Statistics of rare events in disordered conductors

Igor E. Smolyarenko
Lyman Laboratory of Physics, Harvard University, Cambridge, Massachusetts 02138
and NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540

Boris L. Altshuler
NEC Research Institute, 4 Independence Way, Princeton, New Jersey 08540
(Received 1 August 1996)

Asymptotic behavior of the distribution functions of eigenstate intensities and current relaxation times in disordered conductors is studied in the weak disorder limit by means of an optimal fluctuation method. It is argued that this method is more appropriate for the study of rare events in three-dimensional conductors than the approaches based on nonlinear σ models because it is capable of correctly handling fluctuations of the random potential with large amplitude as well as the short-scale structure of the corresponding solutions of the Schrödinger equation. It also helps to clarify the physical picture of such events in one and two dimensions. For two- and three-dimensional conductors, the asymptotics of the distribution functions obtained by this method differ, in some cases significantly, from previously established results. [S0163-1829(97)00316-0]

I. INTRODUCTION

It has been well known for at least a decade[1,2] that a complete description of the Anderson localization transition in disordered mesoscopic conductors must necessarily be formulated in terms of the full distribution functions of conductance g and/or other quantities characterizing the sample. As a consequence, even well into the metallic regime it should be possible to observe the onset of localization by studying deviations of asymptotic tails of the distribution functions from their behavior in the infinite conductance limit. Such a study was performed analytically in Refs. 3 and 4 within the framework of the nonlinear σ model[5] based on the replica technique.[6] It was demonstrated in Ref. 4 that in two spatial dimensions the tails of the distributions of such quantities as conductance, local density of states, current relaxation times, etc., are all described by rather similar logarithmically normal asymptotes—whereas Gaussian distributions would be expected for $g\to\infty$. Recently, Muzykantskii and Khmelnitskii[7] proposed a more straightforward and elegant method to obtain these asymptotes which utilizes the optimal fluctuation method in conjunction with the supersymmetric version of the nonlinear σ model.[8] In a series of subsequent publications the method was employed to study the statistics of eigenfunction amplitudes in weakly localized two-dimensional conductors[9] and the distribution of the local density of states.[10] An extension of the method to quasi-one-dimensional and three-dimensional systems was achieved in Refs. 10–13.

The common feature of all the results obtained so far with the use of various σ models is that the large-z asymptotic behavior of a distribution function $\mathcal{P}(z)$ has the following form for two-dimensional $(d=2)$ and three-dimensional $(d=3)$ systems:

$$\mathcal{P}(z)\sim\exp(-C_d\ln^d z). \tag{1}$$

In open samples, the argument z of the distribution function may have the meaning, e.g., of the current relaxation time τ_r, or the normalized local density of states at the Fermi energy $\rho(E_F,\mathbf{r})/\nu_d$. In closed samples it can be associated with the local wave-function intensity $t=V|\psi_E(\mathbf{r})|^2$ of an energy E eigenstate. ν_d is the average d-dimensional density of states and V is the volume of the sample. One exception is the $e^{-\ln^2}$ law obtained in Ref. 12 for the distribution of wave-function amplitudes in *three-dimensional* samples. We will discuss a possible origin of this difference in Sec. IV.

The coefficients C_d in general depend on the strength of disorder and, for $d\leqslant 2$, on the sample size L. In the two-dimensional case the calculations based on σ models give $C_2=\beta\pi^2\nu_2 D/2\ \ln(L/l)$,[4,9,10,12,13] where p_F is the Fermi momentum, $D=\frac{1}{2}lv_F$ is the electron diffusion constant, l is the mean free path (which is assumed to be much larger than the electron wavelength p_F^{-1}), and v_F is the Fermi velocity. β is a numerical coefficient which, depending on the symmetry of the ensemble of random potentials, takes the following values: $\beta=1$ in time-reversal-invariant systems (the orthogonal symmetry class), and $\beta=2$ when symmetry with respect to time reversal is completely broken (the unitary symmetry class). In Ref. 13 only the unitary ensemble was considered. We will argue below that there exist corrections to this value of C_2 which may become essential in sufficiently small (or sufficiently clean) systems when the inequality $p_FL\gg(p_Fl)^2$ is violated. Such a situation may be easily realized, for example, in the experiments of Ref. 14.

In the one-dimensional case all eigenstates are localized, and the distribution of wave-function amplitudes in a closed sample has a simple exponential form $\exp(-4lt/V)$.[15] Current relaxation time τ_r in an open sample, on the other hand, is characterized by a much broader logarithmically normal distribution

$$\mathcal{P}(\tau_r)\sim\exp(-C_1\ln^2\tau_r\Delta), \tag{2}$$

where $C_1 = l/2L$ for a sample of length $2L$, and $\Delta = 1/\nu_d V$ is the mean level spacing.[13,16]

In Refs. 10 and 12 saddle-point solutions of the supersymmetric nonlinear σ models were obtained for the three-dimensional case. However, conventional nonlinear σ models used in Refs. 4, 7, 9, 10, and 12 are low-''energy'' effective field theories in which the role of energy is played by the diffusion operator $D\nabla^2$. As such they are only applicable to describing phenomena that can be characterized as diffusive—the mean free path l is the smallest relevant length scale in these theories. Formally this is expressed as a requirement that the scale of spatial variations of the fundamental variables of the theory—the Q matrices—must be much larger than l. The optimal fluctuations of the Q matrices computed in Refs. 10 and 12 were found to vary rapidly over distances $\sim l$ in three-dimensional systems, which made it impossible to obtain rigorous results. The coefficient C_3 was estimated in Ref. 10 to be of the order of $(p_F l)^2$.

In an attempt to overcome the limitations of diffusive σ models and account exactly for the spatial variation of the optimal fluctuation in three-dimensional systems on the scale of the mean free path, a generalized version of the model (ballistic nonlinear σ model) was introduced in Ref. 17 and used in Ref. 13 to compute the distribution of current relaxation times. The functional form of Eq. (1) was reproduced in that calculation and a numerical value $\pi/9\sqrt{3}(p_F l)^2$ for C_3 was found. We will argue based on the calculations presented below that all the results for the three-dimensional systems quoted above underestimate the probability $\mathcal{P}$ of occurrence of rare events, with a single exception: in the limit $\ln t > p_F l$ (which is admissible in the three-dimensional case), Ref. 12 overestimates this probability.

In this paper we propose an alternative method for the investigation of the distribution function asymptotics. Instead of integrating out the disorder degrees of freedom and *then* looking for a saddle point of the resulting effective field theory (the nonlinear σ model), we suggest that large-z behavior of the distribution functions is governed by a saddle point of the original theory based on the Schrödinger equation for a particle in a random potential. The distribution functions are expressed as functional integrals over both electronic degrees of freedom and realizations of the random potential. The saddle points of these functional integrals correspond to optimal fluctuations, i.e., the highest-probability configurations of disorder that let the electronic eigenstates at a given energy E have the desired property—for example, an anomalously large intensity $V|\psi_E(\mathbf{r})|^2$ at some point $\mathbf{r}$. In our view, this approach possesses the advantage of being applicable to systems of arbitrary dimensionality, as well as being able to take into account nonsemiclassical effects. It is also conceptually simpler, and makes the physical origin of the results much more transparent. In many respects our approach is similar to the ideas utilized in Refs. 18 and 19 to study the tails of the density of states in doped semiconductors. We were able to reproduce the general $\ln\mathcal{P} \sim -\ln^d z$ behavior of the distribution functions for $d=2$ and 3 as well as the log-normal form of the distribution of the current relaxation times for $d=1$. In the two- and three-dimensional cases we obtain values for the coefficients C_d [see Eq. (5)] that differ from their previously published estimations.

Our calculations are performed in the limit of the largest possible values of the eigenstate intensities t, constrained only by the condition $t < V$. To determine the low-t limit of applicability of the results and investigate their crossover to the universal Porter-Thomas distribution, the study of the fluctuations around the saddle point is necessary. Such a study is in progress, and will be reported elsewhere.

The rest of the paper is organized as follows. In Sec. II we briefly describe the method and present the main results. A detailed derivation of the basic equations and the analysis of the saddle point solutions is deferred until Sec. III. Section IV is devoted to the discussion of our results and their relation to those obtained previously. A short summary and a list of open questions can be found in Sec. V.

II. ASYMPTOTES OF THE DISTRIBUTION FUNCTIONS

A typical wave function in a metallic sample is spread more or less uniformly throughout the sample volume $V \sim L^d$, so that its amplitude does not differ much from the average value of $1/\sqrt{V}$ anywhere in the sample. In the metallic (or weakly localized) regime such states account for the bulk of the distribution of the local density of states $\rho(\mathbf{r})$ or of the current relaxation times τ_r in open samples. The chances of observing anomalously large values of these quantities are related to the probability of finding an ''anomalously localized state.'' Such a state would be characterized by an amplitude reaching a value much larger than the average at some point $\mathbf{r}$ inside the sample. In what follows we will concentrate on the distributions of wave-function amplitudes in closed samples and current relaxation times in open ones. Other distributions, e.g., that of the local density of states, can be more or less straightforwardly derived within the framework of the same formalism. It is important to note that in the metallic regime the same type of states is responsible for optimally achieving large values of current relaxation times and wave-function amplitudes. In the Anderson insulator regime, typified by purely one-dimensional samples, this is no longer true. While the states responsible for long current relaxation times do have a peak in the bulk of the sample, they are not the same states that dominate the distribution function of eigenstate intensities (see below, Sec. III A).

The problem can be formulated in the following way. Let us consider a spherical (in three dimensions) or a disk-shaped (in two dimensions) conductor of radius L, and compute the probability that an eigenstate ψ of energy E (which we take to be equal to the Fermi energy E_F) has an amplitude $\sqrt{t/V}$ in the center of the sample ($\mathbf{r}=0$), with $t \gg 1$. The distribution of the disorder potential is assumed to be uncorrelated Gaussian,

$$\mathcal{W}[U(\mathbf{r})] = \mathcal{N}_U \exp\left(-\frac{\pi \nu_d \tau}{2} \int U^2(\mathbf{r}) d\mathbf{r} \right), \qquad (3)$$

where $\mathcal{N}_U$ is the normalization constant and τ is the mean free time. Such a probability is then naturally expressed as

$$\mathcal{P}(t) = \langle \delta(V|\psi(0)|^2 - t) \rangle_U, \qquad (4)$$

where $\langle\ \rangle_U$ denotes the averaging with the weight $\mathcal{W}$ over all possible configurations of U, and $\psi(\mathbf{r})$ is the solution of the

Schrödinger equation with the Hamiltonian $H=H_0+U(\mathbf{r})$ and energy E. $H_0=\hat{\mathbf{p}}^2/(2m)$ is a Hamiltonian of free particles with a mass m, and $\hat{\mathbf{p}}$ is the canonical momentum operator.

