2000

Comprehensive study of Fröhlich polaron

A Mishchenko

N Prokof'ev
University of Massachusetts - Amherst, prokofev@physics.umass.edu

B Svistunov

A Sakamoto

Comprehensive study of Frohlich polaron

A.S. Mishchenko[1,2], N.V. Prokof'ev[1,3], A. Sakamoto[2], and B.V. Svistunov[1]
[1] *Russian Research Center "Kurchatov Institute", 123182 Moscow, Russia*
[2] *Department of Applied Physics, The University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan*
[3] *Department of Physics, University of Massachusetts, Amherst, MA 01003, USA*
(February 1, 2008)

A detailed study of the Fröhlich polaron model is performed on the basis of diagrammatic quantum Monte Carlo method [1]. The method is further developed both quantitatively (performance) and qualitatively (new estimators), and is enhanced by spectral analysis of the polaron Green function, within a novel approach. We present the best up to date results for the binding energy, and for the first time make available precise data for the effective mass, including the region of intermediate and strong couplings. We look at the structure of the polaron cloud and answer such questions as the average number of phonons in the cloud and their number/momentum distribution. The spectral analysis reveals non-trivial structure of the spectral density at intermediate and large coupling: the spectral continuum features pronounced peaks that we attribute to unstable excited states of the polaron.

PACS numbers: 71.38.+i, 02.70.Lq, 05.20.-y

I. INTRODUCTION

The polaron problem (for an introduction see Ref. [2]) has originally emerged in the solid state physics as a problem of electron moving in a (dielectric) medium. It became clear, however, that this problem is of essential general-physical interest, as a model of a quantum object strongly coupled to an environment. Starting from the work of Landau [3], the polaron problem has been attracting a permanent attention, serving as a testing ground of new non-perturbative methods.

The most popular model in the polaron problem is the so-called Fröhlich Hamiltonian describing an electron coupled to non-dispersive (optical) phonons of a dielectric medium via its polarization (Plank's constant and electron mass are set equal to unity) [4]:

$$H \;=\; H_{\rm e} + H_{\rm ph} + H_{\rm e\text{-}ph}\;, \tag{1.1}$$

$$H_{\rm e} \;=\; \sum_{\mathbf{k}} \frac{k^2}{2}\, a^\dagger_{\mathbf{k}} a_{\mathbf{k}}\;, \tag{1.2}$$

$$H_{\rm ph} \;=\; \sum_{\mathbf{q}} \omega_q\, b^\dagger_{\mathbf{q}} b_{\mathbf{q}}\;, \tag{1.3}$$

$$H_{\rm e\text{-}ph} \;=\; \sum_{\mathbf{k},\mathbf{q}} V(\mathbf{q})\, \left(b^\dagger_{\mathbf{q}} - b_{-\mathbf{q}}\right)\, a^\dagger_{\mathbf{k}-\mathbf{q}} a_{\mathbf{k}}\;, \tag{1.4}$$

$$V(\mathbf{q}) \;=\; i\, \left(2\sqrt{2}\alpha\pi\right)^{1/2}\, \frac{1}{q}\;. \tag{1.5}$$

In Eqs. (1.1-1.5), $a_{\mathbf{k}}$ and $b_{\mathbf{q}}$ are the annihilation operators for the electron with momentum $\mathbf{k}$ and for the phonon with momentum $\mathbf{q}$, respectively, $\omega_q \equiv \omega_0$ is the $\mathbf{q}$-independent phonon frequency, which can be set equal to unity without loss of generality, α is a dimensionless coupling constant.

Despite a lot of work addressed to the Fröhlich Hamiltonian, the model is still far from being completely understood. In the most interesting region - at intermediate and large values of α, almost all available treatments are of variational character. Hence, these treatments, even if consistent with each other, can not guarantee quantitative and qualitative reliability of the results. Moreover, some treatments are known to be in qualitative disagreement with the others. As a characteristic example, note that certain approaches suggest that the polaron states at small and large α's are

of qualitatively different nature, and there should occur a sort of phase transition in the parameter α [5–11] (see, however, discussion in Ref. [12]). The study of such important issues as electron Z-factor and the structure of the polaronic cloud was also restricted to the perturbation theory and variational treatments at small momenta [13–16]. It remained unclear what are the limits of applicability of these results, and whether they correctly describe the physics of polarons in the most important range of intermediate α.

Recently, a method of diagrammatic quantum Monte Carlo (MC) was developed, which is very efficient for the polaron-like problems [1]. The method allows direct simulation of entities specified in terms of (positive definite) diagrammatic expansions. In Ref. [1] the polaron Green function was simulated, and the results were used to extract the polaron spectrum.

In the present paper, we employ the diagrammatic Monte Carlo scheme of Ref. [1] for a detailed study of the Fröhlich model. We significantly enhanced the original scheme by (i) introducing N-phonon Green functions (with $2N$ external phonon lines), which are simulated in one and the same MC process with the ordinary (0-phonon) Green function, (ii) developing a powerful procedure of spectral analysis of the Green function. The N-phonon Green functions allow us to consider the structure of the phonon cloud and facilitate obtaining polaron parameters at large α, where the polaron is essentially a many-phonon object. In particular, direct estimators for the energy, effective mass, group velocity, and Z-factors can be constructed. The spectral analysis of the Green function gives the most complete information about the polaron, including the possibility to reveal stable and metastable excited states, if any.

The paper is organized as follows. In Section II we introduce the set of Green functions, describe the corresponding diagrammatic series, and discuss how they are related to the polaron parameters. In Section III we describe qualitatively the Monte Carlo procedure [the quantitative discussion of the updates is given in the Appendix A]. The calculated properties of the polaron (energy, effective mass, structure of the polaronic cloud, ets.) are presented in Section IV. In Section V we analyze excited states of the polaron by restoring the spectral density of the Green function by a novel method of the spectral analysis [detailed description of the method is presented in Appendix B].

II. GREEN FUNCTIONS AND DIAGRAMS

In this section we introduce basic entities and establish their relations, which will be utilized in the rest of the paper.

We start with the standard Green function of the polaron in the momentum ($\mathbf{k}$) – imaginary-time (τ) representation:

$$G(\mathbf{k},\tau) \;=\; \langle \mathrm{vac}|\, a_{\mathbf{k}}(\tau) a^{\dagger}_{\mathbf{k}}(0)\,|\mathrm{vac}\rangle \;,\;\; \tau \geq 0 \;, \tag{2.1}$$

$$a_{\mathbf{k}}(\tau) \;=\; e^{H\tau} a_{\mathbf{k}} e^{-H\tau} \;. \tag{2.2}$$

Here $|\mathrm{vac}\rangle$ is the vacuum state.

The physical information that $G(\mathbf{k},\tau)$ contains is clear from the expansion

$$G(\mathbf{k},\tau) \;=\; \sum_{\nu} |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac}\rangle|^2 \, e^{-(E_\nu(\mathbf{k})-E_0)\tau} \;, \tag{2.3}$$

where $\{|\nu\rangle\}$ is a complete set of eigenstates of the Hamiltonian H in the sector of given $\mathbf{k}$, i.e. $H\,|\nu(\mathbf{k})\rangle = E_\nu(\mathbf{k})\,|\nu(\mathbf{k})\rangle$, $H\,|\mathrm{vac}\rangle = E_0\,|\mathrm{vac}\rangle$. Since in our model $E_0 = 0$, we omit it below. Rewriting Eq. (2.3) as

$$G(\mathbf{k},\tau) \;=\; \int_0^\infty d\omega \, g_{\mathbf{k}}(\omega)\, e^{-\omega\tau} \;, \tag{2.4}$$

$$g_{\mathbf{k}}(\omega) \;=\; \sum_{\nu} \delta(\omega - E_\nu(\mathbf{k})) \, |\langle \nu | a^{\dagger}_{\mathbf{k}} | \mathrm{vac}\rangle|^2 \;, \tag{2.5}$$

one defines the spectral function $g_{\mathbf{k}}(\omega)$ which has poles (sharp peaks) at frequencies corresponding to stable (metastable) particle-like states. Hence, if at a given $\mathbf{k}$ there exists a stable polaron with the energy $E(\mathbf{k})$, the spectral function reads

$$g_{\mathbf{k}}(\omega) \;=\; Z^{(\mathbf{k})}\, \delta(\omega - E(\mathbf{k})) \;+\; \ldots \;, \tag{2.6}$$

where

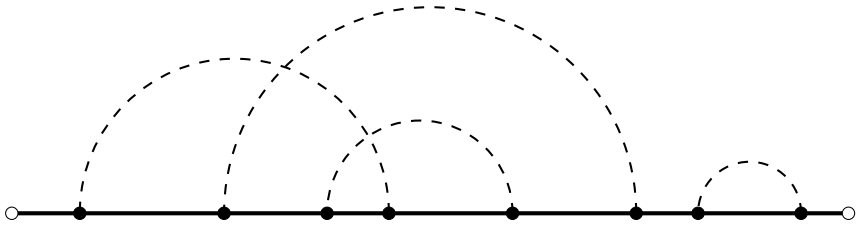

FIG. 1. A 0-phonon diagram.

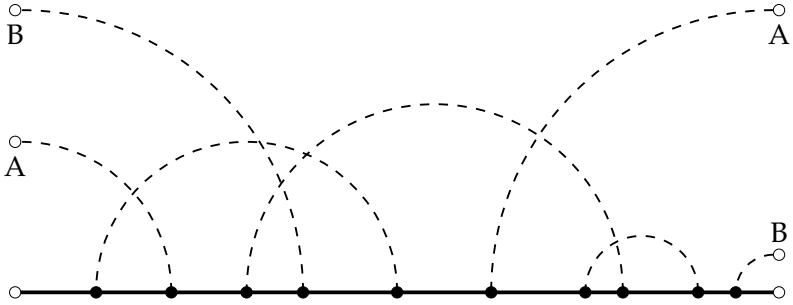

FIG. 2. A 2-phonon diagram. Note, that disconnected phonon lines always appear in pairs (labels A and B), both lines in a pair having the same momentum.

$$Z^{(\mathbf{k})} \;=\; |\,\langle\ \text{polaron}\ (\mathbf{k})\ |\ \text{free electron}\ (\mathbf{k})\ \rangle\,|^2\ . \tag{2.7}$$

Moreover, if the polaron state is the ground state, its energy and Z-factor are "projected out" by the Green function behavior at long times:

$$G(\mathbf{k}, \tau \gg \omega_0^{-1}) \;\rightarrow\; Z^{(\mathbf{k})}\, e^{-E(\mathbf{k})\tau}\ . \tag{2.8}$$

Along with the standard polaron Green function (2.1), it is reasonable to introduce the N-phonon Green function

$$G_N(\mathbf{k}, \tau;\, \mathbf{q}_1, \ldots, \mathbf{q}_N) \;=\; \langle \mathrm{vac}|\, b_{\mathbf{q}_N}(\tau) \cdots b_{\mathbf{q}_1}(\tau)\, a_{\mathbf{p}}(\tau) a^\dagger_{\mathbf{p}}(0)\, b^\dagger_{\mathbf{q}_1}(0) \cdots b^\dagger_{\mathbf{q}_N}(0) |\mathrm{vac}\rangle\ ,\ \ \mathbf{p} \;=\; \mathbf{k} - \sum_{j=1}^{N} \mathbf{q}_j\ . \tag{2.9}$$

Relations (2.3-2.8) are readily generalized to the case of N-phonon Green function. In particular, the N-phonon Z-factor for the stable (groundstate) polaron with momentum $\mathbf{k}$,

$$Z_N^{(\mathbf{k})}(\mathbf{q}_1, \ldots, \mathbf{q}_N) \;=\; |\,\langle\ \text{polaron}\ (\mathbf{k})\ |\ \text{free electron}\ (\mathbf{p})\ +\ \text{free phonons}\ (\mathbf{q}_1, \ldots, \mathbf{q}_N)\ \rangle\,|^2 \tag{2.10}$$

[the momentum $\mathbf{p}$ is defined as in Eq. (2.9)], is given by

$$G_N(\mathbf{k}, \tau \gg \omega_0^{-1};\, \mathbf{q}_1, \ldots, \mathbf{q}_N) \;\rightarrow\; Z_N^{(\mathbf{k})}(\mathbf{q}_1, \ldots, \mathbf{q}_N)\, e^{-E(\mathbf{k})\tau}\ . \tag{2.11}$$

Our MC procedure of simulating Green functions will utilize a standard diagrammatic expansion - Matsubara technique at $T = 0$. The diagrams (see Figs. 1, 2) are built of the following elements: (i) free-electron propagator (solid line)

$$G^{(0)}(\mathbf{p}, \tau_2 - \tau_1) \;=\; \exp[-\frac{p^2}{2}(\tau_2 - \tau_1)] \;, \tag{2.12}$$

(ii) phonon propagator (dashed line)

$$D(\mathbf{q}, \tau_2 - \tau_1) \;=\; \exp[-\omega_q(\tau_2 - \tau_1)] \;, \tag{2.13}$$

(iii) vertex factor $V(\mathbf{q})$ ascribed to the vertex formed by a phonon propagator (with the momentum $\mathbf{q}$) and two adjacent electron propagators. External lines of diagrams arise from the operators standing in Eqs. (2.1), (2.9) and their times and momenta are defined accordingly. Momenta of the internal lines are free up to the momentum conservation constraint at each vertex. To obey this constraint, we choose phonon momenta to be free, fixing thus the momenta of the electron lines. Times ascribed to the line ends are subject to the chronologization constraint: The time of the left end is 0, the time of the right diagram end is τ, times of the vertexes must increase along the global electron line, directed from 0 to τ. Integration over all free parameters of the diagrams is assumed, within the domains consistent with the constraints.

Upon the formulation of the diagrammatic rules, it is reasonable to introduce the irreducible N-phonon Green function, $\tilde{G}_N$, which consists only of irreducible diagrams. [By irreducible diagrams we understand those ones that *do not* contain phonon lines decoupled from the electron.] ¿From this definition it is clear that the reducible part of the N-phonon Green function, $G_N - \tilde{G}_N$, is a sum of products of irreducible functions $\tilde{G}_{N'}$, $N' < N$, and free phonon propagators. Therefore, the reducible part does not contribute to the large-τ asymptotics of G_N [because of extra factors $e^{-\omega_0 \tau}$ coming from the disconnected phonon propagators], and, in particular, Eq. (2.11) holds true for $\tilde{G}_N$ as well.

It is usefull then to consider the following function:

$$P(\mathbf{k}, \tau) \;=\; G(\mathbf{k}, \tau) + \sum_{N=1}^{\infty} \int d\mathbf{q}_1 \cdots d\mathbf{q}_N \, \tilde{G}_N(\mathbf{k}, \tau; \, \mathbf{q}_1, \ldots, \mathbf{q}_N) \tag{2.14}$$

[Note, that if $\tilde{G}_N \to G_N$, this expression would be singular at $\tau \to 0$ because of the divergence of the integrals for disconnected phonon propagators]. The function P is readily calculated by our MC procedure (see next sections) and, according to (2.8), (2.11), satisfies the relation

$$P(\mathbf{k}, \tau \gg \omega_0^{-1}) \;\to\; e^{-E(\mathbf{k})\tau} \;. \tag{2.15}$$

Here we took into account the completeness relation

$$Z_0^{(\mathbf{k})} + \sum_{N=1}^{\infty} \int d\mathbf{q}_1 \cdots d\mathbf{q}_N \, Z_N^{(\mathbf{k})}(\mathbf{q}_1, \ldots, \mathbf{q}_N) \;=\; 1 \;. \tag{2.16}$$

Eqs. (2.15), (2.16) imply that the lowest level $E(\mathbf{k})$ is non-degenerate. Otherwise one should introduce degeneracy factors to the right-hand sides of the equations.

III. DIAGRAMMATIC QUANTUM MONTE CARLO

In the previous work [1,17] it has been shown how to sum convergent (and arbitrary otherwise) diagrammatic series numerically without systematic errors. In this section we outline the basic numerical procedure of evaluating series for various Green functions described in the previous section [the details being discussed in the Appendix A], and introduce a number of estimators, that render the evaluation process significantly more efficient.

