Compressing local vertex functions from the multipoint numerical renormalization group using quantics tensor cross interpolation

Markus Frankenbach, Marc K. Ritter, Mathias Pelz, Nepomuk Ritz, Jan von Delft, and Anxiang Ge
Arnold Sommerfeld Center for Theoretical Physics, Center for NanoScience, and Munich Center for Quantum Science and Technology, Ludwig-Maximilians-Universität München, 80333 Munich, Germany

(Received 2 July 2025; accepted 27 August 2025; published 8 October 2025)

The multipoint numerical renormalization group (mpNRG) is a powerful impurity solver that provides accurate spectral data useful for computing local, dynamic correlation functions in imaginary or real frequencies nonperturbatively across a wide range of interactions and temperatures. It gives access to a local, nonperturbative four-point vertex in imaginary and real frequencies, which can be used as input for subsequent computations such as diagrammatic extensions of dynamical mean-field theory. However, computing and manipulating the real-frequency four-point vertex on large, dense grids quickly becomes numerically challenging when the density and/or the extent of the frequency grid is increased. In this paper, we compute four-point vertices in a strongly compressed quantics tensor train format using quantics tensor cross interpolation, starting from discrete partial spectral functions obtained from mpNRG. This enables evaluations of the vertex on frequency grids with resolutions far beyond the reach of previous implementations. We benchmark this approach on the four-point vertex of the single-impurity Anderson model across a wide range of physical parameters, both in its full form and in its asymptotic decomposition. For imaginary frequencies, the full vertex can be represented to an accuracy on the order of 2×10^{-3} with maximum bond dimensions not exceeding 120. The more complex full real-frequency vertex requires maximum bond dimensions not exceeding 170 for an accuracy of $\lesssim 2\%$. Our work marks another step toward tensor-train-based diagrammatic calculations for correlated electronic lattice models starting from a local, nonperturbative mpNRG vertex.

DOI: 10.1103/jx7h-lsqk

I. INTRODUCTION

In the study of strongly correlated systems, correlations at the two-particle level play a key role. A powerful framework for computing two-particle correlation functions is given by quantum field theory approaches such as the functional renormalization group (fRG) [1] or (closely related [2–4]) the parquet equations [5]. While these methods formally provide exact and unbiased equations at the four-point (4p) level, solving them in practice requires some approximations. A common choice is the perturbative parquet approximation, which limits the applicability of these methods to weak interactions. In order to apply these diagrammatic methods to correlated electronic lattice systems in the physically relevant strong interaction regime, it has been proposed to combine them with dynamical mean-field theory (DMFT) [6]. DMFT approximates the self-energy to be local, i.e., momentum-independent, thereby neglecting spatial correlations but capturing local correlations nonperturbatively [7]. In the form of DMFT+fRG (DMF^{2}RG) [8] or the dynamical vertex approximation [9,10], the fRG or the parquet equations can, in principle, be used to self-consistently add nonlocal correlations on the two-particle level on top of the local DMFT result.

However, such calculations entail two numerical challenges: the solution of the impurity model arising in the self-consistent DMFT loop and, subsequently, solving the fRG or parquet equations for frequency- *and* momentum-dependent vertices. The present work is concerned with the *interface* between these two steps, i.e., the conversion of local four-point spectral functions obtained from an impurity solver to a 4p vertex. An impurity solver that yields such 4p spectral functions is the multipoint numerical renormalization group (mpNRG)[11,12]. This extension of the numerical renormalization group (NRG) [13,14] is capable of computing both imaginary- and real-frequency local correlation functions up to the four-point level in the form required for a subsequent diagrammatic extension of DMFT [15,16]. Just as NRG, which has been the gold standard for solving impurity problems on the two-point (2p) level for decades [17,18], mpNRG can be applied to a wide range of parameters, including large interactions and low temperatures. A central ingredient to mpNRG are spectral representations of time-ordered correlation functions in the frequency domain [11]. These represent correlators as convolutions of formalism-dependent but system-independent kernels with formalism-independent but system-dependent partial spectral functions (PSFs). While the former are known analytically, the latter are obtained from their respective Lehmann representations, using the eigenenergies and (discarded) eigenstates obtained from mpNRG.

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

The local 4p vertex can be computed using the symmetric improved estimator (sIE) technique [15], which avoids the numerically unstable amputation of 2p Green's functions.

An appealing feature of spectral representations is that the same set of PSFs can be used to obtain imaginary- and real-frequency correlation functions. However, even when energy conservation is exploited, the 4p vertex is a huge, three-dimensional object. Hence its computation from PSFs on a large, dense grid quickly becomes challenging or even unfeasible due to its huge memory footprint. Furthermore, performing calculations with such vertices as required in fRG or parquet calculations poses a major challenge [19–21].

It is thus highly desirable to represent 4p vertices in a *compressed format* that reduces the computational cost of the operations occurring in diagrammatic calculations. A promising candidate for compression is the quantics tensor train (QTT) representation [22,23] of multivariate functions, which has recently proven useful in various areas of physics [24–29]. Its first application in the context of many-body theory was in Ref. [25], a study that demonstrated the compressibility of correlation functions and used QTT-based algorithms to solve the Schwinger-Dyson and Bethe-Salpeter equations. Furthermore, the QTT representation has been employed successfully in imaginary-frequency parquet calculations for the Hubbard atom and the single-impurity Anderson model (SIAM), using the parquet approximation [28].

These recent developments and the need for efficient representations of 4p vertices motivate this work: We use mpNRG to compute the local vertex of the SIAM as a function of real and imaginary frequencies in QTT format and investigate its compressibility across a broad range of physical parameters. The reason for studying the SIAM is its natural appearance in a DMFT treatment of the Hubbard model and the fact that it can be solved accurately using (mp)NRG. To compute the vertex of the SIAM, we employ the quantics tensor cross interpolation (QTCI) algorithm [30–33], which iteratively constructs a QTT by sparse sampling of the target function. This sampling-based interpolation enables evaluation of the mpNRG vertex on grids much larger and much denser than those accessible with the previous state of the art [15]. For appropriate error tolerances, the maximum bond dimensions (ranks) of the resulting QTTs are within a range where diagrammatic calculations, such as those presented in Ref. [28], should be feasible, even for real frequencies.

This paper is organized as follows: In Sec. II, we recapitulate how the 4p vertex of the SIAM can be obtained in imaginary or real frequencies from PSFs. Additionally, we briefly explain key features of the QTCI algorithm and how it is employed in this work. In Sec. III, we show that imaginary- and, in particular, real-frequency vertices are representable by low-rank QTTs within a reasonable error margin. In Sec. IV, we provide an outlook on how the results of this work may be used to perform diagrammatic calculations for lattice models in QTT format.

II. METHODS

This section explains how to compute 4p vertices in QTT format. This is achieved in two steps: First, we convolve PSFs with formalism-dependent frequency kernels to obtain correlation functions (cf. Ref. [11]). In a second step, the sIE scheme is employed [15] to extract the 4p vertex from various correlators and self-energies in a numerically stable fashion. The vertex is computed both in its asymptotic decomposition [34], which the sIE naturally yields, and in its "full" form.

A. Partial spectral functions

The input to our calculations is given by the PSFs

$$\mathcal{S}[\boldsymbol{\mathcal{O}}](\boldsymbol{\omega}) = \int \frac{\mathrm{d}^\ell t}{(2\pi)^\ell}\, \mathrm{e}^{\mathrm{i}\boldsymbol{\omega}\cdot\boldsymbol{t}} \left\langle \prod_{i=1}^{\ell} \mathcal{O}^i(t_i) \right\rangle, \tag{1}$$

depending on a tuple $\boldsymbol{\mathcal{O}} = (\mathcal{O}^1, \ldots, \mathcal{O}^\ell)$ of operators in the Heisenberg picture and ℓ frequency arguments $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_\ell)$. By $\langle \mathcal{O} \rangle = \mathrm{Tr}[\mathrm{e}^{-\beta H}\mathcal{O}]/Z$, we denote the thermal expectation value, with the partition function $Z = \mathrm{Tr}[e^{-\beta H}]$ at inverse temperature $\beta = 1/T$. Time translation invariance implies

$$\mathcal{S}[\boldsymbol{\mathcal{O}}](\boldsymbol{\omega}) = \delta(\omega_{1\ldots\ell})\, S[\boldsymbol{\mathcal{O}}](\boldsymbol{\omega}), \tag{2}$$

with the shorthand $\omega_{1\ldots\ell} = \sum_{i=1}^{\ell} \omega_i$, thus making $S[\boldsymbol{\mathcal{O}}]$ a function of $\ell - 1$ independent frequencies. In this work, we are primarily interested in the case $\ell = 4$, i.e., three-dimensional PSFs.

The PSFs carry the formalism-independent information that is specific to the model itself. The *same* set of PSFs can thus be used to compute Matsubara and Keldysh correlators by convolution with formalism-specific kernels. In this work, PSFs were computed using mpNRG as described in Ref. [12]. This yields PSFs on a discrete, $(\ell - 1)$-dimensional logarithmic energy grid for all relevant operator tuples $\boldsymbol{\mathcal{O}}$. This yields the representation

$$S[\boldsymbol{\mathcal{O}}](\boldsymbol{\omega}) = \sum_{\boldsymbol{\epsilon}} S[\boldsymbol{\mathcal{O}}](\boldsymbol{\epsilon})\, \delta(\boldsymbol{\omega} - \boldsymbol{\epsilon}), \tag{3}$$

where the peak weights $S[\boldsymbol{\mathcal{O}}](\boldsymbol{\epsilon})$ and energies $\boldsymbol{\epsilon}$ are obtained as output of the mpNRG computation. The energies $\boldsymbol{\epsilon}$ are binned into a Cartesian product of logarithmic grids.

