Error mechanisms and rates in tunable-barrier single-electron turnstiles and charge-coupled devices

Neil M. Zimmerman[a] and Emmanouel Hourdakis
National Institute of Standards and Technology, Electricity Division, Electronics and Electrical Engineering Laboratory, Technology Administration, U.S. Department of Commerce, Gaithersburg Maryland 20899

Yuki Ono, Akira Fujiwara, and Yasuo Takahashi
NTT Basic Research Laboratories, 3-1, Morinosato Wakamiya, Atsugi-shi 243-0198, Japan

(Received 12 April 2004; accepted 19 July 2004)

Si-based single-electron tunneling (SET) devices have of late become an important alternative to the metal-based ones, both for ultralarge scale integration (ULSI) electronics and for electrical metrology. We have very recently been designing, fabricating, and measuring SET turnstiles, pumps, and charge-coupled devices using tunable barriers in silicon. Having shown the potential of these devices, we wish to understand the error mechanisms which may manifest themselves, and to predict the level of these errors, in order to decide how feasible these devices will be. In this paper, we devote a substantial amount of analysis to the consideration of the "dynamical" error mechanism. This particular error considers how electrons split up as the barrier is raised, or alternatively how the Coulomb blockade is formed. We then consider a wide variety of other errors, including thermal, frequency, leakage, and heating errors. We show the dependence of the error rate on each of those mechanisms, and predict maxima or minima for the corresponding parameters. In the conclusion, we discuss the various advantages Si-based turnstiles or pumps would offer with respect to the metal-based ones. © *2004 American Institute of Physics*. [DOI: 10.1063/1.1791758]

I. INTRODUCTION AND MOTIVATION

In the past decade, single-electron tunneling (SET) devices[1] have been proposed for applications in both integrated microelectronics and in electrical metrology. In both cases, these applications often depend on the ability of SET pumps or turnstiles to move charge in units of just one e. For instance, for use in logic circuits,[2] the common element is a memory node which has a counted number of electrons placed onto it. In metrology, the obvious application is as a current source, with a value of $I=ef$; here, I is the current, e is the electron's charge, and f is the frequency with which electrons are clocked through the device. In both of these fields of applications, it is clear that one very important characteristic of the device is the rate of errors; in particular, we are most interested in the probability that in each cycle, the device passes more than or less than the exact number of electrons desired.

For example, the conventional acceptable relative bit error rate for an ultralarge scale integration (ULSI) chip is 10^{-10} during an operational period of 10 years.[3] Given a typical conservative estimate for clock rate of 100 MHz, and the number of circuit elements as 100 000, this yields a limitation on the bit error rate per device of less than 10^{-31}![3] The case where this is most limiting for SET-based logic is for "charge bits,"[2] where one bit is represented by only a single electron; in this case, we require a probability of error per cycle of the charge source that is less than this number. Clearly, a theoretical investigation of the mechanisms and values of errors will be important if SET devices are ever to be used in ULSI microelectronics.

Applications of SET devices in electrical metrology are less exacting; the typical desired relative error rate is less than 10^{-8}. Since there has already been a substantial amount of work, both theoretical and experimental, on errors in SET pumps used for metrology, we will devote some length in this introduction to that field. We note in advance that most of this detailed work has been on devices based on metal oxide tunnel junctions, and is thus not directly applicable to Si-based devices.

A. Metrological aplications

By far the best reported work is the result that in SET pumps made with Al/AlO_x tunnel junctions, the error when shuttling one electron forward and then back repeatedly can be as low as about 10^{-8}.[4] However, because one source of errors comes from running the pump too fast, the frequency is limited to a few tens of MHz; this corresponds to a maximum current of a few picoamperes. This value of current is too small to be useful as a direct current standard; instead, NIST has pursued a capacitance standard whose basis is the measurement of the voltage across a capacitor when a counted number of electrons is placed on to a capacitor plate.[5]

For a variety of reasons, it would be interesting and useful to have a current standard with a much larger value of current but still based on the fundamental charge of the electron. One possible approach to this is similar to the metal pumps and turnstiles, but using Si-based materials. There are

[a]301-975-5887; electronic mail: neilz@mailaps.org: http://www.eeel.nist.gov/811/femg/set.html

two major potential advantages to this alternative with respect to the metal devices: one is that the value of the capacitance in the Si devices is typically a factor of 10 or more smaller; this increases the maximum speed of the device by the same ratio, and thus the maximum current. The second is that it may be possible to parallelize a large number of pumps together. However, a major conceptual difficulty with this approach is the long-term charge offset drift in SET devices.[6] Recently, we have shown that at least in one class of Si-based SET devices, the long-term offset drift is at least a factor of 1000 smaller, so that it would no longer pose a problem to the parallelization of a large number of SET pumps.[7] Thus, the higher speed of one Si device, and the potential of parallelization, are both motivations for the pursuit of Si-based SET charge standards [turnstiles, pumps, or charge-coupled devices (CCDs)], in the hope of making a large-value current standard.

B. Devices based on tunable barriers

Recently, we have been experimentally pursuing a variety of Si-based SET pumps, turnstiles, and CCD devices.[8–11] For some of these devices, there are tunnel junctions defined by fabrication, whose resistance values are fixed; this is the typical means by which we make the pumps and turnstiles. For the CCD-type devices, the tunnel junctions have tunable resistance values. For both classes of devices, however, the barrier between lead and islands is always tuned by means of a gate which is used to deplete certain regions of the silicon channel.[8–10] This tunable barrier is a common feature of all the devices which we have pursued.

For the purpose of this paper, we define these terms as follows (see Fig. 1): Each of these devices has a source-drain voltage U, an island potential V_{isl}, and gate voltages applied to form the tunnel barriers V_{G}. A turnstile is a Coulomb blockade-based device which has a nonzero U, but a time-independent V_{isl}; the resistances of the tunnel barriers are always well above the resistance quantum $R_{\mathrm{Q}}=h/e^2$, and are modulated in time by varying V_{G}. A pump is similar, but with $U=0$, and with the addition of control of the island potential via a time-dependent V_{isl}. A CCD is also similar to the turnstile, except that the tunnel barriers will have resistances well below R_{Q} during the parts of the cycle when electrons are being transferred.

There has been previous experimental work on tunable-barrier devices, although not in silicon. The first work[12] demonstrated current plateaus which depended on a microwave frequency. A second publication[13] demonstrated a device with improved fabrication, leading to an error of about 10^{-3}. In a similar device with an optically driven SET pump,[14] the error rate was about 10^{-4}.

There has also been a fair amount of previous theoretical work on tunable-barrier devices. The most relevant work[15] was intended for mechanical single-electron shuttles. It considers a master equation approach; the main results is that the contact time (when the barrier resistance is small) should be long compared to the RC time for optimum operation. The analysis in this present work fits in a similar theoretical framework. Other work includes: In one paper, the oscillating barrier gives rise to photon-assisted tunneling and sidebands.[16] However, the effect of these is suppressed in devices with a sufficiently large Coulomb blockade energy which is generally numerically true in our devices. Much of the theoretical work has been concerned with blockade in mesoscopic devices, where the phase coherence of the electron is important; this work is generally based on Ref. 17. In our devices, the phase coherence is not important, so these effects do not manifest themselves. Finally, in an analysis of a device with two barriers that have sinusoidal modulations,[18] the authors showed that the phase where the electron tunnels is not the phase where the barrier is lowest. While interesting, this is not directly applicable to our devices, where the barriers are kept at a fixed low value for a nonzero time.

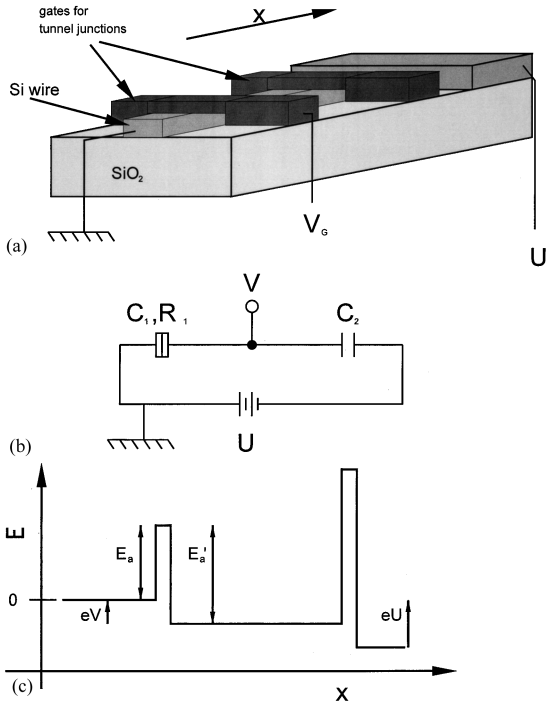

FIG. 1. (a) Physical sketch of gates and wire for turnstile, or CCD-type device. The upper gate is not shown, for clarity. (b) Simple circuit with one tunnel junction and one blocking capacitor. (c) Energy diagram for the upper part of the circuit in (b), showing the potential of the island, and the activation barrier E_{a}. For many devices, the capacitive coupling of the upper gate to the island is much larger than from the source or drain, so that V is defined by the voltage of the upper gate.

C. Statement of the problem

In this paper, we attempt to analyze in detail the error mechanisms and error rates in the various phases of operation of the turnstiles and CCDs (we leave the consideration of pumps, which are more complicated, to subsequent work), with a hope that will lead us towards schemes to optimize the performance.

