

# Quantum impurity models using superpositions of fermionic Gaussian states: Practical methods and applications

Samuel Boutin  and Bela Bauer

Station Q, Microsoft Corporation, Santa Barbara, California 93106 USA



(Received 10 May 2021; revised 30 July 2021; accepted 3 August 2021; published 24 August 2021)

The coherent superposition of nonorthogonal fermionic Gaussian states has been shown to be an efficient approximation to the ground states of quantum impurity problems [Bravyi and Gosset, *Commun. Math. Phys.* **356**, 451 (2017)]. We propose a practical approach for performing a variational calculation based on such states. Our method is based on approximate imaginary-time equations of motion that decouple the dynamics of each Gaussian state forming the ansatz. It is independent of the lattice connectivity of the model and the implementation is highly parallelizable. To benchmark our variational method, we calculate the spin-spin correlation function and Rényi entanglement entropy of an Anderson impurity, allowing us to identify the screening cloud and compare to density matrix renormalization group calculations. Secondly, we study the screening cloud of the two-channel Kondo model, a problem difficult to tackle using existing numerical tools.

DOI: [10.1103/PhysRevResearch.3.033188](https://doi.org/10.1103/PhysRevResearch.3.033188)

## I. INTRODUCTION

Quantum impurity models—systems of a few strongly interacting degrees of freedom coupled to a large bath of noninteracting fermions—constitute an important class of problems in condensed matter physics. Despite the small number of interacting modes involved, this class of problems can exhibit rich many-body physics phenomena. The archetypical phenomenon is the Kondo effect, where even weak interactions can lead to strong nonperturbative corrections to the ground state [1]. Such models also appear as effective models in many embedding methods, such as dynamical mean-field theory [2] that solve extended quantum many-body systems by approximately mapping them to quantum impurity problems.

Over the years, various numerical methods have been developed to tackle quantum impurity problems. A particularly successful approach is Wilson’s numerical renormalization group (NRG) [1] and its extensions [3], which have allowed to study this class of problems in the thermodynamic limit. A related set of variational methods, based on the density matrix renormalization group (DMRG) [4] has also been used extensively and compared to NRG [5,6]. Finally, quantum Monte Carlo methods have been successfully applied to systems where the sign-problem is mild, see, e.g., Refs. [7,8]. Despite these methods being very powerful, they each come with their limitations. For example, NRG is limited by an exponential scaling in the number of degrees of freedom of the impurity and the number of channels in the noninteracting bath; DMRG scales more favorably in the size of the impurity, but (since it does not exploit the noninteracting nature of the bath) retains

an exponential scaling with the number of channels in the bath, rendering it very challenging to study, e.g., mesoscopic problems with several leads. Quantum Monte Carlo methods, on the other hand, typically suffer from severe sign problems for multiorbital systems.

A natural question is whether a well-chosen class of variational states could exploit the structure of quantum impurity models to circumvent the limitations of these established approaches. For example, this was pursued for some models using a non-Gaussian variational approach in Refs. [9–11]. Recently, this question was affirmatively answered more generally by proving that the ground states of any quantum impurity problem can be approximated by a superposition of nonorthogonal fermionic Gaussian states [12]. This is obviously the case when the number of states in the superposition—which we will refer to as the rank of the ansatz—grows exponentially with the full system size, as the states then form a complete many-body basis. More interestingly, the rigorous mathematical bounds of Ref. [12] demonstrate that the minimal rank to obtain a good approximation of the ground state scales only with the size of the impurity and the desired precision, while being independent of the size of the bath.

Having chosen this superposition of Gaussians (SGS) ansatz for the study of quantum impurity models, the challenge is to devise practical algorithms to perform numerically efficient computations. Here, we focus on the problem of finding the lowest-energy state within the variational manifold. Multiple generic approaches exist for performing the energy minimization within a variational space, such as gradient descent and imaginary time evolution [13,14]. In particular, formal solutions for related ansatzes were originally developed under the name of resonating Hartree-Fock [15–17]. However, due to the large number of variational parameters in the SGS ansatz and the presence of several nonlinear constraints, their numerical implementation can become prohibitively costly for large systems.

---

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

In this paper, we propose a simpler and numerically less costly path towards energy minimization within the variational manifold of SGS states based on several key approximations to the imaginary time equations of motion. First, at each step we project the dynamics onto the subspace orthogonal to the one spanned by the current set of Gaussian states forming the SGS. Furthermore, we alternate the evolution of the coefficients of the coherent superposition of states and the (normalized) Gaussian states themselves. This allows us to decouple the equations of motion for each Gaussian state at each step in the evolution. While the projection of these equation of motions onto the variational manifold does not exactly correspond to imaginary time evolution, we show that under this evolution, the energy is nonincreasing. The variational state therefore converges to a local energy minimum within the manifold and can thus be used to study ground-state properties of quantum impurity models (we note that a guarantee on convergence to a global minimum typically cannot be given for variational algorithms).

To illustrate the power of this approach, we apply it to two canonical impurity models: the single-impurity Anderson model [18] and the two-channel Kondo model [19]. The former has been studied using a variety of methods and is well-understood both analytically and numerically, thus allowing us to confirm the validity of our method. We find that using comparable computational resources, our method is able to achieve an error in the ground-state energy that is about one order of magnitude better than DMRG. The two-channel Kondo model, on the other hand, is much more challenging to study numerically, and real-space correlation functions for fermionic leads had previously eluded numerical simulations. Instead, prior numerical studies have required either mapping to related problems in the same universality class [20] or focusing on quantities that can be calculated using the local dynamics of the impurity [21].

The SGS ansatz considered in this paper can be seen as a generalization of the generalized Hartree-Fock (GHF) method [22,23], which aims to find the approximate ground state of a system using a variational minimization over the field of fermionic Gaussian states. While exact for a noninteracting system, GHF corresponds to a mean-field approximation for interacting systems. Additional approaches related to our paper are known as multicomponent or multireference Hartree-Fock-Bogoliubov methods, where a superposition of Gaussian states is obtained by projecting a single symmetry-breaking Gaussian state in a chosen symmetry sector [24,25]. We note furthermore that our method can be viewed as a generalization of configuration interaction (CI) methods commonly used in quantum chemistry calculations; however, in contrast to typical CI methods, we use *nonorthogonal* Gaussian states rather than orthogonal Slater determinants and optimize the choice of Gaussian basis states along with their coefficients, rather than using a fixed basis (for example, based on the Hartree-Fock solution or single-particle excitations).

The remainder of this paper is structured as follow. In Sec. II, we set the notation and describe the structure of the SGS ansatz. In Sec. III, we describe a generic minimization procedure for finding a variational approximation to the ground state and its numerical implementation. In Sec. IV, as a first demonstration of the method, we study the screening

cloud of a single impurity Anderson model and benchmark our results using DMRG. Finally, in Sec. V we extend the calculations of the previous section by considering the two-channel Kondo model.

## II. ANSATZ AND PROBLEM STRUCTURE

We start by describing the structure of fermionic quantum impurity models and of the SGS ansatz. We also introduce the covariance matrix formalism for fermionic Gaussian states, which will be used throughout this paper.

### A. Generic quantum impurity model

We consider a lattice model of  $N$  fermionic degree of freedoms with Hamiltonian  $H = H_2 + H_4$ . We choose to work within a formalism of Majorana operators, noting that any fermion problem (both with and without particle-number conservation) can be rewritten in this form. The noninteracting part of the Hamiltonian is given by

$$H_2 = i \sum_{k,l=1}^{2N} A_{k,l} c_k c_l, \quad (1)$$

where  $A$  is a real and skew-symmetric matrix, and  $c_k = c_k^\dagger$  is a set of  $2N$  Majorana operators that obey standard anticommutation relations  $\{c_j c_k\} = 2\delta_{j,k}$ . The interacting part of the Hamiltonian reads

$$H_4 = \sum_{k,l,p,q}^{2N} U_{k,l,p,q} c_k c_l c_p c_q, \quad (2)$$

where the rank-4 tensor  $U$  is skew-symmetric with respect to the exchange of any neighboring indices. The interaction involves at most  $M \ll N$  distinct Majorana operators making  $U$  a sparse tensor. We remark that no assumption with regards to the lattice connectivity is made in our model definition. Although we focus here and below on quartic interaction terms, there is no fundamental limitation to including interaction terms involving a larger number of operators.

The ground state of the quadratic Hamiltonian  $H_2$  will be a Slater determinant if  $H_2$  conserves the number of particles or, more generally, a fermionic Gaussian state [26] in the case where the  $U(1)$  symmetry is broken and only the parity of the number of particles is preserved. As Slater determinants constitute a subset of the fermionic Gaussian states, we will focus our discussion on Gaussian states. Although there might be a slight numerical overhead associated with working in this enlarged class of states, it has the advantage of naturally allowing the treatment of (mean-field) superconductivity.

### B. Covariance matrix formalism

We now briefly introduce key aspects of the covariance matrix formalism that we use to describe Gaussian states [12]. We defer technical details of the computations with covariance matrices to Appendix A.

Any fermionic Gaussian state  $|\phi\rangle$  obeys a Wick theorem and thus can be fully described by the expectation values of fermion bilinears. It is convenient to introduce the covariance

matrix (CM)

$$\Gamma_{k,l} = \frac{-i}{2} \langle \phi | [c_k, c_l] | \phi \rangle, \quad (3)$$

where  $k, l \in 1 \dots 2N$ . This matrix is real and skew-symmetric by construction. The expectation value of any product of Majorana operators can then be calculating as the Pfaffian of a submatrix of  $\Gamma$ . For a normalized pure state the elements of the CM are subject to the constraint  $\Gamma^2 = -\mathbb{1}$ .

Since the covariance matrices  $\Gamma$  are matrices of expectation values, they are invariant under a gauge transformation  $|\phi\rangle \rightarrow e^{i\theta}|\phi\rangle$ , with  $\theta$  a real number. For calculations involving multiple Gaussian states  $|\phi_\mu\rangle$  ( $\mu = 1, 2, \dots$ ), it is often necessary to fix this gauge freedom. Following Ref. [12], this can be achieved by choosing a reference state  $|\phi_0\rangle$  and taking  $\langle \phi_0 | \phi_\mu \rangle$  to be real and positive for all  $\mu$ . Overlaps of Gaussian states and matrix elements can be obtained using the respective covariance matrices of the states. See Appendix A and Ref. [12] for the relevant expressions.

### C. Sum of Gaussian states ansatz

Following Ref. [12], the variational ansatz considered in this paper is formulated as<sup>1</sup>

$$|\psi\rangle = \sum_{\mu=1}^R \lambda_\mu |\phi_\mu\rangle, \quad (4)$$

with  $\{|\phi_\mu\rangle\}$  a set of nonorthogonal Gaussian states,  $\lambda_\mu$  complex scalar amplitude, and  $R$  the rank of the ansatz. We emphasize that allowing for *nonorthogonal* states in Eq. (4) is crucial to the efficiency of the ansatz. This is because the linear combination of several Gaussian states is in general not Gaussian; therefore, one cannot orthogonalize the set of vectors  $|\phi_\mu\rangle$  while preserving their Gaussianity.<sup>2</sup>

The variational state defined in Eq. (4) is characterized by the set of covariance matrices and amplitudes  $\{\Gamma^\mu, \lambda_\mu | \mu = 1 \dots R\}$ . This corresponds to  $O(RN^2)$  variational parameters, subject to two normalization constraints. First, the normalization of the variational state requires

$$\sum_{\mu,\nu} \lambda_\mu \lambda_\nu^* G_{\mu,\nu} = 1, \quad (5)$$

where we introduce  $G_{\mu,\nu} = \langle \phi_\mu | \phi_\nu \rangle$  the Gram matrix characterizing the overlap between states. Second, we take each Gaussian state to be pure and normalized leading to the constraint on each covariance matrix  $(\Gamma^\mu)^2 = -\mathbb{1}$ . The sets of parameters obeying these constraints form the variational manifold. Our aim is to find the state  $|\psi_0\rangle$  that minimize the energy within this manifold.