B. Matsubara formalism

A Matsubara correlator $G(\mathrm{i}\boldsymbol{\omega})$ depending on ℓ operators $(\mathcal{O}^1, \ldots, \mathcal{O}^\ell)$ can be expressed via $\ell!$ PSFs $S[\boldsymbol{\mathcal{O}}_p]$, as was shown in Sec. II C of Ref. [11]:

$$G(\mathrm{i}\boldsymbol{\omega}) = \sum_p G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{p,\boldsymbol{\epsilon}} \boldsymbol{\zeta}^p K(\mathrm{i}\boldsymbol{\omega}_p - \boldsymbol{\epsilon}_p) S[\boldsymbol{\mathcal{O}}_p](\boldsymbol{\epsilon}_p). \tag{4}$$

The frequencies $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_\ell)$ are restricted to discrete fermionic or bosonic grids, depending on the type of the respective operator $\mathcal{O}^i$. Similarly to Eq. (2), frequency conservation $\omega_{1\ldots\ell} = 0$ is understood. The sum $\sum_p$ is over all permutations of ℓ elements, permuting frequency arguments and operators accordingly. Using the shorthand $\bar{i} = p(i)$, we can then write

$$\boldsymbol{\omega}_p = (\omega_{p(1)}, \ldots, \omega_{p(\ell)}) = (\omega_{\bar{1}}, \ldots, \omega_{\bar{\ell}}), \tag{5}$$

$$\boldsymbol{\mathcal{O}}_p = (\mathcal{O}_{p(1)}, \ldots, \mathcal{O}_{p(\ell)}) = (\mathcal{O}_{\bar{1}}, \ldots, \mathcal{O}_{\bar{\ell}}). \tag{6}$$

Depending on whether p transposes an even or odd number of fermionic operators, a sign factor $\boldsymbol{\zeta}^p = \pm 1$ is required in

Eq. (4). The summands $G_p(\mathrm{i}\boldsymbol{\omega}_p)$ are termed *partial correlators*. Most importantly, Eq. (4) also introduces the Matsubara frequency kernel K, which reads

$$K(\boldsymbol{\Omega}_p) = \begin{cases} \prod_{i=1}^{\ell-1} \Omega^{-1}_{\bar{1}\dots\bar{i}} & \text{if } \prod_{i=1}^{\ell-1} \Omega_{\bar{1}\dots\bar{i}} \neq 0, \\ -\frac{1}{2}\big[\beta + \sum_{\substack{i=1\\ i\neq j}}^{\ell-1} \Omega^{-1}_{\bar{1}\dots\bar{i}}\big] \prod_{\substack{i=1\\ i\neq j}}^{\ell-1} \Omega^{-1}_{\bar{1}\dots\bar{i}} & \text{if } \exists j : \Omega_{\bar{1}\dots\bar{j}} = 0, \end{cases} \tag{7}$$

where $\Omega_j = \mathrm{i}\omega_j - \epsilon_j$ [cf. Eq. (4)] and $\Omega_{\bar{1}\dots\bar{i}} = \mathrm{i}\omega_{\bar{1}\dots\bar{i}} - \epsilon_{\bar{1}\dots\bar{i}}$ with $\mathrm{i}\omega_{\bar{1}\dots\bar{i}} = \sum_{j=1}^{i} \mathrm{i}\omega_{\bar{j}}$ and $\epsilon_{\bar{1}\dots\bar{i}} = \sum_{j=1}^{i} \epsilon_{\bar{j}}$. The Definition (7) assumes that at most one of the partial sums $\Omega_{\bar{1}\dots\bar{j}}$ vanishes, which is the case if there is at most one bosonic Matsubara frequency (this is always true in the present work). The first case in Eq. (7) is called regular kernel, and the second case anomalous kernel. In the most important situation, $\ell = 4$ and $\prod_{i=1}^{\ell-1} \Omega_{\bar{1}\dots\bar{i}} \neq 0$, the spectral representation of a partial correlator for permutation p is simply given by

$$G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{\epsilon_{\bar{1}},\epsilon_{\bar{1}\bar{2}},\epsilon_{\bar{1}\bar{2}\bar{3}}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}) \prod_{i=1}^{3} (\mathrm{i}\omega_{\bar{1}\dots\bar{i}} - \epsilon_{\bar{1}\dots\bar{i}})^{-1}. \tag{8}$$

Equation (8) assumes that the PSFs $S[\mathcal{O}_p]$ are parametrized in partially summed energies $(\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}})$, which is always the case for our data (see Ref. [12]).

C. Keldysh formalism

We now turn to the relation between PSFs and correlators in the Keldysh formalism [35–37]. For details and derivations, see Sec. II D of Ref. [11]. Keldysh ℓ-point correlators $G^{\boldsymbol{k}}(\boldsymbol{\omega})$ carry a Keldysh index $\boldsymbol{k} = (k_1, k_2, \dots, k_\ell)$ with $k_i \in \{1, 2\}$. Their spectral representation is analogous to Eq. (4):

$$G^{\boldsymbol{k}}(\boldsymbol{\omega}) = \frac{2}{2^{\ell/2}} \sum_p G^{\boldsymbol{k}}_p(\boldsymbol{\omega}_p), \tag{9a}$$

$$G^{\boldsymbol{k}}_p(\boldsymbol{\omega}_p) = \sum_{\boldsymbol{\epsilon}} \boldsymbol{\zeta}^p K_b^{\boldsymbol{k}_p}(\boldsymbol{\omega}_p, \boldsymbol{\epsilon}_p) S[\mathcal{O}_p](\boldsymbol{\epsilon}_p). \tag{9b}$$

It involves ℓ real frequencies $\boldsymbol{\omega}$ that satisfy $\omega_{1\dots\ell} = 0$. The broadened Keldysh frequency kernel $K_b^{\boldsymbol{k}_p}$ is a linear combination of the broadened, fully retarded kernels $K_b^{[\lambda]}$:

$$K_b^{\boldsymbol{k}_p}(\boldsymbol{\omega}_p, \boldsymbol{\epsilon}_p) = \sum_{\substack{\lambda=1\\ k_{\bar{\lambda}}\text{even}}}^{\ell} (-1)^{\lambda-1+k_{\bar{1}\dots\overline{\lambda-1}}} \cdot K_b^{[\lambda]}(\boldsymbol{\omega}_p, \boldsymbol{\epsilon}_p), \tag{10a}$$

$$K_b^{[\lambda]}(\boldsymbol{\omega}_p, \boldsymbol{\epsilon}_p) = \prod_{j=1}^{\ell-1} \lim_{\gamma_0\to 0^+} \int_{\mathbb{R}} \mathrm{d}\omega'_{\bar{1}\dots\bar{j}} \frac{\delta_b(\omega'_{\bar{1}\dots\bar{j}}, \epsilon_{\bar{1}\dots\bar{j}})}{\omega_{\bar{1}\dots\bar{j}} - \omega'_{\bar{1}\dots\bar{j}} + \mathrm{i}\gamma^{\lambda}_{0,j}}, \tag{10b}$$

where $\delta_b(\omega'_{\bar{1}\dots\bar{j}}, \epsilon_{\bar{1}\dots\bar{j}})$ is a broadened version of the Dirac-δ function appearing in Eq. (3). This broadening ensures a smooth structure of the kernel, free from unphysical poles or δ peaks. The imaginary shifts $\mathrm{i}\gamma^{\lambda}_{0,j}$ in Eq. (12) are defined as

$$\mathrm{i}\gamma^{\lambda}_{0,j} = \begin{cases} \mathrm{i}\gamma_0 \cdot (\ell - j), & j \geqslant \lambda, \\ -\mathrm{i}\gamma_0 \cdot j, & j < \lambda. \end{cases} \tag{10c}$$

While the factors $\ell - j$ and j in Eq. (10c) can be disregarded in the limit $\gamma_0 \to 0^+$, they remain relevant for the linear broadening. Details on the broadening procedure can be found in Appendix A and Sec. VI of Ref. [12].

D. Symmetric improved estimators: From correlators to the vertex

In principle, the one-particle irreducible 4p vertex can be obtained simply by amputating the four external 2p propagators ("legs") of the connected impurity Green's function $G_{\mathrm{con}}[d_{\sigma_1} d^\dagger_{\sigma_2} d_{\sigma_3} d^\dagger_{\sigma_4}]$ (cf. Sec. III A). In practice, however, this leads to pronounced numerical artifacts, especially at asymptotically large frequencies, where both functions decay to zero. A numerically stable scheme that avoids direct amputation is the sIE technique introduced in Ref. [15]. In addition, this method yields the vertex in its asymptotic decomposition [34]. This decomposition separates the contributions that decay only in one or two frequency directions from the genuinely three-dimensional core vertex Γ_{core}, which asymptotically decays in all three frequencies,

$$\begin{aligned}\Gamma(\omega, \nu, \nu') = {} & \Gamma_{\mathrm{core}}(\omega, \nu, \nu') \\ & + \sum_{r=a,p,t} \big[\mathcal{K}^r_2(\omega_r, \nu_r) + \mathcal{K}^r_{2'}(\omega_r, \nu'_r) + \mathcal{K}^r_1(\omega_r)\big] \\ & + \Gamma_0. \end{aligned} \tag{11}$$

The functions $\mathcal{K}^r_1$, $\mathcal{K}^r_2$, and $\mathcal{K}^r_{2'}$ (not to be confused with the kernels K defined above) only depend on one or two frequencies if parametrized in their native channel r. They are one- (1D) and two-dimensional (2D) contributions to the two-particle reducible vertex in channel r. This channel can be the antiparallel (a), parallel (p), or the transverse (t) channel. These are also known as the particle-hole, particle-hole crossed, and particle-particle channels, respectively [15]. By Γ_0 we denote the frequency-independent bare vertex. Note that the sIE method does not provide a decomposition of Γ_{core} into two-particle reducible contributions $\mathcal{K}^r_3$ and a two-particle irreducible term.