In Fig. 1, we show a generic device which will form part of the conceptual framework for this paper. The devices are typically made using silicon-on-insulator wafers, and have a silicon wire running underneath the various gates, which will both deplete to form tunnel junctions (shown) as well as invert the silicon wire (not shown). Generally, only one of

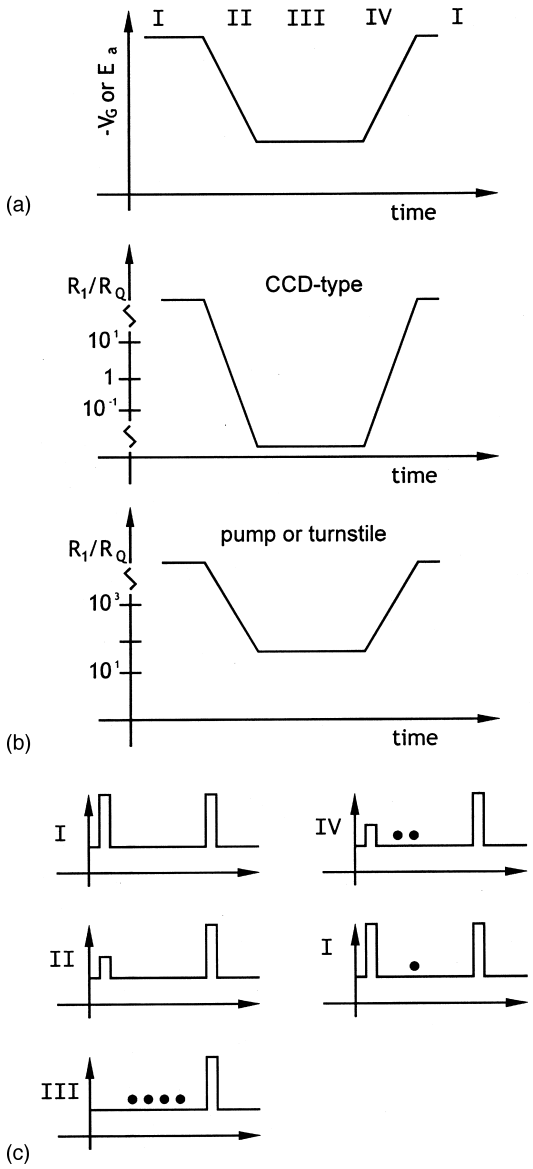

FIG. 2. Illustration of the time dependence of various quantities, as the barrier for the first gate is lowered and then raised. Note that one complete cycle of the full device would include a similar set of steps for the second barrier. (a) Sketch of the gate voltage applied to the barrier region vs time. Note that, for a device that has electron transport, a more negative gate voltage impedes such transport, and ultimately forms a tunnel barrier. (b) Top: Time dependence of the resistance of the tunnel junction vs time for the CCD-type device; note that the value goes well below R_Q when the gate voltage is low. Bottom: Time dependence of the resistance of the tunnel junction vs time for the pump or turnstile; note that the value always stays well above R_Q. (c) Configuration of barriers, and charges on island, as a function of time. In phase III, with the barrier low, the number of charges on the island is indeterminate. In phase IV, as the barrier is rising, the number of charges is moving towards the energetically favored number of one. In phase I again, the number of charges is the desired number; errors occur if a number other than this is "locked in" because the barrier is raised too quickly.

the two (or multiple) junctions is conducting at any time; thus, we show an electrical circuit with one tunnel junction having both capacitance and resistance, and the other junction being only a pure capacitor. In the bottom section of this figure, we show a generic energy diagram for the same device, with the activation barrier E_a, bias voltage U, and voltage difference V between the lead and the island. As described in the caption, the island potential V can arise from either the source-drain bias voltage or from a capacitively-coupled voltage from the upper gate, depending on the relative capacitances; usually, the upper gate capacitance dominates. For the latter case, we can see the relationship between the two voltages,

$$V = \frac{C_{UG}}{C_\Sigma} U_{UG} - eN/C_\Sigma, \tag{1}$$

where C_{UG} is the capacitance between the upper gate and the island, C_Σ is the total island capacitance, U_{UG} is the voltage on the upper gate, and N is the number of extra electrons on the island.

In Fig. 2, we show the typical time dependence for one of the two gates. During phase I, the barrier and thus the resistance of this junction is very high; this would be either the resting phase of the device, or the phase during which the other of the two barriers is active. During phases II, III, and IV, this junction is active, with the amplitude of the activation barrier being decreased substantially. For the CCD-type device, as shown in the top panel of part B, the resistance of the junction goes to a value small compared to R_Q. In contrast, for the pump or turnstile as shown in the bottom panel, the resistance of the junction stays large compared to the resistance quantum when the gate voltage or activation barrier is at a minimum. Part C shows the time dependence of the number of charges: the number is indeterminate in phase III, and reaches the desired number, 1, in the last phase. Errors occur when the number in the final phase is wrong, because the wrong number became "locked in" during phase IV, as the barrier is rising.

In the remainder of this paper, we wish to analyze and calculate the various error mechanisms and their concomitant rates. Much of this work concentrates on the errors in phase IV, which we will call the "dynamical error." We analyze the dynamical error in some detail, and develop recommendations for optimizing the behavior (in particular, being able to run the device as fast as possible). The analysis in the other phases depends on simpler arguments, some of which come from previous work by others. We compile all of these errors, in order to give a list of mechanisms, error rates, and limits on operating parameters such as frequency, voltage, and temperature.

II. THEORY OF DYNAMICAL ERRORS

In this section, we will develop the formalism necessary to analyze one specific part of the barrier height cycle. This part is the raising of the barrier, denoted as phase IV in Fig. 2; we will call errors occurring during this part of the cycle dynamical errors. Although this section is general, some of the specific mechanisms we develop specifically for the CCD-type device, where the resistance between the island and the outside starts off much less than R_Q.

Our general approach will be using a master equation; in the following few sections, we will develop the rates for various error mechanisms between the desired state and un-

desired states. Following that, we will develop the master equation approach, and examine some limiting approximations.

A. Combining thermal hopping and quantum mechanical tunneling

1. Thermal over-the-barrier hopping and Coulomb blockade

In phase IV, the energy barrier between the island and the lead is changing from very small to large. Because of this, during this phase there can be substantial motion due to thermal excitation of carriers over the barrier.

The combination of thermal over-the-barrier hopping and quantum mechanical tunneling, in the context of the Coulomb blockade, has not been considered in detail to date. There has been one theoretical prediction that a SET transistor, for a particular set of conditions, will have a current which depends on gate voltage with thermal hopping, just as for tunneling.[19] There has also been one experiment which seemed to confirm the theory, although it did not obey all the conditions specified in the theory.[20]

There has been no *a priori* theory developed for single charge transfer in the context of both thermal hopping and quantum mechanical tunneling. It is clear that, as the experimental development of tunable-barrier charge sources continues, the development of such a fundamental theory is becoming of greater importance.

In the absence of such a theory, we wish to develop a phenomenological approach that allows us to estimate the rate of desired and undesired (error) charge transfers in the context of both thermal hopping and tunneling. To do so, we will first estimate the resistance of a tunnel junction due to both hopping and tunneling, and then given this resistance as the single time-dependent parameter, use the standard theory of Coulomb blockade to calculate the rate of charge transfer.

2. Resistance R versus gate voltage V_G for both thermal hopping and quantum mechanical tunneling

We must note that, in actuality, there is no linear resistance for thermal hopping, where the rate of transfers depends exponentially on the height of the barrier. This will be important in the context of SET turnstiles or CCDs, where there is a nonzero bias voltage across the junction. However, in the absence of a fundamental theory, we will use one simple way to estimate the resistance of a junction for both hopping and tunneling, following a standard linear approximation:[19]

$$R = G^{-1} = \frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\left\{\int_{-\infty}^{+\infty} d\epsilon\left[\frac{-\delta n(\epsilon)}{\delta\epsilon}\right]\mathcal{T}(\epsilon)\right\}^{-1}. \tag{2}$$

Here, ϵ is the energy of the incoming electron, $\mathcal{T}(\epsilon)$ is the transmission of the junction, n is the Fermi function, and $N_{\text{chann}} \approx wtk_F^2$ is the number of transverse states in the channel; w, t, k_F are the width of the silicon wire, thickness of the inversion layer, and the Fermi wave vector, respectively. The simplest approximation we can make to obtain the transmission is to assume a square barrier of height E_a; then, using the WKB approximation,

$$\begin{aligned}\mathcal{T} &= 1, \quad \epsilon > E_a \\ &= e^{-[(E_a-\epsilon)/V_G^0]^{1/2}}, \quad \epsilon < E_a,\end{aligned} \tag{3}$$

where $V_G^0 = \hbar^2/2mL^2$, and m, L are the carrier mass and thickness of the barrier, respectively. Here, the first line corresponds to thermal over-the-barrier hopping, and the second line to tunneling. For an order of magnitude estimate, if we assume the mass of free electrons, $L = 0.1\ \mu$m, we obtain $V_G^0 = 40\ \mu\text{eV} \approx 0.4\ \text{K} \times k_B$ (V_G^0 is so small because L is so long). From Eq. (2), we get

$$\begin{aligned}R &= \frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\left[\int_0^{E_a} d\epsilon\left(\frac{-dn(\epsilon)}{d\epsilon}\right)e^{-[(E_a-\epsilon)/V_G^0]^{1/2}} + \int_{E_a}^{+\infty} d\epsilon\left(\frac{-dn(\epsilon)}{d\epsilon}\right)\right]^{-1} \\ &= \frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\left[\int_0^{E_a} d\epsilon\left(\frac{-dn(\epsilon)}{d\epsilon}\right)e^{-[(E_a-\epsilon)/V_G^0]^{1/2}} - n(\infty) + n(E_a)\right]^{-1} \\ &= \frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\left[\int_0^{E_a} d\epsilon\left(\frac{-dn(\epsilon)}{d\epsilon}\right)e^{-[(E_a-\epsilon)/V_G^0]^{1/2}} + \frac{1}{1+e^{E_a/kT}}\right]^{-1}.\end{aligned} \tag{4}$$

This simplifies in several limits for T and E_a:

$$\begin{aligned}R &= \left(\frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\right)2, \quad E_a \ll kT \\ &= \left(\frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\right)e^{E_a/kT}, \quad kT \ll E_a \ll kT\frac{kT}{eV_G^0} \\ &= \left(\frac{1}{N_{\text{chann}}}\frac{h}{2e^2}\right)2e^{\sqrt{(E_a/V_G^0)}}, \quad E_a \gg kT\frac{kT}{eV_G^0}.\end{aligned} \tag{5}$$

The middle line shows the temperature-dependent resistance corresponding to thermal activation, and the bottom line shows the temperature-independent tunneling resistance.