Before turning to the full minimization problem, we first consider the following simpler problem. Given a set of Gaussian states  $\{|\phi_\mu\rangle\}$  we wish to find the amplitudes  $\{\lambda_\mu\}$ , which

minimize the energy  $E = \langle \psi | H | \psi \rangle$ . These optimal amplitudes, leading to a normalized state with the lowest energy within the subspace, can be obtained by diagonalizing the Hamiltonian projected onto the subspace spanned by the set of Gaussian states. This leads to the generalized eigenvalue problem [12,15]

$$h\lambda = EG\lambda \quad (6)$$

where  $h$  is the  $R \times R$  matrix with elements  $h_{\alpha,\beta} = \langle \phi_\alpha | H | \phi_\beta \rangle$ . In the case of orthogonal states  $G = \mathbb{1}$  and this reduces to a regular eigenvalue problem. This standard result will be at the core of the minimization approach introduced in Sec. III as we will alternate between updating the amplitudes by solving Eq. (6) and updating the covariance matrices assuming fixed amplitudes.

### III. PROJECTED EQUATIONS OF MOTION FOR ENERGY MINIMIZATION

A generic approach for finding the ground state of a quantum system is imaginary time evolution. Starting from an initial state  $|\psi\rangle$  and evolving according to the imaginary time Schrödinger equation

$$\partial_\tau |\psi\rangle = -[H - E_\psi] |\psi\rangle, \quad (7)$$

where  $E_\psi = \langle \psi | H | \psi \rangle$ , allows to reach the ground state in the  $\tau \rightarrow \infty$  limit as long as the initial state has finite overlap with the ground state. In the case where  $|\psi\rangle$  is a variational state, the equations of motion must be projected back onto the part of the variational manifold orthogonal to  $|\psi\rangle$  in order to best approximate the dynamics of the system and preserve the norm of the state. Different methods were introduced to perform this projection [27]. In the case of the time-dependent variational principle (TDVP), this projection requires the inversion of the Gram matrix of the tangent states (obtained by taking the derivative of  $|\psi\rangle$  with respect to each variational parameter), which can be very large [13]. For the parametrization considered in this paper, this does not appear to be a scalable approach as it would require repeated operations on matrices of dimension  $RN^2 \times RN^2$  leading to an  $O(R^2N^6)$  numerical complexity.

In this section, we instead derive simplified projections of the imaginary time equation of motion for the SGS states. Our approach can be understood as a parametrized energy descent. We derive equations of motion for the covariance matrices  $\Gamma^\mu$  as a function of an external parameter  $s$  such that the energy  $E(s) = \langle \psi(s) | H | \psi(s) \rangle$  decreases monotonically and converges to a local energy minima as  $s \rightarrow \infty$ . As we do not pretend the followed approach to be sufficient to recover the system dynamics, we denote the evolution parameter as  $s$  to distinguish the resulting equations from imaginary time (denoted  $\tau$  above).

#### A. Path of energy descent

As eluded to at the end of Sec. II, at any instant  $s$ , one can separate the Hilbert space in two instantaneous subspaces, where one (referred to as “parallel subspace” below) is spanned by the set  $\{|\phi_\mu(s)\rangle\}$  of Gaussian states forming the ansatz, and the other is the orthogonal complement

<sup>1</sup>We follow the convention that Greek indices run over the labels of the states forming the ansatz (i.e.  $\alpha, \beta \dots \in 1 \dots R$ ), while latin indices run over the Majorana operator labels (i.e.  $k, l, \dots \in 1, \dots 2N$ ).

<sup>2</sup>Of course, one may choose a (larger) set of  $S$  orthogonal Gaussian states  $|\omega_\nu\rangle$  such that  $\text{span}\{|\phi_\mu\rangle\} \subseteq \text{span}\{|\omega_\nu\rangle\}$ ; however, this would jeopardize the computational efficiency since in general  $S \gg R$ .

(“orthogonal subspace” below). The energy minimization in the parallel subspace is easily performed by choosing amplitudes satisfying Eq. (6) and we will thus focus here on the orthogonal subspace assuming fixed amplitudes.

We consider as a starting point the imaginary time Schrödinger equation. However, instead of projecting on a subspace orthogonal to the instantaneous state  $|\psi(s)\rangle$  leading to Eq. (7), we project onto a subspace orthogonal to the set  $\{|\phi_\mu(s)\rangle\}$  of Gaussian states

$$\partial_s |\psi(s)\rangle = -\Pi^\perp(s)H|\psi(s)\rangle, \quad (8)$$

where we have introduced the projector

$$\Pi^\perp(s) = \mathbb{1} - \sum_{\alpha,\beta} |\phi_\alpha(s)\rangle [G^{-1}(s)]_{\alpha,\beta} \langle \phi_\beta(s)|. \quad (9)$$

The inverse Gram matrix in the above equation ensures that  $(\Pi^\perp)^2 = \Pi^\perp$  and is necessary since the Gaussian states considered are generally nonorthogonal (but linearly independent, insuring that  $G$  is nonsingular). This equation of motion preserves the norm of the state since  $\Pi^\perp(s)|\psi(s)\rangle = 0$  leading to  $\partial_s \langle \psi(s)|\psi(s)\rangle = 0$ .

The projection in Eq. (8) constitutes a first approximation to the equations of motions and its justification is twofold. First, assuming optimal state amplitudes, any dynamics lowering the energy should be orthogonal to the instantaneous parallel subspace. However, the resulting equation of motion is only approximate due to the additional implicit projection onto the Gaussian state manifold. This projection will become explicit through the use of the Wick theorem in Sec. III B. Second, the projection to the orthogonal subspace ensures that the different Gaussian states do not collapse to a single state. Indeed, without it, all states would collapse to a mean-field approximation of the ground state independently.

Inserting the definition of the ansatz into the projected equation of motion Eq. (8), we obtain

$$\sum_\mu \lambda_\mu \partial_s |\phi_\mu\rangle = - \sum_\mu \lambda_\mu \Pi^\perp(s)H|\phi_\mu(s)\rangle. \quad (10)$$

In order to move forward, we decouple the equation of motions of the different Gaussian states by postulating that there exists an effective Hamiltonian  $B_\mu$  such that

$$\partial_s |\phi_\mu\rangle = -B_\mu |\phi_\mu\rangle, \quad (11)$$

and which satisfies Eq. (10). Upon inspection, one can find that

$$B_\mu = w_\mu \Pi^\perp H |\psi\rangle \langle \psi| + \text{H.c.} \quad (12)$$

decouples the equations, with  $w_\mu$  a complex scalar. Taking  $w_\mu = [\lambda_\mu \tilde{\lambda}_\mu^*]^{-1}$ , with the renormalized amplitude  $\tilde{\lambda}_\mu = \langle \phi_\mu | \psi \rangle$ , satisfies Eq. (10) confirming the validity of the decoupling scheme. Although this is not a unique choice, this approach ensures a monotonic decrease of energy, i.e.,  $\partial_s \langle \psi | H | \psi \rangle \leq 0$ , upon simultaneous integration of Eq. (11) for all states.

More generally, ensuring a path of energy descent upon the evolution of a given state  $|\phi_\mu\rangle$  leads to the constraint  $\text{Re}[w_\mu \tilde{\lambda}_\mu^* \lambda_\mu] > 0$ . Hence, taking  $w_\mu = \tilde{\lambda}_\mu \lambda_\mu^* / |\tilde{\lambda}_\mu \lambda_\mu|$ , possibly up to a real and positive multiplicative factor for each state,

is sufficient to decrease the energy of the variational state upon evolution.

This decoupling scheme is the second major approximation to the equations of motion used in this paper. While exact for a generic many-body basis, the decoupling is approximate in the case where the evolution is projected on a constrained variational manifold. As is shown numerically in Secs. IV and V, the optimization nevertheless converges towards the ground state.

## B. Equation of motion for the covariance matrices

In order for the above equations of motion to be useful, there must exist an efficient numerical implementation of them. We now derive the counterpart of Eq. (11) in the covariance matrix (CM) formalism.

Taking the derivative of the CM, as defined in Eq. (3), with respect to the evolution parameter  $s$  we obtain (taking  $k \neq l$ )

$$\partial_s \Gamma_{k,l}^\mu = -\langle \phi_\mu | i c_k c_l | \partial_s \phi_\mu \rangle + \text{c.c.}, \quad (13)$$

where c.c. denotes the complex conjugate. Inserting Eq. (11) and the definition of  $\Pi^\perp$  one obtains after some algebra

$$\begin{aligned} \partial_s \Gamma_{k,l}^\mu = & i \sum_\beta \lambda_\beta \tilde{\lambda}_\beta^* w_\mu \left[ \langle \phi_\mu | c_k c_l H | \phi_\beta \rangle \right. \\ & \left. - \sum_\gamma \langle \phi_\mu | c_k c_l | \phi_\gamma \rangle [G^{-1} h]_{\gamma,\beta} \right] + \text{c.c.} \end{aligned} \quad (14)$$

In the special case of a rank 1 ansatz, Eq. (14) falls back onto the equation of motion for imaginary time evolution in the GHF approximation as derived for example in Ref. [23].

We now specialize to the quantum impurity model with quartic interacting defined in Sec. II A. In order to rewrite the differential equation purely as a matrix equation, we introduce the complex skew-symmetric matrices [12]

$$\Delta^{\alpha,\beta} = [i(\Gamma^\alpha - \Gamma^\beta) - 2\mathbb{1}](\Gamma^\alpha + \Gamma^\beta)^{-1}, \quad (15)$$

which allows to compute easily matrix elements between different Gaussian states  $\langle \phi_\beta | c_k c_l | \phi_\alpha \rangle = i G_{\beta,\alpha} \Delta_{k,l}^{\alpha,\beta}$  ( $k \neq l$ ). More generally matrix elements involving  $n$  distinct Majorana operators are proportional to the Pfaffian of an  $n \times n$  submatrix of  $\Delta^{\alpha,\beta}$  (see Appendix A for complete expressions). Similarly to GHF, we introduce generalized Fock matrices

$$F_{k,l}^{\alpha,\beta} = A_{k,l} + 6 \sum_{m,n} U_{k,l,m,n} \Delta_{m,n}^{\alpha,\beta}, \quad (16)$$

which for  $\alpha = \beta$  ( $\Delta^{\alpha,\alpha} = \Gamma^\alpha$ ) falls back on the standard Fock matrix for GHF [23].

With these definition, the matrix differential equation for each CM takes the form

$$\begin{aligned} \partial_s \Gamma^\mu = & \sum_\beta \lambda_\beta \tilde{\lambda}_\beta w_\mu \left[ i 2 G_{\mu,\beta} [\Delta^{\beta,\mu}, F^{\beta,\mu}] \right. \\ & + 2 G_{\mu,\beta} (\Delta^{\beta,\mu} F^{\beta,\mu} \Delta^{\beta,\mu} + F^{\beta,\mu}) \\ & \left. - h_{\mu,\beta} \Delta^{\beta,\mu} + \sum_\gamma G_{\mu,\gamma} [G^{-1} h]_{\gamma,\beta} \Delta^{\gamma,\mu} \right] + \text{c.c.}, \end{aligned} \quad (17)$$

where the first and second line are respectively reminiscent of the equation of motion for real and imaginary time evolution of a Gaussian state [23]. The third line ensures normalization and cancels in the case  $R = 1$ . We have verified numerically that Eq. (17) preserves the norm and the purity of the state, which requires

$$(\partial_s \Gamma^\mu) \Gamma^\mu + \Gamma^\mu (\partial_s \Gamma^\mu) = 0, \quad (18)$$

due to the constraint  $\Gamma^2 = -\mathbb{1}$  for a pure state. Equation (17) constitutes one of the main results of this paper.

### C. Numerical implementation

In order to find a good approximation of the ground state, we alternately evolve the covariance matrices for a small step  $s \rightarrow s + \delta s$  and update the amplitudes by solving Eq. (6). The computational complexity of each of these steps is  $O(R^2 N^3)$  (assuming  $R \ll N$ ). Since in each step, each Gaussian state can be evolved separately, one can easily parallelize over them and achieve a significant speedup. Appendix B presents additional details on the numerical implementation.