The requirement that E must be an eigenvalue of H was enforced explicitly in Ref. 9 by introducing the corresponding δ function into the definition of $\mathcal{P}$. In the alternative approach proposed here—the direct optimal fluctuation method—this requirement is easier to impose, if necessary, at a later stage in the calculations through appropriate boundary conditions for the saddle point equations.[20]

Rewriting Eq. (4) as a constrained functional integral and introducing Lagrange multipliers to enforce the constraints (see Sec. III for details), one can demonstrate that the resulting ''action'' $\mathcal{A}$ possesses a saddle point, and the leading contribution to $\ln\mathcal{P}(t)$ is given by $-\mathcal{A}_0$, the value of the action at that point. With exponential accuracy the results for the distribution function of eigenstate intensities are

$$\mathcal{P}(t)\sim\exp[-\kappa(p_F l)\ln^3 t] \quad (3\mathrm{D}), \tag{5a}$$

$$\mathcal{P}(t)\sim\exp\left(-\pi^2\nu_2 D\frac{\ln^2 t}{\ln(L/r_0)}\right) \quad (2\mathrm{D}), \tag{5b}$$

where κ is a number which we estimate as 3×10^{-3}. The length scale r_0 in Eq. (5b) is of the order of the electron wavelength p_F^{-1}. In two- (2D) and three-dimensional (3D) conductors anomalously long current relaxation times τ_r in open samples are due to the states which are also characterized by anomalously large intensity t somewhere in the sample. As a result, asymptotic behavior of the distribution of τ_r has the same form as Eq. (5) with $\tau_r\Delta$ in place of t [Eqs. (40) and (58)]. In contrast, in the one-dimensional case the states responsible for these two effects have different structures leading to different probability distributions for t and τ_r (see Sec. III A below).

In both two- and three-dimensional cases, our answers have a different l dependence compared to the results obtained in Refs. 4, 9, 10, 12, and 13. In the three-dimensional case $\ln\mathcal{P}$ is proportional to the *first* power of $p_F l$, while all the results derived from the nonlinear σ model in three dimensions lead to a $(p_F l)^2$ dependence. Since $p_F l\gg1$, Eq. (5a) indicates a *substantially increased* probability of observing a rare event as compared to the previous results.[10,13]

Similarly, in the two-dimensional case the logarithm in the denominator contains L/r_0 instead of the L/l established in Ref. 4, and reproduced with minor modifications in Refs. 9, 10, 12, and 13. This difference leads only to small corrections, unless $p_F L$ becomes of the order of $(p_F l)^2$.

These asymptotes correspond to the optimal configurations of the potential $U(\mathbf{r})$, which are essentially Bragg mirrors,

$$U(\mathbf{r})\propto\frac{\sin 2p_F r}{r^{d-1}}, \tag{6}$$

as $r\to\infty$. The $r\to0$ behavior is more complicated. In one- and two-dimensional systems the $r\to0$ region does not contribute to $\ln\mathcal{P}$ in the leading in $\ln t$ order. In three dimensions the shape of the optimal fluctuation at small distances is a narrow potential well surrounded by a wider potential barrier

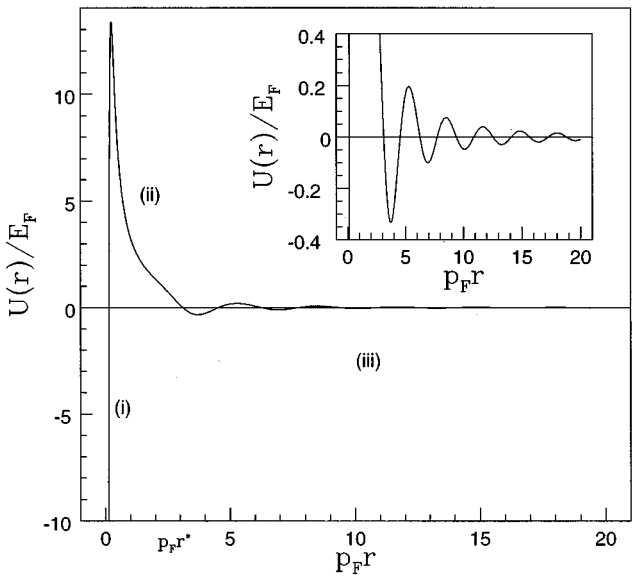

FIG. 1. Profile of an optimal configuration of the potential in three dimensions as a function of the radial coordinate.

(Fig. 1) such that the whole structure supports a narrow resonance in the s-wave channel. Anomalously large values of t are achieved by combining this local resonance with Bragg reflection at larger distances.

The role of local resonances in producing large values of t for any d in the interval $2\leqslant d<4$ can be illustrated by the following simplified version of the direct optimal fluctuation method. Let us approximate the shape of the resonance-producing potential as $U(\mathbf{r})=U_0\theta(r-a)\theta(b-r)$, where $\theta(r)$ is the step function, and U_0 and b are optimization parameters.[21] The constant a is determined from the condition that there is only one resonance in the potential well and it has a given energy E_F [$a\sim p_F^{-1}=(2mE_F)^{-1/2}$]. The wave function of the resonant state in such a potential exponentially decays in the interval $a<r<b$ (for $U_0>E_F$). Assuming that the amplitude of the wave function for $r>b$ is close to its average value $1/\sqrt{V}$, we can estimate $t\equiv V\psi^2(0)\sim\exp\{2\sqrt{2m(U_0-E_F)}(b-a)\}$. Anticipating the final result $b\sim p_F^{-1}\ln t$, we can neglect a in the exponent. We then have

$$U_0\approx E_F+\frac{1}{2m}\left(\frac{\ln t}{2b}\right)^2, \tag{7}$$

and, for $a\ll b$,

$$\mathcal{A}=\frac{\pi\nu_d\tau}{2}\int U^2(\mathbf{r})d\mathbf{r}=\frac{\pi\nu_d\tau}{2}V_d U_0^2 b^d, \tag{8}$$

where $V_d=\pi^{d/2}/\Gamma(d/2+1)$ is the volume of a d-dimensional sphere of unit radius. Substituting Eq. (7) into Eq. (8) and minimizing with respect to b, we find

$$b=p_F^{-1}\left(\frac{4-d}{4d}\right)^{1/2}\ln t, \tag{9}$$

and it then follows that $U_0=4E_F/(4-d)$. The resulting estimate for the asymptotic behavior of the distribution function reads

$$\mathcal{P}(t)\sim\exp\{-K_d(p_F l)\ln^d t\}, \tag{10a}$$

$$K_d=\frac{1}{2d}\left(\frac{1}{4\pi}\right)^{d-1}\left(\frac{d}{4-d}\right)^{(4-d)/2}V_d^2. \tag{10b}$$

In the three-dimensional case Eq. (10) differs only by a number from a more rigorous estimate quoted in Eq. (5b), suggesting that local resonances indeed play an important role in the three-dimensional case. As it should be, the coefficient $K_3\approx 0.032$ is *larger* than κ since it would be naïve to expect a box-shaped potential to be optimal.

In the two-dimensional case, large values of t are not optimally produced by local resonances: according to Eq. (5b) the coefficient in front of $\ln^2 t$ is proportional to $[\ln(L/r_0)]^{-1}$, i.e., it depends on the sample size.

The hypothesis that log-normal (in one and two dimensions) and $\exp(-C_3\ln^3 t)$ (in three dimensions) asymptotics of the distribution functions may be produced by Bragg reflection locking the states in (or out of) certain regions of the sample was put forth by Muzykantskii and Khmelnitskii in Ref. 13. They explicitly demonstrated that in a one-dimensional sample of length $2L$ an electron of energy E_F can be trapped for a time $\tau_r\gg 1/\Delta$ by a potential of the form $U(x)=[p_F\ln(\tau_r\Delta)/mL]\cos(2p_F x)$. Such a configuration appears with the probability

$$\mathcal{P}(\tau_r)\sim\exp\left(-\frac{l}{2L}\ln^2(\tau_r\Delta)\right), \tag{11}$$

which is the same as the log-normal long-time asymptotics of the exact solution for the conductance.[16] Our results are consistent with this hypothesis in higher dimensions.

The most unexpected feature of our results is the fact that a *single* configuration of the potential, i.e., a single compact region of the configurations space, seems to be responsible for the tails of the distribution functions. The alternative would be for a large number of more or less equal-weight configurations to be able to produce the required wave function amplitude $\sqrt{t/V}$, and the total probability $\mathcal{P}(t)$ would then be given by the sum of the weights of these configurations. An argument against such an alternative is the fact that the results obtained by simply singling out the most probable type of configurations can lead to higher probabilities (significantly higher in the three-dimensional case) than the σ-model results[10,13] which correspond to a sum over all configurations.[22]

III. DIRECT OPTIMAL FLUCTUATION METHOD

In order to enforce the condition that ψ in Eq. (4) is a solution of the Schrödinger equation we rewrite Eq. (4) formally as

$$\mathcal{P}(t)=\int\mathcal{D}U(\mathbf{r})\mathcal{W}[U(\mathbf{r})]\mathcal{N}_\psi[U]\int\mathcal{D}\psi(\mathbf{r})\times\prod_{\mathbf{r}'}\delta((H-E)\psi(\mathbf{r}'))\delta(V|\psi(0)|^2-t), \tag{12}$$

where $\mathcal{N}_\psi[U]$ is a U-dependent normalization constant defined by

$$\mathcal{N}_\psi[U]\int\mathcal{D}\psi(\mathbf{r})\prod_{\mathbf{r}'}\delta((H-E)\psi(\mathbf{r}'))=1. \tag{13}$$

Introducing auxiliary variables $\chi(\mathbf{r})$ and λ associated, respectively, with the first and the second δ functions in Eq. (12) enables us to represent the distribution function as unconstrained functional integrals

$$\mathcal{P}(t)=\mathcal{N}_U\int\mathcal{D}U(\mathbf{r})\mathcal{N}_\psi\int\mathcal{D}\psi(\mathbf{r})\int\mathcal{D}\left(\frac{\chi(\mathbf{r})}{2\pi}\right)\times\int\frac{d\lambda}{2\pi}e^{-\mathcal{A}[U,\psi,\chi,\lambda]}, \tag{14a}$$

$$\mathcal{A}[U,\psi,\chi,\lambda]=\frac{\pi\nu_d\tau}{2}\int d\mathbf{r}\;U^2(\mathbf{r})-i\lambda(V|\psi(0)|^2-t)-i\int d\mathbf{r}\;\chi(\mathbf{r})\left(\frac{\hat{\mathbf{p}}^2}{2m}+U(\mathbf{r})-E\right)\psi(\mathbf{r}). \tag{14b}$$

The next step is to make a saddle-point approximation in Eq. (14a), utilizing the fact that t is large. The stationarity condition for the action Eq. (14b) leads to the following set of equations:

$$\pi\nu_d\tau U(\mathbf{r})-i\chi(\mathbf{r})\psi(\mathbf{r})=0, \tag{15a}$$

$$\left(\frac{\hat{\mathbf{p}}^2}{2m}+U(\mathbf{r})-E\right)\psi(\mathbf{r})=0, \tag{15b}$$

$$\left(\frac{\hat{\mathbf{p}}^2}{2m}+U(\mathbf{r})-E\right)\chi(\mathbf{r})=-2\lambda\sqrt{Vt}\,\delta(\mathbf{r}), \tag{15c}$$

$$V|\psi(0)|^2-t=0. \tag{15d}$$

In deriving the saddle-point equations we neglect the U dependence of the normalization constant $\mathcal{N}_\psi$. The reasoning leading to this approximation is explained in detail in the Appendix.

