

Self-consistent-field ensembles of disordered Hamiltonians: Efficient solver and application to superconducting films

Matthias Stosiek,¹ Bruno Lang,² and Ferdinand Evers¹

¹*Institute of Theoretical Physics, University of Regensburg, D-93040 Regensburg, Germany*

²*Institute of Applied Informatics, University of Wuppertal, D-42119 Wuppertal, Germany*



(Received 25 September 2019; accepted 13 February 2020; published 13 April 2020)

Our general interest is in self-consistent-field (scf) theories of disordered fermions. They generate physically relevant subensembles (“scf ensembles”) within a given Altland-Zirnbauer class. We are motivated to investigate such ensembles (i) by the possibility to discover new fixed points due to (long-range) interactions; (ii) by analytical scf theories that rely on partial self-consistency approximations awaiting a numerical validation; and (iii) by the overall importance of scf theories for the understanding of complex interaction-mediated phenomena in terms of effective single-particle pictures. In this paper we present an efficient, parallelized implementation solving scf problems with spatially local fields by applying a kernel-polynomial approach. Our first application is the Bogoliubov-deGennes theory of the attractive- U Hubbard model in the presence of on-site disorder; the scf fields are the particle density $n(\mathbf{r})$ and the gap function $\Delta(\mathbf{r})$. For this case, we reach system sizes unprecedented in earlier work. They allow us to study phenomena emerging at scales substantially larger than the lattice constant, such as the interplay of multifractality and interactions or the formation of superconducting islands. For example, we observe that the coherence length exhibits a nonmonotonic behavior with increasing disorder strength already at moderate U . With respect to methodology our results are important because we establish that partial self-consistency (“energy-only”) schemes as typically employed in analytical approaches tend to miss qualitative physics such as island formation.

DOI: [10.1103/PhysRevB.101.144503](https://doi.org/10.1103/PhysRevB.101.144503)

I. INTRODUCTION

The symmetry classification of disordered metals as it has been devised by Altland and Zirnbauer is nowadays considered to be complete [1–3]. The classification is fundamental in the sense that all *generic* ensembles of random Hamiltonians have been covered. The classifying criterion is the presence or absence of one of the four elementary symmetries: time-reversal, spin-rotation, sublattice (chiral), and particle-hole (Bogoliubov-deGennes type).

Based on an (incomplete) analogy to the conventional Landau-Ginzburg-Wilson theories of classical phase transitions, there was a widespread misunderstanding with many researchers at the late 1980s and early 1990s that a classification based on symmetry (and topology) alone would (more or less) determine the phase diagrams and the associated critical points as well. In other words, the symmetry classification was largely identified with a classification of universality classes, i.e., of all nonequivalent quantum field theories (low-energy action functionals) that describe a disordered electron system. Therefore, it came as a surprise for the larger part of the community when models of disordered fermions had been found that formally belong to the same symmetry class but nevertheless exhibit different phase diagrams.

It is perhaps fair to say that despite of the progress in the symmetry classification, we are still far from a systematic understanding of all universality classes and phase diagrams that systems of disordered fermions could exhibit. One could rephrase by saying that the generic ensembles of random

Hamiltonians covered in the 10-fold way possess physically relevant subensembles that exhibit their own phase diagrams and critical fixed points. The power-law random-banded matrices (PRBM) constitute a well-studied example [4]. It offers a laboratory for criticality that can be addressed relatively easily with analytical and numerical techniques [4,5].

A. General motivation for investigating scf ensembles

The appearance of criticality in the PRBM ensemble is a synthetic property; it is imposed by putting long-range (power-law) correlations into the hopping amplitudes of a tight-binding Hamiltonian. It therefore is interesting to explore properties of other ensembles that also exhibit long-range correlations in the Hamiltonian matrix elements but of a kind that is self-generated and in this sense “emergent.” Plausible candidates for such Hamiltonians are effective single-particle systems that appear in self-consistent-field (scf) theories of interacting fermions. A prototypical example could be the Hartree-Hamiltonian of a disordered wire or film; it carries a long-range correlated on-site potential due to a weakly screened Coulomb interaction.

Quite generally, we have in mind fermionic Hamiltonians.

$$\hat{H}_{\text{sc}} = \frac{1}{2} \sum_{xy} [h_{xy}[n, \Delta] c_x^\dagger c_y + \Delta_{xy}[n, h] c_x^\dagger c_y^\dagger + \text{H.c.}]; \quad (1)$$

the matrix $h_{xy}[n, \Delta]$ is a functional of the density matrix n and the pairing fields Δ . The self-consistency condition inherent to generic mean-field theories requires that the fields n and

Δ are expectation values of operators \hat{n} , $\hat{\Delta}$ to be calculated employing, among other ingredients, such as density matrices or exchange-correlation kernels, also \hat{H}_{sc} . Thus, scf conditions are implied,

$$n = \langle \hat{n} \rangle_{h,\Delta}, \quad \Delta = \langle \hat{\Delta} \rangle_{n,h} \quad (2)$$

that h , Δ , and n obey.

A microscopic randomness will enter \hat{H}_{sc} , e.g., via h incorporating random on-site energies or hopping amplitudes. The set of random Hamiltonians introduced thereby follows the conventional symmetry classification. However, only the subset of all members of a given symmetry class that happens to comply with Eq. (2) forms the *scf ensemble*.

We believe that scf ensembles, their physical and mathematical properties, constitute a fundamental research topic that may not yet have received the amount of attention it deserves. Our belief bases on two observations: (i) The elements of scf ensembles certainly tend to exhibit nontrivial correlations in their matrix elements h_{xy} and Δ_{xy} . If correlations happen to be strong enough, e.g., sufficiently long ranged, then new phases with novel critical behavior can be expected to emerge. (ii) Mean-field theories are important because they provide a tractable reference point for a perturbative analysis of interaction effects. Thus, they are a generic encounter in all theories of disordered fermions that try to incorporate interactions. To reveal, in particular, the impact of quantum fluctuations a thorough understanding of the mean-field reference point would certainly seem helpful.

We give examples for occurrences of scf ensembles:

(1) *Hartree theory (H)*: The obvious example to define ensembles of self-consistent Hamiltonians would be the Hartree theory. In this case, $\Delta = 0$ and the field n in Eqs. (1) and (2) should be identified with the particle density $n(\mathbf{r})$.

(2) *Hartree-Fock theory (HF)*: $\Delta = 0$; n resembles the density matrix $n(\mathbf{r}, \mathbf{r}')$ and h the Fock operator.

(3) *Density-functional theory (DFT)*: In the orthodox flavor $\Delta = 0$, n represents the particle density $n(\mathbf{r})$ and h becomes the Kohn-Sham-Hamiltonian. Roughly speaking, DFT differs from HF due to the presence of correlations in $h[n]$.

(4) *Bogoliubov-deGennes-Hamiltonian (BdG)*: The basic Hamiltonian is given in Eqs. (1) and (2).

Our short list is far from exhaustive and further examples could be given. For instance, we recall that many spin systems have faithful representations in terms of fermionic network models that also could be dressed with self-consistency requirements, like self-consistent fluxes.

1. Remarks

(i) The investigation of scf ensembles is a very challenging endeavor. The difficulty is that each disorder configuration requires us to find its own self-consistent fields h and Δ . The solution of the scf cycle is very difficult to do with analytical techniques. But also numerically it is demanding already at moderate system sizes of a few thousand sites. Consequently, the number of studies including full self-consistency appears to be limited. In Table I we list contributions most relevant to us.

(ii) A more general perspective can be developed that operates with self-consistency constraints on the Green's function

TABLE I. Studies of mean-field Hamiltonians in the literature that have been performed with full self-consistency. Abbreviations: integer quantum Hall effect (IQHE), Hartree-Fock (HF), time-dependent Hartree-Fock, density of states (DoS), inverse participation ratio (IP), multifractal (mf), local density approximation (LDA) for density-functional theory in Kohn-Sham (KS) formulation, localization length exponent (ν), linear system dimension (L), Bogoliubov-deGennes-type pairing terms (BdG), distribution function of local SC gaps [$P(\Delta)$], two dimensional (2D), correlation function of various observables A and B [$\Phi_{AB}(\mathbf{q})$].