In this work, we consider the single-impurity Anderson model with interaction $H_{\mathrm{int}} = U d^\dagger_\uparrow d_\uparrow d^\dagger_\downarrow d_\downarrow$, where $d^\dagger_\sigma$ creates an electron with spin σ on the impurity; see Sec. III A for details. We denote the self-energy of the impurity Green's function $G[d_\sigma, d^\dagger_{\sigma'}](\nu)$ as $\Sigma^{\sigma\sigma'}(\nu)$. In the absence of a magnetic field, it satisfies $\Sigma^{\sigma\sigma'}(\omega) = \delta^{\sigma\sigma'}\Sigma(\omega)$. Following Ref. [15], Γ_{core} can be obtained as

$$\begin{aligned}\Gamma_{\mathrm{core}}(\boldsymbol{\omega}) = {} & \sum_{a_i\in\{d,q\}} Y_{a_1}(\omega_1) Y_{a_3}(\omega_3)\, G_{\mathrm{con}}[a_1, a^\dagger_2, a_3, a^\dagger_4](\boldsymbol{\omega}) \\ & \cdot Y_{a_2}(\omega_2) Y_{a_4}(\omega_4), \end{aligned} \tag{12}$$

where we introduced an auxiliary operator $q = [d, H_{\rm int}]$, and

$$Y_{a_i}(\omega_i) = \begin{cases} -\Sigma\big((-1)^{i-1}\omega_i\big), & a_i = d, \\ X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, & a_i = q \end{cases} \tag{13}$$

is a 2×2 matrix acting on the ith Keldysh index of $G_{\rm con}$. In the Matsubara formalism, X is replaced by scalar unity. In practice, the main workload in computing $\Gamma_{\rm core}$ at a given frequency $\boldsymbol{\omega}$ is the evaluation of all $2^4 = 16$ connected correlators, each comprised of $4! = 24$ partial correlators. The quantities $\mathcal{K}_1^r$ and $\mathcal{K}_{2^{(\prime)}}^r$ can be computed using an analogous approach presented in Appendix B.

E. Quantics tensor cross interpolation

Having summarized the evaluation of the four-point vertex in its asymptotic decomposition, we next discuss the QTCI method [30–33], which we used to obtain vertex functions in the form of QTTs. Recently, it has been shown in the Matsubara formalism that this representation is well suited for efficient diagrammatic calculations [28]. We discuss only the basics of QTCI here. For a detailed introduction, we refer to Ref. [38].

Let us begin with the quantics representation [22,23]. Consider a 1D function $f(\omega)$ defined on a discrete, equidistant grid $\{\omega_0, \ldots, \omega_{2^R-1}\}$ consisting of 2^R points with $\omega_m \in \mathbb{R}$. The grid index m of a point ω_m can be written in binary representation

$$m = \sum_{\ell=1}^{R} 2^{R-\ell}\sigma_\ell, \quad \sigma_\ell \in \{0, 1\}, \tag{14}$$

so that m can be identified with the R-tuple $(\sigma_1, \ldots, \sigma_R)$. Hence, the mapping $m \mapsto f(\omega_m)$ can be viewed as an R-leg tensor $F_{\sigma_1\cdots\sigma_R} = f(\omega_{m(\{\sigma_\ell\})})$. This so-called quantics encoding can be generalized to higher-dimensional functions, in particular to functions $f(\omega, \nu, \nu')$ of three frequency arguments. The frequencies lie on a Cartesian product of 1D grids, each of size 2^R. We use a binary encoding

$$(\omega_i, \nu_j, \nu'_k) = ((\sigma_{11}, ..., \sigma_{1R}), (\sigma_{21}, ..., \sigma_{2R}), (\sigma_{31}, ..., \sigma_{3R})), \tag{15}$$

with the binary variables $\sigma_{n\ell}$ labeled by $n = 1, 2, 3$ for ω, ν, ν'. The function f can then be represented by a tensor with $3R$ indices:

$$F_{\boldsymbol{\sigma}} = F_{\sigma_{11}\sigma_{21}\sigma_{31}\cdots\sigma_{1R}\sigma_{2R}\sigma_{3R}} = f(\omega_i, \nu_j, \nu'_k). \tag{16}$$

Importantly, note that the tensor indices in Eq. (16) have been *interleaved*, such that the indices corresponding to the same length scale $2^{R-\ell}$ in different variables $\sigma_{1\ell}, \sigma_{2\ell}, \sigma_{3\ell}$ are adjacent. Alternatively, triples $(\sigma_{\ell 1}, \sigma_{\ell 2}, \sigma_{\ell 3})$ of legs can be fused to single legs $\tilde{\sigma}_\ell = \sum_{n=1}^{3} 2^{n-1}\sigma_{\ell n}$, which yields the *fused* representation of f as an R-leg tensor:

$$\widetilde{F}_{\tilde{\boldsymbol{\sigma}}} = \widetilde{F}_{\tilde{\sigma}_1\cdots\tilde{\sigma}_R} = f(\omega_i, \nu_j, \nu'_k). \tag{17}$$

The second ingredient of QTCI is the tensor cross interpolation (TCI) algorithm [30–32,38], which approximates tensors $F_{\boldsymbol{\sigma}} = F_{\sigma_1\cdots\sigma_L}$ (with $L = R$, $2R$, or $3R$ for one-, two-, or three-dimensional functions, respectively) using tensor trains constructed from a sampled subset of all tensor elements. If a low-rank factorization of the tensor exists, the number of samples taken is much smaller than the number of elements of the full tensor. This way, the cost of generating all tensor elements, exponentially large in R, can be avoided.

More precisely, the TCI algorithm seeks to find a tensor train

$$F^{\rm QTCI}_{\sigma_1\cdots\sigma_L} = \sum_{\alpha_1\cdots\alpha_{L-1}} \big[M_1^{\sigma_1}\big]_{1\alpha_1}\big[M_2^{\sigma_2}\big]_{\alpha_1\alpha_2}\cdots\big[M_L^{\sigma_L}\big]_{\alpha_{L-1}1} \tag{18}$$

that minimizes the elementwise error

$$\varepsilon_{\boldsymbol{\sigma}}[F] = \frac{\big|F^{\rm QTCI}_{\boldsymbol{\sigma}} - F_{\boldsymbol{\sigma}}\big|}{\max_{\boldsymbol{\sigma}'}|F_{\boldsymbol{\sigma}'}|}. \tag{19}$$

Here, α_ℓ are virtual bond indices with ℓ-dependent bond dimensions, $\alpha_\ell = 1, ..., \chi_\ell$. The maximum bond dimension, $\chi = \max \chi_\ell$, is called the *rank* of $F^{\rm QTCI}_{\boldsymbol{\sigma}}$. The maximum in Eq. (19) is estimated using all sampled entries of $F_{\boldsymbol{\sigma}}$. The TCI algorithm optimizes the tensors $[M_\ell^{\sigma_\ell}]_{\alpha_{\ell-1}\alpha_\ell}$ iteratively, progressively sampling $F_{\boldsymbol{\sigma}}$, until no $\boldsymbol{\sigma}$ is found where the error $\varepsilon_{\boldsymbol{\sigma}}$ exceeds a given tolerance τ. During this process, the bond dimensions χ_ℓ are increased dynamically to improve the accuracy of the tensor train representation. Finding a tensor train representation of $F^{\rm QTCI}_{\boldsymbol{\sigma}}$ with TCI has a computational cost of $\mathcal{O}(R\chi^3)$.

In conjunction with the quantics representation, TCI can be employed to approximate not only tensors, but also functions defined on discrete grids by tensor trains. Once the bond dimensions are saturated, i.e., no longer increase with R, the computational cost of QTCI scales linearly in R. This translates to an exponential resolution of the target function at linear cost. Of course, the function is only approximated within an error margin given by Eq. (19). The next section details how we used QTCI to compress 4p vertex functions on exponentially large grids.

F. Implementation details

To obtain the core and full vertices in the Matsubara and Keldysh formalisms as QTTs, we apply QTCI to functions that evaluate $\Gamma_{\rm core}(\boldsymbol{\omega})$ and $\Gamma(\boldsymbol{\omega})$ on individual frequency points $\boldsymbol{\omega}$ to be specified on demand by the TCI algorithm. The frequencies $\boldsymbol{\omega}$ reside on an equidistant grid of 2^{3R} points. For Matsubara grids, the grid spacing is set by the temperature, and the extent of the grid can be increased exponentially by increasing R. For Keldysh vertices, which are functions of continuous frequencies, one may exponentially increase either the density of grid points or the extent of the grid, or both, by increasing R. It is important not to precompute the vertices on a dense grid, as this precomputation step would incur costs scaling as $\mathcal{O}(2^{3R})$. By avoiding precomputation, R can be increased to yield grid sizes and/or grid densities beyond those attainable by conventional means. The TCI algorithm samples a *sparse* set of $\mathcal{O}(\chi^2 R)$ points $\boldsymbol{\omega}$, which is much smaller than 2^{3R}, the total number of grid points for the 4p vertex functions considered in this work. For this application, function evaluation during sampling is the dominant cost as opposed to the $\mathcal{O}(\chi^3 R)$ cost of computing prrLU factorizations (see Sec. 3.3 of Ref. [38]). More specifically, the computational effort is dominated by the evaluation of partial 4p correlators. These enter $\Gamma_{\rm core}$ and Γ via the full correlators appearing in Eq. (12). In this section, we discuss how our code evaluates partial

correlators in an efficient way. Readers only interested in our results on the compressibility of Matsubara and Keldysh vertices can move on to Sec. III.

1. Matsubara vertices

In the Matsubara case, evaluating the regular part of the 4p correlator, Eq. (8), constitutes the majority of computational cost. Using the shorthand $k_{\omega_{\bar{1}\ldots\bar{i}}\epsilon_{\bar{1}\ldots\bar{i}}} = \mathrm{i}\Omega_{\bar{1}\ldots\bar{i}} = (\mathrm{i}\omega_{\bar{1}\ldots\bar{i}} - \epsilon_{\bar{1}\ldots\bar{i}})^{-1}$, we can rewrite Eq. (8) as

$$G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{\epsilon_{\bar{1}},\epsilon_{\bar{1}\bar{2}},\epsilon_{\bar{1}\bar{2}\bar{3}}} k_{\omega_{\bar{1}}\epsilon_{\bar{1}}} k_{\omega_{\bar{1}\bar{2}}\epsilon_{\bar{1}\bar{2}}} k_{\omega_{\bar{1}\bar{2}\bar{3}}\epsilon_{\bar{1}\bar{2}\bar{3}}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}). \tag{20}$$

Since $(\omega_{\bar{1}}, \omega_{\bar{1}\bar{2}}, \omega_{\bar{1}\bar{2}\bar{3}})$ and $(\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}})$ live on finite frequency grids, the kernels $k_{\omega_{\bar{1}\ldots\bar{i}}\epsilon_{\bar{1}\ldots\bar{i}}}$ can be viewed as matrices. A typical grid size for the spectral function peaks $(\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}})$ is $70 \times 70 \times 70$, while the largest grids for Matsubara frequencies $(\omega_{\bar{1}}, \omega_{\bar{1}\bar{2}}, \omega_{\bar{1}\bar{2}\bar{3}})$ used in this work have $2^{12} = 4096$ points in each dimension. Thus, $k_{\omega_{\bar{1}\ldots\bar{i}}\epsilon_{\bar{1}\ldots\bar{i}}}$ can be precomputed and stored for all relevant grid sizes.

