Electric-field-induced memory effects in granular films

R. E. Cavicchi* and R. H. Silsbee
Laboratory of Atomic and Solid State Physics, Cornell University, Ithaca, New York 14853-2501
(Received 6 July 1987; revised manuscript received 2 May 1988)

We report on a study of memory effects in granular film systems. Experiments by Lambe and Jaklevic and by Adkins *et al.* are examined. In the former, a granular film is embedded in a capacitor, and the memory effect is an anomaly in the low-temperature capacitance-voltage plot centered at the bias in which the capacitor was cooled. In the latter, the granular film forms one plate of a capacitor, bias is applied between the capacitor plates, and the memory effect is an anomaly in the low-temperature plot of the granular film's conductance versus bias, centered about the cooling bias. In reviewing these experiments, it is found that there is lacking a theoretical understanding which correctly addresses the roles of electrostatic and electrochemical potentials in establishing the memory. Experiments are carried out on a Lambe-Jaklevic device which provide insights into the effect and supplement the existing experimental results. Two theoretical models are proposed. One, based on the idea of the Coulomb glass, predicts results inconsistent with some aspects of the experimental results. The other is related to the ideas in the original Lambe-Jaklevic and Adkins *et al.* works but resolves issues raised in considering their models and is consistent with all experimental results.

I. INTRODUCTION

Metal island films are known to display memory effects consistent with some kind of freezing-in of a charge distribution. Lambe and Jaklevic[1] performed measurements on a "tunnel capacitor," Fig. 1(a), in which an island film of particles, typical diameter ≈ 100 Å, is embedded in a capacitor and is separated from the lower plate by a thick oxide, and from the upper plate by a thin tunnel barrier. When the sample is cooled from room temperature to 4.2 K, the capacitance shows an oscillation as a function of bias applied to the sample V_S. This is shown in Fig. 2(a), a plot of $\Delta C = C - C_0$ versus V_S, where C is the measured device capacitance and C_0 is the expected zero-bias capacitance in the absence of the memory signal. Remarkably, if the sample is cooled with a static bias applied during the cooling, the feature is centered not on zero bias, but on the cooling bias V_{SC}. The magnitude of the signal may be characterized by the quantity $\Delta C(V_{SC})$ and is found to be independent of V_{SC}. Adkins *et al.*[2] measured the in-plane conductance of a granular film evaporated on a dielectric substrate as a function of the voltage applied to a field electrode on the underside of the substrate, Fig. 1(b). If the sample is cooled in a field electrode bias V_{SC}, the low-temperature film conductance displays a minimum at V_{SC}. Shown in Fig. 2(b) is a plot of the ratio $\Delta\sigma/\sigma$ versus V_S for a sample cooled with $V_{SC}=0$, where $\Delta\sigma$ is the difference between the measured device conductance σ and the conductance σ_0 expected in the absence of the memory signal. In both experiments, the sample "remembers" the bias it was cooled in, displaying a memory which persists for long times at low temperatures.

The similarity of the experimental results in the Lambe-Jaklevic and Adkins *et al.* experiments suggests that the same physics is responsible for both memory effects. However, the physical mechanism which produces these memory effects is not well understood. The memory may involve Coulomb interactions between charges on different particles, interactions between oxide defect charges, and particle electrons, or both. We have investigated the problem with theoretical models and experiments, focusing on the Lambe-Jaklevic device where we need not be concerned with the many-body problem of charge transport in a granular film.

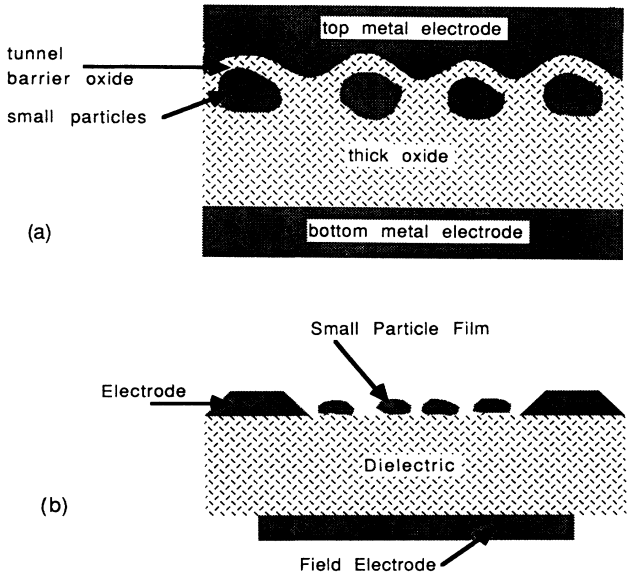

FIG. 1. The sample configurations of (a) Lambe-Jaklevic—a tunnel capacitor, and (B) Adkins *et al.*—a granular film field-effect device.

 ©1988 The American Physical Society

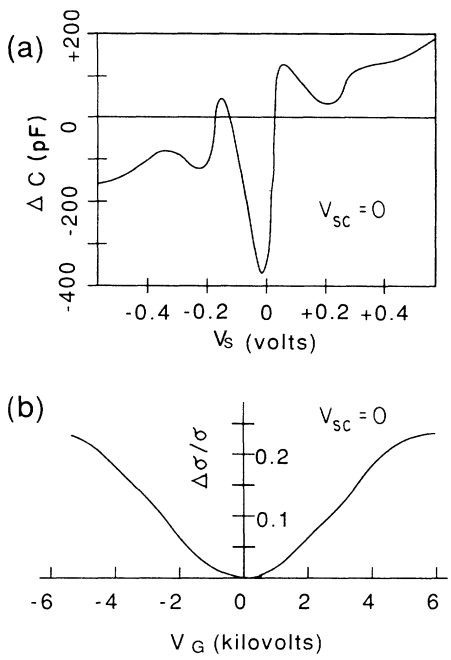

FIG. 2. Memory effect in (a) Lambe-Jaklevic[1] and (b) Adkins *et al.* devices.[2] In both cases $V_{SC}=0$ V.

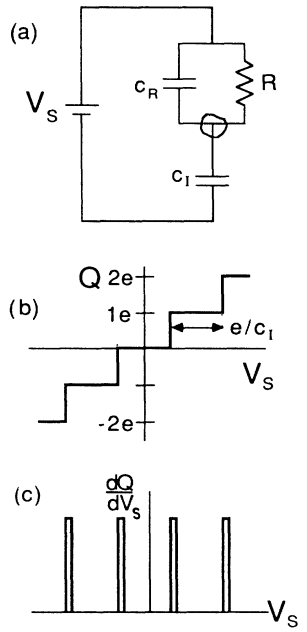

FIG. 3. (a) Circuit model for a small metal particle in either a Lambe-Jaklevic or Adkins *et al.* device. (b) Charge added to the particle as a function of applied bias at zero temperature with $V_{SD}=0$. (c) Incremental capacitance deduced from (b).

In Sec. II we review the ideas and experimental results of Lambe-Jaklevic and Adkins *et al.* Additional results from experiments carried out by us on a Lambe-Jaklevic type of device are also summarized. Then we consider two possible models. One, described in Sec. III, assumes the memory results from interparticle Coulomb interactions, and is based on the idea of a Coulomb glass. However, it is not in accord with some aspects of the experimental results. The other, presented in Sec. IV, is a polarization model based on ideas of the preceding two groups, but giving a more detailed treatment of the roles of the electrochemical and electrostatic potentials in the problem. Finally, a discussion and summary of our results is given in Sec. V.

II. REVIEW OF EXPERIMENTS

A. Lambe-Jaklevic tunnel capacitor

Figure 3(a) shows a circuit model for one of the small metal particles in the Lambe-Jaklevic device in which c_R and c_I are the capacitance between the particle and the upper and lower plate, and tunneling is represented by the resistance R. What is unusual about this circuit is that the geometrical capacitances c_R and c_I are so small that the discreteness of the electron charge becomes important. This is illustrated in Fig. 3(b), a plot of the net charge Q, in units of a single-electron charge, accumulated on the particle as a function of V_S at zero temperature. The step length is given by e/c_I. For a 100 Å diameter particle, c_I is on the order of a micropicofarad. This means that the change in V_S, associated with a single-electron transfer, is on the order of 0.1 V. The incremental capacitance, dQ/dV_S, deduced from the charging staircase of Fig. 3(b), is a periodic sequence of δ functions

$$e\delta(V+e/2c_I+ne/c_I)$$

where n is an integer [Fig. 3(c)].

In the discussions that follow, it will often be convenient to refer to potentials at the particle rather than across the device. We represent the applied bias V_S, cooling bias V_{SC}, and step length e/c_I by quantities that are scaled down by a factor, $\gamma=c_I/(c_R+c_I)$:

$$V\equiv\gamma V_S\ ,\qquad V_C\equiv\gamma V_{SC}\ ,\qquad \frac{e}{c}\equiv\gamma\frac{e}{c_I}\ . \tag{1}$$

We shall use the prefix S in the subscript to denote voltages applied to the device; the absence of S in the notation for voltages will mean voltages at the particle.

For samples we have used in our experiments, γ is typically 0.2. The step length referred to the particle e/c plays a crucial role in the discussion. It is the change in electrostatic potential ϕ of the particle upon the addition of one electron and is called the charging voltage. We also note that the mean energy level spacing δ is about two orders of magnitude smaller than e/c. We believe that δ plays no role in the memory effect, but it can be of importance in the dynamics of charge transport at very low temperatures.[3]

We shall also find it convenient to express the measured device capacitance C as the sum of a geometrical component C_G and a particle-electrode tunneling component C_P. In a model device with N particles having

identical c_I, c_R, and R, C_G is given by

$$C_G=\gamma N c_R+C_N \ , \tag{2}$$

where C_N is the capacitance between the top and bottom electrodes arising from regions in the granular film where there are no particles. C_P is obtained by summing over the incremental capacitance contributions of the individual particles

$$C_P=\sum_j \gamma \frac{dQ}{dV_S}(j) \ . \tag{3}$$

Assuming a linear R in the circuit model, appropriate only if $V_S \gg e/c_I$, C_P is instead given by

$$C_P=Nc_I\frac{1-\omega^2R^2c_Rc_I/\gamma}{1+(\omega Rc_I/\gamma)^2}-Nc_R\gamma \ , \tag{4}$$

where ω is the measuring frequency in radians per second. C_P may be determined experimentally from $C(f)-C(\infty)$, where $f=\omega/2\pi$.