3. Resistance R versus gate voltage V_G, Including leakage

There is clearly an upper limit to the resistance as a function of the gate voltage. This does not affect any of the numerical results in calculating the effects of the dynamical error; we include it to avoid the unphysical limit of an infinite resistance. For instance, the theory predicts that if E_a is 0.1 V, $R \approx 10^{22}\ \Omega$. This limit is clearly the leak resistance R_{leak}, which is the maximum resistance of an electrostatically-produced tunnel junction; in Ref. 21, the leakage resistance was estimated to be $10^{20}\ \Omega$, from measuring a leakage time of greater than 10 000 s. Thus, we obtain the gate voltage-dependent resistance of

$$R(V_G) = \frac{RR_{\text{leak}}}{R + R_{\text{leak}}}, \tag{6}$$

where R is the value from Eq. (4).

Now, having calculated $R(V_G)$, we can use this single time-dependent parameter to calculate the rates of motion.

B. Calculation of rates of motion

We wish to consider the error mechanism in phase IV that comes from number fluctuations on the island. In phase III and the beginning of phase IV, as seen in Fig. 2(b), it is clear that because R is much less than R_Q, the number of carriers on the island is not a good quantum number. This means that there will be an extra error mechanism which comes from the increased probability that these number fluctuations will cause an unwanted result: the number of electrons on the island will end up (after the barrier is raised) at a value other than the desired one.

In order to evaluate this increased probability, we need to obtain two results: first, we need to calculate the rates to change the number of electrons on the island by one, as a function of the time (or equivalently, of the barrier height). For example, Γ^{01} is the rate for a change from zero electrons to one. We desire a calculation of the rates over a range of resistances R from much less than R_Q to much greater than R_Q. Second, we will use these time-dependent rates in a master equation to calculate the probabilities of ending up with zero, one,... electrons on the island; the sum of the probabilities of other than one electron will be the error.

1. Rates of motion for small and large barriers

Unfortunately, there is no simple calculation that contains the rates of motion over the range from much less than R_Q to much greater than R_Q. There have been nonperturbative calculations over the entire range, but these only compute the renormalized Coulomb blockade energy,[22,23] not the rate of motion.

Thus, we wish to obtain an approximate relation for this rate which agrees in the limits of small and large resistance with the results of previous perturbative calculations. It is clear that the treatment given in this section is phenomenological; a more fundamental, *a priori* approach would be desired in the future. Two results which help us are as follows.

Weak tunneling ($R \gg R_Q$). The "orthodox" equation for the tunneling rate is[1]

$$\Gamma = \frac{1}{e^2 R} \frac{\Delta E}{1 - e^{-\beta \Delta E}}. \tag{7}$$

Here, we explicitly include the Coulomb blockade by using the expression for the energy change which includes this contribution:[24]

$$\Delta E^{\pm} = \frac{\mathrm{e}}{C_\Sigma}(-\mathrm{e}/2 \mp Ne \pm C_2 U), \tag{8}$$

where $C_\Sigma = C_1 + C_2$, and where we have neglected stray background charge; also, the ± refers to the addition or subtraction of one electron from the island. Because of the form of the equation, the rate is exponentially suppressed for motion in the "wrong" direction (i.e., towards a higher energy state). Also, in the context of pumping electrons, it is clear that U will be set at a particular value. We will define the desired state as having one excess electron ($N=1$), so that often we will want $U = e/C_2$; this makes $\Delta E^{\pm}$ most negative, and thus makes motion out of the desired state unlikely. More generally, we will restrict $|U - e/C_2| < e/2C_2$.

Strong tunneling ($R \ll R_Q$). We are interested in the fluctuations in the charge due to the fact that N is no longer a good quantum number. We argue on physical grounds that, in this limit, the leading term for the rate of motion in and out should simply be $1/RC$; here, we are explicitly assuming that at any instant in time, the number of electrons on the island can change randomly, and that the time required to do this is simply the relaxation time RC. Thus, we conclude that in this limit the rate can be written as

$$\Gamma = \frac{1}{RC}. \tag{9}$$

We wish to obtain a function which smoothly approximates the rates in both limits of R. Using a simple exponential smoothing function with a crossover at R_Q,[25] we obtain a rate for the number of fluctuations which is

$$\Gamma = \frac{1}{e^2 R} \frac{\Delta E}{1 - e^{-\beta \Delta E}} \frac{1 - e^{-R/R_Q}}{1 + e^{-R/R_Q}} + \frac{1}{RC} \frac{2e^{-R/R_Q}}{1 + e^{-R/R_Q}}. \tag{10}$$

2. Energy barrier E_a versus gate voltage V_G

We wish to obtain the energy barriers E_a and E_a' as portrayed in Fig. 1(c). The parameter which is controlled in the experiment is V_G, and this is the quantity we wish to use to parametrize the equations. In order to do so, we need to obtain the energy barrier E_a in terms of V_G.

We use standard results from the silicon microelectronics field for this purpose. In particular, it is well known that in Si/SiO_2 metal-oxide-semiconductor capacitors, the surface potential ψ_S closely follows the gate voltage V_G;[26] in fact, for a wide range of gate voltages in the direction of accumulation (larger tunneling barrier), the surface potential is approximately equal to the gate voltage for voltages small compared to the band gap (1.1 eV). Thus, we will make the simplifying approximation that $E_a = |eV_G|$ and $E_a' = |eV_G| + e\mathrm{V}$, where we have assumed that the gate voltage is measured with respect to the flatband voltage in the left lead.

3. Total rate of motion Γ

Finally, having developed all of the above formalism and approximations, we can now derive the rate of motion, as a function of the gate voltage or barrier resistance. We combine Eqs. (4), (10), and (6) to obtain (for motion onto the island)

TABLE I. List of activation barriers E_a and energy changes ΔE for various conditions. (a) shows the energies to move toward the desired state, (b) shows the energies to move away from the desired state. In each part, the upper set of quantities are for small gate voltages, and the lower set for large gate voltages. As shown in the figures and discussed in the text, it is sufficient to consider only the smaller range.

		U	Activation barrier E_a	$0->1$ Change ΔE	Activation barrier E_a	$2->1$ Change ΔE
(a)	$eV_G<e^2/C_\Sigma\|C_2U/e-2\|$	general $U=e/C_2$	$\|eV_G\|$	ΔE^+ $+e^2/2C_\Sigma$	CCD (no barrier)	
	$eV_G>e^2/C_\Sigma\|C_2U/e-2\|$	general $U=e/C_2$	$\|eV_G\|$	ΔE^+ $+e^2/2C_\Sigma$	$\|eV_G\|+eV_{N=2}$ $\|eV_G\|-e^2/C_\Sigma$	ΔE^- $+e^2/2C_\Sigma$
		U	Activation barrier E_a	$1->0$ Change ΔE	Activation barrier E_a	$1->2$ Change ΔE
(b)	$eV_G<e^2/C_\Sigma\|C_2\ U/e-2\|$	general $U=e/C_2$	$\|eV_G\|+eV_{N=1}$ $\|eV_G\|$	ΔE^- $-e^2/2C_\Sigma$	$\|eV\|_{N=2}$ e^2/C_Σ	impossible
	$eV_G>e^2/C_\Sigma\|C_2\ U/e-2\|$	general $U=e/C_2$	$\|eV_G\|+eV_{N=1}$ $\|eV_G\|$	ΔE^- $-e^2/2C_\Sigma$	$\|eV_G\|$ $\|eV_G\|$	ΔE^+ $-e^2/2C_\Sigma$

$$\Gamma(V_G)=\frac{1}{e^2R(V_G)}\frac{\Delta E^+}{1-e^{-\beta\Delta E+}}\frac{1-e^{-R(V_G)/R_Q}}{1+e^{-R(V_G)/R_Q}}+\frac{1}{R(V_G)C_\Sigma}\frac{2e^{-R(V_G)/R_Q}}{1+e^{-R(V_G)/R_Q}},\tag{11}$$

where $R(V_G)$ is given by Eqs. (4) and (6), and ΔE by Eq. (8). For purposes of discussion, we label the two terms in this equation as Γ_3 and Γ_4. For motion off the island, we would replace ΔE^+ by ΔE^-.

The formalism developed above is clearly not an *a priori* theory, and development of such a fundamental theory would clearly be desirable. However, we believe that this relationship for the dynamical error rate captures the essential physics necessary to analyze this error mechanism.

4. Barriers for specific situations

We need to examine the activation barrier and energy change for motion onto and off of the island. In Fig. 3, we show the potential energy diagrams for $U_{UG}=e/C_{UG}$; here, we have now specifically considered the gate (denoted "UG," or upper gate) voltage and capacitance, because this capacitance is usually much bigger than the capacitance to the drain. This choice of the upper gate voltage makes the $N=1$ state optimally preferred, which one can see by substituting in Eq. (8). Here, we have used the result that

$$V=\frac{C_{UG}}{C_\Sigma}U_{UG}-eN/C_\Sigma=e/C_\Sigma-eN/C_\Sigma=(1-N)e/C_\Sigma.\tag{12}$$

We note that in this figure and in Table I, we have assumed that (1) the barrier does not change as the electron moves; (2) for $N=1$, the appropriate energy barrier to use is with V as in Eq. (12). This latter assumption means that we are suppressing consideration of dynamic effects such as the image charge[27,28] on the barrier.

We note one interesting thing from this figure: when the barrier is smaller than the Coulomb blockade energy, $|eV_G|<e^2/C_\Sigma$, there is no stable state for $N=2$ or higher. This is because the potential of the island rises above the top of the activation barrier in this case. It is not clear exactly what this means in reality; for instance, does the potential of the island truly rise by the Coulomb blockade energy when it is higher than the barrier? This is an interesting regime which, to our knowledge, has not been investigated either experimentally or theoretically. For the purposes of this paper, we will simply assume that the barrier is as drawn.