Below we will investigate only those solutions of Eqs. (15) which possess full rotational symmetry (or reflection symmetry in the one-dimensional case). While it is likely that even for rotationally invariant boundary conditions there exist solutions which break this symmetry, we will assume that they are not optimal, i.e., that they correspond to extremum points other than the global minimum of $\mathcal{A}$. The mean free time enters these equations only through the $\pi\nu_d\tau$ factor in Eq. (15a). It can be absorbed into redefinitions of $\chi(\mathbf{r})$ and λ, and therefore the above assumption does not depend on the disorder strength.

Eliminating χ in favor of U and ψ and switching to dimensionless variables $v(r)=U(|\mathbf{r}|)/E_F$, $r=p_F|\mathbf{r}|$, we arrive (in the absence of magnetic field) at the following system of equations (primes denote differentiation with respect to r):

$$y(r)=\ln'[r^{(d-1)/2}\psi(r)], \tag{16a}$$

$$y'(r)+y^2(r)+\frac{1}{4r^2}\delta_{d,2}+1=v(r), \tag{16b}$$

$$2y(r)[r^{d-1}v(r)]-[r^{d-1}v(r)]'=\bar{\lambda}, \tag{16c}$$

where

$$\bar{\lambda}=\frac{2i\lambda t}{\pi\nu_d\tau}\frac{p_F^d}{E_F^2 S_d}, \tag{17}$$

$\delta_{d,2}$ is the Kronecker symbol, and $S_d=dV_d$ is the area of a d-dimensional sphere of unit radius.

Equations (16a) and (16b) together are equivalent to the Schrödinger equation for the radial part of the wave function, while Eq. (16c) provides the connection between the wave function and the potential that is necessary to achieve a given value of t at a minimal ''cost'' in terms of the weight function $\mathcal{W}$. λ can be viewed as a parameter effecting a Legendre transformation of $\mathcal{A}$, so (λ,t) or, alternatively, $(\lambda,\ln t)$, form a pair of conjugate variables.

It follows from Eq. (16c) that $\bar{\lambda}$ must be real, so the saddle-point value of λ is purely imaginary. The saddle-point action $\mathcal{A}_0$ is expressed in terms of the solution $v(r)$ of Eqs. (16) as

$$\mathcal{A}_0=\frac{\pi\nu\tau}{2}\frac{E_F^2}{p_F^d}S_d\int_0^L r^{d-1}v^2(r)dr. \tag{18}$$

The choice of boundary conditions to Eqs. (16) is not entirely straightforward, and we would like to motivate it by appealing to an analogy with ordinary one-dimensional Schrödinger equation. In the variational formulation of quantum mechanics the Schrödinger equation appears as a stationarity condition for the functional

$$\mathcal{F}=E+\int_0^L dx\ \psi^*(x)(\hat{H}-E)\psi(x), \tag{19}$$

where $\hat{H}$ is the Hamiltonian and E is the particle energy. In Eq. (19), ψ must satisfy two boundary conditions [e.g., $\psi(0)=\psi(L)=0$] and also the normalization condition $\int_0^L dx|\psi|^2=1$ which follows from varying $\mathcal{F}$ with respect to E. The overall number of conditions for this second-order equation is three, which makes the problem overdefined. As a result, solutions exist only for a discrete set of values of E. One of these values would correspond to the ground state and thus to an absolute minimum of $\mathcal{F}$, while the rest would simply correspond to its stationary points. In the $L\to\infty$ limit the set of allowed E values becomes infinitely dense, so that E effectively becomes a continuous variable.

An alternative way to look at the Schrödinger equation (often useful in numerical calculations[26]) is to consider E as an extra unknown function satisfying a trivial equation

$$\frac{d}{dx}E=0. \tag{20}$$

One then has effectively three first-order equations with three extra conditions imposed on them. However, now these equations are nonlinear, and the three conditions do not specify the solution uniquely. Rather, there exists a family of solutions corresponding to various quantum states of the system, one of which minimizes the functional $\mathcal{F}$. Again, in the $L\to\infty$ limit the family of solutions becomes continuous.

Returning to the problem of specifying a set of conditions for Eqs. (16), we see that Eq. (16c) is a generalization of Eq. (20). λ plays the role of an eigenvalue, so we can formally supplement Eqs. (16) with $\lambda'=0$, increasing the number of equations to four. The following set of conditions must be imposed on this system of equations.

First, a boundary condition is provided by Eq. (15d),

$$\psi(0)=\left(\frac{t}{V}\right)^{1/2}. \tag{21a}$$

The normalization requirement

$$S_d\int_0^L r^{d-1}dr\,\psi^2(r)=1 \tag{21b}$$

provides the second condition.

Another boundary condition is derived from regularity requirements on $v(r)$. Since $v(r)$ can have at most a square integrable singularity as $r\to 0$, it follows that, in three dimensions,

$$\lim_{r\to 0}(\sqrt{r}\psi'(r))=0. \tag{21c}$$

In two dimensions ψ' must be finite at $r=0$, and for a one-dimensional system the corresponding requirement is $\psi'(0)=0$. Finally, a fourth condition may be provided by another boundary condition on ψ, e.g.,

$$\psi(L)=0. \tag{21d}$$

However, the problem of minimizing $\mathcal{A}$ is not solved uniquely by the set of equations (16) together with conditions (21). As in the ordinary quantum mechanical case, we are faced with a family of solutions only one of which actually corresponds to the absolute minimum $\mathcal{A}_0$ of the action $\mathcal{A}$. Finding this ''ground-state'' solution requires a parametrization of the whole discrete family of solutions which presents a difficult task. It can nevertheless be somewhat simplified by employing the following observation. For large enough L the set of solutions of Eqs. (16) may be approximately considered continuous. It is then possible to replace condition (21d) with a free boundary condition. The result is that the family of allowed solutions now admits a continuous parametrization, simplifying the task of determining the minimum of the functional $\mathcal{A}$. In practice this parametrization depends on the dimensionality d of the sample, and we will consider the cases of $d=1$, 2, and 3 separately. Employing a continuous parametrization which effectively corresponds to relaxing the boundary condition Eq. (21d) means not enforcing the condition that E_F is an eigenvalue of the Hamiltonian. As was already mentioned,[20] such an approximation leads only to errors that are beyond the exponential accuracy of the optimal fluctuation method. Such errors can therefore be safely neglected.

The general formalism presented above can be easily modified to describe the distribution of the current relaxation times. A crucial feature of this method is that it is not necessary to include the term proportional to λ into the action $\mathcal{A}$ from the very beginning. Without it, the right-hand side of Eq. (15c) is simply equal to zero and $\bar{\lambda}$ reappears in Eq.

(16c) as a constant of integration. The particular requirements on the wave function ψ with respect to which the action $\mathcal{A}$ is minimized are all hidden in the set of conditions Eqs. (21). To compute the distribution of the current relaxation times τ_r in the saddle-point approximation, one only has to modify the conditions set in Eqs. (21). Their explicit form depends on the geometry of the sample and external leads, and has to be established separately for samples of different dimensionality.

A. One-dimensional wire

We consider first a purely one-dimensional disordered wire of length $2L \gg l$. Exact results for this case obtained in Refs. 15 and 16 make it possible to use it as a ''laboratory'' for testing our method. From the computational point of view this case is simpler than those of the two- and three-dimensional conductors because the solution of Eqs. (16) can be written in a closed form. For any typical realization of the random potential all the eigenstates are exponentially localized in the one-dimensional case with a localization length being of the order of the mean free path l. This fact introduces complications which are absent when $d>1$ and disorder is weak. In the metallic samples in higher dimensions (see below, Secs. III B and III C) the same type of ''anomalously localized states'' is responsible both for very large eigenstate intensities in closed samples and very long current relaxation times in open ones. In the Anderson insulator regime, however, the states responsible for these two effects are rather dissimilar.

1. Distribution of current relaxation times

An anomalously long relaxation time τ_r is due to the states whose amplitude is suppressed at the edges of the sample beyond the typical value $e^{-L/l}/\sqrt{l}$. With the normalization integral $N=\int_{-L}^{L}|\psi(x)|^2 dx$ set equal to 1, the relaxation time in the regime $\tau_r\Delta \gg 1$ is inversely proportional to the eigenstate intensity at the edges,[13]

$$\tau_r\Delta \sim \frac{1}{L\psi^2(L)}. \tag{22}$$

Since exponentially large changes in this intensity lead only to algebraic changes in the intensity of the same state at its maximum, we can choose the ratio

$$\theta=\frac{\psi(0)}{\psi(L)} \tag{23}$$

as a parameter of the distribution function instead of τ_r (assuming that the state is peaked at $x=0$). With exponential accuracy, minimizing $\mathcal{A}$ with τ_r fixed is equivalent to minimizing with θ kept fixed, and the Jacobian of the transformation from θ to τ_r gives only a prefactor.

The system of Eqs. (16) takes the form

$$v(x)=1+y^2(x)+y'(x),$$

$$2y(x)v(x)-v'(x)=\tfrac{1}{2}\overline{\lambda}\,\mathrm{sgn}x, \tag{24}$$

where x is the dimensionless coordinate along the wire, and primes denote differentiation with respect to x. Note that according to Eq. (16a), y is simply the logarithmic derivative of the wave function when $d=1$.

Eliminating $v(x)$ and integrating once, we arrive (for $x>0$) at

$$(y')^2=(y^2+1)^2+\overline{\lambda}y+C. \tag{25}$$

Here C is a constant of integration which parametrizes the stationary points of $\mathcal{A}$ as described at the end of Sec. III A. Its value is fixed by a minimization procedure which is formally expressed as

$$\mathcal{A}_0=\min_C \mathcal{A}(C). \tag{26}$$

Although Eq. (25) is exactly integrable in terms of elliptic functions, it is sufficient to make a perturbative expansion in $\overline{\lambda}$ and C in order to recover the leading order dependence on τ_r. Rewriting Eq. (25) as

$$dx=-\frac{dy}{R(y)}, \quad R(y)=\sqrt{(y^2+1)^2+\overline{\lambda}y+C}, \tag{27}$$

we obtain the period T of the wave-function oscillations

$$T=2\int_{-\infty}^{\infty}\frac{dy}{R(y)} \tag{28}$$

and the logarithm of the ratio of the wave-function amplitudes at $x=0$ and $x=L$:

$$\ln\theta=\ln\frac{\psi(0)}{\psi(L)}=2\frac{L}{T}\int_{-\infty}^{\infty}\frac{y\,dy}{R(y)}. \tag{29}$$

We assume that the length of the sample corresponds to an integer number of periods. Therefore, since $y(0)=0$, it also follows that $y(-L)=y(L)=0$. The relative error introduced into Eq. (29) by neglecting the fractional part of L/T is $O(1/L)$.

Expanding the right-hand side of Eqs. (28) and (29) in both $\overline{\lambda}$ and C, we obtain $-\pi\overline{\lambda}/8=(T/2L)\ln\theta$ and $T=2\pi+O(C)$. For action $\mathcal{A}$ we then find from Eq. (18) to the lowest order (with $S_d=2$),

$$\mathcal{A}=\pi\nu_1\tau\frac{E_F^2}{p_F}\frac{2\pi L}{8T}(\overline{\lambda^2}+3C^2). \tag{30}$$

It is thus obvious that, at least to the lowest order in $1/L$, the minimum of $\mathcal{A}$ is achieved when $C=0$. Combining the results for the asymptotic tail of the distribution function with exponential accuracy we obtain

$$\mathcal{P}(\theta)\sim\exp\left\{-\frac{2l}{L}\ln^2\theta\right\}. \tag{31}$$

From Eq. (16a) and the fact that the change in the wave-function amplitude over one period (increase for $x<0$ or decrease for $x>0$) does not depend on x, we deduce the exponential form of the wave-function envelope

$$\psi(x)\propto e^{-(|x|/L)\ln\theta}. \tag{32}$$

Equation (22) now gives $\ln\tau_r\Delta=\ln\theta^2$, and the distribution in Eq. (31) is seen to coincide with the exact answer of Altshuler and Prigodin.[16]

The applicability of the above formula is restricted to the region $L/l<\ln\tau_r\Delta<p_FL$. Here the first inequality ensures that the localization length $L/\ln\theta$ is smaller than l, while the second inequality is necessary for the validity of the expansion in λ and C.