Dim.	Mean field	Observables	Parameters	Ref.
IQHE	HF	Thouless numbers, IP	$L \approx 10$	[6]
	TDHF	“Kubo conductivity”		[7]
3D	HF	DoS, IP	$L = 10$	[8]
3D	HF	DoS, IP	$L = 24$	[6]
3D	HF	mf-dim D_2	$L = 10$	[9]
3D	HF	DOS, mf-dim D_2	$L = 18$	[10]
3D	BdG	$\Delta(\mathbf{r})$, LDoS	$L = 50(3D)$	[11]
	<i>s</i> -wave			
3D	DFT	mf-spectrum $f(\alpha)$	$L = 22$	[12]
3D	LDA	ν with	factor of	[13,14]
		KS-states	two in L	
2D	BdG	DoS, $P(\Delta)$	$L = 24$	[15,16]
	<i>s</i> -wave			
2D	BdG	DoS, $P(\Delta)$	$L = 38$	[17]
	<i>s</i> -wave	LDoS		
2D	BdG	DoS, $\Delta(\mathbf{r})$	$L = 36$	[18–20]
	<i>s</i> -, <i>d</i> -wave			
2D	BdG	$P(\Delta)$, $\Delta(\mathbf{r})$	$L = 25$	[21–24]
	<i>s</i> -wave	$\Phi_{AB}(\mathbf{q})$		
2D	BdG	$\Delta(\mathbf{r})$	$L = 12$	[25]
	<i>s</i> -wave			

$G(E)$ rather than on the elements of the Hamiltonian. The generalization becomes nontrivial when the self-energy $\Sigma(E)$ picks up an energy dependence. As a prototypical example we mention the GW theory. It constitutes an electronic-structure method that builds on the Hedin equations approximating them by ignoring vertex corrections [26]. In its full flavor the theory features a Green's function that satisfies a self-consistent set of equations defined by a truncated diagrammatic expansion [27,28].

(iii) A potential classification scheme of scf ensembles will involve concepts very different from the one designed by Altland and Zirnbauer (AZ). To see this, we recall that AZ distinguish 10 classes according to presence or absence of discrete symmetries. In contrast, the scf requirement as formulated in (2) invokes parameter-bound kernels. Hence, *a priori* the number of scf ensembles is not limited and an impression might arise according to which the scf ensembles carry a degree of arbitrariness and therefore are less fundamental. To address this reservation against the basic concept, we recall that there is a very special set of scf theories which is standing out; these scf theories share with the parent field theory they derive from the basic symmetries and in this sense are conserving [29]. Therefore, a classification of *conserving scf ensembles* goes together with the basic program of condensed-matter theory, which is to identify and understand

the fixed-point theories that are possible within a given AZ symmetry class.

B. Motivation for numerics and challenges

The numerical challenge that the scf ensembles pose as compared to simulations of noninteracting fermions is that for each sample the scf equation [Eq. (2)] has to be solved in an iterative fashion. Since the ensemble average requires solving hundreds of samples, typically, the computational cost for such studies is extensive. Presumably, this is the main reason why numerical studies of scf ensembles have been performed infrequently in the past, despite their obvious fundamental relevance.

Thus motivated, we here present an implementation of the scf problem that allows us to achieve relevant system sizes at an affordable numerical cost. The interplay of disorder-induced quantum interference and mean-field interactions can be studied on length scales that exceed the lattice constant by two orders of magnitude.

1. Reduction of scaling: KPM

The computationally demanding step limiting the code performance is the calculation of the scf fields, n , Δ , and h , that need to be evaluated in every iteration cycle of the self-consistency process. In the case of Hartree-Fock theory, for instance, this implies the reconstruction of the density matrix from a given Fock operator. In straightforward implementations the Hamiltonian is diagonalized in each iteration cycle to feed eigenvalues and eigenvectors into the right-hand side of (2); the cost is $\mathcal{O}(N_{\text{bf}}^3)$ operations, where N_{bf} is the dimension of the single-particle Hilbert space [30].

Consider the Hartree approximation: The matrix diagonalization appears, because the trace

$$n(\mathbf{r}) = \text{tr}[f(\mu - \hat{H}_H)\hat{n}(\mathbf{r})] \quad (3)$$

contains the Fermi-Dirac function, $f(\mu - \hat{H}_H)$, of a matrix valued argument that conventionally is evaluated in the basis of eigenstates of \hat{H}_H .

What many suggestions for $\mathcal{O}(N_{\text{bf}}^x)$ solvers, $x < 3$, of the self-consistency problem have in common is that they employ an alternative approach for trace computation that avoids a diagonalization of \hat{H}_H and therefore can be more efficient, in principle, than $\mathcal{O}(N_{\text{bf}}^3)$. One of the well-established options is the kernel polynomial method (KPM) [31]. The conceptual idea behind this approach is to expand $f(x)$ into a rapidly converging series of Chebyshev polynomials, $T_l(x)$, that are obtained recursively:

$$\begin{aligned} n(\mathbf{r}) &= \sum_{l=0}^{N_C} a_l \text{tr} T_l(\hat{H}_s)\hat{n}(\mathbf{r}) \\ &= \sum_{l=0}^{N_C} a_l \sum_{b=1}^{N_{\text{bf}}} \langle b | T_l(\hat{H}_s)\hat{n}(\mathbf{r}) | b \rangle, \end{aligned} \quad (4)$$

where \hat{H}_s denotes an appropriately scaled Hamiltonian \hat{H}_H and $|b\rangle$ is a suitable basis in which \hat{H}_s is sparse (N_C , order of the Chebyshev expansion; a_r , known expansion coefficients; N_{bf} , number of basis functions). As is seen from (4), the evaluation of the trace is of order $\mathcal{O}(N_C N_{\text{bf}}^2 N_{\text{nz}})$. N_{nz} denotes the number

of nonzero entries of \hat{H}_s per row. For a dense matrix we have $N_{\text{nz}} = N_{\text{bf}}$, while for a very sparse matrix $N_{\text{nz}} \approx N_{\text{bf}}^0$. For example, the BdG case, we have $N_{\text{nz}} = 2d + 2$, with $2d$ denoting the number of nearest neighbors on a cubic lattice in d dimensions.

2. Signatures of implementation

We have implemented a matrix-free KPM solver of the self-consistency problem (1) and (2) for the situation where the self-consistent fields are local in real space $n(\mathbf{r})$ and $\Delta(\mathbf{r})$, as is the case for the Hartree approximation and the Bogoliubov theory of s -wave superconductors. It operates at zero and nonzero temperature and is optimized for accelerated convergence for averages over the phase-space of disordered scf ensembles.

The KPM aspect of our implementation is similar to other variants described in earlier work. They have been proven useful in applications of the BdG equation for nanostructures with one or very few impurities but have not been applied to disordered samples. They differ in details from our implementation; Covaci *et al.* [11] and Nagai *et al.* [32,33] also use KPM to perform traces. In addition, Covaci *et al.* [11] also have employed a matrix-free implementation. While these authors expand the Green's function employing the Lorentz kernel, we expand the spectral function where the Jackson kernel typically has better convergence properties [31].

C. Application to dirty superconductors

Motivated by experiments on the superconductor-insulator transition, e.g., Baturina *et al.* [34] and Sacépé *et al.* [35], the attractive- U Hubbard model with on-site disorder has been studied recently in several computational [17,21–25,36–38] and analytical works [39–41].

Important insights have been gained within the framework of BdG theory. The most striking findings include (i) the granularity of the pairing amplitude (“islands”) emergent on the scale of the coherence length even for short-range disorder [16]; (ii) the parametric decoupling of the spectral gap from the mean pairing amplitude at large disorder: While the first remains relatively large the second decays to zero [16]. (iii) A parameter regime was predicted where the typical size of pairing amplitude is increased as compared to the clean limit, so disorder has a pronounced tendency to *enhance* superconductivity. The mechanism was explored in 3D near the Anderson transition [39,40] but also in 2D samples with short- and long-range interactions [42–44]. Several predictions are broadly consistent with numerical results obtained on a honeycomb lattice [17]. (iv) At large interactions the coherence length was reported to exhibit a nonmonotonous behavior with increasing disorder strength [24].

Despite the progress the current situation is not fully satisfying: On the one hand, computational studies of BdG-Hamiltonians have been limited to system sizes L that do not allow us to study the most interesting regime of length scales where the coherence length exceeds the lattice spacing: $\xi \gg a$. While analytical approaches, on the other hand, operate in this regime, they rely on partial self-consistency in order to become tractable.

Motivated by this observation, we investigate as a first application of our technology the BdG-problem of disordered superconductors focussing on s -wave pairing in thin films at $T = 0$. The full parameter plane of disorder, W , and interaction, U , is considered in which we study the distribution function and autocorrelations of the local gap function, $\Delta(\mathbf{r})$, as our main observable. Our computational machinery allows us to cover the full parameter space from the extreme regimes, which have been addressed computationally before to the analytically tractable weak coupling limit. In this effort we observe the formation of islands in large regions of the parameter space, as far as we know for the first time, on mesoscopic scales considerably exceeding the lattice constant. Regimes are included with parameters relatively close to the one where strong inhomogeneity has been observed experimentally [45].

