ_B are all equal to 0.63 eV for UL=0, 1, 2 nm. The spectral currents are divided into three parts: thermal current $\tilde{I}_{therm}$ (current with energy above Φ_B), tunneling current $\tilde{I}_{tunnel}$ and the combination of both which is called mixed current $\tilde{I}_{mixed}$. From the figure, we can see that $\tilde{I}_{tunnel}$ and $\tilde{I}_{mixed}$ are surpassed

as UL increases. When the devices are switched on (Fig. 4 (d)-(f)), Φ_B changes to 0 eV, and only $\tilde{I}_{therm}$ contributes to the total current. In the 1-nm-UL case, the peak of $\tilde{I}_{therm}$ is nearly 5 times larger than that without UL (9.7 vs 2.2 μA/eV), which leads to higher I_{on} (1021 vs 203 μA/eV).

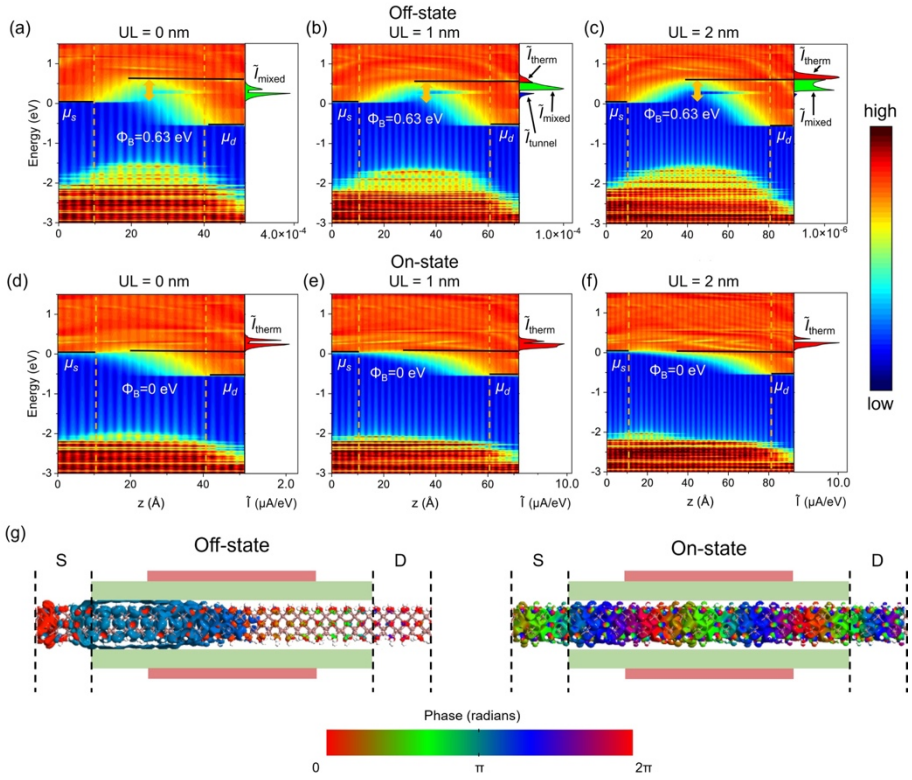


Figure S3. Local device density of states and spectral currents of 3-nm-L_g ultrathin In_2O_3 TFTs at (a-c) off-and (d-f) on-states with UL from 0 to 2 nm for HP applications, respectively; (g) Transmission eigenstates of 3-nm-L_g 1-nm-UL ultrathin In_2O_3 TFTs at the on- and off- states for HP applications.

Appendix E

Strain engineering is an effective approach to improve the performance of the Si MOSFETs. Research has shown that the drive current is increased by approximately 4.5 times in Si pMOSFET and approximately 2 times in Si nMOSFET by applying strain [54,55]. This modulation occurs because strain alters the atomic distances and interactions within the lattice, which in turn modifies the electronic properties, such as effective mass and bandgap. Here, we apply the uniaxial strain along

the x-direction (transport direction) on In_2O_3. The strain effect is reflected by the change in unit cell $a_{strain} = a_0(1+\alpha)$, in which a_0 and a_{strain} are the lattice constant before and after applying strain.

