Analysis of terrace-width distributions using the generalized Wigner surmise: Calibration using Monte Carlo and transfer-matrix calculations

Hailu Gebremariam, Saul D. Cohen,* Howard L. Richards,† and T. L. Einstein‡
Department of Physics, University of Maryland, College Park, Maryland 20742-4111, USA
(Received 12 September 2003; revised manuscript received 18 December 2003; published 12 March 2004)

Measurement of terrace-width distributions (TWD's) of vicinal surfaces is used routinely to find the dimensionless strength $\tilde{A}$ of the elastic repulsion between steps. For sufficiently strong repulsions, the TWD can be described by a Gaussian about the mean step spacing, but controversy has arisen on the correct prefactor in the relation of the TWD variance to $\tilde{A}$. Instead of the various Gaussian approximations, we have advocated for several years that the TWD be fit with the generalized Wigner distribution, essentially a gamma distribution in the normalized squared TWs. The basis for this idea stems from a mapping of the step model to the Sutherland model of fermions in one dimension. While several applications to experiment have been successful, definitive comparison of the various approximations requires high-quality numerical data. We report transfer matrix and extensive Monte Carlo simulations of terrace-step-kink models to support our contentions. Our work includes investigation of finite-size effects and of the breakdown of the continuum picture for values of $\tilde{A}$ larger than in typical experiments.

DOI: 10.1103/PhysRevB.69.125404 PACS number(s): 68.35.Md, 05.40.−a, 68.37.Ef, 68.35.Bs

I. INTRODUCTION

The last decade has witnessed *quantitative* measurement of the widths ℓ of terraces on vicinal surfaces. The resulting data for the terrace width distribution (TWD) provides arguably the best way to gauge the strength of the elastic repulsion between steps, specifically the coefficient A (which has units of energy times length) of the elastic repulsion A/ℓ^2. In standard analyses two simplifications are routinely made: (1) The elastic repulsion is taken to act just in the $\hat{x}$ direction, perpendicularly to the mean step direction $\hat{y}$ [in "Maryland notation"]. This formulation should be a good approximation when the step meandering is relatively modest, with variation scale in $\hat{y}$ that is "slow" (large) compared to ℓ. The preceding formulation of the repulsion as a function of ℓ implicitly uses this ansatz. (2) Individual steps can be well characterized in terms of their stiffness $\tilde{\beta}$, which is inversely proportional to their diffusivity. This requires that $x(y)$ for a step be single valued and that $x'(y)$ be small. Typically, one invokes the step continuum approximation, in which a continuum approximation is made along the $\hat{y}$ direction.[1] Thereafter, A appears only in the form of a *dimensionless interaction strength*

$$\tilde{A} \equiv A\tilde{\beta}(k_B T)^{-2}, \tag{1}$$

where $\tilde{\beta}$ is the step stiffness.

To characterize a TWD obtained from experiment or numerical simulation, it is typically deemed sufficient to specify just its variance σ^2 or some similar measure of the width of the single peak. At least when $\tilde{A}$ is not small, this procedure is adequate since the TWD's shape can be adequately approximated by a Gaussian. Justification of this form can readily be derived from a mean-field-like (Gruber-Mullins) approach,[2,3] which indicates that the variance varies inversely with $\tilde{A}^{1/2}$. Two recent theories from other perspectives produce the same proportionality, but with different proportionality constants. Subsequently, we recognized that the TWD might better be described using the generalized Wigner surmise rather than a Gaussian. In actual experimental systems $\tilde{A}$ is typically between 0 and 10, sometimes in the teens.[4,5] (While occasional values up to nearly 4000 have been reported[6] for $\tilde{A}$, our belief is that values above about 20 indicate anomalous behavior.) Exact theoretical results for the TWD (with $\tilde{A} \geqslant 0$) are available only for $\tilde{A}=0$ and $\tilde{A}=2$.[7,8] Hence, resolution of the controversy about the optimal viewpoint of TWDs requires high-quality numerical simulation.

In this paper we report the results of simulations over several years to produce data adequate to confront the controversy convincingly. Our primary method has been to do long Monte Carlo calculations for large lattices using the simplest reasonable model for step fluctuations. This work is complemented by calculations using transfer-matrix methods. Our primary focus is on the predictions of the variance made by the various theoretical approximations; we analyze the results in a way that highlights the differences between the predictions. This work is a considerable advance over earlier preliminary remarks on this problem.[9] In previous papers we discussed relevant analytic results for the various approximations,[10] applications of the generalized Wigner distribution to experimental data,[5,9,11] and how these studies fit in general research on random matrix theory.[12] In this paper we limit our discussion to aspects bearing on our theme of using numerical studies of standard models to test the relative merits of the various approximations for the TWD.

In the following section, we offer a succinct review of the physics of the various approaches and the key formal results. Readers familiar with earlier (and more expanded) expositions of most of this material can use the equations and two newly compiled tables to refresh their recollections. In Sec. III, we describe the models used for the computations and

 ©2004 The American Physical Society

TABLE I. Tabulation of the prefactors K_X that relate the variance to the inverse root of the dimensionless interaction strength $\tilde{A}$ [cf. Eq. (3)], or to $\tilde{A}_{\rm eff}$ [so $(\varrho/2)^2$] if the "eff?" entry is Y. The square brackets for the Wigner listing recall that Eq. (5) differs from Eq. (2) underlying the other entries.

Model	Approximation	Reference	X	eff?	$K_{X({\rm NN})}$	$K_{X({\rm all})}$
Gruber-Mullins	Single active step	2	GM	N	$1/\sqrt{48}\doteq 0.144$	$\sqrt{15/8\pi^4}\doteq 0.139$
Grenoble	Neglect entropic interactionn	14,15	E0	N	$\sqrt{2/3\pi^2}\doteq 0.260$	$\frac{1}{\pi}\int_0^{2\pi}\frac{1-\cos\phi}{\phi(2\pi-\phi)}\doteq 0.247$
Grenoble, modified	Average entropic interactionn	14,15,10	EA	Y	$\sqrt{2/3\pi^2}\doteq 0.260$	$\frac{1}{\pi}\int_0^{2\pi}\frac{1-\cos\phi}{\phi(2\pi-\phi)}\doteq 0.247$
Saclay	Roughening theory [VGL (Ref. 18)]	16,20,21	R	Y		$2/\pi^2\doteq 0.203$
Wigner	[gamma-like, not Gaussian distribution]	7,10	W	[Y]		[1/4]

the calculational procedures we followed. Section IV, which presents our results, is the heart of the paper. It begins with our numerical evidence favoring the use of the Wigner expression rather than a Gaussian to describe the TWD and proceeding to examine finite-size effects, the breakdown of the continuum picture, and temperature effects. The concluding section briefly recapitulates our findings and their implications.

II. REVIEW OF FORMAL BACKGROUND

A. Gaussian approximations to TWD's

It is convenient to divide ℓ by its average value, thus constructing the dimensionless parameter $s\equiv\ell/\langle\ell\rangle$. Then the TWD $P(s)$ is not just normalized but has unit mean. Moreover, for steps with just A/ℓ^2 repulsions, $P(s)$ is expected to be insensitive to $\langle\ell\rangle$, so that the rescaled TWD's for different misorientations of the same surface are then expected to coincide.[13] It was long conventional to describe TWD's by Gaussian distributions:

$$P(s)\approx P_G(s)\equiv\frac{1}{\sigma_G\sqrt{2\pi}}\exp\left[-\frac{(s-1)^2}{2\sigma_G^2}\right]. \tag{2}$$

Use of a Gaussian distribution can be justified in several ways, assuming the elastic repulsion between the steps is strong enough to confine the motion of each step to a parabolic well near its mean position. The venerable Gruber-Mullins (GM) treatment, which allows only one step of the vicinal surface to meander, finds that[2,3]

$$\sigma^2=K_X\tilde{A}^{-1/2}, \tag{3}$$

where the subscript X anticipates that different proportionality constants, tabulated in Table I, will result from different approximation schemes, each of which is given a mnemonic label X.