Our observation might indicate that islands play a crucial role for the stability of the superconducting phase in actual experiments. Namely, islands imply localized Cooper pairs and therefore a diminishing of the phase stiffness. In other words, islands go together with enhanced phase fluctuations that destabilize long-range superconducting order. This connection between island formation and stability has been emphasized before [16,25].

Calculating the autocorrelation function of the spectral gap, $|\Delta(\mathbf{q})|^2$, we can extract a characteristic inverse length scale $\xi^{-1}(W, U)$ with the physical meaning of a correlation length. We study ξ within the full phase diagram. Interestingly, concomitantly with island formation we find an enhanced BdG coherence length. A similar observation if only at very large interaction strength, $U = 5$, has been made by Seibold *et al.* [24]. To what extent the enhancement of ξ is an artifact of mean-field theory that is removed when adding phase fluctuations remains to be seen.

Finally, like earlier authors [16] we also pay a special attention to the sensitivity of the behavior of computational observables to approximations made in the self-consistency procedure. We find that the island formation when observed in moderate parameter regions is a characteristic hallmark of full self-consistency; it escapes partial (“energy-only”) self-consistent schemes. We conclude that the renormalization of wave functions associated with full self-consistency will probably be an important ingredient of a qualitative theory of the superconductor-insulator transition.

II. MODEL AND IMPLEMENTATION

A. BdG-treatment of Hubbard model

We consider the attractive- U Hubbard model [46] on the square lattice in two dimensions within the mean-field (BdG-type) approximation:

$$\begin{aligned}\hat{H}_{\text{BdG}} &= \hat{H}_0 + \hat{H}_I \\ \hat{H}_0 &= -t \sum_{(i,j),\sigma} \hat{c}_{i,\sigma}^\dagger \hat{c}_{j,\sigma} + \text{H.c.} + \sum_{i=1, \sigma}^{N_{\text{bf}}} (V_i - \mu) \hat{n}_{i,\sigma} \quad (5) \\ \hat{H}_I &= -\frac{U}{2} \sum_{i=1, \sigma}^{N_{\text{bf}}} n(\mathbf{r}_i) \hat{n}_{i,\sigma} - \sum_{i=1}^{N_{\text{bf}}} \Delta(\mathbf{r}_i) \hat{c}_{i,\uparrow} \hat{c}_{i,\downarrow} + \text{H.c.},\end{aligned}$$

with local occupation number $n(\mathbf{r}_i) = \sum_{\sigma} \langle \hat{n}_{i,\sigma} \rangle$, pairing amplitude $\Delta(\mathbf{r}_i) = \langle \hat{c}_{i,\downarrow}^\dagger \hat{c}_{i,\uparrow}^\dagger \rangle$, $U > 0$ and random potential $V_i \in [-W, W]$ drawn from a box distribution. We mention that here we have identified the number of basis functions N_{bf} with the number of lattice sites exploiting rotational invariance in spin space. We employ periodic boundary conditions and work at $T = 0$; the chemical potential μ is adjusted to fix the filling factor to $n = \sum_{i=1, \sigma}^{N_{\text{bf}}} \frac{\langle \hat{n}_{i,\sigma} \rangle}{N_{\text{bf}}} = 0.875$ [47].

This Hamiltonian is diagonalized by a Bogoliubov transformation,

$$\gamma_{n,\uparrow}^\dagger = \sum_{i=1}^{N_{\text{bf}}} [u_n(\mathbf{r}_i) \hat{c}_{i,\uparrow}^\dagger + v_n(\mathbf{r}_i) \hat{c}_{i,\downarrow}], \quad (6)$$

$$\gamma_{n,\downarrow}^\dagger = \sum_{i=1}^{N_{\text{bf}}} [u_n(\mathbf{r}_i) \hat{c}_{i,\downarrow}^\dagger - v_n(\mathbf{r}_i) \hat{c}_{i,\uparrow}]. \quad (7)$$

The particle and hole wave functions $u_n(\mathbf{r}_i)$ and $v_n(\mathbf{r}_i)$ are determined solving the BdG equation

$$\begin{pmatrix} h & \Delta \\ \Delta^* & -h^* \end{pmatrix} \begin{pmatrix} u_n(\mathbf{r}_i) \\ v_n(\mathbf{r}_i) \end{pmatrix} = \epsilon_n \begin{pmatrix} u_n(\mathbf{r}_i) \\ v_n(\mathbf{r}_i) \end{pmatrix}, \quad (8)$$

where the physical sector corresponds to $\epsilon_n > 0$ and

$$hu_n(\mathbf{r}_i) = -t \sum_{\hat{\delta}} u_n(\mathbf{r}_i + \hat{\delta}) + \left[V_i - \mu - U \frac{n(\mathbf{r}_i)}{2} \right] u_n(\mathbf{r}_i), \quad (9)$$

$$\Delta u_n(\mathbf{r}_i) = \Delta(\mathbf{r}_i) u_n(\mathbf{r}_i); \quad (10)$$

the sum over $\hat{\delta}$ is over the lattice sites neighboring \mathbf{r}_i . The scf conditions for the density $n(\mathbf{r})$ and the gap function $\Delta(\mathbf{r})$ read

$$\Delta(\mathbf{r}_i) = U \langle \hat{c}_{i,\uparrow}^\dagger \hat{c}_{i,\downarrow} \rangle = U \sum_n u_n(\mathbf{r}_i) v_n^*(\mathbf{r}_i), \quad (11)$$

$$n(\mathbf{r}_i) = \sum_{\sigma} \langle \hat{c}_{i,\sigma}^\dagger \hat{c}_{i,\sigma} \rangle = 2 \sum_n |v_n(\mathbf{r}_i)|^2. \quad (12)$$

We assume self-consistency to be attained if the relative change per iteration cycle in $\Delta(\mathbf{r}_i)$ is at each site \mathbf{r}_i smaller than α . Typical values we take are $\alpha = 0.1\%$, 0.5% , 1% , 3% . Note that the average change α_{avg} per iteration cycle is much smaller than α , e.g., for a typical sample at moderate disorder $W = 2$ we have $\alpha_{\text{avg}} = 0.014\%$, 0.025% , 0.05% , 0.1% .

B. Matrix-free implementation of sparse-matrix vector product

To speed up a single self-consistency iteration we optimize the Chebyshev expansion. Its performance critical part constitutes of the recursive action of the Hamiltonian on a basis vector, Eq. (4). An implementation of the sparse-matrix vector product custom tailored to our system is crucial for an optimal performance. The sparse-matrix vector multiplication is memory bound, i.e., the performance is limited by the time it takes to fetch data from memory. For this reason we devised a self-written “matrix-free” matrix vector product that outperforms standard state-of-the-art sparse-matrix vector multiplication libraries.

The idea is the following: Conventional sparse matrix packages keep all nonzero elements, i.e., value and index, in memory. Matrix-free implementations become efficient if

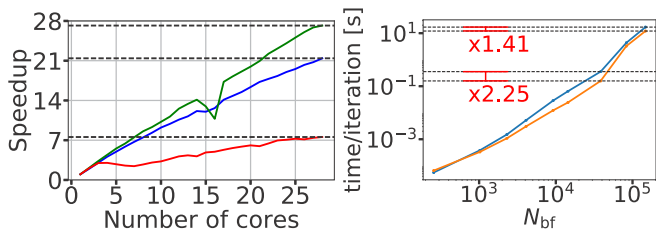


FIG. 1. Benchmarking intranode parallelization and code performance. Left: Speedup with the number of cores per process for different system sizes. The performance dips (green and blue: near 16; red: near 7, 14, 21, 25) with the rising number of cores we assign to a hardware issue related to caching [parameters: $L = 96$ (blue, center trace), $L = 192$ (green, top trace), and $L = 288$ (red, bottom trace).] Right: Performance check comparing the matrix-free implementation (orange, bottom trace) with standard `mkl_sparse_d_mm` of the MKL Sparse BLAS library (blue, top trace). One iteration corresponds to one sparse matrix-vector product. The ratio of the timings of the MKL and matrix-free algorithms is shown in red at $L = 192$ and $L = 384$.

many of the nonzero elements have identical values storing only the different values that occur.

With matrix-free implementations the graph of the Hamiltonian has to be hard-coded into the matrix-vector product routine. For our Hamiltonian the amount of memory load operations of matrix data is reduced by a factor of 6 reflecting the number of nonzero elements per row of \hat{H}_{BDG} . In addition, the integer indices corresponding to the matrix graph do not have to be loaded. Altogether, this leads to a reduction of data to be loaded by a factor of 9 [48]. We mention that recently a library has been made available that automatizes the implementation of such a matrix-free matrix-vector product for a given Hamiltonian [49].