This model calculation can be used to illustrate the following points. First, to the lowest order in λ, we have

$$y(x)=\cot(x+\varphi)+\frac{\lambda}{4}\sin^2(x+\varphi), \tag{33}$$

where φ is a phase shift, and

$$v(x)=\tfrac{1}{2}\lambda\sin 2(x+\varphi), \tag{34}$$

so the shape of the optimal configuration of the potential is indeed a Bragg mirror.

Second, the sign of λ is opposite to that of $\ln\theta$, so negative values of λ correspond to the wave-function amplitude decreasing from the center of the sample outwards. The same will hold true in two and three dimensions. The amplitude of the oscillating potential is constant throughout the sample. A similar observation was made in Ref. 13: in a one-dimensional wire the gradients of the supersymmetric density matrix corresponding to the optimal solution were found to be x independent.

2. Distribution of eigenstate intensities

The amplitude of a typical localized state in its domain of localization is $1/\sqrt{l}\gg 1/\sqrt{2L}$. It is not sensitive to the details of the structure of the random potential outside this domain. Thus, to increase the amplitude of such a state beyond the $1/\sqrt{l}$ value, it is enough to adjust the random potential inside this domain to achieve a ''doubly localized'' structure of the state: starting from the center of the domain it would decay first with an exponent $1/\xi\gg 1/l$, and then revert to a typical $1/l$ decay.[23] Such a structure in principle cannot be described by the saddle-point equations alone, because it requires combining an optimal fluctuation inside the central ''bump'' with typical potentials elsewhere in the sample. Nevertheless an estimate of the distribution function can be obtained by solving the saddle-point equations in the region corresponding to the central ''bump'' (whose size must be determined self-consistently) and matching the solution to the expected typical exponential decay outside this region. Note that such a solution redistributes the weight of the wave function inside the size l region of the sample around the bump, but leaves the exponential tails outside almost unaffected.

The central bump can be created by the Bragg mirror Eq. (34) which extends over the region of size $2B\xi$, where B is a numerical factor to be determined later. We then have for the relative change of the wave function amplitude over half a period,

$$\ln\frac{|\psi(x)|}{|\psi(x+\pi)|}=\int_{-\infty}^{\infty}\frac{y\,dy}{R(y)}. \tag{35}$$

This change must be equal to π times the inverse localization length ξ. This gives $|\lambda|=8/\xi$. On the other hand, the wave-function amplitude at the origin $\psi^2(0)$ can be estimated as $1/\xi$ under the assumption that the dominant contribution to the normalization integral comes from the central bump. This assumption is valid if $e^{-B}<\xi/l$ or $B\sim\ln(l/\xi)$. Replacing L with $B\xi$ in Eq. (30), for the logarithm of the distribution function we obtain $\ln\mathcal{P}=-2l\psi^2(0)\ln(l\psi^2(0))$. This estimate gives a somewhat faster decay of $\mathcal{P}$ than the exact result $\ln\mathcal{P}=-4l\psi^2(0)$.[15] This is not surprising since, as was explained at the beginning of this subsection, typical configurations which lie beyond the scope of the saddle-point equations play an important role in the one-dimensional case.

Note that the configurations which dominate the distribution of relaxation times produce large values of eigenstate intensity with a much smaller probability than the ''doubly localized'' structure just described. Indeed, for these configurations we obtain $\psi^2(0)=\ln\theta/L$, leading to $\mathcal{P}\sim\exp\{-2lL\psi^4(0)\}$, which is much less than $\exp\{-2l\psi^2(0)\ln[l\psi^2(0)]\}$.

B. Two-dimensional conductors

In the absence of localization, even states whose wave function is strongly peaked somewhere in the interior of the sample have the dominant contribution to the normalization integral come from the bulk of the sample rather than from the vicinity of the peak. This statement ceases to be correct only when t reaches the values of the order of $V\sim L^d$. Thus, for states with $t\ll V$, the normalization integral is independent of t. Such states do not contribute to the $\tau_r\Delta\gg 1$ tail of the distribution of the relaxation times because their intensity at the boundary is close to the typical value $1/L^d$.[24]

For states which are peaked so strongly at some point $\mathbf{r}_0$ that the normalization integral is dominated by the contributions from its vicinity the situation is reversed: t is always of the order of V, while the intensity at the boundary can be made smaller than $1/V$ leading to $\tau_r\gg 1/\Delta$. In both regimes the quantity θ defined similarly to Eq. (23) as

$$\theta=\frac{\psi(0)}{\sqrt{L}\,\psi(L)}=\frac{\psi(0)}{L^{(d-1)/2}\psi(L)} \tag{36}$$

serves as a convenient parameter of the distribution function. Indeed, either $\psi(L)$ (in the first regime) or $\psi(0)$ (in the second regime) stays constant and becomes a convenient ''reference point.'' The factor $L^{(1-d)/2}=1/\sqrt{L}$ is introduced to cancel out the overall $r^{(1-d)/2}$ dependence of rotationally invariant wave functions in d dimensions. θ is always large in the asymptotic region $t\gg 1$.

In dimensions higher than one the system of equations (16) cannot be integrated exactly in terms of the standard functions. Nevertheless, numerical methods combined with the asymptotic analysis allow us to investigate fully the behavior of its solutions in various regimes. In two dimensions the solution can be obtained as an asymptotic expansion in λ/r. The values of t for which this expansion breaks down turn out to lie close to the limiting value πL^2, and are thus not very interesting. The leading terms of the expansion of $v(r)$ are

$$v(r)\approx\frac{\lambda}{2r}\sin(2r+2\varphi)+\frac{\eta}{r}\sin^2(r+\varphi). \tag{37}$$

In this expression η is a constant of integration which plays a role analogous to the role of C in the one-dimensional case—that of a minimization parameter. Just as in the one-dimensional case, the minimum of $\mathcal{A}$ is reached when $\eta=0$. We will give a perturbative proof of this statement at the end of this subsection. With $\eta=0$ the asymptotic expansion of $y(r)$ has the form

$$y(r)\approx\cot(r+\varphi)+\frac{\overline{\lambda}}{4r}\sin^2(r+\varphi). \tag{38}$$

Integrating $y(r)$ in the sense of the principal value, we obtain

$$\ln\theta=-P\int_{r_0}^{L}y(r)dr+\ln\theta_1\approx\frac{|\overline{\lambda}|}{8}\ln\frac{L}{r_0}+\ln\theta_1, \tag{39}$$

where the length scale r_0 at which the asymptotic expansion breaks down is defined as $r_0\sim\max(1,|\overline{\lambda}|)$, and $\ln\theta_1$ represents the contribution to the integral from distances $r\lesssim r_0$. This contribution can always be neglected with logarithmic accuracy. Inverting Eq. (39) we obtain $|\overline{\lambda}|=8\ln\theta/\ln(L/r_0)$. The dimensionless integral in the saddle-point action $\int_0^L v^2 r\,dr$ evaluates to $8\ln^2\theta/\ln(L/r_0)$, and we find

$$\mathcal{P}(\theta)\sim\exp\left\{-\pi^2\nu_2 D\frac{\ln^2(\theta^2)}{\ln(L/r_0)}\right\}. \tag{40}$$

The envelope of the wave function corresponding to the solution described by Eq. (37) is

$$\psi(r)\sim\frac{\psi(0)}{\sqrt{r}}(r_0/r)^{\ln\theta/\ln(L/r_0)}, \tag{41}$$

where we have approximated $\sqrt{r_0}\psi(r_0)\approx\psi(0)$. For $r_0\sim1$ the error introduced by this approximation leads only to $O(1)$ corrections to the large logarithm $\ln t$.

The normalization integral

$$N=2\pi\psi^2(0)\int_{r_0}^{L}dr\left(\frac{r_0}{r}\right)^{2\ln\theta/\ln(L/r_0)} \tag{42}$$

has two distinct regimes. When θ^2 is greater than L/r_0 the integral is dominated by small distances, and the dependence of t on θ is weak. In this regime t has a value close to πL^2, and $\ln\theta^2$ is identified with $\ln\tau_r\Delta$, so that Eq. (40) describes the asymptotic behavior of the distribution of relaxation times.

In the opposite case we have $N=2\pi\psi^2(0)r_0^{\alpha}L^{1-\alpha}/(1-\alpha)$, where $\alpha=\ln\theta^2/\ln(L/r_0)$. Up to irrelevant constants we then have $\ln t\approx\ln\theta^2$, and the distribution function acquires the form of Eq. (5b).

These results have essentially the same form as those obtained in Refs. 4 and 11–13 using the σ-model formalism. This fact suggests that both the approach employed here and the saddle-point solution of the σ model represent *the same* saddle point of the underlying theory.

They differ, however, in two important aspects. First, the logarithm in the denominator is cut off at distances ~1 (wavelength p_F^{-1} in conventional units) rather than at $r\sim l$, as in Ref. 12. We believe that this difference simply reflects the fact that in order to describe this saddle point exactly one has to include short-wavelength degrees of freedom which are neglected in the derivation leading to the σ model.

The second difference may turn out to be an indication of a deeper problem. Our calculation is performed explicitly for the case when there are no perturbations breaking the symmetry with respect to time reversal. Nevertheless, the numerical coefficient in the exponent coincides with the answer obtained in Ref. 12 for the *unitary* case ($\beta=2$), when the symmetry with respect to time reversal is completely broken. In order to see what effect violation of time-reversal invariance would have in our approach, we explicitly introduce magnetic field into the system of equations (15). Choosing the direction of the field $\mathbf{H}$ along the $\hat{z}$ axis (perpendicular to the two-dimensional sample), and writing $\mathbf{A}=\frac{1}{2}\mathbf{H}\times\mathbf{r}$, we can immediately see that under the assumption of circular symmetry of the solution ψ, only the terms quadratic in $\mathbf{A}$ survive. For weak enough fields $HL^2\ll\phi_0$ (where ϕ_0 is the magnetic flux quantum) these terms can be neglected because they only lead to exponential decay of the wave function on a scale larger than the sample size. Thus the set of equations (16) is unchanged. On the other hand, the fields of this magnitude are sufficient to suppress the Cooperon contribution in the σ-model calculation, leading to a crossover between the orthogonal and unitary symmetry classes. Therefore our calculations reproduce *exactly* the σ-model answer in the case of broken time-reversal invariance (the unitary symmetry class), but they predict a *faster* decay of the distribution function at large t (or τ_r) in the orthogonal case.

It well may be that in contrast with the unitary case the optimal fluctuations in the T-invariant two-dimensional systems are not rotationally invariant, and correspond to action that is one-half of the action for the best rotationally invariant fluctuation leading to Eq. (40). However, before the existence of such solutions of Eqs. (15) is demonstrated, it is impossible to exclude a possibility that the σ-model approach to the statistics of rare events in two-dimensional systems is not completely reliable. We believe that this issue deserves further consideration.