In rederiving Gaussian TWD's from different perspectives, two groups ascertain that the variance in Eq. (3) using $K_{\rm GM}$ underestimates (for given $\tilde{A}$) the true variance: The Grenoble group[14,15] start with the insight that the *entropic* repulsion becomes relatively less important as the energetic repulsion increases (because large energetic repulsions diminish the chance of neighboring steps approaching each other), becoming negligible at very large $\tilde{A}$ (hence the notation $X={\rm E0}$). If *both* steps bounding a terrace fluctuate independently, then the variance of the TWD should be the *sum* of the variances of the fluctuations of each step, i.e., *twice* the variance obtained in the Gruber-Mullins picture. This factor is reduced modestly by corrections due to the (anti)correlations[16] of neighboring steps. If entropic repulsions are included in an average way (denoted $X={\rm EA}$) (Ref. 10) rather than discarded, the range applicability of the asymptotic form for σ^2 is greatly extended. Explicitly, $\tilde{A}$ is replaced in Eq. (3) by an effective interaction strength $\tilde{A}_{\rm eff}$ obtained from the cubic term of the expansion of the projected free energy of a vicinal surface as a function of misorientation slope:[17]

$$\frac{\tilde{A}_{\rm eff}}{\tilde{A}}\equiv\frac{1}{4\tilde{A}}(\sqrt{4\tilde{A}+1}+1)^2\sim 1+\tilde{A}^{-1/2}+\frac{\tilde{A}^{-1}}{2}+\cdots. \tag{4}$$

Then Eq. (3) is modified to $\sigma^2\sim K_{\rm EA}\tilde{A}_{\rm eff}^{-1/2}$, with values for $K_{\rm EA}$ the same as $K_{\rm E0}$ (cf. Table I).

From a capillary-wave treatment in the $\hat{x}$ as well as $\hat{y}$ directions[18,19] and with correlation functions drawn from the celebrated VGL (Villain-Grempel-Lapujoulade[18]) theory of roughening (therefore denoted $X=R$) on vicinals, the Saclay group[16,20,21] obtained the "$\tilde{A}_{\rm eff}$ modification" of Eq. (3).

Note finally that the TWD is narrower, and so K_X is smaller (cf. Table I), when all steps, rather than just nearest neighbors (NN), are taken to have A/ℓ^2 repulsions. Since the entropic interaction *ipso facto* just involves NN steps, the difference is larger for the Grenoble approximations than for Gruber-Mullins, but still relatively unimportant. In experiments, one typically obtains the variance by measuring the width of the TWD, which is proportional to $\tilde{A}_{[\rm eff]}^{-1/4}$. Thus, a seemingly large 10% variation in $\tilde{A}$ translates to a few percent variation in the width, well within experimental error. Conversely, this hinders the accuracy with which $\tilde{A}$ can be estimated from experiment.

In short, since the various approaches leading to Gaussian TWD's make different fundamental approximations, the predicted relationships between the TWD width and $\tilde{A}$ differ notably; a good fit of a TWD by a Gaussian does not imply an unambiguous estimation of $\tilde{A}$.

B. Wigner approximation to TWD's: CGWD

The exact results for TWD's come from mapping the spatial configurations of steps to the world lines of spinless fermions in the Calogero-Sutherland model.[22,23] When this model is integrable, the fluctuating positions correspond to the *energy* eigenvalues of random matrices for ensembles with orthogonal ($\tilde{A}=-1/4$), unitary ($\tilde{A}=0$), or symplectic ($\tilde{A}=2$) symmetry.[7,8] The TWD then corresponds to the joint probability density function of the eigenvalues.

For those three cases the TWD's can be very well approximated by the so-called Wigner surmise,[8,24]

$$P_\varrho(s)=a_\varrho s^\varrho \exp(-b_\varrho s^2), \tag{5}$$

where the constants a_ϱ deriving from normalization of $P(s)$ and b_ϱ associated with unit mean are

$$a_\varrho=\frac{2\left[\Gamma\left(\frac{\varrho+2}{2}\right)\right]^{\varrho+1}}{\left[\Gamma\left(\frac{\varrho+1}{2}\right)\right]^{\varrho+2}} \quad \text{and} \quad b_\varrho=\left[\frac{\Gamma\left(\frac{\varrho+2}{2}\right)}{\Gamma\left(\frac{\varrho+1}{2}\right)}\right]^2 \tag{6}$$

for the values $\varrho=1$, 2, and 4, respectively.

For arbitrary ϱ there is no symmetry-based justification of the Wigner distribution of Eq. (5). Nonetheless, we have argued that it provides a viable, arguably optimal interpolation scheme between the two special values of ϱ and also out to the Grenoble expression for nearly infinite repulsion;[9,10] we have also used it successfully to analyze experimental data.[5,9,11] For brevity, we refer hereafter to this set of formulas, Eqs. (5) and (6), as the CGWD (continuum generalized Wigner distribution). Recently some of us have been able to derive[25] the CGWD by considering the two-particle Calogero model[22] (harmonically bound interacting spinless fermions on a line) (rather than the infinite-particle Calogero-Sutherland model[23] of fermions on a circle).

The variance σ_W of the CGWD can be expressed simply in terms of b_ϱ. From Eq. (6) it follows that

$$\sigma_W^2=\frac{\varrho+1}{2b_\varrho}-1\approx\frac{1}{2\varrho}-\frac{3}{8\varrho^2}+\frac{3}{16\varrho^3}-\frac{7}{384\varrho^4}+\mathcal{O}(\varrho^{-5}) \tag{7}$$

for large values of ϱ, as given in Eq. (A8) of Ref. 11. Based on the mapping of the step problem onto the Sutherland Hamiltonian,[23] we recognize that

$$\tilde{A}=\frac{\varrho}{2}\left(\frac{\varrho}{2}-1\right), \quad \Rightarrow \varrho=2\sqrt{\tilde{A}_{\text{eff}}}. \tag{8}$$

Thus, when $\tilde{A}$ is large, we find in essence that $K_W=1/4$ in the modified Eq. (3); K_W agrees rather well with the modified Grenoble result K_{EA} (esp. $K_{\text{EA(all)}}$).

To accentuate the difference between the various methods of estimating $\tilde{A}$, we shall find it useful below [cf. Fig. 2(b) as well as Fig. 3] to scrutinize $\sigma^2\varrho$. From Eq. (7) we find that according to the CGWD, this quantity is a slowly varying, monotonically increasing function of ϱ and so of $\tilde{A}$, with an asymptotic value of 1/2. In contrast, for schemes which obey $\sigma^2=K_X\tilde{A}_{\text{eff}}^{-1/2}$, $\sigma^2\varrho$ is just a constant $2K_X$.

From Eqs. (7) and (8) (inverted), one can estimate the variance from $\tilde{A}$, but experimentalists usually seek the reverse. An excellent estimate[11] of $\tilde{A}_{\text{W}}$ from the variance, based on series expansion of Eq. (7), is

$$\tilde{A}\approx\frac{1}{16}\left[(\sigma^2)^{-2}-7(\sigma^2)^{-1}+\frac{27}{4}+\frac{35}{6}\sigma^2\right], \tag{9}$$

with all four terms needed to provide a good approximation over the full physical range of $\tilde{A}$. The Gaussian methods described earlier essentially use just the first term of this expression and adjust the prefactor. When $\tilde{A}$ is not weak (basically when $\tilde{A}$ is above 2—see Ref. 11 for details), $P_\varrho(s)$ approaches a Gaussian. Thus, in that limit, to determine the value of ϱ (and thence $\tilde{A}$) underlying experimental data, a reasonable alternative to fitting with Eq. (5) is to fit with a Gaussian [albeit with poorer χ^2 than for a fit with $P_\varrho(s)$], extract the variance, and then apply Eq. (9). Alternatively, one can take advantage of ''canned'' statistical software packages by straightforwardly recasting Eq. (5) as a standard gamma distribution function in the random variable $b_\varrho s^2$ with an exponent $(\varrho+1)/2-1$.[26]

C. Synopsis of prior theoretical results

Before proceeding to the numerics, let us summarize what we know. At $\tilde{A}=0$ and 2, the exact solutions lead to variances of about 0.1800 (Refs. 7, 8, and 13) and—as some of us have recently shown[27]—0.1041, respectively. (More such information is tabulated in Refs. 9 and 10.) The variances of the CGWD are nearly identical: as listed in Table II, they are slightly too small for small $\tilde{A}$ and eventually become barely too large by $\tilde{A}=2$. At $\tilde{A}=0$ the Gaussian approximations are fundamentally invalid. At $\tilde{A}=2$ Table II shows that the Saclay and GM estimates are somewhat too low whereas the Grenoble estimate is far too high. For very large $\tilde{A}$, the argument underlying the modified Grenoble viewpoint becomes compelling. The CGWD variance approaches it nicely.[28] The Saclay estimate of σ^2 here is significantly too small. Thus, we know before starting that the CGWD provides the best estimate at $\tilde{A}=2$ and approaches the correct limit for very large $\tilde{A}$.[9,10] Qualitatively it certainly captures the global behavior of variance as a function of $\tilde{A}$. Our chief concern is how accurately the CGWD estimate embodied in Eq. (9) reproduces the ''crossover'' behavior.

