Random telegraph signals caused by a single dopant in a metal oxide semiconductor field effect transistor at low temperature

Kouta Ibukuro,[1, a)] Joseph William Hillier,[1] Fayong Liu,[1] Muhammad Khaled Husain,[1] Zuo Li,[1] Isao Tomita,[1, 2] Yoshishige Tsuchiya,[1] Harvey Nicholas Rutt,[1] and Shinichi Saito[1, b)]

[1)] *School of Electronics and Computer Science, University of Southampton. University Road, Southampton, SO17 1BJ, United Kingdom.*

[2)] *Department of Electrical and Computer Engineering, National Institute of Technology, Gifu college. 2236-2 Kamimakuwa, Motosu, Gifu, 501-0495, Japan.*

(Dated: 1 May 2020)

While the importance of atomic-scale features in silicon-based device for quantum application has been recognised and even the placement of a single atom is now feasible, the role of a dopant in the substrate has not attracted much attention in the context of quantum technology. In this paper, we report random telegraph signals (RTSs) originated from trapping and detrapping of an electron at a donor in the substrate of a p-type Metal-Oxide-Semiconductor Field-Effect-Transistor (MOSFET). RTSs, not seen when the substrate was grounded, were observed when a positive bias was applied to the substrate. The comprehensive study on the signals observed reveals that the nature of the RTSs is discrete threshold voltage variations due to the change in depletion layer width depending on the charge state of a single dopant, neutral or positively charged.

I. INTRODUCTION

The importance of atomic-scale features, such as a dopant[1–6] and a trap state[7–13] has been widely recognised in the context of silicon (Si) quantum technology, such as nano-electronic circuits[1–4,14–16] information processing[6–9,17–19], hardware security[12], bio-sensing[13] and metrology[5,10,11,20,21]. As the size of Si metal-oxide-semiconductor (MOS) field-effect-transistors (FETs) approaches to the physical limit, electronic circuits based on single-atom devices were proposed[14,15]. Use of solitary dopants as a fundamental building block of electronic circuits could be a disruptive solution to maintain the rate of device scaling[1–4,16,22,23]. With regards to quantum information processing, spin qubits based on Si are considered to be promising, due to its weak spin-orbit coupling and abundance of non-magnetic isotopes[17–19]. Realising spin-qubits with single-implanted donors are interesting, as they can provide a discrete energy level with larger energy separation due to stronger quantum confinement than the one realised by a quantum dot (QD) defined by patterning or field-effect[6]. Trap states in Si devices have been considered as an impediment to reliable performance of complementary-MOS (CMOS) FETs[24–30]. However, a few attempts have been made to perform single-spin manipulation based on trap states in a standard MOSFET for spin-qubits, which proved to be successful[7–9]. Also, the variation caused by trap states can be used as a fingerprint of a device for hardware security[12], while a trap state in a liquid-gated FETs for bio-sensing offers enhanced sensitivity to the change in pH of the solution[13]. Finally, such a trap state is known to be useful for a single-electron pump (SEP)[10,11] for quantum metrology[20,21]. SEP is a periodically driven single-electron transistor with tunable potential barriers, outputting drain current of $I_{\mathrm{d}} = ef$, where e is the elementary charge and f is the frequency at which the device is driven[20,21]. A SEP that takes advantage of a trap state has achieved 7.4GHz operation with an uncertainty of 20 parts-per-million (ppm)[11], approaching the metrological requirement of an electric current standard. Without doubt, engineering of such an atomic-scale structures embedded in a Si device will continue to play a crucial part in future quantum applications.

The first challenge in utilising atomic-scale features is to find their signatures in transport characteristics of the devices[31–34]. We recently proposed characterisation of conventional Si MOSFETs with a long integration time at low temperature, which exhibits Coulomb diamonds (CDs)[31] and random telegraph signals[32] (RTSs) in current-voltage (I-V) characteristics. RTSs are discrete threshold voltage (V_{th}) variations over time[25,35], and two V_{th} states are supposed to correspond to an empty and occupied state of a charge trap at Si-SiO_2 interface or inside amorphous SiO_2[32]. The physical origin of the CDs observed in the MOSFETs were, on the other hand, attributed to remote surface roughness of polycrystalline Si (poly-Si)[31]. In this sense, single electron phenomena caused by trap states and structural disturbance in standard MOSFETs have been explored. However, towards application envisaging single-atom circuits[1–4,16] and spin-qubits[6], the use of a single dopant is more desirable and suitable.

In order to achieve single-carrier manipulation using a solitary dopant in a conventional MOSFETs, we focused on the dopants in the substrate, and the dopant ionisation profile was electrically tuned in a systematic manner by applying reverse bias to the substrate. By doing this, we observed RTSs caused by a trapping and detrapping of an electron at a donor in the substrate (well) of a p-type Si MOSFET at low temperatures. The device was initially characterised at 3.8K while the substrate was grounded, and it showed a CD in I-V characteristics, indicating that the effective width of the channel is of the order of 10 nm. Also, no RTSs were observed, which means that no trivial charge traps in the oxide or at the interface were present at this bias condition. However, RTSs started to be

[a)] Electronic mail: ki1m17@soton.ac.uk
[b)] Electronic mail: S.Saito@soton.ac.uk

observed by applying positive bias to the substrate, which can be seen as discrete V_{th} shifts in I-V characteristics. Based on this observation, a physical model to explain this RTS based on trapping and detrapping of an electron at a donor was proposed, where the hole transport was significantly affected by the stochastic change in depletion layer width depending on the charge state of the donor, charge neutral or ionised. The occupancy of the trap level and average lifetimes of two I_d states were systematically controlled by gate voltage (V_g), and the dependence of average RTS lifetimes on V_g suggests that the physical origin of the RTS could be different from trap states at the interface or in the gate oxide[25,36,37]. The temperature was gradually raised to 25K in order to investigate the dependence of the RTSs on temperature, showing almost no dependence until 16K before the faster RTS switching was observed at higher temperature. This indicates that the trapping and detrapping of an electron is achieved via quantum mechanical tunnelling at low temperature[38].

This paper also quantitatively discusses the impact of applying reverse bias to the substrate onto the dopant ionisation profile, based on the result of split capacitance-voltage (C-V) measurements with reverse substrate bias[25,39]. Application of substrate voltage have been widely used in the context of Si quantum devices[9,10,40], though the detailed study on its effect on carrier transport has not been performed. Width of the depletion layer was calculated as a function of V_g and voltage applied to the well (V_{well}), and the depletion layer was constantly extended as V_{well} increased, as expected. Effective hole mobility was also calculated as a function of effective electric field[24,41], and the degradation of hole mobility was prominent at low effective electric field, where impurity Coulomb scattering is the dominant scattering mechanism. However, regardless of the values of V_{well}, the hole mobility approaches to the universal mobility curve at higher effective electric field, where Si-SiO_2 interface scattering dominates the mobility degradation[24,41]. This confirms that the primary effect of applying reverse bias to the substrate is to widen the depletion layer, rather than to increase the electric field across the MOS structure[24,40].

II. MOSFET CHARACTERISTIC AT ROOM TEMPERATURE AND LOW TEMPERATURE

A p-type MOSFET fabricated using a standard 65nm technology was characterised at various temperatures (from 300K to 3.8K) with a Janis pulse-tube refrigerator and a Keysight B1500 semiconductor parameter analyser. Figure 1 (a) and (b) show schematics of the cross-sectional and birds-eye views of the device, respectively. The channel length (L) and width (W) were 50nm and 10μm, respectively, and the capacitive effective thickness (t_{eff}) is 2.4nm. The device was mounted on and wire-bonded to a cryogenic sample holder, and the p-type handle layer was insulated from the holder by cryogenic varnish. Electrical contact to the n-type region in the substrate, well, was achieved from the dedicated pad connected to the well (Figure 1 (a)). Phosphorous was used to form the well.

From our previous works[31,32], the presence of a QD was

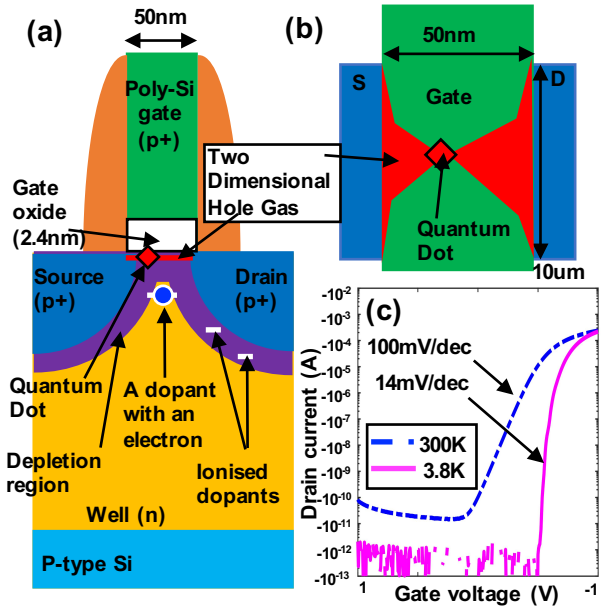

FIG. 1. Schematics and transfer characteristics at 300K and 3.8K of the device. (**a**) Schematic of the cross section of the device. (**b**) Schematic of the device viewed from top. (**c**) Transfer characteristics of the device at 300K (blue broken line) and at 3.8K (solid pink line).

expected in our device, resulting in narrower effective channel width than the actual dimension of the device (Figure 1(b)). Figure 1(c) show the transfer characteristics (drain current (I_d) against gate voltage (V_g)) of the device at both 300K and 3.8K when the drain voltage (V_d) was -50mV and source was grounded (V_s = 0mV). At 300K, the subthreshold slope ($S = dV_g/d\log I_d$) was 100mV/decade, while off current (I_{off}) was around -10pA. As the temperature decreased to 3.8K, S became steeper (14mV/decade) and I_{off} was less than -10fA, which was the expected behaviour of CMOS transistors at low temperatures[24,31]. However, if the channel was considered to be uniform in the subthreshold region at 3.8K, S should have decreased linearly as the temperature decreased, as $S = 2.3kT\left(1 + C_{dep,m}/C_{ox}\right)/e$ indicates, where k is Boltzmann constant, T is temperature, e is elementary charge, $C_{dep,m}$ is the maximum depletion layer capacitance and C_{ox} is the oxide capacitance[24]. This means that the hole channel is considered to be non-uniform and the dominant current path is much narrower than the actual width of the device.