As an aside, we note that considering  $P$  interaction terms with  $m > 4$  Majorana operators could, in the worst case, lead to additional operations of computational complexity  $O(PR^2 N^2 m^3)$ . However, in the case of a quantum impurity model where  $m, P \ll N$ , this should not affect the scaling of the overall computational complexity of the method. We note that the above estimate is an upper bound on the complexity obtained by calculating the right-hand side of Eq. (14) separately for each term of weight  $m$  and each matrix element. Taking advantage of the skew-symmetric matrix structure of the CM might allow further algorithmic improvement. For example, in the case of  $P = N^2$  quadratic operators ( $m = 2$ ) the naive estimate lead to complexity  $O(R^2 N^4)$  while the matrix formulation of Eq. (17) is reduced to  $O(R^2 N^3)$ .

In order to reduce the risk of converging to a local energy minima, we use in the remainder of this paper the following strategy. Starting from the mean field (GHF) solution we gradually increase the rank of the ansatz by 1 after every  $n$  iterations. The additional Gaussian state is chosen by applying a random special orthogonal transformation  $Q$  to one of the current states of the ansatz with a high amplitude. For  $n$  sufficiently large and  $\|Q - \mathbb{1}\|_F$  sufficiently small we have found numerically this approach to converge towards a good approximation of the ground state. This approach is needed due to our choice of optimal amplitudes at each step of the algorithm. If all states are added at once, there is a strong risk to converge to a solution where some of the amplitudes vanish, effectively leading to an approximate ground state of a lower rank.

## IV. ANDERSON IMPURITY AND THE SCREENING CLOUD

As a benchmark of the method, we study real-space properties of the single impurity Anderson model (SIAM) [18]. This well-known model was previously studied numerically using a multitude of methods including NRG [28] and DMRG [29,30]. We will use the latter method to benchmark our results and confirm their validity.

We consider a 1D lattice where the SIAM Hamiltonian takes the form  $H = H_0 + H_I$ , with

$$H_0 = -t \sum_{r=2}^{L-1} (d_{r-1}^\dagger d_r + \text{H.c.}) - \mu \sum_{r=1}^{L-1} n_r, \quad (19)$$

describing a bath of  $L - 1$  free fermions with hopping parameter  $t$  and chemical potential  $\mu$ , and the impurity Hamiltonian

$$H_I = -t' (d_0^\dagger d_1 + \text{H.c.}) + \epsilon_d n_0 + U n_{0,\uparrow} n_{0,\downarrow}. \quad (20)$$

Here,  $d^\dagger$  are fermionic creation operators,  $n_r = d_r^\dagger d_r = \sum_\sigma d_{r,\sigma}^\dagger d_{r,\sigma}$  is the number operator, and when omitted, spin indices are summed. The hopping between the impurity (site 0) and the first site of the lead is  $t'$  and  $U > 0$  is a repulsive interaction. We focus on the particle-hole symmetric point at half-filling where  $\epsilon_d = -U/2$  and  $\mu = 0$ .

In the strong interaction limit  $U > t'$ , charge fluctuations are suppressed on the impurity site creating an effective spin 1/2 impurity. The Hamiltonian then maps to the Kondo model with effective (weak) coupling strength  $J = 8\Gamma/U$ , where  $\Gamma = (t')^2/t$  is the broadening of the impurity energy level by the leads [31]. The ground state is then a singlet state formed by the impurity spin and the collective spin of a delocalized so-called *Kondo cloud* of electrons from the leads. One can associate to the cloud a length scale  $\xi_K \sim v_F/T_K$  where  $v_F$  is the Fermi velocity and

$$T_K \sim \sqrt{U\Gamma} \exp\left(-\frac{\pi U}{8\Gamma}\right) \quad (21)$$

is the Kondo temperature estimated using the perturbative renormalization group [32].

### A. Convergence and comparison to DMRG

As a first test of the convergence of the method, we compute the variance of the Hamiltonian in the variational ground state,

$$\delta_H^2 = \langle \psi | (H - \langle H \rangle)^2 | \psi \rangle, \quad (22)$$

where  $\langle H \rangle = \langle \psi | H | \psi \rangle$ . Figure 1(a) presents  $\delta_H^2$  as a function of the ansatz rank and for different number of iterations of the variational minimization (see Sec. III C) for a system size  $L = 500$ . The variance  $\delta_H^2$  should converge to zero as the variational state converges to an eigenstate of the system ( $\delta_H^2 = 0$  for any eigenstate of  $H$ ). We estimate in Fig. 1(b) the converged ground-state energy  $E^*$  using a linear extrapolation of the variational energy as  $\delta_H^2 \rightarrow 0$  for  $\delta_H^2 \lesssim 2 \times 10^{-5}$ .

As a second test of the validity of our variational SGS state results, Fig. 2 compares the energies obtained using the SGS ansatz to energies obtained using state-of-the-art DMRG simulations [33]. In both cases, we plot the results as a function of the refinement parameter of the ansatz, which is  $R$  for the SGS ansatz and the bond dimension  $m$  for DMRG. For the parameters considered, which were chosen to span similar computation times, the SGS-based method reaches lower energies than DMRG.

### B. Screening cloud—Spin-spin correlation

To better check the validity of the variational solution, we calculate the ground-state equal-time spin-spin correlation

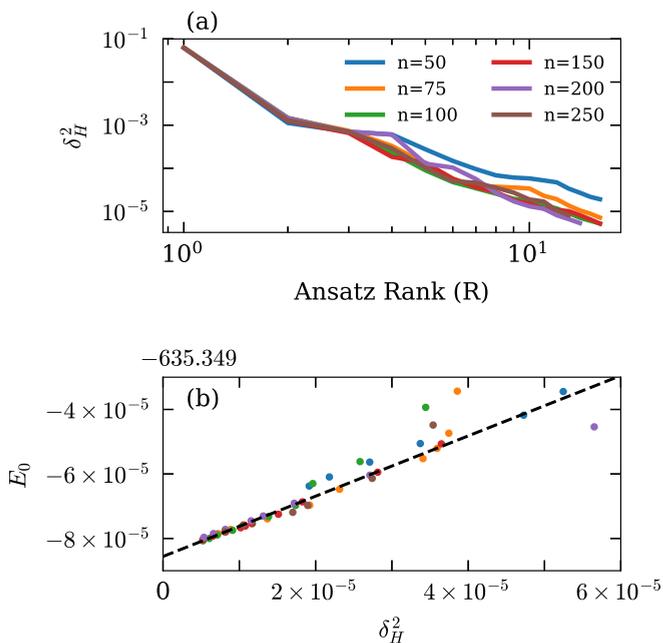


FIG. 1. (a) Convergence of the variance of the Hamiltonian  $\delta_H^2$  [Eq. (22)] as a function of the number of iterations  $n$  per Gaussian state (see main text). (b) Extrapolation using a linear fit (dashed-black curve) of points with  $\delta_H^2 < 2 \times 10^{-5}$ . The obtained intercept,  $E_0^* = -635.3490861$ , estimates the energy of the ground state up to a precision given by the standard error of the fit  $\sigma_{E_0^*} = 4 \times 10^{-7}$ . We consider a system of length  $L = 500$  at the particle-hole symmetric point ( $\mu = 0$ ,  $\epsilon_d = -U/2$ ) with interaction strength  $U = t = 1$  and coupling  $U/\Gamma = 5$ .

function between the impurity and the sites along the lead

$$C(r) = \langle \mathbf{S}(0) \cdot \mathbf{S}(r) \rangle, \quad (23)$$

where the spin operator is  $\mathbf{S}(r) = \frac{1}{2} \sum_{s,s'} d_{r,s}^\dagger \boldsymbol{\sigma}_{s,s'} d_{r,s'}$ .

From a low-energy expansion of the bath operators [34] away from the impurity ( $k_F r \gg 1$ , with  $k_F$  the Fermi wave

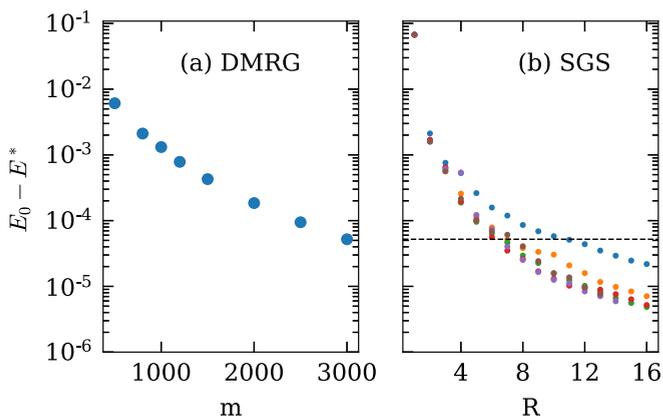


FIG. 2. Convergence of the ground-state energy obtained from (a) DMRG and (b) the SGS ansatz respectively as a function of the bond dimension  $m$  and the rank of the SGS ansatz. Dashed-black horizontal line is the lowest energy obtained using DMRG. The ground-state energy estimate  $E_0^*$  is obtained from the extrapolation of Fig. 1(b). Parameters and legend correspond to those of Fig. 1.

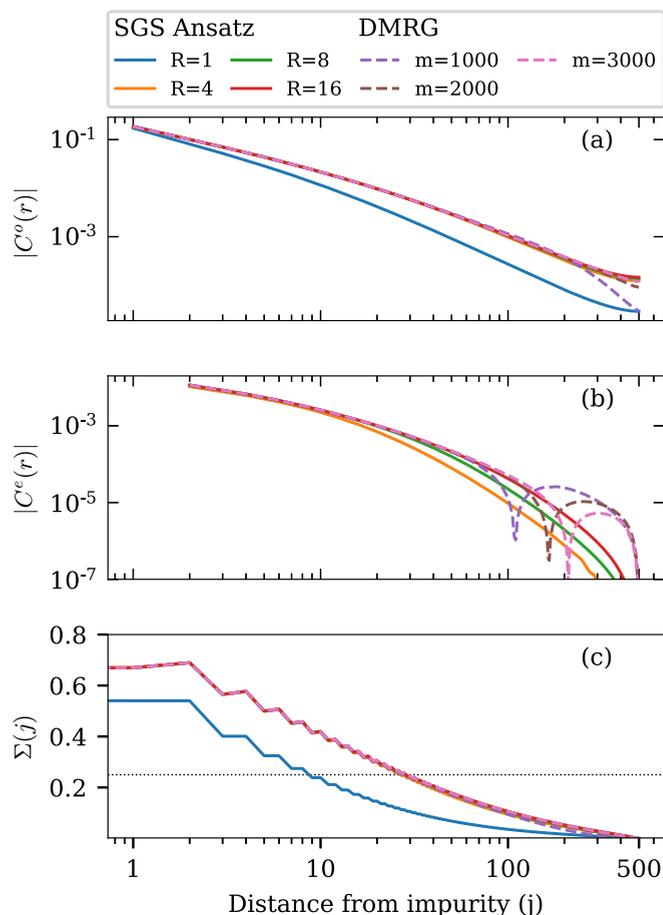


FIG. 3. Comparison of the spin-spin correlation function  $C(r)$  calculated using DMRG (dashed curves) and the SGS ansatz (solid curves). (a) Odd sites and (b) even sites. The mean-field solution ( $R = 1$ , blue-solid curve) is absent from panel (b) as  $C^e(r)$  is identically zero up to numerical precision. (c) Integrated correlation function as defined in Eq. (25). See Fig. 1 for parameters.

vector), the correlation function is expected to be the sum of uniformly decaying and oscillating functions

$$C(r) = C_U(r) + C_{2k_F}(r) \cos(2k_F r) \quad (24)$$

where, at zero temperature,  $C_U(r)$  and  $C_{2k_F}(r)$  are smoothly decaying functions. As we focus on a half-filled lattice where  $k_F = \pi/2a$  (with  $a$  the lattice constant), the correlation function is the sum of uniform and staggered ( $\cos(2k_F r) \rightarrow (-1)^r$ ) contributions.