TABLE II. Comparison of various continuum-limit analytic approximations for the variance [using tabulated results (Ref. 10)] with the values σ_*^2 available at special values of $\tilde{A}$ from exact solutions. For $\tilde{A}=2$ we have recently calculated that $\sigma_*^2 \doteq 1.041$ (Ref. 27). For asymptotically large $\tilde{A}$, the Grenoble viewpoint is used for σ_*^2; however, the continuum picture does not describe adequately the discrete nature of physical systems and numerical models for (unphysically) large $\tilde{A}$.

$\tilde{A}$	$-1/4$	0	2	$\to\infty$
ϱ	1	2	4	$\to 2\tilde{A}^{1/2}$
$\sigma^2_{\text{Wigner}}/\sigma_*^2-1$	$-4\frac{1}{2}\%$	-1%	$+0.4\%$	$+1\%$
$\sigma^2_{\text{EA(all)}}/\sigma_*^2-1$			$+68\%$	0%
σ_R^2/σ_*^2-1			-3%	-4%
$\sigma^2_{\text{GM(all)}}/\sigma_*^2-1$		$[-27\%]$	-5.8%	-44%

III. MODELS AND NUMERICAL METHODS

A. TSK model and variants

For our numerical work we use models that are *discrete* in the $\hat{y}$ as well as the $\hat{x}$ directions.[3,16,21,29] For simplicity we consider a vicinal simple cubic lattice with unit lattice constant. Periodic boundary conditions are imposed in both directions. The length of the system in the $\hat{y}$ direction is denoted by L_y and the number of steps by N; $L_x \equiv N\langle \ell \rangle$. The position of nth step, relative to its mean position, is written $x_n(y)$.

For numerical simulations, the most elementary model that contains the necessary physics is the terrace-step-kink (TSK) model. In the TSK model the only thermal excitation is kinks of energy ϵ along the steps: the terraces have no adatoms or vacancies. Then $x_n(y)$ must be a single-valued function and the Hamiltonian contains term $\epsilon|x_n(y+1)-x_n(y)|$. The stiffness $\tilde{\beta}_{\text{TSK}}$ of an isolated step is $2k_BT\sinh^2(\epsilon/2k_BT)$.[29] The noncrossing constraint requires $x_{n+1}(y) > x_n(y)$. In addition to the resulting entropic repulsion, there is an energetic term due to elastic repulsions: $A/[x_{n+1}(y)-x_n(y)]^2$. For our numerical simulations, we almost always limit this repulsion to neighboring steps, a common simplification in Monte Carlo;[3,20] we did some test runs to assess the change produced by also including second-neighbor steps.

The above-noted noncrossing constraint forbids overlap of edge elements from adjacent steps but does allow corners to touch. In addition to standard edge exclusion, we separately consider the case of ''corner exclusion'': configurations with $x_{n\pm1}(y+1)=x_n(y)$ forbidden. Obviously, this distinction is most significant for weak $\tilde{A}$ and for small $\langle \ell \rangle$.

In performing the transfer matrix calculations discussed in Sec. III C, we use the restricted terrace-step-kink model (RTSK), a variant of the TSK model with the additional constraint $|x_{n+1}(y)-x_n(y)| \leq 1$. In the RTSK model, the step stiffness $\tilde{\beta}$ is $(k_BT/2)[2+\exp(\epsilon/k_BT)]$.[29] Since the fluctuations are limited in the RTSK model, the ratio $\tilde{\beta}_{\text{RTSK}}/\tilde{\beta}_{\text{TSK}}$ must perforce be above one. Explicitly,

$$\frac{\tilde{\beta}_{\text{RTSK}}}{\tilde{\beta}_{\text{TSK}}} = \frac{1+2\exp(-\epsilon/k_BT)}{[1-\exp(-\epsilon/k_BT)]^2} \to 1+4e^{-\epsilon/k_BT}+\cdots \tag{10}$$

differs insignificantly from unity at low temperatures. The ratio increases monotonically, eventually becoming $3(k_BT/\epsilon)^2$ to leading order at temperatures well above those at which the assumptions underlying the TSK and especially the RTSK models become unphysical. While $\tilde{\beta}_{\text{TSK}}$ decays monotonically with increasing T, $\tilde{\beta}_{\text{RTSK}}$ has a minimum around $T=0.7$ (as the diffusivity saturates); however, $\tilde{\beta}_{\text{RTSK}}/(k_BT)^2$ does still decay monotonically. In any case, for the pair of temperatures used most often in numerical work, $k_BT/\epsilon=0.5$ and 0.84, we note that $\tilde{\beta}_{\text{RTSK}}/\tilde{\beta}_{\text{TSK}}=1.70$ and 3.32, respectively.

B. Monte Carlo approaches

We have studied the TSK model using standard Monte Carlo methods.[30] We check that the equilibration time from initially straight steps is several times the autocorrelation time as measured in equilibrium. Corresponding to attachment/detachment-limited transport, update attempts occur independently at randomly chosen single sites (along step n at coordinate y). Most of our results are based on the conventional, straightforward Metropolis algorithm.[31] We also coded the ''refusal-free'' n-fold way algorithm,[32,33] which is much more efficient at low T (Ref. 30) or large $\tilde{A}$. The main disadvantage of the n-fold way is that the tabulation of energy classes becomes too cumbersome for large $\langle \ell \rangle$. While helpful in our preliminary runs, the number of classes needed for multistep step interactions overwhelmed the advantages in our production runs.

In our simulations, the temperature was generally set so that $k_BT/\epsilon=0.5$, even though the actual value is about 0.2 for the copper vicinals we have studied most extensively,[5,11] in order that the dynamics be reasonably rapid. In pursuing convincing numerical data, we performed three levels of simulations. First we carried out a (rather extensive) set of preliminary runs with $N=10$ and $\langle \ell \rangle = 6$.[9]

Next, we used systems with $N=40$ and $\langle \ell \rangle = 12$, so that $L_x=480$, a preliminary discussion of which appeared earlier.[9] To insure that L_y, the size of the system along the mean step direction, is much larger than y_{coll}, the characteristic distance between close approaches at this temperature, we generally set $L_y=250$: For the TSK model, $y_{\text{coll}}=\langle \ell \rangle^2\tilde{\beta}/4k_BT=(\langle \ell \rangle/2)^2\sinh^2(\epsilon/2k_BT)$,[29] which is about 50 at $k_BT/\epsilon=0.5$ and $\langle \ell \rangle = 12$. We doubled $\langle \ell \rangle$ from the initial runs to assure that discreteness effects[11] played a negligible role for physical values of $\tilde{A}$. In some cases of unphysically large $\tilde{A}$ (above 50), it was necessary to use $\langle \ell \rangle = 18$. This issue is discussed further in Sec. IV C.

Third, in our most extensive calculations, we used $\langle \ell \rangle = 10$ and, in most cases, $L_y=2000$ and $N=100$. (For $\tilde{A}=50$, we found it safest to increase to $L_y=3000$ and $N=200$.) We used a standard high-quality random-number generator [Ran3 (Ref. 34)] and averaged over 100 runs using

different initial seeds. In these runs the variance reached its steady-state value after about 3000 Monte Carlo steps (MCS); we started ''taking data'' after 10 000 MCS, recording results every 10 MCS until reaching 30 000 MCS.