As first noted by Giaever and Zeller,[4] in their study of small particles in a tunnel junction, if the work function of the particle W_p is different from that of the electrode W_e, there will be a small residual difference $V_D(0)$ between the Fermi levels of particle and electrode at zero bias. Figure 4 shows how this offset comes about at zero temperature and zero bias. For the sake of clarity, we leave out the factor $-e$, multiplying all potentials indicated on this energy diagram. In Fig. 4(a) the tunneling interaction is initially imagined to be turned off. The vacuum levels are aligned and the Fermi levels differ by an amount

$$V_D(0)=\Delta W=W_p-W_e \ .$$

When the tunneling interaction is turned on, the Fermi levels are brought into closer alignment by the transfer of electrons, Figs. 4(b) and 4(c). With each transfer, the particle's potential changes by e/c. This transfer process ends with $V_D(0)$ reduced to a value between $\pm e/2c$, Fig. 4(c). Lambe and Jaklevic[1] noted that if a given particle has a nonzero $V_D(0)$, its periodic capacitance pattern will not be centered about zero but will be offset from zero by some amount $V_D(0)/\gamma$.

At this point we wish to precisely define the important variables having to do with the electrostatic and electrochemical potentials at the particle. We define the quantity $V_D(V)$ as the difference in Fermi levels, or, equivalently, for the temperature regime of these experiments, electrochemical potentials, between the particle and the nearby electrode at a bias V. The bias dependence of V_D is given by

$$V_D(V)=\{[V_D(0)+V+e/2c]\bmod e/c\}-e/2c \ . \tag{5}$$

Charge transfer occurs at the thresholds $V_D(V)=\pm e/2c$, or at $V=[V_D(0)+e/2c]\bmod e/c$, and the capacitance pattern as a function of V_S will be symmetric about a voltage $V_D(0)/\gamma$. We define the electrostatic potential between a particle and nearby electrode as a distinct quantity

$$\phi=V+Q/c+\delta\phi_{\rm pol} \ ,$$

where $\delta\phi_{\rm pol}$ is a possible contribution from an oxide polarization to ϕ.

Because the work functions of different crystal faces of a metal differ by amounts on the order of tenths of volts, one expects a similar variation in ΔW from particle to particle depending on a particle's shape and orientation. This variation is large compared with the charging voltage e/c at the particle, about 20 meV for a 100 Å diameter particle. To see how this large variation in work functions affects the capacitance measurement in the device, consider an ensemble of particles with identical e/c, and with a broad distribution of work functions, with $V=0$, Fig. 5(a). With the tunneling interaction turned off, the distribution of work function differences $n_W(\Delta W)$ is equivalent to a distribution of V_D among the particles which we call $n'(V_D;0)$ where the zero represents an applied bias of zero volts, and the prime signifies that the tunneling interaction is off. The peak of the distribution is located at $\overline{W}_p$-W_e, where $\overline{W}_p$ is the mean work function of the particles and depends on the metal which composes the particles. If the particles and electrode are the same metal, then $\overline{W}_p-W_e=0$. When the tunneling interaction is turned on, charge transfer occurs as described in Fig. 4, and the Fermi level of a given particle is translated by steps of e/c until it lies within a band of width e/c centered about $V_D=0$—the "Coulomb zone." The collapse of the broad distribution leads to a distribution, $n(V_D;0)$ which, if the width of the W_p distribution is large compared with e/c, will be uniform between $\pm e/2c$ and equal to $n_u \equiv N/(e/c)$, Fig. 5(b). Giaever and Zeller[4] assumed $n(V_D;0)$ was constant to obtain good fits

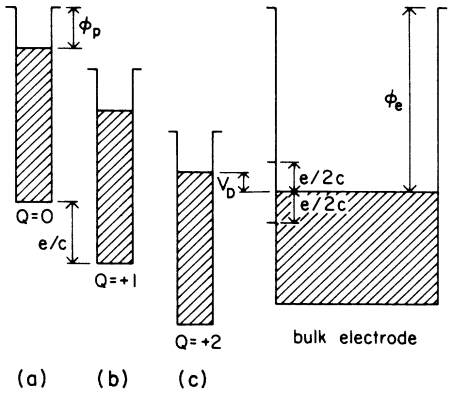

FIG. 4. Alignment of the Fermi levels of a single particle with the bulk electrode. (a) The tunneling interaction is imagined to be initially turned off. Switching on this interaction causes the Fermi levels to align via the transfer of charge (b) and (c). With each transfer, the particle's potential drops by e/c. Finally, in (c), the situation is reached where no more charge transfers occur. The Fermi levels remain separated by the residual amount V_D, determined by the work functions of particle and electrode, the charging voltage, and any random static potentials in the oxide.

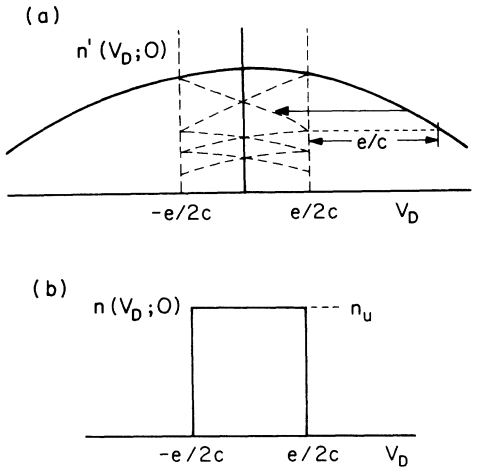

FIG. 5. The translation of a broad distribution of V_D into a zone of width e/c for an ensemble of particles with a given e/c. In (a) the bold line indicates the distribution of V_D before the tunneling interaction is turned on. This is equivalent to the distribution of W_P-W_e among particles with different W_p. When the tunneling interaction is switched on, particles are translated into the zone as described in Fig. 4. The arrow indicates an e/c-wide piece of the boldface curve being translated into the zone. (b) The resulting pileup of such pieces (dashed line) causes the distribution $n(V_D;0)$ to be uniform and equal to n_u within the zone and zero outside of it.

to their conductance data. However, a constant $n(V_D;0)$ implies a constant distribution of offsets $V_D(0)/\gamma$ between $\pm e/2c_I$, and therefore the periodic structure for individual particles is washed out in the sum Eq. (3).

To understand how a nonuniform distribution of V_D arises during cooling and leads to a capacitance feature centered about the cooling bias V_C, Lambe and Jaklevic gave an explanation summarized here. Consider an ensemble of particles with differing V_D but with identical e/c. Let $V_D(V_C)$ be the difference in Fermi levels between particle and electrode at bias V_C. The application of a bias V_C results in a distribution $n(V_D;V_C)$ that is shifted in V_D by the amount V_C with respect to $n(V_D;0)$, with charge transfer moving those particles which are biased out of the "Coulomb zone" back in at the other end of the distribution

$$n(V_D;V_C)=n([(V_C+V_D+e/2c)\bmod e/c]-e/2c;0)\ . \tag{6}$$

With a broad distribution of work functions, the distribution of V_D is expected to be uniform and equal to n_u. According to Lambe and Jaklevic, as the sample rests in a given bias some polarization process occurs which tends to align the Fermi levels of the particles with that of the electrode, giving a peak in $n(V_D;V_C)$ at $V_D=0$. If this polarization responds rapidly to bias changes, the peak shifts as the bias is changed and no evidence of it is seen experimentally. If the sample is cooled, however, so that the thermally-activated polarization processes have slowed to the point that the response time to bias changes is longer than any relevant experimental time $\tau_{\rm exp}$, then the polarization is effectively frozen. The distribution $n(V_D;V_C)$ will have a permanent peak at $V_D=0$ like that shown in 6(a), unaffected by changes in the applied bias.

At a bias V, it is only those particles at the extremes, $\pm e/2c$, of the distribution $n(V_D;V)$ which contribute to the incremental capacitance C_p so that $C_p \alpha n(e/2c;V)$. If the sample has been cooled at a bias V_C, leaving an associated peak in $n(V_D;V_C)$ at $V_D=0$, then the peak will contribute an excess capacitance only when the bias is set such that $V=V_C+e/2c+ne/c$, bringing the peak to the end of the distribution $n(V_D;V)$. The relationship between $C_P(V)$ and $n(V_D;V_C)$ can be expressed as

$$C_P(V)=e\gamma^2 n([(V-V_C-e/2c)\bmod e/c];V_C)\ ; \tag{7}$$

C_P is thus a periodic mapping of $n(V_D;V_C)$. Note that if $n(V_D;V_C)=n_u$, then $C_P=\gamma N c_I \equiv C_{P,u}$, the zero-frequency limit of Eq. (4). Figure 6(b) shows $C_P(V)$ resulting from $N(V_D;0)$ in Fig. 6(a).