To complete the derivation of the distribution function presented in this subsection, we outline the proof of the statement that $\eta=0$ is the optimal choice. The solution ψ of equations (15) can be written as

$$\psi(r)=\frac{A(r)}{\sqrt{r}}\sin[r+\varphi(r)].$$

Unlike the $\eta=0$ case, the phase φ does not have a finite limit as $r\to\infty$. The logarithmic derivative $\tilde{y}(r)$ of the amplitude function $A(r)$ can be expanded in an asymptotic Fourier series of the form

$$\tilde{y}(r)\sim\sum_{n=1}^{\infty}\frac{1}{r^n}\left\{y_{(n,0)}+\sum_{m=1}^{\infty}\{y_{(n,m)}^{(c)}\cos m[r+\varphi(r)]+y_{(n,m)}^{(s)}\sin m[r+\varphi(r)]\}\right\}. \tag{43}$$

Substituting Eq. (43) and a corresponding expansion for $\varphi'(r)$ into Eqs. (15), we find that $y_{(1,0)}=\bar{\lambda}/8$ is independent of the constant of integration η. Integrating $y(r)$ over r we obtain $y_{(1,0)}\ln(L/r_0)$, and therefore the relation between $\ln\theta$ and $\bar{\lambda}$ established in Eq. (39) is also η independent to the leading order in $\bar{\lambda}$. On the other hand, adding the term $(\eta/r)\sin^2(r+\varphi)$ to v can only increase the value of the integral $\int_0^L v^2 r\,dr$ because the cross-term in the expansion of the square integrates to zero. It then follows that the minimum of the integral for a given value of θ is achieved by setting $\eta=0$. Note, however, that, as in the one-dimensional case, this proof is perturbative: it relies on the possibility to expand the solutions in powers of $\bar{\lambda}/r$.

C. Three-dimensional case

In the three-dimensional case, it can be demonstrated self-consistently that large values of t correspond to large negative values of $\bar{\lambda}$. As a result, there exists a range of values of r where expansion in $\bar{\lambda}/r^{d-1}=\bar{\lambda}/r^2$ is impossible. A typical solution which was obtained numerically using the so-called relaxation method[26] is shown in Fig. 1. One can distinguish three asymptotic regimes: (i) $r\ll r_1$, (ii) $r_1\ll r\ll r^*$, and (iii) $r\gg r^*$; it will be shown below that $r_1=1/|\bar{\lambda}|$ and $r^*=\sqrt{|\bar{\lambda}|/2}$. The first region corresponds to a potential well, and the second one to a potential barrier. Taken together, these two regions support a resonance in the s-wave channel at the energy E_F. However, besides a potential-well—potential-barrier combination, resonant scattering can be also caused by a weak periodic potential (Bragg reflection), and that is what the third region corresponds to. An interesting consequence of the solution presented in Fig. 1 is that the optimal way to achieve large values of t in three dimensions is to combine the two effects in the ''right'' proportion.

Analytically the three regions in Fig. 1 are described by the following asymptotic formulas. In the first region (at small r) $v(r)$ behaves as

$$v_1(r)\sim\bar{\lambda}/r+\bar{\lambda}^2\ln r+(1+c+\tfrac{7}{12}\bar{\lambda}^2), \tag{44}$$

where c is an arbitrary constant which cannot be determined from the boundary conditions Eq. (21). Note that the singularity is weak enough so that it produces only a finite contribution to the saddle-point action $\mathcal{A}_0$. The behavior of the wave function in this regime is given by

$$y(r)\approx\frac{1}{r}+\tfrac{1}{2}\bar{\lambda}+\tfrac{1}{3}\bar{\lambda}^2 r\,\ln r+cr, \tag{45}$$

which corresponds to

$$\psi(r)\approx\psi(0)[1+\bar{\lambda}r/2+O(r^2\ln r)].$$

The expansion in Eq. (44) breaks down for $r\sim 1/|\bar{\lambda}|$, which gives the approximate value for r_1. In the second region the solution can be obtained with the help of the semiclassical approximation, $y^2(r)\approx v(r)$, and the result is

$$v_2(r)\approx\left(\frac{|\bar{\lambda}|}{2r^2}\right)^{2/3}. \tag{46}$$

Finally, in the third region, where the $\bar{\lambda}/r^2$ expansion works, $v(r)$ has the asymptotic form

$$v_3(r)\sim\frac{\bar{\lambda}}{2}\frac{\sin(2r+2\varphi)}{r^2}+\eta\frac{\sin^2(r+\varphi)}{r^2}+O\left(\frac{1}{r^3}\right). \tag{47}$$

The constants η and φ in this expression are analogous to their counterparts in the two-dimensional case. The phase variable φ again has the meaning of the wave-function phase shift; it is finite due to the rapid decay of the potential at large r irrespective of the value of η. Constants η and φ are not independent: they both can be regarded as functions of c. Either η or c can be chosen as a minimization parameter.

We have not been able to establish the analytical dependence $\eta(c)$; however, based on a numerical analysis the qualitative features of this dependence can be described as follows. For a given $\bar{\lambda}$ there exists a critical value $c_0(\bar{\lambda})$ such that if $c>c_0$, the third (oscillatory) region never develops. Instead, $v(r)$ exhibits a singularity at some finite value of r, leading to a divergent integral in $\mathcal{A}_0$. So the values $c>c_0$ correspond to unphysical solutions. Exactly at the critical value c_0 the oscillations are also absent, and $v(r\to\infty)\approx 1+\bar{\lambda}^2/(4r^4)$. For slightly smaller c, $\eta(c)$ is large and positive ($\gg|\bar{\lambda}|$), and the oscillations appear only after a more or less protracted intermediate regime in which $v(r)\approx 1$. An estimate of the point r^* at which the onset of the oscillatory behavior occurs can be obtained by noting that the oscillations of $v(r)$ are driven [through Eq. (16)] by oscillations of the wave function. Therefore the amplitude of the oscillations in Eq. (47) cannot significantly exceed 1, so as to preserve the oscillatory—rather than exponentially damped—behavior of the wave function. This requirement leads to

$$r^*\approx\frac{1}{\sqrt{2}}(\bar{\lambda}^2+\eta^2)^{1/4}. \tag{48}$$

As c decreases further, η monotonically decreases as well, eventually covering the whole $(+\infty,-\infty)$ interval.

When η is positive or small negative ($\eta\gtrsim-|\bar{\lambda}|$), the behavior of $v(r)$ and $y(r)$ in the regions (i) and (ii) depends on η (or c) only very weakly so that this dependence can be ignored in computing the contribution of these regions to $\mathcal{A}$. Larger negative values of η start affecting the length of region (ii), and may lead to an emergence of a hybrid regime in which v oscillates nonharmonically with an $r^{-4/3}$ envelope.

In what follows we will assume that similarly to the two-dimensional case the minimal value $\mathcal{A}_0$ is achieved by setting $\eta=0$, even though the breakdown of the asymptotic expansion in $\bar{\lambda}/r^2$ at small distances makes it impossible to construct an analytical argument for this statement analogous to the proofs for one- and two-dimensional systems. This assumption is borne out by numerical analysis. It is certainly obvious that positive values of η can never be optimal because of a corresponding ''costly'' $v\sim 1$ region. As for large negative values, they are ruled out by the fact that a pronounced ''hybrid'' regime never appears in numerical solutions. Setting $\eta=0$ reduces Eq. (48) to $r^*\approx\sqrt{|\bar{\lambda}|/2}$.

With $\eta \approx 0$, both $v_2(r)$ and $v_3(r)$ reach values ~ 1 at $r = r^*$. While approximation schemes devised for $v \gg 1$ and $v \ll 1$ break down around r^*, we can calculate contributions θ_1 and θ_2 to $\theta = \psi(0)/L\psi(L) = \theta_1\theta_2$ from regions $r < r^*$ [(i) and (ii)] and $r > r^*$ (iii) separately. Similarly to the two-dimensional case, the L factor in the denominator of θ is introduced to cancel the overall $1/r$ dependence of the wave functions, which is an artifact of spherically symmetric boundary conditions.

(i) and (ii) We combine together the first and the second regions because the wave function amplitude does not change appreciably in the very short region (i). Region (ii) corresponds to a stretched exponential decay of ψ. Integrating $y(r) = -\sqrt{v_2(r)}$ over r from $1/\sqrt{|\bar{\lambda}|}$ to $r^* \approx \sqrt{|\bar{\lambda}|}/2$, we obtain

$$\ln\theta_1 \approx \int_{1/\sqrt{|\bar{\lambda}|}}^{\sqrt{|\bar{\lambda}|}/2} dr \left(\frac{|\bar{\lambda}|}{2r^2} \right)^{1/3} \approx 3r^*. \tag{49}$$

(iii) Using Eq. (47) with $\eta = 0$, we find

$$y(r) \sim \cot(r+\varphi) + \frac{\bar{\lambda}}{4r^2}\sin^2(r+\varphi). \tag{50}$$

The first term in this expression stems from the oscillations of the wave-function, while the second one describes the decrease ($\bar{\lambda} < 0$) of the wave-function envelope from the center of the sample outwards. The second contribution $\ln\theta_2$ is given by the principal value of the integral $\int y\, dr$,

$$\ln\theta_2 \approx -P\int_{\sqrt{|\bar{\lambda}|}/2}^{L} y(r)dr \approx r^*/4. \tag{51}$$

Adding the two contributions, we establish the relation between $\bar{\lambda}$ and θ,

$$\ln\theta = \frac{13}{4}r^* = \frac{13}{4}\left(\frac{|\bar{\lambda}|}{2} \right)^{1/2}, \tag{52}$$

verifying the self-consistency of the assumption $|\bar{\lambda}| \gg 1$ made at the beginning of this subsection.

The dimensionless integral in the saddle-point action is also evaluated separately in the combined regions (i) and (ii) and in region (iii):

$$\int_0^{r^*} r^2 v^2 dr = 3r^{*3} = \frac{192}{13^3}\ln^3\theta \quad [\text{regions (i) and (ii)}], \tag{53a}$$

$$\int_{r^*}^{L} r^2 v^2 dr = r^{*3}/2 = \frac{32}{13^3}\ln^3\theta \quad [\text{region (iii)}]. \tag{53b}$$

Both contributions turn out to be proportional to the same power of $\ln\theta$, indicating that local resonances and Bragg reflection play equally important roles in the formation of anomalously large wave function intensities. Combining the results, we obtain

$$\mathcal{P}(\theta) \sim \exp\left\{ -\frac{56}{2197} p_F l\, \ln^3\theta \right\}. \tag{54}$$

The envelopes of the wave function in the two regions are

$$\psi(r) \sim \frac{\psi(0)}{r}\exp\{-3r^{*2/3}r^{1/3}\} \quad \left(\text{for } \frac{1}{r^{*2}} < r < r^* \right) \tag{55a}$$

and

$$\psi(r) \sim \frac{\psi(0)}{r}\exp\{r^{*2}/4r - \tfrac{13}{4}r^*\} \quad (\text{for } r > r^*). \tag{55b}$$

The normalization integral is given by

$$N = \pi\psi^2(0)\left\{ \frac{1}{3r^{*2}} + 4Le^{-(13/2)r^*} \right\}. \tag{56}$$

When $L \gg \theta^2$ (with more realistic boundary conditions this inequality will become $L^3 \gg \theta^2$), the normalization integral is dominated by the contribution from region (iii). Then $\psi^2(0) \propto \theta^2$, and with exponential accuracy we finally obtain

$$\mathcal{P}(t) \sim \exp\{-\tfrac{7}{2197}(p_F l)\ln^3 t\}. \tag{57}$$

Of course, the separation into regions $r < r^*$ and $r > r^*$ is approximate. It is possible that the crossover region $r \sim r^*$ makes a contribution of the same order of magnitude as the two asymptotic regions (ii) and (iii). Thus the number $\frac{7}{2197} \approx 3.2 \times 10^{-3}$ can only be considered as an order-of-magnitude estimate of the coeficient κ introduced in Sec. II. It must be mentioned that convergence to the asymptotic form of Eq. (57) is extremely slow because of the stringency of the requirement $r^* = \frac{4}{13}\ln\theta \gg 1$.