Table I shows the corresponding activation barrier and energy E_a and $\Delta E^\pm$, for motion between the three states we will consider ($N=0,1,$ or 2).

We show in Figs. 4 and 5 some simple examples of the rates defined in Eq. (11). These graphs show the resistance and the rates as a function of gate voltage when the potential U is set such that $C_{UG}U_{UG}/e=1$; as mentioned above, this is the condition in which it is optimally favorable for one electron to tunnel onto the island due to the Coulomb blockade.

In these graphs, we have chosen the temperature so that the exponential factor in Eq. (7) is small enough such that the error rate Γ_3^{10} is small compared to the desired rate Γ_3^{01}. This is one example of a general feature: what we desire is that the forward rate Γ_{tot}^{01} is large compared to the backward or error rate Γ_{tot}^{10}.

We are interested in the probability that the number of electrons on the island at the end of phase IV is not the correct number; what we ultimately desire is that this probability is less than some relative error rate ϵ_0. One way of quantifying this is to note that a minimal condition for achieving a relative error rate no larger than ϵ_0 is that the rates for the undesired motion be at most ϵ_0 times the rates for the desired motion. If we set $\epsilon_0=10^{-8}$, we can see that we need $-\Delta E^-> 20kT$, or $e^2/C_\Sigma>40kT$. Note that this condition may change somewhat if we operate at a voltage other than the optimum, $U=e/C_2$.

We note that the important region of change of the gate voltage is at low voltages, where the resistance is not much greater than R_Q. In particular, our estimate for the rate of number fluctuations for a low barrier Γ_4 falls to an accept-

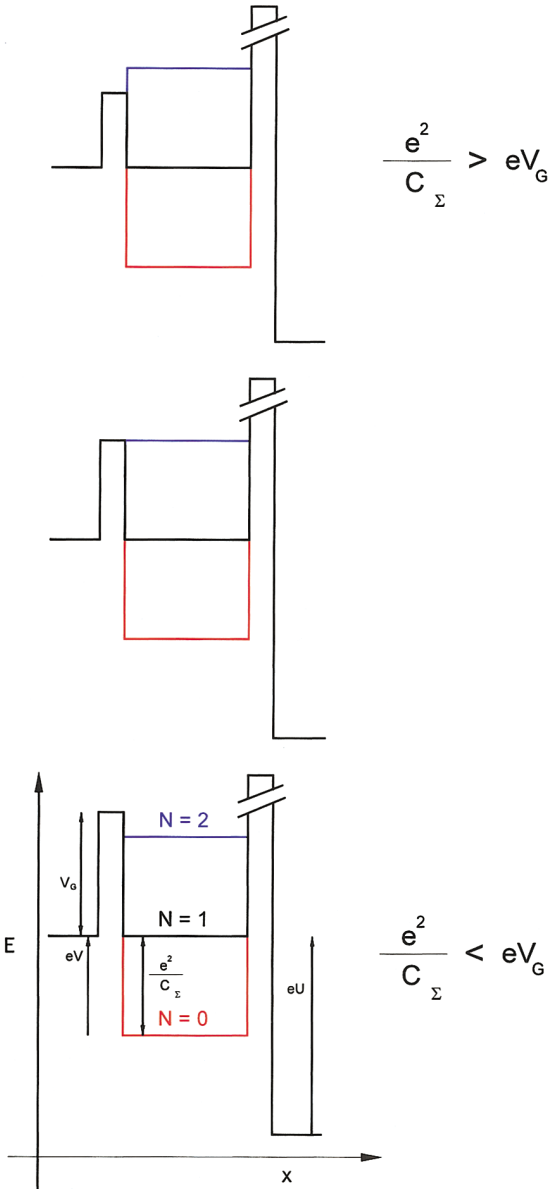

FIG. 3. (Color) Potential energy diagrams for the left lead, island, and right lead, for three cases. Top, $e^2/C_\Sigma > |eV_G|$; middle, $e^2/C_\Sigma = |eV_G|$; bottom, $e^2/C_\Sigma < |eV_G|$. We have assumed $C_\Sigma \approx C_2 \approx C_{UG}$. In each case, three potentials are shown, corresponding to 0, 1, or 2 extra electrons on the island. See Table I for the corresponding activation barriers and energy changes. This figure shows the results for $U=e/C_2$ or $U_{UG}=e/C_{UG}$, which makes the energy changes $\Delta E^{\pm}$ optimum for Coulomb blockade to favor the $N=1$ state.

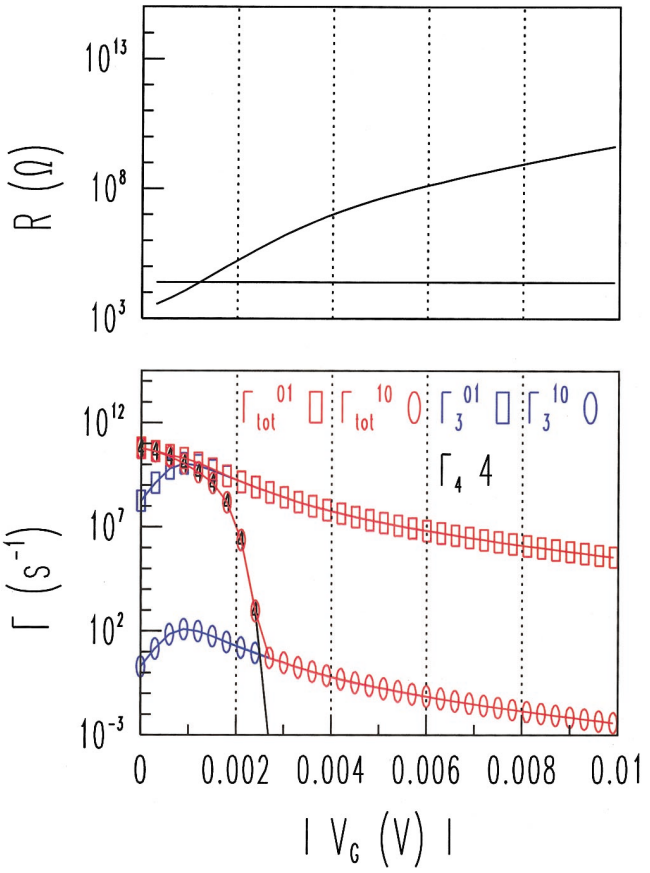

FIG. 4. (Color) Upper: The tunneling resistance vs the gate voltage, for the parameters shown. The length of the forbidden region underneath the gate is $L=0.03$ μm, which determines the dependence of the resistance on gate voltage. For comparison to the approximations in Eq. (5), we note that $k_BT=4\times10^{-4}$ eV and $k_BT(k_BT/eV_G^0)=4\times10^{-3}$ eV; thus, the resistance is dominated by thermally activated motion for smaller V_G. Lower: Rates of motion vs gate voltage for the same range as in the upper curve. Γ_4 [second term in Eq. (11)] is large at low gate voltages, when $R<R_Q$; at large voltages, Γ_3 [first term in Eq. (11)] is larger. The energy change for the motion onto the island (rates Γ^{01}) ΔE^+ is positive, corresponding to a desired tunneling event towards the desired state, conversely, the energy change for rates Γ^{10} is negative, inhibiting motion away from the desired state. Since ΔE^+ is positive and ΔE^- negative, Γ_3^{01} is much larger than both Γ_3^{10} and the thermal rate Γ_{otb}; the same is true of Γ_{tot}^{01} vs Γ_{tot}^{10}. Note that $e^2/C_\Sigma|2-C_2U/e|=0.016$ eV, and thus all of the important rate changes (all of the action in this graph) occur for $|eV_G|<e^2/C_\Sigma|2-C_2U/e|$, or $|eV_G|<e^2/C_\Sigma$ (i.e., the energy barriers are as indicated in the top panel of Fig. 3 over this whole range). Parameters: $T=4$ K, $U=0.032$ V, $C_\Sigma=10$ aF, $N_{chann}=10$, $V_G^0=0.000\ 046$ eV, $\Delta E^{01}=-0.008$ eV, $\Delta E^{10}=-0.008$ eV, $V=0.016$ eV, $e^2/C_\Sigma|2-C_2U/e|=0.016$ eV.

ably low value in this particular simulation at about 0.0025 V. Although the exact details of how the error rate Γ_4 depends on V_G are not known, it is clear that it will fall with some dependence as R increases.

C. Master equation

It is very clear from the foregoing that the rates, both desired and undesired, depend quite sensitively on the details of the device parameters. In addition, in general it may be difficult to discriminate via device geometry between tunneling and thermal over-the-barrier hopping. Thus, it is not possible in this work to directly calculate the dynamical error exactly from the device geometry. Instead, what we hope to do is to give insight into the general trends as well as develop the formalism, allowing calculation of optimum strategies for particular devices in the future.