Figure 3 compares the correlation functions computed using the approximate ground states obtained with SGS ansatz and DMRG. To plot more easily the highly oscillating function  $C(r)$ , we introduce the correlation function on the even (e) and odd (o) sites, denoted as  $C^e(r)$  and  $C^o(r)$ , respectively. For the odd sites [Fig. 3(a)], the uniform and staggered parts of the correlation function are both negative, leading to a larger amplitude than the even sites [Fig. 3(b)] where the contributions have opposite signs. Focusing first on the odd sites, the mean-field solution ( $R = 1$ , solid blue curve) differ qualitatively from the higher-precision DMRG results as expected. Modestly increasing the rank of the SGS ansatz, we recover the same behavior as high-precision DMRG. Small

discrepancies between the two methods are observed far from the impurity where the amplitude of the correlation function is small. These differences are consistent with the expected precision of both methods and we expect that further increasing the values of  $m$  and  $R$  would reduce these differences. Similarly, for the even sites, both methods agree close to the impurity. Notably, the DMRG results shows a change of sign of  $C^e(r)$  far from the impurity. This effect has been previously observed in other DMRG studies of the SIAM [29,30] and appears to disappear as the bond dimension of the MPS is increased. No such effect is observed for the SGS state, thus suggesting that this feature in the DMRG results is an artifact of the truncation of the MPS bond dimension.

As a final comparison between methods, we consider in Fig. 3(c) the integrated correlation function

$$\Sigma(x) = 1 + \sum_{r=1}^x \frac{C(r)}{C(0)}, \quad (25)$$

which allows characterization of the screening cloud of the impurity [29]. In the ground state, the total spin  $\mathbf{S}_{\text{tot}}^2$  is expected to be zero for even  $L$ , leading to the sum rule  $\Sigma(L-1) = 0$ . While this sum rule is not explicitly enforced, it is approximately respected and the violation converges towards zero as the rank of the ansatz is increased [ $\Sigma(L-1) \approx 2 \times 10^{-6}$  for  $R = 16$ ]. To conclude this section, we use the SGS ansatz to study the screening cloud of the Anderson impurity for a large system of  $L = 1000$  sites and different coupling strengths  $\Gamma = (t')^2/t$ . Figure 4(a) shows  $-C^o(r)$  for three different couplings between the impurity and the bath. In the Kondo regime, the correlation function  $C^o(r, \Gamma)$  is expected to collapse to a universal function  $\tilde{C}(r/\xi_K)$  through the relation  $\tilde{C}(r/\xi_K) = \xi_K(\Gamma)C^o(r, \Gamma)$  with  $\xi_K(\Gamma)$  a coupling-dependent length scale. Figure 4(b) present the scaling collapse of ten different curves for the odd distances from the impurity  $C^o(r)$ .<sup>3</sup> As expected from previous numerical studies [28] and analytical calculations [34] the scaled correlation function decays following a power law  $\tilde{C}(r/\xi_K) \propto (r/\xi_K)^{-\nu}$  with a crossover in the exponent  $\nu$  near  $r/\xi_K = 1$ . From power-law fits, we extract the exponents  $\nu \sim 0.76$  for  $r \ll \xi_K$  and  $\nu \sim 1.84$  for  $r \gg \xi_K$ . Given the sensitivity of the results of the fit on the range of parameters considered, this results are consistent with the previously established scenarios of a crossover from  $\nu = 1$  to  $\nu = 2$  near  $r \sim \xi_K$ .

Finally, Fig. 4(c) presents the impurity screening length scale as extracted using three different methods. First, the parameters  $\xi_K(\Gamma)$  were obtained from the scaling collapse of the correlation function  $C^o(r, \Gamma)$  (blue disks). Second, as a comparison, we also plot the length scale  $\xi_c(\Gamma)$  (orange crosses) at which the integrated correlation function falls below a given threshold:

$$\Sigma(\xi_c) = c \quad (26)$$

<sup>3</sup>The scaling collapse is obtained by minimizing the square of the distance on a log-log scale between the numerical data points and a polynomial fit to  $\xi_K(\Gamma)C^o(r)$  as a function of  $r/\xi_K(\Gamma)$  for all  $\Gamma$  values considered.

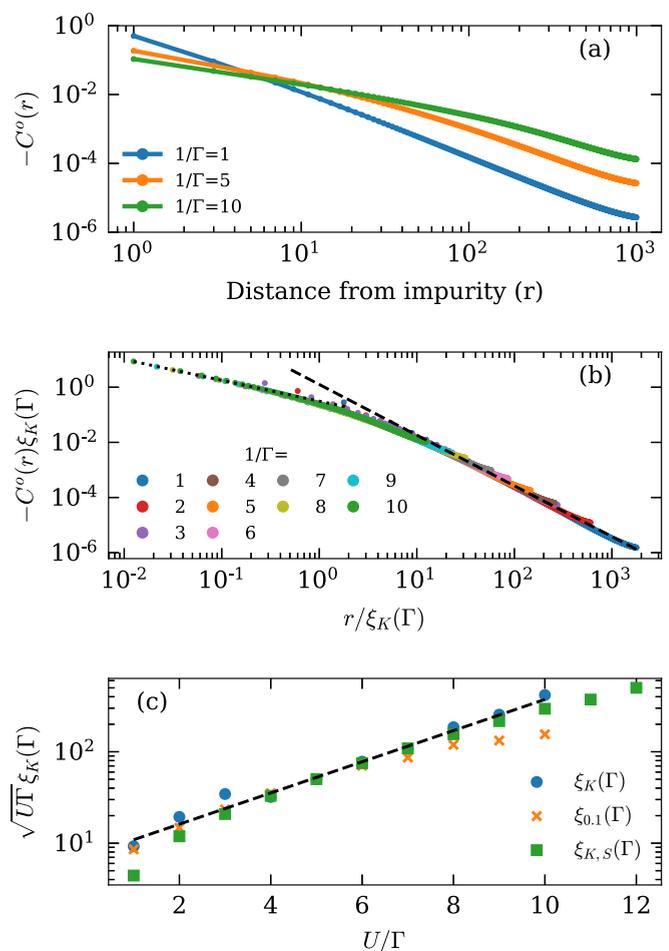


FIG. 4. (a) Spin-spin correlation function of a SIAM with  $L = 1000$  (ansatz Rank  $R = 14$ ) on odd sites for coupling strengths  $\Gamma = 1.0$  (dark-blue curve),  $\Gamma = 0.2$  (orange curve) and  $\Gamma = 0.1$  (green curve). (b) Scaling collapse of  $C^o(r)$ . Dashed (dotted)- black curve is a fit of the form  $A(r/\xi_K)^{-\nu}$  to the data with  $r > 10\xi_K$  ( $r < \xi_K/2$ ), with exponent  $\nu \approx 1.84$  ( $\nu \approx 0.76$ ). (c) Length scales  $\xi_K$  extracted from the scaling collapse of panel (b) (blue disks),  $\xi_{0.1}$  extracted from the integrated correlation function [orange crosses, cf Eq. (26)] and  $\xi_{K,S}$  extracted from the scaling collapse of the impurity entropy (green squares). See Sec. IV C for a discussion of the impurity entropy. Black dashed line is a fit of the expected functional form  $A \exp[\pi U/8\Gamma]$  to  $\sqrt{U\Gamma}\xi_K(\Gamma)$ . As  $\xi_K(\Gamma)$ ,  $\xi_{0.1}(\Gamma)$ , and  $\xi_{K,S}$  are related up to a scaling parameters, we set  $\xi_K(1/5) = \xi_{0.1}(1/5) = \xi_{K,S}(1/5)$ . To reduce finite-size effect we append to the end of the bath an additional  $L_W = 20$  sites with an exponentially decreasing hopping parameter  $t_n = \Lambda^{-n}t$  with  $\Lambda = 1.2$  and  $t = 1$ .

for a threshold parameter  $c \in (0, 1)$  [29]. Although simpler, this second method has the disadvantage of being sensitive to finite-size effects and convergence, as illustrated by the saturation of  $\xi_{0.1}$  (orange crosses) for smaller couplings. These effects are reduced in the case of the scaling collapse approach as it takes into account the correlation function calculated at all odd sites. As a third method, we show the length scale  $\xi_{K,S}(\Gamma)$  (green square) obtained from the scaling collapse of the impurity contribution to the entanglement entropy (near the impurity). This method will be described in the following section. As scaling collapse methods determine the screening

length only up to a global prefactor, we scale the data sets such that all three methods result in the same length scale for  $1/\Gamma = 5$ .

In the Kondo regime ( $\Gamma/U \ll 1$ ), we verify that the length scale follows the expected scaling of Eq. (21) with the prefactor obtained through a fit (dashed black line) to the blue disks. In the regime of intermediate coupling strength all methods follow the expected exponential scaling. Away from this regime, in the case of weak Kondo coupling ( $U \gg \Gamma$ ), different level of sensitivity to finite-size effect and convergence lead to an underestimated screening length compared to the expected exponential scaling. Similarly, in the strong coupling regime ( $U \sim \Gamma$ ) the mapping from the SIAM to the Kondo model is no longer valid and deviations are expected.

### C. Impurity entropy

As a second probe of the SGS variational ground state, we consider the contribution of the Anderson impurity to the entanglement entropy. This quantity offers a different approach to study the screening cloud of an impurity [35–38].

We consider a bipartition of the sites in subsystems A and B and compute the second Rényi entropy

$$S_2(A, |\psi\rangle) = -\log \text{Tr}(\rho_A^2), \quad (27)$$

where  $\rho_A = \text{Tr}_B |\psi\rangle\langle\psi|$  is the reduced density matrix for subsystem A. Throughout, subsystem A is formed of the  $l$  first sites of the lead and the impurity. While in the case of a single Gaussian state the entropy can easily be computed from the decomposition in normal modes of the covariance matrix [39,40], the presence of coherences between the different Gaussian states forming the ansatz in the density matrix modify this calculation. In Appendix C we use the fermionic coherent state formalism [41] to derive expressions of  $S_2(A, |\psi\rangle)$  for an SGS state of arbitrary rank.

Even though  $S_2$  differs from the more common von Neumann entropy  $S_1$ , many of the same universal properties can be extracted. In particular, using conformal field theory (CFT) calculations, the constant contribution of the boundary to the entropy was shown to be independent of the order of the Rényi entropy at criticality (see, e.g., Refs. [42,43]). In addition the form of the leading corrections to the scaling of the entropy with subsystem size  $l$  due to irrelevant boundary operators was shown to be independent of the order  $n$  of the Rényi entropy [44,45].