C. Improved convergence of scf cycle

We improve the code performance by reducing the number of iterations needed until the convergence of the scf cycle. The main idea applies, e.g., when scanning the parameter space at fixed U for increasing disorder strength W . At strength W_1 a converged solution Ψ_1 is found for a given disorder realization. Thereafter, a sample at larger strength $W_2 > W_1$ is generated by rescaling the disorder in the first sample by a factor of W_2/W_1 . Then Ψ_1 will be used to initialize the scf cycle for the second sample.

D. Scaling and design considerations

As almost all runtime is spent on the recursive matrix vector products, the code lends itself very well to being split in an efficient low-level (i.e., C) kernel embedded in a high-level (i.e., python) code that implements the rest of the self-consistency cycle in a convenient way with negligible loss of performance. The kernel has been optimized for both threading and vectorization. In Fig. 1 we show benchmarks performed on a compute node with two 14-core Haswell Xeon Processor E5-2697 v3; we monitor the time spent for performing a single sparse matrix-vector product. Figure 1 (left) illustrates the efficiency of our intranode (OpenMP)

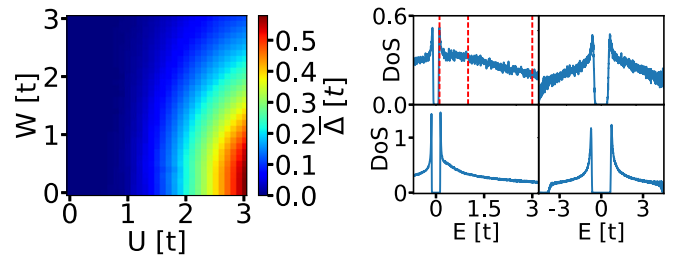


FIG. 2. Left: Disorder averaged gap $\overline{\Delta(\mathbf{r})}$ in the U - W parameter plane. Parameters: $L = 64$; $N_E = 500$, $\alpha = 1\%$. Right: Density of states for typical samples shown at four characteristic points. The red lines indicate the energies at which the LDoS is investigated in Fig. 3. Parameters: $W = 0.5$ (bottom) and 1.5 (top) and $U = 1.5$ (left) and 3.0 (right), $L = 192$; $N_C = 6144$, $\alpha = 3\%$.

parallelization. For the investigated system sizes $L < 288$ the memory-bound runtime limit is not yet reached as is evidenced by the high speedup through parallelization. This makes it very advantageous to perform calculations in this size regime, where parallelization can still be utilized effectively. Figure 1 (right) compares our matrix-free implementation with the standard MKL. As is seen from the data, our matrix-free implementation is advantageous already at system sizes as small as $N_{bf} \sim 1000$ sites. Note that at such small system sizes even full diagonalization routines can compete. As a technical remark we mention that, in principle, the matrix-free code should always be faster as compared to MKL implementations. The crossover size originates from our decision to use python as a platform, which leaves an interface to a C-based kernel. This interface is plagued with a small overhead that becomes negligible beyond the crossover size.

An additional level of parallelism is obtained by running the expansion of different basis vectors independently on different nodes. The average over the disorder ensemble is performed via farming. This internode parallelization scales almost perfectly.

III. RESULTS: BDG-STUDY OF DISORDERED SUPERCONDUCTORS

A. Mesoscopic fluctuations of LDoS and local gap function

As a first application of our technology, we investigate statistical properties of $\Delta(\mathbf{r})$ and of the local density of states (LDoS), $\rho(\mathbf{r}, E)$, throughout the U - W plane. To give a first impression we display in Fig. 2 (left) the gap function averaged over a suitable ensemble of disordered samples, $\overline{\Delta(U, W)}$; the overline indicates the ensemble average with N_E disorder configurations, typically $N_E \approx 700$ – 800 samples. The data have been obtained on a square lattice and should be compared with an analogous plot produced on the honeycomb lattice by Potirniche *et al.* [17]. The gap enhancement seen for the honeycomb lattice is not reproduced in Fig. 2. This is somewhat surprising, perhaps, because the phenomenon on the honeycomb lattice has been interpreted in terms of analytical results from quantum-field theory [44], which also should apply to the square lattice.

Also displayed in Fig. 2 (right) is the density of states, $\rho(E) = \int_{L^2} d\mathbf{r} \rho(E, \mathbf{r})$, calculated for four samples in

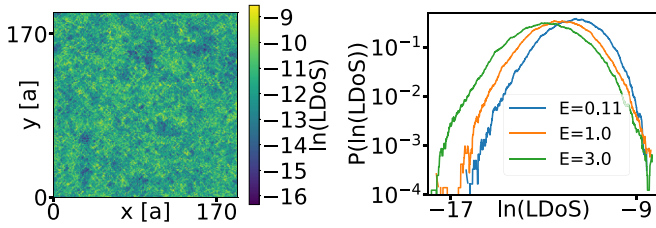


FIG. 3. Distribution of the local density of states (LDoS), $\rho(E, \mathbf{r})$. Left: Spatial distribution for a typical sample at peak energy of DoS ($E = 0.11$, cf. Fig. 2). Right: Corresponding distribution function of LDoS, \mathcal{P}_{ld} , at energies $E = 0.11, 1.0, 3.0$ illustrating the flow of the distribution with E . In Fig. 2 the corresponding DoS can be found (parameters: $W = 1.5, U = 1.5$; energy resolution 0.013 ; $N_C = 6144, \alpha = 3\%$).

representative regions of the parameter plane. At weak disorder the spectral gap and the coherence peaks are readily identified. Notice that only in the limit of weak disorder do the spectral gap and $\bar{\Delta}$ scale with each other [16].

To characterize the statistical properties of physical observables we focus in the following on autocorrelation and distribution functions. We will compare numerical findings with predictions from analytical theories and, in particular, study the sensitivity of qualitative results on modifications in the scf conditions.

1. Distribution functions of LDoS and local gaps

a. LDoS. We begin the statistical analysis with the spatial fluctuations of the LDoS, $\rho(E, \mathbf{r})$. Figure 3 (left) displays an example showing how the LDoS is spatially distributed over a typical sample with moderate disorder and interaction, $W \gtrsim U \gtrsim 1$. The logarithmically broad distribution of the LDoS is readily identified. The corresponding distribution function, \mathcal{P}_{ld} , is displayed in Fig. 3 (right). It takes a log-normal form, already familiar for disordered films with size smaller than a localization length, see, e.g., Eq. (4.101) in Mirlin [50].

With increasing energy the distribution shifts to smaller values, which is merely reflecting the decrease of the DoS $\rho(E)$, also seen in Fig. 2 (right). In contrast, the width of $\rho(E, \mathbf{r})$ is seen to grow. We assign this growth to the fact that the LDoS constitutes an average taken over a fixed-size energy window η . The number of eigenfunctions in the averaging window is estimated as $\rho(E)\eta L^2$ and therefore changes in energy if the DoS does. It is larger for energies near the coherence peak as compared to the bulk and for that reason the width of \mathcal{P}_{ld} should be expected to be reduced.

The LDoS distribution has been investigated analytically at temperatures above the critical temperature T_c [51]. Our observations are broadly consistent with these results, since it is reported that the distribution develops a pronounced non-Gaussian character on decreasing the temperature. For a more quantitative comparison, simulations at finite temperatures are required that are underway [52].

b. Local order parameter. The logarithmically broad distribution of the LDoS is concomitant with a similarly broad distribution of the local gap function \mathcal{P}_{lg} (Fig. 4). The evolution of the latter function with interaction strength is very interesting. As long as disorder, W , and interaction, U ,

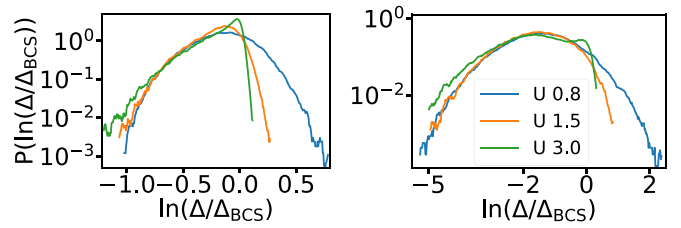


FIG. 4. Distribution of the local gap function, \mathcal{P}_{lg} , with interaction strength U for a typical sample with $L = 192$ at weak disorder, $W = 0.5$ (left plot), and stronger disorder, $W = 2.0$ (right plot). As reference energy the pairing amplitude of the clean system, $\Delta_{\text{BCS}}(U)$ has been chosen (parameters: $U = 0.8, 1.5, 3.0$; $N_C = 8192, 3072, 1024, \alpha = 0.1\%$).

are weak the distribution of the local order parameter is close to Gaussian and in this sense roughly following the statistics of the LDoS, see Fig. 4 (left). The typical value is seen to be very close to the pairing amplitude of the clean system, $\Delta_{\text{BCS}}(U)$. However, with growing U the weight of untypically large values of Δ is seen to be suppressed rapidly, while the weight of untypically small values is rather resilient.

