Supporting Information

Jiadong Yao[1], Wenxuan Guo[1], Yali Liu[1], Xinyue Niu[1], Mengge Li[1], Xiaoxiang Wu[1], Ying Yu[1], Tianjian Ou[1], Jian Sha[1] and Yewu Wang[1,2]*

[1]*Department of Physics, Zhejiang Province Key Laboratory of Quantum Technology and Devices & State Key Laboratory of Silicon Materials, Zhejiang University, Hangzhou 310027, P. R. China.*

[2]*Collaborative Innovation Centre of Advanced Microstructures, Nanjing University, Nanjing 210093, P. R. China*

Section 1: The characterizations of device

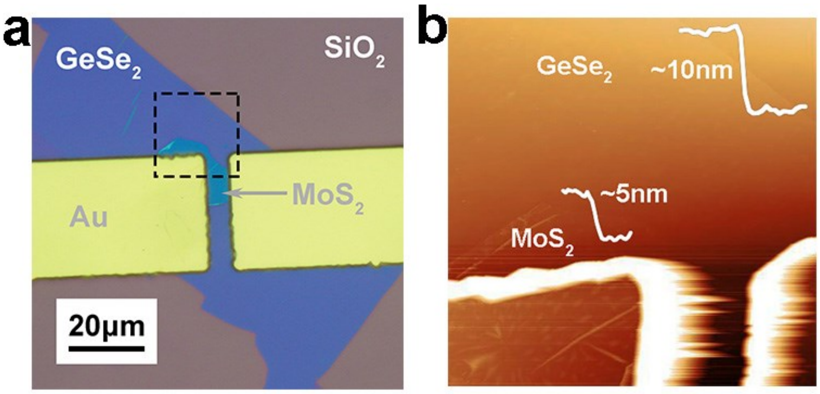

Figure S1. (a) Optical image of the device based on $MoS_2/GeSe_2$ van der Waals heterojunction. (b) Atomic force microscope image of the area marked with dashed rectangle in (a). From the scan lines (white), the thickness of MoS_2 and $GeSe_2$ are 5 nm and 10 nm respectively.

Section 2: Calculation of time constants

Here, we used a simple model[1] to estimate the time scale of charge transfer under the perpendicular gate voltage. When the back gate is switched from 0 V to a strong enough positive gate (for example, +80 V) directly, the band of $GeSe_2$ is tilted and a triangular barrier is formed. At this moment, the electrons in the MoS_2 channel is

* To whom correspondence should be addressed. E-mail: yewuwang@zju.edu.cn.

possible to transfer over the barrier under the effect of gate induced electric field. The change of 2D carrier density n is deduced by the equation below

$$dn = Tn_{3D}vdt \tag{1}$$

where T is the transmission coefficient of electrons over the barrier, n_{3D} is the 3D density of electrons in the MoS_2 channel and v is the velocity of electrons. The relationship between n_{3D} and n is that $n_{3D}=n/D$, where D is the thickness of MoS_2 channel. Therefore, the time constant of this exponential decay could be obtained to be

$$\tau = \frac{D}{v\,T} \tag{2}$$

The velocity v could be obtained by $v=\mu E$, where μ is the mobility and E is the gate induced electric field. First, the mobility is calculated from the transfer curve with the formula below

$$\mu = \frac{dI_{ds}}{dV_g}\frac{L}{WV_{ds}}\frac{1}{C} \tag{3}$$

Where L and W are the length and width of the device, C is the total capacitance of $GeSe_2$ and SiO_2 layer in parallel. The mobility is obtained to be 18.4 $cm^2V^{-1}s^{-1}$ according to the data in the Figure 2a in the main text. The electric field E in the MoS_2 channel is calculated with the relation $E\,\varepsilon=E_I\,\varepsilon_I$, where the electric field E_I in the $GeSe_2$ layer under the gate voltage V_g can be obtained using the Eq. (1) in the main text and the dielectric constants of MoS_2 and $GeSe_2$ are $\varepsilon=12.8$ and $\varepsilon_I=7.0$.