As for the correlation function in Sec. IV B, for a lattice model at half-filling, the entanglement entropy is the sum of a uniform and a staggered contribution,

$$S_2(l, \Gamma, L) = S_2^U(l, \Gamma, L) + (-1)^l S_2^A(l, \Gamma, L). \quad (28)$$

The uniform and staggered contributions to the entropy can be extracted using a local polynomial interpolation [36]. The impurity entropy  $S_{\text{imp}}(l, \Gamma, L)$  is then obtained by the subtraction

$$S_{\text{imp}}(l, \Gamma, L) = S_2^U(l, \Gamma, L) - S_2^{0,U}(l, L) \quad (29)$$

where  $S_2^{0,U}(l, L)$  is the uniform part of the entropy in the absence of the impurity. To make the calculation of  $S_{\text{imp}}$  more explicit, Fig. 5(a) presents the raw data for a system of length  $L = 400$  and  $\Gamma = 1/12$ . The corresponding impurity entropy

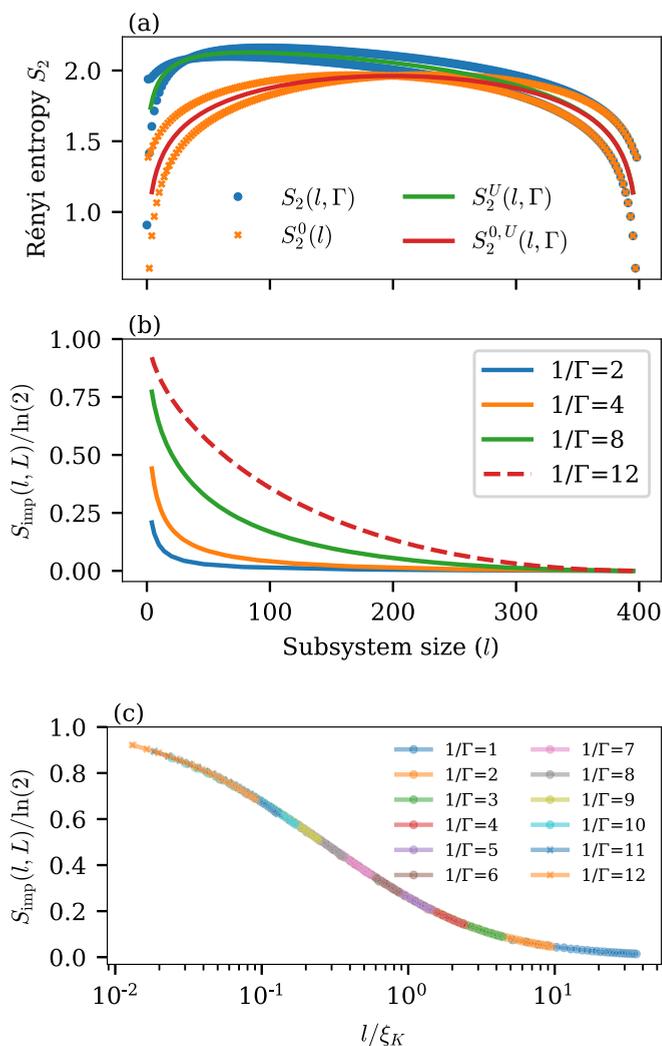


FIG. 5. (a) Raw data for the Rényi entropy  $S_2(A)$  with (blue disks) and without (orange crosses) the impurity ( $L = 400$ ,  $\Gamma = 1/12$ ,  $R = 8$ ). Solid curves are the uniform part of the entropy extracted using a local polynomial interpolation [36]. (b) Impurity entropy for different coupling strengths. Dashed-red curve corresponds to the subtraction of the green and red curves of panel (a). (c) Scaling collapse of  $S_{\text{imp}}$  close to the impurity. To reduce finite-size effects, we consider only the first 25 sites. See Fig. 4(c) for the extracted screening length  $\xi_{K,S}(\Gamma)$ .

$S_{\text{imp}}$  is the dashed-red curve in Fig. 5(b). As expected, for small couplings ( $\Gamma \ll 1$ ) and close to the impurity ( $l \ll \xi_K$ ),  $S_{\text{imp}} \sim \ln(2)$  indicating the entanglement of the impurity with the lead. Consistent with Fig. 4(c), the characteristic length scale over which the  $S_{\text{imp}}$  decays increases as  $\Gamma$  is reduced.

To extract more quantitatively the screening length  $\xi_K$ , we again consider a scaling collapse of the data. The impurity entropy was previously found to be amenable to scaling using [35]

$$S_{\text{imp}}(l, \Gamma, L) = S_{\text{imp}}\left(\frac{l}{\xi_K(\Gamma)}, \frac{l}{L}\right). \quad (30)$$

Figure 5(c) shows this scaling collapse where, focusing on points near the impurity where ( $l \leq 25$  and  $L = 400$ ), we

consider  $l/L \sim 0$ . The correlation length extracted from the collapse  $\xi_{K,S}(\Gamma)$  is shown in Fig. 4(c) (green squares  $\xi_{K,S}$ ). In the regime where the mapping from the SIAM to the Kondo model is valid ( $U/\Gamma \gg 1$ ) and  $\xi_K \ll L$ , we obtain a good agreement between the different methods to extract the Kondo screening length.

## V. TWO-CHANNEL KONDO EFFECT

As a second application of the variational method, we consider the two-channel Kondo (2CK) impurity model. We focus on the symmetric regime where the impurity is coupled with the same strength  $J$  to both channels. In this regime, the presence of a second channel of free electrons lead to very different ground-state properties from the single channel case and the physics of the model includes an intermediate coupling fixed point with non-Fermi liquid behavior [19].

We focus on the ground state of the lattice version of the symmetric two-channel Kondo Hamiltonian

$$H_{2CK} = -t \sum_{\alpha,j=2}^{L-1} (d_{j-1,\alpha}^\dagger d_{j,\alpha} + \text{H.c.}) + H_J + H_{BC}, \quad (31)$$

where  $\alpha = 1, 2$  is the channel index,  $H_J$  is the coupling between the impurity spin and the channels of free fermions and  $H_{BC}$  accounts for the choice of boundary condition away from the impurity. We consider directly an antiferromagnetic spin-spin interaction of coupling strength  $J > 0$  between the impurity spin and the first site of each channel:

$$H_J = J \sum_{\alpha=1,2} \mathbf{S}_{\text{imp}} \cdot \mathbf{S}_\alpha(1), \quad (32)$$

where  $\mathbf{S}_\alpha(r) = \frac{1}{2} \sum_{s,s'} d_{r,\alpha,s}^\dagger \boldsymbol{\sigma}_{s,s'} d_{r,\alpha,s}$  is the fermionic spin operator. As the SGS ansatz is fermionic, we take the impurity to be a fermionic site with fixed single-occupancy.<sup>4</sup> We also introduce the unitless coupling parameter  $g \equiv \rho_0 J$  with  $\rho_0 = 1/4t$  the density of state at the Fermi Level.

For finite-size systems, the two-channel Kondo model is known to exhibit important differences between the case where the total number of sites is even or odd [20]. In order to preserve the symmetry between the two channels and have an even number of sites (including the impurity), we introduce an additional site at the opposite end of the chain (creation operator  $f^\dagger$ ) coupling the two channels,

$$H_{BC} = -t \sum_{\alpha=1,2} (f^\dagger d_{L-1,\alpha} + \text{H.c.}), \quad (33)$$

leading to a total of  $2L$  sites in the model.

### Intermediate coupling fixed point and screening cloud

The 2CK model exhibits an intermediate coupling fixed point, as well as a duality between the weak and strong cou-

<sup>4</sup>As the charge of the fermionic impurity is conserved by  $H_{2CK}$ , the ground state of this model with a fermionic impurity will be the same as in the case where the system is explicitly projected to the single-occupancy subspace (spin-1/2 impurity). We have verified numerically that charge fluctuations at the impurity site vanish in the variational ground state.

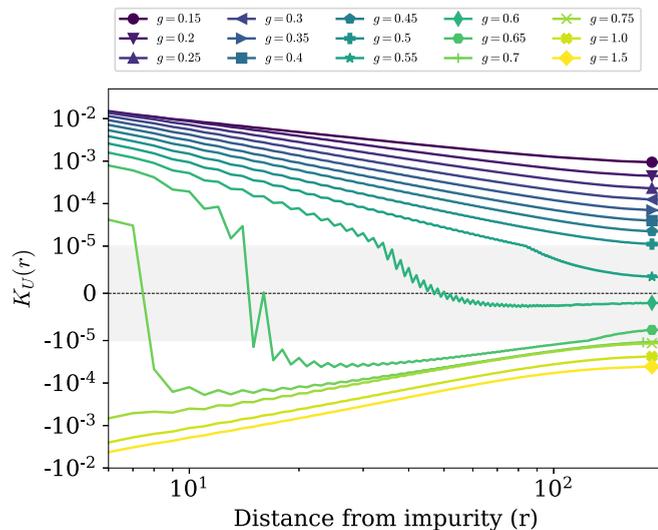


FIG. 6. Uniform part of the spin-spin correlation function  $K(r)$  [see Eq. (34)] on a symmetric log axis. A sign change of the correlation function can be observed for  $g \sim 0.7$  consistent with the crossing of the 2CK intermediate coupling fixed point (see main text). Each channel is half-filled with hopping parameter  $t = 1/2$  and  $L = 200$  sites. An SGS variational state with rank  $R \geq 25$  is used for all coupling strengths. The shaded area indicate the linear part of the symmetric logarithmic vertical axis. A sequential color scale is used to identify curves from weak ( $g = 0.15$ , top purple curve) to strong ( $g = 1.5$ , bottom yellow curve) coupling regimes. A single marker is used to further identify each curve.

pling regimes [46]. In order to explore this physics in the SGS variational ground state, we consider the spin-spin correlation function

$$K(r) = \sum_{\alpha=1,2} (\mathbf{S}_{\text{imp}} \cdot \mathbf{S}_\alpha(r)), \quad (34)$$

where  $r$  is the distance to the impurity and  $\alpha$  is the channel index. We separate again the correlation function in a uniform and a staggered part  $K(r) = -K_U(r) - \cos(2k_F r) K_{2k_F}(r)$ .

Figure 6 shows the uniform part of the correlation function  $K_U(r)$  for various unitless coupling strengths  $g = 0.15 - 1.5$  on a symmetric logarithmic axis. Considering first  $K_U$  close to the impurity, we observe a reduction of correlations as the coupling strength is increased. In absolute value,  $|K_U(r)|$ , reaches a minimum for  $g^* \sim 0.7$  followed by a sign change of  $K_U(r)$  near the impurity. We associate this point where correlations are minimal (smallest screening cloud) as the intermediate coupling strength fixed point of the 2CK model. This result is in agreement with previous NRG results for the Kondo temperature  $T_{2CK}$  [46].

In the weak coupling regime ( $g \lesssim g^*$ ), the Kondo screening length of the 2CK model is to leading order  $\xi_{2CK} \propto e^{1/g}$  [34,46]. In order to extract  $\xi_{2CK}(g)$ , we use a scaling collapse of  $K(r)$ . We focus on the staggered part of the correlation function in the weak coupling regime where correlations are larger. To limit finite-size effects, we consider only correlations up to a distance  $L/2$  from the impurity.

Figure 7(a) presents a scaling collapse following the same procedure as in Sec. IV. For  $g = 0.15 - 0.75$ , all correlation functions can be collapsed on a universal scaling function

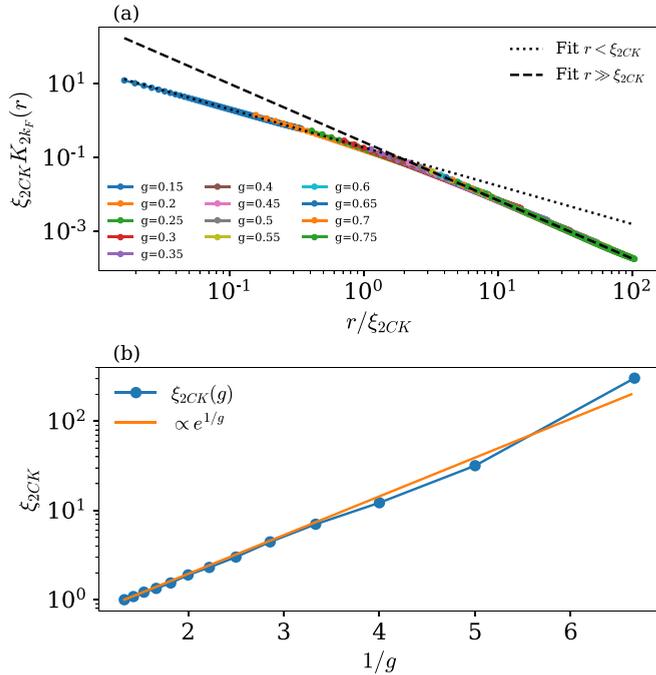


FIG. 7. (a) Scaling collapse of  $K_{2k_F}(r)$ , the staggered part of the spin-spin correlation function for various unitless coupling strengths  $g = 0.15 - 0.75$ , with  $L = 200$  and, to limit finite-size effects,  $r \leq L/2$ . Black dashed (dotted) is a fit of the form  $A(r/\xi_K)^{-\nu}$  to the data far (close) to the impurity with exponent  $\nu = 1.58$  (1.04). (b) Screening length  $\xi_{2CK}$  extracted from the scaling collapse. Orange line is  $\propto e^{1/g}$  with the proportionality factor set by scaling the first point of the curve.

exhibiting a crossover between two different power laws near  $r \sim \xi_{2CK}$ . Using power-law fits, we obtain the exponent  $\nu = 1.04$  near the impurity and  $\nu = 1.58$  far from the impurity (dotted and black dashed lines). An exponent  $\nu < 2$  is coherent with the expected non-Fermi liquid behavior of an overscreened multichannel Kondo model. In particular, for  $r \gg \xi_{2CK}$  the scaling  $K_{2k_F}(r) \propto (r/\xi_{2CK})^{-1.5}$  was calculated in Ref. [34].