For increasing disorder and weak U more and more sites develop a pairing well below the clean limit, $\Delta(\mathbf{r}) \ll \Delta_{\text{BCS}}$, consistent with observations made in Ref. [16]. Eventually, the shoulder is seen to dominate [Fig. 4 (right)] and the distribution \mathcal{P}_{lg} becomes bimodal. It features a peak near Δ_{BCS} and a logarithmically distributed background. The bimodal shape of \mathcal{P}_{lg} is apparent also from Lemarié *et al.* [22] where it is seen at very large interaction, $U = 5$.

2. Autocorrelations of gap function and coherence length

We consider the disorder averaged spatial autocorrelator $\Phi_{\text{lg}}(\mathbf{q}) = |\overline{\Delta(\mathbf{q})}|^2$ of the pairing function $\Delta(\mathbf{r})$ in Fourier space. At weaker disorder the correlation function displays a peak at $(\pi/a, \pi/a)$, Fig. 5. It originates from us choosing the filling fraction 0.875 which is close to the commensurate value unity and thus should be seen as a signature of the square lattice; it disappears at stronger disorder, e.g., at $W = 2.5$. The same signature manifests in Fig. 6 where we show $\Phi_{\text{lg}}(\mathbf{q})$ along two directions in \mathbf{q} space, $(\pi/a, 0)$ and $(\pi/a, \pi/a)$: As already obvious from Fig. 5, at wave numbers of order of the inverse lattice spacing, a^{-1} , and low W the correlator exhibits

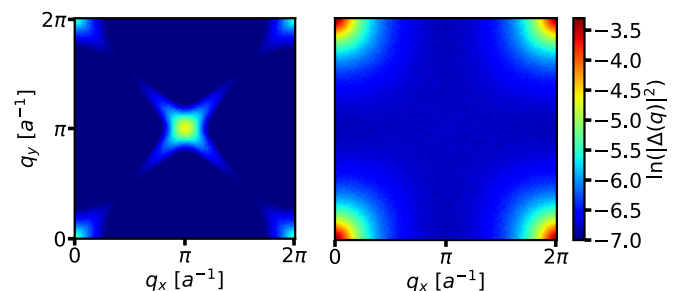


FIG. 5. The gap autocorrelation function $\Phi_{\text{lg}}(\mathbf{q}) = |\overline{\Delta(\mathbf{q})}|^2$ in logarithmic representation for $L = 192$ and $U = 1.5$ at two values of disorder, $W = 0.5$ (left) and $W = 2.5$ (right); $N_E \approx 900\text{--}1000$, $\alpha = 0.1\%$, $N_C = 1024$.

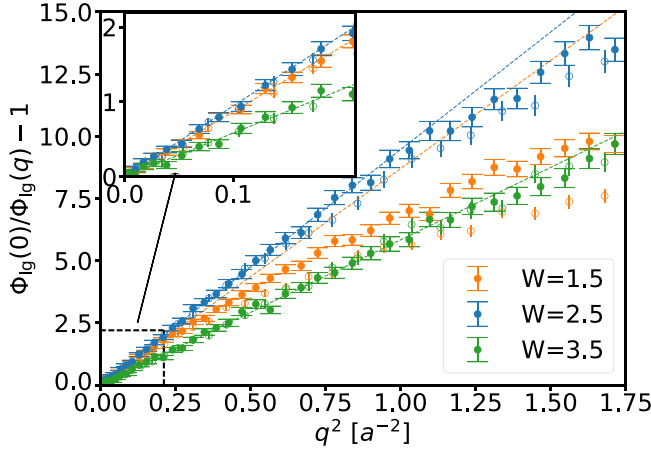


FIG. 6. The nontrivial part of the inverted normalized gap auto-correlation function $\Phi_{\text{lg}}^{-1}(\mathbf{q}) = |\Delta(\mathbf{q})|^{-2}$ evolving with W at fixed U . Φ_{lg} is shown averaged over equivalent directions $(\pi/a, 0)$, $(0, \pi/a)$ (full symbols) and $(\pi/a, \pi/a)$, $(\pi/a, -\pi/a)$ (open). The inset shows a blowup of the small wave-number regime where open and closed symbols collapse, so all traces are isotropic (parameters: $U = 1.5$, $W = 1.5$ (orange), 2.5 (blue), 3.5 (green), $L = 192$; $N_C = 1024$, $N_E \approx 600$ – 1000 , $\alpha = 0.5\%$).

pronounced deviations from isotropy reflected by the collapse of open and closed symbols.

Notwithstanding anisotropy at $q \approx a^{-1}$, in the limit of small wave numbers $q \ll a^{-1}$ the correlator $\Phi_{\text{lg}}(q)$ is isotropic and with good accuracy we have

$$\frac{\Phi_{\text{lg}}(0)}{\Phi_{\text{lg}}(q)} = 1 + (q\xi)^2 + \dots, \quad (13)$$

where $\Phi_{\text{lg}}(0) := \Phi_{\text{lg}}(q \rightarrow 0)$, given for different W in Fig. 7. $\Phi_{\text{lg}}(q)^{-1}$ behaves nearly quadratically over the whole momentum range where $\Phi_{\text{lg}}(q)$ exhibits isotropic behavior. Both the increase of $\Phi_{\text{lg}}(0)$ [as approximated by $\Phi_{\text{lg}}(\pi/L, 0)$] with disorder and the characteristic length ξ have been displayed in Fig. 7.

To attain ξ we have used a linear fit of $\Phi_{\text{lg}}(0)/\Phi_{\text{lg}}(q^2)$ in the isotropic regime. As with the range of this regime also the number of data points increases considerably with W , the uncertainty, i.e., the size of the error bars, of ξ is seen to

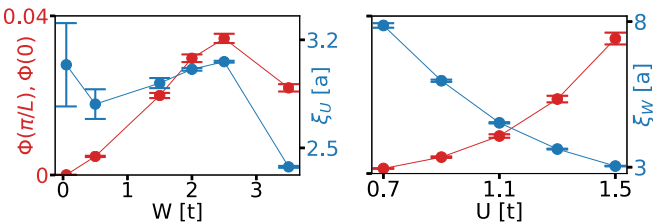


FIG. 7. Variation of $\Phi_{\text{lg}}(\pi/L, 0)$ and $\Phi_{\text{lg}}(0)$ (red) and the correlation length, ξ (blue), with increasing disorder (left, $U = 1.5$) and increasing interaction (right, $W = 2.5$). $\Phi_{\text{lg}}(0)$ coincides with $\Phi_{\text{lg}}(\pi/L)$ within the symbol size as portrayed here. The error bars depict the ensemble average error. The uncertainty due to cutoff α for ξ_W is discussed in the Appendix [parameters (left): $N_C = 1024$, $N_E \approx 600$ – 1000 , $\alpha = 0.1\%$; parameters (right): $N_C = 16\,384$, $N_E \approx 500$, $\alpha = 3\%$].

decrease with rising W in Fig 7 (left). $\xi_U(W)$ exhibits a local nonmonotonicity on its way from the clean to the dirty limit; the nonmonotonic decay is readily seen also from the original data Fig. 6. This peculiar behavior should be interpreted in connection with the formation of superconducting islands. It occurs in the same parameter range and may relate to a percolation transition. Our data show that the nonmonotonous shape, which was found in Ref. [24] albeit at unrealistically strong interactions $U = 5$, carries over all the way into the physically more relevant regime of intermediate parameter values.

IV. IMPACT OF SELF-CONSISTENCY

We return to a central theme of our interest, which is the impact of self-consistency on the calculation of physical observables.

A. Partial (energy-only) self-consistency scheme

The full BdG problem is specified by the set of equations (8)–(12). It is highly complicated, e.g., because the scf conditions (11) and (12) are nonlinear. As is frequently done in such situations, the full scf problem is replaced by a simplified variant exhibiting partial self-consistency.

Various possibilities for such simplifications are conceivable. The scheme we here describe is inspired by analytical calculations performed by Feigelman *et al.* [39,40]. The overall procedure can be considered a generalization of BCS theory that allows for an inhomogeneous order parameter. To bring the self-consistency requirement into the familiar BCS form, additional approximations besides the mean-field decoupling are necessary.