Suppose that we are interested in the function $Q(\{y\})$, which depends on a set of variables $\{y\}$, and which is given in terms of a series of integrals with an ever increasing number of integration variables:

$$Q(\{y\}) \;=\; \sum_{m=0}^{\infty} \sum_{\xi_m} \int dx_1 \cdots dx_m \, \mathcal{D}_m(\xi_m, \{y\}, x_1, \ldots, x_m) \;. \tag{3.1}$$

Here ξ_m indexes different terms/diagrams of the same order m. The term $m = 0$ is understood as a certain function of $\{y\}$. Both external $\{y\}$ and internal $\{x_i\}$ variables are allowed to be either continuous or discrete; in the latter case integrals are understood as sums. Diagrammatic MC process is a numeric procedure based on the Metropolis principle

[18] that samples various diagrams in the parameter space $(\{y\}, m, \xi_m, \{x\}_m)$ and collects statistics for $Q(\{y\})$ in such a way that the final result converges to the exact answer. The process has very much in common with the Monte Carlo simulation of a distribution given by a multi-dimensional integral. Nevertheless, there is an essential difference associated with the fact that integration multiplicity in the expansion Eq. (3.1) is varying.

Summing the series for $Q(\{y\})$ is the process of sequential stochastic generation of diagrams described by functions $\mathcal{D}_m$, Eq. (3.1). The MC process consists of a number of elementary updates falling into two qualitatively different classes: (I) those which do not change the type of the diagram (change the values of arguments in $\mathcal{D}_m$, but not the function itself), and (II) those which do change the diagram order. The set of elementary updates and their implementation is problem-specific; the only necessary requirements are ergodicity, i.e., given two arbitrary diagrams it takes finite number of updates to transform one to another, and detailed balance, i.e., diagrams contribute to the statistics according to ratio of their $\mathcal{D}$-functions. In the Appendix A we describe a set of updates which we find to work very efficiently. However, considering enormous freedom in constructing various updates, we have little doubts that they may not be further improved.

Though Green function contains complete information associated with the polaron spectrum, and the accumulation of the histogram for G is straightforward, it is reasonable to introduce certain direct estimators, including one for the Green function itself. These estimators substantially enhance the accuracy of calculations and/or allow collecting more information during one MC run (by "spreading" the data to the different values of the external parameters).

A. Estimators for effective mass, group velocity, and energy

We start with the family of estimators that are constructed in accordance with the following standard MC rule. Suppose we have some quantity A specified by the diagrammatic expansion

$$A = \sum_\nu \mathcal{D}_\nu^{(A)} , \tag{3.2}$$

where $\mathcal{D}_\nu^{(A)}$'s are the diagrams for A which are parametrized by internal variables denoted by the unified index ν, and the summation over ν is understood as the summation over discrete variables and integration over the continuous ones. Suppose next that all $\mathcal{D}_\nu^{(A)}$'s are positive definite and we have a MC process of generation of random ν with probability density given by $\mathcal{D}_\nu^{(A)}$. Then, if some quantity B is specified by a (similar) diagrammatic expansion

$$B = \sum_\nu \mathcal{D}_\nu^{(B)} , \tag{3.3}$$

the estimator for the ratio B/A is given by

$$\frac{B}{A} = \left(\sum_{\mathrm{MC}_A\{\nu\}} Q_\nu \right) \Big/ \sum_{\mathrm{MC}_A\{\nu\}} 1 , \tag{3.4}$$

$$Q_\nu = \frac{\mathcal{D}_\nu^{(B)}}{\mathcal{D}_\nu^{(A)}} . \tag{3.5}$$

Here $\mathrm{MC}_A\{\nu\}$ means the set of ν's generated during the MC run. Commonly, the quantities A and B in Eq. (3.4) are, respectively, the partition function and an observable. We, however, will use this relation in a somewhat different context. For one thing, in our case A and B can correspond to one and the same Green function, but at different values of external parameters (say, momentum or coupling constant). This way we are able to obtain results for a number of different values of the external parameters from a MC process for just one fixed set of parameters. We can also directly calculate derivatives with respect to the external parameters. To this end we should analytically take the corresponding limit from both sides of Eq. (3.4).

Let us obtain estimators for the effective mass, m_*, and group velocity, $\mathbf{v}(\mathbf{k}) = \partial E(\mathbf{k})/\partial \mathbf{k}$. First we note that $(\tau \to \infty,\ \lambda \to 0)$

$$\frac{P(\mathbf{k} + \lambda \hat{\mathbf{e}}, \tau)}{P(\mathbf{k}, \tau)} \to \begin{cases} \exp(-\lambda^2 \tau / 2m_*) , & k = 0 , \\ \exp(-\lambda \hat{\mathbf{e}} \mathbf{v}(\mathbf{k}) \tau) , & k \neq 0 , \end{cases} \tag{3.6}$$

where $\hat{\mathbf{e}}$ is a unit vector. Considering the denominator and the numerator of the left-hand side of Eq. (3.6) as A and B (respectively), we can take advantage of Eqs. (3.4) and (3.5). The function Q is given by

$$Q \,=\, \prod_j \exp\{-\frac{1}{2}[(\mathbf{p}_j + \lambda\hat{\mathbf{e}})^2 - \mathbf{p}_j^2](\Delta\tau)_j\} \,. \tag{3.7}$$

Here j numerates free-electron propagators of a given diagram for $P(\mathbf{k},\tau)$; $\mathbf{p}_j$ is the momentum corresponding to the propagator j, and $(\Delta\tau)_j$ is the length of this propagator. Eq. (3.7) immediately follows from the fact that the series for $P(\mathbf{k}+\lambda\hat{\mathbf{e}},\tau)$ can be obtained from the series for $P(\mathbf{k},\tau)$ by adding the momentum $\lambda\hat{\mathbf{e}}$ to all free-electron propagators. As we are interested only in the limit $\lambda \to 0$, we can expand Eq. (3.7) in powers of λ:

$$Q \,=\, 1 - \lambda\tau(\hat{\mathbf{e}}\overline{\mathbf{p}}) - \frac{\lambda^2}{2}\tau + \frac{\lambda^2}{2}\tau^2(\hat{\mathbf{e}}\overline{\mathbf{p}})^2 + \mathcal{O}(\lambda^3) \,, \tag{3.8}$$

where $\overline{\mathbf{p}}$ is the mean electronic momentum of the given diagram

$$\overline{\mathbf{p}} \,=\, \frac{1}{\tau}\sum_j \mathbf{p}_j\,(\Delta\tau)_j \,. \tag{3.9}$$

Comparing Eq. (3.8) with the corresponding expansions of the right-hand side of Eq. (3.6), we arrive to the following estimators

$$\langle\,\overline{\mathbf{p}}\,\rangle_{\rm MC} \;\to\; \mathbf{v}(\mathbf{k}) \qquad (\tau\to\infty) \,, \tag{3.10}$$

$$1 - \frac{\tau}{3}\,\langle\,(\overline{\mathbf{p}})^2\,\rangle_{\rm MC} \;\to\; \frac{1}{m_*} \qquad (\tau\to\infty) \,, \tag{3.11}$$

where $\langle\ldots\rangle_{\rm MC}$ means MC averaging in accordance with Eq. (3.4).

A special care should be taken for treating the time of the Green function as an external parameter of the diagrams (in the sense adopted in this section). The problem is that relations (3.4-3.5) imply that the internal parameters of diagrams $\mathcal{D}_\nu^{(A)}$ and $\mathcal{D}_\nu^{(B)}$ have one and the same domain of definition, otherwise the ratio (3.5) is not correctly defined. Meanwhile, the domain of internal times of diagrams directly depends on the external time. To circumvent this problem, one can introduce scaled internal times by simple relation $\tau_i = \tau\tilde{\tau}_i$, where τ is the external time (length of diagram in time), τ_i is an internal time variable (position in time of an electron-phonon vertex), and $\tilde{\tau}_i \in [0,1]$ is the corresponding scaled time variable with the domain of definition independent of τ.

Now it is easy to obtain a direct estimator for the polaron energy. To this end we start from the relation

$$\frac{P(\mathbf{k},(1+\lambda)\tau)}{P(\mathbf{k},\tau)} \;\to\; e^{-\lambda E(\mathbf{k})\tau} \qquad (\tau\to\infty) \tag{3.12}$$

and proceed analogously to Eqs. (3.6-3.11). In this case for the function Q we have

$$Q \,=\, (1+\lambda)^N \left(\prod_j \exp[-\lambda\frac{p_j^2}{2}(\Delta\tau)_j]\right)\left(\prod_s \exp[-\lambda\omega_0(\Delta\tau)_s]\right) \,, \tag{3.13}$$

where indexes j and s stand for the electron and phonon propagators, respectively, and N is the number of integrations over times (or, equivalently, number of interaction vertexes) in a given diagram. Then, in the limit $\lambda \to 0$, we expand the right-hand sides of Eqs. (3.12-3.13) up to terms proportional to λ, and in accordance with Eqs. (3.4-3.5) arrive to the estimator

$$\frac{1}{\tau}\left\langle \sum_j \frac{p_j^2}{2}(\Delta\tau)_j \,+\, \sum_s \omega_0(\Delta\tau)_s \,-\, N \right\rangle_{\rm MC} \;\to\; E(\mathbf{k}) \qquad (\tau\to\infty) \,. \tag{3.14}$$

B. Reweighting

We will also employ the reweighting technique [19], which allows one to utilize the statistics being generated for some given set of external variables, ξ, for calculations at a different set, ξ'. In terms of the diagrammatic Monte Carlo, this technique is based on the relation

$$\sum_{\mathrm{MC}\{\nu\}|_{\xi'}} Q_\nu(\xi') \;=\; \sum_{\mathrm{MC}\{\nu\}|_{\xi}} \frac{\mathcal{D}_\nu(\xi')}{\mathcal{D}_\nu(\xi)}\, Q_\nu(\xi')\;, \tag{3.15}$$

where Q_ν is any quantity summed over MC statistics. [We omitted superscript A at $\mathcal{D}_\nu$ since it is not relevant here.] The relation (3.15) follows from the fact that the MC statistics for the set ξ' involves the same (in the sense of structure and the values of internal parameters) diagrams as the statistics for the set ξ. The difference is only due to a different probability to generate a diagram with the set ξ rather than ξ'. This difference can be taken into account analytically by the corresponding ratio, which immediately leads to (3.15).

In our case, typical external parameters are the interaction constant and the polaron momentum: $\xi = (\alpha, \mathbf{k})$. The corresponding ratio of the diagrams is

$$\frac{\mathcal{D}_\nu(\alpha', \mathbf{k}')}{\mathcal{D}_\nu(\alpha, \mathbf{k})} \;=\; \left(\frac{\alpha'}{\alpha}\right)^{N/2} \prod_j \exp\{-[(\mathbf{k}'-\mathbf{k})^2 + 2\mathbf{p}_j(\mathbf{k}'-\mathbf{k})](\Delta\tau)_j/2\} \;=\; \left(\frac{\alpha'}{\alpha}\right)^{N/2} e^{-[(\mathbf{k}'-\mathbf{k})^2/2 + \overline{\mathbf{p}}(\mathbf{k}'-\mathbf{k})]\tau}\;. \tag{3.16}$$

This relation allows to get many points at different α' and $\mathbf{k}'$, at no extra cost in CPU time, while performing MC at a given set of α and $\mathbf{k}$ (cf. Ref. [19]).

C. Exact estimator for Green's function

Calculation of the Green function by means of histogram, though is simple and natural, involves an apparent shortcoming, associated with the finite width of the histogram cell. There is always a competition between the decreasing systematic error by making the size of the cell smaller, and the increasing statistical accuracy, which requires increasing the size of the histogram cell. This problem can be solved by introducing an *exact* (free of systematic errors) estimator for the Green function, as follows from a generic consideration presented below.

Given some function $A(\xi_0)$ of an external variable/set of variables ξ_0, specified with the (positive definite) diagrammatic expansion

$$A(\xi_0) \;=\; \sum_\nu \mathcal{D}_\nu(\xi_0) \equiv \int d\xi \sum_\nu \mathcal{D}_\nu(\xi)\delta(\xi-\xi_0)\;, \tag{3.17}$$

[considering a general case, we do not assume that the domain of definition of ν is independent of ξ] and having arranged a MC process of generating configurations $\{\nu, \xi\}$ with the probability density proportional to $\mathcal{D}_\nu(\xi)$, we would like to construct an estimator a_{ξ_0} the average of which over the MC process gives (up to a global normalization factor) the function $A(\xi_0)$. Let us look for a_{ξ_0} in the following form

$$a_{\xi_0}(\nu,\xi) \;=\; \begin{cases} q(\nu)\mathcal{D}_\nu(\xi_0)/\mathcal{D}_\nu(\xi), & \text{if } \xi \in \Gamma_0 \text{ and } \mathcal{D}_\nu(\xi) \neq 0\;, \\ 0\;, & \text{otherwise}\;. \end{cases} \tag{3.18}$$

Here Γ_0 is some finite domain in the space of variable ξ including the point ξ_0, $q(\nu)$ is some function to be defined later. [We adopt a convenient and consistent with the MC procedure convention that $\mathcal{D}_\nu(\xi) \equiv 0$, if ξ is out of the range of definition of the corresponding diagram.] ¿From Eq. (3.18) we have

$$\langle\, a_{\xi_0}\,\rangle_{\mathrm{MC}} \;\equiv\; C\sum_\nu \int d\xi\, a_{\xi_0}(\nu,\xi)\, \mathcal{D}_\nu(\xi) \;=\; C\sum_\nu q(\nu)\, \mathcal{D}_\nu(\xi_0) \int_{\xi\in\Gamma_0,\ \mathcal{D}_\nu(\xi)\neq 0} d\xi \tag{3.19}$$

where

$$C^{-1} \;=\; \sum_\nu \int d\xi\, \mathcal{D}_\nu(\xi) \tag{3.20}$$

is the normalization factor for the distribution of the random pairs (ν, ξ) induced by the series (3.17). ¿From Eq. (3.19) it is seen that if we choose

$$q^{-1}(\nu) \;=\; \int_{\xi\in\Gamma_0,\ \mathcal{D}_\nu(\xi)\neq 0} d\xi \tag{3.21}$$

[note that, according to (3.18), the definition of $q(\nu)$ is relevant only when $\mathcal{D}_\nu(\xi_0) \neq 0$, and that $q^{-1}(\nu) \neq 0$, since at least small neighborhood of the point ξ_0 contributes to the integral], then

$$\langle\, a_{\xi_0} \,\rangle_{\rm MC} \;=\; CA(\xi_0)\;. \tag{3.22}$$

The particular form of the estimator for the Green function is readily obtained by identifying ξ with τ, and noting that the ratio of diagrams standing in Eq. (3.18) is given in this case by Eq. (3.13). As the domain of definition of any diagram with respect to τ is independent of the diagram structure: $\tau \in [0, \infty]$, the factor q in Eq. (3.18) is simply proportional to the inverse size of the interval Γ_0. The choice of Γ_0 for each particular τ_0 is arbitrary, being a matter of taste and convenience.

D. Improved estimators for phonon statistics

Collecting statistics for the phonon cloud can be significantly improved by a trick described below.

We start with noting that in the case of $\tau \to \infty$, which is relevant to the ground-state properties, the set of all N-phonon diagrams possesses a certain symmetry. To reveal this symmetry, we transform diagrams to the circular representation by the rules illustrated in Fig. 3. The transformation involves "gluing" the outer ends of the electron line, as well as the outer ends of pairs of (corresponding to each other) external phonon lines. It is easy to check that the procedure is consistent with the definitions of propagators, Eqs. (2.12, 2.13). In the limit of large τ, which we are interested in, the probability to find a phonon propagator with length $> \tau/2$ is vanishingly small. That is why we may omit orientation, understanding the time length of phonon propagator as the length of the smallest arc between its ends. With the same accuracy, in the limit $\tau \to \infty$, we may ignore the pairs of phonon lines like the one shown in Fig. 3(c), which do not have unambiguous circular counterparts.

It is worth noting that circular diagrams naturally occur in a finite-temperature technique (cf., e.g., Ref. [20]), where the circumference τ has a meaning of inverse temperature.

Within the circular representation the symmetry necessary for constructing improved estimators is clearly seen. Indeed, a circular diagram represents a *whole class* of plain diagrams, due to its independence of the position of the point corresponding to the ends of the plain diagram.

Thus, having generated a certain plain diagram and associating it with the corresponding circular diagram, one effectively produces a whole class of diagrams to be included into the statistics.