We implemented two methods to speed up the threefold contractions in Eq. (20): (1) compressing the kernels and (2) performing one kernel contraction as a preprocessing step.

(1) Compressing the kernels is the more general of the two methods, in that it has a smaller memory footprint [<1 GB per full correlator $G(\mathrm{i}\boldsymbol{\omega})$ for $R = 12$ and $\tau = 10^{-3}$]. The idea is to compress the kernels $k_{\omega\epsilon}$ by exploiting their low-rank structure [39,40]: We apply singular value decomposition (SVD) to each $k_{\omega\epsilon}$ and discard singular values below a given cutoff S_{cut}, resulting in the approximation

$$k_{\omega\epsilon} \approx \sum_a U_{\omega a} S_a V^{\dagger}_{a\epsilon}. \tag{21}$$

We then contract the singular values $S_a \geqslant S_{\mathrm{cut}}$ and the right-hand isometries $V^{\dagger}$ with the PSF by performing ϵ sums to obtain a smaller rank-3 tensor A, thus reducing the cost of the threefold summation in Eq. (20):

$$G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{a_1a_2a_3} A_{a_1a_2a_3} \prod_{i=1}^{3} U_{\omega_i a_i}, \tag{22a}$$

$$A_{a_1a_2a_3} = \sum_{\epsilon_{\bar{1}},\epsilon_{\bar{1}\bar{2}},\epsilon_{\bar{1}\bar{2}\bar{3}}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}) \prod_{i=1}^{3} S_{a_i} V^{\dagger}_{a_i\epsilon_{\bar{1}\ldots\bar{i}}}. \tag{22b}$$

The computations (21) and (22b) are performed during preprocessing prior to the QTCI run. Note that this treatment of the low-rank Matsubara kernels is closely related to the so-called intermediate representation (IR) of Matsubara Green's functions (see Refs. [41,42]). The cutoff S_{cut} should be chosen as to introduce an error significantly below the TCI tolerance τ in all target quantities. While one can bound the error in Eq. (22a), e.g., using the Cauchy-Schwarz inequality, these estimates were found to be very conservative. We observed that setting $S_{\mathrm{cut}} = 10^{-2}\,\tau$ leads to errors more than 2 orders of magnitude below the TCI tolerance when evaluating correlators and vertices. This is shown in Fig. 1, where we plot the accuracy of Matsubara (and Keldysh) core vertex evaluations for different TCI tolerances τ and numbers of quantics bits R. In Matsubara, the accuracy improves with increasing R. This is because, for a fixed SVD cutoff, fewer singular values are discarded for larger R.

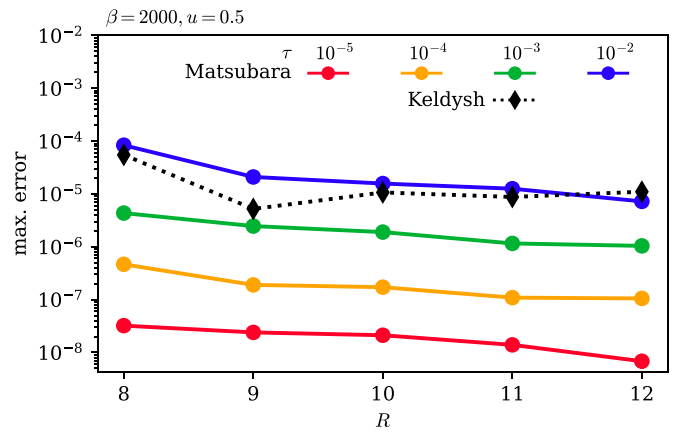

FIG. 1. Maximum error of Matsubara (circles) and Keldysh (diamonds) core vertex evaluations with SVD truncations as described in the main text. The error is measured relative to the maximum of the respective vertex function. By τ we denote the target TCI tolerance and use a cutoff of $S_{\mathrm{cut}} = 10^{-2}\tau$ for Matsubara. For Keldysh, we choose an SVD cutoff of 10^{-6} times the largest singular value and find that this yields results that are sufficiently accurate for a tolerance of $\tau = 10^{-3}$. We show maximum errors over 64 000 sampling points for the Matsubara core vertex and 2×10^6 sampling points for the more complicated Keldysh core vertex. All errors are well below the respective target tolerance τ.

A further speedup can be achieved by realizing that even though all singular values S_a in Eq. (21) are larger than S_{cut}, their products appearing in Eq. (22b) can become negligibly small. Ordering S_{a_i} by decreasing magnitude, we therefore discard all entries $A_{a_1a_2a_3}$ where $a_1 + a_2 + a_3$ is larger than some integer N:

$$G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{\sum_i a_i \leqslant N} A_{a_1a_2a_3} \prod_{i=1}^{3} U_{\omega_i a_i}. \tag{23}$$

A similar truncation is useful when constructing IR 4pt Green's functions (see Ref. [43]). To determine N for a prescribed tolerance τ, we estimate the contribution from terms with $\sum_i a_i > N$ via the Cauchy-Schwarz inequality:

$$\begin{aligned} &\left| \sum_{\sum_i a_i > N} A_{a_1a_2a_3} \prod_{i=1}^{3} U_{\omega_i a_i} \right| \\ &\quad \leqslant \sqrt{\sum_{\sum_i a_i > N} \left|A_{a_1a_2a_3}\right|^2} \prod_{i=1}^{3} \max_{\omega_i} \sqrt{\sum_{a_i} \left|U_{\omega_i a_i}\right|^2} \\ &\quad = \sqrt{\sum_{\sum_i a_i > N} \left|A_{a_1a_2a_3}\right|^2}. \end{aligned} \tag{24}$$

The second factor in the second line of Eq. (24) is equal to one, since the U's are isometries. Hence, Eq. (24) provides a simple bound on the error in $G_p(\mathrm{i}\boldsymbol{\omega})$, which is independent of

the frequency at hand. We choose N such that

$$\sqrt{\sum_{\sum_i a_i > N} \left|A_{a_1 a_2 a_3}\right|^2} \leqslant \frac{\tau}{10} \max_{\boldsymbol{\omega}} |G(\mathrm{i}\boldsymbol{\omega})|, \tag{25}$$

which ensures an error one order of magnitude below the TCI tolerance in the full correlator G. While this error in principle occurs *per* partial correlator G_p, the criterion (25) was observed to yield sufficient accuracy. Overall, this first method gives a substantial speedup compared to directly performing the contractions in Eq. (20): It accelerates the evaluation of the full 4p impurity correlator $G[d_\uparrow, d_\uparrow^\dagger, d_\uparrow, d_\uparrow^\dagger]$ at $\beta = 2000$ and $u = 0.5$ (cf. Sec. III A) at a single frequency in an $R = 12$ quantics grid using an SVD cutoff of $S_{\mathrm{cut}} = 10^{-5}$ by more than a factor of 60. This observation simply reflects the strong compresssibilty of the Matsubara kernels.

(2) A more straightforward way to speed up pointwise evaluations of partial correlators [Eq. (20)] is to precompute one of the three contractions before running QTCI. This yields an object depending on the variables $(\omega_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}})$,

$$B_p(\omega_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}) = \sum_{\epsilon_{\bar{1}}} k_{\omega_{\bar{1}}\epsilon_{\bar{1}}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}). \tag{26}$$

B_p gives access to G_p via

$$G_p(\mathrm{i}\boldsymbol{\omega}_p) = \sum_{\epsilon_{\bar{1}2}\epsilon_{\bar{1}23}} k_{\omega_{\bar{1}\bar{2}}\epsilon_{\bar{1}\bar{2}}} k_{\omega_{\bar{1}\bar{2}\bar{3}}\epsilon_{\bar{1}\bar{2}\bar{3}}} B_p(\omega_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}). \tag{27}$$

In this approach, we only have two kernel contractions in each evaluation G_p, but have to store the intermediates [Eq. (26)] for all partial correlators. Their size grows linearly in the grid frequency grid size, i.e., as 2^R where R is the number of quantics bits. For example, if $R = 12$ and the PSFs live on a $70 \times 70 \times 70$ logarithmic grid (which results from 2×6 decades of energy bins with eight points per decade and discarding zeros in the PSFs), each full correlator consumes 12.9 GB of memory. On an $R = 12$ grid at $\beta = 2000$ and $u = 0.5$, the precomputation also yields a speedup of about a factor of 60. But in contrast to the compression of kernels [Eq. (21)], this speedup is independent of the TCI tolerance τ. Overall, method (2) is recommended as long as its memory demands can be met, because it evaluates correlators in a numerically exact fashion.