1. General considerations

The situation we encounter is as follows: at the beginning of phase IV, the probability of finding an undesired number of electrons on the island is relatively large, either because (for the CCD-type device) the rates of forward and backward motion are the same when $R<R_Q$, or (for the turnstile) because we are starting with a state that has $N=0$. As time progresses through phase IV, because the bias voltage is set to favor the state with $N=1$, the probability of finding one electron on the island increases towards 1, and the probabili-

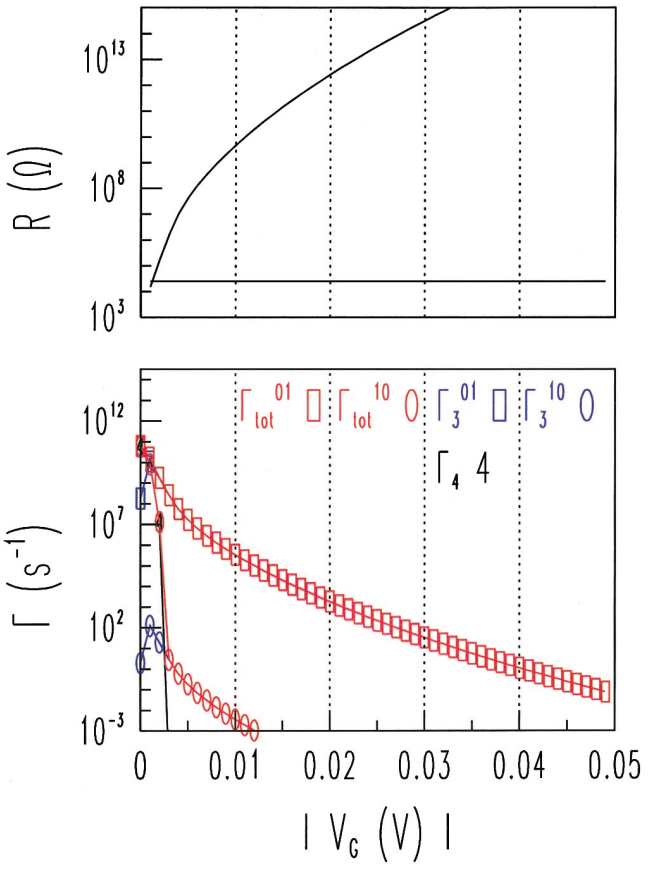

FIG. 5. (Color) Similar to the previous graph, but over a large gate voltage range. Note that the rates Γ_3^{01} and Γ_{tot}^{01} are substantially larger than those of the reverse direction. Above about 0.06 V, the resistance saturates at R_{leak}, and the rates saturate also. Parameters: same as the previous figure.

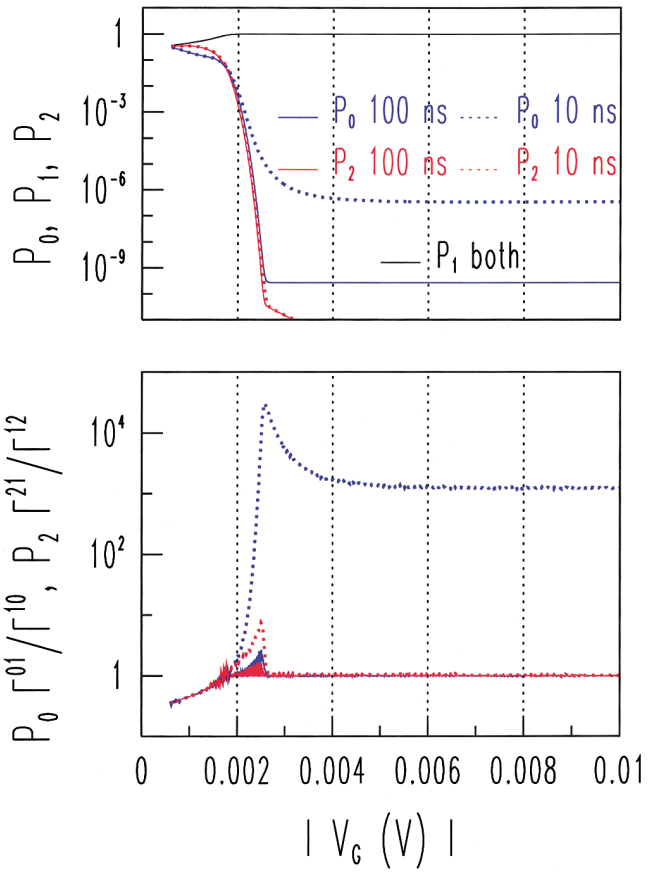

FIG. 6. (Color) Probabilities as a function of time, or gate voltage, for a linear gate voltage ramp. The same parameters are used as in Fig. 4. The ramp is from 0 to −0.2 V, over a time of either 10 or 100 ns. Upper: Probabilities P_0, P_1, P_2 for the number of electrons on the island are 0, 1, or 2, respectively. The desired probability P_1 goes to approximately 1 for all conditions. The probabilities for an undesired error, P_0 and P_2, rapidly decrease in value over the same voltage range (≈0.001 to 0.003 V) for which the undesired rate Γ^{10} falls rapidly in Fig. 4. For a relatively slow ramp time of 100 ns, both undesired probabilities fall below our criterion. However, for a faster ramp time of 10 ns, the probabilities of having an error with zero electrons on the island after the barrier is raised saturates at a substantially larger value than our desired error rate. This is an example of the dynamical error, because the barrier has been raised so quickly that the wrong number of electrons is locked into the island. Lower: Here we plot the two probabilities for undesired outcomes, multiplied by the rates as indicated. For the three cases where the undesired probabilities fall to an acceptably low level, the ratios saturate at 1.0. However, for the ramp time of 10 ns, the ratio does not reach 1.0; this indicates the basic cause for this locking in: P_0 cannot follow the decrease in Γ^{10} quickly enough. Parameters: same as the previous figure, with $|V_{G\ \mathrm{ramp}}|=0.2$ V.

ties of finding a number other than one decrease towards zero, in an approximately exponential fashion. In the rest of this section, we will specialize to the case of the CCD; the other case is similar.

The first question we must ask is: how many different states should we consider? Again this is a detailed question which in general must be answered for each specific device. However, it seems reasonable for our general discussion to only consider three states, those with $N=0,1,2$. We will call the probabilities of finding that number of electrons on the island at any time as $P_0(t), P_1(t), P_2(t)$. The next question we must consider is: what are the values of these probabilities at the beginning of phases IV, time $t=0$? This is also a detailed question which must be answered for each specific device; thus, we make the simplest possible approximation, which is that $P_0(0)=P_1(0)=P_2(0)=1/3$. Finally, in the standard way for a master equation, we obtain the rate of change of these probabilities as

$$\frac{dP_0(t)}{dt} = -\Gamma_{tot}^{01}P_0(t) + \Gamma_{tot}^{10}P_1(t),$$

$$\frac{dP_1(t)}{dt} = -(\Gamma_{\mathrm{tot}}^{10} + \Gamma_{\mathrm{tot}}^{12})P_1(t) + \Gamma_{\mathrm{tot}}^{01}P_0(t) + \Gamma_{\mathrm{tot}}^{21}P_2(t), \quad (13)$$

$$\frac{dP_2(t)}{dt} = -\Gamma_{\mathrm{tot}}^{21}P_2(t) + \Gamma_{\mathrm{tot}}^{12}P_1(t).$$

2. Frequency limit—linear voltage ramp

In Fig. 6, we show the results of the master equation for the same parameters as in Fig. 4. Here, we assume that the gate voltage on the barrier is raised linearly in time, from 0 to 0.2 V, over two possible ramp times t_{ramp} as indicated. These ramp times are chosen because they are in the relevant range for our devices, and also because they show the basics of the dynamical error for the parameters we have used.

In order to get this, we need to obtain the rate Γ^{21}; since there is no barrier for motion in this direction, we can use a standard result from the transit time in CCDs,[26] which yields $\Gamma^{21}=1/\tau=\mu V/L^2\approx 10^{-12}\ \mathrm{s}^{-1}$.

We note that the probability for an error which leaves zero electrons on the island after the gate is raised saturates

at an undesired large value for the shorter ramp time. This is an example of the locking in of the wrong number of electrons on the island. The lower panel of Fig. 6 gives us insight into why this occurs. We consider the first line of Eq. (13), over the relevant voltage range from 0.001 to 0.003 V. In this range, the value of the probability for the desired outcome, P_1, is close to 1, so that we can write

$$\frac{dP_0(t)}{dt} = -\Gamma^{01}P_0(t) + \Gamma^{10}; \tag{14}$$

here, we have suppressed the subscript "tot." In the region of interest, what is occurring is that Γ^{10} is rapidly decreasing, as the resistance rises above R_Q, and thus the Coulomb blockade turns on. In turn, P_0 is attempting to follow the rapid decrease of Γ^{10}. We can see from Eq. (14) how this occurs: as Γ^{10} decreases, the first term on the right-hand side drives P_0 down by providing a negative sum to the time derivative of P_0. However, if Γ^{01} decreases too quickly, because the gate voltage has been ramped too quickly, the first term cannot drive down the probability quickly enough, and so it gets locked in to a relatively large value. Thus, this equation tries to provide negative feedback, and in particular tries to maintain the sum on the right-hand side at a small value. We note this is equivalent to satisfying detailed balance:

$$P_0/P_1 = \Gamma^{10}/\Gamma^{01}. \tag{15}$$

We can see the feedback in the lower panel of Fig. 6: for the one case (P_0 for 10 ns), where the probability of an error saturates at a relatively large value, the ratio plotted shows that this feedback fails to occur.

Given the guidance from this analysis, we can also estimate in a straightforward way what the limit on the ramp time or frequency of operation will be in general. In the gate voltage region of interest, this failure occurs when $\Gamma^{01}P_0 \gg \Gamma^{10}$. Thus, we can approximate

$$\frac{dP_0(t)}{dt} = -\Gamma^{01}(t)P_0(t), \tag{16}$$

or

$$\ln P_0(t) = -\int_0^{t_{\text{ramp}}} dt\ \Gamma^{01}(t). \tag{17}$$

From the first term of Eq. (11), we can see that the time dependence of $\Gamma^{01}(t)$ comes from the gate voltage dependence of the resistance R, predominately as $1/R(V_G)$; also, in the gate voltage region of interest, an approximation for R comes from the middle line of Eq. (5). Thus, we can write

$$\Gamma^{01}(t) \propto 1/R = \Gamma_0 e^{-|eV_G|/kT}, \tag{18}$$

and

$$\begin{aligned}\ln P_0(t) &= -\int_0^{t_{\text{ramp}}} dt\ \Gamma^{01}(t) \\ &= -t_{\text{ramp}}/|V_{\text{G ramp}}|\int_0^{|V_{\text{G ramp}}|} dV_G\Gamma^{01}(V_G) \\ &= -t_{\text{ramp}}/|V_{\text{G ramp}}|\Gamma_0\int_0^{|V_{\text{G ramp}}|} e^{-|eV_G|/kT} \\ &= -kT/e\ t_{\text{ramp}}/|V_{\text{G ramp}}|\Gamma^{01}(R=R_Q),\end{aligned} \tag{19}$$

where we have defined the start of the integral at the point where the resistance goes above R_Q, and $V_{\text{G ramp}}$ is the value of $V_G(t_{\text{ramp}})$, when the ramp is finished.