The actual sample will consist of a distribution of e/c, thereby washing out the sharp feature associated with the bumps in $C_p(V)$ at $V=V_C\pm e/2c$, yielding a capacitance hole that has a width Γ corresponding to an average e/c for the system as indicated by the dashed line in Fig. 6(b). If the distribution of e/c in the sample is narrow enough,

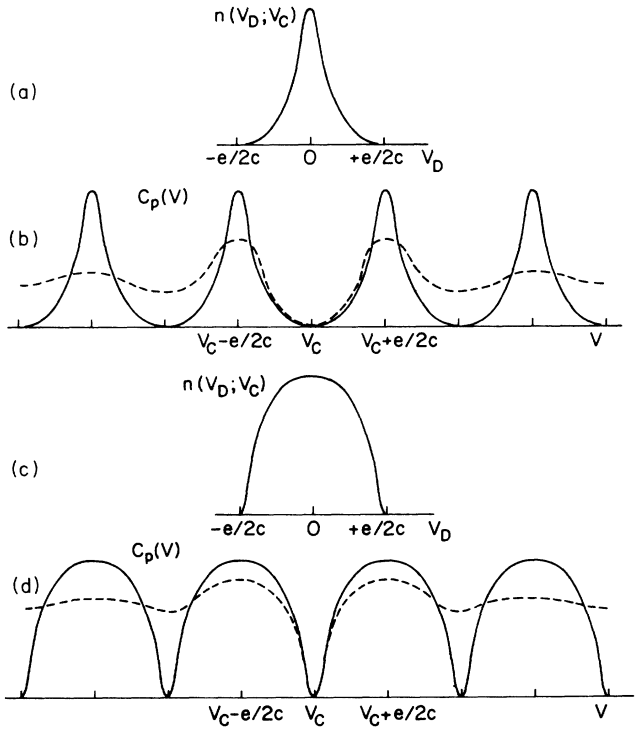

FIG. 6. Capacitance resulting from $n(V_D;V_C)$. In (a) and (b) the distribution is modeled as a peak at $V_D=0$ after Lambe-Jaklevic while in (c) and (d) we model it as having holes at $\pm e/2c$. (a) and (c), $n(V_D;V_C)$ frozen-in at the cooling bias. (b) and (d) resulting C_P vs V. A superposition of ensembles of particles with different e/c is indicated by the dashed line.

as in the Lambe-Jaklevic data of Fig. 2(a), additional damped oscillations may be seen on either side of V_C before the dephasing of the different periodicities is complete.

Later in this paper we present a model in which $n(V_D;V_C)$ is made nonuniform by the creation of a hole at the ends of the distribution, $\pm e/2c$, as in Fig. 6(c) rather than by a peak at the center. In this case, the sharp feature in the capacitance-bias characteristic is at V_C [Fig. 6(d)], so that a distribution of e/c will not tend to wash out the sharp feature. Thus, after averaging over a distribution of e/c, a narrow hole in the end of the distribution can, in principle, produce a Γ that is narrower than e/c, while a narrow peak in the center of the distribution cannot.

The maximum deviation of $n(V_D;V_C)$ from n_u in Fig. 6(c) occurs at $V_D=\pm e/2c$. We define a quantity

$$H \equiv n_u - N(e/2c;V_C)$$

to be the depth of the hole in $n(V_D;V_C)$. H is related to the maximum deviation $\Delta C(V_C)$ of C_P from $C_{p,u}$ by

$$\Delta C(V_C) \equiv C_{p,u} - C_P(V_C) = e\gamma^2 H \ , \tag{8}$$

where the latter equality follows from Eq. (7).

The Lambe and Jaklevic discussion suggests a polarization process in the oxide between particle and electrode which reduces V_D towards zero for all of the particles. However, in considering candidate polarization processes, it is important to distinguish between the electrochemical potential, i.e., V_D, which is crucial in the discussion of the electron transfer, and the electrostatic potential ϕ, which defines the fields which may induce remnant polarization of the insulator. The particle shown in Fig. 4(c) illustrates this point. Two electrons have been transferred from the particle to bring the particle's Fermi level to within $e/2c$ of the Fermi level of the bulk electrode. This has left the particle positively charged. ϕ is represented in the diagram by the difference between vacuum levels of the particle and electrode. A polarization responding to this electrostatic potential will raise the potential energy for an electron on the particle, reducing the difference between vacuum levels, thereby increasing V_D. Thus, for some particles, the polarization actually increases the difference in Fermi levels between particle and electrode, while for other particles, the polarization may reduce this difference. Positive and negative shifts of V_D are equally likely, giving no tendency, on average, to bring the Fermi levels of the particles into juxtaposition with that of the electrode.

If the polarization process involves thermal or photoexcitation of electrons into the conduction band, it is ϕ, together with the kinetics of the electron trap, which will determine where the electrons go. Electrostatic potentials also govern the motion of ionic charge in the oxide. We know of no polarization process which responds only to the electrochemical potential difference and is consistent with all of the aspects of the experiment. The puzzle, then, is to find a mechanism by which the medium response, driven by electrostatic potential changes, can respond in a systematic fashion to the relative positions of Fermi levels. The answer is found in focusing on the behavior of those particles which are at the *ends* of the V_D distribution rather than on finding a mechanism to provide a peak at the center of the distribution, as suggested by Lambe and Jaklevic.

To gain a better understanding of the phenomenon, we have both duplicated the Lambe-Jaklevic experiments and carried out additional experiments on a tunnel capacitor. The sample-making procedure we have used for our experiments is described in another paper.[3] We summarize some of the most significant results herein.

1. Frozen-in memory hole

A tunnel capacitor cooled in a bias, V_C, displays a capacitance signal like that in Fig. 2(a). The oscillation $\Delta C(V_S)$ is centered about the cooling bias. Depending on the distribution of e/c in the sample, the oscillation may be strongly damped and only appear as a hole without the side lobes associated with subsequent oscillations. The magnitude of the signal $\Delta C(V_C)$ is independent of the cooling bias.

2. Square-wave cooling

Lambe and Jaklevic also found that two holes could be produced by cooling the sample in a 1 Hz square-wave bias, alternating between V_{SA} and V_{SB}. At low temperature, they observed capacitance holes at V_{SA} and V_{SB}. $\Delta C(V_S)$ for each of the holes has half the amplitude of a single hole. The frequency of the square wave is observed not to influence the position or depth of the holes; we have seen this effect using square-wave frequencies up to 20 kHz.

3. Hole associated with temperature range

We have found that a memory hole may be associated with a temperature cooling range. If the sample is cooled from T_3 to T_2 in a bias V_{SA}, and then from T_2 to T_1 in a bias V_{SB}, the sample will display two holes: a broad hole at V_{SA}, and a narrower hole at V_{SB} (Fig. 7). If the sample is annealed to temperature T_2, the hole at V_{SB} will be erased, but the hole at V_{SA} will remain.

4. Magnitude of memory hole

To characterize the magnitude of the memory hole $\Delta C(V_C)$, one needs to determine C_P. As indicated by Eq. (4), this can be done from the frequency dependence of C. Lambe-Jaklevic report on capacitance measurements at only one frequency. The tunneling barrier in their samples is thin enough that the rate that characterizes electron transfers between a particle and the nearby electrode, $\nu=\gamma/2\pi Rc_I$ in the circuit analogue of Fig. 3, is large compared to the frequency of the ac measuring signal for all particles in the device. We have been interested in studying the dynamics of the tunneling process[3,5] and have produced samples with tunnel barriers thick enough that our available frequencies (100 Hz–50 kHz) span the range $f<\nu$ to $f>\nu$.

Figure 8 shows a comparison of our experimental re-

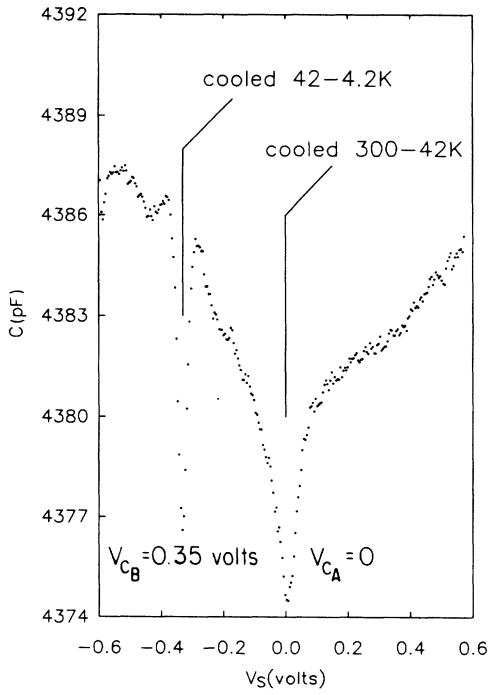

FIG. 7. $C(V_S)$ at 4.2 K for a sample cooled from 300 to 42 K in a bias $V_{SC}=0$ V, and from 42 to 4.2 K in a bias $V_{SC}=-0.35$ V.

sults for $C(f)$ and $\Delta C(V_C,f)$ at 4.2 K. We note that samples without particles show almost no dependence of C with frequency and no memory effects. $C(f=\infty)$ is about 1760 pF. The scale for the right-hand ordinate axis for the memory hole plot was chosen so that one point—at 150 Hz—would lie on the sample capacitance curve. The observation that as a result of this single choice of scale factor all the memory hole data lie on the curve $C_P(f)$, supports the statement that the frequency dependent capacitance and capacitance memory hole in the device can be attributed to particles characterized by the same distribution of tunneling times. The scale factor is a characterization of the magnitude of the memory hole in terms of the total capacitance contribution of the particles. For the sample in Fig. 8, the magnitude of the memory hole is about 5% of the capacitance of all the particles.

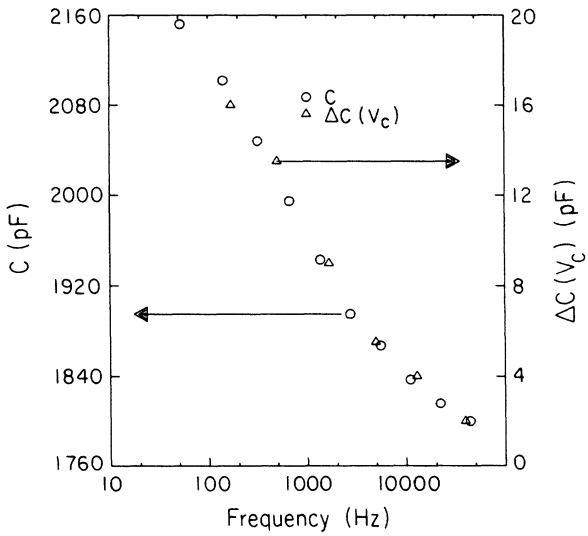

FIG. 8. Frequency dependence of device capacitance C (circles) and capacitance memory hole depth $\Delta C(V_C)$ (triangles).