2. Keldysh vertices

In the Keldysh formalism, evaluating partial correlators $G_p^{\boldsymbol{k}}$ [cf. Eq. (9b)] also comes down to threefold contractions of a three-dimensional PSF with kernel matrices. This can be seen by rewriting the kernel $K_b^{[\lambda]}$ [Eq. (10b)] as a product of one-dimensional kernels evaluated at frequencies $\omega_{\bar{1}\ldots\bar{i}}$:

$$K_b^{[\lambda]}(\boldsymbol{\omega}_p, \boldsymbol{\epsilon}_p) = \prod_{i=1}^{3} k_b^{[\lambda,i]}(\omega_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}}), \tag{28a}$$

$$k_b^{[\lambda,i]}(\omega_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}}) = \lim_{\gamma_0 \to 0^+} \int_{\mathbb{R}} \mathrm{d}\omega'_{\bar{1}\ldots\bar{i}} \frac{\delta_b(\omega'_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}})}{\omega_{\bar{1}\ldots\bar{i}} - \omega'_{\bar{1}\ldots\bar{i}} + \mathrm{i}\gamma_{0,i}^{\lambda}}. \tag{28b}$$

Equation (9b) can then be written as

$$G_p^{\boldsymbol{k}}(\boldsymbol{\omega}_p) = \sum_{\substack{\lambda=1 \\ k_{\bar{\lambda}}\text{-even}}}^{4} (-1)^{\lambda-1+k_{\bar{1}\ldots\overline{\lambda-1}}} \cdot G_p^{[\lambda]}(\boldsymbol{\omega}_p), \tag{29a}$$

$$G_p^{[\lambda]}(\boldsymbol{\omega}_p) = \sum_{\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}) \times \prod_{i=1}^{3} k_b^{[\lambda,i]}(\omega_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}}). \tag{29b}$$

The ensuing contractions [Eq. ((29a) to be performed for $\lambda = 1, ..., 4$ are analogous to Eq. (20). However, interpolating the complex structure of the Keldysh vertex requires more evaluations compared to its Matsubara counterpart. At the same time, the memory cost of a precomputation analogous to Eq. (26) becomes prohibitive for large ($R \gtrsim 12$) frequency grids, since it must be applied to $G_p^{[\lambda]}$ for $\lambda = 1, ..., 4$ and for each partial correlator. For these reasons, the optimization of Eq. (29a) needs to go beyond the compression scheme for the Matsubara kernels from Eqs. (21) and (22b). To this end, we exploit the fact that the structure of the 1D kernels $k_b^{[\lambda,i]}(\omega_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}})$ becomes simpler at large frequencies $\omega_{\bar{1}\ldots\bar{i}}$: We divide the $\omega_{\bar{1}\ldots\bar{i}}$ grid into n_L equally sized intervals $I_1^i, \ldots, I_{n_L}^i$, with $n_L = 2^3$ as a default. Then, for each dimension i and each interval I_j^i, we SVD-decompose the restricted kernel

$$k_b^{[\lambda,i]}(\omega_{\bar{1}\ldots\bar{i}}, \epsilon_{\bar{1}\ldots\bar{i}})\Big|_{\omega_{\bar{1}\ldots\bar{i}} \in I_j^i} \approx \sum_{a_i} U^{ij}_{\omega_{\bar{1}\ldots\bar{i}} a_i} S^{ij}_{a_i} V^{\dagger ij}_{a_i \epsilon_{\bar{1}\ldots\bar{i}}}, \tag{30}$$

discarding singular values that are at least 6 orders of magnitude smaller than the largest singular value. This strategy of partitioning the $\omega_{\bar{1}\ldots\bar{i}}$ grid prior to the SVD truncation allows us to discard more singular values in outer intervals, where the kernel is more compressible. Next, for each triple of intervals (I_k^1, I_l^2, I_m^3), we contract the corresponding singular values and right-hand isometries into the PSF. While this entails precomputing n_L^3 3-leg tensors of the form

$$(A^{klm})_{a_1 a_2 a_3} = \sum_{\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}} (SV^\dagger)^{1k}_{a_1\epsilon_{\bar{1}}} (SV^\dagger)^{2l}_{a_2\epsilon_{\bar{1}2}} (SV^\dagger)^{3m}_{a_3\epsilon_{\bar{1}23}} S[\mathcal{O}_p](\epsilon_{\bar{1}}, \epsilon_{\bar{1}\bar{2}}, \epsilon_{\bar{1}\bar{2}\bar{3}}), \tag{31}$$

it yields a substantial speedup in evaluations of $G^{\boldsymbol{k}}(\boldsymbol{\omega})$: For $\beta = 2000$, $u = 0.5$, $\omega_{\max} = 0.65$ (cf. Sec. III A), and $R = 12$, this scheme is about a factor of 150 faster than a naive kernel contraction. This speedup refers to an average over 2×10^5 evaluations on random frequency points, since the compressibility of the kernels depends on the intervals (I_k^1, I_l^2, I_m^3) the frequencies $(\omega_{\bar{1}}, \omega_{\bar{1}\bar{2}}, \omega_{\bar{1}\bar{2}\bar{3}})$ belong to. Indeed, truncated isometries U^{ij} pertaining to the outermost intervals usually have about 5 times fewer rows than those of the inner intervals. That Keldysh core vertex evaluations using the above scheme are sufficiently accurate (i.e., to more than 10^{-3}, see Sec. III C) is verified in Fig. 1.

Having explained the optimization of vertex evaluations, we turn to the settings chosen in the QTCI routine. All of our code is written in Julia (versions 1.9.4 and 1.10.3), us-

ing the TCI package `TensorCrossInterpolation.jl`, the quantics utilities `QuanticsGrids.jl`, and the QTCI package `QuanticsTCI.jl` of the tensor4all collaboration [38,44]. The latter exposes the `quanticscrossinterpolate` routine, which is the entry point of the QTCI algorithm and offers various settings: We used the default `:backandforth` sweep strategy and the `:fullsearch` pivot search strategy. The increase in computational cost entailed by a full pivot search was accepted to ensure a reliable interpolation. For Matsubara objects, the interleaved representation was chosen to obtain maximum memory compression. In the Keldysh case, the interleaved representation exhibited convergence problems: After 80 sweeps (with $\leqslant 5$ sweeps until convergence being common), a QTT with an error significantly exceeding the tolerance was obtained. Switching to the fused representation solved this problem. This is due to the fact that a two-site update in a three-dimensional fused representation corresponds to a six-site update in the interleaved representation, which implies more extensive sampling of the target function. Another choice worth mentioning is that of initial pivots: For Keldysh vertices, it was sufficient to use the grid center as the only initial pivot. In Matsubara, the same choice occasionally lead to premature termination of the TCI algorithm, resulting in a QTT representation that was missing relevant features. We therefore chose 125 initial pivots forming a cube at Matsubara frequencies $(\omega_i, \nu_{j-1}, \nu'_{k-1})$ with $i, j, k \in \{-2, ..., 2\}$. On fermionic grids, the cube is thus centered around $\nu_{-1} = \nu'_{-1} = -\pi T$. This choice of initial pivots ensures that the sharp Matsubara vertex structure around the origin is properly sampled. Finally, since vertex evaluations are the bottleneck of our QTCI compressions, a significant speedup can be achieved via multithreading. The samples evaluated during a two-site optimization step (see Sec. 4.3 of Ref. [38]) are independent of one another and can therefore be evaluated in parallel.

We tested our code for evaluating vertex functions with the sIE scheme against the prior MATLAB implementation used in Ref. [15]. We found numerically exact agreement with a normalized discrepancy $<10^{-13}$ for the Matsubara quantities. In Keldysh, the maximum discrepancy in the core vertex between our Julia code and the MATLAB code of Ref. [15] is about $0.002 \cdot ||\Gamma_{\text{core}}||_\infty$ (with the supremum norm $||\cdot||_\infty$). This discrepancy can be attributed to small differences in the broadening implementation, mainly the interpolation of the broadened kernel from a logarithmic to a linear grid (cf. Appendix A). This discrepancy is one order of magnitude smaller than the error introduced by the arbitrariness inherent in the choice of broadening parameters. As an additional test, our code was used to generate Keldysh vertex data to check the fulfillment of exact diagrammatic relations of mpNRG data [16].

To conclude this section, Fig. 2 compares the singular value spectra of regular Matsubara kernels [Eq. (7)] and broadened, fully retarded Keldysh kernels [Eq. (28a)] at different temperatures. As expected, the singular values of both Matsubara and Keldysh kernels decay significantly faster at higher temperatures. In Keldysh, this is due to the temperature-dependent linear broadening γ_L (see Appendix A). Moreover, the singular values of Keldysh kernels decay much more slowly than their Matsubara counterparts at the same temperature. This reflects the more complex structure, i.e., lower compressibility, of Keldysh vertices.

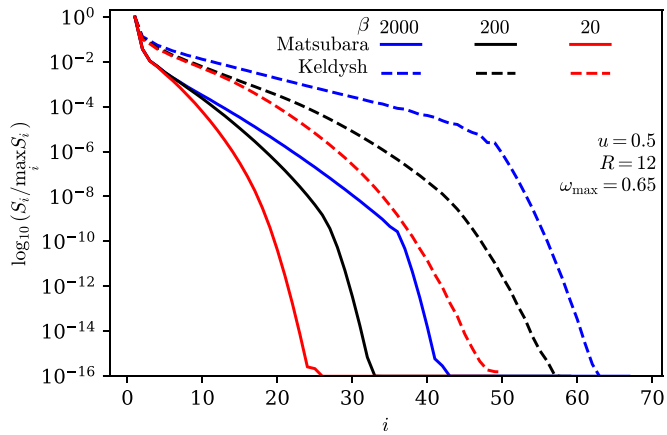

FIG. 2. Singular values S_i of regular Matsubara kernels [Eq. (7), solid line] and broadened, fully retarded Keldysh kernels [Eq. (28a), dashed line]. We show kernels at inverse temperatures $\beta \in \{20, 200, 2000\}$ and interaction $u = 0.5$. The frequency grids are bosonic with 2^{12} points, with the Keldysh grid ranging from -0.65 to $\omega_{\max} = 0.65$.

III. RESULTS

In this section, we show how QTCI performs in compressing the 4p vertex of the single-impurity Anderson model in the Matsubara (Sec. III B) and Keldysh (Sec. III C) formalisms. We discuss the benefits of its QTT representation compared to storing the vertex on dense frequency grids, considering both the core and the full vertex [Γ_{core} and Γ in Eq. (11)]. The asymptotic contributions are discussed in Appendix B. The two most relevant numerical parameters are the number R of quantics bits, corresponding to a grid with 2^R points in each dimension and the maximum bond dimension χ, which serves as a measure for compressibility.

A. Single-impurity Anderson model

The Hamiltonian of the single-impurity Anderson model [46] reads

$$H = \sum_\sigma \epsilon_d\, n_\sigma + U n_\uparrow n_\downarrow + \sum_{k\sigma} \epsilon_k c^\dagger_{k\sigma} c_{k\sigma} + \sum_{k\sigma} V_k \big(d^\dagger_\sigma c_{k\sigma} + \text{H.c.}\big), \quad n_\sigma = d^\dagger_\sigma d_\sigma, \tag{32}$$

where $d^\dagger_\sigma$ with spin $\sigma \in \{\uparrow, \downarrow\}$ creates an electron in an interacting, single-orbital impurity. $c^\dagger_{b\sigma}$ creates an electron in a noninteracting bath, coupled to the impurity via a hybridization term V_k. Electrons on the impurity site interact with the interaction strength U. Since the $c_{k\sigma}$ electrons occur only quadratically, they can formally be integrated out, yielding a frequency-dependent hybridization function $\Delta(\nu)$ as an additional quadratic term for the d electrons. We choose the hybridization function as

$$\Delta(\nu) = \frac{\Delta}{\pi} \ln \left| \frac{\nu + D}{\nu - D} \right| - i\Delta\, \theta(D - |\nu|), \tag{33}$$

TABLE I. Kondo temperatures $T_K = T_K(U, \Delta)$ with inverses $\beta_K = 1/T_K$ for different parameter sets. The Kondo temperature was computed via the Bethe ansatz solution of the SIAM (see, e.g., Ref. [45]). All quantities have been rounded to three significant digits.