C. Transfer matrix approach with RTSK

Since the finite-size limitation on L_y hampers, at least in the earlier stages, our exploration of low T or large $\tilde{A}$ (even with the n-fold way), we have also used the transfer-matrix (TM) method[35] to study TWD's.[11] This method *ipso facto* has infinite L_y; L_x is limited by the size of matrices that available computers can handle, on use of symmetries or sparse-matrix methods to limit the computational power needed, and on the viability of simplifying approximations to enhance such reductions. Invariably, L_x must be smaller than in the Monte Carlo runs. In order to achieve the benefits of a sparse matrix, we restrict our transfer-matrix studies to the RTSK model. We note that the numerical transfer-matrix technique was applied independently in Ref. 21 to the TWD problem.

To construct the TM, we first enumerate all valid combinations of N step positions in a periodic system of width L_x, giving the set of basis states $|i\rangle$. We calculate the energy E_i within each basis state $|i\rangle$ (from step-step interactions) and the energy E_{ij} from interactions between basis states $|i\rangle$ and $|j\rangle$ (from kinks). Then we can construct for each advance in $\hat{y}$ a symmetric TM **T** from the various Boltzmann weights:

$$\mathbf{T}_{ij} = \exp\left[-\left(E_{ij} + \frac{E_i + E_j}{2} \right) \Big/ k_B T \right]. \tag{11}$$

We reduce the dimensionality of **T** by using the invariance of the Hamiltonian under reflection, translation, and cyclic permutation of the steps.

As $L_y \to \infty$, Z^{1/L_y}, where Z is the partition function, approaches the (nondegenerate) largest eigenvalue of **T**, Λ_0.[36–39] The corresponding eigenvector, $|\Lambda_0\rangle$,[40] can be used to find the TWD by (cf. also Ref. 21)

$$P(s) = \sum_i |\langle i|\Lambda_0\rangle|^2 P_i(s), \tag{12}$$

where $P_i(s)$ is the distribution of terrace widths within basis state $|i\rangle$. While carrying out the computation expressed in Eq. (12), it is simple to find the second moment of $P(s)$ and thence to estimate $\tilde{A}$, e.g., using Eq. (9). Alternatively, one can fit a Wigner distribution to the computed (discrete) $P(s)$. While non-negligible, the differences between these two approaches is rather insignificant, comparable to the effects of temperature discussed below.

Obtaining quantitative information about the thermodynamic limit from TM analysis requires an extrapolation to $N \to \infty$ from small values of N. For thermodynamic quantities derived from the one or two largest eigenvalues of the TM, there is extensive literature on finite-size scaling,[41] although even in these cases many subtleties can confound the extrapolation. Here we must extract σ^2 from the information contained in the eigenvector $|\Lambda_0\rangle$. We have carried out power-law and exponential extrapolations; in either case, we have three equations (from the data for $n=3$, 4, and 5) in three unknowns (constant term, prefactor, and power exponent or exponential prefactor). The results are shown in Fig. 1. According to the discussion in Ref. 21, the power-law procedure is more appropriate. The exponential extrapolation can then be used as a rough gauge of the uncertainty.

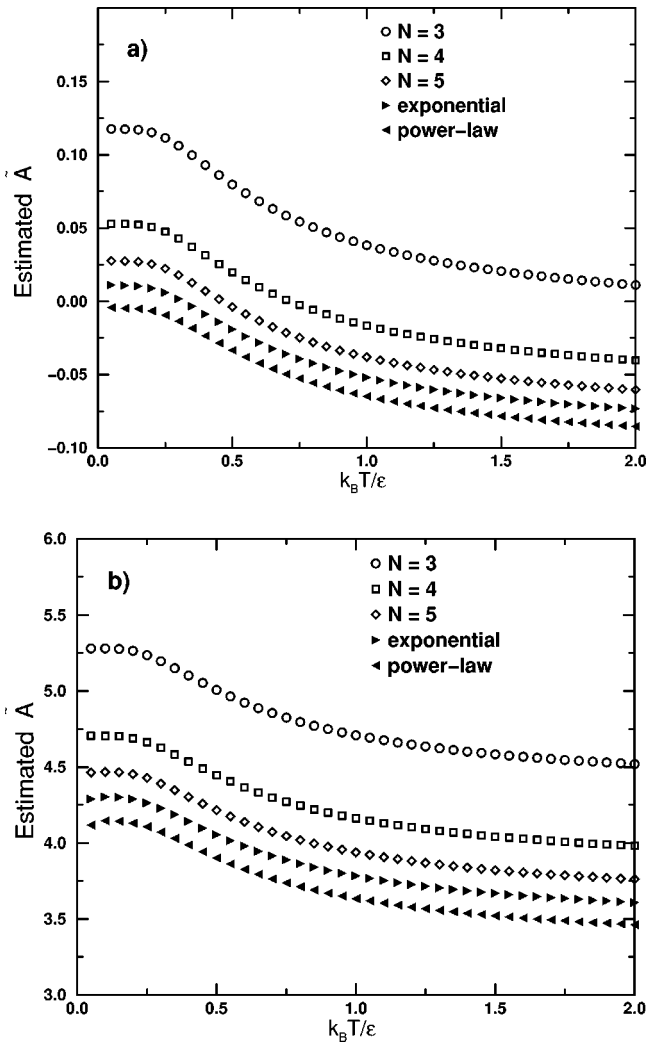

FIG. 1. The temperature dependence of the RTSK model for (a) noninteracting steps ($\tilde{A}=0$) and (b) a moderate nearest-step interaction ($\tilde{A}=4$), for $\langle \ell \rangle = 5$. The *estimated* value of $\tilde{A}$ comes from a least-squares fit of the Wigner distribution to the TWD given by the numerical TM method. In addition to TM results for three, four, and five steps, extrapolations based on assumed power-law and exponential convergence are shown.

IV. RESULTS AND DISCUSSION

A. Main result: Overall superiority of CGWD

The main result of this paper is displayed in Fig. 2(a). The various predictions of the variance are plotted vs $\tilde{A}$. A logarithmic scale is used for the horizontal axis so as not to give undue visual emphasis to larger values of $\tilde{A}$ nor to blur the

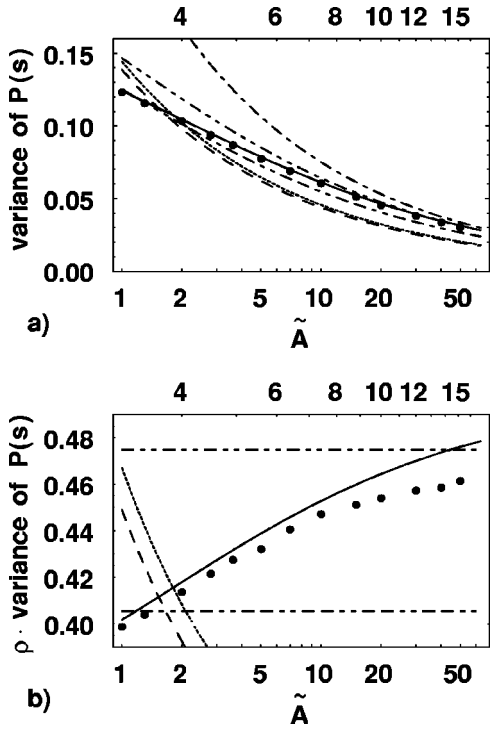

FIG. 2. (a) Plot of the variance σ^2 as a function of $\tilde{A}$ on a logarithmic scale, plotted for the CGWD ["Wigner distribution" or *W*] (light solid curve) and for the modified Grenoble or EA (short-long dashed for NN step interactions only; short-short-long dashed for all steps interacting), Saclay or R (short-short-long-long dashed curve), and Gruber-Mullins (short dashed for NN step interactions only; long dashed for all steps interacting) Gaussian approximations. Monte Carlo data are shown as ●'s, with statistical errors less than the size of the symbols. (b) Replotting with the variance multiplied by ϱ to highlight differences between the various approximations, using the same coding as in the upper panel. The ordinates of the horizontal lines are twice the corresponding values of K_X in Table I.

region of rapid variation for small (but nonvanishing) $\tilde{A}$, for which an exact calibration point exists. The Wigner result is essentially given by Eq. (9). The physical values of $\tilde{A}$ range from near 0 up to the mid teens. More than ordinary A/ℓ^2 elastic repulsions are presumably involved in the rare cases in which larger values are observed. There are relatively few reports of small but nonzero values of $\tilde{A}$. We suspect that one reason is that any of the Gaussian approximations manifestly fail in this regime, so that before the recognition of the utility of the Wigner distribution, one could not deal quantitatively with small $\tilde{A}$.[42]

To heighten the contrast, we replot in Fig. 2(b) the curves of Fig. 2(a) using as the ordinate the variance multiplied by ϱ (i.e., $2\tilde{A}_{\text{eff}}^{1/2}$); with such rescaling, the Saclay and the modified Grenoble predictions appear as horizontal lines. Clearly in this plot the numerical data increases systematically with $\tilde{A}$, much like the CGWD.