In practice, the procedure is as follows. Let index $i = 1, \ldots, N_e$ label electron propagators in the circular diagram; we thus split the diagram into N_e pieces each having duration $\Delta\tau_i$ in time, $\sum_{i=1}^{N_e} \Delta_i = \tau$. When the circular diagram is cut anywhere on the interval $\Delta\tau_i$ we obtain a contribution to $\tilde{G}_{N_i}(\mathbf{k}, \tau; \mathbf{q}_1^{(i)}, \ldots, \mathbf{q}_{N_i}^{(i)})$, where N_i is the number of phonon propagators which are cut along with the electron propagator on this interval, and $\{\mathbf{q}_i\}$ are their momenta. An estimator for the integrated N-phonon Z-factor is found then to be

$$Z_N^{(\mathbf{k})} \equiv \int \ldots \int \prod_{j=1}^{N} d\mathbf{q}_j\, Z_N^{(\mathbf{k})}(\mathbf{q}_1, \ldots, \mathbf{q}_N) = \left\langle \sum_{i=1}^{N_e} \frac{\Delta\tau_i}{\tau} \delta_{N_i,N} \right\rangle_{MC}\;, \tag{3.23}$$

i.e., due to time invariance of the circular representation each interval contributes to the statistics according to its duration in time. The one-phonon distribution function within the N-phonon states manifold is given by

$$\begin{aligned} F_N^{(\mathbf{k})}(\mathbf{q}) \equiv \int \ldots \int \prod_{j=2}^{N} d\mathbf{q}_j\, Z_N^{(\mathbf{k})}(\mathbf{q}, \mathbf{q}_2, \ldots, \mathbf{q}_N) &\equiv \frac{1}{N} \sum_{l=1}^{N} \int \ldots \int \prod_{j\neq l}^{N} d\mathbf{q}_j\, Z_N^{(\mathbf{k})}(\mathbf{q}_1, \ldots, \mathbf{q}_l = \mathbf{q}, \ldots, \mathbf{q}_N) \\ &= \left\langle \sum_{i=1}^{N_e} \frac{\Delta\tau_i}{\tau}\, \delta_{N_i,N}\, \frac{1}{N_i} \sum_{j=1}^{N_i} \delta(\mathbf{q}_j - \mathbf{q}) \right\rangle_{MC}\;, \end{aligned} \tag{3.24}$$

Obviously, $Z_N^{(\mathbf{k})} = \int d\mathbf{q} F_N^{(\mathbf{k})}(\mathbf{q})$. Summing over all N we obtain the one-phonon distribution function

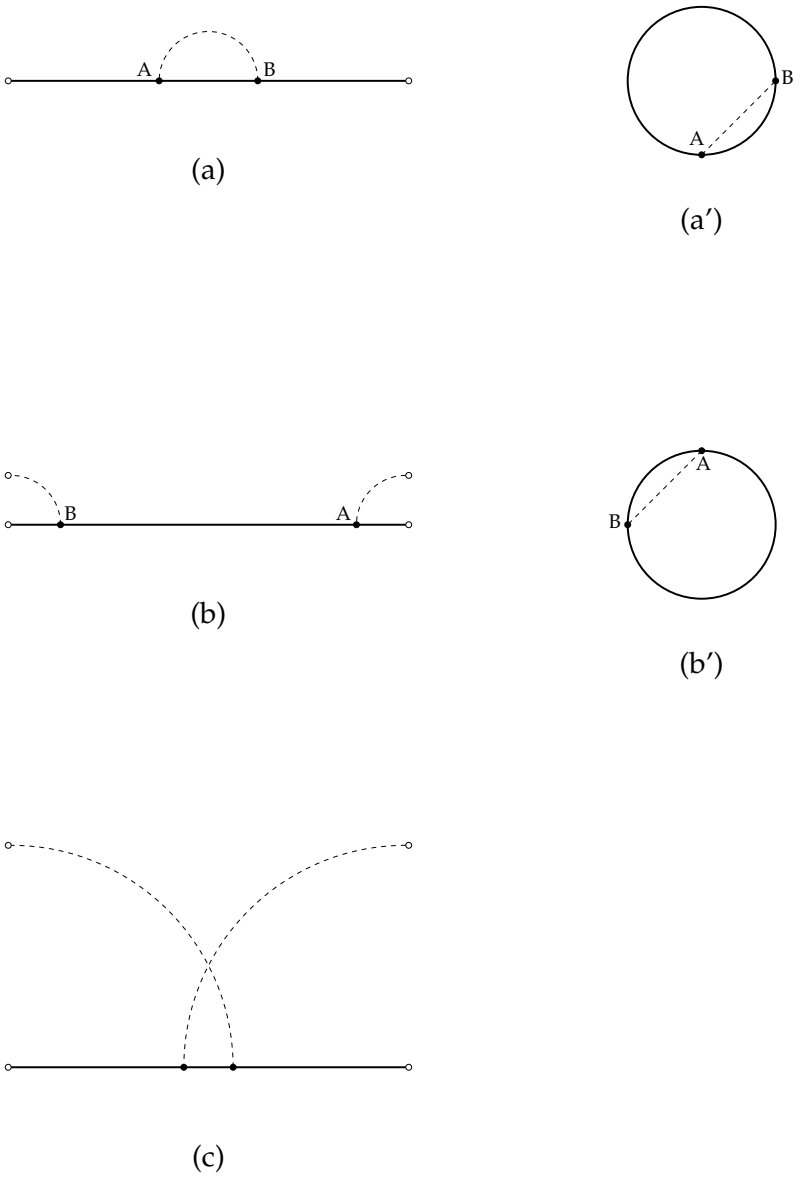

FIG. 3. Correspondence between the standard and circular representations of diagrams.

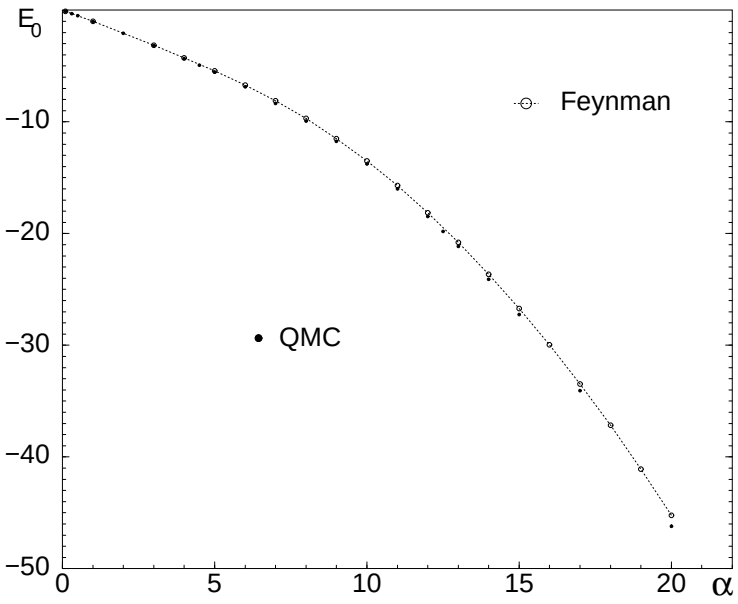

FIG. 4. Bottom of the polaron band E_0 as a function of α. The error bars are much smaller than the point size.

$$F^{(\mathbf{k})}(\mathbf{q}) = \sum_{N=1}^{\infty} F_N^{(\mathbf{k})}(\mathbf{q}) \ . \tag{3.25}$$

IV. NUMERIC RESULTS

A. Groundstate energy and effective mass

In Fig. 4 we present our results for the bottom of the band E_0 as a function of α in a wide region of coupling strengths. The data are compared with Feynman's variational treatment [13] to demonstrate the remarkable accuracy of the Feynman's approach to the polaron energy. We thus conclude that just a simple extrapolation between the second-order perturbative result $E_0 = -\alpha - 1.26(\alpha/10)^2$ and Feynman's strong-coupling variational estimate $E_0 = -\alpha^2/3\pi - 2.83$ [such an extrapolation is very close to Feynman's variational treatment in the whole range of α's] yields quite satisfactory approximation for $E_0(\alpha)$.

In Fig. 5 accurate data for the effective mass are presented up to $m_* \sim 1000$ [For larger m_*'s the Fröhlich model is not realistic anyway since the phonon dispersion becomes relevant]. At $\alpha \leq 9$ the statistics were collected at integer values of α with reweighting (in accordance with the procedure described in the previous section) to corresponding finite intervals (see the plot). At $\alpha > 9$ the reweighting procedure proved to be ineffective.

Let us compare the data for m_* with the small- and strong-coupling analytic results. At small α's, the formula $m_* = (1 - \alpha/6)^{-1}$, known to coincide with the perturbation expansion up to the second order [4,16,14], works well up to $\alpha \approx 2$. In contrast to the case of E_0, the strong-coupling limit [21] $m_* = \alpha^4/48$ drastically overestimates the effective mass in the whole range of physically interesting α's. Feynman's variational technique works better, but still with a considerable deviation (up to 50%) in the region $5 < \alpha < 10$.

Almost the same degree of accuracy gives variational treatment by Feranchuk, Fisher, and Komarov [11]. An important point about this treatment, however, is that it suggests a phase transition from "light" to "heavy" polaron at α close to 7.5, which should lead to an especially rapid increase of m_* just after $\alpha = 7.5$. In this connection, we note that our curve $m_*(\alpha)$ is essentially smooth and does not suggest any sort of phase transition or sharp cross-over. [The results of more deep numeric study of the possibility of the phase transition are presented in the next (sub)sections.]

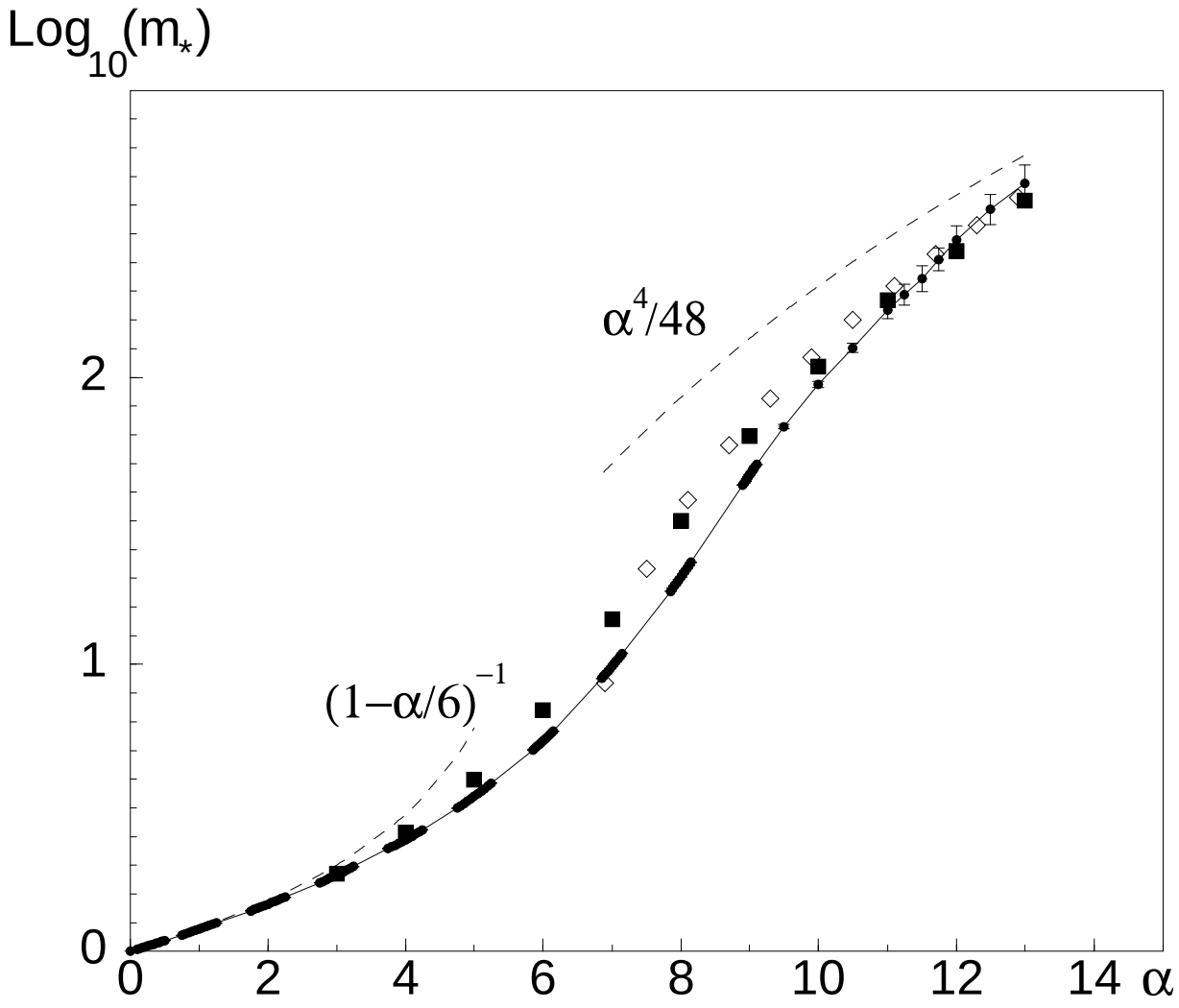

FIG. 5. Effective mass as a function of coupling parameter. Our MC data (circles interpolated by solid line) are compared with perturbation theory and strong-coupling-limit results (dashed lines), Feynman's approach (squares), and Feranchuk et. al. variational approach (dimonds).

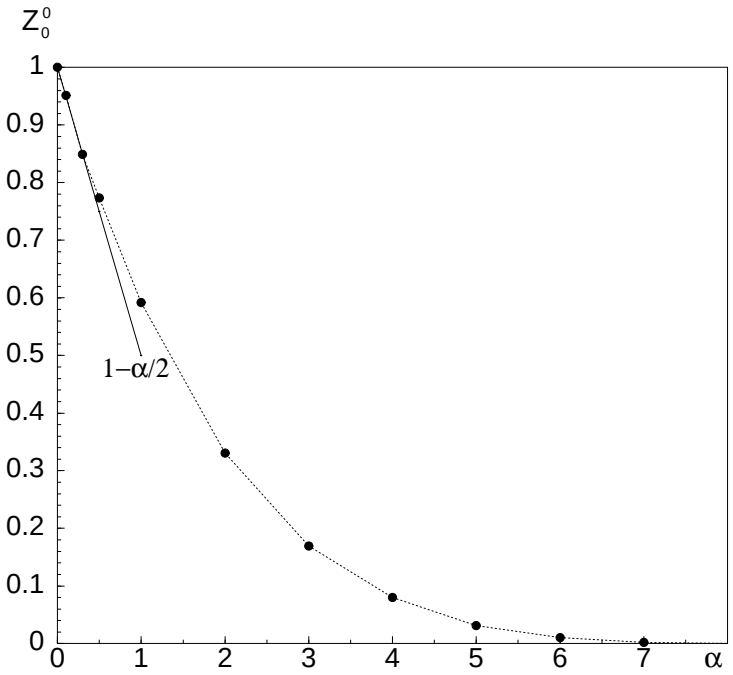

FIG. 6. The bare-electron $Z_0^{(0)}$-factor for the ground state as a function of the coupling strength; filled circles are the MC data (with error bars smaller than point size), and the solid line is the perturbation theory result (4.1).

B. Structure of polaronic cloud

In this subsection we present our results for the electron Z-factor, or $Z_0^{(\mathbf{k})}$, for $\mathbf{k} = 0$ as a function of the coupling strength in the region of small and intermediate α (for $\alpha > 10$, the bare electron weight in the polaron ground state becomes vanishingly small), and for $\alpha = 1$ as a function of momentum up to the end-point. We study also, how the distribution of phonons in the ground state, $Z_N^{(0)}$, and the average number of phonons, $\bar{N} = \sum_{N=1}^{\infty} N Z_N^{(0)}$, evolve with α. Finally, we show how the physics of the end-point is seen in the transformation of the one-particle distribution function $F_{\mathbf{k}}(\mathbf{q})$, Eq. (3.25), and how it allows to identify the relevant self-energy diagram.

For small α and $\mathbf{k} = 0$ the leading behavior is readily obtained from the perturbation theory

$$
\begin{aligned}
F_1^{(0)}(\mathbf{q}) &= \frac{\sqrt{2}\alpha}{4\pi^2} \, \frac{\sin\theta}{[q^2/2+1]^2} dq \, d\theta \, d\varphi \\
Z_0^{(0)} &= \alpha/2 \\
Z_0^{(0)} &= 1 - \alpha/2 \, ,
\end{aligned}
\tag{4.1}
$$

We have verified that for $\alpha < 1$ the perturbative results (4.1) are describing the data rather accurately [see also Figs. 6, 8, and dashed curves (connecting filled circles) in Fig. 11].