U	Δ	u	T_K	$\beta_K = 1/T_K$
0.05	0.0318	0.5	4.14×10^{-2}	24.2
0.05	0.0159	1.0	9.58×10^{-3}	104
0.05	0.0106	1.5	3.57×10^{-3}	280
0.05	0.00530	3.0	3.36×10^{-4}	2980
0.05	0.00318	5.0	2.06×10^{-5}	48 400

with a box-shaped imaginary part, characterized by the bandwidth $2D$ and the hybridization strength $\Delta \in \mathbb{R}$. Moreover, we set $\epsilon_d = -U/2$, which leads to a particle-hole symmetric Hamiltonian.

In the following, energy, temperature, and frequencies are measured in units of half the bandwidth $D = 1$. The interaction strength is specified by the dimensionless quantity $u = U/\pi\Delta$. Our analysis covers a wide parameter range from weak ($u = 0.5$) to very strong ($u = 5.0$) interactions and moderate ($\beta = 20$) to low ($\beta = 2000$) temperatures. The corresponding Kondo temperatures are given in Table I. (It should be noted that the two datasets for $\beta = 20$ and $\beta = 200$ have $u = 0.5004$ rather than $u = 0.5$. This minor difference changes the Kondo temperature by less than a factor of 1.002.) To parametrize the vertex, different frequency conventions and index orderings can be used. Both are listed in Appendix D. Finally, note that the spin structure of the vertex $\Gamma^{\sigma_1\sigma_2\sigma_3\sigma_4}$ can be simplified by exploiting the SU(2) spin symmetry of the SIAM in the absence of a magnetic field. Only components of the form

$$\Gamma^{\sigma\sigma'} = \Gamma^{\sigma\sigma\sigma'\sigma'} \tag{34}$$

are needed. Moreover, we have $\Gamma^{\downarrow\downarrow} = \Gamma^{\uparrow\uparrow}$ and $\Gamma^{\downarrow\uparrow} = \Gamma^{\uparrow\downarrow}$ by spin flip symmetry, such that only $\Gamma^{\uparrow\uparrow}$ and $\Gamma^{\uparrow\downarrow}$ remain independent. The same applies to $\Gamma^{\sigma\sigma'}_{\text{core}}$.

B. mpNRG vertex functions: Matsubara formalism

Let us first consider the QTCI compression of the Matsubara core and full vertices. An important input to the QTCI algorithm is the specified error tolerance τ [see Eq. (19)]. When compressing vertices from mpNRG, the choice of tolerance should be based on the accuracy of the PSFs. Based on benchmark results of Refs. [11,12,15], we expect the mpNRG vertex to be reliable to roughly $10^{-3} \cdot ||\Gamma_{\text{core}}||_\infty$, where the error is partially systematic (as opposed to pure white noise). It should be emphasized that this is only an estimate and inherent errors in mpNRG (due to discretization of the noninteracting bath and discarding high-energy eigenstates during iterative diagonalization) are different from those stemming from TCI. A tolerance significantly below $\tau = 10^{-3}$ may be desirable for two reasons: First, to avoid errors [Eq. (19)] *larger* than our mpNRG accuracy estimate of 10^{-3}: After all, a local error [Eq. (19)] below the tolerance is only expected within the set of pivots that have been sampled by TCI—and even for these, the tolerance is not strictly guaranteed

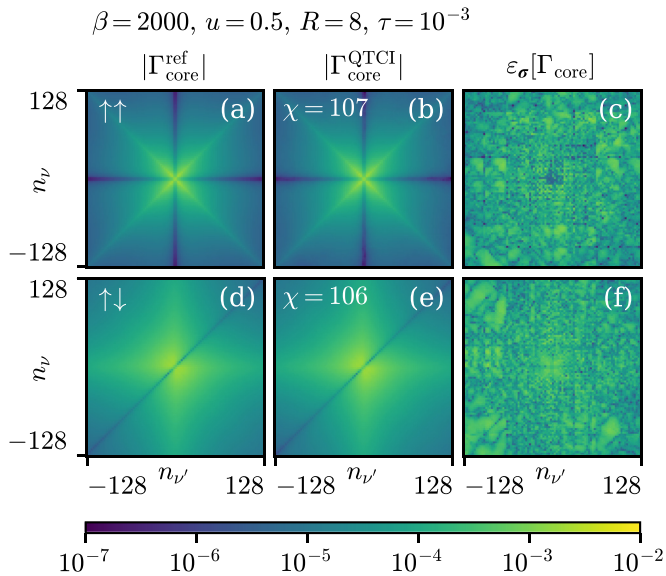

FIG. 3. QTCI compression of the Matsubara core vertex $\Gamma_{\text{core}}(\omega, \nu, \nu')$ in the p channel at $\beta = 2000$, with $R = 8$ and tolerance $\tau = 10^{-3}$. Heatmaps show the $\log_{10}$ absolute value of $\Gamma^{\uparrow\uparrow}_{\text{core}}$ in panels (a) and (b) and $\Gamma^{\uparrow\downarrow}_{\text{core}}$ in panels (d) and (e) on the slice $\omega = 0$. n_ν and $n_{\nu'}$ enumerate the fermionic Matsubara frequencies ν, ν'. Left column: Reference data $\Gamma^{\text{ref}}_{\text{core}}$. Center column: QTCI representation $\Gamma^{\text{QTCI}}_{\text{core}}$. Right column: Normalized error $\varepsilon_{\sigma}[\Gamma_{\text{core}}] \lesssim 1.58 \times 10^{-3}$ defined in Eq. (19). We reproduce key features of the vertex on a large frequency box with a comparatively low QTT rank of $\chi = 107$ and $\chi = 106$, respectively.

by the TCI routine used here (see Sec. 4.3.1 of Ref. [38]), which breaks full nesting conditions. Lowering the tolerance increases the confidence that the required accuracy has been reached even outside the sampled set. The second motivation is to assess the potential of our approach for situations where more precise input data are available. We shall therefore investigate tolerances ranging from 10^{-2} to 10^{-5}.

The vertex functions $\Gamma_{\text{core}}(\omega, \nu, \nu')$ and $\Gamma(\omega, \nu, \nu')$ to be represented in QTT format here generally have prominent structures around the origin, along the frequency axes, and along the diagonals [47]. This is exemplified in Fig. 3, which shows slices of Γ_{core} at fixed bosonic frequency $\omega = 0$. The inverse temperature is $\beta = 2000$. We compare reference data with the QTT representation of the vertex for a TCI tolerance of $\tau = 10^{-3}$. The reference was obtained by evaluating the vertex only on the two-dimensional slice shown in Fig. 3. A logarithmic color scale has been chosen to expose imperfections of the TCI approximation. Figure 3 illustrates how QTCI represents important features of the vertex in a strongly compressed format: For a $256 \times 256 \times 256$ ($R = 8$) frequency grid, we have ranks of $\chi = 107$ for $\Gamma^{\uparrow\uparrow}_{\text{core}}$ and $\chi = 106$ for $\Gamma^{\uparrow\downarrow}_{\text{core}}$. This translates to memory footprints reduced by factors of 92 (from 268 to 2.9 MB) and 89 (from 268 to 3.0 MB), respectively.

A systematic account of the compressibility of Γ_{core} is provided in Fig. 4. It shows [panel (a)] ranks and [panel (b)] memory consumption of the resulting QTTs, as a function of tolerance τ and grid size. The grid size is governed by the number R of quantics bits in each dimension. Two

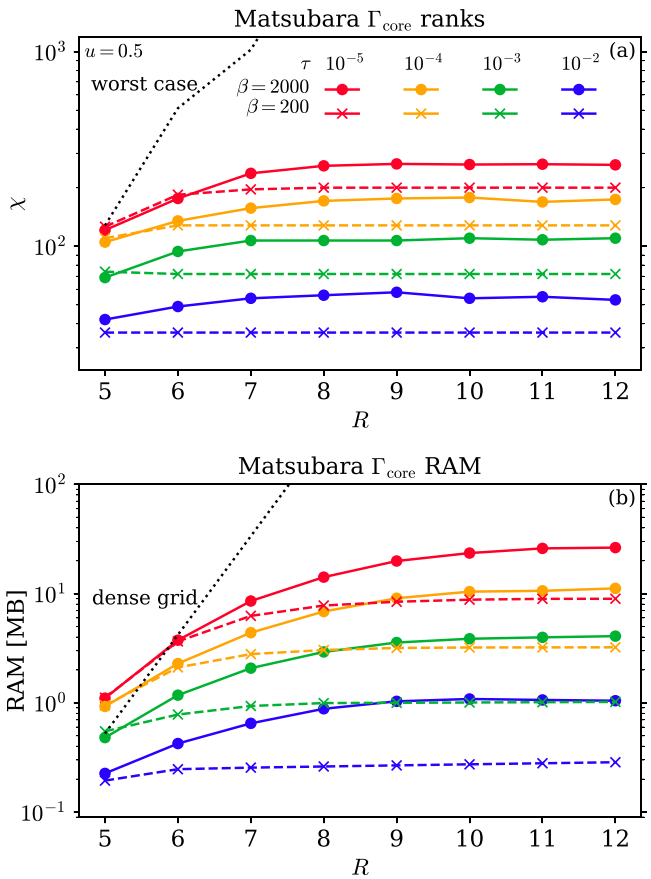

FIG. 4. (a) Rank and (b) random access memory (RAM) usage of the interleaved QTT representation of the Matsubara core vertex $\Gamma^{\uparrow\uparrow}_{\rm core}$ in the p channel vs frequency grid size for different tolerances. The grid has 2^R points in each frequency argument. For the target tolerance of $\tau = 10^{-3}$, ranks saturate at $\chi \approx 100$. Dotted worst-case lines in panels (a) and (b) indicate the maximum rank of a $3R$-leg QTT [hence the even-odd alternation in the worst case of panel (a)] and the RAM requirements of dense grids with 2^{3R} points, respectively.

different datasets at $\beta = 200$ and $\beta = 2000$ are represented by crosses and circles, respectively. The QTT ranks do not exceed 250 (red circles, $\tau = 10^{-5}$), with compressions for the target accuracy of $\tau = 10^{-3}$ (green) *saturating* with respect to R at $\chi = 96$ even for the low-temperature data. This rank saturation reflects the simple asymptotic structure of $\Gamma_{\rm core}$.