5. *Fixed-temperature hole development*

Lambe and Jaklevic observed that if a sample was subjected to a bias V_{SA}, different from the cooling bias V_{SC}, for a time τ_{SA} long compared to the time it takes to sweep the bias in a C-V measurement τ_S, a subsequent C-V measurement revealed that a small capacitance hole had developed, centered about V_{SA}. Meanwhile, the hole at V_{SC} had decreased in amplitude. An alternative way of observing this time-dependent effect is to apply a step voltage from V_{SC} to V_{SA} and to observe the C as a function of time t after the step. The experiment monitors the development of the new hole by measuring the decrease of capacitance at its center V_{SA}. An advantage of this technique is that one may observe processes that occur on a time scale that is short compared to the time it takes to sweep the bias in a Lambe-Jaklevic measurement. Figure 9 shows a typical capacitance decay taken at 0.14 K. This sample had a $\Delta C(V_C)$ of about 10 pF when cooled from room temperature to 4.2 K. The signal shown in Fig. 9 represents an average of 16 traces using a square-wave bias of magnitude $V_{SC}-V_{SA}=0.1$ V at 0.2 Hz, taken with the aid of a signal-averaging oscilloscope. For comparison, also shown in Fig. 9 are attempted fits to

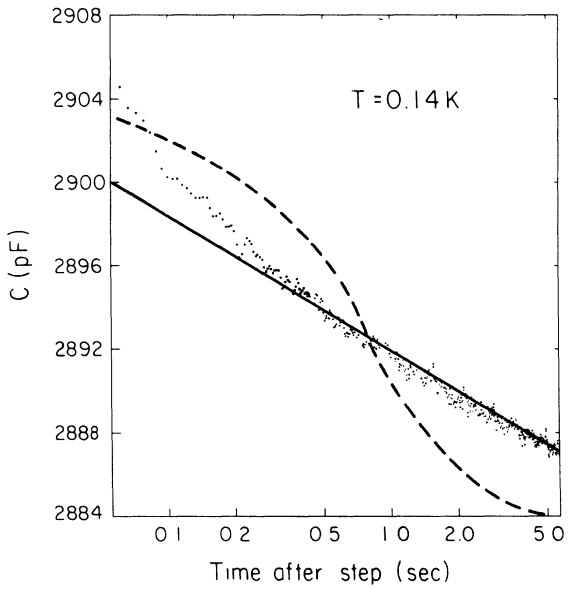

FIG. 9. Device capacitance $C(V_A)$ vs log$t$ when the bias is switched from V_{SC} to V_{SA} at 0.14 K. The magnitude of the voltage step V_{SC}-V_{SA} is 0.1 V. Data represent an average of 16 traces. Also shown are attempted fits to an exponential (dashed line) and logarithmic time dependence (solid line).

$$C = C_{K1}\exp(-t/\tau_K) + C_{K2}$$

with $C_{K1} = 10$ pF, $C_{K2} = 2884$ pF, and $\tau_K = 1$ sec, and to a log$t$ dependence with a slope C_S of 6 pF per decade. The data agree qualitatively better with the latter fit. The characteristic slope C_S depends on temperature and V_{SA}. At fixed temperature, C_S increases from zero at $V_{SA} = V_{SC}$ to its maximum value for $V_{SA} - V_{SC} > 0.05$ V. C_S increases with decreasing temperature until it saturates at a maximum value below 2 K.

6. *Lowest temperature for hole freezing*

There exists a minimum temperature $T_{\min}$ below which it is not possible to freeze a memory hole by further cooling. If a sample is cooled between temperatures T_2 and T_1 in a bias V_C, and $T_2 < T_{\min}$, there will be no memory hole at V_C. $T_{\min}$ is typically a few degrees Kelvin.

B. The Adkins field-effect device

The experiments of Adkins *et al.*[2] can be described in a way closely analogous to those of Lambe and Jaklevic, using the same variables. In Fig. 3(a), c_R corresponds to the capacitance between a particle and the rest of the film, while c_I corresponds to the capacitance between the particle and the field electrode, and the resistor represents tunneling in the film plane. In the charging staircase, Fig. 3(b), Q corresponds to the number of electrons accumulated on a particle as a function of field electrode bias. As before, if V_S is the bias applied to the sample, the step length is e/c_I. To describe potentials between the particle and the surrounding film, the scale transformation by a factor $c_I/(c_R + c_I)$ is again useful and we can refer to the bias and charging voltage by the quantities defined in Eq. (1). As in the Lambe-Jaklevic experiment, there will be a distribution of work functions W_p among the particles. We can again, if somewhat less precisely, define a quantity V_D as being the difference between the Fermi level of a particle and the mean Fermi level of its surrounding particles. By the arguments given in connection with Fig. 5, one may expect that V_D will be uniformly distributed between $\pm e/2c$ in the absence of the memory effect.

The conductivity σ in the plane is assumed to be via thermally activated hopping. In their paper, Adkins *et al.* demonstrate that, if for all the particles the charging staircase were centered about $V=0$ ($V_D=0$), the difference in activation energy between $V=0$ and $V = e/2c$ would give rise to an enormous change in conductivity with V. If all the capacitances were the same, the effect would be periodic in V. The same arguments used in the discussion of the Lambe-Jaklevic memory hole, however, may be used here to show that one expects two effects—the distribution of particle size and the uniform distribution of V_D—to completely wash out this enormous periodic conductance. Nevertheless, the experiments show a small (few percent) dip in the conductance at the cooling bias rather than zero.

Adkins *et al.* explain the observed effect in terms of the development of a polaron in the dielectric as a response to the mean charge state of a particle during the cooling. At low temperature, when the field electrode bias is equal to the cooling bias, the particle is in its "ground state." Moving the bias away from V_{SC} reduces the excitation energy needed to add a charge to the particle, thereby increasing the number of carriers in the film and the conductivity. However, the experiment of Lambe-Jaklevic, in which cooling the sample in a square wave results in two memory holes, suggests that it is not a mean charge state that is responsible for the memory hole. Furthermore, the argument does not include a discussion of the role of electrochemical potential V_D, and thus how to handle particles like those in Fig. 4(c), where the induced polarization would actually reduce the activation energy.

III. COULOMB-GAP MODEL

The Coulomb glass model is a many-body problem concerned with how n_1 charged particles are distributed among n_2 sites where $n_2 > n_1$. The particles interact with one another via Coulomb repulsion. The system is glass-like in that configurations for the assignment of the n_1 particles on the n_2 sites that are nearly equal in total energy may be widely separated in phase space. For the system to move in phase space from one state to another of equal or lower energy, may require passing through states of much higher energy. If the thermal energy available is much less than this effective energy barrier, the time constant for the system to make such a move can be prohibitively long. The system can be stable in a local energy minimum point in phase space, even though there are existing lower-energy states. Efros and Shklovskii[6] gave a simple argument to show that in the ground state the density of states for adding an electron to the system has a gap at the Fermi level.

One possible explanation for the memory effects in the granular films is that the system is behaving like a Coulomb glass. In this view, the N particles act as sites for the electrons. The cooling bias defines the single-site energies V_D (V_C). There are many possible distinct configurations for the assignment of charge to the particles, which vary in total energy because of the complex interplay of the Coulomb interactions between charges on different particles and the distribution of single-site energies. At high temperatures, there is enough thermal energy to cause fluctuations in the charge states of the individual particles. These fluctuations allow the system to gain access to a subset of the many nearly equivalent configurations. As the temperature is lowered, the charge states of fewer and fewer particles are fluctuating, allowing the system less freedom in reaching alternative low-energy configurations. The glass, upon cooling, freezes into one of a number of local minima in the space of available configurations, the particular minimum depending upon the detailed history of bias and temperature.

If, by changing the bias from V_C to V, an additional electron is added to the system at low temperature, its addition to a particle destroys the careful local arrangement of charges designed to minimize the Coulomb energy. If,

because of the low temperature, no rearrangement of charge is allowed, the energy of the state with the added charge will be higher than that of the true ground state at the altered bias by the order of the Coulomb interaction between particles. This extra energy, or Coulomb gap, suppresses the transfer of charge. In the Lambe-Jaklevic experiment, the gap causes a reduced capacitance. For the Adkins experiment, since transport is via excited states of the system, the gap will cause a reduction in the conductivity. If variable range hopping is the transport mechanism, the conductivity of the granular film is predicted to display a crossover from a $\exp(-T^{-1/3})$ to $\exp(-T^{-1/2})$ dependence.[7]

We note that Coulomb-gap effects are more likely to be observable in the Adkins experiment than in the Lambe-Jaklevic tunnel capacitor. This is because in the Lambe-Jaklevic experiment, the presence of image charges in the nearby tunneling electrode causes the effective interaction between particles to be dipole-dipole. Nevertheless, Davies[8] has worked out the theory for this system and has determined that there is still a weak Coulomb gap.