We here derive equations for partial self-consistency starting from the mean-field Hamiltonian Eq. (5). We express the field operators employing as a basis the eigenstates $\psi_l(\mathbf{r}_i)$ of the noninteracting part of \hat{H}_{BdG} , i.e., \hat{H}_0 :

$$\hat{d}_{l,\sigma} = \sum_{i=1}^{N_{\text{bf}}} \hat{c}_{i,\sigma} \psi_l(\mathbf{r}_i), \quad \hat{d}_{l,\sigma}^\dagger = \sum_{i=1}^{N_{\text{bf}}} \hat{c}_{i,\sigma}^\dagger \psi_l(\mathbf{r}_i). \quad (14)$$

The corresponding eigenvalues of ψ_l are denoted ξ_l and will be measured with respect to the Fermi-energy E_F . Expressing \hat{H}_{BdG} in \hat{d} , \hat{d}^\dagger we obtain

$$\begin{aligned} \hat{H}_{\text{BdG}} = & \sum_{l=1,\sigma}^{N_{\text{bf}}} \xi_l \hat{d}_{l,\sigma}^\dagger \hat{d}_{l,\sigma} + \sum_{l,m,n,o,\sigma} M_{lmno} \langle \hat{d}_{l,\sigma}^\dagger \hat{d}_{n,\sigma} \rangle \hat{d}_{m,\sigma}^\dagger \hat{d}_{o,\sigma} \\ & - U \sum_{l,m,n,o=1}^{N_{\text{bf}}} M_{lmno} \langle \hat{d}_{n,\downarrow} \hat{d}_{o,\uparrow} \rangle \hat{d}_{l,\uparrow}^\dagger \hat{d}_{m,\downarrow}^\dagger + \text{H.c.}, \end{aligned} \quad (15)$$

where an abbreviation

$$M_{lmno} = \sum_i \psi_l^*(\mathbf{r}_i) \psi_m^*(\mathbf{r}_i) \psi_n(\mathbf{r}_i) \psi_o(\mathbf{r}_i), \quad (16)$$

has been introduced.

The main approximate step in partial self-consistency is to neglect all terms with more than two indices

$$M_{lmno} = \begin{cases} M_{ln}, & \text{if } l = m \text{ and } n = o \\ 0, & \text{otherwise} \end{cases}, \quad (17)$$

together with the Hartree term. The simplified mean-field Hamiltonian then reads

$$\hat{H}_{\text{BdG}}^s = \sum_{l=1, \sigma}^{N_{\text{bf}}} \xi_l \hat{d}_{l, \sigma}^\dagger \hat{d}_{l, \sigma} + \sum_{l=1}^{N_{\text{bf}}} \Delta_l \hat{d}_{l, \uparrow}^\dagger \hat{d}_{l, \downarrow}^\dagger + \text{H.c.}, \quad (18)$$

with an s -wave pairing strength

$$\Delta_l = -U \sum_{m=1}^{N_{\text{bf}}} M_{lm} \langle \hat{d}_{m, \uparrow} \hat{d}_{m, \downarrow} \rangle. \quad (19)$$

The Hamiltonian (18) is structurally equivalent to the BCS Hamiltonian in the sense that the kinetic term and Δ_l are diagonal in the same (real-space) basis; Cooper pairs form within a Kramer's doublet. The corresponding BCS gap equation reads

$$\Delta_l = \frac{U}{2} \sum_{m=1}^{N_{\text{bf}}} M_{lm} \frac{\Delta_m}{\sqrt{\Delta_m^2 + \xi_m^2}}. \quad (20)$$

Converting back to real-space we have

$$\Delta(\mathbf{r}_i) = \frac{U}{2} \sum_{l=1}^{N_{\text{bf}}} \frac{\Delta_l}{\sqrt{\Delta_l^2 + \xi_l^2}} \psi_l^2(\mathbf{r}_i). \quad (21)$$

The advantage of the partial (“energy-only”) scf scheme is that the pairing amplitude can be calculated solely from the eigenstates and eigenvalues of the noninteracting reference Hamiltonian \hat{H}_0 . This comes at the expense of ignoring changes in the wave functions related to pairing and the inhomogeneous Hartree shift.

B. Effects of self-consistency schemes on the local-gap distribution

We compare the results of full and energy-only self-consistency schemes for the local pairing amplitude $\Delta(\mathbf{r})$ for the Anderson problem in 2D and 3D.

1. Two dimensions

Figure 8 shows a spatial distribution of $\Delta(\mathbf{r})$ as obtained for typical sample at intermediate interaction and three disorder values. The calculation with full self-consistency, Fig. 8 (left) column exhibits a clear tendency toward the formation of superconducting islands. In contrast, with energy-only self-consistency, right column, a rather homogeneous speckle pattern is found missing any indications of island formation. Hence, already by inspecting individual samples we expect that distribution functions of physical observables will depend in a qualitative way on the applied scf scheme in broad parameter regions.

In order to highlight the effect of screening, we have displayed in Fig. 8 also the results of an intermediate scf scheme. It operates in an energy-only mode but adopts for the disorder the effective single particle potential (“screened” potential) as it is obtained as a result from the full scf calculation. As is seen from Fig. 8, in the center column the first indications of islands emerge, but the contrast is still largely underestimated. This result underlines the importance of full consistency in the scf-procedure.

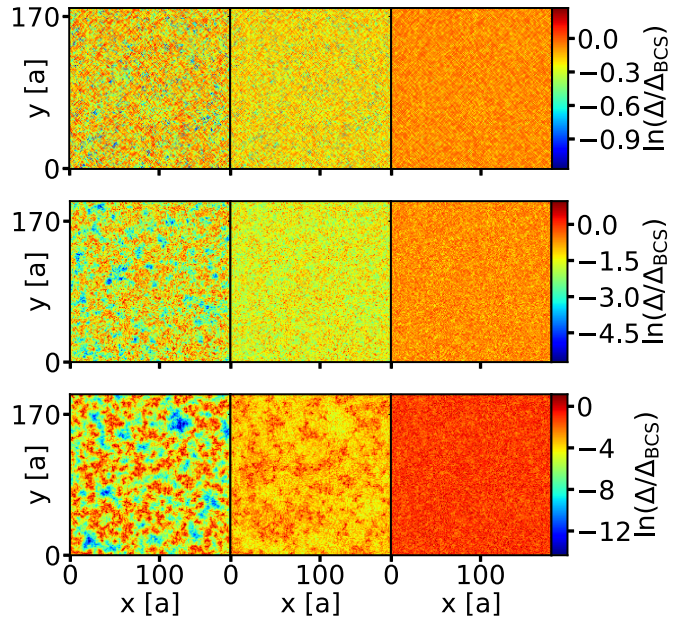


FIG. 8. Evolution of islands with disorder increasing from top to bottom, $W = 0.5, 2.0, 3.5$. Different self-consistency schemes are compared. Left column: Full self-consistency. Center column: Energy-only self-consistency with inhomogeneous Hartree shift. Calculation is done with the single-particle (“screened”) potential as it results from the full scf calculation, left. Right column: Energy-only scheme. The energy-only data have been calculated employing full diagonalization (parameters: $U = 1.5$; $N_C = 1024$, $\alpha = 0.5\%$).

2. Three dimensions

In analogy to the 2D case, we compare the results of full and energy-only self-consistency schemes for the local pairing amplitude $\Delta(\mathbf{r})$ in 3D. All 3D results have been computed with a conventional full diagonalization solver.

In the noninteracting 3D Anderson problem there is a phase transition at a critical disorder strength W_c , where all states become localized. For a disorder strength below W_c there exists an energy E_c , the mobility edge, which separates a fully localized band from a band of extended states. We note that as the Anderson Hamiltonian is symmetric around $E = 0$ this is also true for the mobility edge. We refer to B. Bulka [53] for the phase diagram.

Our interest is in the importance of self-consistency in the presence of attractive on-site interactions close to the mobility edge in the insulating band. For comparability with authors that have considered an energy-only approach in this context before [40], we choose a Gaussian disorder distribution

$$p(V_i) = \frac{1}{\sqrt{2\pi}W} \exp\left[-\frac{V_i^2}{2W^2}\right] \quad (22)$$

of the random onsite energies V_i in Eq. (9).

Figure 9 shows the spatial distribution of $\Delta(\mathbf{r})$ of a typical sample as obtained for moderate interaction and disorder strength and chemical potential in the localized band. The chosen filling factor $n = 0.3$ corresponds to a chemical potential of $\mu \approx -6$ in the fully self-consistent case. The mobility edge without interactions is located at $E_c \approx -5.5$ for the disorder strength $W = 4$ that is considered here [53].