In light of recent work by Rohshap *et al.* [28], these results are very promising: There, the authors demonstrate that self-consistent parquet calculations with maximum bond dimensions of 200 are feasible on a single CPU (cf. Sec. VI B of Ref. [28]). While the calculations in Ref. [28] were performed in a different parameter regime of the SIAM, their computational cost is determined by the bond dimensions of the QTTs involved. Our results therefore suggest that QTT-based parquet calculations with an NRG Matsubara vertex as input will be feasible. Figure 4 further shows that TCI ranks of $\Gamma_{\rm core}$ do not significantly increase beyond a grid size of $R = 8$. In this region of saturated ranks, both memory usage and run-time required for manipulations of the vertex such as convolutions or frequency transformations scale logarithmically in the grid size (linearly in R) [33]. In this regime, the QTT representation yields an exponential reduction in computational cost compared to dense grids. Lowering the tolerance (thus increasing χ) comes at a run-time cost of $O(\chi^4)$ for the most expensive manipulations performed in Ref. [28] (see Sec. V D there).

The linear scaling in R generically allows for exponentially cheap reduction of discretization errors (for continuous variables) or errors due to finite-size domains (for discrete variables) [33]. For Matsubara vertices, the asymptotic structure contains terms Γ_0, $\mathcal{K}_1$, $\mathcal{K}_2$, and $\mathcal{K}_{2'}$ that are independent of some of the frequencies [see Eq. (11)], implying that the function does not decay to zero at infinity [34]. One might be tempted to conclude that this makes any finite box representation invalid without high-frequency extrapolation. In practice, vertex functions are used in evaluating diagrammatic equations such as expectation values of observables, Bethe-Salpeter equations, or the Schwinger-Dyson equation. In all of these cases, the vertex is embedded in a frequency summation or integral with single-particle propagators that *do* approach zero asymptotically for high frequencies. The error generated in such summations and integrals is thus the relevant criterion for frequency box size, as was shown in Ref. [28] for a parquet approach, including Bethe-Salpeter equations. There, the error of a QTCI-based self-consistent calculation of the density channel irreducible vertex was shown to improve from 10^{-1} for $R = 5$ to 10^{-3} for $R = 9$. Frequency grids larger than $R = 8$ are hence relevant. We are considering significantly lower temperatures ($\beta D \in \{200, 2000\}$, $U/\pi\Delta = 0.5$ here versus $\beta D = 100$, $U/\pi\Delta \approx 0.51$ in Ref. [28]). Since $\Gamma_{\rm core}$ becomes more complicated at these low temperatures

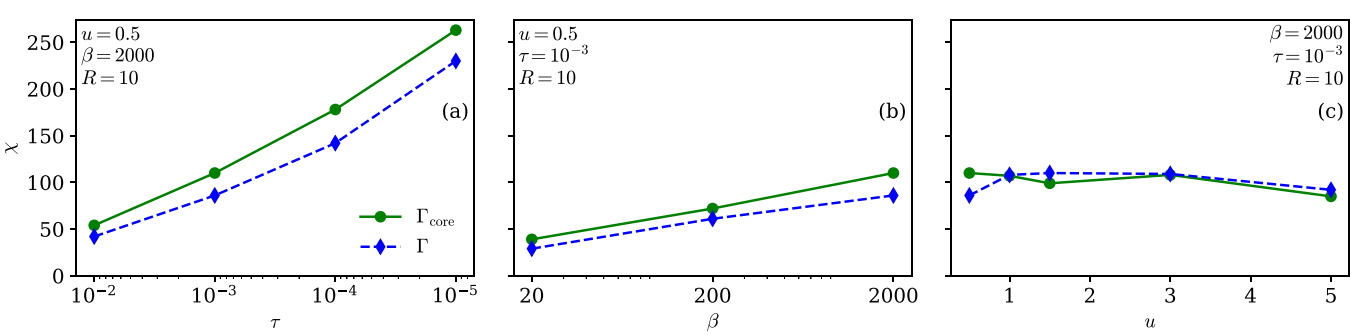

FIG. 5. Rank of the Matsubara core vertex $\Gamma^{\uparrow\uparrow}_{\rm core}$ and full vertex $\Gamma^{\uparrow\uparrow}$ in the p channel vs (a) TCI tolerance τ, (b) inverse temperature β, and (c) interaction strength u.

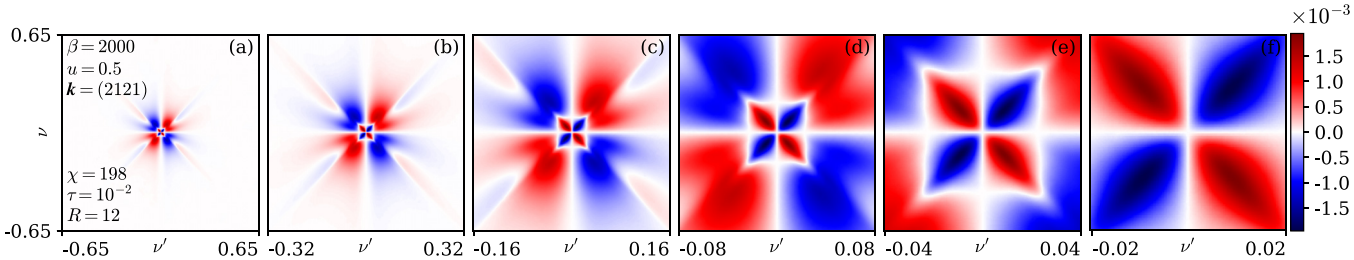

FIG. 6. Imaginary part of the Keldysh core vertex $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}$, at $\omega = 0$, compressed using $R = 12$ and $\tau = 10^{-2}$. On this slice, $\text{Re}(\Gamma^{2121,\uparrow\uparrow}_{\text{core}})$ is a factor of 25 smaller than $\text{Im}(\Gamma^{2121,\uparrow\uparrow}_{\text{core}})$. The QTT rank is $\chi = 198$. The QTT representation allows us to zoom in by a factor of 2^5 (from left to right) while retaining a sharp resolution throughout.

[cf. Fig. 5(b)], we expect large grids to be even more relevant in NRG + parquet calculations.

To assess the range of applicability of our approach, we also examined how the QTT rank for Γ_{core} as well as the full vertex Γ depends on the desired tolerance, inverse temperature β, and interaction strength u. The results for the $\uparrow\uparrow$ flavor and an $R = 10$ grid are summarized in Fig. 5. The full and core vertices show a similar increase in rank with the TCI tolerance [Fig. 5(a)], since Γ_{core} contains precisely the complex three-dimensional structure of the full vertex. Consistent with previous results on random pole-based Matsubara correlators (cf. Fig. 8(b) in Ref. [48]), the TCI ranks increase with β, though only logarithmically [see Fig. 5(b)]. Finally, Fig. 5(c) shows the ranks versus the interaction strength. The key finding is that both vertices remain strongly compressible with ranks $\leqslant 110$ when increasing u from the perturbative regime ($u \ll 1$) to very strong coupling ($u = 5$). Since the y axis ranges only from 85 to 110, the observed variation in ranks with u does not carry much significance. Overall, Fig. 5 suggests that parquet calculations with an mpNRG vertex as input will be feasible across a wide range of parameters.

According to Eq. (11), the full vertex also contains lower-dimensional contributions K^r_1 and $K^r_{2^{(\prime)}}$. In Appendix B, we verify that they have very small TCI ranks ($\chi \lesssim 20$ for $\tau = 10^{-3}$) compared to Γ_{core} and Γ. Moreover, we focused on the $\uparrow\uparrow$ component of the vertex in the p channel. In Appendix C, we discuss how the ranks of Γ_{core} and Γ depend on the frequency channel and spin component.

C. mpNRG vertex functions: Keldysh formalism

We now turn to computations of the Keldysh vertex in QTT format. In contrast to the Matsubara vertex, this object gives direct access to real-frequency dynamic response functions, but is a significantly more complicated function on a continuous domain of real frequencies. Faithfully capturing its structure on a finite grid while keeping the computational cost in check is very challenging. This has been achieved in Refs. [20,21], but requires tedious manual tuning of nonlinear grids. The evaluation of mpNRG Keldysh vertices on nonlinear grids is discussed in Appendix E. Our QTCI-based approach allows us to automatically capture features on different length scales on an extremely fine equidistant grid. Our grid for ω contains 0, while ν and ν' live on a grid that is offset from 0 by half a grid spacing. An alternative choice would be to include 0 in all three grids. In QTCI, we can refine the grid until all features are represented up to a given tolerance, so that shifting the ν and ν' grids by half a grid spacing does not make a difference. The resolution attained with QTCI is exemplified in Fig. 6, showing a QTT representation of $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}$ on a slice at $\omega = 0$ using $R = 13$ quantics bits. The TCI tolerance was set to 10^{-2}. All panels show the same slice, but zoom in by factors of 2 moving from left to right. The rightmost panel still exhibits a sharp resolution after a 32-fold magnification. As a further illustration, Fig. 7 compares the TCI-compressed vertex (center) to the reference (left), showing the normalized error on the right. The slices are taken again at $\omega = 0$ and for $u = 0.5$ (top row) and $u = 1.5$ (bottom row). Overall, we see that TCI resolves the core vertex to 1% precision with ranks of 184 and 187, respectively.