The extracted screening length  $\xi_{2CK}(g)$  is shown in Fig. 7(b) and compared to the expected exponential scaling (orange curve). Small deviations from the expected exponential behavior is expected for smaller values of  $g$  where finite-size effects become important ( $\xi_{2CK} \gtrsim L$ ). This result is also in agreement with previous DMRG calculations for an effective spin chain representation of the symmetric 2CK model where  $\xi_{2CK}$  was extracted from a scaling collapse of the impurity entanglement entropy [20].

## VI. CONCLUSIONS

In this paper, we have developed a practical method for finding the variational ground state of an impurity problem using a coherent superposition of fermionic Gaussian states ansatz. The approach has a computational complexity  $O(R^2N^3)$ , thus scaling polynomially with the rank  $R$  of the ansatz and the number of fermionic modes  $N$ . Combined with the favorable scaling of the accuracy with rank  $R$ , as guaranteed by the results of Ref. [12], this gives a powerful approach

to study the ground state of quantum impurity models. In particular, the approach is independent of spatial locality or lattice connectivity, allowing for more flexibility than methods such as DMRG. In addition, its implementation is highly parallelizable allowing for further speed improvements.

In order to verify the method we have first studied the single impurity Anderson model. Comparing the ground-state energy obtained using DMRG to the one obtained using the SGS ansatz, we have shown that an SGS state of rank  $R \sim 6$  can rival with high-precision DMRG calculations with a large bond dimension  $m = 3000$ . Further increasing the ansatz rank up to  $R = 16$  allows to improve the precision of the ground-state energy estimate by an order of magnitude. To demonstrate the quality of the variational ground state, we have also carefully examined the spin-spin correlation functions and the impurity entanglement entropy, and found excellent agreement with analytical and previous numerical results.

To highlight the potential of the method, we have also studied the two-channel Kondo model. Studying again a spin-spin correlation function, we found signatures of the intermediate coupling fixed point in the form a sign change of the uniform part of the correlation function, and were able to confirm the expected exponential scaling of the Kondo screening length. We find a power-law decay of the correlation function far from the impurity with exponent  $\nu \approx 1.58$ , which is close to the expected non-Fermi liquid behavior with exponent  $\nu = 1.5$  predicted by CFT calculations [34]. These results showcase the power of the SGS ansatz for the study of real-space properties of multichannel impurity models, a space of applications previously ill-covered by current standard numerical methods.

An open problem is whether there are computationally efficient ways of extracting the impurity Green's function (in real or imaginary time or frequency) within this class of ansatz states. Such a method would allow this approach to be integrated as impurity solver in embedding methods, such as the dynamical mean-field theory, which requires the impurity Green's function to achieve self-consistency. Furthermore, it appears possible to further reduce the scaling of the method by replacing the Gaussian covariance matrices by Gaussian fermionic matrix product states [47], which could in principle further reduce the scaling with the number of fermionic modes  $N$ .

## ACKNOWLEDGMENTS

DMRG calculations were performed using the ITensor Library [33]. Numerical calculations of Pfaffians were performed using the PFAPACK package [48]. The authors thank A. W. W. Ludwig for insightful discussions.

## APPENDIX A: OVERLAP AND MATRIX ELEMENTS OF FERMIONIC GAUSSIAN STATES

To make the description in this paper self-contained, we will now summarize the formulas needed for calculating overlaps and matrix elements of nonorthogonal Gaussian states. These formulas are used to obtain the final form of the projected equation of motions Eq. (17) in Sec. III B. Derivations

of these formulas and extensions to the case of orthogonal states can be found in Ref. [12].

Let  $|\phi_1\rangle, |\phi_2\rangle$  be two nonorthogonal Gaussian states. Using Eq. (3), these states can be described by covariance matrices  $\Gamma_1$  and  $\Gamma_2$ . In term of these real skew-symmetric matrices, the norm of the overlap of the states is

$$|\langle\phi_1|\phi_2\rangle|^2 = \left| \text{Pf} \left[ \frac{\Gamma_1 + \Gamma_2}{2} \right] \right|. \quad (\text{A1})$$

As discussed in Sec. IIB, the covariance matrix formalism describes Gaussian states up to a global phase factor. In order to calculate overlaps between arbitrary Gaussian states, one need to fix this gauge degree of freedom. Following Ref. [12], we fix the phase of each state with respect to a reference state  $|\phi_0\rangle$  such that  $\langle\phi_0|\phi_\mu\rangle = |\langle\phi_0|\phi_\mu\rangle|$  where  $\mu = 1, 2 \dots$ . With this gauge freedom fixed, the overlap of arbitrary Gaussian states can then be computed as

$$\langle\phi_1|\phi_2\rangle = \frac{\text{Pf}[\frac{1}{2}(\Gamma_1 + \Gamma_2)] \text{Pf}[\frac{1}{2}(\Delta^{1,2} + \Gamma_0)]}{\sqrt{\text{Pf}[\frac{1}{2}(\Gamma_1 + \Gamma_0)] \text{Pf}[\frac{1}{2}(\Gamma_2 + \Gamma_0)]}}, \quad (\text{A2})$$

where we introduced the complex skew-symmetric matrix

$$\Delta^{1,2} = [i(\Gamma_1 - \Gamma_2) - 2\mathbb{1}](\Gamma_1 + \Gamma_2)^{-1}, \quad (\text{A3})$$

and the denominator of Eq. (A2) is  $\langle\phi_0|\phi_1\rangle\langle\phi_2|\phi_0\rangle$ . This expression is used to calculate the Gram matrix  $G$  needed for solving Eq. (17).

Similarly, the Wick theorem for Gaussian states can be extended to compute matrix elements of arbitrary product of Majorana operators. Taking a product of Majorana operators with distinct indices  $\mathcal{I} = i_1, i_2, \dots, i_n$ , the matrix element is

$$\langle\phi_2|c_{i_1}c_{i_2}\dots c_{i_n}|\phi_1\rangle = \langle\phi_2|\phi_1\rangle \text{Pf}[i\Delta^{1,2}(\mathcal{I})], \quad (\text{A4})$$

where  $\Delta^{1,2}(\mathcal{I})$  is a  $n \times n$  submatrix of  $\Delta^{1,2}$  obtained by taken the row and columns corresponding to the indices in  $\mathcal{I}$ . Finally, in order to compute Eq. (17), one needs to calculate the matrix elements of the Hamiltonian defined by Eqs. (1) and (2). Using Eq. (A4) and taking advantage of the skew-symmetric structure of  $A$  and  $U$ , the matrix element  $h_{\alpha,\beta} = \langle\phi_\alpha|H_2 + H_4|\phi_\beta\rangle$  is given by

$$h_{\alpha,\beta} = G_{\alpha,\beta} \text{Tr}[\mathcal{F}^{\beta,\alpha} \Delta^{\beta,\alpha}], \quad (\text{A5})$$

where we introduced the matrix [23]

$$\mathcal{F}_{k,l}^{\beta,\alpha} = A_{k,l} + 3 \sum_{m,n} U_{k,l,m,n} \Delta_{m,n}^{\beta,\alpha}. \quad (\text{A6})$$

We note that the interaction part of Eq. (A6) differs by a factor of 2 from the matrix  $F$  defined in Eq. (16). This difference is in agreement with the GHF results ( $R = 1$  case) obtained in Ref. [23].

## APPENDIX B: NUMERICAL SOLUTION OF EQ. (17)

In this Appendix, we provide additional details on the numerical implementation of the variational method. There are two main approaches to solving Eq. (17) numerically. The first is to solve the equation using a generic differential equation solver, for example taking to first order

$$\Gamma^\mu(s + \delta s) = \Gamma^\mu(s) + \delta s \partial_s \Gamma^\mu(s), \quad (\text{B1})$$

with  $\delta s$  the step size. The drawback of this approach is that the accumulation of small numerical errors due to the finite step size will rapidly lead to covariance matrices, which do not represent normalized and pure states. One must then frequently correct the normalization by decomposing the covariance matrix in the canonical form

$$\Gamma^\mu = R \bigoplus_{j=1}^N \begin{pmatrix} 0 & \lambda_j \\ -\lambda_j & 0 \end{pmatrix} R^T, \quad (\text{B2})$$

with  $R$  a real orthogonal matrix and rescaling the coefficients  $\lambda_j$  to unity.

A second approach is to rewrite Eq. (B1) as an orthogonal transformation [23]

$$\Gamma^\mu(s + \delta s) = R(s) \Gamma^\mu(s) R^T(s) + O(\delta s^2), \quad (\text{B3})$$

with the orthogonal matrix

$$R(s) = \exp \left\{ \frac{\delta s}{2} [\Gamma^\mu(s), \partial_s \Gamma^\mu(s)] \right\}, \quad (\text{B4})$$

where we used  $(\Gamma^\mu)^2 = -\mathbb{1}$  and neglected terms quadratic in  $\delta s$ . This second approach is numerically more expensive than the former due to the required matrix exponentiation. However, it preserves normalization up to machine precision. Our numerical experiments showed this second approach to perform better in some cases as it allows for taking larger time steps and one need not frequently correct the normalization of the covariance matrices using costly canonical form decomposition.

## APPENDIX C: ENTANGLEMENT ENTROPY OF SGS STATES

In this Appendix, we derive the expression used to compute the order 2 Rényi entropy  $S_2(|\psi\rangle, A)$  in Sec. IV C. We consider a bipartition of an SGS state  $|\psi\rangle$  in parts A and B. Defining the density matrices  $\rho_\alpha = |\phi_\alpha\rangle\langle\phi_\alpha|$ , the reduced density matrix for subsystem A is

$$\rho_A = \text{Tr}_B(|\psi\rangle\langle\psi|) = \sum_{\alpha,\beta} \lambda_\alpha \lambda_\beta^* \frac{\text{Tr}_B[\rho_\alpha \rho_\beta]}{G_{\alpha,\beta}}, \quad (\text{C1})$$

since the states in the SGS ansatz are nonorthogonal and thus  $G_{\alpha,\beta} = \langle\phi_\alpha|\phi_\beta\rangle \neq 0$ . In order to evaluate Eq. (C1), we expand the density matrix in a fermionic coherent state basis [41]. The derivation is sketched below with the final result given by Eq. (C16).

Using Eq. (C1), the Rényi entropy of order 2 of an SGS state is

$$S_2(|\psi\rangle, A) = -\ln \left\{ \sum_{\alpha,\beta,\gamma,\delta} T_{\gamma,\delta}^{\alpha,\beta} \right\}, \quad (\text{C2})$$

where we introduced the rank-4 tensor

$$T_{\gamma,\delta}^{\alpha,\beta} = \frac{\lambda_\alpha \lambda_\beta^* \lambda_\gamma \lambda_\delta^*}{G_{\alpha,\beta} G_{\gamma,\delta}} T_{\gamma,\delta}^{\alpha,\beta}, \quad (\text{C3})$$

which obeys the relations  $T_{\gamma,\delta}^{\alpha,\beta} = (T_{\delta,\gamma}^{\beta,\alpha})^*$  and  $T_{\gamma,\delta}^{\alpha,\beta} = T_{\alpha,\beta}^{\gamma,\delta}$  and where  $T_{\gamma,\delta}^{\alpha,\beta} = \text{Tr}_A(\text{Tr}_B(\rho_\alpha \rho_\beta) \text{Tr}_B(\rho_\gamma \rho_\delta))$ . The final result, as a function of the covariance matrices of the Gaussian states, is given by inserting Eq. (C28) in the above equation.