The presence of a QD can be directly verified from 2D contour plots of differential conductance (dI_d/dV_g) against V_g and V_d, shown in Figure 2. I_d was measured twice as a function of $|V_g|$ (from 0.5V to 0.7V with 1mV increments (forward sweep), and from 0.7V to 0.5V with 1mV decrements (reverse sweep)) with fixed V_d to check for the absence of hysteresis, and V_d was varied from -30mV to 30mV with 0.2mV increments. Differential conductance was then calculated and the result for the forward sweep is shown in Figure 2. Compliance value of 1μA was set to limit I_d, which corresponds to two zero conductance region with $|V_d|$ = 30mV and $|V_g|$ = -0.7V. Hysteresis was not observed. Several CDs were ob-

served, which are highlighted by dotted lines in Figure 2. Two CDs with charging energy of about 5meV (CD1 and CD2) were followed by a sequence of CDs with smaller charging energy (CD3), 1 or 2meV, as $|V_g|$ increased. Two CDs with charging energy of 1 or 2meV were also observed at $|V_g| >$ 0.65V, which were labelled as CD4 in Figure 2. The presence of CDs, particularly CD1 and CD2, confirms that the dominant transport was through the QD after the transistor turned on. From the charging energy of the dot, the size of the QD can be estimated[14,31]. The charging energy (E_c) is firstly converted to the coupling capacitance of the QD, C_Σ by $E_c = e^2/C_\Sigma$, which gives C_Σ = 32aF. C_Σ can be decomposed into the coupling capacitance of the QD to source (C_S = 15.0aF), drain (C_D = 5.42aF) and gate (C_G = 11.62aF), from the gradient of the CDs[14,31]. Finally, the size of the QD (S_{QD}) can be estimated from $C_G = C_{eff}S_{QD}$ assuming that the C_{eff} is solely determined by t_{eff}, $C_{eff} = \varepsilon_{ox}/t_{eff}$ = 1.44μF/cm^2. ε_{ox}=3.9ε_0 is permittivity of SiO_2 and ε_0 is permittivity of vacuum, 8.854$\times$10^{12}F/m. S_{QD} is 0.808$\times$10^{-15}m^{-2}, and the diameter of the QD ($d_{QD} = \sqrt{4S_{QD}/\pi}$) is about 32nm when the shape of the QD is assumed to be rounded[31]. The origin of the QD can possibly be remote surface roughness due to a poly-Si grain in the gate electrode[31,32]. The shrinkage of CDs (CD3 and CD4 are smaller than CD1 and CD2) can be attributed to the change in inversion layer thickness[31].

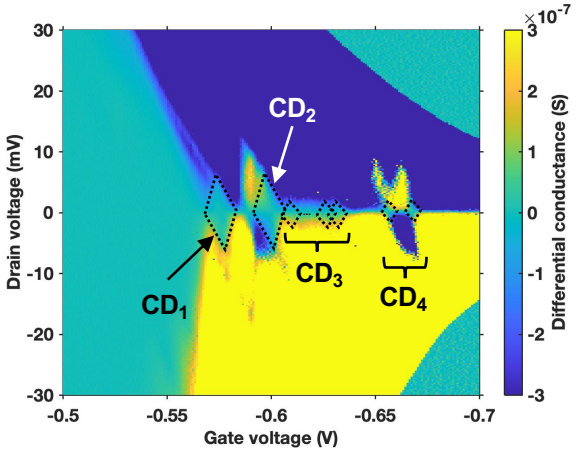

FIG. 2. Differential conductance (dI_d/dV_g) against V_g and V_d at 3.8K. V_{well} was 0V.

These two observations can confirm the transport mechanism in our device at the subthreshold region at low temperature. As $|V_g|$ increased and the device began to operate at the subthreshold region, holes would start to flow through a QD first, since it would provide the current path with low potential energy. After $|V_g|$ exceeded threshold voltage (V_{th}), on the other hand, a 2DHG was uniformly formed under the gate dielectric, and the same on current (I_{on}) of 0.2mA at 300K was achieved (Figure 1(c)). In the subthreshold region located between those two schemes, the carrier transport is thought to be described as a hybrid of the two mechanisms, where holes were predominantly travelling through a narrow, weak link between source and drain involving the QD, resulting in the standard MOSFET characteristic being modulated by Coulomb blockades from the QD (Figure 2), as well as the gentle subthreshold swing (Figure 1(c)).

III. EFFECT OF SUBSTRATE BIAS ON DOPING IONISATION PROFILE

In order to investigate how dopants in the substrate affect the hole transport in the channel, positive substrate bias ($V_{well} > 0$) was applied such that the dopant ionisation profile would be systematically changed. Application of positive substrate bias means that the p-n junction between source/drain and the substrate was reversely biased, leading to more dopants becoming ionised and therefore fewer mobile holes being introduced from source/drain at a given $|V_g|$ to satisfy charge neutrality[24];

$$Q_{tot} = Q_{dep} + Q_{inv} \quad (1)$$

$$Q_{dep} \rightarrow Q'_{dep} = Q_{dep} + \Delta Q_{dep} \quad (2)$$

$$Q_{inv} \rightarrow Q'_{inv} = Q_{inv} - \Delta Q_{inv}, \quad (3)$$

where Q_{tot} is the total charge induced at a certain gate voltage, Q_{dep} is the ionised dopants in the substrate (depletion layer charges), Q_{inv} is the mobile holes introduced from source/drain (inversion layer charges) and $\Delta Q_{dep}(>0)$ and $\Delta Q_{inv}(>0)$ are the increase and decrease in Q_{dep} and Q_{inv}, respectively, due to the reverse bias being applied to the substrate. This effect is called the "body effect"[24], and $|I_d|$ was expected to decrease as V_{well} increased, which is equivalent to a positive $|V_{th}|$ shift[24].

In order to experimentally obtain Q_{dep} and Q_{inv} as a function of V_g and V_{well}, split capacitance-voltage (C-V) measurements with positive substrate bias[42] were performed. From this split C-V measurements, one can quantitatively estimate how many dopants were additionally ionised by the positive V_{well}. This characterisation was performed at room temperature using a Cascade probe station and the B1500. Two devices with (W,L) = (10μm, 10μm) (device B) and (10μm, 4μm) (device C) were measured, which were in the same wafer as the one measured at low temperature (device A), such that the extracted parameters from this measurement can be used for the interpretation of the measurement result obtained from device A. The use of two MOSFETs with different gate lengths is known to be helpful to eliminate parasitic capacitance and resistance[43]. Transfer characteristics with V_d = -50mV were obtained prior to the C-V measurements, and the result from Device B was subtracted from the one from Device C to obtain I_d. Figure 3 shows drain conductance ($g_d = I_d/V_d$) against V_g, while V_{well} was varied from 0mV to 500mV with 50mV increments. V_g was swept from 1V to -1V with 1mV decrements, and the results were shown from 0V to -1V in linear scale to highlight the shift in $|V_{th}|$. $|V_{th}|$ was 0.34V when the substrate was unbiased (V_{well} = 0V), and the shift in $|V_{th}|$ towards positive $|V_g|$ was clearly observed as V_{well} increased.

Then, in order to characterise capacitance of the devices, AC signal with a frequency of 100kHz and an amplitude of 100mV was applied on top of V_g, which was swept from

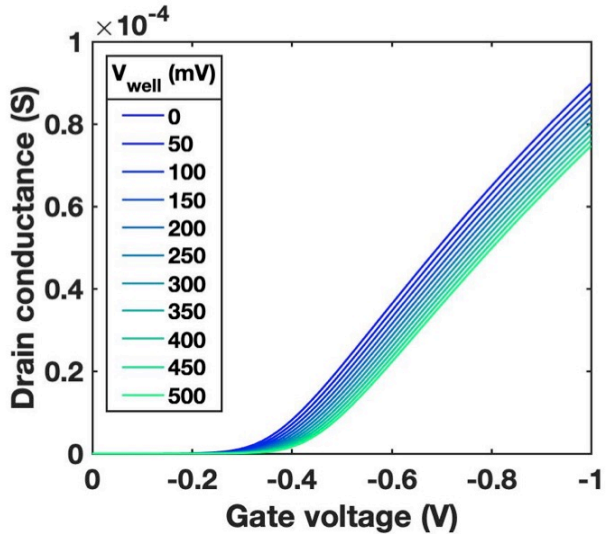

FIG. 3. Drain conductance ($g_d = I_d/V_d$) as a function of V_g with varied V_{well} at 300K. V_d was -50mV.

2V to -1V with 1mV decrements. The out-of-phase current was measured at (1) source and drain together (Figure 4 (a)) (2) substrate (Figure 4 (b)) (3) source, drain and substrate altogether, to characterise (1) inversion layer capacitance, $C_{inv} = dQ_{inv}/dV_g$ (capacitance due to mobile holes) (2) depletion layer capacitance, $C_{dep} = dQ_{dep}/dV_g$ (capacitance due to ionised donors) (3) total capacitance, $C_{tot} = dQ_{tot}/dV_g$ (capacitance due to both ionised donors and mobile holes), while (1) V_{well} was applied from 0V to 500mV with 50mV increments (2) $V_s = V_d$ were applied from 0V to -500mV with -50mV decrements (equivalent to positive V_{well}) (3) N/A (source, drain and substrate were all used to measure current and substrate bias cannot be applied) in order to see the effect of positive substrate bias on the C-V profile. The capacitance measured from device B in those configurations was subtracted by the capacitance from device C and normalised by the difference in the transistor area between the two devices, 10μm by 6μm. To clarify, C_{inv}, C_{dep} and C_{tot} are given as capacitance per unit area with the unit being μF/cm^2. From this split C-V measurements, Q_{inv}, Q_{dep} and Q_{tot} are calculated as a function of V_g, V_{well};

$$Q_{inv}(V_g, V_{well}) = \int_{V_{FB}}^{V_g} C_{inv}(V_g^*, V_{well}) dV_g^* \quad (4)$$

$$Q_{dep}(V_{gs}, V_{well}) = \int_{V_{FB}}^{V_{gs}} C_{dep}(V_{gs}^*, V_{well}) dV_{gs}^*, \quad (5)$$

where $V_{gs} = V_g - V_s$ is gate voltage measured from source.

Figure 4 shows C_{inv}, C_{dep} at various V_{well} values and C_{tot} against V_{gs}. C_{inv} were plotted by solid lines with a colour gradient from blue (V_{well} = 0mV) to light green (V_{well} = 500mV), while C_{dep} were plotted by solid lines with a colour gradient from red ($V_s = V_d$ = 0mV, equivalent to V_{well} = 0mV) to purple ($V_s = V_d$ = -500mV, equivalent to V_{well} = 500mV). Orange empty circles shows C_{tot} against V_{gs}, matching the result of C_{dep} with V_{well} = 0mV in the accumulation region and C_{inv} with V_{well} = 0mV in the inversion region.