It would seem, then, that the idea of a Coulomb glass in the granular film could explain the reductions in capacitance and conductance observed in the experiments when $V=V_C$. We have sought a means of distinguishing the Coulomb glass model from other models. One prediction, peculiar to the Coulomb glass model, has to do with the nonergodic properties of the glass. Consider the Lambe-Jaklevic experiment with the sample in the ground state with bias V_C at zero temperature. Add electrons one at a time by moving the bias away from V_C. Added electrons will go onto the particles offering the lowest-energy state available. However, because $T=0$ no rearrangement of charge can occur, and so the system cannot reach the true ground state corresponding to the system with the new number of electrons. If the number of electrons added to the system is on the order of the number of sites (particles), all vestiges of the correlations characterizing the ground-state charge distribution will have disappeared. Suppose the bias is now returned suddenly to V_C, i.e., in a time short compared to the transfer time constant. Instead of the electrons being removed from the particles in the reverse order of their addition, the order of removal in response to a step change in potential will be primarily determined either randomly or by the details of the tunneling rates of the various particles, and not by the energetic relations which defined the order of slow filling of the sites. Since rearrangement processes are forbidden at $T\sim 0$, the system is most unlikely to return to the particular ground state corresponding to V_C. The effects associated with the system in the ground state, reduced capacitance in the Lambe-Jaklevic experiment and reduced conductivity in the Adkins experiment, will no longer be present.

We have performed this test on a Lambe-Jaklevic device. At low temperature we slowly sweep the bias away from V_C and then suddenly step it back to V_C. We observe that the capacitance hole remains. Indeed, the reported stability of the memory hole, despite repeated sweeps of the bias to voltages much greater than e/c away from V_C in the Lambe-Jaklevic and Adkins experiments, indicates that the Coulomb glass does not dominate the physics of the devices. Once the glass was moved away from its ground state by a slow change of bias away from V_C, it is questionable that even with a slow return of the bias back to V_C it could then retrace its path through phase space to find the ground state. It might be interesting to search for some loss of the memory hole with voltage sweeps in a device of the Adkins *et al.* type as a small effect atop the larger memory effect which is probably due to oxide processes as discussed later. The observation of an irretrievable loss of a portion of the capacitance hole after an initial sweep away from the cooling bias would provide more convincing evidence for the Coulomb glass state than the hard-to-establish crossover of the power law of the conductivity exponential.

IV. POLARIZATION MODEL

The model presented here for the memory effects is based on the idea of a relaxation of the dielectric medium surrounding a particle, and is qualitatively related to the ideas discussed in the experimental work of Lambe-Jaklevic and Adkins *et al.* The model resolves the question raised in the discussion of Fig. 4(c): What roles do the electrochemical and electrostatic potentials V_D and ϕ play in the creation of the memory hole? It also provides a context for understanding the experimental results described above for the Lambe-Jaklevic devices.

In this model, the dielectric is treated as homogeneous but nonideal. The nonideality is the freezing of polarizability upon cooling, presumably due to motion or reorientation of defect or impurity species present in the dielectric in low concentration. On the size scales (~ 100 Å) of concern here, the defects can hardly be thought of as homogeneously distributed; one might expect, for example, a single defect to reside in the tunneling barrier between the particle and electrode. In the Appendix we describe a model in which the particle's V_D is influenced by a neighboring two-state defect, and in which the energies of the two defect states are altered by the electron occupancy of the particle. The results of the model are similar to those predicted by the homogeneous dielectric model, although some significant differences do arise. However, existing experiments are unable to distinguish between the models. We choose the homogeneous model for presentation below as the more straightforward to develop.

Crucial to the model is the quantized nature of the charge transport process which determines the electrostatic fields to which the medium is to respond. To illustrate this point, consider the case which involves a particle large enough that its charging voltage may be neglected and in which the work functions of particle and electrode differ by ΔW. The electric field between the particle and electrode will be equal to $\Delta W/d$, independent of the bias applied to the device, where d is the separation between particle and electrode. The large particle is able to align its Fermi level with that of the electrode by the transfer of charge. Now, for a small particle, the bias must change by e/c in order to induce a single charge

transfer. This gives rise to an electric field that increases linearly with bias until a charge transfer occurs, where the potential drops by e/c and the field by $(e/c)/d$. The field-bias characteristic for a small particle is the sawtooth of Fig. 10, where the average electric field is $\Delta W/d$. The field at zero bias is determined by the number of charges transferred to bring the Fermi levels into alignment to within $e/2c$ and by any random static fields in the oxide. Random static fields play a role that is similar to the distribution of work functions, and for clarity will be deleted from the following treatment.

The dielectric medium develops a polarization $\mathbf{P}=(\epsilon-1)\mathbf{E}/4\pi$ in response to the electric field from the charged particle. The dielectric response may be decomposed into two components. One part involves mechanisms, e.g., electronic or ionic polarization, for which the equilibrium configuration may surely be reached at any temperature in response to local charges in the electrostatic field. It makes no contribution to the memory effect. The other part describes a dielectric process responsible for the "freezing-in" of remnant polarization in the dielectric. The rate at which these processes respond to changes in the electric field is assumed to decrease strongly with reduced temperature. We describe this part of the dielectric response in terms of a time and temperature-dependent contribution to the dielectric constant $\epsilon_f(T,t)$ which is the response of these processes in the medium to a step electric field change. The processes in the medium may respond with a broad distribution of time constants τ. For example, one simple model for $\epsilon_f(T,t)$ is the relation

$$\epsilon_f(T,t)=\int_0^\infty g(T,\tau)(1-e^{-t/\tau})d\tau\ , \tag{9}$$

where $g(T,\tau)$ defines the contribution to the dielectric constant at temperature T and at $t=\infty$ of processes with relaxation time τ. For thermally activated polarization processes, τ will depend exponentially on $1/T$. Processes with $\tau>\tau_{\rm exp}$, a time characteristic of the experiment in question, are said to be "frozen" and will not contribute to the polarization of a medium in an experiment. If the medium is cooled from a temperature T_2 to T_1 in the presence of an electric field $\mathbf{E}$, those processes which are fast ($\tau<\tau_{\rm exp}$) at T_2 and slow at T_1 leave a frozen polarization in the medium proportional to the field in which the medium is cooled:

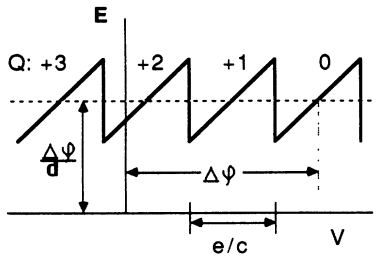

FIG. 10. Electric field between particle and tunneling electrode as a function of applied bias. The e/c-wide ramps on the sawtooth are regions of fixed charge on the particle, indicated by Q for a particle like that in Fig. 4.

$$\delta\mathbf{P}_f(T_1,T_2;\tau_{\rm exp})=\delta\epsilon_f(T_1,T_2;\tau_{\rm exp})\mathbf{E}/4\pi\ , \tag{10}$$

where

$$\delta\epsilon_f(T_1,T_2,\tau_{\rm exp})=\int_{T_1}^{T_2}\frac{d\epsilon_f}{dT}(\tau_{\rm exp},T)dT\ . \tag{11}$$

Below the lower temperature T_1, this contribution to the total polarization is fixed and unaffected by changes in the applied field on time scales of less than $\tau_{\rm exp}$. The existence of dielectric processes which can give rise to this dielectric response is supported by the observed decrease in the dielectric constant with decreasing temperature and decreasing frequency in capacitors formed without particles.

A. Frozen-in memory hole

For simplicity we shall focus on the dielectric between the particle and nearby electrode. Freezing processes in material between the particle and distant electrode contribute to the memory hole in a similar way. Consider an ensemble of particles with the same e/c, with a distribution of work functions that is broad compared to e/c. Figure 11(a) is a potential energy diagram of such a system with an imposed bias $V_C=0$ with the tunneling interaction turned off. Particles with different work functions are displaced horizontally for clarity. When the tunneling interaction is turned on, Fig. 11(b), charge transfer is induced to bring the Fermi levels of the particles all to within a band of width e/c that is centered about the Fermi level of the bulk electrode, as described previously. It is essential to note that the electrostatic fields in this system are quantized to the values $(Qe/c)/d$ where Q is the number of electrons transferred to the particle. It is this field, dependent upon the charge state of the particle but not upon the position of its V_D within the $\pm e/2c$ band, which drives the polarization of the medium. Note again the apparent paradox that the field driving the medium polarization is independent of small differences from particle to particle in Fermi level or work function, yet it is the polarization induced by this field which is supposed to shift the V_D distribution in such a way as to bring the Fermi levels into a closer match with that of the substrate. To see how $n(V_D;0)$ is affected by the freezing-in of polarization, it is useful first to imagine the particles at zero temperature and the dielectric at finite temperature. Cool the dielectric between the temperatures T_2 and T_1. As a result of the cooling and the local electrostatic field of a particle, some polarization is frozen-in. Let V_D' represent the difference in Fermi levels between particle and electrode before the polarization develops and V_D be this difference after the polarization develops. The polarization will alter the electrostatic potential ϕ by an amount $\delta\phi_{\rm pol}=4\pi d\,\delta P$ with

$$\delta P=\delta\epsilon_f(Qe/c)/d\ .$$

The shift in electrochemical potential $V_D'-V_D=\delta\phi_{\rm pol}$ is as indicated by the vertical arrows in Figs. 11(b) and 12(a). As a result of the polarization potential shift, an additional charge transfer will be induced for those particles with V_D' near the Coulomb threshold to a state of in-

creased charge magnitude; the threshold is at $-e/2c$ for negatively charged particles and at $+e/2c$ for positively charged particles. The polarization process leaves gaps, indicated by the double-ended arrows, at the ends of the V_D distribution: at $+e/2c$ for negatively charged particles and $-e/2c$ for positively charged particles, Fig. 12(b). The gap is a consequence of the step increase in the magnitude of the electric field experienced by those particles for which a charge transfer is induced by the developing polarization. The magnitude of the gap, $\Delta V_D(T_1,T_2,\tau_{\rm exp})$, is related to the difference in $\delta\phi_{\rm pol}$ between neighboring charge states, Q and $Q+1$ by

$$\Delta V_D(T_1,T_2,\tau_{\rm exp})=\delta\phi_{\rm pol}(Q+1)-\delta\phi_{\rm pol}(Q)=\frac{e}{c}\delta\epsilon_f(T_1,T_2,\tau_{\rm exp})\ . \tag{12}$$

Here, $\tau_{\rm exp}$ is the sweep time in the capacitance-voltage measurement.