### 1. Coherent state operator expansion

Setting first the required notation, we introduce for each fermionic mode the coherent states

$$|\eta\rangle_i = (1 - \eta_i \hat{d}_i^\dagger)|0\rangle_i, \quad \langle \bar{\eta}|_i = \langle 0|_i(1 - \hat{d}_i \bar{\eta}_i), \quad (\text{C4})$$

where  $\eta_i, \bar{\eta}_i$  ( $i = 1, \dots, N$ ) are Grassman variables obeying the usual algebra

$$\begin{aligned} \eta_i^2 = \bar{\eta}_i^2 = 0 \\ \eta_i \bar{\eta}_j = -\bar{\eta}_j \eta_i, \\ \eta_i \eta_j = -\eta_j \eta_i. \end{aligned} \quad (\text{C5})$$

To lighten the notation we introduce the states  $|\eta\rangle = \bigotimes_i |\eta\rangle_i$  and  $\langle \bar{\eta}| = \bigotimes_i \langle \bar{\eta}|_i$ , as well as the shorthands  $\bar{\eta} \cdot \eta = \sum_i \bar{\eta}_i \eta_i$  for products and  $d^N \eta = \prod_{j=1}^N d\bar{\eta}_j d\eta_j$  for differentials. In addition, we make the dependency on barred variables implicit such that  $F(\eta, \bar{\eta}) \rightarrow F(\eta)$ .

Following the results of Ref. [41], any operator  $\mathcal{O}$  can be represented by the integral

$$\mathcal{O} = \int d^N \eta \chi(\mathcal{O}, \eta) F(\eta), \quad (\text{C6})$$

where  $\chi(\mathcal{O}, \eta)$  is the characteristic function of the operator

$$\chi(\mathcal{O}, \eta) = \text{Tr}[\mathcal{O}D(\eta)], \quad (\text{C7})$$

and the operator  $F(\eta)$  is

$$F(\eta) = \int d^N \psi e^{\frac{1}{2} \bar{\eta} \cdot \eta - \bar{\psi} \cdot \psi + \psi \cdot \bar{\eta} - \eta \cdot \bar{\psi}} |\psi\rangle \langle -\bar{\psi}|. \quad (\text{C8})$$

Following the language of quantum optics  $D(\eta) = \exp\{d^\dagger \cdot \eta - \bar{\eta} \cdot d\}$  is the fermionic analog of the bosonic displacement operator.

In this paper, the operator  $\mathcal{O}$  in Eq. (C6) is the density matrix of a fermionic Gaussian state [26,39]

$$\rho = \frac{1}{2^N} \prod_{j=1}^N (1 - i\lambda_j b_{2j-1} b_{2j}), \quad (\text{C9})$$

with the canonical modes  $b_j = \sum_i R_{i,j} c_i$  with  $R \in SO(2N)$ . The rotation matrix  $R$  and the eigenvalues  $\lambda_j$  are defined through the normal form decomposition of the covariance matrix introduced in Eq. (B2). Using this decomposition and rewriting the displacement operator in this canonical basis, the trace in Eq. (C7) can be evaluated mode per mode leading to the characteristic function

$$\chi(\rho, \eta) = \exp \left\{ \frac{-i}{2} (\bar{\eta}^T \quad \eta^T) V_N^T \Gamma V_N \begin{pmatrix} \bar{\eta} \\ \eta \end{pmatrix} \right\}, \quad (\text{C10})$$

where the  $2N \times 2N$  matrix  $V_N$  makes the rotation of the Grassman operators from a complex fermion basis to a real Majorana basis:

$$V_1 \begin{pmatrix} \bar{\eta}_1 \\ \eta_1 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} \bar{\eta}_1 + \eta_1 \\ i(\bar{\eta}_1 - \eta_1) \end{pmatrix}. \quad (\text{C11})$$

The characteristic function Eq. (C10) allows to relates the Grassman variable representation used, e.g., in Refs. [12,26] to expressions involving the density matrix through Eq. (C6).

### 2. Reduced density matrix

We now turn to deriving the necessary expressions for evaluating numerically Eq. (C1). To this end, we need to compute partial traces of the form  $\text{Tr}_B(\rho_\alpha \rho_\beta)$ , where  $\text{Tr}_B$  is the partial trace over subsystem  $B$ , which is composed of  $N_B$  modes such that  $N = N_A + N_B$ .

Denoting states and Grassman variables associated with the subsystem by a subscript, the partial trace in the coherent state basis takes the form

$$\text{Tr}_B(\rho_\alpha \rho_\beta) = \int d^{N_B} \psi_B e^{-\bar{\psi}_B \cdot \psi_B} \langle -\bar{\psi}_B | \rho_\alpha \rho_\beta | \psi_B \rangle. \quad (\text{C12})$$

Using Eq. (C6) and noting that the  $F$  operator is separable such that  $F(\eta) = F(\eta_A)F(\eta_B)$ , we obtain

$$\begin{aligned} \text{Tr}_B(\rho_\alpha \rho_\beta) = \int d^N \eta \chi(\rho_\alpha, \eta) F(\eta_A) \int d^N \lambda \chi(\rho_\beta, \lambda) \\ \times F(\lambda_A) K(\eta_B, \lambda_B) \end{aligned} \quad (\text{C13})$$

where we introduced the kernel

$$K(\eta_B, \lambda_B) = \int d^{N_B} \psi_B e^{-\bar{\psi}_B \cdot \psi_B} \langle -\bar{\psi}_B | F(\eta_B) F(\lambda_B) | \psi_B \rangle. \quad (\text{C14})$$

Using Eq. (C8) and evaluating the resulting Gaussian integral we obtain

$$K(\eta_B, \lambda_B) = 2^{N_B} \exp \left\{ \frac{1}{2} (\bar{\eta}_B \cdot \lambda_B - \bar{\lambda}_B \cdot \eta_B) \right\}, \quad (\text{C15})$$

where we used for example  $\langle -\bar{\psi} | \eta \rangle = \exp(-\bar{\psi} \eta)$ .

Inserting Eq. (C15) in Eq. (C13), we now isolate the integrals over subsystem  $B$

$$\text{Tr}_B(\rho_\alpha \rho_\beta) = \int d^{N_A} \eta_A d^{N_A} \lambda_A F(\eta_A) F(\lambda_A) \Xi_{\alpha,\beta}(\eta_A, \lambda_A) \quad (\text{C16})$$

with the effective characteristic function for the subsystem  $A$

$$\Xi_{\alpha,\beta}(\eta_A, \lambda_A) = \int d^{N_B} \eta_B d^{N_B} \lambda_B \chi(\rho_\alpha, \eta) \chi(\rho_\beta, \lambda) K(\eta_B, \lambda_B). \quad (\text{C17})$$

In order to take advantage of the structure of Eq. (C11), we formalize the basis change of Eq. (C11) by introducing the new Grassman variables  $\theta_i = \sum_j (V_N)_{i,j} (\bar{\eta}, \eta)_j$  with the differentials transforming as  $d\bar{\eta}_j d\eta_j = \frac{-i}{2} d\theta_{2j-1} d\theta_{2j}$ . Using  $(\tilde{\cdot})$  to indicate functions with arguments in the rotated basis, Eq. (C17) takes the form

$$\tilde{\Xi}_{\alpha,\beta}(\theta_A, \phi_A) = \left( \frac{-1}{2} \right)^{N_B} \int D\theta_B D\phi_B \tilde{\chi}(\rho_\alpha, \theta) \tilde{\chi}(\rho_\beta, \phi) e^{\theta_B^T \cdot \phi_B}, \quad (\text{C18})$$

where  $D\theta_B = \prod_{j=2N_A+1}^{2N} d\theta_j$ . In the case  $N_B = N$ , i.e., when tracing over the whole system, we recover the result of Ref. [26].

Introducing the block structure of the covariance matrix

$$\Gamma^\mu = \begin{pmatrix} \Gamma_A^\mu & \Gamma_{AB}^\mu \\ \Gamma_{BA}^\mu & \Gamma_B^\mu \end{pmatrix} \quad (\text{C19})$$

where  $\Gamma_A, \Gamma_B$  are skewsymmetric and  $\Gamma_{AB} = -\Gamma_{BA}^T$ , one can evaluate the Gaussian integrals resulting in

$$\tilde{\Xi}_{\alpha,\beta}(\theta_A, \phi_A) = 2^{-N_B} \text{Pf}[M_B^{\alpha,\beta}] \exp \left\{ -\frac{i}{2} (\theta_A^T \quad \phi_A^T) \left[ \begin{pmatrix} \Gamma_A^\alpha & 0 \\ 0 & \Gamma_A^\beta \end{pmatrix} + \Sigma_A^{\alpha,\beta} \right] \begin{pmatrix} \theta_A \\ \phi_A \end{pmatrix} \right\}, \quad (\text{C20})$$

where we introduced the  $4N_J \times 4N_J$  skewsymmetric matrices ( $J \in \{A, B\}$ )

$$M_J^{\alpha,\beta} = \begin{pmatrix} \Gamma_J^\alpha & i\mathbb{1} \\ -i\mathbb{1} & \Gamma_J^\beta \end{pmatrix}, \quad \Sigma_A^{\alpha,\beta} = \begin{pmatrix} \Gamma_{AB}^\alpha & 0 \\ 0 & \Gamma_{AB}^\beta \end{pmatrix} [M_B^{\alpha,\beta}]^{-1} \begin{pmatrix} \Gamma_{AB}^\alpha & 0 \\ 0 & \Gamma_{AB}^\beta \end{pmatrix}^T. \quad (\text{C21})$$

Finally, one can transform back to the original basis to obtain  $\Xi(\eta_A, \bar{\eta}_A, \lambda_A, \bar{\lambda}_A)$ , by taking  $\theta_A = V_N(\bar{\eta}_A, \eta_A)$  and  $\phi_A = V_N(\bar{\lambda}_A, \lambda_A)$ .

### 3. Trace of product of reduced density matrices

Building on the results of the previous section, we now turn to the calculation of the trace needed to evaluate Eq. (C3). Dropping the subscript  $A$  when there is no confusion and using Eq. (C16) the trace of the product of reduced density matrices takes the form

$$I_{\gamma,\delta}^{\alpha,\beta} = \int d^{N_A} \eta d^{N_A} \lambda d^{N_A} \xi d^{N_A} \nu \Xi_{\alpha,\beta}(\eta, \lambda) \Xi_{\gamma,\delta}(\xi, \nu) \mathcal{K}(\eta, \lambda, \xi, \nu), \quad (\text{C22})$$

where we introduce the function

$$\mathcal{K}(\eta, \lambda, \xi, \nu) = \int d^{N_A} \psi e^{-\bar{\psi} \cdot \psi} \langle -\bar{\psi} | F(\eta) F(\lambda) F(\xi) F(\nu) | \psi \rangle. \quad (\text{C23})$$

Inserting the definition of  $F$  and performing five Gaussian integrals, one obtains

$$\mathcal{K}(\eta, \lambda, \xi, \nu) = 2^{N_A} e^{\frac{1}{2}(\bar{\eta} \cdot \lambda - \bar{\lambda} \cdot \eta)} e^{\frac{1}{2}(\bar{\xi} \cdot \nu - \bar{\nu} \cdot \xi)} e^{\frac{1}{2}(\bar{\eta} \cdot \nu - \bar{\nu} \cdot \eta)} e^{\frac{1}{2}(\bar{\lambda} \cdot \xi - \bar{\xi} \cdot \lambda)} e^{\frac{1}{2}(\bar{\xi} \cdot \eta - \bar{\eta} \cdot \xi)} e^{\frac{1}{2}(\bar{\nu} \cdot \lambda - \bar{\lambda} \cdot \nu)}. \quad (\text{C24})$$

As all the coherent states have been eliminated, we rotate back to the real Majorana basis introduced in Eq. (C18) such that

$$I_{\gamma,\delta}^{\alpha,\beta} = 2^{-4N_A} \int D\theta D\phi D\theta' D\phi' \tilde{\Xi}_{\alpha,\beta}(\theta, \phi) \tilde{\Xi}_{\gamma,\delta}(\theta', \phi') \tilde{\mathcal{K}}(\theta, \phi, \theta', \phi'). \quad (\text{C25})$$