C_{inv} shows its dependence on V_{well} in the inversion region ($V_{th} > V_g$), corresponding to the threshold voltage shift due to the body effect[24]. At a given V_g, C_{inv} is lower when positive V_{well} was applied, for example V_{well} = 500mV;

$$C_{inv}(V_g, V_{well} = 0V) > C_{inv}(V_g, V_{well} = 500mV). \quad (6)$$

Therefore,

$$Q_{inv}(V_g, V_{well} = 0V)$$
$$= \int_{V_{FB}}^{V_g} C_{inv}(V_g^*, V_{well} = 0V) dV_g^* \quad (7)$$
$$> \int_{V_{FB}}^{V_g} C_{inv}(V_g^*, V_{well} = 500mV) dV_g^* \quad (8)$$
$$= Q_{inv}(V_g, V_{well} = 500mV). \quad (9)$$

This result is consistent with the physical picture described in equation (3).

As V_{well} increased, accumulation C-V curves are also shifted towards positive V_{gs}, meaning that flatband voltage (V_{FB}) is shifted by applying positive V_{well}. This is reasonable, as the flatband condition is determined solely by the difference between V_{gs} and $V_{well} - V_s$, and therefore as $V_{well} - V_s$ increased, more V_{gs} needs to be applied in order to achieve the same flatband condition. This parallel V_{FB} shift resulted in the increase in maximum depletion layer charge $Q_{dep,m}$ (Q_{dep} at $|V_{gs}| \gg |V_{th}|$) by applying positive V_{well}, for example 500mV;

$$Q_{dep,m}(V_{well} = 0V)$$
$$= \int_{V_{FB}=1.01V}^{-\infty} C_{dep}(V_{gs}^*, V_{well} = 0V) dV_{gs}^* \quad (10)$$
$$< \int_{V_{FB}=1.51V}^{-\infty} C_{dep}(V_{gs}^*, V_{well} = 500mV) dV_{gs}^* \quad (11)$$
$$= Q_{dep,m}(V_{well} = 500mV), \quad (12)$$

which is consistent with equation (2). Table I shows the calculated $Q_{dep,m}$ at various V_{well} values, which proves the inequality (11).

TABLE I. Maximum depletion layer charge at different V_{well} ($V_g = -1V$)

V_{well} (mV)	0	50	100	150	200	250	300	350	400	450	500
$Q_{dep,m}$ (μC/cm^2)	0.404	0.413	0.423	0.431	0.440	0.449	0.458	0.466	0.473	0.480	0.487

So far, from the split C-V measurements, the overall trend expected from body effect (equation (2) and (3)) was experimentally proven. Furthermore, the depletion layer width can be quantitatively obtained as a function of V_g and V_{well}. This is equivalent to analyse the ionisation profile of donors in the substrate as V_g and V_{well} changes, which gives further insight into the effect of applying V_{well}. To this end, the following model was employed[24];

$$Q_{dep}(V_{gs}, V_{well}) = eN_d W_{dep}(V_{gs}, V_{well}), \quad (13)$$

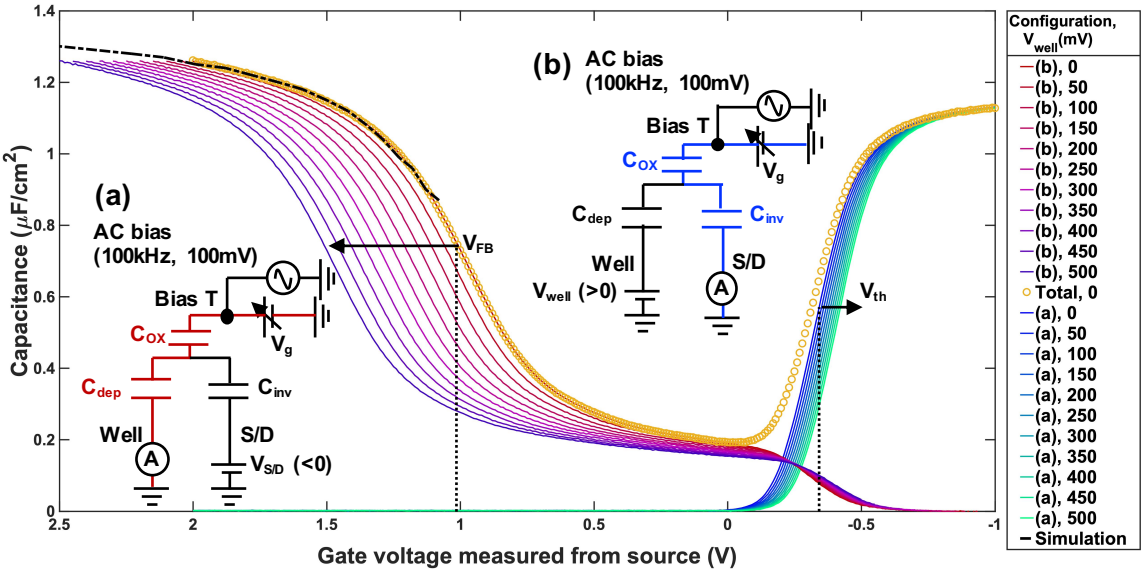

FIG. 4. The results of the split C-V measurements with varied V_{well} at 300K. Two insets ((**a**) and (**b**)) describe the measurement setups to characterise C_{inv} and C_{dep}, respectively. C_{inv} against V_{gs} with different V_{well} values (from 0V to 500mV with 50mV increments) are plotted by solid lines with a colour gradient from blue to light green. C_{dep} against V_{gs} with different V_{well} values (from 0V to 500mV with 50mV increments) are plotted by solid lines with a colour gradient from red to purple. Orange empty circles show total capacitance C_{tot} ($=C_{inv}+C_{dep}$) at V_{well} = 0V. A simulated C-V curve using analytical quantum mechanical model was plotted in a black broken line.

which is valid when $|V_{gs}| > |V_{FB}|$. This model means that the total amount of Q_{dep} in the n-well per unit area can be described by the product of the substrate doping (donor) concentration (N_d) and the depletion layer width (W_{dep}). The spatial variation of N_d is ignored in this model and the precise value of N_d is not assumed *apriori*. Firstly, N_d is determined from $Q_{dep,m}(V_{well} = 0V)$ and maximum depletion layer width, $W_{dep,m}$. After $|V_{gs}|$ exceeds the threshold voltage ($|V_{th}|$), the depletion layer width is expected to be constant, known as maximum depletion layer width[24], which is given below;

$$W_{dep,m} = \sqrt{\frac{4\varepsilon_{Si}kT\ln(N_d/n_i)}{e^2N_d}}, \qquad (14)$$

where $\varepsilon_{Si} = 11.9\varepsilon_0$ is permittivity of Si and n_i is intrinsic carrier density of Si, 10^{11}cm^{-3}. This $W_{dep,m}$ was assumed to be valid only when $V_{well} = 0V$, and was used as a boundary condition to obtain N_d using $Q_{dep,m}(V_{well} = 0V)$;

$$Q_{dep,m}(V_{well} = 0V) = eN_d\sqrt{\frac{4\varepsilon_{Si}kT\ln(N_d/n_i)}{e^2N_d}}, \qquad (15)$$

which gives $N_d = 5.97\times10^{17}\text{cm}^{-3}$. This means that the average distance between dopant atoms are 12nm, which is in the same scale as the gate length of the device. Using this N_d, W_{dep} was calculated as a function of V_{gs} and V_{well} using equation (13), and plotted in Figure 5. W_{dep} against V_{gs} were plotted by solid lines with a colour gradient from red (V_{well} = 0mV) to purple (V_{well} = 500mV). As expected, W_{dep} becomes almost constant after $|V_{gs}|$ exceeds threshold voltage (-0.34V) at any V_{well}. The inset in Figure 5 shows W_{dep} at V_{gs} = -0.5V as a function of V_{well}. This shows that in this model the rate of depletion layer widening upon the application of V_{well} (dW_{dep}/dV_{well}) is almost constant about 20nm/V.

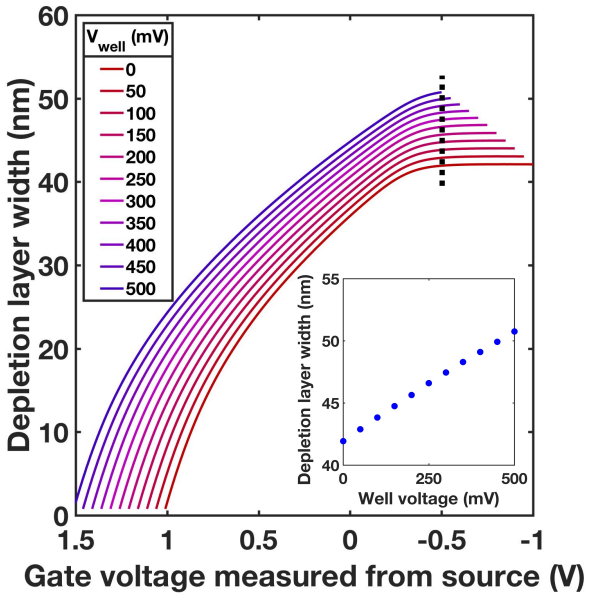

FIG. 5. Depletion layer width (W_{dep}) as a function of V_{gs} with varied V_{well} at 300K. (Inset) W_{dep} at V_{gs} = -0.5V as a function of V_{well} (values along the dotted line in the main figure).

By analysing the change in depletion layer width due to application of V_{well}, it can be concluded that reverse substrate bias can indeed systematically alter the donor ionisation pro-

file. The increase in ionised dopants in the substrate does not only reduce the minority carriers in the inversion layer, but also enhance interaction between holes and ionised dopants by remote Coulomb scattering, resulting in the reduction of hole mobility. This can be experimentally confirmed by calculating effective hole mobility ($\mu_{\rm eff} = \frac{L}{W}\frac{g_{\rm d}}{Q_{\rm inv}}$) against effective electric field ($E_{\rm eff} = \frac{1}{\varepsilon_{\rm Si}}(Q_{\rm dep} + \frac{Q_{\rm inv}}{3})$) at various $V_{\rm well}$ values[41]. Solid lines with a colour gradient from blue to light green in Figure 6 show the result, and the black dotted line is the universal mobility curve ($\mu_{\rm univ} = 185/(1 + \frac{E_{\rm eff}}{0.45})$) for Si-$SiO_2$ interface. As can be seen from Figure 6, the degradation of hole mobility is more prominent at low $E_{\rm eff}$ values, where the dominant scattering mechanism is remote Coulomb scattering between holes and ionised donors. At high $E_{\rm eff}$, however, regardless of $V_{\rm well}$ values, all the mobility curves overlapped and approached to the universal mobility curve. This means that the scattering at the Si-SiO_2 interface, the dominant scattering mechanism at high $E_{\rm eff}$, has not increased much upon the application of $V_{\rm well}$, and the enhanced interface scattering does not account much for the degradation of hole mobility. The mobility degradation has been mentioned previously, and the mechanism was considered to be enhanced Si-SiO_2 interface scattering from qualitative discussion[40,44]. Our quantitative result reveals that the enhanced Coulomb scattering is the dominant mechanism to explain the mobility reduction by applying $V_{\rm well}$.