The data in Fig. 6 make it clear that perturbation theory may not be trusted for $\alpha > 1$ when the bare-electron state in the polaron wave function is no longer the dominant contribution, e.g., $Z_0^{(0)}(\alpha = 3) < 0.2$. The bare electron Z-factor vanishes rather rapidly for $\alpha > 3$ (the dependence $Z_0^0(\alpha)$ is faster than exponential) and becomes $< 10^{-5}$ for $\alpha \geq 10$. We do not attempt to fit the data to the particular functional dependence since we believe that in the interval $3 < \alpha < 10$ the polaron state undergoes a smooth transformation between weak and strong coupling limits.

In Fig. 7 we show the distribution of multi-phonon states in the polaron cloud at $k = 0$. We see how it gradually evolves from the perturbation theory case into the strong-coupling regime. For $\alpha > 10$ the data may be fit to the Gaussian distribution, but at smaller values of α the distribution is essentially asymmetric - it decays faster for $N > \bar{N}$ than for $N < \bar{N}$. We note, that even for $alpha = 17$ the phonon number distribution is rather large which means that the polaronic cloud is essentially a superposition of states with different N. The effects resulting from this fact are outside the scope of the variational Ψ^4-theory and may, for example, account for the considerable deviation of the effective mass discussed above.

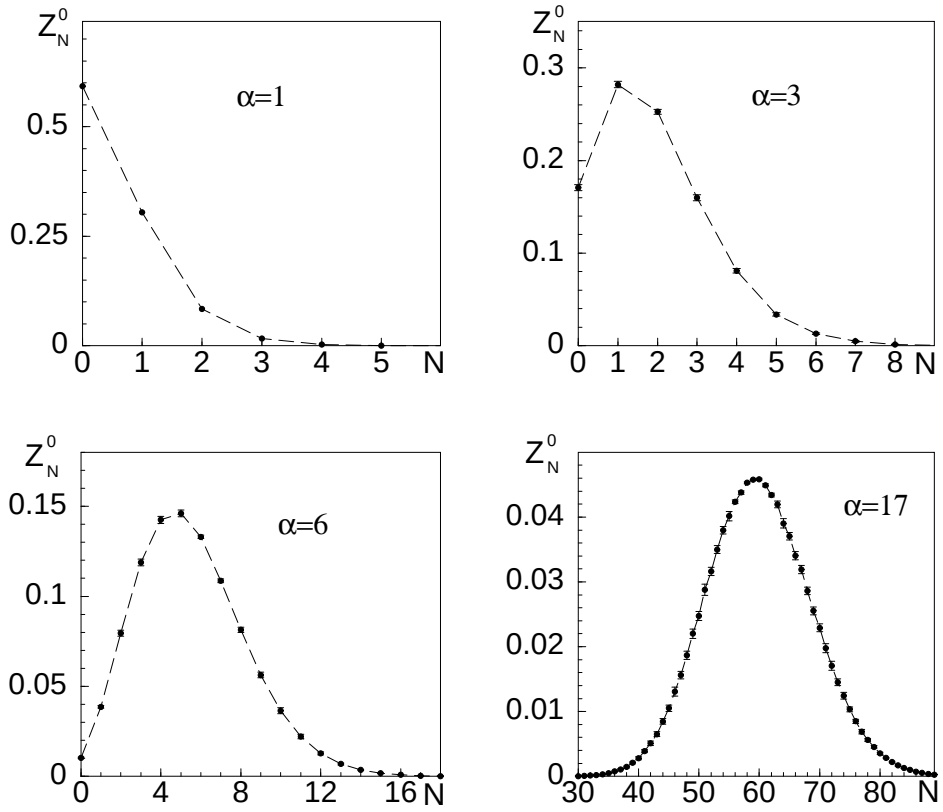

FIG. 7. Partial contributions of N-phonon states to the polaron ground state for various values of α. Error bars are shown, but are typically smaller than the point size. (The dashed lines are to guide an eye.)

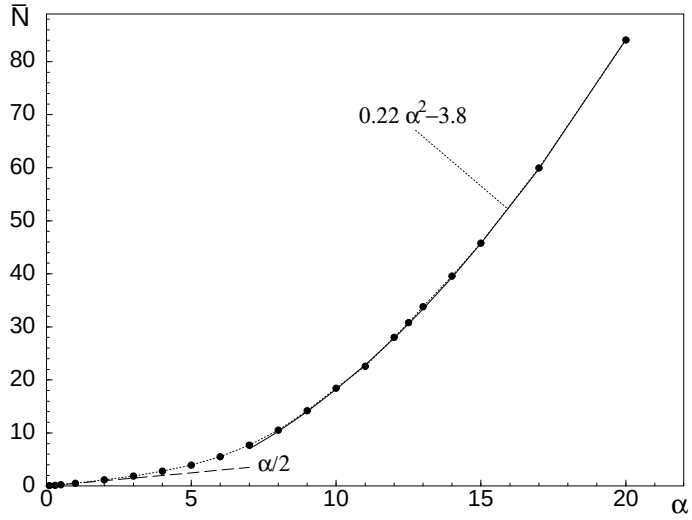

FIG. 8. The average number of phonons in the polaron ground state as a function of α. Filled circles are the MC data (error bars are smaller than the point size), the dashed line is the perturbation theory result (4.1), and the solid line is the parabolic fit for the strong coupling limit.

It is well known that polaron models often support the self-trapping phenomenon, when the ground state changes in a relatively narrow interval of parameters from light- to heavy-mass state with a sharp increase in the number of phonons contributing to the polaronic cloud. The same phenomenon was advocated for the Fröhlich model by a number of authors [5–11], including the statement that more than one stable polaron state exists in the region of intermediate α. Clearly, if there were a sharp transformation of the polaronic ground state, it would have been immediately seen in the phonon statistics. Such a transformation might be not visible on the energy plot if the hybridization matrix element between the two competing states is not small, and their energy derivatives, $dE/d\alpha$, are close to each other at the point of crossing. However, if we are to speak about *different* polaronic states, then (almost by definition) their structure has to undergo an abrupt change with α. Our data on the bare electron Z-factor Z_0^0 and phonon distribution functions Z_N^0 are evolving smoothly with α and thus prove continuous formation of the self-trapped state.

To further support this conclusion, we plot in Fig. 8 the dependence $\bar{N}$ *vs* α. The crossover between the perturbative result $\bar{N} \sim \alpha/2$ of Ref. [16] and the strong-coupling limit, where $\bar{N} \sim 0.22\alpha^2$, demonstrates no sign of the level crossing picture. As a side remark we note that the result of Ref. [16] which predicted that perturbation theory for $\bar{N}$ works well in the intermediate range $1 < \alpha \leq 6$ is not true. In fact, this law breaks down along with the perturbation theory.

Consider now the evolution of the polaronic cloud with momentum as we approach the end point. [The dispersion curve $E(k)$ featuring the end point at momentum k_c of the form $E(k \to k_c) = E_0 + \omega_0 - (k - k_c)^2/2m_c$ was calculated in Ref. [1].] Although the polaronic state is stable for $E(k) - E(0) < \omega_0$ the bare electron weight vanishes as $k \to k_c$, see Fig. 9. From this plot we estimate $k_c(\alpha = 1) \approx 1.83$. This figure also makes it clear to what degree the earlier result that $Z_0^{(\mathbf{k})}$ is momentum-independent [22] works.

With the numerical tools at hand it is possible to "visualize" the physics of the end point. In accordance with the generic Pitaevskii theory of the end point [23], in the vicinity of the end point the polaron can be considered as a (weakly) bound state of phonon, carrying almost all the momentum of the state, and a polaron with almost zero momentum. This physics is transparent from the comparison between the statistics of N-phonon states in the ground state and at $k \to k_c$ shown in Fig. 10. Evidently, the two curves can be matched by shifting the ground-state distribution by one, i.e. $Z_N^{(k_c)} \approx Z_{N+1}^0$ (the maximum of the $Z_N^{(k=1.79)}$ curve is a little depleted because of the remaining finite weight Z_0^0 since $k < k_c$). It means that the polaronic state near the end point is a superposition of bound states of the phonon with momentum around k_c and a polaron at the band bottom.

Since Fig. 10 does not tell us explicitly what are the parameters of the extra phonon present in the polaronic cloud at $k \to k_c$, we plot in Figs. 11 12 normalized distribution functions of phonon momenta (in $|q|$ and in the angle

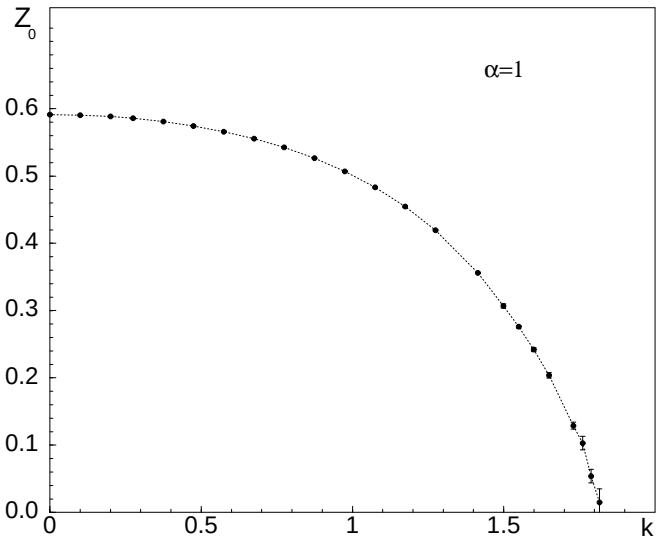

FIG. 9. The bare-electron Z-factor as a function of the polaron momentum up to the end point.

between $\mathbf{q}\hat{\mathbf{k}}$). It is obvious from these figures that the extra phonon momentum is concentrated around $\mathbf{k}$.

V. SPECTRAL ANALYSIS

The spectral function $g_{\mathbf{k}}(\omega)$ (2.5) provides important information about the system since it has poles (sharp peaks) at frequencies corresponding to stable (metastable) particle-like states. Besides, since the probability of absorption of a free electron with the momentum $\mathbf{k}$ into a polaron state is proportional to the spectral function, the latter can be measured experimentally by angle-resolved inverse photoemission spectroscopy.

In spite of many elaborated treatments of the properties of the polaron, the knowledge about high-energy part of the polaron spectrum is mostly limited by attempts to calculate the spectral density either by perturbation theory approaches or at strong coupling limit [24]. As both the Green function asymptotic behavior and the machinery of estimators provides information about ground state properties only, the spectral density is indispensable for the study of excited states of the system.

The spectral analysis, i.e. solving the equation (2.4), was performed by a novel method [detailed description of the method and testing examples are presented in Appendix II]. The most important features of the method are that it avoids distortion of equation by nonlinear terms and does not suffer from systematic errors caused by preassigned discretisation of the ω-space.

To perform a joint check of the diagrammatic Monte Carlo approach and the method of spectral analysis, we compared the spectral densities obtained by our numeric calculations and by perturbation theory for zero temperature. The analytic expression for the high-energy part ($\omega > 0$) of the spectral density could be obtained for the arbitrary interaction potential $V(\mid \mathbf{q} \mid)$ which depends on the modulus $\mid \mathbf{q} \mid$ of the phonon momentum. For zero polaron momentum, $\mathbf{k} = 0$, the the imaginary part of the linear in α self-energy part $\Sigma(0, \omega > 0)$ is

$$\mathrm{Im}\Sigma(0,\omega) = -\frac{1}{\sqrt{2}\pi}\sqrt{\omega - 1}\mid V(\sqrt{2(\omega-1)})\mid^2 \theta(\omega - 1). \tag{5.1}$$

(Here θ is the theta-function.) Then, using the relation $g_{\mathbf{0}}(\omega) = -\mathrm{Im}G_{\mathbf{k}=0}(\omega)/\pi$ and keeping only linear with respect to α terms one gets

$$g_{\mathbf{0}}(\omega > 0) = \frac{1}{\sqrt{2}\pi^2}\frac{\sqrt{\omega-1}}{\omega^2}\mid V(\sqrt{2(\omega-1)})\mid^2 \theta(\omega - 1). \tag{5.2}$$

The expression for the low-energy part ($\omega < 0$) of the spectral density depends on the specific form of the interaction potential and we consider the perturbation theory result for the short-range interaction

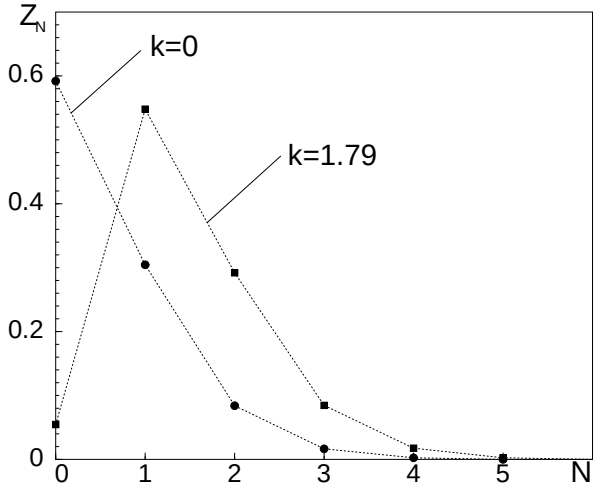

FIG. 10. Partial weights Z_N of the N-phonon states in the structure of the polaronic state for $\alpha = 1$ at the band bottom and near the end point. (The dashed lines are to guide an eye.)

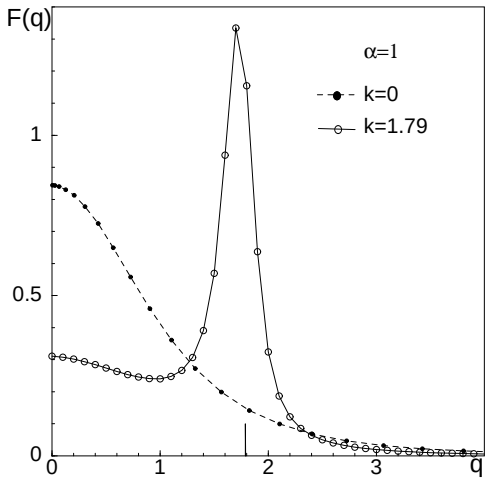

FIG. 11. Phonon distribution functions in q-modulus for the ground state (filled circles) and close to the end-point for $k = 1.79$ (open circles). The momentum k_c is indicated by a bar at the q-axis. (The lines are just linear interpolations between accurate numeric points.)

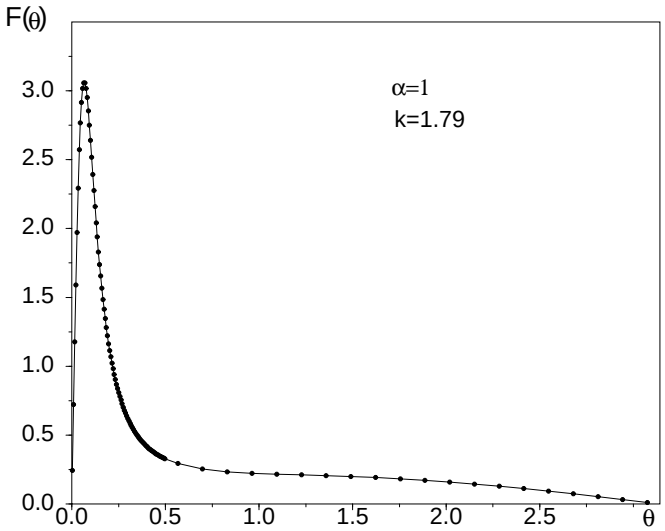

FIG. 12. Phonon distribution functions in the angle between the vectors $\mathbf{q}$ and $\mathbf{k}$ close to the end-point for $k = 1.79$.