For the transmission coefficient T of electrons over the barrier, calculation based on thermal emission or Fowler-Nordheim tunneling is performed. For thermal emission, there is

$$T_{TE} = e^{-\frac{U}{k_B T}} \tag{4}$$

Where the U is the height of barrier seen by electrons and equals to the conduction band offset here, k_B is the Boltzmann constant and T is the temperature. So $T_{TE} = 9.27\times10^{-16}$ at the room temperature T=300K. For Fowler-Nordheim tunneling, the transmission coefficient could be found[2]

$$T_{FN} = \exp\left(-\frac{4}{3}\sqrt{\frac{2m^*}{(h/2\pi)^2}\frac{(U-\varepsilon)^{3/2}}{eE_1}}\right) \tag{5}$$

Where U is the barrier height, $m^*=0.6m_0$ [3] is the effective mass of electron, h is the Plank constant, E_1 is the electric field of tunneling layer ($GeSe_2$) and ε is the energy of electrons which is much less than the barrier height here. In the case of $V_g = +80\ V$, $T_{FN} = 6.27\times10^{-14}$. Finally, we obtained the estimate of the time constant at $V_g = +80\ V$

$$\tau_{TE} = 37s \tag{6}$$

And

$$\tau_{FN} = 0.5s \tag{7}$$

Section 3: The performance of the memory device

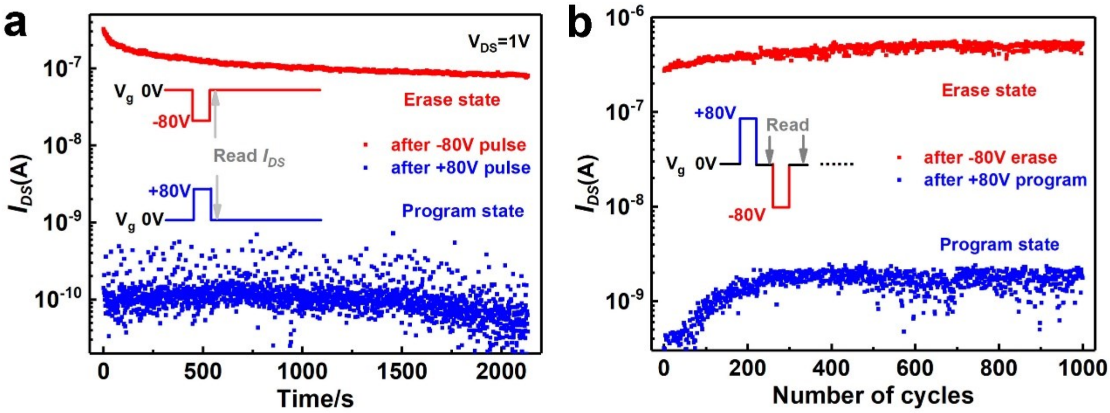

Figure S2. (a) The stability of the two-resistance state. The gate voltage V_g is biased to +80 V (-80 V) first and then reset to 0 V, the time dependence of current in the two states is recorded. (b) Cycle robustness of the device. The V_g is biased to +80 V first so that the device is set to the program state, and after the V_g is reset to 0 V the current is read to confirm the resistance. Then, the V_g is biased to -80 V to set the device to erase state, the current is read again to confirm this resistance after V_g returns to 0 V. The cycle is repeated for thousands of times to test the robustness.

References

[1]A. S. Kumar, M. Wang, Y. Li, R. Fujita, and X. P. A. Gao, ACS Appl. Mater. Inter, **12**(41), 46854 (2020).

[2]M. Depas, B. Vermeire, P. W. Mertens, R. L. Vanmeirhaeghe, and M. M. Heyns, Solid-State Electron., **38**(8), 1465 (1995).

[3]K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, Phys. Rev. Lett., **105**(13), 136805

(2010).