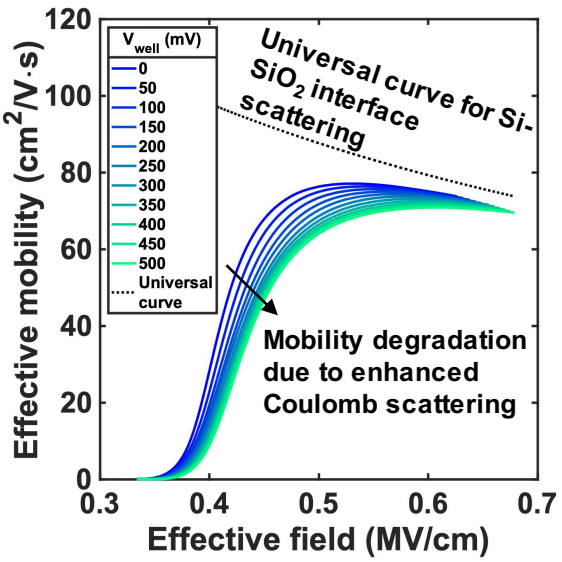

FIG. 6. Effective mobility ($\mu_{\rm eff}$) against effective electric field ($E_{\rm eff}$) with varied $V_{\rm well}$ at 300K. The degradation of mobility was observed when $E_{\rm eff} < 0.5$MV/cm with higher $V_{\rm well}$ values, while $\mu_{\rm eff}$-$E_{\rm eff}$ curves approach to the universal mobility curve for Si-SiO_2 interface, indicating that the mobility degradation by applying substrate bias is due to enhanced Coulomb scattering.

The reliability of the C-V measurements was assured from good agreement with the numerical simulation based on analytical quantum mechanical model[45]. This model assumes the potential exponentially depends on the distance from the Si-SiO_2 interface, which allows one to find an exact solution of Schrödinger equation. The solution is then feed-backed into Gauss's law, which determines the parameter of the exponential potential, resulting in self-consistent Poisson-Schrödinger equation[45]. The result of C-V simulation is shown in Figure 4 as a black dashed line, which agreed well with the experimental result with equivalent oxide thickness (EOT) of 2nm being used as a fitting parameter. Other input parameters were $N_{\rm d} = 6\times10^{17}{\rm cm}^{-3}$ and $V_{\rm FB} = 1.01$V, which were determined from equation (15) and maximum curvature of the C-V curve in accumulation region, respectively. The difference between $t_{\rm eff} = 2.4$nm and EOT = 2nm is attributed to the quantum confinement near the Si-SiO_2 interface, which can be converted into width of wavefunction in Si, $t_{\rm Si} = (t_{\rm eff} - {\rm EOT}) \times \varepsilon_{\rm Si}/\varepsilon_{\rm ox}$ = 1.2nm, which is reasonable[31,45,46].

IV. RANDOM TELEGRAPH SIGNALS TRIGGERED BY REVERSE SUBSTRATE BIAS

After investigating the effect of substrate bias on the dopant ionisation profile in the substrate, detailed I-V scans with positive $V_{\rm well}$ were performed to see its effect on the hole transport in the device. Figure 7 (a) to (f) show 2D contour plots of $I_{\rm d}$ as a function of $V_{\rm g}$ and $V_{\rm d}$ at $V_{\rm well}$ = 0V, 100mV, 200mV, 300mV, 400mV and 500mV. $V_{\rm g}$ and $V_{\rm d}$ were swept from -0.5V to -0.7V and 30mV to -30mV with 1mV and 0.2mV decrements, respectively, and the results within -0.7V < $V_{\rm g}$ < -0.6V and -15mV < $V_{\rm d}$ < 15mV are displayed. Even at current range of sub-μA, the presence of CD4 can be clearly seen. Also, as $V_{\rm well}$ increased, $|V_{\rm th}|$ shifted towards higher $|V_{\rm g}|$, as expected. Another notable feature in Figure 7 (c) to (e) is that a few discrete current peaks were observed, highlighted by arrows in the figures. Such current peaks were not observed when $V_{\rm well}$ = 0V (Figure 7 (a)) and the number of current peaks seen in Figure 7 (a) and (f) ($V_{\rm well}$ = 100mV and 500mV) is much less than Figure 7 (c), (d) and (e) ($V_{\rm well}$ = 200mV, 300mV and 400mV). This indicates that the peaks could only be observed around a certain $V_{\rm well}$ value. The current peaks were observed for both positive and negative $V_{\rm d}$, and the ones in negative $V_{\rm d}$ ($V_{\rm d} < 0$) are addressed in the following discussion. To clarify the nature of the current peaks, $I_{\rm d}$-$V_{\rm g}$ curves with $V_{\rm well}$ = 200mV and $|V_{\rm d}|$ varying from 7mV to 13mV with 1mV increments are displayed in Figure 8. Forward sweeps (increasing $|V_{\rm g}|$) are shown with solid lines, while reverse sweep (decreasing $|V_{\rm g}|$) are shown with dotted lines. Discretised threshold voltage shifts of about 1.5mV were observed around $|V_{\rm g}|$ = 0.65V in Figure 8, which coincides with the current peaks observed in 2D contour plots. The threshold voltage shifts continued to be observed with higher $|V_{\rm d}|$. The inset of Figure 8 shows $I_{\rm d}$-$V_{\rm g}$ curve with $V_{\rm well}$ = 200mV and $|V_{\rm d}|$ = 30mV, showing the same threshold voltage shift.

To confirm that the observed threshold voltage shift is caused by a single carrier[47–49], $|I_{\rm d}|$ was monitored over 1000s (time domain characteristic). The interval of measurements was 100ms, and the integration time (included in the interval) was 20ms. Figure 9 (a) to (e) show the time domain characteristic at various $|V_{\rm g}|$ (from large $|V_{\rm g}|$ to small $|V_{\rm g}|$), while

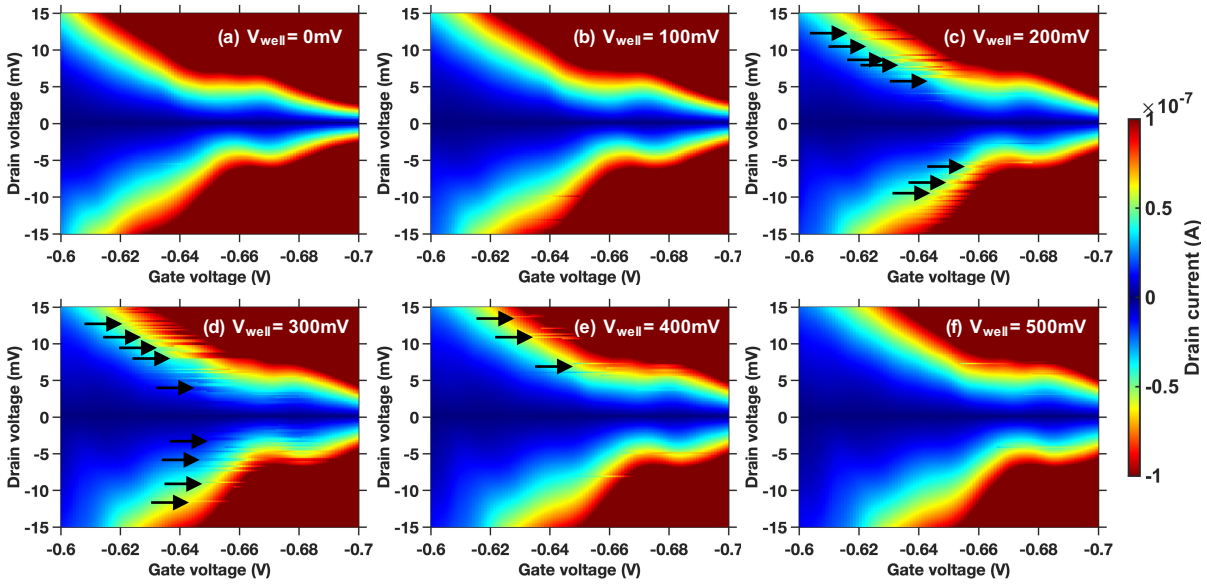

FIG. 7. 2D contour plots of drain current (I_d) as a function of V_g and V_d with V_{well} being (**a**) 0mV, (**b**) 100mV, (**c**) 200mV, (**d**) 300mV, (**e**) 400mV and (**f**) 500mV at 3.8K. Current peaks were highlighted by black arrows in (**c**), (**d**) and (**e**). It took 5s for V_g to be swept from -0.6V to -0.7V.

Figure 9 (f) to (j) show corresponding histograms (probability to observe a certain current value) for each time trace. $|V_d|$ was fixed at 30mV to increase the signal to noise ratio between the high $|V_{th}|$ state (= the low $|I_d|$ state) and the low $|V_{th}|$ state (= the high $|I_d|$ state). When $|V_g|$ was 0.62V (Figure 9 (e) and (j)), $|I_d|$ was stable around 0.143μA. The random switching between two states started to be observed at $|V_g|$ = 0.639V, which can also be identified from the two peaks in the histogram (Figure 9 (j)). The probability of the two states was almost equal at $|V_g|$ = 0.645V, and the low current state was the dominant state at $|V_g|$ = 0.655V, indicating that the probability to observe one of the states are well controlled by the gate voltage. These $|V_g|$ values coincide with the range of $|V_g|$ in Figure 8 where the threshold voltage shifts were observed, meaning that the nature of the shifts were RTSs. Also, as the switching was between only two states, the fhysteresis in $|I_d|$ observed in Figure 8 was due to trapping and de-trapping of a single carrier and no multiple charge states were involved[47–49]. At $|V_g|$ = 0.725V, $|I_d|$ was stable around 5.02μA and the random switching between two current states were not observed.

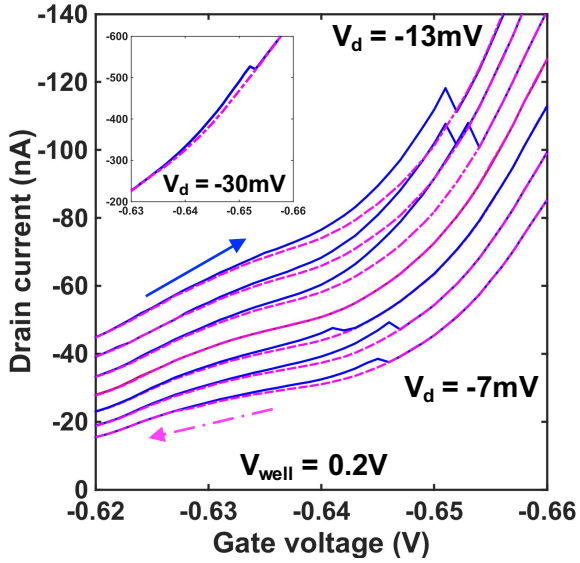

FIG. 8. I_d as a function of V_g with V_{well} = 200mV and varied V_d of (main) -7mV to -13mV with 1mV decrements and (inset) -30mV at 3.8K. Blue solid lines show the results for forwards sweeps ($|V_g|$ increased from 0.62V to 0.66V), while pink broken lines show the results for reverse sweeps ($|V_g|$ decreased from 0.66V to 0.62V). It took 72s for $|V_g|$ to be swept from 0.62V to 0.66V.