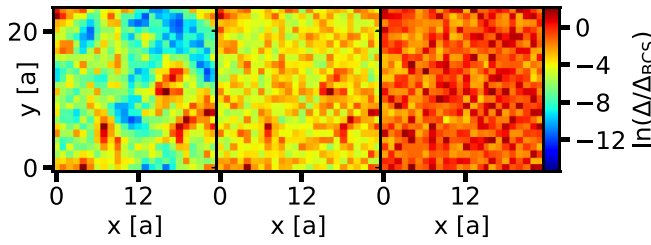


FIG. 9. Corresponding plot to Fig. 8 in 3D. A representative 2D slice of a sample is shown. The data have been calculated employing full diagonalization (parameters: $L = 24$, $W = 4.0$, $U = 2.5$, $n = 0.3$; $\alpha = 0.5\%$).

As in the 2D case, the field obtained within the fully self-consistent calculation shows a pronounced formation of islands [Fig. 9 (left)]. The energy-only scheme in analogy to our 2D results exhibits a rather homogeneous spatial distribution [Fig. 9 (right)]. The results of the energy-only scheme with “screened” potential shown in Fig. 9 (center) again show first indications of island development with dramatically underestimated contrast. This highlights the importance of full self-consistency also in 3D.

To what extent the conclusions of earlier theoretical works that consider this scenario [39,40] are affected remains to be seen.

C. Effects of self-consistency on gap autocorrelator

Figure 10 shows data analog to Fig. 6, now with energy-only self-consistencies. As is obvious already from individual sample, Fig. 8, the contrast parametrized by $\Phi_{\text{lg}}(0)$ is much smaller as compared to the case of full self-consistency given in Fig. 5. As one would expect from Fig. 8, the contrast with screened potential [Fig. 8 (right)] exceeds the bare scheme [Fig. 8 (left)] considerably.

The most striking and perhaps unexpected feature, however, is a qualitative difference. In the full scf calculation, $\Phi_{\text{lg}}(q)$ follows Eq. (13) and exhibits a well-defined parabolic shape in the vicinity of small wave numbers. This feature is not reproduced within energy-only schemes. The bare scheme does not exhibit an appreciable curvature up to $q \approx a^{-1}$, so the coherence length with energy-only self-consistency, ξ_{co} ,

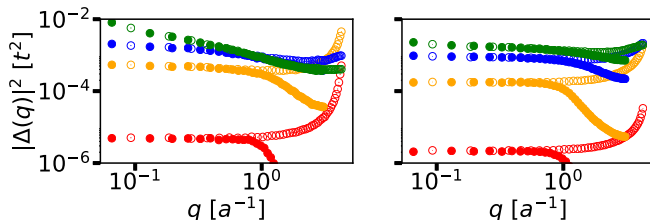


FIG. 10. Gap autocorrelation function $\Phi_{\text{lg}}(\mathbf{q}) = \overline{|\Delta(\mathbf{q})|^2}$ calculated employing two different energy-only self-consistency schemes. Φ_{lg} is shown along directions $(\pi/a, 0)$ (full symbols) and $(\pi/a, \pi/a)$ (open); traces for four different disorder values are shown, $W = 0.05, 0.5, 1.5, 2.5$, from bottom to top. Left: Energy-only self-consistency with screened potential. Right: Energy-only self-consistency (parameters: $U = 1.5$, $L = 96$; $N_E = 1000$, $\alpha = 0.1\%$).

vanishes. In contrast, within the screened scheme $\Phi_{\text{lg}}(q)$ does not show clear saturation at small wave numbers within the range of q values accessible. We thus interpret these results as a strong indications that wave-function renormalization as it occurs within the full scf scheme is crucial for understanding those aspects of qualitative physics that hinge on long-range spatial correlations.

V. CONCLUSIONS

We have implemented an efficient solver of scf Hamiltonians that is based on the kernel polynomial method. An application to disordered s -wave superconducting films has been presented that employs the Bogulubov-deGennes approximation. The statistical properties of the local density of states and of the local gap function $\Delta(\mathbf{r})$ have been studied. In this context our computational machinery proves useful since system sizes can be accessed significantly exceeding the ones that have been achieved in the earlier work. We thus can study the crossover in disorder strength W and interaction strength U from the strongly coupled into the perturbative regime, where analytical methods apply and can provide conceptual insights.

Along this way three key observations have been made. (i) Superconducting islands form in large regions of the U - W phase space and thus appear to be a typical encounter already at intermediate interaction and disorder strength. (ii) Presumably related to island formation, the (mean-field) correlation length exhibits a nonmonotonous variation when sweeping from very weak to strong disorder. (iii) Island formation is a hallmark of wave-function renormalization in the sense that islands do not form with partial (“energy-only”) self-consistency. To investigate into possible consequences of this observation for analytical treatments of the superconductor-insulator transition we leave as a topic for future research.

As a concluding remark we note that the BdG-Anderson problem and the associated ensemble of self-consistent random Hamiltonians is a particular representative of a very large class of random matrices that satisfy a self-consistency constraint (“scf ensembles”). Presumably, because of the considerable challenges that such ensembles imply for analytical and computational treatments very little is known about them. We take the observations that have been reported for the BdG ensembles, in this work as well as by the earlier authors, as a strong indication that much is there to be discovered.

ACKNOWLEDGMENTS

We are grateful to Soumya Bera, Igor Burmistrov, Christoph Strunk, and Thomas Vojta for numerous inspiring discussions; we also express our gratitude to Ivan Kondov for sharing mathematical and computational expertise. Support from the DFG under Grants No. EV30/11-1, No. EV30/12-1, and No. SFB-1277 (Projects A03 and B01) is acknowledged. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. [54] for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre [55]. This work was performed on the supercomputer ForHLR funded by the Ministry

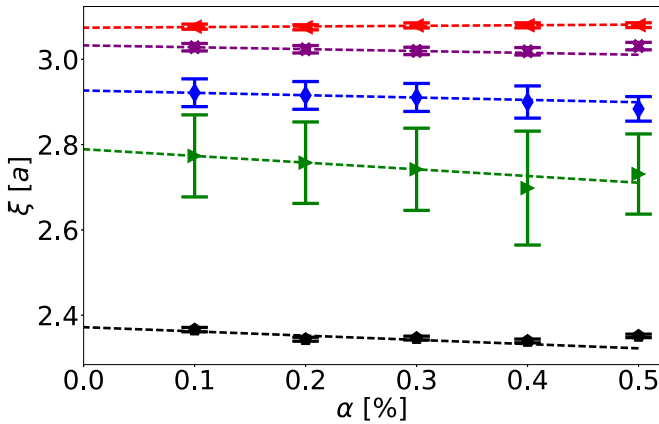


FIG. 11. Development of ξ with cutoff α for disorder strengths from bottom to top $W = 3.5$ (black), 0.5 (green), 1.5 (blue), 2.0 (purple), and 2.5 (red), error bars depict the uncertainty stemming from the ensemble average. The dashed lines show a linear fit accounting for the three smallest α values (parameters: $U = 1.5$, $L = 192$; $N_E \approx 600$ – 1000).

of Science, Research and the Arts Baden-Württemberg and by the Federal Ministry of Education and Research.

APPENDIX

Self-consistency cutoff discussion

In Fig. 11 the dependence of ξ on α at fixed W is shown. The data demonstrate good convergence behavior of

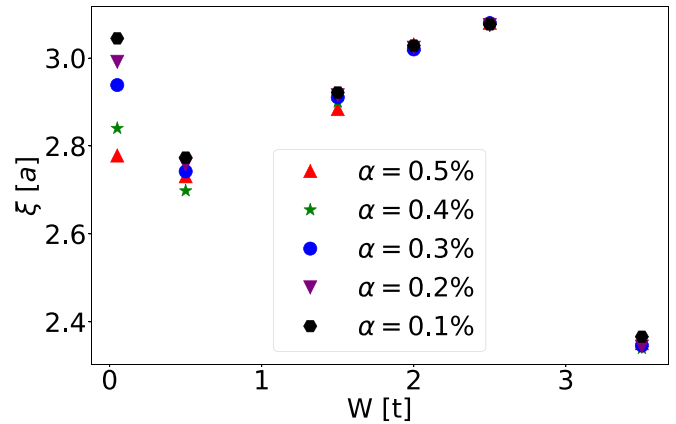


FIG. 12. Replot of the data Fig. (11) to illustrate the (converged) variation of ξ with W .