When we now allow the particles to be at nonzero temperature, their charge states are subject to thermal fluctuations. The probability of a particle being excited to its neighboring charge state of lowest energy is given by the Boltzmann factor,

$$\exp\left[-e\left(\frac{e}{2c}-|V_D'|\right)\Big/kT\right]\ .$$

The dielectric medium around a particle will therefore experience the switching electric field from the particle's fluctuating charge state. This has the effect of causing the frozen-in hole in $n(V_D;0)$ to have a width Γ given by

$$\Gamma(T_f,T_f+dT)=\frac{kT_f}{e} \tag{13}$$

for polarization freezing between T_f+dT and T_f. Since the area of the hole must be conserved, the depth of the hole in $n(V_D;0)$, $H(T_f,T_f+dT)$ will be

$$H=\frac{\delta\epsilon_f(T_f,T_f+dT,\tau_{\rm exp})}{kT_f/e}N\ . \tag{14}$$

The argument given above for the case $V_C=0$ may be easily generalized for nonzero cooling bias. In this case the gaps develop in the distribution $n(V_D;V_C)$ at $V_D=\pm e/2c$ to produce a distribution like that in Fig. 6(c). From this model the roles of electrochemical and electrostatic potentials V_D and ϕ are resolved. The electrochemical potential defines the biases at which electron transfers occur. At a given cooling bias, the polarization which develops and freezes-in may cause a charge transfer only in particles whose V_D is near the Coulomb threshold. The charge transfer causes a step change in ϕ,

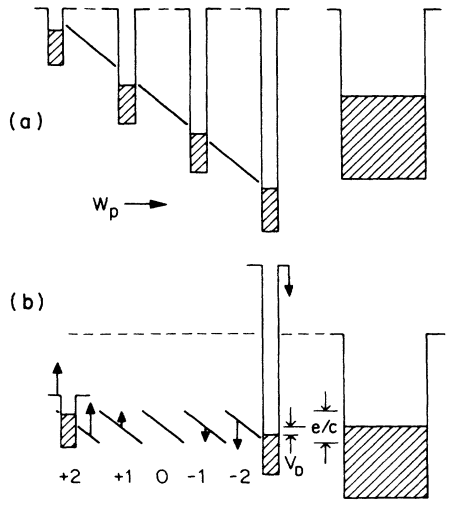

FIG. 11. Potential energy diagram of an ensemble of particles with a unique e/c, but a broad distribution of W_p (a) before and (b) after the tunneling interaction is turned on. (a) The vacuum levels of all particles are aligned with the vacuum level of the electrode. Horizontal position in the diagram reflects a specific value of W_p. (b) All particles have V_D within $\pm e/2c$ of the electrode Fermi level. Subsequent slow polarization of the medium in response to the change in charge state will cause the potential energy to add an electron for positively charged particles to increase by an amount $\phi_{\rm pol}$ up arrows) and negatively charged particles to decrease (down arrows).

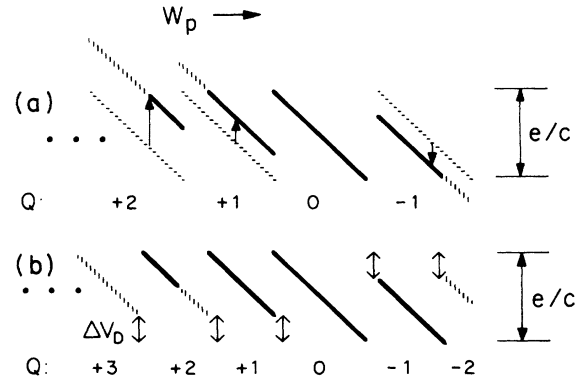

FIG. 12. Development of gaps in $n(V_D;0)$ for an effective medium model. (a) An abbreviated version of Fig. 11(b) where the freezing polarization changes the energy of particles according to their charge state. As in Fig. 11, horizontal position in the diagram reflects a specific value of W_p. Horizontally-hatched lines indicate Fermi levels before development of polarization. The resulting state after polarization develops is indicated with a boldface line for particles that remain within the Coulomb zone, and a vertically-dashed line for particles in which a charge transfer is induced by the polarization potential. (b) The situation after the polarization-induced transfers occur. The absence, represented by double-ended arrows, of positively charged particles with $V_D<-e/2c+\Delta V_D$ and negatively charged particles with $V_D>e/2c-\Delta V_D$ results in gaps in $n(V_D;0)$ distribution near $\pm e/2c$.

increasing the magnitude of the electric field. This in turn, increases the polarization, moving the particle's electrochemical potential away from threshold. We propose, then, that the better expression of the effect of the medium polarization in producing the memory hole is "to provide V_D shifts away from threshold" rather than "to shift potentials to bring the Fermi levels into a better match."

We may now apply this model to understand the other experimental results mentioned earlier.

B. Square-wave cooling

For cooling in a single bias V_C the memory hole is made at the ends of $n(V_D;V_C)$, near the Coulomb thresholds at $\pm e/2c$. For particles with V_D near zero, the polarization shifts as many particles into a range of V_D as away from it. During a square-wave cooling where the cooling bias alternates between V_A and V_B, in general, particles with V_D near $\pm e/2c$ for one bias are away from threshold at the other bias. During the half of the cycle when the voltage is V_A, freezing polarization will cause gaps to develop in $n(V_D;V_A)$ near $V_D=\pm e/2c$. During the other half of the cycle, when the voltage is V_B, the freezing polarization causes gaps in $n(V_D;V_B)$ near $V_D=\pm e/2c$. The gap which has developed in $n(V_D;V_A)$ now appears as a dip somewhere inside the distribution $n(V_D;V_B)$. The freezing polarization shifts just as many particles into this region of V_D as away from it. When the voltage returns to V_A, the development of the gaps in the ends $n(V_D;V_A)$ continues. The freezing of each hole is analogous to the freezing of a single hole, but since only half of the polarization that freezes is effective in producing a given hole, each hole is of only half the depth of the single hole. The frequency of the square wave is irrelevant to the argument as long as it is below the inverse of the particle's charging time constant and well above the inverse of the cooling time. The holes in the distribution $n(V_D;V_A)$ or $n(V_D;V_B)$ will be measured in the capacitance-bias sweep according to Eq. (7) with either V_A or V_B replacing V_C.

If the square-wave bias were of amplitude exactly equal to e/c, this argument would suggest that one would obtain one hole of the same depth as for single bias cooling. Actually, the distribution of e/c in the sample assures that this condition cannot be met for all particles, though there will be some interference between the two capacitance features unless the amplitude is much greater than e/c.

C. Holes associated with freezing temperatures

This effect is similar to the square-wave cooling experiment. If the sample is cooled from temperature T_3 to T_2 in a bias V_A, gaps develop in the distribution $n(V_D;V_A)$ near $V_D=\pm e/2c$. When cooled from T_2 to T_1 in bias V_B, these gaps are now a dip somewhat in the middle of $n(V_D;V_B)$, and new gaps develop near $V_D=\pm e/2c$ for $n(V_D;V_B)$. The model predicts a hole width $\Gamma=kT_f/e$ at freezing temperature T_f. Thus the memory hole frozen in between T_3 and T_2 will be wider than that frozen in between T_2 and T_1. The data of Fig. 7 agree with this prediction. However, a quantitative effort to fit a theory to the shapes of the two holes must take into account the distribution of particle sizes. Moreover, the appropriate distribution changes with temperature. This is a result of the fact that because of the periodicity of the capacitance for fixed e/c, a hole wider than e/c cannot be made, so when $kT_f/e > e/c$, the hole is washed out. Therefore, only particles with $e/c > kT/e$ can contribute to a memory hole. Indeed, the data of Fig. 7 can also be explained with the assumption of a broad distribution of e/c. As the temperature is reduced, particles with smaller and smaller e/c contribute to the memory hole, making the net hole width narrower. We have not undertaken a theoretical fit to the data, but point out the possibility of making a memory hole that is narrow compared to the average e/c in the system as described in Sec. II, in connection with Fig. 6.

D. Magnitude of the effect

The effect that is seen at 4.2 K is the superposition of holes frozen-in through the entire temperature range of cooling. Equation (14) gives the depth H of the hole in $n(V_D;0)$ for polarization freezing near T_f. The accumulated capacitance hole

$$\Delta C(V_C,T_2,T_1,\tau_{\rm exp}) ,$$

created by cooling the sample in a single bias V_C from temperature T_2 to T_1, is obtained by combining Eqs. (8), (11), and (14):

$$\Delta C(V_C,T_2,T_1,\tau_{\rm exp})=\int_{T_1}^{T_2}\frac{e^2/c}{kT}\frac{d\epsilon_f}{dT}(\tau_{\rm exp},T)dTC_{P,u} . \tag{15}$$

To obtain an upper-bound estimate for the magnitude, we replace $d\epsilon_f/dT(\tau_{\rm exp},T)$ with the constant $\epsilon_\varnothing(T_2)-\epsilon_\varnothing(T_1)/T_2-T_1$, where $\epsilon_\varnothing(T)$ is the experimentally measured dielectric constant of a sample without particles at temperatures T. The resulting estimate is conveniently expressed as

$$\frac{\Delta C(V_C,T_2,T_1)}{C_{P,u}}\sim\frac{e^2/c}{k(T_2-T_1)}\frac{C_\varnothing(T_2)-C_\varnothing(T_1)}{C_\varnothing(T_1)}\epsilon_\varnothing(T_1)\ln\frac{T_2}{T_1} , \tag{16}$$

where $C_\varnothing(T)$ is the capacitance of a sample without particles at temperature T. The dependence on the ratio of T_1 and T_2 is in rough accord with the experimental observations, an example of which is the approximately equal magnitude of the two holes in Fig. 7. Between 77 and 4.2 K we observe a $\Delta C_\varnothing/C_\varnothing$ of $\frac{1}{50}$; $\epsilon_\varnothing$ is about 5; and $e^2/c/k\Delta T$ is of order unity. We obtain an upper bound of 30%. The fact that this result is about six times larger than the observed effect indicates that only a fraction of the temperature-dependent dielectric constant is involved in the freezing of polarization.