Over this region, we want the probability of having zero electrons to be less than the acceptable error rate ϵ_0: $\ln P_0 < \ln \epsilon_0$ or

$$t_{\text{ramp}} > \frac{V_{\text{G ramp}}e|\ln \epsilon_0|}{kT\Gamma^{01}(R_Q)}. \tag{20}$$

As a numerical example, if we consider the rates in Fig. 4, with $\epsilon_0 = 10^{-8}$, we can see that at $R \approx R_Q$, the desired rate Γ^{01} is about 1×10^{12} s^{-1}. With $T=4$ K, we thus obtain the result that the minimum ramp time or time for this phase T^{IV} is 6 ns, for a linear ramp from 0 to 0.2 V. This is equivalent to a maximum frequency of 160 MHz (just for this phase).

3. Frequency limit—optimized ramp

Clearly, a simple linear ramp over the whole of phase IV is not an optimum shape for the ramp of the gate voltage. It is clear that, as discussed above, a detailed optimization of the rate needs detailed knowledge of the device parameters. However, we again wish to give some general guidance on how to optimize the ramp.

For instance, one obvious strategy is to ramp quickly for $R < R^*$, then stop for a certain time, then ramp as fast as possible above this. Here, R^* is the value of $R(t)$ where the Coulomb blockade is fully active (the corresponding V_G^* is $\approx$0.0025 V in Figs. 4 and 6). In this case, when V_G is stopped at V_G^*, from Eq. (16) with $\Gamma^{01}(t)$ a constant, $P_0 \propto e^{-\Gamma^{01}t}$ will fall to an acceptable value in a time of $t^* \approx \ln \epsilon_0\ \Gamma^{01}(R^*)$.

To get an approximate answer, we can assume that the Coulomb blockade is fully developed by the time $R = 10R_Q$. In this case, from Eq. (10), we can estimate that this waiting time is $t^* \approx |\ln \epsilon_0|R^*C_\Sigma \approx 10|\ln \epsilon_0|R_QC_\Sigma$. As an example, we show three possible optimized ramps in Fig. 7. The red curve shows the optimum waiting voltage $V_G^* = 0.0025$ V, and demonstrates that only a very short waiting time is required to equilibrate. The black and blue curves show the deleterious effect of stopping too soon or too late, even with a much longer t^*.

Thus, to optimize the gate voltage ramp, we can do the following.

(1) Estimate junction and gate capacitances from the SETT behavior.
(2) With one barrier low, measure the conductance of the other barrier as a function of gate voltage.

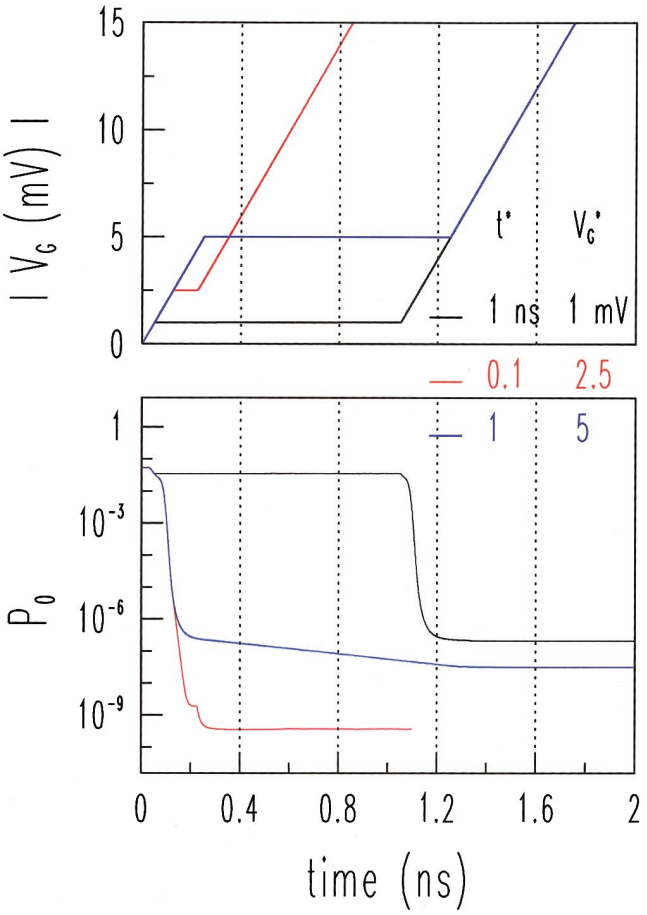

FIG. 7. (Color) Probabilities as a function of time, for a gate voltage ramp with a waiting time. This shows the effect of an optimized ramp. Parameters are the same as the previous two figures, with the gate voltage ramps as indicated. In all cases, the ramp rate dV_G/dt is the same as in the 10 ns curves of the previous figure. Note that in previous figure, this ramp time without a waiting time resulted in a locking in of the wrong the number of electrons. Upper: Three representative possible gate voltage ramps. The difference between them is the waiting time t^* and waiting gate voltage V_G^*. The optimum case (red) has a very short waiting time (0.1 ns), the other two have a long time (1 ns). Lower: Probability P_0 of an error. The optimum case, which stops at $V_G^*=0.0025$ V, reaches an acceptable dynamical error probability in a very short time; as noted in the text, this is the value where the Coulomb blockade becomes fully active. The other two cases show nonoptimum possibilities. The black curve shows the effect of stopping too soon: the error rate Γ_{tot}^{10} is still too large, and so P_0 does not fall. The blue curve, on the other hand, has a value of V_G^* which is too large: at this point, the value of Γ_{tot}^{01} is too small to drive P_0 down to the desired level in 1 ns.

(3) Estimate the optimum bias voltage U, and V_{G}^*; during the ramp, stop at $V_{\text{G}}=V_{\text{G}}^*$ for a time of $t^* \approx 10|\ln \epsilon_0|R_{\text{Q}}C_\Sigma$.

D. Summary of dynamical error considerations

Generality of results. The tunneling and activation rates depend very sensitively on the device parameters. Thus, we suspect that it will not be possible in practice to simulate the rates and thus optimize device operation from considerations of the device geometry alone. Rather, we believe that empirical measurements of the electrical characteristics of one barrier must be made before any such attempts to optimize.

In contrast, the general results that we have obtained, for frequency limits, and for the optimization procedure, depend on only two quite reasonable assumptions.

(1) The Coulomb blockade turns on rapidly, as a function of $t, R,$ or V_{G}, after $R>R_{\text{Q}}$.
(2) The Coulomb blockade depends only on the barrier resistance, independent of whether the motion is tunneling through or activation over the barrier.

Temperature limit. Just as in any other SET device, the thermal energy must be low enough to allow operation of this device. In particular, a crude rule of thumb is that we require $kT<(1/40)e^2/C_\Sigma$ to achieve a relative error less than $\epsilon_0 \approx 10^{-8}$.

Frequency limit. Because of the possibility of locking in the undesired state, the gate voltage cannot be ramped at an arbitrarily fast rate. Rather, we are limited to minimum ramp times as follows:

$$T^{\text{IV}} > \frac{V_{\text{G ramp}} e|\ln \epsilon_0|}{kT\Gamma^{01}(R_{\text{Q}})} \quad \text{[nonoptimized (easy)]},$$
$$T^{\text{IV}} > 10|\ln \epsilon_0|R_{\text{Q}}C_\Sigma \quad \text{[optimized (not easy)]}. \tag{21}$$

III. THEORY OF OTHER MECHANISMS

A. Leakage through other barrier

1. Rates

In the preceding section, we considered the situation where the gate voltage and the resistance of one barrier was varied, while the resistance of the other barrier was considered to be infinite. We now consider the "leakage" error due to the finite resistance of the second barrier, R_{leak}. The mechanism and the magnitude of the error will depend on whether R_{leak} is larger or smaller than R_{Q}, and on the phase of the cycle.

We will consider both the tunneling through the single junction formed by the second barrier, as well as the cotunneling through both the barriers. As in the preceding section, for single-junction tunneling, we have

$$\Gamma = \frac{1}{e^2R}\frac{\Delta E}{1-e^{-\beta\Delta E}}, \quad R>R_{\text{Q}}$$
$$= \frac{V}{eR}, \quad R<R_{\text{Q}}. \tag{22}$$

We will consider in this discussion of leakage only events that are energetically favorable in the Coulomb blockade regime, because the energy change obeys $|\Delta E| \approx E_C \gg kT$. Thus, we have for two successive single-junction events causing a net error,

$$\Gamma_{\text{leak}} = \frac{|\Delta E|}{e^2R_\Sigma}, \quad R>R_{\text{Q}}$$
$$= \frac{V_{SD}}{eR_\Sigma}, \quad R<R_{\text{Q}}. \tag{23}$$

Here, $R_\Sigma = R(t) + R_{\text{leak}}$ is the total resistance of both junctions.

By definition, cotunneling can only occur when both barriers are large compared to R_Q. In this case, the rate is[29]

$$\Gamma_{\text{cot}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{RR_{\text{leak}}}\frac{1}{E_C^2}[(eV_{\text{SD}})^2 + (2\pi kT)^2] \times (eV_{\text{SD}})\frac{1}{1 - e^{-\beta eV_{\text{SD}}}}, \tag{24}$$

with R the resistance of the active barrier, and other symbols as defined previously. We note that this result is numerically only valid for small bias; its usage here is only to obtain an approximate error rate. Since for the turnstile or CCD, the voltage bias across the device is about $V_{\text{SD}} \approx E_C \gg kT$, the last fraction is approximately 1, so

$$\Gamma_{\text{cot}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{RR_{\text{leak}}}\frac{(eV_{\text{SD}})^3}{E_C^2}. \tag{25}$$

Phase I. In this phase, both barriers have resistances equal to R_{leak}, so $R_\Sigma = 2R_{\text{leak}}$. In this case, there are two possible modes for errors, given the voltage bias across the device.