As shown in Fig. 3, we find similar results using the transfer-matrix method to study the RTSK model for $\langle \ell \rangle = 5$. While three decades of $\tilde{A}$ are displayed, the middle one

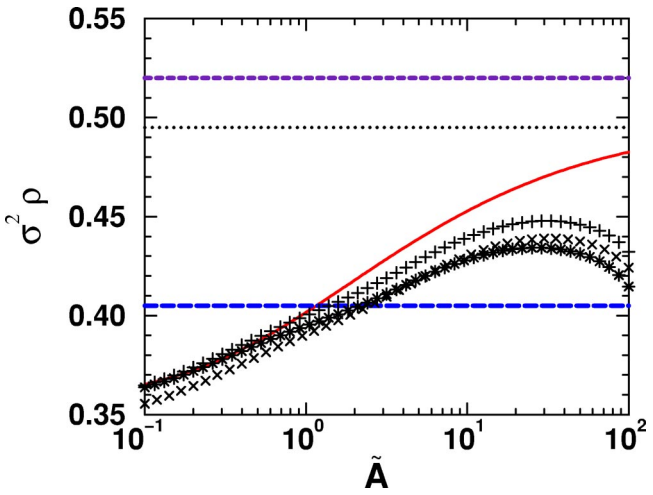

FIG. 3. Plot of the product $\sigma^2\varrho$ vs $\tilde{A}$ on a logarithmic scale, as in Fig. 2(b), also at $k_BT/\epsilon = 1/2$. The symbols are obtained from TM computations for the RTSK model with $\langle \ell \rangle = 5$. The ×'s are for four steps ($N=4$) while the +'s are for $N=5$, with nearest-neighbor (NN) step-step interactions. The *'s include next nearest in addition to NN interactions. The thin solid curve is the Wigner distribution. The thick solid and dashed horizontal lines give the EA(NN) and R approximations.

is of greatest interest. The points for $N=5$ are consistently slightly higher than those for $N=4$ and in close proximity to the CGWD curve, much like the Monte Carlo data in Fig. 2(b). The convergence appears rapid, seemingly to values modestly above the CGWD, as might be expected since only nearest-neighbor repulsions are considered. (See below for the remaining curve.)

B. Convergence in lattice width

For selected cases, we performed extensive tests of the convergence of our estimate of the variance with increasing lattice width and number of steps, as described at the end of Sec. III B. Figure 4 summarizes our findings for the special

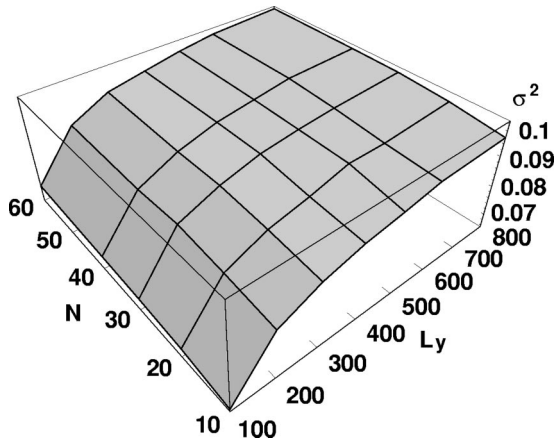

FIG. 4. Finite-size dependence of the variance for the exactly solvable case $\tilde{A}=2$, with $k_BT/\epsilon = 1/2$ and $\langle \ell \rangle = 10$. Each intersection of the grid lines on the graphed surface represents a selected pair of N and L_y. Error bars are of order 10^{-3}, too small to depict.

TABLE III. Finite-size extrapolation of variance data plotted in Fig. 4 for $\tilde{A}=2$ and $\langle \ell \rangle = 10$. [For reference, $\sigma_W^2 = 0.1045$ and $\sigma_{\text{exact}}^2 = 0.105$ (Ref. 27)]. The top part shows fits to the L_y dependence, for $L_y = 200$, 300, 400, 500, 600, 800, assuming $\sigma^2 = \sigma_\infty^2 - B/L_y^\alpha$ for various N. In the bottom part, the results of similar fits for fixed L_y and varying N are listed. The error bars are the "asymptotic" standard error.

N	σ_∞^2	B	α
10	0.104 ± 0.004	1.03	0.72 ± 0.18
20	0.106 ± 0.005	1.11	0.74 ± 0.29
30	0.104 ± 0.001	2.30	0.91 ± 0.01
40	0.106 ± 0.002	2.46	0.90 ± 0.18
60	0.106 ± 0.002	1.82	0.85 ± 0.17
L_y	σ_∞^2	B'	α'
500	0.0974 ± 0.0002	0.191	1.61 ± 0.24
600	0.124 ± 0.183	0.036	0.09 ± 0.60
800	0.1015 ± 0.0009	0.033	0.76 ± 0.22

case of $\tilde{A}=2$, for which we know the variance exactly. In Table III are the results of the associated extrapolations in L_y for fixed N and in N for fixed L_y. The extrapolations in L_y, using $\sigma^2 = \sigma_\infty^2 - B/L_y^\alpha$, are reasonably well behaved. The value of α ranges between 0.7 and 0.9, compared with the expectation from Le Goff *et al.*[21] that it be 1. Likewise, the extrapolated value is in decent agreement with the variance calculated for the exact solution for an infinite number of continuum steps. In contrast, the extrapolations in N are more variable, with sizable variations in α and in the extrapolated values. This thwarted our attempt to extrapolate in terms of a single length variable $(L_y^2 + cN^2)^{1/2}$, where c is an adjustable parameter.

The values reported in the preceding section, in particular in Fig. 2(a), are based on extrapolations to infinite L_y for several different values of L_y, assuming that $\sigma_\infty^2 - \sigma_{L_y}^2 = B/L_y^\alpha$. For example, for $\tilde{A}=2$ and $N=10$, we found in this nonlinear least-square fit that $\alpha = 0.91\ldots$ and $B = 2.3\ldots$. The resulting value of σ_∞^2 was within 0.001 of the value predicted by Eq. (7). (If instead we assumed this value from the outset, we could do a linear fit of $\sigma_W^2 - \sigma_{L_y}^2$ vs B/L_y^α to find $\alpha = 0.90\ldots$ and $B = 2.2\ldots$.)

We note the recent interest in the finite-size dependence of interfaces in thin films, as discussed authoritatively in Ref. 43. There, however, only one interface is present, the film is three dimensional, and the primary focus is on the film thickness, somewhat analogous to $\langle \ell \rangle$.