First, we take α from -5% to +2% to study the influence of strain on bandgap E_g and effective mass m^* as shown in Figure S4 (a) and (b). The negative and positive signs represent compression and tension, respectively. When compressive strain is applied, the E_g of ultrathin In_2O_3 changes slightly (from 2.27 to 2.25 eV), and there is also a slight reduction in m^* (from $0.34m_0$ to $0.337m_0$). E_g and m^* also show slight variations in the tension case until α is set as 2%, after which they sharply decrease to 1.82 eV and $0.335m_0$, respectively. The I_{on} of the 4-nm-L_g 1-nm-UL In_2O_3 TFTs with different α to investigate the strain effect on the device performance as shown in Figure S4 (c). I_{on} changes slightly in the compression case due to the slightly changes in bandgap and effective mass. The situation is different when the devices are extended. I_{on} initially increases with the rising strain and then gradually decreases. The most significant enhancement of I_{on} is observed when α is equal to 1%, the corresponding increase is 97%. In brief, a 1% tensile strain can improve the I_{on} of the In_2O_3 TFTs but I_{on} will decrease if $\alpha > 1\%$. When 1% tensile strain is applied, the effective mass remains almost unchanged, and the bandgap decreases slightly compared to the original structure. This reduction in bandgap may increase carrier concentration, thereby enhancing device performance. However, when tensile strain exceeds 1%, significant lattice distortion occurs, affecting the crystal structure. This excessive strain leads to a substantial reduction in the bandgap, which increases the off-state current and ultimately causes a decline in overall device performance. The more detailed underlying mechanism needs further discussion.

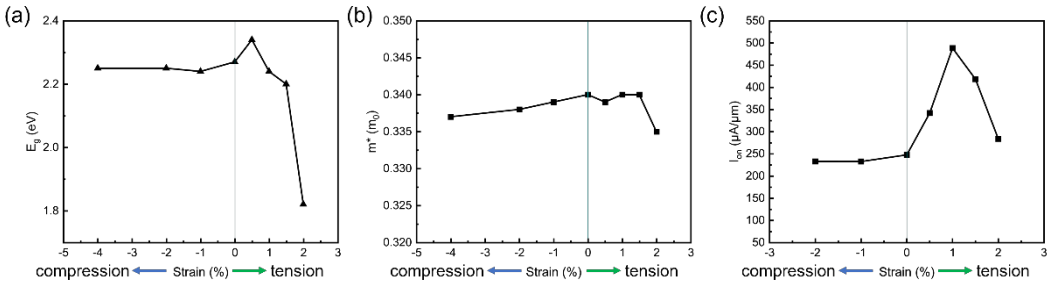


Figure S4. The change of (a) bandgap; (b) effective mass for 0.7-nm thickness In_2O_3 and (c) on-state currents for the 4-nm-L_g 1-nm-UL In_2O_3 TFTs under compression and tension strain.

Appendix F

Based on the MATLAB Simscape model, four types of pseudo-CMOS design inverters are characterized to evaluate the performance of In_2O_3 TFTs on circuit level. The configurations of In_2O_3 TFTs in these inverters are all 5-nm-L_g 1-nm-UL with 1-μm width. The simulated voltage transfer curves (VTCs) and the corresponding circuit diagrams are shown in Figure S5 (a)-(d). The voltage gains of these inverters at a supply voltage V_{dd} of 0.64 V are listed in Table 2. Among these four types of inverters, DL, PEL, and PDL are capable of achieving rail-to-rail operation, which refers to the switching from V_{ss} to V_{dd} except for LEL. PDL and DL inverters both have a high voltage gain of 5, but the former (PDL inverters) contains four transistors, resulting in higher power consumption.