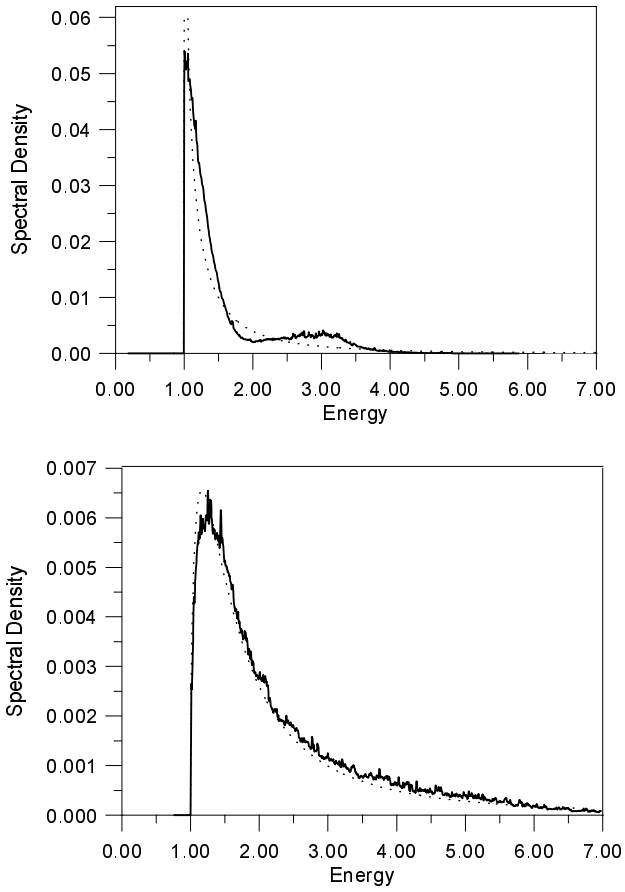

FIG. 13. The comparison of the numeric results (solid lines) and the perturbation theory curves (dashed lines) for the spectral density of Frohlich model with $\alpha = 0.05$ (upper panel) and the short-range interaction model with $\alpha = 0.05$ and $\kappa = 1$ (lower panel).

$$V(\mid \mathbf{q} \mid) \;=\; i\,\left(2\sqrt{2}\alpha\pi\right)^{1/2}\,\frac{1}{\sqrt{q^2+\kappa^2}}\;, \tag{5.3}$$

which reduces to the Frohlich one when $\kappa \to 0$. The low-energy part is the delta-functional peak

$$g_{\mathbf{0}}(\omega<0) = \frac{\alpha}{(\kappa+\sqrt{2})^2}\delta\left(\omega+\alpha\frac{\sqrt{2}}{\kappa+\sqrt{2}}\right). \tag{5.4}$$

The comparison of our numeric results for the low-energy part of the Frohlich polaron ($\kappa = 0$)spectral density for $\alpha = 0.05$ and Eq. (5.4) demonstrates a perfect agreement (with the accuracy 10^{-4} for the polaron energy and Z-factor), whereas our results for the high-energy part (upper panel in Fig. 13) significantly deviate from the analytic curve. This is not surprising since for Frohlich polaron the perturbation theory expression is diverging as $\omega \to \omega_0$ and, therefore the perturbation theory breaks down. To test the case when perturbation theory is obviously valid we set $\kappa = 1$ and obtained a perfect agreement for both the low- and high-energy parts of $g(\omega)$ (lower panel in Fig. 13). We note that the high-energy part of $g(\omega)$ is successfully restored by our method despite the fact that the total weight of the feature is less than 10^{-2} for $\alpha = 0.05$.

One can note that the main deviation of the actual spectrum of Frohlich polaron from the perturbation theory result is the extra broad peak in the actual spectral density at $\omega \sim 3.5$. To study this feature we calculated $g(\omega)$ for $\alpha = 0.5$, $\alpha = 1$, and $\alpha = 2$ (see Fig. 14). Note, that the peak is seen for higher values of the interaction constant and its weight grows with α. Near the threshold, $\omega = 1$, the spectral density demonstrates the square-root dependence $\sim \sqrt{\omega - 1}$ (see the insert).

To trace the evolution of the peak at higher values of α we calculated the spectral density for $\alpha = 4$, $\alpha = 6$, and $\alpha = 8$ (see Fig. 15). At $\alpha = 4$ the peak at $\omega \sim 3.5$ already dominates in the spectral density. Moreover, a distinct high-energy shoulder appears at $\alpha = 4$, which transforms into a broad peak at $\omega \sim 9$ in the spectral density for $\alpha = 6$. The spectral density for $\alpha = 8$ demonstrates further redistribution of the spectral weight between different maxima without significant shift of the peak positions. One can also see that there is a high-energy shoulder which is, probably, the precursor of another peak which would appear for higher values of the interaction constant.

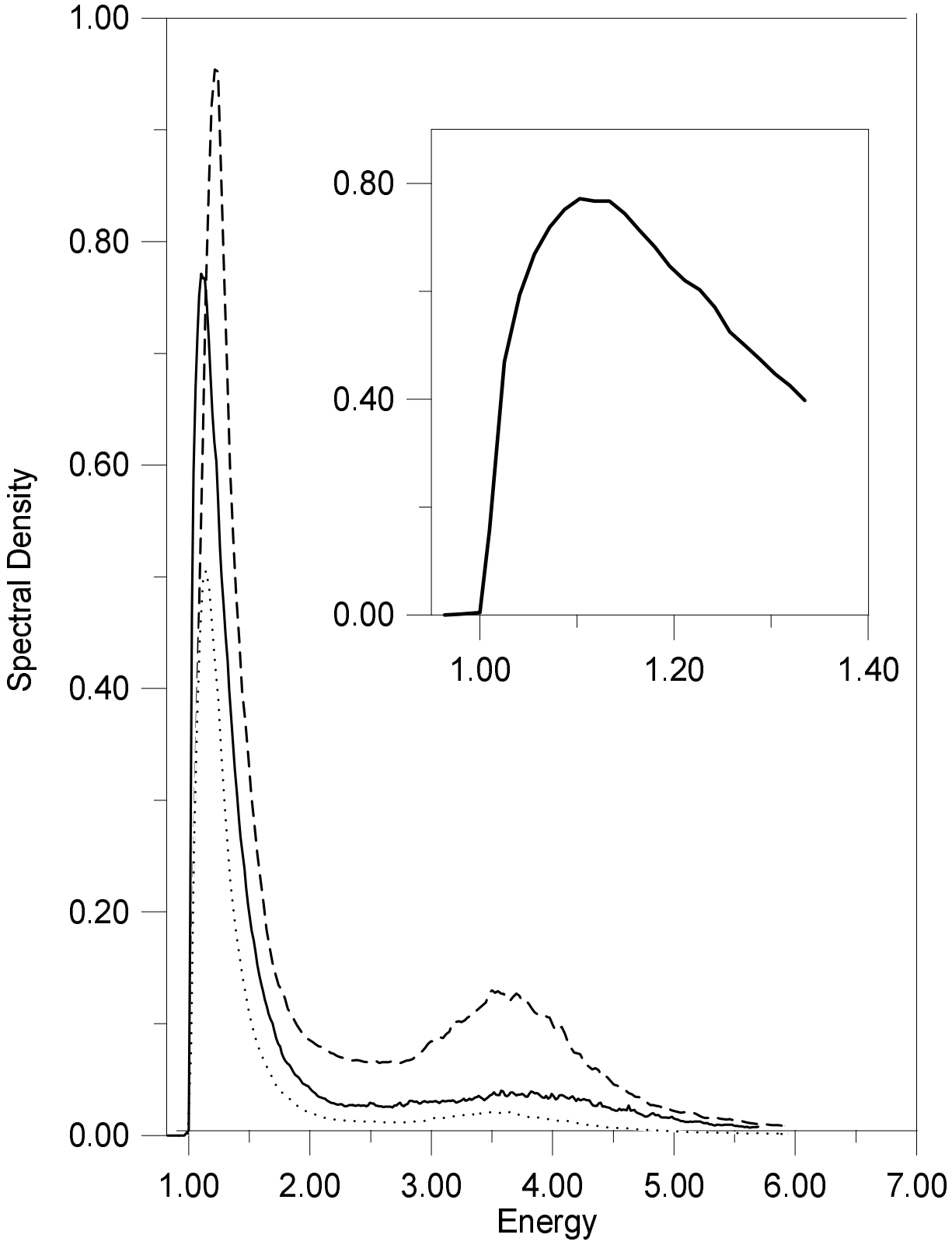

FIG. 14. The spectral density of Frohlich polaron for $\alpha = 0.5$ (dotted line), $\alpha = 1$ (solid line) and $\alpha = 2$ (dashed line), with energy counted from the position of the polaron. The initial fragment of the spectral density for $\alpha = 1$ is shown in the insert.

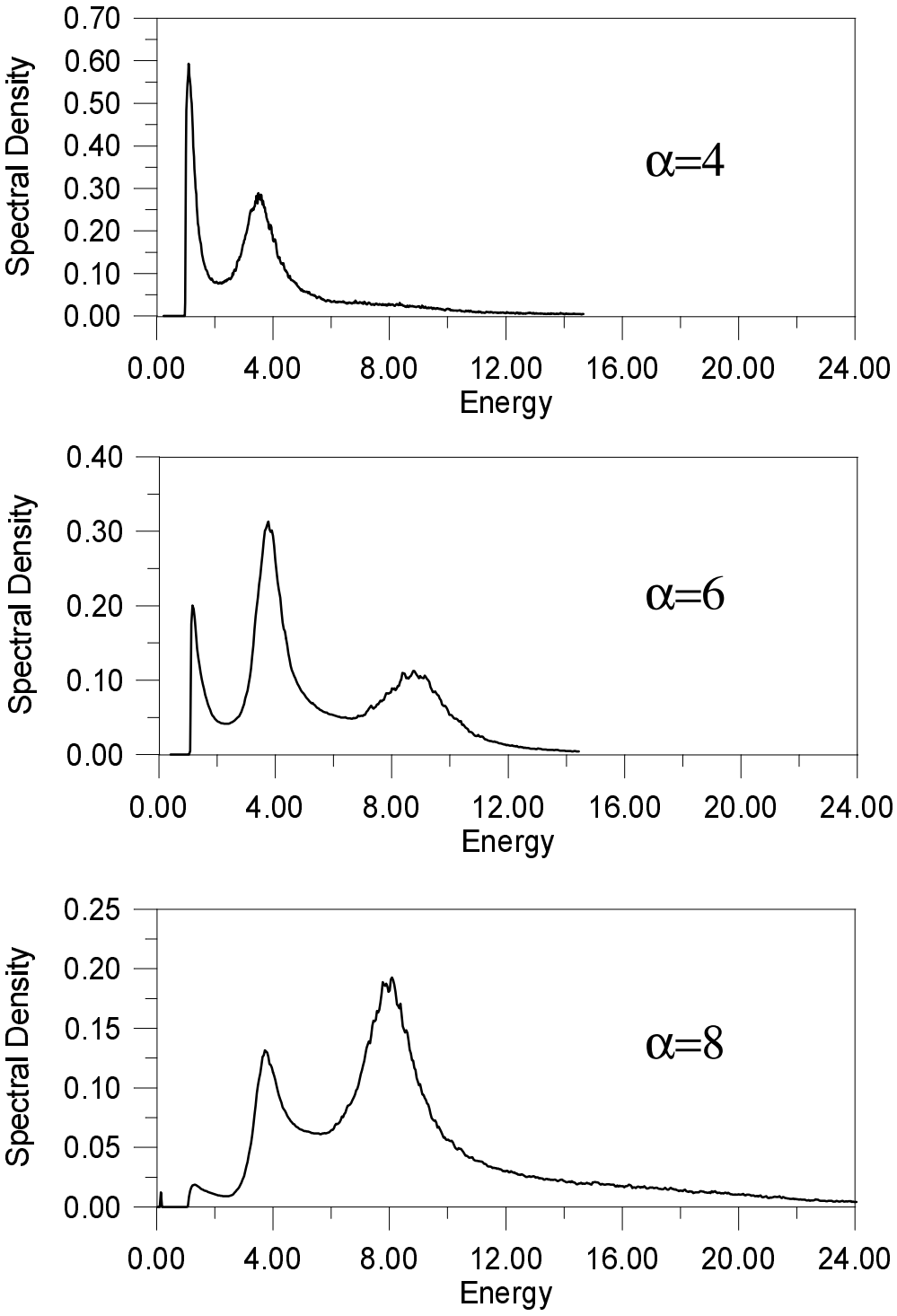

FIG. 15. Evolution of spectral density with α in the cross-over region from intermediate to strong couplings. (The polaron ground state peak is shown only for $\alpha = 8$. Note, that the spectral analysis still resolves it, despite its very small weight $< 10^{-3}$.) The energy is counted from the position of the polaron.

The excited states of the polaron were studied within the frameworks of different approaches [25–29] by calculating optic absorption spectra. The light absorption is associated with the transitions from the polaron ground state ϕ_0 with $\mathbf{k} = 0$ and $E = E_0$ to the excited states f with E_f which are characterized by the presence of a finite number of real phonons along with the polaron. The optic absorption spectrum at the frequency ω is proportional to the transition probability

$$P(\omega) = 2\pi \sum \langle \phi_0 \mid \hat{O} \mid f \rangle \langle f \mid \hat{O} \mid \phi_0 \rangle \delta(E_0 - E_f + \omega). \tag{5.5}$$

(Here $\hat{O} = \mathbf{Er}$ is the electric dipole interaction, $\mathbf{E}$ is the electric field.)

It was shown in the weak-coupling limit [25,26] that the optic absorption spectrum has a broad peak with the onset separated from the polaron state by the optic phonon energy. Our calculations confirm (see Fig. 14) that there are no metastable excited states of the polaron in the weak-coupling regime.

On the other hand, in the strong-coupling limit the existence of the metastable relaxed excited state (RES), i.e. the state where the lattice readapts to the new electronic configuration and the polaron-lattice system is in the local minimum of the total energy, was predicted [27–29]. This state manifests itself as a sharp peak in the absorption spectrum which is located at the frequency equal to the energy difference $E_{\mathrm{RES}} - E_0$. To check the existence of RES one can study the spectral density (2.5) since although the matrix elements of transition probability (5.5) and spectral density (2.5) are different, both functions have to demonstrate sharp peaks at the energies of the metastable excited states. From Fig. 15 we conclude that there is no metastable excited state because the width of the peaks is comparable with the excitation energy, i.e. with the distance from the polaron ground state. Moreover, according to the strong-coupling approaches [27], the excitation energy of the RES state is proportional to α^2, whereas peak positions in $g(\omega)$ with respect to E_0 do not change with α.

The variational treatment developed in Ref. [11] suggests that in a certain region of α there may exist two different *stable* states of the polaron [the corresponding equations for variational parameters have two solutions]. Our numeric study can shed light on this situation.

To start with, let us discuss what one could observe would the two states really exist. If at some point $\alpha = \alpha_*$ there occurs a level crossing so that the ground state switches from one state to another, and the two states differ essentially in the number of phonons and/or in the effective mass, one would expect at the point α_* a sharp change of these quantities. The change should be almost jump-like in the case of small hybridization between the two states, and look like a smooth cross-over otherwise. Even in the case of sufficiently strong hybridization, one may distinguish between two qualitatively different cases: (i) the case when the level separation is less than ω_0 [and thus both states are stable against decay], and (ii) the case when the upper level is in the continuum, and therefore is unstable.

Strictly speaking, in the case (ii) one can invoke the second level only in some quantitative sense, since there is no qualitative difference between the case (ii) and a situation with only one polaron state. These quantitative features could be associated with the non-monotonic behavior of the derivatives (with respect to α) of the effective mass and/or the mean number of phonons, and, of course, another peak in the spectral density.

Our study of the spectral densities shows that the case (i) is not realized because we found that $g(0 < \omega < 1) = 0$. Therefore, there is no excited stable state in the energy gap between the ground state energy and incoherent continuum. Instead, there are several many-phonon unstable states at energies $E_f - E_0 \sim 1$, ~ 3.5, and ~ 8.5. One can speculate that these states reveal themselves in variational approaches and can be mistreated as quasi-stable states of the polaron. It should be emphasized, however, that the situation does not resemble that of the level-crossing at all, since we do not observe non-monotonic behavior of the derivatives (with respect to α) of the effective mass and/or the mean number of phonons.

VI. ACKNOWLEDGMENTS

We are indebted to N. Nagaosa for inspiring discussions and valuable remarks. This work was supported by the Russian Foundation for Basic Research (Project 99-02-17288) and by Priority Areas Grants and Grant-in-Aid for COE research from the Ministry of Education, Science, Culture and Sports of Japan.