V. DISCUSSION; SINGLE ELECTRON TRAPPING AND DE-TRAPPING IN THE SUBSTRATE

In the previous sections, the I-V and C-V characteristics of pMOSFETs at room temperature and low temperature were introduced (Section II, III), before unexpected observation of RTSs at 3.8K by the application of positive substrate voltage (V_{well}) were shown in detail (Section IV). The biggest difference from conventional RTSs observed in CMOS devices[24–30] is that the RTSs were only observed when positive bias was applied to the substrate such that the depletion layer was further widened from its maximum width[24,25]. Although the modification of RTSs parameters, such as $\Delta|I_d|$, lifetime of the two current state (τ_{High} and τ_{Low}), by applying V_{well} was previously reported[25,39], RTSs triggered by applying V_{well} have not been observed so far. As seen in Section III, the primary

effect of applying reverse substrate bias is to increase ionised donors in the substrate and widen the depletion layer, which result in a positive $|V_{th}|$ shift. Therefore, we attribute the origin of the RTS to a dopant atom ionised by the application of positive V_{well}, capturing and re-emitting an electron and shifting $|V_{th}|$ accordingly. The presence and absence of a single electron in the substrate and resulting change in the depletion layer width could have a significant impact on the hole transport, as the effective width of the channel was limited by the size of the QD, about 32nm, at low temperature (Figure 1). Figure 10 illustrates the proposed physical model to explain the mechanism of the RTS. When V_{well} was grounded (0V), a dopant atom was below the Fermi energy (E_F) and therefore filled with an electron, meaning that it could not affect the hole transport in the inversion layer (Figure 10 (a)). As V_{well} increased, the depletion layer extended and the dopant level was subsequently raised, resulting in the dopant level aligned with E_F (Figure 10 (b) and (c)). At this condition, both situations where the dopant was ionised or filled with an electron were energetically equally favourable, such that the switching between those two charge states would occur. Further increase in V_{well} would result in complete ionisation of the dopant, and the dopant could not influence the transport anymore.

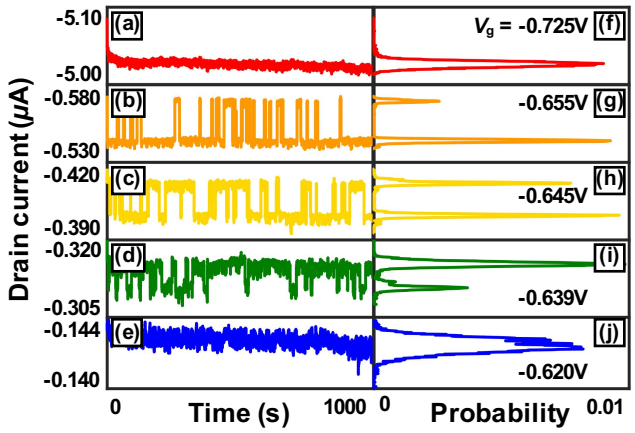

FIG. 9. Time domain characteristics of the random telegraph signals at 3.8K. (**a** – **e**) I_d against time (1000s) with V_g being (**a**) -0.725V, (**b**) -0.655V, (**c**) -0.645V, (**d**) -0.639V and (**e**) -0.620V, while V_d and V_{well} of -30mV and 200mV were applied, respectively. (**f** – **j**) I_d against the probability to observe a certain I_d values ($P(I_d)$) within 1000s (histograms). (**f**) is the histogram for the time trace shown in (**a**), (**g**) for (**b**), (**h**) for (**c**), (**i**) for (**d**) and (**j**) for (**e**).

In order to justify this model, we estimate how much V_{well} is required to ionise one dopant under the QD, δV_{well}. The extension of depletion layer, δW_{dep}, is defined such that the volume under QD ($S_{QD}\delta W_{dep}$) contains one donor, $S_{QD}\delta W_{dep}N_d = 1$. This results in δW_{dep} = 2.1nm, and therefore $\delta V_{well} = \delta W_{dep}(dW_{dep}/dV_{well})^{-1}$ is about 0.1V, which is of the same orders of magnitude with V_{well} causing the RTSs. Considering the uncertainty associated with the size of the QD, this estimation is in good agreement with our observation.

This claim can also be supported from the direction of $|V_{th}|$ shift due to RTSs. As $|V_g|$ increased, the dominant current state shifted from the high $|I_d|$ state to the low $|I_d|$ state, corresponding to the positive threshold voltage shift (Figure 9). This can be explained by the ionisation of a single dopant due to the increased $|V_g|$. The dopant was initially well below E_F, before being brought into E_F by applying positive V_{well}. Then, as $|V_g|$ increased, the conduction band became bent further and the dopant level became in resonant with E_F, resulting in an electron escaping and re-entering the donor level by quantum mechanical tunnelling or thermal activation. This ionisation contributed to the further widening of the depletion layer, leading to the positive $|V_{th}|$ shift, consistent with the observation (Figure 9). Further increase in $|V_g|$ would result in complete ionisation of the dopant, and RTSs would not be observed anymore.

VI. TRAPPING AND DETRAPPING PROCESS OF AN ELECTRON

In the previous section, the physical origin of the RTS was suggested to be a dopant in the substrate, and two current states, high $|I_d|$ state and low $|I_d|$ state, were attributed to be two charge states, charge neutral and ionised, respectively. In order to study the mechanism of the RTS further, the statistics of the signal were investigated in detail[25]. Figure 11 shows the occupancy of each of the two current states, which is the probability to observe the high $|I_d|$ state (N_{High}) or the low $|I_d|$ state (N_{Low}) against gate voltage. N_{High} and N_{Low} were defined as follows[33,34];

$$N_{High} = \frac{1}{I_T}\int_{\langle I_d\rangle}^{\infty} P(|I_d|)\mathrm{d}|I_d| \tag{16}$$

$$N_{Low} = \frac{1}{I_T}\int_{-\infty}^{\langle I_d\rangle} P(|I_d|)\mathrm{d}|I_d|, \tag{17}$$

where $P(|I_d|)$ is the probability to observe a certain $|I_d|$ value (Figure 11), $\langle I_d\rangle = \frac{1}{2}(I_{d,High} + I_{d,Low})$ is the average $|I_d|$ value of the high $|I_d|$ state ($I_{d,High}$) and the low $|I_d|$ state ($I_{d,Low}$), and $I_T = \int_{-\infty}^{\infty} P(|I_d|)\mathrm{d}|I_d|$ is a normalising factor. As $|V_g|$ increased, N_{High} decreased and N_{Low} increased, which is consistent with the transfer characteristic (Figure 8) as well as time domain measurement and their corresponding histograms (Figure 9). At $|V_g|$ = 0.645V, N_{High} and N_{Low} became almost equal, meaning that at this bias condition the high $|I_d|$ state and the low $|I_d|$ state were energetically almost equally favourable.

Asymmetry was found in N-$|V_g|$ characteristic (Figure 11) around $|V_g|$ = 0.645V. RTSs cannot be detected when $|V_g|$ was smaller than 0.639V, where the high $|I_d|$ state dominates (about up to 80%). However, RTSs were still observed when $|V_g|$ were more than 0.66V, where the dominant low $|I_d|$ state, exceeded 80%. This asymmetry can be attributed to nonlinear $|I_d|$ - $|V_g|$ characteristics[34]. Firstly, $|I_d|$ is not only susceptible to RTSs, but also disturbed by analogue noise such as shot noise[50], thermal noise[51,52], negative bias temperature instabilities (NBTI)[53], noise from electrical component of this systems (cables, adapters etc), which would be overlaid onto the time traces of $|I_d|$ and widen the probability distribution of $|I_d|$ around the mean value. For RTSs to be detected, the

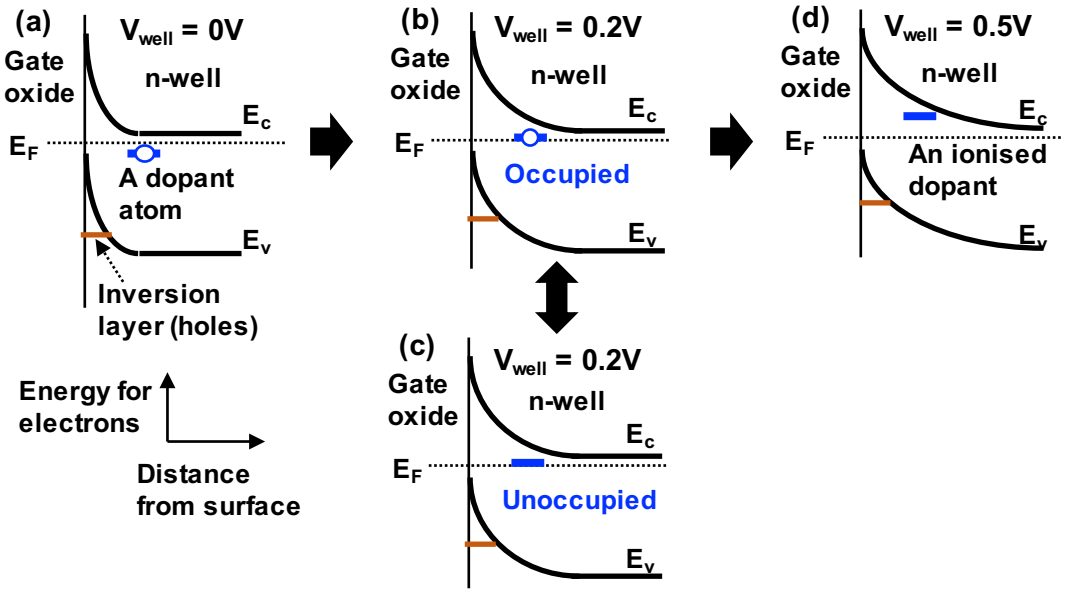

FIG. 10. Energy band diagrams to explain the physical model of the random telegraph signals observed in our device. (**a**) A dopant located at the outside of the maximum depletion layer and lower than Fermi level (E_F) could not influence the transport of inversion layer charges, holes in this case. (**b**) and (**c**) As V_{well} increased, the depletion layer extended and the dopant level moved upwards, resulting in the energy level of the dopant aligned with E_F. This means that the both charge states of the dopant, occupied by an electron (charge neutral (**b**)) or unoccupied (ionised (**c**)), were both energetically equally favourable, and the two states stochastically switched over time. When the dopant became ionised, the depletion layer underneath the hole channel, which is a narrow weak link with the width of 30nm, was further widened, leading to positive, discrete threshold voltage ($|V_{th}|$) shift. (**d**) Further increase in V_{well} resulted in the complete ionisation of the dopant and the RTS were not observed.