C. Deviation from continuum at very large $\tilde{A}$

All the discussed approximations rely on a continuum, capillary-wave-type approximation of steps, so that $\tilde{A}$ is the only significant parameter in determining the TWD. When the discrete nature of a real vicinal surface is considered, e.g., in numerical simulations, temperature enters as a separate variable. This behavior was noted implicitly in an early study of noninteracting steps[3] and was a major focus of very recent work by Le Goff *et al.*[21]

1. "Toothlike" configurations

In Fig. 3, the turning down of the TM-derived variances at $\tilde{A}$ greater than about 40 signals the beginning of a striking change in the scaling behavior. (In a similar plot at $k_B T/\epsilon = 0.84$ in Ref. 25, the peak occurs at $\tilde{A} \approx 60$.) The step continuum picture begins to falter, and the low-energy excitations are no longer capillary waves. Instead, for strong step-step repulsions, discrete "toothlike" single-move excitations dominate. If just nearest steps interact, the associated excitation energy is A times

$$\Delta(\langle \ell \rangle) \equiv \frac{1}{(\langle \ell \rangle + 1)^2} + \frac{1}{(\langle \ell \rangle - 1)^2} - 2\frac{1}{\langle \ell \rangle^2} \approx \frac{6}{\langle \ell \rangle^4}. \quad (13)$$

(Similarly, the contribution from the repulsion between nth neighbors has the same form as in expression (13), but with $n\langle \ell \rangle$ replacing $\langle \ell \rangle$.) For large $\tilde{A}$ it is no longer true that the only dependence on A is through $\tilde{A}$, and the variance no longer varies roughly as ϱ^{-1} (or, equivalently, as $\tilde{A}_{\text{eff}}^{-1/2}$). Specifically, for asymptotically large repulsions we expect that

$$\sigma^2 \propto \exp[-A\Delta(\langle \ell \rangle)/k_B T] = \exp[-\tilde{A}\Delta(\langle \ell \rangle) k_B T/\tilde{\beta}]. \quad (14)$$

With TM calculations for $\langle \ell \rangle = 5$ we have verified that Eq. (14) holds to at least three-significant-figure accuracy over a decade of $\tilde{A}$. (We note that the excitation can involve multiple in-phase "teeth" without increasing the net interstep repulsion. The degree to which these multistep excitations contribute depends on $\epsilon/k_B T$, further emphasizing the breakdown of scaling just with $\tilde{A}$. Furthermore, these arguments can be extended to cases with noninteger values of $\langle \ell \rangle$, though one must be more careful in dealing with finite-size issues.)

The value of $\tilde{A}$ at which the continuum picture starts to fail naturally depends strongly on $\langle \ell \rangle$, since that provides a measure of the relative discreteness of the integer deviations of the steps. We can estimate crudely the value of $\tilde{A}$ for which discreteness alters the scaling behavior of σ^2: In the Gruber-Mullins approach, the root-mean-square displacement of a step is $\langle \ell \rangle/(48\tilde{A})^{1/4}$. The crossover then should occur when this rms distance is of order unity, i.e., $\tilde{A} \sim \langle \ell \rangle^4/48$. A TM study (at $k_B T/\epsilon = 0.84$) with $N=4$ shows that the peak in a plot of $\sigma^2 \varrho$ vs $\tilde{A}$ occurs at $\tilde{A} \approx 50$ for $\langle \ell \rangle = 6$ but at $\tilde{A} \approx 200$ for $\langle \ell \rangle = 10$. These two values are roughly consistent with this prediction, though the dependence on $\langle \ell \rangle$ seems less than quartic (but decidedly more than quadratic). Since the crossover occurs for unphysically strong $\tilde{A}$, we do not explore it in exhaustive detail. Suffice it to say that the present results are consistent with the argument in Ref. 11 that discreteness is relatively insignificant

over the physical range of $\tilde{A}$ so long as $\langle \ell \rangle$ is at least 4 (although for large $\tilde{A}$ it would be safer to have $\langle \ell \rangle \geq 5$).

To investigate further this unphysical limit, we generated configurations for very large values of $\tilde{A}$ up to 1000. Even there, the discussions typically extended over several lattice spacings in the $\hat{y}$ direction. With $N=20$, no obvious signatures of ''long-wavelength phonons'' were evident, though we did not pursue a detailed analysis of the structure factor.

2. Roughening transition to facet at large $\tilde{A}$

A major difference between discrete and continuum models is that the former admit a roughening transition (while the latter is always rough). To estimate for what value of $\tilde{A}$ this transition should occur, we now seek the connection between the Calogero-Sutherland Hamiltonian used in our work and the exactly solvable VGL Hamiltonian underlying the work of the Saclay group. In the VGL model[18] the step-step repulsion vanishes for $\ell \geq \langle \ell \rangle$, takes the value w_n for $\ell = \langle \ell \rangle - 1$, and is infinite for smaller separations.

A straightforward way to make the correspondence between the two models is to adopt the Gruber-Mullins viewpoint, expand the potential due to neighboring steps at $+\langle \ell \rangle$ and $-\langle \ell \rangle$, and use Eq. (13) to show $w_n \approx A\Delta(\langle \ell \rangle) \approx 6A/\langle \ell \rangle^4$. We likewise equate the VGL discrete-Gaussian parameter W_0 and our absolute solid-on-solid kink energy ϵ and take the low-temperature limit of the stiffness (in either the TSK or RTSK model). Then the VGL condition[18] for the roughening transition $(w_n/T_R)\exp(W_0/T_R)=2$ translates to

$$\tilde{A}_R = \langle \ell \rangle^4/6. \tag{15}$$

Given the severity of the approximations invoked, the factor of 6 is unlikely to be precisely correct, but presumably $\tilde{A}_R$ should scale like $\langle \ell \rangle^4$. The implications are that for $\tilde{A} < \tilde{A}_R$ we are at temperatures above the roughening temperature (as is implicit in the idea of a vicinal surface). For $\tilde{A} > \tilde{A}_R$ a facet can develop, altering fundamentally the long-range correlations. Note that already for $\langle \ell \rangle = 3$, $\tilde{A}_R$ is about 14, above the typical maximum physical $\tilde{A}$; for $\langle \ell \rangle = 4$, $\tilde{A}_R$ is well above the physical range of $\tilde{A}$.

In this regard, it is worth recalling that in Ref. 11 some of us showed that $\langle \ell \rangle = 4$ is the threshold for the difference between continuum and discrete to become negligible. Thus, experimental vicinal surfaces should never be faceted in the regime in which it is sensible to apply the analysis in this paper.

D. Temperature dependence of the TWD

In Fig. 1 we also see a residual temperature dependence in the RTSK model. As for the TSK model, the stiffness $\tilde{\beta}_{\mathrm{RTSK}}$ is known for the RTSK model with an isolated step; using $\tilde{\beta}_{\mathrm{RTSK}}$ [and Eq. (1)], we vary A to maintain a constant $\tilde{A}$ while varying the temperature. As discussed in Sec. III C, the TWD is derived from the numerical TM for each temperature and then fitted to the CGWD $P_{\varrho}(s)$ of Eq. (5).

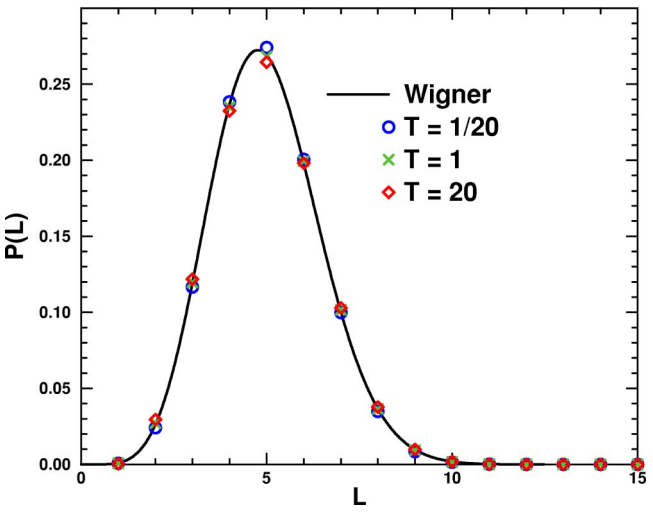

FIG. 5. The temperature dependence of the RTSK model for $N=5$ interacting steps ($\tilde{A}=4$). The TWD's for $T[(\epsilon/k_{\mathrm{B}})]=0.05$, 1, and 20 differ very little from each other and from the Wigner distribution.