APPENDIX A: UPDATING PROCEDURES

1. Updates of class I

Updating procedures of this class are the simplest. They mimic standard rules of simulating a given distribution function $\mathcal{D}_m$. In the present case we are dealing with quite a number of variables having different physical meaning: external variables $\{y\}$ include τ, N, α, and k, and internal variables describe the topology of the diagram (index ξ_m), times of electron-phonon vertices and momenta of phonon propagators. From this list of variables follows a set of updates simulating multi-dimensional distribution $\mathcal{D}_m$.

a. Vertex shift in time

We choose at random any interaction vertex inside the graph (we exclude the diagram closing points which are updated separately), and change its time position from τ_v to τ_v' on the interval (τ_1, τ_2) between the nearest left- and right-neighbor vertices, i.e., $\tau_1 < \tau_v,\ \tau_v' < \tau_2$. Let the incoming and outgoing electron momenta for the selected vertex are $\mathbf{p}$ and $\mathbf{p}+\mathbf{q}$. The normalized probability density to find the vertex at time τ_v' is a simple exponential function

$$W(\tau_v') = \frac{\Delta E\, e^{-(\tau_v'-\tau_1)\Delta E}}{1-e^{-(\tau_2-\tau_1)\Delta E}}\ , \tag{A1}$$

where $\Delta E = E(\mathbf{p}) - E(\mathbf{p}+\mathbf{q}) \mp \omega_p$ depending on whether the updated vertex is the left or right end of the corresponding phonon propagator, which allows a trivial solution of the equation

$$\int_{\tau_1}^{\tau_v'} W(s)ds = r\ , \tag{A2}$$

in the form

$$\tau_v' = \tau_1 - \frac{\ln\left(1 - r[1-e^{-(\tau_2-\tau_1)\Delta E}]\right)}{\Delta E}\ . \tag{A3}$$

Here and below r is the random number homogeneously distributed on the unit interval. Since the new variable is selected according to the exact probability density the acceptance ratio for this update is unity.

b. Change of transferred momentum angle

We choose at random any phonon propagator except those attached to the diagram ends (propagators attached to the diagram ends appear in *pairs* with equal momenta, thus single propagator updates do not apply to them) and change its momentum $\mathbf{q} \to \mathbf{q}'$ so that $|\mathbf{q}| = |\mathbf{q}'|$. Let the propagator connects vertices at times τ_1 and τ_2. Evaluating the average electron momentum between these vertices

$$< \mathbf{p} >_{\tau_1,\tau_2} = \frac{\int_{\tau_1}^{\tau_2} \mathbf{p}(\tau)d\tau}{\tau_2-\tau_1}\ , \tag{A4}$$

and introducing vector $\mathbf{p}_0 =< \mathbf{p} >_{\tau_1,\tau_2} + \mathbf{q}$, we may write the probability density to find azimuthal and polar angles φ, θ between vectors $\mathbf{q}$ and $\mathbf{p}_0$ as

$$W(\varphi,\theta) \sim \sin(\theta)\exp\left\{-\frac{\tau_2-\tau_1}{m}\, p_0 q\cos\theta\right\}\ , \tag{A5}$$

This result is a trivial consequence of the quadratic dispersion law for the bare electron spectrum. Clearly, the new azimuthal angle is selected at random ($\varphi = 2\pi r$), and $\cos\theta$ is selected according to the simple exponential function in complete analogy with Eqs. (A1) and (A3) up to trivial change of notations. The acceptance ratio is thus unity.

c. Change of transferred momentum modulus

In this procedure propagators are selected as explained in the previous subsection, but now we change the modulus of the transferred momentum while keeping the polar and azimuthal angles between the vectors $\mathbf{p}_0$ and $\mathbf{q}$ fixed. The probability density now reads

$$W(q) \sim V(q)q^2 \exp\left\{-\frac{\tau_2-\tau_1}{2m}\left[q-p_0\cos\theta\right]^2\right\} \sim \exp\left\{-\frac{\tau_2-\tau_1}{2m}\left[q-p_0\cos\theta\right]^2\right\} , \tag{A6}$$

where we have used explicitly the property of the Fröhlich model that $V(q)q^2$ is q-independent. By tabulating the inverse error-function we ensure fast numerical solution of the equation $\mathrm{erf}(z) = r$, or $z = \mathrm{erf}^{-1}(r)$, and thus generation of the new value $q = z\sqrt{2m/(\tau_2-\tau_1)} + p_0\cos\theta$ with acceptance ratio unity.

d. Change of diagram structure

We select at random any nearest-neighbor pair of vertices inside the graph (again, the diagram closing vertices are excluded) and exchange the assignment of the phonon propagators between these vertices. Namely, if the original momentum transfer was $\mathbf{q}_1$ in vertex 1 and $\mathbf{q}_2$ in vertex 2, we suggest to change these momenta to $\mathbf{q}_2$ and $\mathbf{q}_1$ correspondingly. The acceptance ratio for this procedure depends on whether we are dealing with the left ($c = 1$) or right ($c = -1$) ends of the phonon propagators

$$R = e^{-\tau[E(\mathbf{p}+c_1\mathbf{q_1}-c_2\mathbf{q_2})-E(\mathbf{p})-\omega_0(c_1-c_2)]} , \tag{A7}$$

where τ is the time difference, and $\mathbf{p}$ is the electron momentum between the selected vertices. Clearly, this procedure effectively changes the topology of bosonic lines while keeping fixed their momenta.

e. Change of diagram length in time

This procedure is done in two variants (almost identical to the procedure of shifting the vertex position in time). Consider the case when no artificial potential except the chemical potential is used. In the first variant we select the new time difference τ between the positions of the right diagram end at its left nearest neighbor vertex according to the probability density

$$W(\tau') = \Delta E\, e^{-\tau'\Delta E} , \tag{A8}$$

where $\Delta E = E(\mathbf{p}) + N_z\omega_0 - \mu$, and $\mathbf{p}$ is the momentum of the last electron propagator, μ is the chemical potential, and N_z is the number of phonon propagators attached to the diagram right end (obviously, N_z is the same for the diagram left end). In the second variant we select new time differences between the positions of all nearest-neighbor pairs of vertices. For each such a pair the probability density is still given by Eq. (A8), where in the most general case N_z must be understood as the number of phonon propagators which are cut when the diagram is cut anywhere between the selected pair of vertices. Notice, that the second variant requires much longer computation time; thus if the typical diagram order is very large it must be applied less frequently. In both variants the acceptance ratio is unity.

There is a bottle-neck in the time decay of the electron P-function, which does not allow efficient sampling of both long-time and short-time behavior and causes normalization problems at large α, namely, $P(\tau)$ drops to almost zero vaalue at short times, and then climbs back to $P \sim 1$ before it settles to the asymptotic decay (2.15).

There is however a general prescription of how to eliminate such difficulties by using the so-called "guiding function" [30] or fictitious potential renormalization. This method was successfully applied recently [31] to the problem of tunneling transition amplitudes, where one is bound to collect reliable statistics which varies by orders and orders of magnitude between different points in time. The idea is to modify the statistics of suggested diagrams by introducing the acceptance ratio

$$R = A_{\mathrm{fic}}(\tau_{\mathrm{new}})/A_{\mathrm{fic}}(\tau_{\mathrm{old}}) , \tag{A9}$$

and accordingly multiplying all MC estimators in the time domain by $1/A_{\mathrm{fic}}(\tau)$, where the fictitious potential $A_{\mathrm{fic}}(\tau)$ is arbitrary. Note, that in Eq. (A9) we are dealing with the external variable τ - the diagram length in time. In

the present case the best choice would be $A_{\rm fic} \sim 1/P(\tau)$. We achieve this goal by self-consistently adjusting $A_{\rm fic}$ to $1/P_{MC}(\tau)$ after a certain large number of updates during the thermalization stage (here $P_{MC}(\tau)$ is the statistical result for $P(\tau)$). After thermalization stage we start collecting new statistics for $P(\tau)$ and keep $A_{\rm fic}$ fixed.

f. Change of coupling constant

Since the diagram weight depends on the coupling constant as α^{N_p} where N_p is the number of phonon propagators in the diagram, all we need to do is to select new value of α with this power-law probability density. Normalized probability density is obtained by restricting allowed values of α to a certain parameter range. Acceptance ratio is unity.

g. Change of external momentum

Given the average electron momentum of the diagram $\bar{\mathbf{p}} =< \mathbf{p} >_{0,\tau}$, see Eq. (A4), with external momentum $\mathbf{k}$, we define vector $\mathbf{p}_0 = \mathbf{k} - < \mathbf{p} >_{0,\tau}$ and write the probability density to select new external momentum $\mathbf{k}'$ as

$$W(\mathbf{k}') \sim \exp\left\{-\frac{\tau}{2m}(\mathbf{k}' - \mathbf{p}_0)^2\right\} \ . \tag{A10}$$

As before the new variable is seeded according to this probability density utilizing the tabulated error-function [see subsection A 1 c and Eq. (A6)] and thus is always accepted.

One note is in order here. Although one is allowed to change the coupling constant and the external momentum in a single MC process, it seems more efficient to keep these variables fixed instead of spreading the statistics over some range in the $(\mathbf{k}, \alpha)$ parameter space. However, the knowledge of the relative weights according to which a given diagram contributes to the statistics of various α and $\mathbf{k}$ may be utilized in collecting statistics for the finite neighborhood of the point $(\mathbf{k}_0, \alpha_0)$ used in a given MC simulation. Obviously, reliable results for points other than $(\mathbf{k}_0, \alpha_0)$ are obtained only provided that for typical diagrams the relative weights are of order unity. As explained in the text, this knowledge is also used in deriving estimators for the effective mass and group velocity of the polaron.

2. Updates of class II

These updates are in the heart of the method since they change the diagram order. The generic rules for constructing them are as follows [17]. Let the update $\mathcal{A}$ transform a diagram $\mathcal{D}_m(\xi_m, y, x_1, \ldots, x_m)$ into $\mathcal{D}_{m+n}(\xi_{m+n}, y, x_1, \ldots, x_m, x_{m+1}, \ldots, x_{m+n})$, and, correspondingly, its counterpart $\mathcal{B}$ perform the inverse transformation. For n new variables we introduce vector notation: $\vec{x} = \{x_{m+1}, x_{m+2}, \ldots, x_{m+n}\}$. The update $\mathcal{A}$ involves two steps. First, it *proposes* a change, selecting a new diagram, $\mathcal{D}_{m+n}$, and a particular value of $\vec{x}$, which is seeded with a certain normalized distribution function $W(\vec{x})$. There are no requirements strictly fixing the form of $W(\vec{x})$, but to render the algorithm most efficient, it is desirable that $W(\vec{x})$ be chosen as close as possible to $\mathcal{D}_{m+n}(\mathbf{x})$, i.e., to the actual statistical probability density of $\vec{x}$ in the new diagram. Upon proposing the modification, the update is accepted, with probability, $P_{\rm acc}(\vec{x})$, or rejected. The update $\mathcal{B}$, removing variable $\mathbf{x}$, is accepted with probability $P_{\rm rem}(\vec{x})$). For the pair of complementary updates to be balanced, the following Metropolis-like prescription should be fulfilled [17]:

$$P_{\rm acc}(\vec{x}) \ = \ \begin{cases} R(\vec{x})/W(\vec{x}), & \text{if } R(\vec{x}) < W(\vec{x}) \ , \\ 1 \ , & \text{otherwise} \quad , \end{cases} \tag{A11}$$

$$P_{\rm rem}(\vec{x}) \ = \ \begin{cases} W(\vec{x})/R(\vec{x}), & \text{if } R(\vec{x}) > W(\vec{x}) \ , \\ 1 \ , & \text{otherwise} \quad , \end{cases} \tag{A12}$$

where

$$R(\vec{x}) \ = \ \frac{p_{\mathcal{B}}}{p_{\mathcal{A}}} \, \frac{\mathcal{D}_{m+n}(\xi_{m+n}, y, x_1, \ldots, x_m, \vec{x})}{\mathcal{D}_m(\xi_m, y, x_1, \ldots, x_m)} \tag{A13}$$

and $p_{\mathcal{A}}$ and $p_{\mathcal{B}}$ are the probabilities of selecting updates $\mathcal{A}$ and $\mathcal{B}$, which, in principle, may differ. To solve the polaron problem and account for any possible diagram it is sufficient to have two pairs of complementary processes of type II which are described in detail below.

a. Adding/removing phonon propagators to the diagram

Consider the algorithm for the process increasing the number of internal phonon propagators (i.e., excluding those attached to the diagram closing points) by one. This update is done in two variants which differ in the probability densities according to which the new propagator parameters are suggested. First we select the time position τ_1 for the left-hand end of the extra phonon propagator. This is done by choosing at random (with equal probabilities) one of the free-electron propagators, and by taking for τ_1 any time (with equal probability density) within this propagator. Then we select the transferred momentum and propagator length in time using the distribution function

$$W(\mathbf{q},\tau) = \frac{\omega_0}{4\pi q_0} e^{-\tau\omega_0(1+q/q_0)^2} \tag{A14}$$

where $q_0^2/2 = \omega_0$, i.e., we first seed $|q|$ according to $W_1(q) == 1/[q_0(1+q/q_0)^2]$ (and isotropic around the point $\mathbf{q}=0$), and then τ according to $W_2(q|\tau) = \omega_0(1+q/q_0)^2 e^{-\tau\omega_0(1+q/q_0)^2}$. Since the typical length of the phonon propagator in time depends on how close is the polaron momentum to the dispersion law end-point, we also use another variant of seeding new variables $\mathbf{q}$ and τ, namely, we factorize the distribution function into $W(\mathbf{q},\tau) = W_1(q'=|\mathbf{q}-\mathbf{k}|)W_3(\tau)$ (i.e., isotropic around the point $\mathbf{k}$), where $W_3(\tau) = \Omega e^{-\tau\Omega}$ and $\Omega \ll \omega_0$ down to $\Omega \sim 0.01\omega_0$ close to the end-point.

We underline, that the above choices are motivated by the physics of the problem, in particular, if the combination $V^2(q)q^2$ was some power law function of q (e.g., when the interaction vertex is non-singular at small momentum or even goes to zero as $q \to 0$) one would better have to choose $W_1(q \to 0) \propto V^2(q)q^2$ to ensure that nowhere in the accessible parameter region the acceptance ratio (see below) is singular.

Now the proposing stage is completed, and we are ready to perform accept/reject step, following the above prescription, Eq. (A11). The corresponding function $W(\vec{x})$ ($\vec{x} \equiv \{\tau_1, \tau_2, \mathbf{q}\}$) reads (for the first version)

$$W(\vec{x}) \;=\; \frac{\omega_0\, e^{-\tau\omega_0(1+q/q_0)^2}}{4\pi\tau_0 q_0}\;, \tag{A15}$$

where τ_0 is the length of the free-electron propagator, where the point τ_1 is selected. As mentioned already, this form of W is by no means the unique one. Apart from the factor $p_{\mathcal{B}}/p_{\mathcal{A}}$ which will be discussed later, the ratio (A13) is now completely defined since

$$\frac{\mathcal{D}_{m+n}(\xi_{m+n}, y, x_1, \ldots, x_m, \vec{x})}{\mathcal{D}_m(\xi_m, y, x_1, \ldots, x_m)} = \frac{2\sqrt{2}\pi\alpha}{(2\pi)^3} e^{-(\tau_2-\tau_1)[\omega_0 + E(<\mathbf{p}>_{\tau_1,\tau_2}-\mathbf{q}) - E(<\mathbf{p}>_{\tau_1,\tau_2})]}\;. \tag{A16}$$

The algorithm for the process $\mathcal{B}$ is to select at random (with equal probabilities) some phonon propagator, and, if it is not attached to the diagram end, with the probabilities given in Eqs. (A12) and (A15) remove it.