amplitude of RTSs ($\Delta|I_d|$) needs to be larger than the deviation around the mean value of each current state, which can be characterised by full width of half maximum (FWHM) of each current state. If $|I_d|$ is a linear function of $|V_g|$, $\Delta|I_d|$ would be a constant value, and to the extent that the value is larger than FWHMs of both states RTSs should be detected, resulting in a symmetrical N-$|V_g|$ characteristic. If $|I_d|$ increases faster than a linear function ($|V_g|^2$ for example), $\Delta|I_d|$ would increase as a function of $|V_g|$[34], and detecting RTSs would be difficult at small $|V_g|$ since FWHM and $\Delta|I_d|$ are comparable. Figure 11 inset shows FWHM of both the high $|I_d|$ state ($FWHM_{High}$) and the low $|I_d|$ state ($FWHM_{Low}$), and $\Delta|I_d|$ against gate voltage. A double Gaussian function was used to fit the probability against $|I_d|$ to obtain $FWHM_{High}$, $FWHM_{Low}$ and $\Delta|I_d|$. While $\Delta|I_d|$ and $FWHM_{High}$, $FWHM_{Low}$ were comparable at $|V_g| = 0.639$V, as $|V_g|$ increased $\Delta|I_d|$ exceeded both $FWHM_{High}$, $FWHM_{Low}$, making it easy to distinguish two current states. Increasing the integration time may be useful to average out the fluctuation around the mean value of each current state.

The characteristic of RTS lifetimes (τ_{High}, time-to-emission and τ_{Low}, time-to-capture, defined in this paper), particularly its dependence on V_g, is considered to reflect the physical origin of the signal[36,37,39]. If the origin of the RTS is a trap in the oxide, the way the average lifetimes ($\langle\tau_{High}\rangle$ and $\langle\tau_{Low}\rangle$) depend on V_g is considered to be asymmetric[36,37,39]. This is because the probability for a carrier to be captured by a trap depends both on the carrier density in the inversion layer and V_g. V_g determines the energy level of a trap with respect to Fermi energy, and as V_g increases the trap level would become lower and it would be predominantly occupied. In addition, if the carrier density is higher, the chance of a carrier to be captured by a trap would increase further. The emission process, on the other hand, only depends on V_g as there is no other electron to be emitted. This difference in capture and emission process causes the difference in behaviour of time-to-capture and time-to-emission against V_g, where time-to-capture strongly depends on V_g while time-to-emission is almost a constant[36,37,39]. Our model is based on a dopant exchanging an electron with n-doped well with fixed density of N_d, suggesting that both $\langle\tau_{High}\rangle$ and $\langle\tau_{Low}\rangle$ should be modified by $|V_g|$. Therefore, by calculating $\langle\tau_{High}\rangle$ and $\langle\tau_{Low}\rangle$ as a function of V_g, the observed RTS can be distinguished from the one caused by a trap in the oxide.

As far as the average values of τ_{High} and τ_{Low} ($\langle\tau_{High}\rangle$ and $\langle\tau_{Low}\rangle$) are concerned, they can be efficiently obtained from the histogram of time differential of $|I_d|$ ($\delta I_d(t) = |I_d(t+\Delta t)| - |I_d(t)|$), $P(\delta I_d)$[33];

$$\langle\tau_{High}\rangle = T_{High} \Big/ \frac{N_{total}}{\delta I_T} \int_{\Delta|I_d|/2}^{\infty} P(\delta I_d) \mathrm{d}\delta I_d \qquad (18)$$

$$\langle\tau_{Low}\rangle = T_{Low} \Big/ \frac{N_{total}}{\delta I_T} \int_{-\infty}^{-\Delta|I_d|/2} P(\delta I_d) \mathrm{d}\delta I_d \,, \qquad (19)$$

where $T_{High} = N_{High}T$, $T_{Low} = N_{Low}T$, N_{total} is the total number of measurement points during T and $\delta I_T = \int_{-\infty}^{\infty} P(\delta I_d)\mathrm{d}\delta I_d$ is a normalisation factor. The denominator

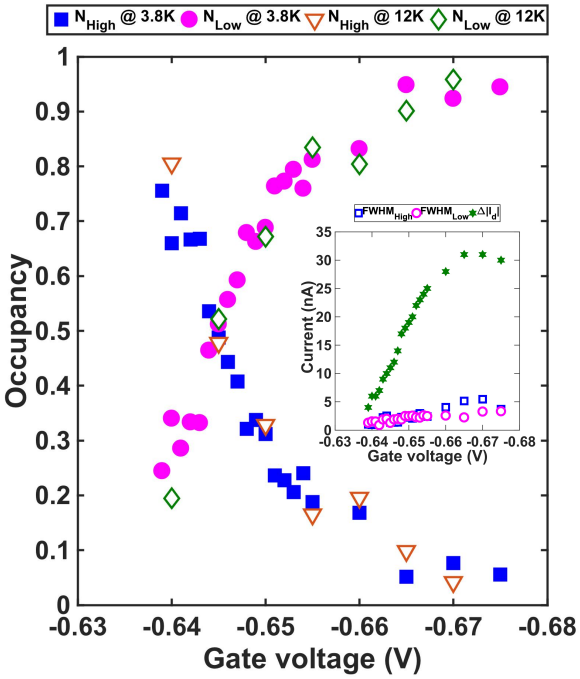

FIG. 11. Occupancy of the two current states (N_{High} and N_{Low}) against V_{g}. N_{High} and N_{Low} at 3.8K were plotted with solid blue square and filled magenta circles, while those at 12K were plotted with triangles with orange lines and diamonds with green lines, respectively. (Inset) Full width of half maximums of the two current states ($\mathrm{FWHM_{High}}$ and $\mathrm{FWHM_{Low}}$) and the amplitude of the random telegraph signals ($\Delta|I_{\mathrm{d}}|$) against V_{g} at 3.8K.

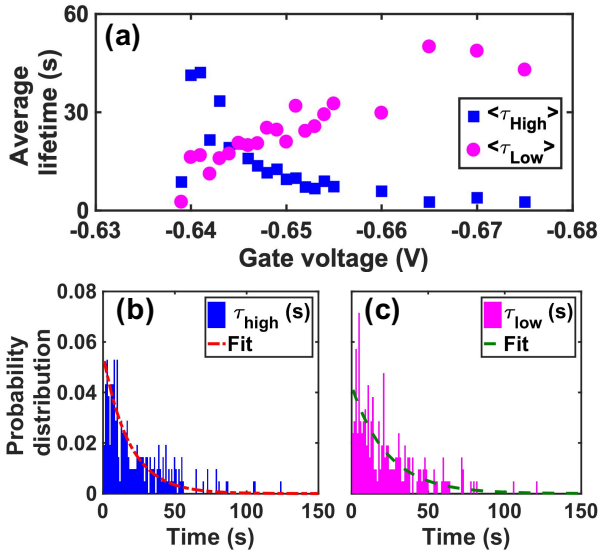

FIG. 12. Study on the lifetime of the RTS observed at 3.8K. (**a**) Average lifetime of high $|I_{\mathrm{d}}|$ state and low $|I_{\mathrm{d}}|$ state ($\langle\tau_{\mathrm{High}}\rangle$ and $\langle\tau_{\mathrm{Low}}\rangle$) against V_{g}. (**b**) and (**c**) Probability distribution of the individual lifetimes of high $|I_{\mathrm{d}}|$ state and low $|I_{\mathrm{d}}|$ state, respectively. The distribution of the data points (the solid blue and magenta bars) follows exponential trend (the red broken line and the green dashed line).

in equation (18) ((19)) is the number of transitions from high (low) $|I_{\mathrm{d}}|$ state to low (high) $|I_{\mathrm{d}}|$ state. Figure 12 (a) displays $\langle\tau_{\mathrm{High}}\rangle$ and $\langle\tau_{\mathrm{Low}}\rangle$ against $|V_{\mathrm{g}}|$. Similar to the N-$|V_{\mathrm{g}}|$ characteristic, two curves ($\langle\tau_{\mathrm{High}}\rangle$-$|V_{\mathrm{g}}|$ and $\langle\tau_{\mathrm{Low}}\rangle$-$|V_{\mathrm{g}}|$) were crossing at $|V_{\mathrm{g}}|$ = 0.645V. While the dependence on $|V_{\mathrm{g}}|$ is different, both $\langle\tau_{\mathrm{High}}\rangle$ and $\langle\tau_{\mathrm{Low}}\rangle$ were modified by $|V_{\mathrm{g}}|$, meaning that the result suggests that the physical origin of the RTS could be a dopant and also different from a trap in the oxide. Further advanced measurement, such as single-electron spin resonance[8] can be a next logical step, which is however out of the scope of this paper.

Figure 12 (b) and (c) are the probability distribution of individual transitions (from the high $|I_{\mathrm{d}}|$ state to the low $|I_{\mathrm{d}}|$ state and from the low $|I_{\mathrm{d}}|$ state to the high $|I_{\mathrm{d}}|$ state, respectively) when $|V_{\mathrm{g}}|$ was fixed to 0.645V. For this particular measurement I_{d} was monitored for 10000s with Δt being 1s so that the longer time trace could be taken without increasing the data points while Δt was short enough to capture the RTS with average lifetimes of 20s. The distributions can both be well approximated by an exponential curve, indicating that there was no periodicity in this signal and also the finite detection bandwidth of the measurement did not affect the statistics of the observed RTS[54].

From the analysis on $\langle\tau_{\mathrm{High}}\rangle$ and $\langle\tau_{\mathrm{Low}}\rangle$ as a function of V_{g}, a dopant in the substrate is considered to be a realistic candidate of the physical origin of the observed RTS. In order to further validate this physical model, finally the device was measured at temperatures up to 25K. RTSs were observed at a similar $|V_{\mathrm{g}}|$ range at higher temperatures as well. The rise in the temperature certainly shifted the threshold voltage, though the statistics of RTS has not been significantly changed, as can be seen from Figure 11. N-$|V_{\mathrm{g}}|$ characteristics at 12K are shown in Figure 11, and similar to the case of 3.8K, N_{High}-$|V_{\mathrm{g}}|$ and N_{Low}-$|V_{\mathrm{g}}|$ curves cross at $|V_{\mathrm{g}}|$ = 0.645V, meaning that both states were observed equally frequently at 12K as well. Then, time domain measurements were taken at $|V_{\mathrm{g}}|$ = 0.645V and different temperatures (5, 7, 10, 11, 12.5, 14, 16, 20 and 25K). Figure 13 shows the average lifetimes of the high and low $|I_{\mathrm{d}}|$ state against inverse of temperature T ($1000/T$), while $|V_{\mathrm{g}}|$ was fixed to 0.645V. This graph indicates that the lifetimes were temperature independent up to 16K, and became faster as the temperature rose to 20 and 25K.