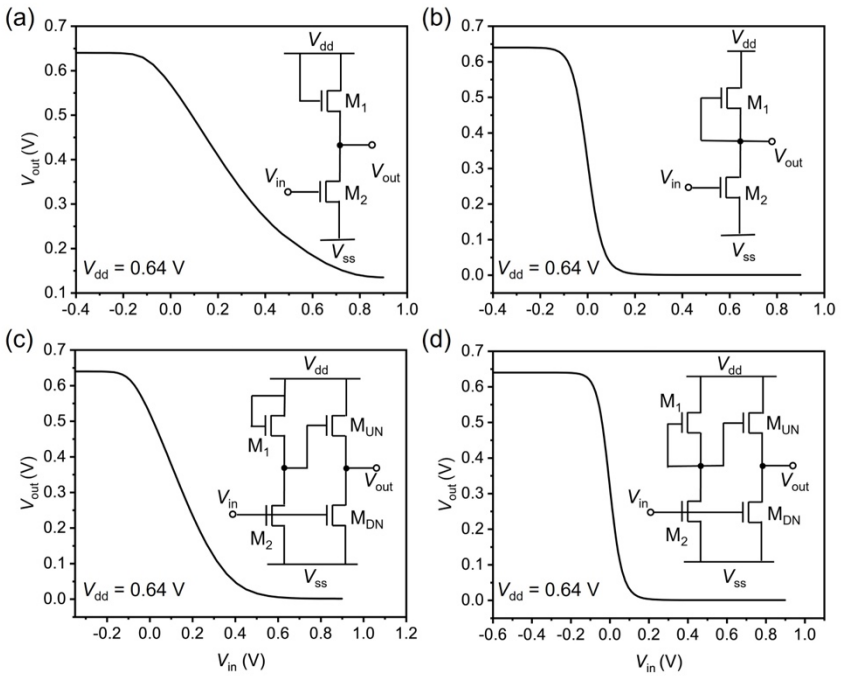


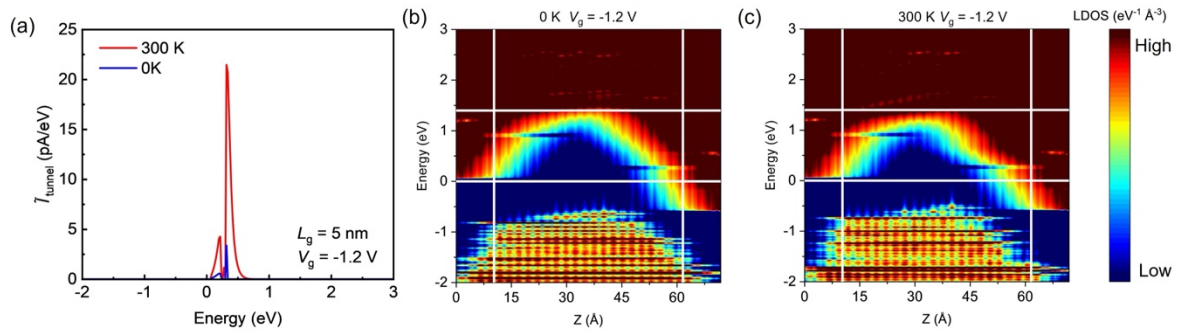
Figure S5. Voltage transfer curves (VTCs) and circuit diagrams of the (a) linear enhancement load inverter LEL; (b) conventional depletion load inverter (DL); (c) pseudo depletion load inverter (PEL); (d) pseudo depletion load inverter (PDL) with V_{dd} of 0.64 V.

Table S2. Performance of the four types of inverters based on the *n*-type 0.7-nm thickness In_2O_3 TFTs with 5-nm-L_g 1-nm-UL.