To complete the description of the sub-processes $\mathcal{A}$ and $\mathcal{B}$, we should define the ratio $p_{\mathcal{B}}/p_{\mathcal{A}}$. It is quite reasonable to select creation and annihilation procedures with equal probabilities. At the first glance it might seem that this immediately leads to $p_{\mathcal{B}}/p_{\mathcal{A}} = 1$, but this is not true. The point is that when we select an electron propagator for placing the point τ_1, we have N_e equal chances, where N_e is the number of free-electron propagators in the diagram being modified [denominator of Eq. (A13)], and when we select a phonon propagator for removing, we have N_{ph} equal chances, where N_{ph} is the number of phonon propagators in the diagram from which we try to remove the propagator [numerator of Eq. (A13)]. These N_e and N_{ph} are straightforwardly related to each other:

$$N_{ph} = (N_e + 1)/2\;. \tag{A17}$$

We thus get

$$\frac{p_{\mathcal{B}}}{p_{\mathcal{A}}} = \frac{2N_e}{N_e+1} = \frac{2N_{ph}-1}{N_{ph}}\;. \tag{A18}$$

(Note a misprint in Ref. [1], where the r.h.s. of Eq. (10) gives $p_{\mathcal{A}}/p_{\mathcal{B}}$ instead of $p_{\mathcal{B}}/p_{\mathcal{A}}$).

b. Adding/removing a pair of phonon propagators attached to diagram ends

This update is done in close analogy with the previous one, except minor changes to which we proceed now. First, we select time positions τ_1 and τ_2 for the left- and right-end propagators according to the probability densities

$W_l(\tau_1) = \Omega e^{-\Omega\tau_1}$ and $W_r(\tau_2) = \Omega e^{-\Omega(\tau-\tau_2)}$. In the first variant of the update $\Omega = \omega_0$ and in the second $\Omega \ll \omega_0$. Also, the phonon momentum is suggested using the same distribution $W_1(q' = q)$ or $W_1(q' = |\mathbf{q} - \mathbf{k}|)$. We thus have

$$W(\vec{x}) = \frac{\Omega^2 e^{-\Omega(\tau+\tau_1-\tau_2)}}{4\pi q_0 (1 + q'/q_0)^2} \ , \tag{A19}$$

and

$$\frac{\mathcal{D}_{m+n}(\xi_{m+n}, y, x_1, \ldots, x_m, \vec{x})}{\mathcal{D}_m(\xi_m, y, x_1, \ldots, x_m)} = \frac{2\sqrt{2}\pi\alpha}{(2\pi)^3} e^{-\omega_0(\tau+\tau_1-\tau_2)} \times$$
$$\begin{cases} e^{-[E(<\mathbf{p}>_{0,\tau_1}-\mathbf{q})-E(<\mathbf{p}>_{0,\tau_1})]\tau_1-[E(<\mathbf{p}>_{\tau_2,\tau}-\mathbf{q})-E(<\mathbf{p}>_{\tau_2,\tau})](\tau-\tau_2)} & \tau_1 < \tau_2 \\ e^{-[E(<\mathbf{p}>_{0,\tau_2}-\mathbf{q})-E(<\mathbf{p}>_{0,\tau_2})]\tau_2-[E(<\mathbf{p}>_{\tau_2,\tau_1}-2\mathbf{q})-E(<\mathbf{p}>_{\tau_2,\tau_1})](\tau_1-\tau_2)-[E(<\mathbf{p}>_{\tau_1,\tau}-\mathbf{q})-E(<\mathbf{p}>_{\tau_1,\tau})](\tau-\tau_1)} & \tau_1 > \tau_2 \end{cases} \ . \tag{A20}$$

The algorithm for the inverse procedure is to select at random (with equal probabilities) a pair of propagators from the list of pairs attached to the diagram end, and with the probabilities given in Eqs. (A12) and (A19) remove it. Since we select procedures inserting and removing pairs of propagators with equal probabilities, we have

$$\frac{p_{\mathcal{B}}}{p_{\mathcal{A}}} = \frac{1}{N_z + 1} \ . \tag{A21}$$

APPENDIX B: METHOD OF SPECTRAL ANALYSIS

1. General background and outline of the method

The problem of restoring positive definite spectral density function $\rho(\omega)$ from known imaginary-time Green function $G(\tau)$ is the problem of solving linear first-type Fredholm equation

$$\int_0^\infty e^{-\tau\omega}\rho(\omega)d\omega = G(\tau) \ , \tag{B1}$$

where the domain of definition of the functions $G(\tau)$ and $\rho(\omega)$ is $[0, \infty]$. The normalization of the Green function $G(0) = 1$ implies the additional constraint

$$\int_0^\infty \rho(\omega)d\omega = 1 \ . \tag{B2}$$

In a realistic situation the Green function is known at a discrete set of times $\{\tau_i\}, i = 1, ..., N$ with some statistic errors at each time point. As is well known, in this case the problem of solving Eq. (B1) belongs to the class of ill-posed problems. The characteristic feature of the ill-posed problem is that the solution of equation (B1) is not unique even when statistic errors are absent, as there is an infinite number of unknown functions $\tilde{\rho}(\omega)$ satisfying (B1). In the case of finite statistic errors one may face a situation when the solution of Eq. (B1) under the constraint (B2) does not exist at all. Therefore, it is natural to formulate the problem as to find an approximate solution $\rho_{\min}(\omega)$ which reproduces G at a finite set of times with smallest deviation $D_{\min}$. The definition of the measure of deviation depends on the method used, and the value of minimal deviation $D_{\min}$ is determined by the magnitude of statistic errors.

There are two fundamental difficulties that are inherent to the spectral analysis. The first one is the well-known saw-tooth instability of the linear Fredholm equation of the first type – an approximate solution $\tilde{\rho}(\omega)$ does not reproduce the true solution $\rho(\omega)$ even if $\tilde{\rho}(\omega)$ generates the Green function

$$\tilde{G}(\tau) = \int_0^\infty e^{-\omega\tau}\tilde{\rho}(\omega)d\omega \tag{B3}$$

which reproduces $G(\tau)$ with any preassigned accuracy. This difficulty is treated usually by the regularization method that smoothes the saw-tooth noise of approximate solution $\tilde{\rho}(\omega)$. The idea of the regularization method is to introduce some nonlinearity into Eq. (B1) that imposes constraints on the derivatives of $\tilde{\rho}(\omega)$. There are two main drawbacks of this method. First, regularization method is unable to restore the spectral density which has sharp features. Second,

due to a distortion of the initial equation by additional regularization terms the approximate solution reproduces the function $G(\tau)$ with relatively high deviation $D \gg D_{\min}$. Hence, the information from the most representative region of the deviations $D \sim D_{\min}$ is lost.

The second difficulty inherent to the problem of solving equation (B1) is that any representation of $\tilde{\rho}(\omega)$ by a preassigned discrete set $\{\rho(\omega_f)\}, f = 1, ..., M$ is the source of uncontrollable systematic errors. For one thing, if the function $\rho(\omega)$ contains a sharp feature with a significant weight at some ω', which does not match the discrete set $\{\omega_f\}$, this feature cannot be reproduced properly and, therefore, the rest of the spectral density can be distorted beyond recognition. Note, that all iteration methods as well as the methods based on solving the nonlinear system of equations use preassigned discretisation of the ω-space.

We present a method of solving equation (B1) that avoids distortion of equation by nonlinear terms and, thus, probes the most representative interval of deviations. Besides, the method does not suffer from systematic errors as it does not involve preassigned discretisation of the ω-space. The idea of the method is to generate by a stochastic procedure a (large enough) set of M positive definite statistically independent approximate solutions $\{\tilde{\rho}_j(\omega)\}, j = 1, ..., M$ with deviation measures $D_j \sim D_{\min}$. And then, taking advantage of the linearity of Eq. (B1), choose the final solution as the average

$$\rho(\omega) \;=\; M^{-1} \sum_{j=1}^{M} \tilde{\rho}_j(\omega) \; . \tag{B4}$$

The reason is that while the particular solution $\tilde{\rho}_j(\omega)$ possesses the saw-tooth instability, the stochastic character of the procedure of particular solution generation should lead to averaging out the saw-tooth noise. Note, that the condition $\tilde{\rho}_j(\omega) > 0$ and constraint (B2) substantially enhance the convergence of the averaging (B4).

The method of generation of a particular solution is based on the optimization of the deviation

$$D[\tilde{\rho}] = \int_0^{\tau_{\max}} \left| G(\tau) - \tilde{G}(\tau) \right| G^{-1}(\tau) d\tau \; . \tag{B5}$$

Here $\tau_{\max}$ is the maximal τ up to which $G(\tau)$ is known. The weight function $G^{-1}(\tau)$ is to efficiently utilize information from the whole range $[0, \tau_{\max}]$, even in the case when the function $G(\tau)$ decreases by orders of magnitude with τ. Note, that we use weight function $G^{-1}(\tau)$ rather than $\tilde{G}^{-1}(\tau)$ to avoid feedback instabilities in generation $\tilde{\rho}(\omega)$.

Our optimization procedure does not involve preassigned fragmentation of the ω-space. The number of parameters used for parametrization of the spectral density function $\tilde{\rho}(\omega)$ is being varied during optimization process, so that any spectral function can, in principle, be reproduced within any preassigned accuracy. The process of generating a particular solution involves a random choice of the initial-configuration parameters and subsequent optimization of the deviation by changing the parameter values, as well as the number of the parameters. The maximal number of continuous parameters and the number of particular solutions M are limited only by the computer performance.

2. Configuration and method of getting independent solution

We parametrize $\tilde{\rho}$ as a sum

$$\tilde{\rho}(\omega) = \sum_{t=1}^{K} \chi_{\{P_t\}}(\omega) \tag{B6}$$

of rectangulars $\{P_t\} = \{h_t, w_t, c_t\}$

$$\chi_{\{P_t\}}(\omega) = \begin{cases} h_t & , \quad \omega \in [c_t - w_t/2, c_t + w_t/2] \; , \\ 0 & , \quad \text{otherwise} \; . \end{cases} \tag{B7}$$

determined by height $h_t > 0$, width $w_t > 0$, and center $c_t > 0$.

A configuration

$$\mathcal{C} = \{\{P_t\}, \, t = 1, ..., K\} \tag{B8}$$

with the constraint

$$\sum_{t=1}^{K} h_t w_t = 1 \tag{B9}$$

defines, according to Eqs. (B6, B3), the function $\tilde{G}(\tau)$ at any time point

$$\tilde{G}_{\mathcal{C}}(\tau) = \begin{cases} 1 & , \quad \tau = 0 \, , \\ 2\tau^{-1} \sum\limits_{t=1}^{K} h_t e^{-c_t \tau} \sinh(w_t \tau/2) & , \quad \tau \neq 0 \, . \end{cases} \tag{B10}$$

To express the deviation (B5) as an analytic function of the values of G and $\tilde{G}$ at the set of times $\{\tau_i\}, i = 1, ..., N$ [where the function $G(\tau)$ is known], we use linear interpolation between closest points.

Note, that the specific type of the functions (B7) is not crucial for the general features of the method although simple form of analytic expressions (B9, B10) is of value for fast performance.

The procedure of obtaining a particular solution $\tilde{\rho}_j(\omega)$ consists of randomly generating some initial configuration $\mathcal{C}_j^{\rm init}$ followed by nondeterministic sequence of configuration changes until deviation satisfies the condition

$$D[\mathcal{C}_j^{\rm fin}] < D_u \sim D_{\rm min} \tag{B11}$$

(D_u is the upper limiting deviation.) for final configuration $\mathcal{C}_j^{\rm fin}$. The nondeterministic character of configuration changes is achieved by random selection of various elementary updates.

3. General features of elementary updates

By elementary update we mean a random change of the configuration, which is either accepted or rejected in accordance with certain rules. There are two classes of elementary updates. The updates of the class I do not alter the number of rectangulars, K, changing only the values of the parameters from a randomly chosen set $\{P_t\}$. The updates of the class II either add a new rectangular with randomly chosen parameters $\{h_{K+1}, w_{K+1}, c_{K+1}\}$, or remove stochastically chosen rectangular t from the configuration. If a proposed change violates constraint (B10) (e.g., a change of w_t or h_t, or any update of the class II), then the necessary change of some other parameter set $\{P_{t'}\}$ is simultaneously proposed, to satisfy the requirement of the constraint.

The updats should keep parameters of a new configuration within domain of definition of configuration $\mathcal{C}$. Formally, the domains of definition of configuration (B8) are $\Xi_{h_t} = [0, \infty]$, $\Xi_{c_t} = [0, \infty]$, $\Xi_{w_t} = [0, 2c_t]$, and $\Xi_K \in [1, \infty]$. However, for the sake of faster convergence, we reduce domains of definition.

As there is no general *a priori* prescription for choosing reduced domains of definition, the rule of thumb is to start with maximal domains and then, after some rough solution is found, reduce the domains to reasonable values suggested by this solution. In particular, since the probability to propose a change of any parameter of configuration is proportional to K^{-1}, it is natural to restrict maximal number of rectangulars $\Xi_K \in [1, K_{\rm max}]$ by some large number $K_{\rm max}$. To forbid rectangulars with extremely small weight, which contribution to $\tilde{G}(\tau)$ is less than statistic errors of $G(\tau)$, one can impose the constraint $h_t w_t \in [S_{\rm min}, 1]$, with $S_{\rm min} \ll K_{\rm max}^{-1}$. When there is some preliminary knowledge that overwhelming majority of integral weight of the spectral function $\rho(\omega)$ is in a range $[\omega_{\rm min}, \omega_{\rm max}]$, one can restrict the domain of definition of the parameter c_t by $\Xi_{c_t} = [\omega_{\rm min}, \omega_{\rm max}]$. Then, to reduce the phase space one can choose $\Xi_{h_t} = [h_{\rm min}, \infty]$ and $\Xi_{w_t} = [w_{\rm min}, \min\{2(c_t - \omega_{\rm min}), 2(\omega_{\rm max} - c_t)\}]$.

While the initial configuration, the update type, and the parameter to be altered are chosen stochastically, the variation of the values of the parameters relevant to the update is optimized to maximize the decrease of $\mathcal{D}$. Each elementary update of our optimization procedure (even that of the class II) is organized as a proposal to change some continuous parameter ξ by randomly generated $\delta\xi$ in a way that the new value belongs to Ξ_ξ. Although proposals with smaller values of $\delta\xi$ are accepted with higher probability it is important, for the sake of better convergence, to propose sometimes changes $\delta\xi$ that probe the whole domain of definition Ξ_ξ. To probe all scales of $\delta\xi \in [\delta\xi_{\rm min}, \delta\xi_{\rm max}]$ we generate $\delta\xi$ with the probability density function $\mathcal{P} \sim (\max(|\delta\xi_{\rm min}|, |\delta\xi_{\rm max}|)/|\delta\xi|)^\gamma$, where $\gamma \gg 1$.

Calculating the deviation measures $D(\xi)$, $D(\xi + \delta\xi)$, $D(\xi + \delta\xi/2)$, and searching for the minimum of the parabolic interpolation, we find an optimal value of the parameter change

$$\delta\xi_{\rm opt} = -B/2A, \tag{B12}$$

where

$$A = 2(D(\xi + \delta\xi) - 2D(\xi + \delta\xi/2) + D(\xi))(\delta\xi)^{-2}, \tag{B13}$$

and

$$B = (4D(\xi + \delta\xi/2) - D(\xi + \delta\xi) - 3D(\xi))\delta\xi. \tag{B14}$$

In the case $A > 0$ and $\xi_{\rm opt} \in \Xi_\xi$ we adopt as the update proposal $\tilde{\delta\xi}$ one of the values $\delta\xi$, $\delta\xi/2$, or $\delta\xi_{\rm opt}$ for which the deviation measure $D(\xi + \tilde{\delta\xi})$ is the smallest. Otherwise, if the parabola minimum is outside Ξ_ξ, one has to compare only deviations for $\delta\xi$ and $\delta\xi/2$.

4. Global updates

The updating strategy has to provide efficient minimization of the deviation measure until criterion (B11) is satisfied. It is highly inefficient to accept only those proposals that lead to the decrease of deviation, since, in a general case, there is an enormous number of deviation local minima $D_{\rm loc}[\mathcal{C}] > D_u$. As we observed it in practice, these multiple minima drastically slow down (or even freeze) the process.

To optimize escape from a local minimum, one has to provide a possibility of reaching a new local minimum with lower deviation through a sequence of less optimal configurations. It might seem that the most natural way of doing this would be to accept sometimes (with low enough probability) the updates leading to the increase of the deviation. However, this simple strategy turns out to be impractical. The reason is that the density of configurations per interval of deviation sharply increases with D. So that the acceptance probability for a deviation-increasing update should be fine-tuned to the value of D. Otherwise, the optimization process will be either non-convergent, or ineffective [if the acceptance probability is, correspondingly, either too large, or too small in some region of D].

A way out of the situation is to perform some sequence of T *temporary* elementary updates of a configuration $\mathcal{C}(0)$

$$\mathcal{C}(0) \to \mathcal{C}(1) \to ... \to \mathcal{C}(r) \to \mathcal{C}(r+1) \to ... \to \mathcal{C}(T)\ , \tag{B15}$$

where the proposal to update the configuration $\mathcal{C}(r) \to \mathcal{C}(r+1)$ is (temporary) accepted with the probability

$$\mathcal{P}_{r\to r+1} = \begin{cases} 1 & , \quad D[\mathcal{C}(r+1)] < D[\mathcal{C}(r)]\ , \\ f\left(D[\mathcal{C}(r)]/D[\mathcal{C}(r+1)]\right) & , \quad D[\mathcal{C}(r+1)] > D[\mathcal{C}(r)]\ . \end{cases} \tag{B16}$$

(Function f satisfies boundary conditions $f(0) = 0$ and $f(1) = 1$.) Then we choose out of the configurations $\{\mathcal{C}(r)\}$ (B15) the one with minimal deviation and, if it is different from $\mathcal{C}(0)$, declare it to be the result of the global update, or, if this configuration turns out to be just $\mathcal{C}(0)$, reject the update.