(1) The electron tunnels from the left onto the island, and then tunnels to the right (favorable).
(2) The electron tunnels from the island to the right lead, and then an electron tunnels from the left lead to the island (unfavorable).

Suppressing the unfavorable case for Γ_{leak}, the probability that the electron (after making the first tunneling event) does not return through the same junction, and therefore produces a net error, is approximately equal to 1. To simplify consideration, we will assume that the voltage drop across both junctions for $N=0$ is the same (as we did in the preceding section). The more general case has a straightforward although more complicated analysis. With a total voltage drop of V_{SD} split equally among the two barriers,

$$\Gamma_{\text{leak}}^{\text{I}} = \frac{|E_C/m|}{2e^2 R_{\text{leak}}}, \qquad \Gamma_{\text{cot}}^{\text{I}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{R_{\text{leak}}^2}\frac{(eV_{\text{SD}})^3}{E_C^2}, \tag{26}$$

where $1/2 < m < 1$, depending on bias.

Phases II and IV. In these phases, the active (left) barrier has a resistance much less than R_{leak}, and which may be less than R_Q (for the CCD), so $R_\Sigma = R_{\text{leak}}$. For $R \ll R_Q$, there is no Coulomb blockade, and so there is no cotunneling; instead, there is a large single-junction rate for tunneling through the inactive barrier. In this case, error rates are

$$\Gamma_{\text{leak}}^{\text{II}} = \frac{|E_C/m|}{e^2 R_{\text{leak}}}, \quad R(t) > R_Q = \frac{V_{\text{SD}}}{eR_{\text{leak}}}, \quad R(t) < R_Q, \qquad \Gamma_{\text{cot}}^{\text{II}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{R(t)R_{\text{leak}}}\frac{(eV_{\text{SD}})^3}{E_C^2} \quad R(t) > R_Q. \tag{27}$$

Phase III. This phase is similar to phases II and IV for the single-junction tunneling rate Γ_{leak}, except that the active barrier resistance $R(t)$ is a constant, R_{low}; again, $R_\Sigma = R_{\text{leak}}$. There is no cotunneling for the CCD, since there is only one tunnel junction. With R_{low} the value of the active barrier for the turnstile in this phase,

$$\Gamma_{\text{leak}}^{\text{III}} = \frac{|E_C/m|}{e^2 R_{\text{leak}}} \quad [\text{turnstile}], \qquad \Gamma_{\text{leak}}^{\text{III}} = \frac{V_{\text{SD}}}{eR_{\text{leak}}} \quad [\text{CCD}], \qquad \Gamma_{\text{cot}}^{\text{III}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{R_{\text{low}}R_{\text{leak}}}\frac{(eV_{\text{SD}})^3}{E_C^2} \quad [\text{turnstile}]. \tag{28}$$

2. Estimates of errors

Since these errors are all independent of past history (in contrast to the dynamical errors), we do not need to consider a master equation, and thus the total error is simply

$$\epsilon^\iota = \int^i dt\Gamma^i, \tag{29}$$

where i represents the time spent in each of the four phases.

Thus, we have the total relative error in one cycle of the turnstile or CCD as

$$\epsilon_{\text{leak}} = \frac{|E_C/m|}{e^2 R_{\text{leak}}}\{T^{\text{I}}/2 + T^{\text{II}} + T^{\text{III}} + T^{\text{IV}}\} \quad [\text{turnstile}],$$

$$\epsilon_{\text{leak}} = \frac{|E_C/m|}{e^2 R_{\text{leak}}}\{T^{\text{I}}/2 + T^{\text{II};R>R_Q} + T^{\text{IV};R>R_Q}\} + \frac{V_{\text{SD}}}{eR_{\text{leak}}}[T^{\text{II};R<R_Q} + T^{\text{III}} + T^{\text{IV};R<R_Q}] \quad [\text{CCD}],$$

$$\epsilon_{\text{cot}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{R_{\text{leak}}}\frac{(eV_{\text{SD}})^3}{E_C^2}\left[\frac{T^{\text{I}}}{R_{\text{leak}}} + \frac{T^{\text{III}}}{R_{\text{low}}} + \int_{\text{II}} dt\frac{1}{R(t)} + \int_{\text{IV}} dt\frac{1}{R(t)}\right] \quad [\text{turnstile}],$$

$$\epsilon_{\text{cot}} = \frac{1}{6\pi^2 h}\frac{R_Q^2}{R_{\text{leak}}}\frac{(eV_{SD})^3}{E_C^2}\left[\frac{T^{\text{I}}}{R_{\text{leak}}} + \int_{\text{II};R>R_Q} dt\frac{1}{R(t)} + \int_{\text{IV};R>R_Q} dt\frac{1}{R(t)}\right] \quad [\text{CCD}]. \tag{30}$$

Here, we have defined T^{I}, T^{II}, T^{III}, T^{IV}as the time spent in phases I, II, III, and IV; we also define $T^{\text{tot}} = T^{\text{I}}/2 + T^{\text{II}} + T^{\text{III}} + T^{\text{IV}}$. We can make some immediate simplifications: Because R_{leak} is so large, the term for phase I in the cotunneling error is so small that we can neglect it (here, we assume that the time spent in phase I, T^{I}, is not much much longer than the time in the other phases). Also, there is less than a factor of 2 difference between $|E_C/m|/e$ and V_{SD}, so we set them equal. In that case, by noting that $R_Q = h/e^2$, we obtain

TABLE II. List of error mechanism, concomitant limits on parameters, and leading corrections to relative error rates for parameters outside of desired ranges.

Mechanism	Equation	Parameter	Error ε	Limit	Limit ($\varepsilon_0=10^{-8}$)	Limit ($\varepsilon_0=10^{-7}$)
Dynamical	20	Time $T^{\mathrm{IV}}_{\min}$ (nonoptimized)	$\varepsilon_0\exp[-T^{\mathrm{IV}}/T^{\mathrm{IV}}_{\min}]$	$eV_{\mathrm{G\ ramp}}\lvert\ln\varepsilon_0\rvert/kT\Gamma^{01}(R_Q)$	>6 ns	>5 ns
		(optimized)		$100\,R_QC_\Sigma$	>25 ps	>22 ps
Thermal		Temperature $T_{\max}$	$\varepsilon_0\exp[T/T_{\max}]$	$e^2/2\lvert\ln\varepsilon_0\rvert C_\Sigma$	<4.4 K	<5.0 K
Leakage	32	Maximum times	$(C_\Sigma R_{\mathrm{leak}})^{-1}[T^{\mathrm{I}}/2+T^{\mathrm{II}}+T^{\mathrm{III}}+T^{\mathrm{IV}}]$	$\varepsilon_0C_\Sigma R_{\mathrm{leak}}$	<10 μs	<100 μs
Frequency	33	$T^{\mathrm{III}}_{\min}$	$\exp[-T^{\mathrm{III}}/C_\Sigma R_{\mathrm{low}}]$	$\lvert\ln\varepsilon_0\rvert C_\Sigma R_{\mathrm{low}}$	>0.1 ps	>0.1 ps
Heating	34	Frequency $f_{\max}$	$\varepsilon_0\exp[(f_{\max}/f)^{1/5}]$	$\varepsilon_0^{-1}A\sigma/k_B(e^2/C_\Sigma k_B)^4$	<200 MHz	<400 MHz

$$\epsilon_{\mathrm{leak}}=\frac{|V_{\mathrm{SD}}|}{eR_{\mathrm{leak}}}T^{\mathrm{tot}}\quad[\text{turnstile}],$$

$$\epsilon_{\mathrm{leak}}=\frac{V_{\mathrm{SD}}}{eR_{\mathrm{leak}}}T^{\mathrm{tot}}\quad[\text{CCD}],$$

$$\epsilon_{\mathrm{cot}}=\frac{1}{6\pi^2}\frac{V_{\mathrm{SD}}}{eR_{\mathrm{leak}}}\left[T^{\mathrm{III}}\frac{R_Q}{R_{\mathrm{low}}}+\int_{\mathrm{II}}dt\frac{R_Q}{R(t)}+\int_{\mathrm{IV}}dt\frac{R_Q}{R(t)}\right]\quad[\text{turnstile}],\tag{31}$$

$$\epsilon_{\mathrm{cot}}=\frac{1}{6\pi^2}\frac{V_{\mathrm{SD}}}{eR_{\mathrm{leak}}}\left[\int_{\mathrm{II};R>R_Q}dt\frac{R_Q}{R(t)}+\int_{\mathrm{IV};R>R_Q}dt\frac{R_Q}{R(t)}\right]\quad[\text{CCD}].$$

We note that, for the turnstile, $R(t)\geqslant R_{\mathrm{Q}}$, and for the CCD, $R(t)\geqslant R_{\mathrm{Q}}/20$; because of this, in both cases, the extra prefactor of $1/6\pi^2$ means that the cotunneling error is smaller, so that the leakage error dominates. Thus, we obtain, noting that it is also approximately true that $|E_C/m|/e\approx V_{\mathrm{SD}}\approx e/C_{\mathrm{UG}}\approx e/C_\Sigma$, that both devices have the same final estimate:

$$\epsilon_{\mathrm{leakage}}=\frac{1}{C_\Sigma R_{\mathrm{leak}}}T^{\mathrm{tot}}\quad[\text{turnstile}]$$

$$=\frac{1}{C_\Sigma R_{\mathrm{leak}}}T^{\mathrm{tot}}\quad[\text{CCD}].\tag{32}$$

We note that these final results are appealing in their simplicity: the error is essentially a fraction of the cycle time divided by the RC time for the device.

We note that the errors in Eq. (32) can be considered as limits on frequency, but also as limits on temperature, because R_{leak} may be thermally activated. However, this limit is easily neglected, because it is likely that we can always increase the terminal value of V_{G} and thus increase the activation barrier, and so increase the value of R_{leak} at any given temperature.