E. Fixed-temperature hole development

As mentioned earlier, the freezing of polarization is probably due to reorientation of defect or impurity states in the oxide. The characteristic rate $1/\tau$ for the reorientation of a two-state defect can be approximated by

$$\frac{1}{\tau}=\frac{1}{\tau_0}[e^{-E_B/kT}+E^{-2s(2ME_B/\hbar^2)^{1/2}}], \tag{17}$$

where M is the defect tunneling mass, and s, the tunneling distance. The left term on the right-hand side of Eq. (17) represents thermal activation while the right term is a Wentzel-Kramers-Brillouin (WKB) tunneling term. Defects responsible for the freezing-in of polarization between T_2 and T_1 have a time constant τ that is short compared to an experimental time $\tau_{\rm exp}$ at the higher temperature and long compared to $\tau_{\rm exp}$ at the lower temperature.

The hole development at fixed temperature occurs in the Lambe-Jaklevic experiment when the sample is subjected to a given bias for a time τ_A, that is long compared to the sweep time τ_S. Defects with τ such that $\tau_S<\tau<\tau_A$ contribute to the capacitance hole at the new bias.

In the capacitance-decay experiment, the hole development is seen in this way: If the sample rests for a long period of time in a bias V_A, the defects go to the lower-energy states that produce the gap in the distribution $n(V_D;V_A)$ at $e/2c$. When a large voltage step is applied, switching the bias from V_A to V_B, $V_A-V_B>e/c$, the capacitance C_P will rise sharply from the low value associated with the hole at $e/2c$ to the average capacitance $C_{P,u}$ away from the hole. After the step, the capacitance will slowly decay as the defects go into the states that produce a gap in the distribution $n(V_D;V_B)$ near $V_D=\pm e/2c$. Substituting Eq. (17) into Eq. (9) and converting the integral over τ to an integral over barrier energies E_B, we obtain for the time-dependent dielectric constant

$$\epsilon_f(T,t)=\int_0^\infty \epsilon_p\left[1-\exp\left[\frac{-t}{\tau_0}(e^{-E_B/kT}+e^{-2s(2mE_B/\hbar^2)^{1/2}})\right]\right]N(E_B)dE_B\ , \tag{18}$$

where $N(E_B)$ is the number of defects with a characteristic barrier energy between E_B and E_B+dE_B, and which contribute ϵ_p to the dielectric constant. Integrating over a uniform distribution of E_B, $N(E_B)=n_0$, yields a time dependence that is logarithmic in the thermally activated limit and nearly logarithmic in the tunneling limit. For the thermally activated regime, the result is

$$\epsilon_f(T,t)=\epsilon_p n_0 kT\ln(t/\tau_0)\ . \tag{19}$$

The capacitance decay $C_P(t)$ is given by

$$C_P(t)=C_{P,u}\left[1-\frac{e^2/c}{kT}\epsilon_f(T,t)\right]\ . \tag{20}$$

Using Eq. (19) we obtain the slope C_S of the logarithmic decay

$$C_S=\frac{dC}{d\log_{10}(t/\tau_0)}=(e^2/c)n_0\epsilon_p C_{P,u}/\ln(10)\ . \tag{21}$$

Using the same model, the magnitude of a capacitance hole $\Delta C(V_C,T_2,T_1)$, frozen between temperatures T_2 and T_1, is obtained by differentiating Eq. (19) with respect to temperature and using the result in Eq. (15). The result is

$$\Delta C(V_C)=(e^2/c)n_0\epsilon_p\ln\frac{T_2}{T_1}\ln\frac{\tau_{\rm exp}}{\tau_0}C_{P,u}\ . \tag{22}$$

The ratio of $\Delta C(V_C)$ to the slope given by Eq. (21) is

$$\ln\frac{T_2}{T_1}\ln\frac{\tau_{\rm exp}}{\tau_0}\Big/\ln(10)\ .$$

This relationship is valid in a temperature window defined at high temperature by the condition $kT<e^2/c$, and at low temperature by the neglect of tunneling of defects in the oxide. Experimentally, due to the distribution of e/c in the sample, this window is too narrow to support a quantitative analysis of the data. A qualitative comparison of this ratio with slope, 6 pF/decade, of the decay seen in Fig. 9 and the hole depth $\Delta C(V_C)\sim 10$ pF seen in the same sample, yields an experimental ratio that is too small by about a factor of 10. This may indicate that the distribution $N(E_B)$ is not constant, but rather increases with decreasing E_B.

F. Lowest temperature for hole freezing

Setting the activated and tunneling rates in Eq. (17) equal, and equal to $1/\tau_{\rm exp}$, determines a characteristic temperature below which it is no longer possible to freeze defect states. This temperature $T_{\rm min}$ is

$$T_{\rm min}=\left[\frac{1}{2s}\right]^2\frac{\hbar^2\ln\frac{\tau_0}{\tau_{\rm exp}}}{2k}\ . \tag{23}$$

Using $\tau_0/\tau_{exp} \sim 10^{-14}$, $s \sim 1$ Å, and for M, the mass of an oxygen ion, we obtain a temperature on the order of a few degrees Kelvin. This rough estimate agrees with experimental results we have obtained on the temperature dependence of the Lambe-Jaklevic memory hole which saturates at about 2 K. These observations are in rough agreement with those of Rogers and Buhrman[8] who see a crossover from activated to tunneling behavior at about 10 K.

V. DISCUSSION AND CONCLUSIONS

We have examined the problem of electric-field memory effects in two different granular film systems with the purpose of understanding the physics of the memory hole-freezing process. A key question raised in considering previous experiments is: How is it that oxide polarization charges, which respond to electrostatic potentials, end up having a systematic effect on the electrochemical potential distribution which has to do with the work functions of the particle and electrode?

The answer, we believe, lies in the quantized nature of the charge transfer process. As the polarization near a particular particle develops in response to a change in bias, the difference in electrochemical potentials of particle and electrode V_D gradually changes. For most of the particles this shift simply moves them uniformly in the V_D distribution $n(V_D;V_C)$ without changing the distribution density. There are some particles, however, for which the difference is increased to the Coulomb threshold at $\pm e/2c$, where a charge transfer results. The charge transfer causes a further large step in the electrostatic potential ϕ and thus, electric field in the dielectric. It is this large step that causes the additional polarization that shifts the particle further into the distribution leaving a gap at the ends of $n(V_D;V_C)$. The effect of the polarization is to shift particles' V_D away from threshold rather than towards zero. The oxide charges are able to sense the Fermi level distribution because those charges near a particle with V_D close to the Coulomb threshold experience a jump in electric field, while those near particles away from the threshold do not.

The model provides a satisfactory qualitative explanation of all experimental results obtained by Lambe-Jaklevic and us on tunnel capacitors:

(a) the existence of the memory hole $\Delta C(V)$ centered at the cooling bias V_{SC};

(b) the observation of two holes, centered about V_A and V_B, when a sample is cooled in a square-wave bias that alternates between V_A and V_B;

(c) the association of memory holes with the temperature window used to create them;

(d) the magnitude of the memory hole $\Delta C(V_C)$;

(e) the approximately logarithmic time dependence of hole development at a fixed temperature;

(f) lowest temperature for hole-freezing T_{min}.

Further experimental work could determine if the memory effect is a useful way to study defects. If a narrower size distribution of particle-electrode capacitances could be produced, the capacitance memory signal frozen in different temperature ranges would give information about the barrier energy and concentration of defects. The capacitance decay signal due to tunneling of defects can provide complementary evidence about the defect tunnel barrier distribution for small barrier heights. In the samples we studied this was a nearly uniform distribution.

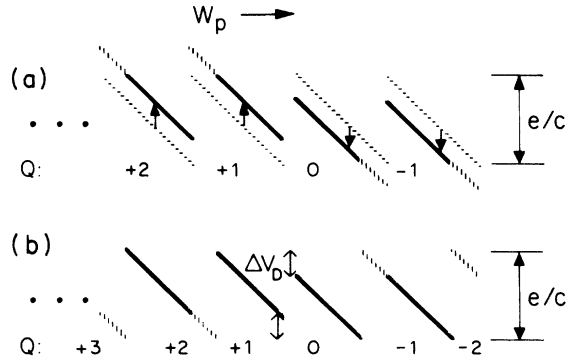

FIG. 13. The microscopic analog of Fig. 12 with an ensemble of particles with distributed W_p (horizontal axis). Each particle is associated with a dipole defect with a crossover field E_c that satisfies $0 < E_c < e/(cd)$. (a) Horizontally-hatched lines indicate state before defect "freezes" where there is no net effect of the defect on $n(V_D;0)$. Arrows indicate shift ΔV_D associated with dipole freezing into the lowest-energy state for the particle-defect-electrode system. This value of E_c causes an upward shift for particles with $Q > 0$ and a downward shift for particles with $Q \leq 0$. Vertically-hatched lines indicate particles for which a charge transfer is induced. (b) After charge transfers have occurred. A gap only develops in $Q = +1$ and $Q = +0$ states.

The other model discussed here, the Coulomb glass, predicts effects that are inconsistent with the major experimental results. In the Lambe-Jaklevic device, the dipole-dipole nature of interparticle interaction greatly reduces the magnitude of Coulomb glass effects. In the Adkins *et al.* device, however, it may be possible to observe the kind of nonequilibrium effects described here. It may also be possible to make a similar structure with a lightly-doped, compensated semiconductor on the insulating side of the transition which might show this type of effect. Experiments of this type have been recently carried out by Monroe *et al.*[10]

ACKNOWLEDGMENTS

This research has been supported by the Semiconductor Research Corporation, the National Science Foundation under Grant No. DMR-80-08546, the Materials Science Center under Grant No. DMR-82-17227, and The Cornell Program on Submicron Structures.