In the opposite regime $L \ll \theta^2$, $\psi^2(0)$ is close to its maximal value ~ 1, and therefore its θ dependence is very weak. On the other hand, similarly to the one-dimensional case, θ^2 becomes proportional to the electric response time τ_r, leading to the distribution of τ_r having a form identical to Eq. (57),

$$\mathcal{P}(\tau_r) \sim \exp\{-\kappa(p_F l)\ln^3 \tau_r \Delta\}. \tag{58}$$

Both Eqs. (57) and (58) differ significantly from the corresponding σ-model results.[12,13] We discuss the possible origins of this difference in Sec. IV.

IV. DISCUSSION

A. General remarks

The main result of the work presented in this paper is that statistics of seldom occurring events in disordered conductors can be successfully studied using the optimal fluctuation method. Previous approaches to the problem have been based on various formulations of the nonlinear σ model, and they invariably seem to require an extension of the σ model to, and sometimes beyond, its limits of validity.

The success of the σ model in describing a wide variety of phenomena in chaotic and disordered systems can be traced to the fact that most such phenomena are semiclassical in nature, and are determined by typical extended quantum states formed by typical fluctuations of the random potential. An investigation of the statistics of rare events, on the other hand, presents quite a different type of problem. The configurations of the random potential that give rise to such events come from a small subset of all possible configura-

tions. It follows then that there are certain disadvantages associated with the σ-model approach to the study of rare events.

(i) First, the assumption that the motion of electrons can be described entirely in semiclassical terms imposes certain restrictions on the types of the potentials over which averaging is performed. As follows from the calculations presented here, in three spatial dimensions there exists a region $r \sim r^*$ where the solution of the Schrödinger equation cannot be obtained semiclassically. As a result, the σ model fails to recognize the existence of this region, and it also misses entirely the local resonance formed at $r < r^*$.

(ii) Second, even when the σ-model calculations succeed in correctly—albeit implicitly—identifying the relevant disorder configurations (as in one- and two-dimensional samples) they do not always produce exact answers because contributions from the short-wavelength degrees of freedom (massive modes) eliminated in the transformation from the fast to slow variables are missed.[28]

B. Direct optimal fluctuation method

It is not immediately evident, however, that a relatively naïve approach based on the direct search for an optimal fluctuation should be more reliable. In order for this to work, the probability of observing a large value of t must be determined by a sum over disorder configurations which all come from a single compact region of the configurations space. The potentials forming this region differ only slightly from some optimal configuration which corresponds to the saddle point. Although we have not proved in this work that the saddle point identified here gives the dominant contribution to the functional integral, ''the preponderance of evidence'' based on the comparison of our results with those obtained using the σ-model would indicate that this is indeed the case.

The starting point of our qualitative analysis is the observation that, apart from the difference in the cutoff scale, our variant of the optimal fluctuation method (the ''direct'' optimal fluctuation method) reproduces identically in the two-dimensional case the σ-model results for the unitary ensemble as well as the one-dimensional result for the distribution of relaxation times. In addition, the shapes of the wave function envelopes obtained by our method essentially coincide with the averaged envelopes obtained by Falko and Efetov in Ref. 12 (see below). Assuming that more than a chance coincidence is involved, it is reasonable to conclude that the optimal configurations of disorder found in this work are the same as the ones that are responsible for the saddle point of the σ model. This conclusion is also supported by the fact that our results are not an artifact of the particular model chosen here. For example, the distribution function in Eq. (5b) is unchanged if the Gaussian fluctuations of the potential have a finite correlation length. Altering the dispersion law in the free Hamiltonian also does not change the basic features of the results, such as the log-normal form of the distribution function or the inverse proportionality of the coefficient C_2 to the logarithm of the system size.[28] In other words, there must exist a correspondence between the saddle points of the theory defined by Eqs. (14) and the saddle points of the σ model. Assuming the existence of such a correspondence we will try to elucidate the origins of the σ-model results by analyzing the corresponding optimal configurations of the potential $U(\mathbf{r})$.

Before proceeding with this analysis, we would like to address briefly the question of the stability of the saddle point described by Eqs. (15). In the three-dimensional case the dominant contribution to the saddle-point action $\mathcal{A}_0$ comes from small distances where the optimal potential is much larger than its typical Gaussian fluctuation, and stability with respect to fluctuations does not pose a problem. In two dimensions, however, the outlying regions of the optimal configuration, where the magnitude of the potential vanishes as $1/r$, must be taken into account. In order to argue that fluctuations do not destroy the saddle point we note that, although we find it convenient to write the Gaussian distribution function $\mathcal{W}[U]$ for the potential in the coordinate representation, it can be written in any orthonormal basis $\{f_n(\mathbf{r})\}$. A typical amplitude of a dimensionless basis function f_n in a typical configuration $U(\mathbf{r})$ is $1/\sqrt{\nu_2 D}$. Let us now choose one of the basis functions, say $f_0(\mathbf{r})$, to be proportional to the optimal solution $v(r)$ given by Eq. (37). Using the normalization condition $\int d\mathbf{r}\, f_0^2(\mathbf{r}) = 1$, we obtain

$$f_0(r) \approx \left(\frac{1}{\pi \ln(L/r_0)} \right)^{1/2} \frac{\sin 2r}{r} \tag{59a}$$

and

$$v(r) \sim \frac{\ln\theta}{\sqrt{\ln(L/r_0)}} f_0(r). \tag{59b}$$

Thus the optimal fluctuation has a much larger amplitude than the typical one as long as $\ln^2\theta \gg [\ln(L/r_0)]/\nu_2 D$. This condition, of course, is just a natural requirement for the validity of the optimal fluctuation method $|\ln \mathcal{P}| \gg 1$.

It remains to be shown, however, that other components of a typical fluctuation $U(\mathbf{r})$ (i.e., those orthogonal to f_0) do not destroy the saddle point. A rigorous investigation of the fluctuations around the saddle point defined by Eqs. (15) will be the subject of a forthcoming publication. Nevertheless, a plausible argument in favor of the stability of this saddle point can be made based on the following observation. The appearance of anomalously localized states due to Bragg reflection can be viewed as a phenomenon analogous to the emergence of a band structure in a periodic lattice. Therefore suppression of this effect by the fluctuations of the random potential is equivalent to the localization transition which destroys the band structure in ordinary periodic lattices. Therefore $L < L_c$, where L_c is the localization length, seems to be a sufficient condition for the stability of the saddle-point solution in the two-dimensional case.

C. Asymptotics of the distribution functions

1. One-dimensional case

We will now try to use the physical intuition afforded by the optimal fluctuation concept to compare the σ-model results of Eqs. (1) and (2) with the distribution functions Eqs. (5), (31), (40), and (58) derived in Sec. III. The first question that can be easily answered is why the distribution function for the electric response times in the one-dimensional case is log-normal instead of a power law [i.e., $\exp(-C_1 \ln \tau_r)$] ob-

tained by a naïve extrapolation of Eq. (1) to $d=1$. In one and two dimensions the optimal configurations found in Sec. III are ''global,'' i.e., the integration in the saddle-point action $\int v^2 d\mathbf{r}$ must be extended over the whole sample. In contrast, in the three-dimensional case the optimal fluctuation, even with the oscillating tail included, is local, so that the above integral converges at large distances. It is well known[29,30] that distribution function tails of the $e^{-\ln^d \tau_r}$ type usually appear as probabilities of optimal fluctuations confined to a finite volume. Thus, in order to explain the ''anomaly'' in the one-dimensional case, we have to understand why a local fluctuation of the random potential necessary to achieve a given value of τ_r has a lower probability than the global one proposed in Ref. 13 and rederived in Sec. III A. A local fluctuation of the potential leading to a large value of τ_r would have to be able either to support a narrow resonance in the bulk of the sample or to suppress the wave function near the edges exponentially. Both scenarios require a large potential barrier—to form a narrow resonance in the first case or to create a classically inaccessible region near the edges in the second one. Assuming, as in Sec. II, a rectangular shape for such a potential barrier, we can repeat the calculation presented there almost verbatim except that t must everywhere be replaced by $\tau_r\Delta$. We then obtain $\ln\mathcal{P}\approx(2/3\sqrt{3})(p_F l)\ln\tau_r\Delta$. In order for this behavior to dominate we must have $p_F l\ln\tau_r\Delta\ll(l/L)\ln^2\tau_r\Delta$ or

$$\ln\tau_r\Delta\gg p_F L. \tag{60}$$

However, from Eq. (9) we see that the corresponding values of b—which determine the size of this local fluctuation of the potential—become larger than the length of the sample $2L$. This is clearly unphysical. Thus despite a slower τ_r dependence of the probability of local resonances, the same value of τ_r has a much larger probability to be produced by an accidentally formed Bragg mirror for all reasonable values of τ_r.

Note also that $\ln\tau_r\Delta\sim p_F L$ corresponds to $v\sim 1$ (or $U\sim E_F$), which invalidates the perturbative expansion of Sec. III A. Essentially, the values of $\ln\tau_r\Delta$ of the order of $p_F L$ or larger correspond to a trivial case: a sample is insulating because the potential is larger than E_F almost everywhere except for a small island in the center where almost all the weight of an eigenstate at E_F is concentrated. In this regime the distinction between global and local fluctuations of the potential becomes blurred.

It is interesting to note that a calculation based on the ballistic σ model performed in Ref. 13 for *quasi*-one-dimensional conductors indicates the existence of a crossover from log-normal to power-law distribution at $\ln\tau_r\Delta\sim L/l$. Although we have not investigated the quasi-one-dimensional case here, it is likely that the physical picture of the interplay between the global and local fluctuations discussed above should not be much different. If that is the case, then it is probable that the crossover found in Ref. 13 has as its underlying cause the same mechanism of local fluctuations becoming comparable in size to the length of the sample. This hypothesis, however, leaves unexplained the difference in scales at which the crossover occurs – L/l in Ref. 13 as opposed to $p_F L$ in the argument presented above. It is possible that the quasi-one-dimensional case brings in some features that are not recognized by the estimates based on the purely one-dimensional model. It should be mentioned, however, that none of the variants of the σ model can provide an adequate description of the effects associated with local resonances, and it is possible that a σ-model estimation of the crossover scale may not be entirely reliable.