We choose the function f in the form

$$f(x) = x^{1+d} \qquad (d > 0)\ , \tag{B17}$$

which leads to comparatively high probabilities to accept small increases of deviation measures and hampers significant enlargements of deviation. Empirically, we found out that the global update procedure is most effective if one keeps parameter $d = d_1 \sim 0$ at the first T_1 steps of sequence (B15) (to leave local minimum) and then changes this parameter to a value $d = d_2 \gg 1$ for the last $T - T_1$ elementary updates (to decrease the deviation measure). In our algorithm the values $T \in [1, T_{\max}]$, $T_1 \in [1, T]$, $d_1 \in [0, 1]$, and $d_2 \in [1, d_{\max}]$ were stochastically chosen for each global update run.

5. Final solution and refinement

After a set of M configurations

$$\left\{C_j^{\rm fin}, j = 1, ..., M\right\} \tag{B18}$$

that satisfy the criterion (B11) is produced, the solution (B4) is obtained by summing up the rectangulars (B7,B18).

We, however, employ a more elaborated procedure, which we call refinement. Namely, we use the set (B18) as a source of $M_{\rm ref}$ new independent starting configurations for further optimization. These starting configurations are generated as a linear combinations of randomly chosen members of the set (B18) with stochastic weight coefficients.

Then, the refined final solution is represented as the average (B4) of $M_{\rm ref}$ particular solutions resulting from the optimization procedure.

The main advantage of such a trick is that initial configurations for optimization procedure now satisfy the criterion (B11) from the very beginning and, thus, upper limiting deviation D_u can be considerably reduced. Moreover, as any linear combination of sufficiently large number R of randomly chosen parent configurations $\left\{C_\eta^{\rm fin}, \eta = 1, ..., R\right\}$ smoothes the saw-tooth noise, the deviation of a summary configuration $C_{\rm ref}^{\rm fin}$ is normally lower than that of each additive one.

6. Elementary updates of class I

(A) Shift of rectangular. Change the center c_t of a randomly chosen rectangular t. The continuous parameter for optimization (B12-B14) is $\xi = c_t$ which is restricted by domain of definition $\Xi_{c_t} = [\omega_{\min} + w_t/2, \omega_{\max} - w_t/2]$.

(B) Change of width without change of weight. Alter the width w_t of a randomly chosen rectangular t without change of the rectangular weight $h_t w_t = const$ and center c_t. The continuous parameter for optimization is $\xi = w_t$ which is restricted by $\Xi_{w_t} = [w_{\min}, \min\{2(c_t - \omega_{\min}), 2(\omega_{\max} - c_t)\}]$.

(C) Change of weight of two rectangulars. Change the heights of two rectangulars t and t' (where t is a randomly chosen and t' is either randomly chosen or closest to t rectangular) without change of widths of both rectangulars. Continuous parameter for optimization is the variation of the rectangular t height $\xi = h_t$. To restrict the weights of chosen rectangulars to $[S_{\min}, 1]$ and preserve the total normalization (B2) this update suggests to change $h_t \to h_t + \delta\xi$ and $h_{t'} \to h_{t'} - \delta\xi w_{t'}/w_t$ with $\delta\xi$ confined to the interval

$$S_{\min}/w_t - h_t < \delta\xi < (h_{t'} - S_{\min}/w_{t'})w_t/w_{t'} \tag{B19}$$

7. Elementary updates of class II

(D) Adding a new rectangular. To add a new rectangular one has to generate some new set $\{P_{\rm new}\} = \{h_{\rm new}, w_{\rm new}, c_{\rm new}\}$ and reduce the weight of some other rectangular t (either randomly chosen or closest) in order to keep the normalization condition (B2). The reduction of the rectangular weight t is obtained by decreasing its height h_t.

The center of the new rectangular is selected at random according to

$$c_{\rm new} = (\omega_{\min} + w_{\min}/2) + (\omega_{\max} - \omega_{\min} - w_{\min})r \tag{B20}$$

As soon as the value $c_{\rm new}$ is generated, the maximal possible width of a new rectangular is given by

$$w_{\rm new}^{\max} = 2\min(\omega_{\max} - c_{\rm new}, c_{\rm new} - \omega_{\min}). \tag{B21}$$

Continuous parameter for optimization $\delta\xi = h_{\rm new} w_{\rm new}$ is generated to keep weights of both new rectangular and rectangular t larger than $S_{\min}$

$$\delta\xi = S_{\min} + r(h_t w_t - S_{\min}) \tag{B22}$$

Then, the value of the new rectangular height $h_{\rm new}$ for given $\delta\xi$ is generated to keep the width of new rectangular within the limits $[w_{\min}, w_{\rm new}^{\max}]$

$$h_{\rm new} = \delta\xi/w_{\rm new}^{\max} + r(\delta\xi/w_{\min} - \delta\xi/w_{\rm new}^{\max}). \tag{B23}$$

(E) Removing a rectangular. To remove some randomly chosen rectangular t, we enlarge the height $h_{t'}$ of some another (either randomly chosen or closest) rectangular t' according to condition (B2). Since such procedure does not involve continuous parameter for optimization, we unite removing of rectangular t with the shift procedure *(A)* of the rectangular t'. Then, the proposal is the configuration with the smallest deviation measure.

(F) Splitting a rectangular. This update cuts some rectangular t into two rectangulars with the same heights h_t and widths $w_{\rm new_1} = w_{\min} + r(w_t - w_{\min})$ and $w_{\rm new_2} = w_t - w_{\rm new_1}$. Since removing a rectangular t and adding of two new glued rectangulars does not change the spectral function we introduce the continuous parameter $\delta\xi$ which describes the shift of the center of a new rectangular with the smallest weight. Second rectangular is shifted into

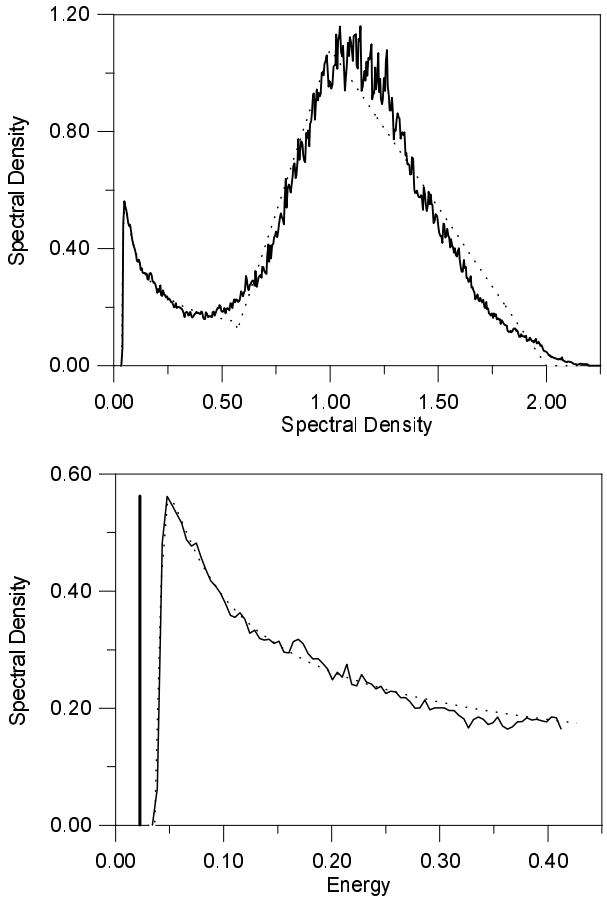

FIG. 16. Model spectral density (dashed line) and the result of spectral analysis (solid line). The position of the delta-function is shown only in the lower panel.

opposite direction to keep the center of gravity of two rectangulars unaltered. The domain of definition Ξ_ξ obviously follows from the parameters of the new rectangulars.

(G) Gluing rectangulars. This update glue two (either randomly chosen or closest) rectangulars t and t' into single new rectangular with the weight $h_{\rm new}w_{\rm new} = h_t w_t + w_{t'}h_{t'}$ and width $w_{\rm new} = (w_t + w_{t'})/2$. The initial center of the new rectangular $c_{\rm new}$ corresponds to the center of gravity of rectangulars t and t'. We introduce a continuous parameter by simultaneously shifting the new rectangular.

8. Tests

To check the accuracy of our approach, we tested it for the spectral density distribution that spreads over large range of frequencies and simultaneously possesses fine structure in low-frequency region. The test spectrum was modeled as the sum of the delta-function with the energy $\varepsilon_\delta = 0.03$ and the weight $Z_\delta = 0.07$, and continuous high-frequency spectral density which starts at the threshold $\varepsilon_{\rm th} = 0.04$. The continuous part of the spectrum $\rho_{\rm con}$ was modeled by the function [In fact, this functional form is predicted by the Pitaevskii theory for $\rho(\omega)$ near the end point.]

$$\rho_{\rm con}(\omega) = \frac{Z_\delta\sqrt{\omega - \varepsilon_{\rm th}}}{2\pi\sqrt{\varepsilon_{\rm gap}}[(\omega - \varepsilon_{\rm th}) + \varepsilon_{\rm gap}]} \tag{B24}$$

(here $\varepsilon_{\rm gap} = \varepsilon_{\rm th} - \varepsilon_\delta$ is a microgap) in the range $\omega \in [\varepsilon_{\rm th}, 0.566]$ and by a triangule at higher frequencies (see the dashed line in the upper panel of Fig. 16). The Green function $G(\tau)$ was calculated from the model spectral density in the $n_{\rm max} = 300$ points $\tau_i = \tau_{\rm max} i^2/n_{\rm max}$ in the time range from zero to $\tau_{\rm max} = 1000$. The restored spectral density reproduces both gross features of high-frequency part (upper panel in Fig. 16) and the fine structure at small frequencies (lower panel of Fig. 16). The energy and the weight of the delta-function was restored with the accuracy 10^{-4}. The final solution was obtained by the averaging (B4) of $M = 1100$ particular solutions.

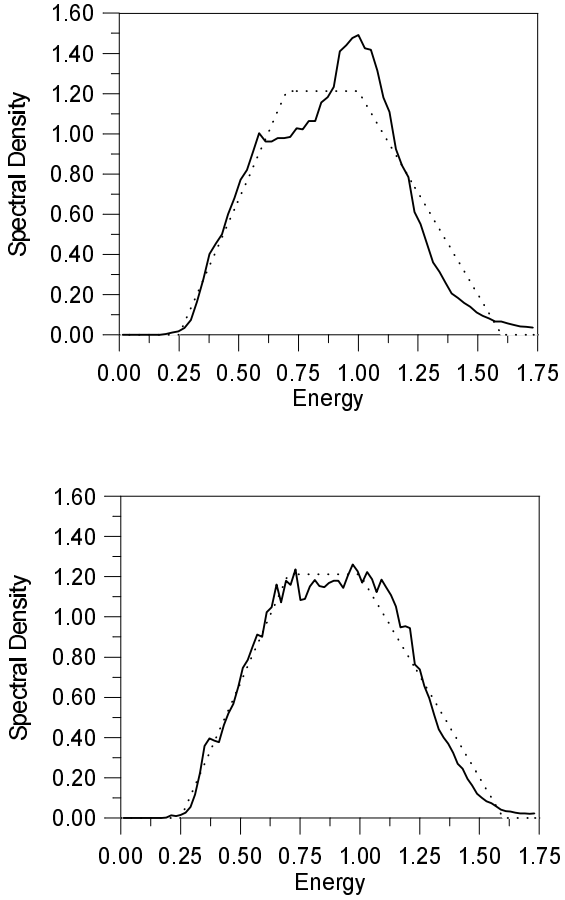

FIG. 17. The model spectrum (dashed lines) and results of spectral analysis (solid lines) for $\eta = 10^{-2}$ (upper panel) and $\eta = 10^{-3}$ (lower panel).

To evaluate the precision which characterizes how typical particular solution $\tilde{\rho}_j(\omega)$ (see (B3)) reproduces the Green function $G(\tau)$, we introduce the maximal relative deviation

$$\eta = \max \left[\frac{\mid G(\tau_i) - \tilde{G}(\tau_i) \mid}{G(\tau_i)} \right] , \; i \in [1, n_{\max}] \; , \tag{B25}$$

which typical value is $\eta = 10^{-4}$ for a particular solution of the spectrum in Fig. 16.

Since the Green function which is obtained from Monte Carlo calculations contains some statistic errors at each time point, the minimal value of parameter η is limited by the quality of the calculated Green function $G(\tau)$. To study the influence of the (uncorrelated) statistic errors we studied the stability of the method against stochastic noise

$$G(\tau_i) \rightarrow G(\tau_i)(1 + \eta r_i) \; , \; i = 1, n_{\max} \; , \tag{B26}$$

introduced by random numbers $r_i \in [0, 1]$. It is seen that the method restores the gross features of the spectrum (position and width) even for rather roughly calculated Green function, with $\eta = 10^{-2}$ (upper panel in Fig. 17), whereas the precision $\eta = 10^{-3}$ is sufficient to resolve the lineshape (lower panel in Fig. 17)).

[1] N.V. Prokof'ev and B.V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998).
[2] J. Appel, in *Solid State Physics*, edited by H. Ehrenreich, F. Seitz, and D. Turnbull (Academic, New York, 1968), Vol. 21.
[3] L.D. Landau, Sov. Phys. **3**, 664 (1933).
[4] H. Fröhlich, H. Pelzer, and S. Zienau, Phil. Mag. **41**, 221 (1950).
[5] E.P. Gross, Ann. Phys. **8**, 78 (1959).
[6] D. Matz and B.C. Burkey, Phys. Rev. B **3**, 3487 (1971).
[7] J.M. Luttinger and C.Y. Lu, Phys. Rev. B **21**, 4251 (1980).
[8] R. Mańka, Phys. Lett. A **67**, 311 (1978); Phys. Stat. Sol. (b) **93**, 53 (1979).
[9] R. Mańka and M. Suffczynski, J. Phys. C **13**, 6369 (1980).
[10] Y. Lépine and D. Matz, Phys. Stat. Sol. (b) **96**, 797 (1979).
[11] I.D. Feranchuk, S.I. Fisher, and L.I. Komarov, J. Phys. C **18**, 5083 (1985).
[12] F.M. Peeters and J.T. Devreese, Phys. Stat. Sol. (b) **112**, 219 (1982).
[13] R.P. Feynman, Statistical Mechanics, Reading, Benjamin (1972)
[14] R.P. Feynman, Phys. Rev. **97**, 660 (1955).
[15] T.D. Schultz, Phys. Rev. **116**, 526 (1959).
[16] T.D. Lee, F.E. Low and D. Pines, Phys. Rev. **90**, 297 (1953).
[17] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Zh. Eksp. Teor. Fiz. **114**, 570 (1998) [Sov. Phys. JETP **87**, 310 (1998)].
[18] N. Metropolis *et al.*, J. Chem. Phys., **21**, 1087 (1953).
[19] A.M. Ferrenberg and R.H. Swendsen, Phys. Rev. Lett. **61**, 2635 (1988); **63**, 1658(E) (1989).
[20] H. Grabert, Phys. Rev. B **50**, 17364 (1994).
[21] L. D. Landau, S. I. Pekar, Zh. Eksp. Teor. Fiz. **18**, 419 (1948).
[22] D. Pines, in *Polarons and Exitons*, eds. C.G. Kuper and G.D. Whitfield, Plenum Press, N.Y., 155 (1962).
[23] L.P. Pitaevskii, Zh. Eksp. Teor. Fiz. **36**, 1168 (1959).
[24] Polarons in Ionic Crystals and Polar Semiconductors, ed. by J. T. Devreese, North-Holland, Amsterdam (1972).
[25] V. Gurevich, I. Lang and Yu. Firsov, Soviet Phys.- Solid State **4**, 918 (1862).
[26] J. Devreese, W. Huybrechts and L. Lemmens, Phys. Stat. Sol. (b) **48**, 77 (1971).
[27] E. Kartheuser, R. Edvard and J. Devreese, Phys. Rev. Lett. **22**, 94 (1969).
[28] J. Devreese, J. De Sitter and M. Goovaerts, Solid State Comm. **9**, 1383 (1971).
[29] J. Devreese, J. De Sitter and M. Goovaerts, Phys. Rev. B **5**, 3367 (1972).
[30] D.M. Ceperley, J. Comp. Phys. **51**, 404 (1983); D.M. Ceperley and B.J. Alder, J. Chem. Phys. **81**, 5833 (1984).
[31] N.V. Prokof'ev, B.V. Svistunov, and I.S. Tupitsyn, Phys. Rev. Lett. **82**, 5092 (1999).