B. Frequency errors

This error results from not waiting long enough for electrons to tunnel, very similar to the metal pumps.[30] Since we separately considered the master equation in phase IV, it is not necessary to separately consider frequency errors in this phase. Also, such errors are unimportant in phases I and II. In phase III, the frequency error in the CCD device is negligibly small, because the resistance is so low. Finally, we can estimate the error in the turnstile in phase III as

$$\epsilon_{\mathrm{freq}}=e^{-T^{\mathrm{III}}/R_{\mathrm{low}}C_\Sigma}.\tag{33}$$

C. Heating errors

This error is more complicated to calculate, because it requires knowledge of the cooling mechanism of the electrons in the device; these mechanisms are not well known at low temperatures.

In general, the low-temperature electron-phonon cooling power has been better studied in metals than in semiconductors. To get an estimate of the effect we follow the general procedure in Niu *et al.*,[31] which we outline in the following: for a cycle frequency $f=1/T_{\mathrm{cycle}}$, the power input to the central island is approximately $P_{\mathrm{in}}=fE_C=fe^2/2C_\Sigma$. In general, the cooling power due to the electron-phonon coupling is $B(T_e^5-T_L^5)$, where T_e and T_L are the electron and lattice temperatures, respectively; B is a constant. For the two-dimensional electron gas in a GaAs/AlGaAs heterostructure, the cooling power is estimated as $\sigma A(T_e^5-T_L^5)$, where A is the area, and $\sigma=30$ fW/K^5/μm^2. Since the temperature rise depends only weakly on the empirical parameter σ as $T_e\propto\sigma^{1/5}$, we believe it is reasonable to use this result for a Si quantum dot.

Setting the incoming power equal to the cooling power, we can obtain the rise in the electron temperature over the lattice temperature. Using an area of $(0.08\ \mu\mathrm{m})^2$, and with a total capacitance of 10 aF, we obtain $T_e=0.09f^{1/5}$, with T_e in units of K, and f in Hz. Using the criterion developed earlier for the maximum temperature to avoid thermal errors, $kT<1/20E_C$, we obtain an estimate of the maximum frequency as $f_{\max}=200$ MHz.

We may also ask the question: What is the effect on the error rate for frequencies higher than this? As we saw above, the relative error grows with temperature as $\epsilon_{\mathrm{therm}}\propto e^{-E_C/kT}$. Thus, if we use a frequency higher than the maximum, we get a relative error which is approximately

$$\epsilon_{\mathrm{heat}}=e^{-20(f_{\max}/f)^{1/5}}.\tag{34}$$

IV. CONCLUSIONS

We refer to Table II, which summarizes most of the results in this paper. In this table, we list the various parameters which are important to reach a desired low relative

error ϵ_0. For each parameter, we show the effect on the error ϵ, as well as the formula for the limit of the parameter, plus numerical values for two possible desired relative error levels.

We note that in general the limits on the parameters are all reasonable; we must of course acknowledge, as described in the foregoing, that the rates for the dynamical error depend very sensitively on the device parameters, and so the values shown for that error mechanism are only approximate estimates. The limit which appears to be most problematic is the minimum time spent in phase IV, for the nonoptimized case. This does not concern us overmuch, for two reasons: one is that this limit can probably be moderately reduced by reducing the temperature, and the second is that it may be possible to substantially reduce it by the optimization procedure outlined above.

Much of this paper has been devoted to trying to understand and calculate the dynamical error mechanism. This mechanism is intimately bound up with the interesting question of how the Coulomb blockade is formed, and in particular how electrons split themselves up as a barrier is formed. We have opted to take the simplest phenomenological approach to trying to understand this question; we believe there may be substantial interesting experiments to be done in this realm, and we invite more fundamental theoretical considerations for this dynamical error.

Another interesting question which is raised herein is that of the interplay between thermal over-the-barrier hopping and Coulomb blockade. As mentioned above, there is one theoretical treatment of this, for the SET transistor.[19] However, we know of no such treatment for the *control* of single-electron motion, and we believe there is room for fruitful analysis in this regime of low, thick tunnel barriers.

Finally, having listed the formulas for and values of the various error mechanisms, we can generally comment on the feasibility of Si-based electron current standards. It is evident from Table II that the putative error mechanisms can in general be controlled at an acceptable level for electrical metrology. Given that, there are substantial advantages to the Si-based devices compared to the metal-based ones that make it desirable to pursue this alternative route towards a high-value current standard. These advantages include the tunable junction resistance, which allows us to avoid the cotunneling error which forces the choice of using many (seven, e.g.) tunnel junctions in the metal pumps.[4,30,32,33]

Also, the Si-based devices inherently have a capacitance which is smaller by about an order of magnitude, allowing them to run significantly faster compared to the frequency error. In addition, this smaller capacitance affords the capability of running the devices at a much higher temperature [in the metal devices, the maximum temperature is about 0.1 K (Ref. 4)]. Finally, the lack of the charge offset drift in at least one class of Si-based SET devices[7] affords the potential of parallelizing a large number of turnstiles or CCDs to achieve a very large current.

ACKNOWLEDGMENTS

It is a pleasure to thank Leonid Glazman (U. Minnesota) for enlightening discussions regarding the interplay of thermally activated motion and Coulomb blockade, Mark Stiles (NIST Gaithersburg) for valuable discussions regarding many of the issues considered in the section on dynamical errors, Eric Vogel (NIST Gaithersburg) for discussions regarding Si device physics, Mark Keller (NIST Boulder) for a critical insight into the analysis of the frequency limit from the dynamical error mechanism, and Mike Kelley (NIST Gaithersburg) for useful comments on the exposition.

[1] H. Grabert and M. H. Devoret, *Single Charge Tunneling* (Plenum, New York, 1992), Vol. 294.
[2] Y. Takahashi, Y. Ono, A. Fujiwara, and H. Inokawa, J. Phys.: Condens. Matter **14**, R995 (2002).
[3] S. Shimano, K. Masu, and K. Tsubouchi, Jpn. J. Appl. Phys., Part 1 **38**, 403 (1999).
[4] M. W. Keller, J. M. Martinis, N. M. Zimmerman, and A. H. Steinbach, Appl. Phys. Lett. **69**, 1804 (1996).
[5] M. W. Keller, A. L. Eichenberger, J. M. Martinis, and N. M. Zimmerman, Science **285**, 1706 (1999).
[6] W. H. Huber, S. B. Martin, and N. M. Zimmerman, *Experimental Implementation of Quantum Computation (IQC'01)*, edited by R. Clark (Rinton, Princeton, NJ, 2001), Vol. 76 (reprint available on our website)..
[7] N. M. Zimmerman, W. H. Huber, A. Fujiwara, and Y. Takahasi, Appl. Phys. Lett. **79**, 3188 (2001).
[8] Y. Ono, N. M. Zimmerman, K. Yamazaki, and Y. Takahasi Jpn. J. Appl. Phys., Part 2 **42**, L1109 (2003).
[9] Y. Ono and Y. Takahasi, Appl. Phys. Lett. **82**, 1221 (2003).
[10] A. Fujiwara and Y. Takahashi, Nature (London) **410**, 560 (2001).
[11] A. Fujiwara, N. M. Zimmerman, Y. Ono, and Y. Takahashi, Appl. Phys. Lett. **84**, 1323 (2004).
[12] L. P. Kouwenhoven *et al.*, Phys. Rev. Lett. **67**, 1626 (1991).
[13] Y. Nagamune *et al.*, Appl. Phys. Lett. **64**, 2379 (1994).
[14] J. M. Jacobson, Y. Yamamoto, and A. Imamo Glu, Europhys. Lett. **26**, 129 (1994).
[15] C. Weiss and W. Zwerger, Europhys. Lett. **47**, 97 (1999).
[16] J. D. White and M. Wagner, Phys. Rev. B **48**, 2799 (1993).
[17] N. S. Wingreen, A. P. Jauho, and Y. Meir, Phys. Rev. B **48**, 8487 (1993).
[18] T. Hatano, K. Fujimoto, and Y. Isawa, Jpn. J. Appl. Phys., Part 1 **34**, 4488 (1995).
[19] K. A. Matveev and L. I. Glazman, Phys. Rev. B **54**, 10339 (1996).
[20] N. Y. Morgan *et al.*, J. Appl. Phys. **89**, 410 (2001).
[21] Y. Takahashi *et al.*, Jpn. J. Appl. Phys., Part 1 **38**, 2457 (1999).
[22] C. P. Herrero, G. Schon, and A. D. Zaikin, Phys. Rev. B **59**, 5728 (1999).
[23] S. V. Panyukov and A. D. Zaikin, Phys. Rev. Lett. **67**, 3168 (1991).
[24] D. V. Averin and K. K. Likharev, *Mesoscopic Phenomena in Solids* (North-Holland, New York, 1991), Vol. 30, Chap. 6, p. 208.
[25] The specific choice of smoothing function does not greatly affect the overall results in this paper. In particular, it numerically affects the results in Figs. 4 and 6, in the region of the crossover. However, it does not affect any of the general conclusions reached in Sec. II D.
[26] S. M. Sze, *Physics of Semiconductor Devices* (Wiley, New York, 1981).
[27] B. N. J. Persson and A. Baratoff, Phys. Rev. B **38**, 9616 (1988).
[28] A. N. Korotkov, Phys. Rev. B **49**, 11508 (1994).
[29] H. Pothier, Ph.D. thesis, L' Universite é Paris, 1991, p 130.
[30] H. D. Jensen and J. M. Martinis, Phys. Rev. B **46**, 13407 (1992).
[31] Q. Niu *et al.*, IEEE Trans. Instrum. Meas. **44**, 564 (1995).
[32] J. M. Martinis, M. Nahum, and H. D. Jensen, Phys. Rev. Lett. **72**, 904 (1994).
[33] M. W. Keller, in *Recent Advances in Metrology and Fundamental Constants*, edited by T. J. Quinn, S. Leschiutta, and P. Tavella (IOS, Amsterdam, 2001), p. 291.