2. Two-dimensional conductors

A discussion of the results obtained by the direct optimal fluctuation method in the two-dimensional case and their counterparts established using the σ-model formalism has already been presented in Sec. III B. Here it seems appropriate to reiterate briefly the following two main points of that discussion. First, it comes as no surprise that when massive modes are taken into account, the short-distance cutoff scale in the logarithm determining the system size dependence of $\ln\mathcal{P}$ becomes of the order of the electron wavelength p_F^{-1} rather than the mean free path l. This leads, however, only to a small relative change in the σ-model result when the inequality $p_F L\gg(p_F l)^2$ is satisfied. On the other hand, the apparent ensemble independence of the distribution function asymptotes obtained by our method is puzzling and requires additional investigation.

3. Three-dimensional case

Derivation of the σ model involves linearization of the spectrum near the Fermi energy. As a result, there is no intrinsic scale in the model that would relate the amplitude of the fluctuating potential to the electron energy. The only scale is provided by the dispersion of the fluctuations of the potential $1/\pi\nu_d\tau$. This limitation of the model does not affect the computation of probabilities of typical events, or of the averages that are dominated by such events, because typical potentials are small compared to the Fermi energy. However, a problem arises when rare large-amplitude fluctuations of the potential $U(\mathbf{r})$ become dominant. The σ model cannot detect the existence of classical turning points around which the semiclassical approximation breaks down. Thus a possible explanation of the log-normal distribution obtained in Ref. 12 for three-dimensional systems is that, within the σ-model approach, the classical turning point at r^* is missed, and the Bragg mirror is effectively assumed to persist until distances of the order of the mean free path l.

We can introduce an (incorrect) cutoff at l into the calculation performed in Sec. III C in order to see what changes in the result will be induced by it. In the notation of Sec. III C, that would correspond to $|\overline{\lambda}|\sim l\ln\theta$, and then $\mathcal{A}_0\propto\int_l^L v^2 r^2 dr\propto|\overline{\lambda}|^2/l\propto l\ln^2\theta$, leading to $\ln\mathcal{P}\sim-(p_F l)^2\ln^2 t$, which coincides with the answer obtained in Ref. 12. The reason that such a cutoff scheme leads to a higher estimate for the probability $\mathcal{P}$ is that it does not correspond to a correct solution of the Schrödinger equation in the region of large potentials $U>E_F$. As a result, the rate of growth of the wave-function amplitude toward the center of the sample is overestimated.

Another discrepancy between our results and those obtained with the help of nonlinear σ models in the three-dimensional case is the difference in powers of $(p_F l)$ in the exponents in Eqs. (1) and (58). We believe that this discrepancy has the same origin as the difference between $\ln(L/l)$

and $\ln(L/r_0)$ in the two-dimensional case: the error introduced by using the mean free path l to determine the cutoff scale. It is interesting to note that while the cutoff procedure used in the diffusive σ-model approach of Ref. 10 is capable of producing only an order-of-magnitude estimate $C_3 \sim (p_F l)^2$, the ballistic σ model,[17,13] while giving an illusion of computing the coefficient C_3 exactly $[C_3=(\pi/9\sqrt{3})(p_F l)^2]$, nevertheless leads to the same extra power of $(p_F l)$. To explain this seemingly paradoxical situation we will first examine the cutoff procedure employed in Ref. 10. It is based on the condition, pointed out in Ref. 7, that in order for the diffusive σ model to be applicable, the spatial gradients of the Q-matrix components cannot exceed $1/l$. Explicitly, for the calculation performed in Ref. 10, this condition reads

$$\left| \frac{d}{dr} \ln \lambda_1 \right| < \frac{1}{l}, \tag{61}$$

where λ_1 parametrizes the noncompact bosonic sector of the Q matrix[8] (not to be confused with the Lagrange multiplier λ used throughout this work). The distance l_* where this condition is violated is used in Ref. 10 as a short-distance cutoff. It was conjectured in Ref. 12 and later confirmed in Ref. 27 that the spatial structure of the saddle-point solution for λ_1 mimics the envelope of anomalously localized states described by the saddle point of the σ model. Therefore the optimal configuration of $(d/dr)\ln\lambda_1(r)$ corresponds to the nonoscillating part of $y(r)$ in the direct optimal fluctuation method.

This correspondence allows us to see directly what effect an artificial short-distance cutoff at l_* would have in our approach. From Eq. (50) we find $l_*^2 \sim l|\overline{\lambda}|$. Introducing such a cutoff at l_* into Eq. (51) we obtain $|\lambda| \sim l_* \ln\theta$, and therefore l_* is estimated as $l_* \sim l \ln\theta$. If then the integral determining the saddle-point action $\mathcal{A}_0$ is also cut off at l_*, we obtain

$$\mathcal{A}_0 \sim l \int_{l_*}^{L} v^2 r^2 dr \sim l \frac{|\overline{\lambda}|^2}{l_*} \sim l^2 \ln^3 \theta,$$

or $(p_F l)^2 \ln^3\theta$ in conventional units, leading to the incorrect answer that was already quoted in Sec. I.

It is thus evident that distances shorter than l_* make an important contribution to the saddle-point action, and this contribution cannot be accounted for by the diffusive nonlinear σ model. It should be emphasized that excluding the short-distance contribution in this way leads to a significantly smaller estimate for the probability of observing a given (large) value of t or τ_r. This can be understood with the help of the following argument. The wave-function amplitude grows substantially between $r \sim l_*$ and $r \sim 1$. Neglecting this growth leads to a need for a faster increase in the wave-function amplitude between L and l_* which can only be achieved by means of a ''costly'' boost in the amplitude of the Bragg mirror. As a result, nonoptimal configurations of the random potential are selected.

Turning now to the calculation performed with the help of the ballistic nonlinear σ-model for the three-dimensional case in Ref. 13, we notice that the distance r_* that separates the ''reaction'' and ''run-out'' zones is of the same order as l_*. The contribution to the saddle-point action from the run-out (diffusive) zone leads to the already quoted $(\pi/9\sqrt{3}) \times (p_F l)^2$ value for C_3, while the reaction zone produces a contribution that has one less power of the large logarithm $\ln\theta$ and is thus neglected. In contrast, in the calculation presented in Sec. III C of this paper, the contribution from distances of the order of $r^* \sim \ln\theta \ll l_*$ dominates the saddle point action and leads to a larger estimate [Eq. (5a)] for the probability of observing anomalously high values of t in the three-dimensional case than the one obtained in Ref. 13. Thus the ballistic generalization of the nonlinear σ model is also not capable of detecting the existence of the scale r^* at which the semiclassical approximation breaks down. Moreover, since the calculation within the framework of the ballistic σ model does not involve any ultraviolet divergencies that would necessitate a short-distance cutoff as in the case of the diffusive σ model, it appears plausible that the ballistic variant of the model is equivalent in this context to introducing an ultraviolet regularization into the theory.

D. Prelocalized wave functions

To complete the comparison of the results obtained by the direct optimal fluctuation method and those found using the nonlinear σ models, we now turn to the issue of the shape of the envelope of anomalously localized states that are responsible for the large-t tails of the distribution function $\mathcal{P}(t)$. In Ref. 12 it was found that in two dimensions such states are characterized by amplitudes decaying outwards according to a power law

$$|\psi(\mathbf{r})|^2 \sim \left(\frac{l}{r}\right)^{\ln t/\ln(L/l)}. \tag{62}$$

This result has to be compared with Eq. (41), which can be rewritten as

$$|\psi(\mathbf{r})|^2 \sim \frac{\psi^2(0)}{r}\left(\frac{r_0}{r}\right)^{\ln t/\ln(L/r_0)}. \tag{63}$$

Apart from the difference in the cutoff scale (l vs r_0) which was discussed at length above, Eq. (63) contains an extra $1/r$ factor in the denominator. It is simply a consequence of the idealized model adopted here with its circularly symmetric boundary conditions. In a more realistic model the optimal wave function would become a superposition of different angular momentum eigenstates. The $1/\sqrt{r}$ behavior of the circularly symmetric component would be canceled in such a superposition. Taking into account the fluctuations around the saddle point would also have the effect of suppressing the $(1/\sqrt{r})^2$ factor in the spatial dependence of the *averaged* envelope of the optimal solution. Note also that Eq. (62) does indeed describe the averaged envelope of anomalously localized states.

In three-dimensional samples the states with anomalously high local amplitudes were found in Ref. 12 to have the envelope

$$|\psi(\mathbf{r})|^2 \sim \exp\left\{-A\left(1-\frac{l}{r}\right)\right\}, \tag{64}$$

where A is a constant which in the leading logarithmic approximation is equal to $\ln t$. Comparing this to Eq. (55) we see, in accordance with the discussion above, that the σ-model result gives the correct functional form $e^{\text{const}/r}$ only for large r, i.e., in the region of space where the optimal potential forms a Bragg mirror. The estimate $l \ln t$ for the constant in the exponential in Ref. 12, however, contains an extra factor of l compared to Eq. (55b) which is a consequence of the choice of the cutoff scale made in Ref. 12. The region of the stretched exponential decay of the wave-function envelope described by Eq. (55a) is missed in the σ-model calculation entirely. As in the two-dimensional case, the extra powers of r in the denominators of Eqs. (55) are a consequence of the artificial rotational symmetry.

It was conjectured in Ref. 12 that these high-amplitude states have a complicated ''snakelike'' structure at short distances. The conjecture was based on the fact that the method employed in Ref. 12 was applicable to amplitudes as high as

$$t \lesssim \frac{V p_F^{d-1}}{l}, \tag{65}$$

rather than a naive $t \lesssim (V/l^d)$ expected from a cutoff at l. We have found no evidence of such behavior here. Since the solutions of the saddle-point equations were assumed to be rotationally invariant from the outset, this cannot be regarded as a conclusive evidence against such a scenario. However, potentially more important is the fact that we did not encounter any limitations on the possible values of t analogous to Eq. (65). Analysis of the fluctuations around the saddle point defined by Eqs. (15), as well as an investigation of the possibility for non-rotationally-invariant solutions of these equations is needed to settle the issue conclusively.

E. Universality

Finally, we would like to make a few remarks concerning the issue of universality. An important consequence of the dominant role played in three dimensions by large local fluctuations of the potential that are responsible for the formation of local resonances is the nonuniversal character of the distribution functions derived in Sec. III C. Indeed, such large-amplitude configurations of the potential can only be optimal if they are not too ''expensive,'' i.e., if their action $\mathcal{A}$ is not too high, compared to the low-amplitude global fluctuations of the potential (Bragg mirrors). Generalizing the distribution of the random potentials to

$$\mathcal{W}_n[U] \propto \exp\left(-\frac{1}{2\sigma} \int U^{2n}(\mathbf{r}) d^d\mathbf{r} \right), \tag{66}$$

where σ is the dispersion, and $n \geqslant 1$, we find through a model calculation similar to the one performed in Sec. II that forming states with large amplitudes at some $\mathbf{r}$ by means of local resonances always leads to $\ln \mathcal{P} \propto -\ln^d t$. For $n < d/2$ the ''cost'' of a corresponding Bragg mirror can be estimated from an appropriate generalization of Eqs. (16), and it also leads to a $\ln^d t$ dependence, as was demonstrated by a detailed study of Gaussian ($n=1$) distribution in three dimensions. Thus local large-amplitude fluctuations of the potential are always important for small enough ($< d/2$) values of n. In the marginal case $n = d/2$, the weight of the Bragg mirror is given by

$$\ln \mathcal{P} \propto -\frac{\ln^d t}{\left(\ln \dfrac{L}{r_0} \right)^{d-1}}, \tag{67}$$

making it a much more probable way of achieving a large local wave-function intensity t. When $n > d/2$ the probability of the corresponding Bragg mirror has a faster than $\ln^d t$ dependence on $\ln t$, but it is compensated by a power of the system size in the denominator,

$$\ln \mathcal{P} \propto -\frac{\ln^{2n} t}{L^{2n-d}}. \tag{68}$$

This situation is realized in the Gaussian case when $d=1$, and it was already discussed in detail. Therefore the seemingly universal character of the log-normal behavior of the tails of the distribution functions is tied to the assumption, which is a usual starting point in the derivation of the σ model, that the random potential has a Gaussian distribution. This assumption is believed not to be crucial for the applicability of the diffusive σ model. Indeed, the diffusion constant involves only the second-order correlator of random potentials. On the other hand, when rare events are considered, large fluctuations of $U(\mathbf{r})$, and therefore the details of its distribution, may become important.