Grouping the variables such that  $\Theta = (\theta, \phi)$ ,  $\Phi = (\theta', \phi')$  and inserting Eqs. (C20) and (C24)

$$I_{\gamma,\delta}^{\alpha,\beta} = \frac{2^{N_B}}{8^N} \text{Pf}[M_B^{\alpha,\beta}] \text{Pf}[M_B^{\gamma,\delta}] \int D\Theta D\Phi e^{\Theta^T X \Phi} \exp \left\{ \frac{-i}{2} [\Theta^T (M_A^{\alpha,\beta} + \Sigma_A^{\alpha,\beta}) \Theta + \Phi^T (M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta}) \Phi] \right\}, \quad (\text{C26})$$

where we introduced the matrix, which couple the  $\Theta$  and  $\Phi$  variables

$$X = \begin{pmatrix} -\mathbb{1}_{2N_A} & \mathbb{1}_{2N_A} \\ i d_{2N_A} & -\mathbb{1}_{2N_A} \end{pmatrix}. \quad (\text{C27})$$

Performing the final Gaussian integrals, we obtain

$$I_{\gamma,\delta}^{\alpha,\beta} = \frac{2^{N_B}}{8^N} \text{Pf}[M_B^{\alpha,\beta}] \text{Pf}[M_B^{\gamma,\delta}] \text{Pf}[M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta}] \text{Pf}[M_A^{\alpha,\beta} + \Sigma_A^{\alpha,\beta} + \sigma_A^{\delta,\gamma}], \quad (\text{C28})$$

with the matrix

$$\sigma_A^{\delta,\gamma} = -X (M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta})^{-1} X. \quad (\text{C29})$$

We note that Eq. (C28) is symmetric under the exchange of indices  $(\alpha, \beta) \leftrightarrow (\gamma, \delta)$  as required for a trace. This can be more explicit by considering the expression for the Pfaffian of a matrix with a block structure

$$\text{Pf}[M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta}] \text{Pf}[M_A^{\alpha,\beta} + \Sigma_A^{\alpha,\beta} + \sigma_A^{\delta,\gamma}] = \text{Pf} \left[ \begin{pmatrix} M_A^{\alpha,\beta} + \Sigma_A^{\alpha,\beta} & iX \\ -iX & M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta} \end{pmatrix} \right]. \quad (\text{C30})$$

Although numerically more costly, the RHS of Eq. (C30) presents the advantage of making no assumption about the existence of the matrix inverse  $(M_A^{\gamma,\delta} + \Sigma_A^{\gamma,\delta})^{-1}$ . Together with Eq. (C2), Eq. (C28) allows to compute the order 2 Rényi entropy of an SGS state of arbitrary rank.

- [1] K. G. Wilson, The renormalization group: Critical phenomena and the kondo problem, *Rev. Mod. Phys.* **47**, 773 (1975).  
 [2] G. Kotliar, S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti, Electronic structure calculations with dynamical mean-field theory, *Rev. Mod. Phys.* **78**, 865 (2006).

- [3] R. Bulla, T. A. Costi, and T. Pruschke, Numerical renormalization group method for quantum impurity systems, *Rev. Mod. Phys.* **80**, 395 (2008).  
 [4] S. R. White, Density Matrix Formulation for Quantum Renormalization Groups, *Phys. Rev. Lett.* **69**, 2863 (1992).

- [5] A. Weichselbaum, F. Verstraete, U. Schollwöck, J. I. Cirac, and J. von Delft, Variational matrix-product-state approach to quantum impurity models, *Phys. Rev. B* **80**, 165117 (2009).
- [6] H. Saberi, A. Weichselbaum, and J. von Delft, Matrix-product-state comparison of the numerical renormalization group and the variational formulation of the density-matrix renormalization group, *Phys. Rev. B* **78**, 035124 (2008).
- [7] E. Gull, A. J. Millis, A. I. Lichtenstein, A. N. Rubtsov, M. Troyer, and P. Werner, Continuous-time monte carlo methods for quantum impurity models, *Rev. Mod. Phys.* **83**, 349 (2011).
- [8] C. Bertrand, S. Florens, O. Parcollet, and X. Waintal, Reconstructing Nonequilibrium Regimes of Quantum Many-Body Systems from the Analytical Structure of Perturbative Expansions, *Phys. Rev. X* **9**, 041008 (2019).
- [9] T. Shi, E. Demler, and J. I. Cirac, Variational study of fermionic and bosonic systems with non-Gaussian states: Theory and applications, *Ann. Phys. (NY)* **390**, 245 (2018).
- [10] Y. Ashida, T. Shi, M. C. Bañuls, J. I. Cirac, and E. Demler, Solving Quantum Impurity Problems in and out of Equilibrium with the Variational Approach, *Phys. Rev. Lett.* **121**, 026805 (2018).
- [11] T. Shi, J. I. Cirac, and E. Demler, Ultrafast molecular dynamics in terahertz-STM experiments: Theoretical analysis using the Anderson-Holstein model, *Phys. Rev. Research* **2**, 033379 (2020).
- [12] S. Bravyi and D. Gosset, Complexity of quantum impurity problems, *Commun. Math. Phys.* **356**, 451 (2017).
- [13] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, Time-dependent Variational Principle for Quantum Lattices, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [14] H. Shi and S. Zhang, Many-body computations by stochastic sampling in Hartree-Fock-Bogoliubov space, *Phys. Rev. B* **95**, 045144 (2017).
- [15] H. Fukutome, Theory of resonating quantum fluctuations in a fermion system: Resonating Hartree-Fock approximation, *Prog. Theor. Phys.* **80**, 417 (1988).
- [16] H. Fukutome, Theory of excitations in a fermion system with large quantum fluctuations: Resonating random phase approximation, *Prog. Theor. Phys.* **81**, 342 (1989).
- [17] N. Tomita, Many-body wave functions approximated by the superposition of spin-projected nonorthogonal slater determinants in the resonating Hartree-Fock method, *Phys. Rev. B* **69**, 045110 (2004).
- [18] I. Affleck, L. Borda, and H. Saleur, Friedel oscillations and the Kondo screening cloud, *Phys. Rev. B* **77**, 180404(R) (2008).
- [19] Ph. Nozières and A. Blandin, Kondo effect in real metals, *J. Phys. France* **41**, 193 (1980).
- [20] B. Alkurtass, A. Bayat, I. Affleck, S. Bose, H. Johannesson, P. Sodano, E. S. Sørensen, and K. Le Hur, Entanglement structure of the two-channel Kondo model, *Phys. Rev. B* **93**, 081106(R) (2016).
- [21] A. K. Mitchell, M. Becker, and R. Bulla, Real-space renormalization group flow in quantum impurity systems: Local moment formation and the Kondo screening cloud, *Phys. Rev. B* **84**, 115120 (2011).
- [22] V. Bach, E. H. Lieb, and J. P. Solovej, Generalized Hartree-Fock theory and the Hubbard model, *J. Stat. Phys.* **76**, 3 (1994).
- [23] C. V. Kraus and J. I. Cirac, Generalized Hartree-Fock theory for interacting fermions in lattices: Numerical methods, *New J. Phys.* **12**, 113004 (2010).
- [24] G. E. Scuseria, C. A. Jiménez-Hoyos, T. M. Henderson, K. Samanta, and J. K. Ellis, Projected quasiparticle theory for molecular electronic structure, *J. Chem. Phys.* **135**, 124108 (2011).
- [25] C. A. Jiménez-Hoyos, T. M. Henderson, T. Tsuchimochi, and G. E. Scuseria, Projected Hartree-Fock theory, *J. Chem. Phys.* **136**, 164109 (2012).
- [26] S. Bravyi, Lagrangian representation for fermionic linear optics, *Quantum Inf. Comp.* **5**, 216 (2005).
- [27] J. Broeckhove, L. Lathouwers, E. Kesteloot, and P. Van Leuven, On the equivalence of time-dependent variational principles, *Chem. Phys. Lett.* **149**, 547 (1988).
- [28] L. Borda, Kondo screening cloud in a one-dimensional wire: Numerical renormalization group study, *Phys. Rev. B* **75**, 041307(R) (2007).
- [29] A. Holzner, I. P. McCulloch, U. Schollwöck, J. von Delft, and F. Heidrich-Meisner, Kondo screening cloud in the single-impurity anderson model: A density matrix renormalization group study, *Phys. Rev. B* **80**, 205114 (2009).
- [30] M. Nuss, M. Ganahl, E. Arrigoni, W. von der Linden, and H. G. Evertz, Nonequilibrium spatiotemporal formation of the Kondo screening cloud on a lattice, *Phys. Rev. B* **91**, 085127 (2015).
- [31] J. R. Schrieffer and P. A. Wolff, Relation between the Anderson and Kondo Hamiltonians, *Phys. Rev.* **149**, 491 (1966).
- [32] I. Affleck, The Kondo screening cloud: What it is and how to observe it, *Perspectives of Mesoscopic Physics: Dedicated to Yoseph Imry's 70th Birthday*, edited by A. Aharony and O. Entin-Wohlman (World Scientific, 2010), pp. 1–44.
- [33] M. Fishman, S. R. White, and E. M. Stoudenmire, The ITensor software library for tensor network calculations, [arXiv:2007.14822](https://arxiv.org/abs/2007.14822).
- [34] V. Barzykin and I. Affleck, Screening cloud in the  $k$ -channel Kondo model: Perturbative and large- $k$  results, *Phys. Rev. B* **57**, 432 (1998).
- [35] E. S. Sørensen, M.-S. Chang, N. Laflorencie, and I. Affleck, Impurity entanglement entropy and the Kondo screening cloud, *J. Stat. Mech.: Theory Exp.* (2007) L01001.
- [36] E. S. Sørensen, M.-S. Chang, N. Laflorencie, and I. Affleck, Quantum impurity entanglement, *J. Stat. Mech.: Theory Exp.* (2007) P08003.
- [37] I. Affleck, N. Laflorencie, and E. S. Sørensen, Entanglement entropy in quantum impurity systems and systems with boundaries, *J. Phys. A: Math. Theor.* **42**, 504009 (2009).
- [38] C. Wagner, T. Chowdhury, J. H. Pixley, and K. Ingersent, Long-Range Entanglement Near a Kondo-Destruction Quantum Critical Point, *Phys. Rev. Lett.* **121**, 147602 (2018).
- [39] A. Botero and B. Reznik, Modewise entanglement of gaussian states, *Phys. Rev. A* **67**, 052311 (2003).
- [40] G. Vidal, J. I. Latorre, E. Rico, and A. Kitaev, Entanglement in Quantum Critical Phenomena, *Phys. Rev. Lett.* **90**, 227902 (2003).
- [41] K. E. Cahill and R. J. Glauber, Density operators for fermions, *Phys. Rev. A* **59**, 1538 (1999).
- [42] H.-Q. Zhou, T. Barthel, J. O. Fjærestad, and U. Schollwöck, Entanglement and boundary critical phenomena, *Phys. Rev. A* **74**, 050305(R) (2006).
- [43] P. Calabrese and J. Cardy, Entanglement entropy and conformal field theory, *J. Phys. A: Math. Theor.* **42**, 504005 (2009).

- [44] E. Eriksson and H. Johannesson, Impurity entanglement entropy in Kondo systems from conformal field theory, *Phys. Rev. B* **84**, 041107(R) (2011).
- [45] E. Eriksson and H. Johannesson, Corrections to scaling in entanglement entropy from boundary perturbations, *J. Stat. Mech.: Theory Exp.* (2011) P02008.
- [46] C. Kolf and J. Kroha, Strong versus weak coupling duality and coupling dependence of the Kondo temperature in the two-channel Kondo model, *Phys. Rev. B* **75**, 045129 (2007).
- [47] N. Schuch and B. Bauer, Matrix product state algorithms for Gaussian fermionic states, *Phys. Rev. B* **100**, 245121 (2019).
- [48] M. Wimmer, Algorithm 923: Efficient numerical computation of the Pfaffian for dense and banded skew-symmetric matrices, *ACM Trans. Math. Software* **38**, 1 (2012).