The resulting estimates of $\tilde{A}$ all have the following generic features: they are approximately temperature independent at low temperatures and decrease approximately monotonically (eventually to an asymptotic value at high temperatures when the assumptions underlying the RTSK model become decidedly unphysical). The details depend on the number N of steps involved in the TM calculation; the smaller the N, the more the finite-size estimate exceeds the expected input value of $\tilde{A}$. Included extrapolations $N \to \infty$, based on assumed power-law and exponential convergences, give a better estimate of the input. The absolute size of the error in the estimate seems to increase with increasing $\tilde{A}$.

The deviation of the estimated value of $\tilde{A}$ from the known, input value does not require a large difference in the corresponding TWD's, as shown in Fig. 5 for $\tilde{A}=4$. Even for these extreme temperatures,[44] the TWD is relatively insensitive.

To check further the robustness of the idea that the TWD depends only on $\tilde{A}$ and not its separate components, we did trial sets of runs at $k_B T/\epsilon = 0.3$, 0.5, and 0.8 for $\tilde{A} = 0$ and 2 (i.e., $\varrho = 2$ and 4, respectively). For both $\tilde{A}$'s, the value of the deduced ϱ at $k_B T/\epsilon = 1/2$ was largest and closest to the input value. It is remarkable that there the variation with temperature is not monotonic. Experimentalists know[45] that simple equilibrium behavior can be observed only over a narrow thermal window: at too cool conditions, slow diffusion precludes establishment of equilibrium conditions, whereas at too warm temperatures, the excitation spectrum becomes much more complicated than in the elementary models invoked by theory. Evidently there is a comparable window of sorts in numerical simulations. At lower temperatures, in addition to the equilibration difficulty, the characteristic length along $\hat{y}$—which is proportional to $\langle \ell \rangle^2 \tilde{\beta}$[29]—grows dramatically (e.g., by a factor of nearly 5 in cooling from 0.5 to 0.3), creating finite-size-induced deviations as depicted in Fig. 4.

At high temperatures, the excitation spectrum can grow more complicated than assumed in the small-deviation expansions implicit in standard analysis.

E. Nearest-neighbor vs long-range repulsions

In the preceding section we have mentioned several times the issue of whether all steps or just neighboring ones suffer a repulsive interaction proportional to the inverse square of their separation. As noted, the Wigner distribution and the Saclay (R) approach intrinsically take all steps to interact, whereas the Gruber-Mullins (GM) and Grenoble (E0 and EA) approaches can be done either way. For the numerical methods, it is much easier to assume just nearest-neighbor interactions. Obviously, for $A=0$ there can be no difference. It is widely taken for granted that the difference is not significant. For example, Ihle *et al.*[15] point out that the width of the TWD, which is what is typically measured and is essentially the root of the variance, changes at most by a few percent, so by less than typical experimental errors. On the other hand, since $\tilde{A}$ depends essentially on the square of the variance (i.e., on the 4th power of the width), the variation can seem more noteworthy. This comment provides a reminder of the difficulty of pinpointing the magnitude of the step repulsion from analysis of TWD's.

In this section we address briefly the effects of including more distant steps in our numerical simulations. Including elastic repulsions with all steps should sharpen the TWD, thereby decreasing the variance, for given $\tilde{A}$. As noted, the difference should vanish at $A=0$ and increase monotonically to nearly 5% when the Grenoble picture applies (cf. Table I). In Fig. 2(b) there is another horizontal line at 0.52, above the plotting window. It is thus remarkable and curious that there is such fine agreement between the Monte Carlo results and CGWD curve when the above arguments would lead to expectations that the Monte Carlo values of the variance for given $\tilde{A}$ would be somewhat larger than those of CGWD, so that if one deduced a value of $\tilde{A}$ from the resulting variance, the result would be somewhat smaller than initially.

In Fig. 2(b) we also include some Monte Carlo results for long-range repulsions, extending out to $N/2$ steps. We find as expected that the difference is small for small $\tilde{A}$ but grows by $\tilde{A}\approx 7$ to a difference comparable to that between the NN and all modified Grenoble approximations. We have not scrutinized the behavior at larger $\tilde{A}$, which involve long equilibration times.

Given the limited size of the systems used in our TM studies, the best we can do is allow second neighbor as well as nearest-neighbor repulsions. In Fig. 3 we see similar behavior, but with an asymptotic difference somewhat less than that between the two modified Grenoble cases and developing more gradually (with increasing $\tilde{A}$) than in the Monte Carlo simulations.

V. CONCLUSION

The numerical studies presented here (esp. in Sec. IV A) show that the CGWD of Eq. (2) is not just an excellent interpolation between the established points at $\tilde{A}=0$ and $\tilde{A}=2$, but also a fine extrapolation expression over the range of physical $\tilde{A}$ and beyond; moreover, it might be viewed as an *interpolation* between $\tilde{A}=2$ and the result of the physically compelling Grenoble viewpoint near $\tilde{A}=\infty$. While the shape of the TWD does approach a Gaussian in the physical regime of moderately strong $\tilde{A}$, the CGWD [via Eq. (9)] provides arguably the best way to extract $\tilde{A}$ from the variance of the TWD and certainly the least ambiguous. Of the Gaussian methods the Saclay (R) scheme is better for moderate $\tilde{A}$ whereas the Grenoble [EA(all)] scheme is better for stronger $\tilde{A}$. The continuum description is a good approximation for terraces at least four atomic spacings wide.

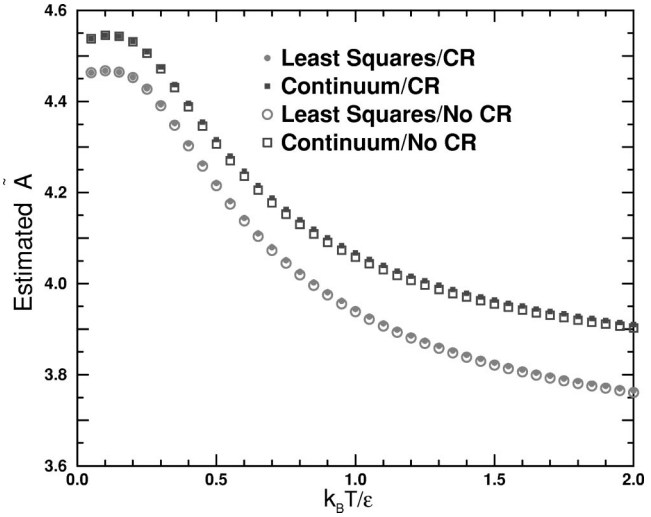

FIG. 6. Plots of estimated $\tilde{A}$ vs k_BT/ϵ for $\langle \ell \rangle = 5$ and $N=5$ with an "input" value of $\tilde{A}=4$. The four sets of data are for corner repulsion included or omitted and for the two different extrapolation methods to deduce $\tilde{A}$ from TM computations.

ACKNOWLEDGMENTS

Work was supported by the NSF-MRSEC at University of Maryland (NSF Grant No. DMR 00-80008) and benefited from ongoing interactions with E.D. Williams, as well as with H. van Beijeren and O. Pierre-Louis. S.D.C. participated in an NSF-MRSEC-sponsored REU program at the outset of this project. Some of the work of S.D.C. and T.L.E. has been partially supported by NSF Grant No. EEC-0085604.

APPENDIX: CORNER REPULSION

In preliminary tests of the free fermion case, $\tilde{A}=0$, we found in fitting with Eq. (5) the TWD produced with the conventional TSK model that the best estimate of ϱ was 1.81, whereas with a TSK model with corner exclusions (TSK-CE model, described in Sec. III A) the best-fit value was the expected $\varrho=2$. Thus, the corner effects, which effectively increase the step repulsion when entropic effects

are important, seem to compensate the effects of discreteness.