Inverter type	Voltage Gain	Rail to Rail
PEL	1.5	Yes
PDL	5.1	Yes
LEL	0.8	No
DL	5.0	Yes

Appendix G

To investigate the influence of the EPC effect on the devices, we extract the tunneling spectral current of 5-nm-L_g In_2O_3 TFTs when V_g = -1.2 V is shown in Figure S6 (a). The phonon-assisted tunneling effect enhances the tunneling current intensity and broadens the energy range. Figure S6 (b) and (c) show the temperature-dependent shift and broadening of the local device density of states (LDDOS) at off-state V_g = -1.2 V, which modifies the barrier height and width. Specifically, when EPC is considered, the interaction between electrons and phonons leads to a broadening and shifting of the electronic states at the band edges because of the fluctuations in the atomic positions. As a result, the conduction band edge is effectively perturbed and additional states within the bandgap become accessible for tunneling. This enhancement of available states leads to an increase in the tunneling current, particularly in the off-state regime, where phonon-assisted tunneling becomes



more pronounced.

Figure S6. (a) The tunneling current of 5-nm-L_g In_2O_3 TFT at gate voltage of -1.2 V; Local device density of states of 5-nm-L_g In_2O_3 TFTs at 0 (b) and 300 K (c) when gate voltage is -1.2 V.

Appendix H

Extension of ITRS standard

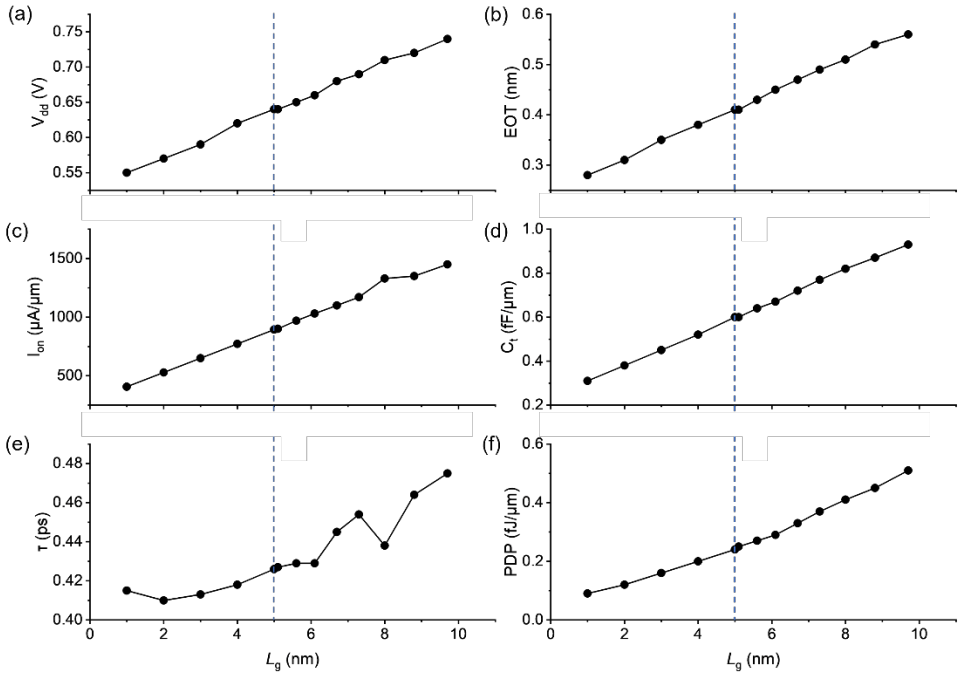


Figure S6. ITRS 2013 standard with figure of merits (FOMs) of (a) supply voltage (V_{dd}); (b) Effective oxide thickness (EOT); (c) On-state current (I_{on}); (d) Total capacitance (Ct); (e) Delay time (τ); (f) Power delay product (PDP). Data below 5-nm-*L*g are the extension of ITRS standard by the linear fit method.