This trend can be understood in a way that the mechanism of the trapping and detrapping of an electron transit from quantum mechanical tunnelling at low temperature to thermal activation at higher temperatures[38];

$$P_{\mathrm{quantum}} = \tau_{\mathrm{q}}^{-1} \tag{20}$$

$$P_{\mathrm{thermal}} = \tau_{\mathrm{th}}^{-1} e^{\frac{-\Delta E}{kT}} \tag{21}$$

$$P_{\mathrm{total}} = \tau_{\mathrm{experiment}}{}^{-1} \tag{22}$$

$$= P_{\mathrm{quantum}} + P_{\mathrm{thermal}} \tag{23}$$

$$= \tau_{\mathrm{q}}^{-1} + \tau_{\mathrm{th}}^{-1} e^{\frac{-\Delta E}{kT}}, \tag{24}$$

where P_{quantum} and P_{thermal} are probability for an electron to become trapped or detrapped via quantum mechanical tun-

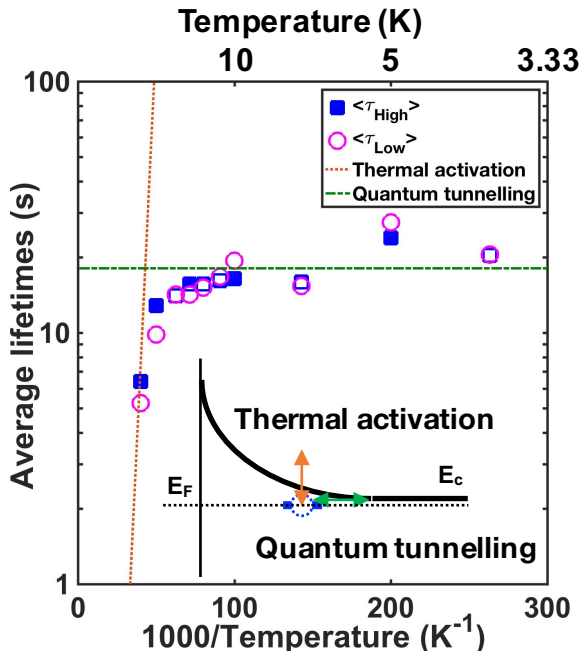

FIG. 13. $\langle\tau_{\mathrm{High}}\rangle$ and $\langle\tau_{\mathrm{Low}}\rangle$ against the inverse of the temperature, 1000/T, at $|V_{\mathrm{g}}| = 0.645$V . $\langle\tau_{\mathrm{High}}\rangle$ are plotted with solid blue squares, while $\langle\tau_{\mathrm{Low}}\rangle$ are plotted with magenta empty circles. The orange dotted line is an Arrhenius plot with an activation energy of 26meV, while the broken green line is the average value of the lifetimes at lower temperature (from 3.8K to 14K). The inset describes the two mechanisms of trapping and detrapping of an electron, quantum mechanical tunnelling and thermal activation, in an energy band diagram.

nelling and thermal activation per 1s, τ_{q} is an attempt interval for the quantum mechanical tunnelling to occur, ΔE is the activation energy for the thermal activation process to occur, τ_{th} is an attempt interval for the thermal activation provided the activation energy is negligibly small, $\tau_{\mathrm{experiment}}$ is experimental data, such as τ_{High} and τ_{Low}. This model means that the total probability of trapping and detrapping of an electron can be described by the sum of the probability of such a carrier exchange to occur via quantum mechanical tunnelling and thermal activation. The inset of Figure 13 describe two mechanisms in an energy band diagram. The orange arrow represents thermal activation, where an electron escapes from a dopant with thermal activation, similar to how free carriers are provided in the conduction band of bulk Si. The green arrow describes the quantum mechanical tunnelling, where an electron tunnel through the potential barrier between a dopant and highly doped region. We used this model to qualitatively explain the temperature dependence of the average lifetimes. The orange dotted line in Figure 13 is an Arrhenius plot with ΔE of 26meV, requirement for a dopant to provide carriers at room temperature, and the green broken line shows the average value of the lifetimes at the temperatures from 3.8K to 14K, 18.04s. These two lines can approximately reproduce the trend observed in Figure 13. This result indicates that the coupling of a single dopant to the electron reservoir (well) at 3.8K is quantum mechanical, indicating that the system in our device can be described by the renowned single impurity Anderson model[55]. Quantum tunnelling rates can also be associated with the analytic solutions in the case of biased double-well systems[56,57].

VII. CONCLUSION

In this paper, observation of random telegraph signals caused by a single dopant in the substrate of a p-type metal-oxide-semiconductor field-effect-transistor was reported. RTSs were initially not observed until the substrate bias was applied such that the depletion layer becomes widened. Trapping and detrapping of an electron changed the depletion layer width under the narrow channel involving a quantum dot, which worked as a sensitive charge sensor. Statistics of individual lifetime of the RTSs obeyed the exponential distribution, and the occupancy of the two current states as well as lifetime were controlled by the gate voltage, as expected. Average lifetimes associated with the discrete current states were modified by gate voltage significantly, indicating that the origin of the signal differs from trap states in the oxide. The temperature dependence of the average lifetime indicates that the tunnelling mechanism transited from quantum mechanical tunnelling to thermal activation as the temperature increased. Engineering of an atomic-scale features and manipulation of a single carrier are crucially important for quantum technology for nanoelectronics[1–4,14–16], metrology[5,10,11,20,21] and even for quantum information processing[6–9,17–19]. We focused on a dopant in the substrate, and realised single-electron manipulations in an industry-grade Si MOSFET in a reasonably easy manner at a relatively higher temperature of 3.8K. Our work proposes an alternative approach to utilise the existing yet to date little considered candidate of such an atomic-scale feature, a solitary dopant in the substrate, for future quantum application.

CONFLICT OF INTEREST STATEMENT

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

AUTHOR CONTRIBUTIONS

K.I, Z.L, M.K.H and S.S prepared samples for measurement. K.I, J.H, F.L, Y.T, H.R and S.S established measurement setup for room temperatures and cryogenic temperatures. K.I, J.H and S.S designed the experiments, and K.I performed measurements. S.S, K.I and I.T proposed the physical model of the observed random telegraph signals. K.I analysed data and drafted manuscript. All authors participated in discussion.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are openly available in ePrints Soton, the University of Southampton Institutional Research Repository (https://doi.org/10.5258/SOTON/D1193)[58].

ACKNOWLEDGMENTS

This work is supported by EPSRC Manufacturing Fellowship (EP/M008975/1), Lloyds Register Foundation International Consortium of Nanotechnology, and the Joint Research Project (e-SI-Amp (15SIB08)). This work is also supported by the European Metrology Programme for Innovation and Research (EMPIR) co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

[1] G. Lansbergen, R. Rahman, C. Wellard, I. Woo, J. Caro, N. Collaert, S. Biesemans, G. Klimeck, L. Hollenberg, and S. Rogge, "Gate-induced quantum-confinement transition of a single dopant atom in a silicon FinFET," Nat. Phys. **4**, 656 (2008).

[2] M. Fuechsle, J. A. Miwa, S. Mahapatra, H. Ryu, S. Lee, O. Warschkow, L. C. Hollenberg, G. Klimeck, and M. Y. Simmons, "A single-atom transistor," Nat. Nanotechnol. **7**, 242–246 (2012).

[3] V. Shorokhov, D. Presnov, S. Amitonov, Y. A. Pashkin, and V. Krupenin, "Single-electron tunneling through an individual arsenic dopant in silicon," Nanoscale **9**, 613–620 (2017).

[4] Z. Durrani, M. Jones, F. Abualnaja, C. Wang, M. Kaestner, S. Lenk, C. Lenk, I. W. Rangelow, and A. Andreev, "Room-temperature single dopant atom quantum dot transistors in silicon, formed by field-emission scanning probe lithography," J. Appl. Phys. **124**, 144502 (2018).

[5] G. Lansbergen, Y. Ono, and A. Fujiwara, "Donor-based single electron pumps with tunable donor binding energy," Nano Lett. **12**, 763–768 (2012).

[6] A. Morello, J. J. Pla, F. A. Zwanenburg, K. W. Chan, K. Y. Tan, H. Huebl, M. Möttönen, C. D. Nugroho, C. Yang, J. A. van Donkelaar, *et al.*, "Single-shot readout of an electron spin in silicon," Nature **467**, 687 (2010).

[7] M. Xiao, I. Martin, and H. Jiang, "Probing the spin state of a single electron trap by random telegraph signal," Phys. Rev. Lett. **91**, 078301 (2003).

[8] M. Xiao, I. Martin, E. Yablonovitch, and H. Jiang, "Electrical detection of the spin resonance of a single electron in a silicon field-effect transistor," Nature **430**, 435 (2004).

[9] K. Ono, G. Giavaras, T. Tanamoto, T. Ohguro, X. Hu, and F. Nori, "Hole spin resonance and spin-orbit coupling in a silicon metal-oxide-semiconductor field-effect transistor," Phys. Rev. Lett. **119**, 156802 (2017).

[10] G. Yamahata, K. Nishiguchi, and A. Fujiwara, "Gigahertz single-trap electron pumps in silicon." Nat. Com. **5**, 5038–5038 (2013).

[11] G. Yamahata, S. P. Giblin, M. Kataoka, T. Karasawa, and A. Fujiwara, "High-accuracy current generation in the nanoampere regime from a silicon single-trap electron pump," Sci. Rep. **7**, 45137 (2017).

[12] T. Tanamoto, Y. Nishi, and K. Ono, "Application of single-electron effects to fingerprints of chips using image recognition algorithms," Appl. Phys. Lett. **115**, 033504 (2019).

[13] J. Li, S. Pud, M. Petrychuk, A. Offenh´usser, and S. Vitusevich, "Sensitivity enhancement of Si nanowire field effect transistor biosensors using single trap phenomena," Nano lett. **14**, 3504–3509 (2014).

[14] H. Grabert and M. H. Devoret, *Single charge tunneling: Coulomb blockade phenomena in nanostructures*, Vol. 294 (Springer US, 1992).

[15] M. Kirihara, K. Nakazato, and M. Wagner, "Hybrid circuit simulator including a model for single electron tunneling devices," Jpn. J. Appl. Phys. **38**, 2028 (1999).