APPENDIX: MICROSCOPIC VERSION

The memory effect depends on the properties of the dielectric medium surrounding the particle. Given the small size of the particle, one may expect local variations in the field and temperature dependence of the dielectric constant to be important in determining how a particle's V_D shifts during cooling. For this reason, we consider a

microscopic model for the freezing-in process. This model may be considered a closely related submodel of the homogeneous dielectric model, as most all of the physics described earlier applies here. In addition, the model provides an appropriate context for discussing experimental results which have to do with the detailed properties of the dielectric. An example is tunneling effects in the dielectric, which must be discussed in terms of a microscopic model.

The oxide surrounding the particle may be expected to contain defects which may have a thermally-activated rotational or translational mobility, for example, a dipole with two stable orientations, or an ion moving between sites. There is also the possibility of thermally- or photoexcited electrons moving to or from trap states. For simplicity, we consider a two-state dipole defect with a moment of magnitude p, which in state A is represented by the vector $-p\hat{i}$, and in state B by the vector $p\hat{i}$, where $\hat{i}$ is a unit vector perpendicular to the plane of the junction pointing from particle to electrode. The defect is assumed to alternate between its two states at a rate

$$\frac{1}{\tau}=\frac{1}{\tau_0}e^{-E_B/kT} , \tag{A1}$$

where $1/\tau_0$ is an "attempt" rate constant and E_B is an activation energy. The defect freezes in one of its two states at a temperature T_f, for which τ is longer than the relevant experimental time $\tau_{\exp}$.

The energy of the two states of the defect in an electric field, $\mathbf{E}\equiv E\hat{i}$ is given by

$$\begin{aligned} U_A &= -pE + W_A , \\ U_B &= pE + W_B , \end{aligned} \tag{A2}$$

where $W_{A,B}$ are site energies having to do with the local configuration for states A and B. The condition $U_A = U_B$ defines a crossover field,

$$\mathbf{E}_c = (W_A - W_B)/2p\hat{i} \equiv E_c\hat{i} .$$

For fields greater than E_c, the defect prefers state A and for fields less than E_c, it prefers state B.

We consider an ensemble of particles with a distribution of work functions, each associated with a defect of moment p and crossover field that satisfies $0 < E_c < e/(cd)$, where d is again the particle-electrode separation. As in Sec. IV A imagining the particles to be at zero temperature and setting $V_C = 0$, we allow the dipoles to orient according to the electric field for each of the charge states of particles in the distribution [Fig. 13(a)]. Only those particles in the two charge states whose oxide fields immediately bracket $\mathbf{E}_c$ [Fig. 13(b)] give rise to a gap in $n(V_D;0)$. This is contrasted with effective medium version of the model in Figs. 12(a) and 12(b), where gaps develop in $n(V_D;0)$ for particles in all charge states (except $Q=0$). The magnitude of the gap is

$$\Delta V_D = \frac{p}{ed}(e/c) . \tag{A3}$$

The quantity p/ed is equal to ϵ_p used in Eq. (18). To calculate the effects of nonzero temperature for $kT/e > e/c$, one can model a particle and a nearby defect as a four-state system, with the particle in one of two charge states, the ground state and the lowest thermally excited state, and the defect in either of its two states, A and B. Define V_D' to be the difference in Fermi levels between particle and electrode in absence of interaction with the defect, and V_D to be the difference in Fermi levels with the interaction turned on and in a frozen state characterized by the Boltzmann distribution over the four levels at the freezing temperature. For a particle with a given $V_D' < 0$, the four states and their energies are

$$\begin{aligned} U_{j,A} &= -pE_j + W_A , \\ U_{j,B} &= pE_j + W_B , \\ U_{j+1,A} &= e(e/2C + V_D') - pE_{j+1} + W_A , \\ U_{j+1,B} &= e(e/2C + V_D') + pE_{j+1} + W_B , \end{aligned} \tag{A4}$$

where j is the charge state of the ground state in the absence of defect polarization, determined by the difference in work functions between particle and electrode, $j+1$ is the lowest excited charge state, and E_j is the electric field between particle and electrode given by $j(e/c/d)$. We can use this model to calculate the depth H of the hole in $n(V_D;0)$ in an ensemble of particles fluctuating between charge states j and $j+1$, with $V_D' < 0$, and associated with defects that have E_C halfway between E_j and E_{j+1}. For these particles, the lowest-energy state of the particle-defect system is when the charge state is j and the defect state is A. A gap should appear at the Coulomb threshold $-e/2c$, as in Fig. 13(b). At $T>0$ some of the defects associated with these particles will freeze in state B. The fraction of these, $f(V_D = -e/2c, n)$ which result in a V_D frozen-in at $-e/2c$, is the number that in absence of the interaction had $V_D' = -e/2c + \Delta V_D$ and is given by

$$\begin{aligned} f\left[V_D = -\frac{e}{2C}, j\right] &= \frac{e^{-U_{j,B}/kT_f} + e^{-U_{j+1,B}/kT_f}}{Z} \\ &= \frac{2e^{-e\Delta V_D/kT_f}}{Z} , \end{aligned} \tag{A5}$$

where Z is the partition function for the four states, and Z, $U_{j,B}$, and $U_{j+1,B}$ are evaluated at $V_D' = -e/2c + \Delta V_D$.

One finds with this model, a hole in the ends of $n(V_D;0)$ of a width Γ and depth H that depend on freezing temperature

$$\begin{aligned} \Gamma &= \Delta V_D \quad kT_j < e\Delta V_D , \\ &= kT_f/e \quad e\Delta V_D < kT_f < e^2/c . \end{aligned} \tag{A6}$$

$$\begin{aligned} H &= (1 - e^{-e\Delta V_D/kT_f}) n_u \quad kT_f < e\Delta V_D , \\ &= \left[\frac{e\Delta V_D}{kT_f}\right] n_u \quad e\Delta V_D < kT_f < e^2/c . \end{aligned} \tag{A7}$$

For the range $e\Delta V_D < kT_f < e^2/c$, the result is essentially the same as that obtained in the effective medium version of the model Eqs. (13) and (14). If kT_f is greater than or

on the order of e^2/c, the four-state model will break down as thermal fluctuations excite the particle to several charge states.

The two versions of the polarization model give somewhat different pictures of which particles are involved in the freezing of a memory hole. In the linear model, all particles near the appropriate end of the V_D distribution are involved in creating the gap in $n(V_D;V_C)$, where in the microscopic picture, only particles in charge states with electric fields near $\mathbf{E}_C$ participate. In general, there will be a distribution of $\mathbf{E}_C$ which depends on the characteristics of the oxide and its defects. The shape of this distribution may have experimental consequences. For example, if $\mathbf{E}_C$ is equal to zero for all defects, one could cool the sample in a bias strong enough to saturate all of the dipoles in the thicker oxide layer between the particle and distant electrode, thereby reducing the hole size. This effect is not seen, however, indicating that the distribution of $\mathbf{E}_C$ is broad.

---

*Present address: AT&T Bell Laboratories, Murray Hill, NJ 07974.

[1]J. Lambe and R. C. Jaklevic, Phys. Rev. Lett. **22**, 1371 (1969).

[2]C. J. Adkins, J. D. J. D. Benjamin, J. M. D. Thomas, J. W. Gardner, and A. J. McGeown, J. Phys. C **17**, 4633 (1984).

[3]R. E. Cavicchi and R. H. Silsbee, Phys. Rev. B **37**, 706 (1988).

[4]I. Giaever and H. R. Zeller, Phys. Rev. Lett. **20**, 1504 (1968).

[5]R. E. Cavicchi and R. H. Silsbee, Phys. Rev. Lett. **52**, 1453 (1984).

[6]A. L. Efros and B. I. Shklovskii, J. Phys. C **8**, L49 (1975).

[7]B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer-Verlag, New York, 1984).

[8]J. Davies (private communication).

[9]C. T. Rogers and R. A. Buhrman, Phys. Rev. Lett. **55**, 859 (1985).

[10]D. P. Monroe, A. C. Gossard, J. H. English, B. Golding, W. H. Haemmerle, and M. A. Kastner, Phys. Rev. Lett. **59**, 1148 (1987).

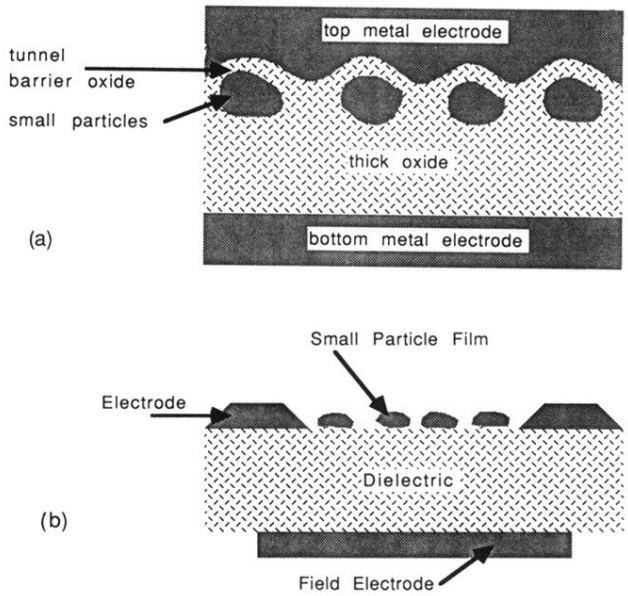

FIG. 1. The sample configurations of (a) Lambe-Jaklevic—a tunnel capacitor, and (B) Adkins *et al.*—a granular film field-effect device.