To check the systematic effects of corner repulsion, we performed transfer-matrix calculations for $\tilde{A}=4$, the results of which are shown in Fig. 6. Working with $\langle \ell \rangle = 5$ and $N=5$, we examined the temperature dependence of the deduced value of $\tilde{A}$. The open circles are the same data as the diamonds in Fig. 1(b). As expected, the effect of corner repulsions increases with increasing temperature, as entropic effects gain relative importance, but evidently their effect is negligible even at this rather modest value of $\tilde{A}$. We also display the deduced values of $\tilde{A}$ with and without corner repulsion when the variance is computed directly from the weighted projection of the ground-state eigenvector, as in Eq. (12). Again, we find that corner repulsion makes a negligible change, but find that how the variance is extracted does make a difference of a few percent in deduced value of $\tilde{A}$. For this set of data, fitting $P(s)$ gives an estimate of $\tilde{A}$ closer to the input value, but it is not certain whether the same would be true if we had used extrapolated data. Since the difference is not substantial on the scale of experimental errors, we did not investigate this matter further.

*Present address: Department of Physics, Columbia University, New York, NY 10027.

†Permanent address: Department of Physics, Texas A&M University–Commerce, Commerce, TX 75429.

‡Corresponding author. Email address: einstein@umd.edu; URL:http://www2.physics.umd.edu/~einstein

[1] This process is often facilitated by viewing the $\hat{y}$ direction as "timelike" in a mapping of the two-dimensional classical model to a $(1+1)$-dimensional quantum problem.

[2] E.E. Gruber and W.W. Mullins, J. Phys. Chem. Solids **28**, 875 (1967).

[3] N.C. Bartelt, T.L. Einstein, and E.D. Williams, Surf. Sci. **240**, L591 (1990).

[4] H.-C. Jeong and E.D. Williams, Surf. Sci. Rep. **34**, 171 (1999).

[5] M. Giesen and T.L. Einstein, Surf. Sci. **449**, 191 (2000).

[6] S. van Dijken, H.J.W. Zandvliet, and B. Poelsema, Phys. Rev. B **55**, 7864 (1997).

[7] M.L. Mehta, *Random Matrices*, 2nd ed. (Academic, New York, 1991).

[8] T. Guhr, A. Müller-Groeling, and H.A. Weidenmüller, Phys. Rep. **299**, 189 (1998).

[9] T.L. Einstein, H.L. Richards, S.D. Cohen, and O. Pierre-Louis, Surf. Sci. **493**, 460 (2001).

[10] T.L. Einstein and O. Pierre-Louis, Surf. Sci. **424**, L299 (1999).

[11] H.L. Richards, S.D. Cohen, T.L. Einstein, and M. Giesen, Surf. Sci. **453**, 59 (2000).

[12] T.L. Einstein, Ann. Henri Poincaré **4** (Suppl. 2), S811 (2003).

[13] B. Joós, T.L. Einstein, and N.C. Bartelt, Phys. Rev. B **43**, 8153 (1991).

[14] O. Pierre-Louis and C. Misbah, Phys. Rev. B **58**, 2259 (1998).

[15] T. Ihle, C. Misbah, and O. Pierre-Louis, Phys. Rev. B **58**, 2289 (1998).

[16] L. Masson, L. Barbier, J. Cousty, and B. Salanon, Surf. Sci. **317**, L1115 (1994).

[17] C. Jayaprakash, C. Rottman, and W.F. Saam, Phys. Rev. B **30**, 6549 (1984).

[18] J. Villain, D.R. Grempel, and J. Lapujoulade, J. Phys. F: Met. Phys. **15**, 809 (1985).

[19] H.-C. Jeong and J.D. Weeks, Surf. Sci. **432**, 101 (1999), and references therein.

[20] L. Barbier, L. Masson, J. Cousty, and B. Salanon, Surf. Sci. **345**, 197 (1996).

[21] E. Le Goff, L. Barbier, L. Masson, and B. Salanon, Surf. Sci. **432**, 139 (1999); E. Le Goff, L. Barbier, and B. Salanon, *ibid.* **531**, 337 (2003).

[22] F. Calogero, J. Math. Phys. **10**, 2191 (1969); **10**, 2197 (1969).

[23] B. Sutherland, J. Math. Phys. **12**, 246 (1971); Phys. Rev. A **4**, 2019 (1971).

[24] F. Haake, *Quantum Signatures of Chaos*, 2nd ed. (Springer, Berlin, 1991).

[25] H.L. Richards and T.L. Einstein, cond-mat/0008089, Phys. Rev. E (to be published); Amber N. Benson, H.L. Richards, and T.L. Einstein, cond-mat/0307626, Phys. Rev. B (to be published).

[26] See, e.g., http://www.itl.nist.gov/div898/handbook/eda/section3/eda366b.htm/

[27] Mehta (Ref. 7) lists the value of σ^2 as 0.105 while σ^2_{Wigner} is 0.104 466 . . . . To clarify this matter we sought higher precision by using the power series for $P(s)$ at small s (cf. Ref. 22) and the long-known asymptotic expression [F.J. Dyson, Commun. Math. Phys. **47**, 171 (1976)] for large s, choosing the crossover value of s between 1.51 and 1.52 so that $P(s)$ had optimally close to unit normalization and mean. Alternatively, we followed B. Dietz and F. Haake, Z. Phys. B: Condens. Matter **80**, 153 (1990), using a Padé approximant to bridge to the asymptotic limit. Both sets of calculations show $\sigma^2 \doteq 0.1041$. In the process, we also calculated higher moments of $P(s)$ [Hailu Gebremariam, T.L. Einstein, and B. Dietz (unpublished)].

[28] At large $\tilde{A}$ discrete effects will produce breakdown of the CGWD estimate. Physical behavior in this regime depends sensitively on boundary conditions and system size. If there is any pinning of the steps, then the low-energy excitation will be "tooth"-like kink-antikink pairs on straight, uniformly spaced steps. If there is no pinning, then there can be long-wavelength (in the $\hat{y}$ direction) in-phase (in the $\hat{x}$ direction) excitations corresponding to long-wavelength phonons.

[29] N.C. Bartelt, T.L. Einstein, and E.D. Williams, Surf. Sci. **276**, 308 (1992).

[30] M.E.J. Newman and G.T. Barkema, *Monte Carlo Methods in Statistical Physics* (Clarendon, Oxford, 1999).

[31] N. Metropolis, A.W. Rosenbluth, M.N. Rosenbluth, A.H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953).

[32] A.B. Bortz, M.H. Kalos, and J.L. Lebowitz, J. Comput. Phys. **17**, 10 (1975).

[33] M.A. Novotny, Comput. Phys. **9**, 46 (1995).

[34] W.H. Press, B.P. Flannery, S.A. Teukolsky, and W.T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing*, 2nd ed. (Cambridge University Press, Cambridge, 1993); also see http://www.nr.com/

[35] C. Domb, Adv. Phys. **9**, 149 (1960); L.K. Runnels and L.L.

Combs, J. Chem. Phys. **45**, 2482 (1966); F.H. Ree and D.A. Chesnut, *ibid.* **45**, 3983 (1966).

[36] H.A. Kramers and G.H. Wannier, Phys. Rev. **60**, 252 (1941); **60**, 263 (1941).

[37] W.J. Camp and M.E. Fisher, Phys. Rev. B **6**, 946 (1972).

[38] M.A. Novotny, J. Math. Phys. **29**, 2280 (1988).

[39] M.P. Nightingale, in *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, Singapore, 1990), p. 287.

[40] We find both Λ_0 and $|\Lambda_0\rangle$ by diagonalizing **T** using the subroutine DNLASO. Source code and documentation for dnlaso.f are available at http://www.netlib.org/laso/. A Lanczos algorithm matrix diagonalization routine [O. Axelsson, *Iterative Solution Methods* (Cambridge University, New York, 1996)], it is most efficient for sparse matrices and can selectively calculate chosen eigenvalues and eigenvectors.

[41] For an excellent set of review papers, see *Finite Size Scaling and Numerical Simulation of Statistical Systems*, edited by V. Privman (World Scientific, Singapore, 1990).

[42] K. Swamy, E. Bartel, and I. Vilfan, Surf. Sci. **425**, L369 (1999).

[43] K. Binder and M. Müller, Int. J. Mod. Phys. C **11**, 1093 (2000).

[44] At $T=20$ the diffusivity ($k_B T/\tilde{\beta}$) in the RTSK model is nearly saturated, making the stiffness unphysically large and so greatly underestimating the variance of the TWD compared to better models such as TSK.

[45] M. Giesen (private communication).