ξ in terms of the cutoff parameter α ; in particular, the α dependency of ξ is seen to be small as compared to the variation with W . Figure 12 replots the data shown in Fig. 11, so the evolution of ξ with W is more clearly illustrated. In particular, it is seen that the nonmonotonic behavior is very well converged in the cutoff α . The stronger change of ξ with α seen at low disorder strengths, e.g., at $W = 0.05$ and 0.5 , is related to the fact that the distribution of local values, $\Delta(\mathbf{r})$, is narrow at small W . In this case, the convergence requirement allowing for a maximal percentage α of change from cycle to cycle has implications for a substantial fraction of all sites; with broad distributions, convergence of most sites will be much better than α .

- [1] M. R. Zirnbauer, *J. Math. Phys.* **37**, 4986 (1996).
- [2] A. Altland and M. R. Zirnbauer, *Phys. Rev. B* **55**, 1142 (1997).
- [3] P. Heinzner, A. Huckleberry, and M. R. Zirnbauer, *Commun. Math. Phys.* **257**, 725 (2005).
- [4] A. D. Mirlin, Y. V. Fyodorov, F.-M. Dittes, J. Quezada, and T. H. Seligman, *Phys. Rev. E* **54**, 3221 (1996).
- [5] A. D. Mirlin and F. Evers, *Phys. Rev. B* **62**, 7920 (2000).
- [6] S.-R. Eric Yang, A. H. MacDonald, and B. Huckestein, *Phys. Rev. Lett.* **74**, 3229 (1995).
- [7] B. Huckestein and M. Backhaus, *Phys. Rev. Lett.* **82**, 5100 (1999).
- [8] F. Epperlein, M. Schreiber, and T. Vojta, *Phys. Rev. B* **56**, 5890 (1997).
- [9] M. Amini, V. E. Kravtsov, and M. Müller, *New J. Phys.* **16**, 015022 (2014).
- [10] H.-J. Lee and K.-S. Kim, *Phys. Rev. B* **97**, 155105 (2018).
- [11] L. Covaci, F. M. Peeters, and M. Berciu, *Phys. Rev. Lett.* **105**, 167006 (2010).
- [12] E. G. Carnio, N. D. Hine, and R. A. Römer, *Physica E* **111**, 141 (2019).
- [13] Y. Harashima and K. Slevin, *Int. J. Mod. Phys. Conf. Ser.* **11**, 90 (2012).
- [14] Y. Harashima and K. Slevin, *Phys. Rev. B* **89**, 205108 (2014).
- [15] A. Ghosal, M. Randeria, and N. Trivedi, *Phys. Rev. Lett.* **81**, 3940 (1998).
- [16] A. Ghosal, M. Randeria, and N. Trivedi, *Phys. Rev. B* **65**, 014501 (2001).
- [17] I.-D. Potirniche, J. Maciejko, R. Nandkishore, and S. L. Sondhi, *Phys. Rev. B* **90**, 094516 (2014).
- [18] A. Ghosal and H.-Y. Kee, *Phys. Rev. B* **69**, 224513 (2004).
- [19] D. Chakraborty, R. Sensarma, and A. Ghosal, *Phys. Rev. B* **95**, 014516 (2017).
- [20] R. Ganguly, I. Roy, A. Banerjee, H. Singh, A. Ghosal, and P. Raychaudhuri, *Phys. Rev. B* **96**, 054509 (2017).
- [21] G. Seibold, L. Benfatto, C. Castellani, and J. Lorenzana, *Phys. Rev. Lett.* **108**, 207004 (2012).
- [22] G. Lemarié, A. Kamlapure, D. Bucheli, L. Benfatto, J. Lorenzana, G. Seibold, S. C. Ganguli, P. Raychaudhuri, and C. Castellani, *Phys. Rev. B* **87**, 184509 (2013).
- [23] T. Cea, D. Bucheli, G. Seibold, L. Benfatto, J. Lorenzana, and C. Castellani, *Phys. Rev. B* **89**, 174506 (2014).
- [24] G. Seibold, L. Benfatto, C. Castellani, and J. Lorenzana, *Phys. Rev. B* **92**, 064512 (2015).
- [25] Y. Dubi, Y. Meir, and Y. Avishai, *Nature* **449**, 876 (2007).
- [26] L. Hedin, *Phys. Rev.* **139**, A796 (1965).
- [27] F. Bechstedt, *Many-Body Approach to Electronic Excitations* (Springer, Berlin, 2015).
- [28] M. J. van Setten, F. Weigend, and F. Evers, *J. Chem. Theory Comput.* **9**, 232 (2013).

- [29] L. P. Kadanoff and G. Baym, *Quantum Statistical Mechanics* (W.A. Benjamin, New York, 1962).
- [30] Once the scf field was found an update of h has to be computed. This computation is efficiently dealt with by employing the fast Fourier transformation (FFT) and therefore not critical. With FFT an operation that formally is $\mathcal{O}(N_{\text{br}}^2)$ can be downgraded to $\mathcal{O}(N_{\text{br}} \ln N_{\text{br}})$.
- [31] A. Weiße, G. Wellein, A. Alvermann, and H. Fehske, *Rev. Mod. Phys.* **78**, 275 (2006).
- [32] Y. Nagai, Y. Ota, and M. Machida, *J. Phys. Soc. Jpn* **81**, 024710 (2012).
- [33] Y. Nagai, Y. Shinohara, Y. Futamura, Y. Ota, and T. Sakurai, *J. Phys. Soc. Jpn.* **82**, 094701 (2013).
- [34] T. I. Baturina, A. Y. Mironov, V. M. Vinokur, M. R. Baklanov, and C. Strunk, *Phys. Rev. Lett.* **99**, 257003 (2007).
- [35] B. Sacépé, C. Chapelier, T. I. Baturina, V. M. Vinokur, M. R. Baklanov, and M. Sanquer, *Phys. Rev. Lett.* **101**, 157006 (2008).
- [36] K. Bouadim, Y. L. Loh, M. Randeria, and N. Trivedi, *Nat. Phys.* **7**, 884 (2011).
- [37] T. Cea, C. Castellani, G. Seibold, and L. Benfatto, *Phys. Rev. Lett.* **115**, 157002 (2015).
- [38] Y. L. Loh, M. Randeria, N. Trivedi, C.-C. Chang, and R. Scalettar, *Phys. Rev. X* **6**, 021029 (2016).
- [39] M. V. Feigel'man, L. B. Ioffe, V. E. Kravtsov, and E. A. Yuzbashyan, *Phys. Rev. Lett.* **98**, 027001 (2007).
- [40] M. V. Feigel'man, V. Kravtsov, L. Ioffe, and E. Cuevas, *Ann. Phys-New York* **325**, 1390 (2010).
- [41] M. V. Feigel'man and L. B. Ioffe, *Phys. Rev. B* **92**, 100509(R) (2015).
- [42] I. S. Burmistrov, I. V. Gornyi, and A. D. Mirlin, *Phys. Rev. Lett.* **108**, 017002 (2012).
- [43] I. S. Burmistrov, I. V. Gornyi, and A. D. Mirlin, *Phys. Rev. Lett.* **111**, 066601 (2013).
- [44] I. S. Burmistrov, I. V. Gornyi, and A. D. Mirlin, *Phys. Rev. B* **92**, 014506 (2015).
- [45] B. Sacépé, T. Dubouchet, C. Chapelier, M. Sanquer, M. Ovadia, D. Shahar, M. V. Feigel'man, and L. Ioffe, *Nat. Phys.* **7**, 239 (2011).
- [46] J. Hubbard, *Proc. R. Soc. Lond. Ser. A* **276**, 238 (1963).
- [47] The filling factor is chosen in a way to be close to half-filling, which favors a high particle-hole overlap, while avoiding the ground-state degeneracy of the superconducting state with a charge density wave state at half-filling.
- [48] The data type for values is double and for the indices is integer. Note that the speed-up to be expected from the matrix-free implementation is less than a factor of 9. This is because not only the matrix but also the basis vectors have to be loaded from memory, so the reduction of memory load operations also depends on how many basis vectors are acted on in parallel.
- [49] A. Pieper, G. Hager, and H. Fehske, [arXiv:1708.09689v2](https://arxiv.org/abs/1708.09689v2) (2017).
- [50] A. D. Mirlin, *Phys. Rep.* **326**, 259 (2000).
- [51] I. S. Burmistrov, I. V. Gornyi, and A. D. Mirlin, *Phys. Rev. B* **93**, 205432 (2016).
- [52] M. Stosiek, I. Burmistrov, and F. Evers (unpublished).
- [53] B. K. B. Bulka, M. Schreiber, *Z. Phys. B* **66**, 21 (1987).
- [54] www.gauss-centre.eu.
- [55] www.lrz.de.