V. OPEN QUESTIONS AND CONCLUSIONS

The most important problem arising from the results of the present study is the need to explain the fact that the asymptotics of the distribution functions in two dimensions derived using the direct optimal fluctuation method do not exhibit any dependence on weak magnetic fields. This feature of our answers must be contrasted with *all* the previous calculations performed in the framework of the nonlinear σ models, in which it was quite obvious that when Cooper modes acquire a mass as a consequence of broken invariance with respect to time reversal, the number of independent components of the Q matrices changes, and that has a profound impact on the results. It is possible that our assumption that rotationally invariant solutions of the saddle-point equations dominate the saddle-point action is not valid in the absence of magnetic field. A more detailed study of the properties of Eqs. (15) will be the subject of a future publication. However, it is impossible to exclude the possibility that the separation of the low-lying exitations of the σ model into the Cooper and diffusion modes may become inexact when the rare configurations of the random potential dominate.

Whether or not the fluctuations around the saddle point can change the leading-order terms in $\ln t$ is also one of the questions that are outside the scope of this work. A peculiar disagreement in the crossover scale from $\ln^2 t$ to $\ln t$ asymptotics of $\ln \mathcal{P}$ between the purely one-dimensional and quasi-one-dimensional cases which was noted in Sec. IV C 1 makes it rather desirable to extend our method to the quasi-1D and quasi-2D geometries.

To conclude, the main results of this work can be summarized as follows. We have demonstrated that the optimal

fluctuation method is a useful tool for the investigation of the statistical properties of anomalous electronic eigenstates in disordered two- and three-dimensional conductors. It was shown that in three dimensions this method is preferable to the nonlinear σ model because the latter does not include effects associated with local resonances which can be formed by the random potential. In the one-dimensional case our approach is shown to reproduce the results obtained earlier with the help of the Berezinskii technique[16] for the distribution of the current relaxation times in open samples. The relevant optimal configurations of the potential coincide with those conjectured in Ref. 13. We have also demonstrated that these configurations are very different from the ones which dominate the asymptotics of the distribution function of the eigenstate intensities in closed samples, thus clarifying the origins of the difference between these two distributions.

In the two-dimensional case the results obtained by the optimal fluctuation method essentially coincide with the σ-model results for the unitary ensemble of random potentials, which we interpret as an indication that the saddle point of the reduced nonlinear σ-model found in Refs. 7, 9, 12, 10, and 13 corresponds to the same saddle point of the full theory as the one that describes the optimal fluctuation of the potential.

ACKNOWLEDGMENTS

The authors would like to thank A. Andreev, D. Fisher, B. Halperin, I. Lerner, H. Li, M. Raikh, and N. Wingreen for useful discussions. We are especially grateful to A. Mirlin for his valuable comments concerning the nature of anomalously localized states in one-dimensional wires. I. S. acknowledges financial support from NSF Grants Nos. DMR 9106237 and DMR 94-16910.

APPENDIX

Introducing an auxiliary field χ as in the main text, we can represent $\mathcal{N}_\psi$ as

$$\mathcal{N}_\psi^{-1}=\int \mathcal{D}\psi\ \mathcal{D}\left(\frac{\chi}{2\pi}\right) \times\exp\left[i\int d\mathbf{r}\ \chi(\mathbf{r})\left(\frac{\hat{\mathbf{p}}^2}{2m}+U(\mathbf{r})-E\right)\psi(\mathbf{r})\right]. \quad \text{(A1)}$$

Performing a $\pi/4$ rotation in the (χ,ψ) space,

$$\chi=(\psi_1+\psi_2)/\sqrt{2},\quad \psi=(\psi_1-\psi_2)/\sqrt{2}, \quad \text{(A2)}$$

and introducing infinitesimal convergence factors, we obtain

$$\mathcal{N}_\psi^{-1}=\int \mathcal{D}\left(\frac{\psi_1}{\sqrt{2\pi}}\right)\mathcal{D}\left(\frac{\psi_2}{\sqrt{2\pi}}\right) \times\exp\left(\frac{1}{2}\,i\int d\mathbf{r}\{\psi_1\hat{O}^{(+)}\psi_1-\psi_2\hat{O}^{(-)}\psi_2\}\right), \quad \text{(A3)}$$

where $\hat{O}^{(\pm)}=(\hat{\mathbf{p}}^2/2m)+U(\mathbf{r})-E\pm i0$. Thus, up to a constant, $\mathcal{N}_\psi$ is equal to a symmetrized spectral determinant:

$$\mathcal{N}_\psi\propto\sqrt{\det\{E-\hat{H}+i0\}\det\{E-\hat{H}-i0\}}=e^{\Re\ \mathrm{Tr}\ \ln(E-\hat{H}+i0)}. \quad \text{(A4)}$$

The variation of the logarithm in the exponent with respect to $U(\mathbf{r})$ is equal to the real part of the Green's function at coinciding arguments $\Re G(\mathbf{r},\mathbf{r};U)$. This quantity probes the whole band and is not sensitive to small changes in U. Therefore, it can be treated as a constant (which we denote as G), so that $\mathcal{N}_\psi$ becomes

$$\mathcal{N}_\psi\propto\exp\left(G\int d\mathbf{r}\ U(\mathbf{r})\right). \quad \text{(A5)}$$

The background potential $\int d\mathbf{r}\ U(\mathbf{r})$ can be absorbed into a redefinition of energies, which justifies the approximation made in deriving Eqs. (15).

It should also be noted that in the σ-model formalism the real part of the Green's function $G(\mathbf{r},\mathbf{r})$ is effectively set to zero under the assumption of an infinite symmetric band. Therefore any corrections to the distribution function that may arise due to $\mathcal{N}_\psi\neq 1$ are at any rate beyond the scope of the σ model.

[1] B. Shapiro, Phys. Rev. B **8**, 4394 (1986).
[2] *Mesoscopic Phenomena in Solids*, edited by B. L. Altshuler, P. A. Lee, and R. A. Webb (North-Holland, Amsterdam, 1991).
[3] F. Wegner, Z. Phys. B **36**, 209 (1980).
[4] B. L. Altshuler, V. E. Kravtsov, and I. V. Lerner, in *Mesoscopic Phenomena in Solids* (Ref. 2), and references therein.
[5] F. Wegner, Z. Phys. B **35**, 207 (1979).
[6] S. F. Edwards and P. W. Anderson, J. Phys. F **5**, 965 (1975).
[7] B. A. Muzykantskii and D. E. Khmelnitskii, Phys. Rev. B **51**, 5480 (1995).
[8] K. B. Efetov, Adv. Phys. **32**, 53 (1983).
[9] V. I. Falko and K. B. Efetov, Europhys. Lett. **32**, 627 (1995).
[10] A. D. Mirlin, Phys. Rev. B **53**, 1186 (1996).
[11] A. D. Mirlin, Pis'ma Zh. Eksp. Teor. Fiz. **62**, 583 (1995) [JETP Lett. **62**, 603 (1995)].
[12] V. I. Falko and K. B. Efetov, Phys. Rev. B **52**, 17 413 (1995).
[13] B. A. Muzykantskii and D. E. Khmelnitskii (unpublished).
[14] A. Kudrolli, V. Kidambi, and S. Sridhar, Phys. Rev. Lett. **75**, 822 (1995).
[15] B. Altshuler and V. Prigodin, Zh. Eksp. Teor. Fiz. **95**, 348 (1989) [Sov. Phys. JETP **68**, 198 (1989)].
[16] B. Altshuler and V. Prigodin, Pis'ma Zh. Eksp. Teor. Fiz. **47**, 36 (1988) [JETP Lett. **47**, 43 (1988)].
[17] B. A. Muzykantskii and D. E. Khmelnitskii, Pis'ma Zh. Eksp.

Teor. Fiz. **62**, 68 (1995) [JETP Lett. **62**, 76 (1995)].
[18] B. I. Halperin and Melvin Lax, Phys. Rev. **148**, 722 (1966).
[19] J. Zittartz and J. S. Langer, Phys. Rev. **148**, 741 (1966).
[20] In the formalism of Falko and Efetov (Ref. 9), the extra δ function in the definition of $\mathcal{P}$ was essential in enabling the mapping of the problem onto the nonlinear σ model. However, the saddle-point solution of the reduced variant of the model (Ref. 9) that describes deviations from the universal Porter-Thomas statistics [see, e. g., F. Haake, *Quantum Signatures of Chaos* (Springer-Verlag, Berlin, 1991)] involves only the variables parametrizing the bosonic sector of the theory. This can be viewed as an indication that the energy-level repulsion is not important in the approximation corresponding to the saddle-point treatment of the nonlinear σ model. In the context of the direct optimal fluctuation method, this observation receives quite a natural explanation. A given value of energy E can be fine tuned to become an eigenvalue by a uniform shift in the optimal configuration of $U(\mathbf{r})$ of the order of the mean level spacing $\Delta = 1/\nu_d V$. Such a shift can result in corrections to $\ln\mathcal{P}$ that are at most $\sim E_F\tau$, much less than the leading contribution when $\ln t$ is large.
[21] Similar configurations of the random potential were proposed by E. Levin and B. Shklovskii [Solid State Commun. **67**, 233 (1988)] as the ''traps'' responsible for exponentially small conductance of low-density electron gas in strong electric fields.
[22] In several cases the probabilities obtained by our method are *smaller* than the corresponding σ-model results. The relation between the two sets of answers and the likely reasons for the differences between them are discussed in Secs. III and IV.
[23] Y. Fyodorov and A. D. Mirlin, Int. J. Mod. Phys. B **8**, 3795 (1994).
[24] It is assumed that for the two- and three-dimensional samples the area of the contact with external leads is comparable in size with the whole surface area of the sample.
[25] E. Kamke, *Differential Equations* (Chelsea, New York, 1971).
[26] W. H. Press *et al.*, *Numerical Recipes in C* (Cambridge University Press, Cambridge, 1988).
[27] A. D. Mirlin, J. Math. Phys. (to be published).
[28] I. E. Smolyarenko (unpublished).
[29] I. M. Lifshitz, Zh. Éksp. Teor. Fiz. **44**, 1723 (1963) [Sov. Phys. JETP **17**, 1159 (1963)].
[30] B. I. Shklovskii and A. L. Efros, *Electronic Properties of Doped Semiconductors* (Springer, Berlin, 1984).