[16] J. Mol, J. Verduijn, R. D. Levine, F. Remacle, and S. Rogge, "Integrated logic circuits using single-atom transistors," Proceedings of the National Academy of Sciences **108**, 13969–13972 (2011).

[17] M. Veldhorst, C. Yang, J. Hwang, W. Huang, J. Dehollain, J. Muhonen, S. Simmons, A. Laucht, F. Hudson, K. M. Itoh, *et al.*, "A two-qubit logic gate in silicon," Nature **526**, 410 (2015).

[18] R. Maurand, X. Jehl, D. Kotekar-Patil, A. Corna, H. Bohuslavskyi, R. Laviéville, L. Hutin, S. Barraud, M. Vinet, M. Sanquer, *et al.*, "A CMOS silicon spin qubit," Nat. Com. **7**, 13575 (2016).

[19] F. A. Zwanenburg, A. S. Dzurak, A. Morello, M. Y. Simmons, L. C. Hollenberg, G. Klimeck, S. Rogge, S. N. Coppersmith, and M. A. Eriksson, "Silicon quantum electronics," Rev. Mod. Phys. **85**, 961 (2013).

[20] S. Giblin, M. Kataoka, J. Fletcher, P. See, T. Janssen, J. Griffiths, G. Jones, I. Farrer, and D. Ritchie, "Towards a quantum representation of the ampere using single electron pumps," Nat. Com. **3**, 930 (2012).

[21] S. Giblin, A. Fujiwara, G. Yamahata, M.-H. Bae, N. Kim, A. Rossi, M. Möttönen, and M. Kataoka, "Evidence for universality of tunable-barrier electron pumps," Metrologia (2019).

[22] G. E. Moore, "Cramming more components onto integrated circuits," Electronics **38**, 114–117 (1965).

[23] P. M. Koenraad and M. E. Flatté, "Single dopants in semiconductors," Nat. Mat. **10**, 91–100 (2011).

[24] Y. Taur and T. H. Ning, *Fundamentals of modern VLSI devices* (Cambridge university press, 2013).

[25] E. Simoen and C. L. Claeys, *Random telegraph signals in semiconductor devices* (IOP Publishing Limited, 2016).

[26] N. Tega, H. Miki, M. Yamaoka, H. Kume, T. Mine, T. Ishida, Y. Mori, R. Yamada, and K. Torii, "Impact of threshold voltage fluctuation due to random telegraph noise on scaled-down SRAM," in *Proc. 46th Ann. Int. Reliab. Phys. Symp. - IRPS08* (IEEE, 2008) pp. 541–546.

[27] H. Kurata, K. Otsuga, A. Kotabe, S. Kajiyama, T. Osabe, Y. Sasago, S. Narumi, K. Tokami, S. Kamohara, and O. Tsuchiya, "Random telegraph signal in flash memory: Its impact on scaling of multilevel flash memory beyond the 90-nm node," IEEE J. Solid-State. Circuits. **42**, 1362–1369 (2007).

[28] X. Wang, P. R. Rao, A. Mierop, and A. J. P. Theuwissen, "Random telegraph signal in CMOS image sensor pixels," in *IEDM Tech. Dig.* (2006) pp. 1–4.

[29] D. Veksler, G. Bersuker, L. Vandelli, A. Padovani, L. Larcher, A. Muraviev, B. Chakrabarti, E. Vogel, D. Gilmer, and P. Kirsch, "Random telegraph noise (RTN) in scaled RRAM devices," in *Proc. 51th Ann. Int. Reliab. Phys. Symp. - IRPS13* (IEEE, 2013) pp. MY–10.

[30] N. Tega, H. Miki, F. Pagette, D. Frank, A. Ray, M. Rooks, W. Haensch, and K. Torii, "Increasing threshold voltage variation due to random telegraph noise in FETs as gate lengths scale to 20 nm," in *Symp. VLSI Technol. Dig. Tech. Papers* (IEEE, 2009) pp. 50–51.

[31] Z. Li, M. K. Husain, H. Yoshimoto, K. Tani, Y. Sasago, D. Hisamoto, J. D. Fletcher, M. Kataoka, Y. Tsuchiya, and S. Saito, "Single carrier trapping and de-trapping in scaled silicon complementary metal-oxide-semiconductor field-effect transistors at low temperatures," Semicond. Sci. Technol. **32**, 075001 (2017).

[32] Z. Li, M. Sotto, F. Liu, M. K. Husain, H. Yoshimoto, Y. Sasago, D. Hisamoto, I. Tomita, Y. Tsuchiya, and S. Saito, "Random telegraph noise from resonant tunnelling at low temperatures," Sci. Rep. **8**, 250 (2018).

[33] F. Liu, K. Ibukuro, M. K. Husain, Z. Li, J. Hillier, I. Tomita, Y. Tsuchiya, H. Rutt, and S. Saito, "Manipulation of random telegraph signals in a silicon nanowire transistor with a triple gate," Nanotechnology **29**, 475201 (2018).

[34] K. Ibukuro, M. K. Husain, L. Zuo, J. Hillier, F. Liu, Y. Tsuchiya, H. N. Rutt, and S. Saito, "Single electron memory effect using random telegraph signals at room temperature," Front. Phys. **7**, 152 (2019).

[35] M. Toledano-Luque, R. Degraeve, P. J. Roussel, L.-Å. Ragnarsson, T. Chiarella, N. Horiguchi, A. Mocuta, and A. Thean, "Fast ramped voltage characterization of single trap bias and temperature impact on time-dependent Vth variability," IEEE Trans. Elec. Dev. **61**, 3139–3144 (2014).

[36] H. Miki, M. Yamaoka, N. Tega, Z. Ren, M. Kobayashi, C. D'Emic, Y. Zhu, D. Frank, M. Guillorn, D.-G. Park, *et al.*, "Understanding short-term BTI behavior through comprehensive observation of gate-voltage dependence of RTN in highly scaled high-κ/metal-gate pFETs," in *Symp. VLSI Technol. Dig. Tech. Papers* (IEEE, 2011) pp. 148–149.

[37] A. Yonezawa, A. Teramoto, T. Obara, R. Kuroda, S. Sugawa, and T. Ohmi, "The study of time constant analysis in random telegraph noise at the sub-

threshold voltage region," in *Proc. 51th Ann. Int. Reliab. Phys. Symp. - IRPS13* (IEEE, 2013) pp. XT–11.
[38]J. H. Scofield, N. Borland, and D. Fleetwood, "Temperature-independent switching rates for a random telegraph signal in a silicon metal–oxide–semiconductor field-effect transistor at low temperatures," Appl. Phys. Lett. **76**, 3248–3250 (2000).
[39]E. Simoen and C. Claeys, "Substrate bias effect on the random telegraph signal parameters in submicrometer silicon p–metal–oxide–semiconductor transistors," J. App. Phys. **77**, 910–914 (1995).
[40]K. Ono, T. Tanamoto, and T. Ohguro, "Pseudosymmetric bias and correct estimation of coulomb/confinement energy for unintentional quantum dot in channel of metal-oxide-semiconductor field-effect transistor," Appl. Phys. Lett. **103**, 183107 (2013).
[41]S. Takagi, A. Toriumi, M. Iwase, and H. Tango, "On the universality of inversion layer mobility in Si MOSFET's: Part I - Effects of substrate impurity concentration," IEEE Trans. Elec. Dev. **41**, 2357–2362 (1994).
[42]R. Van Langevelde and F. M. Klaassen, "Effect of gate-field dependent mobility degradation on distortion analysis in MOSFETs," IEEE Trans. Elec. Dev. **44**, 2044–2052 (1997).
[43]H. Irie and A. Toriumi, "Advanced split-CV technique for accurate extraction of inversion layer mobility in short channel MOSFETs," in *Extended abstracts of the 2005 International Conference on Solid State Devices and Materials, Kobe, Japan* (2005) pp. 864–865.
[44]S. Sun and J. D. Plummer, "Electron mobility in inversion and accumulation layers on thermally oxidized silicon surfaces," IEEE Trans. Elec. Dev. **27**, 1497–1508 (1980).
[45]S. Saito, K. Torii, M. Hiratani, and T. Onai, "Analytical quantum mechanical model for accumulation capacitance of MOS structures," IEEE Elec. Dev. Lett. **23**, 348–350 (2002).
[46]T. Ando, A. B. Fowler, and F. Stern, "Electronic properties of two-dimensional systems," Rev. Mod. Phys. **54**, 437 (1982).
[47]K. Nishiguchi and A. Fujiwara, "Single-electron counting statistics and its circuit application in nanoscale field-effect transistors at room temperature," Nanotechnology **20**, 175201 (2009).
[48]K. Nishiguchi, Y. Ono, and A. Fujiwara, "Single-electron counting statistics of shot noise in nanowire si metal-oxide-semiconductor field-effect transistors," Appl. Phys. Lett. **98**, 193502 (2011).
[49]K. Nishiguchi and A. Fujiwara, "Single-electron stochastic resonance using si nanowire transistors," Jpn. J. Appl. Phys. **50**, 06GF04 (2011).
[50]C. Beenakker and M. Büttiker, "Suppression of shot noise in metallic diffusive conductors," Phys. Rev. B **46**, 1889 (1992).
[51]J. B. Johnson, "Thermal agitation of electricity in conductors," Phys. Rev. **32**, 97 (1928).
[52]H. Nyquist, "Thermal agitation of electric charge in conductors," Phys. Rev. **32**, 110 (1928).
[53]D. K. Schroder and J. A. Babcock, "Negative bias temperature instability: Road to cross in deep submicron silicon semiconductor manufacturing," J. Appl. Phys. **94**, 1–18 (2003).
[54]M. Jenei, E. Potanina, R. Zhao, K. Y. Tan, A. Rossi, T. Tanttu, K. W. Chan, V. Sevriuk, M. Möttönen, and A. Dzurak, "Waiting time distributions in a two-level fluctuator coupled to a superconducting charge detector," Phys. Rev. Research **1**, 033163 (2019).
[55]P. W. Anderson, "Localized magnetic states in metals," Phys. Rev. **124**, 41 (1961).
[56]M. P. Fisher and A. T. Dorsey, "Dissipative quantum tunneling in a biased double-well system at finite temperatures," Phys. Rev. Lett. **54**, 1609 (1985).
[57]H. Grabert and U. Weiss, "Quantum tunneling rates for asymmetric double-well systems with ohmic dissipation," Phys. Rev. Lett. **54**, 1605 (1985).
[58]K. Ibukuro, J. W. Hillier, F. Liu, M. K. Husain, L. Zuo, I. Tomita, Y. Tsuchiya, H. N. Rutt, and S. Saito, "Dataset for random telegraph signals caused by a single dopant in a metal oxide semiconductor field effect transistor at low temperature," `https://doi.org/10.5258/SOTON/D1193`, (University of Southampton, 2020).