This 1% error is comparable to the uncertainty due to the broadening of spectral functions, which supersedes the NRG error of 10^{-3} in the real-frequency case: While there are well-

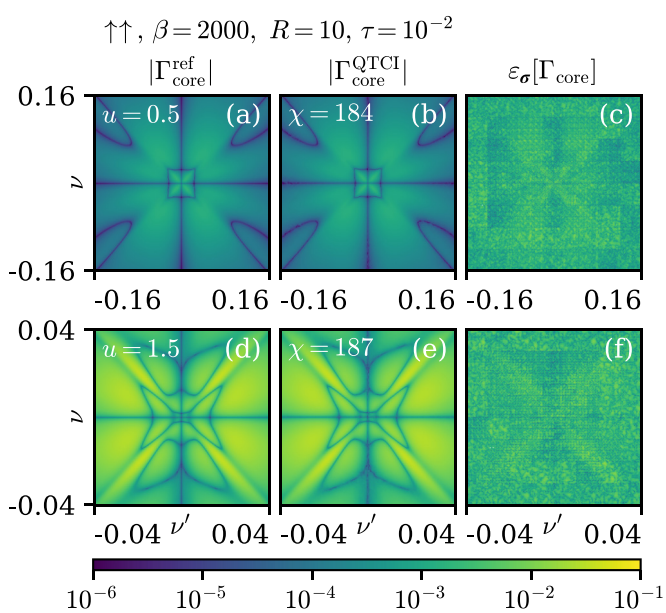

FIG. 7. QTCI compression of the Keldysh core vertex $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}(\omega, \nu, \nu')$ in the p channel at $\beta = 2000$, with $R = 10$ and tolerance $\tau = 10^{-2}$. Heatmaps show the $\log_{10}$ absolute value of the vertex on the slice $\omega = 0$. The interaction strengths are $u = 0.5$ in panels (a)–(c) and $u = 1.5$ in panels (d)–(f). Left column: Reference data $\Gamma^{\text{ref}}_{\text{core}}$. Center column: QTT representation $\Gamma^{\text{QTCI}}_{\text{core}}$. Right column: Normalized error $\varepsilon_{\boldsymbol{\sigma}}[\Gamma_{\text{core}}] \lesssim 1.73\%$ defined in Eq. (19). The QTT representation captures complex features using moderate bond dimensions of $\chi = 184$ and $\chi = 187$, respectively.

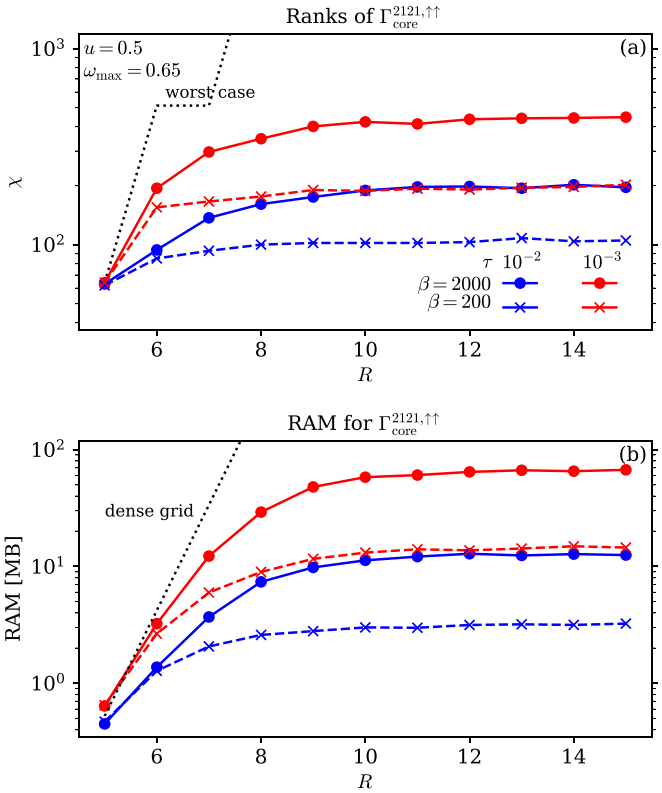

FIG. 8. (a) Rank and (b) RAM usage of Keldysh core vertex $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}(\omega, \nu, \nu')$ in the p channel vs frequency grid size for different tolerances and temperatures. The grid has 2^R points in each frequency argument. For a tolerance of $\tau = 10^{-2}$ (blue), the bond dimension saturates at $\chi \approx 200$. Dotted worst-case lines in panels (a) and (b) indicate the maximum rank of an R-leg QTT (fused representation) and the RAM requirements of dense grids with 2^{3R} points, respectively.

established schemes [12,15] to choose broadening parameters, legitimate choices can vary within a range that causes vertex functions to change by a few percent. Our default tolerance for Keldysh objects is therefore chosen as $\tau = 10^{-2}$. In view of ongoing research aiming to develop an impurity solver less susceptible to broadening artifacts [49], which could be extended to the multipoint case in the future, we extend our investigations down to $\tau = 10^{-3}$.

Figure 8 shows [panel (a)] the rank and [panel (b)] the memory size of the compressed Keldysh core vertex component $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}$ at temperatures $\beta = 200$ and 2000 versus the number of quantics bits R in each dimension. This Keldysh component of $\Gamma^{\boldsymbol{k}}_{\text{core}}$ was found to have the highest bond dimension (see Appendix C, Fig. 12). We set a fixed box size of $\omega_{\max} = 0.65$ [cf. Fig. 6(a)] and increase the resolution with R. We verified that the chosen box size is large enough to capture all relevant structures within the target tolerance $\tau = 10^{-2}$. At this tolerance, the rank shown in Fig. 8 saturates at $\chi = 202$. This rank is again of a magnitude where a self-consistent parquet calculation in the Matsubara formalism was shown to be feasible on a single core in Ref. [28]. The 16 components of the Keldysh vertex can be inferred from just 5 components using complex conjugation and crossing symmetry [50]. Nevertheless, in follow-up computations such as solving the parquet equations, these multiple Keldysh components in contrast to a single Matsubara vertex may necessitate parallelization already for $\chi \approx 200$. Multithreaded or distributed schemes will certainly be required for the most difficult case considered here ($\tau = 10^{-3}$ and $\beta = 2000$), which results in ranks of $\chi \approx 450$. On a different note, the QTT vertex has a vastly reduced memory footprint, as shown in Fig. 8(b): For $\beta = 2000$, $\tau = 10^{-2}$, and $R = 10$, it requires 11.3 MB of memory, compared to 17.1 GB for a dense grid representation; this corresponds to a compression ratio of 1:1513. Although real-frequency diagrammatic calculations for the SIAM are limited by run-time rather than memory [20], this paves the way for investigation of more complicated models with multiple orbitals or momentum dependence, which have prohibitive memory requirements if attempted with dense grids [51–53].

In Fig. 9, we explore the compressibility of the core and full vertices for varying tolerance τ, inverse temperature β, and interaction strength u. As seen before in Fig. 8, lowering the tolerance below 10^{-2} results in a steep increase in the rank. As in the Matsubara formalism, the rank increases with β, but only slowly. Panel (c) reveals a much less predictable behavior: The ranks of both the full and core vertices reach a maximum at $u = 1.0$ and decrease significantly for large u. Moreover, the rank of the full vertex Γ approaches that of Γ_{core} from below with increasing interaction u. This reflects the increasing magnitude of Γ_{core} relative to the asymptotic contributions $\mathcal{K}^r_1$ and $\mathcal{K}^r_{2^{(\prime)}}$: At weak interaction, the magnitude

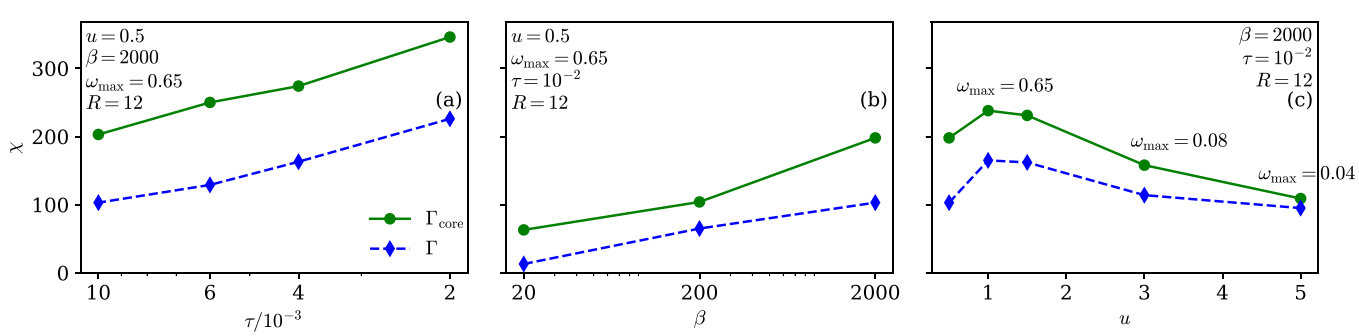

FIG. 9. Rank of Keldysh core vertex $\Gamma^{\uparrow\uparrow}_{\text{core}}$ and full vertex $\Gamma^{\uparrow\uparrow}$ in the p channel vs (a) TCI tolerance τ, (b) inverse temperature β, and (c) interaction strength u. We show data for the $\boldsymbol{k} = (2121)$ Keldysh component. In panel (c), we choose smaller box sizes $\omega_{\max} = 0.08$ and $\omega_{\max} = 0.04$ for $u = 3.0$ and $u = 5.0$, respectively. This is because the extent of the core vertex decreases at these strong interactions.

TABLE II. QTT ranks of the five Keldysh components not related by crossing symmetry or complex conjugation (center columns) compared to the QTT rank of a single tensor that contains all five components (rightmost column). Parameters are $u = 0.5$, $\beta = 2000$, $\omega_{\max} = 0.65$, $\tau = 10^{-2}$, and $R = 8$.

Component $\boldsymbol{k}$	1111	2111	2121	2112	1222	All
Rank χ	130	136	159	126	95	**344**

of the core vertex is much smaller than that of the full vertex, which is dominated by $\mathcal{K}_1^r$ and $\mathcal{K}_{2^{(\prime)}}^r$. Since TCI measures the error relative to the supremum norm of the target function [cf. Eq. (19)], this means that Γ_{core} need not be resolved as accurately at weak interaction. The compression of the asymptotic contributions $\mathcal{K}_1^r$ and $\mathcal{K}_{2^{(\prime)}}^r$ is discussed in Appendix B, together with their Matsubara counterparts. In Appendix C, we discuss how the ranks of $\Gamma_{\text{core}}^{\boldsymbol{k}}$ and $\Gamma^{\boldsymbol{k}}$ depend on flavor, frequency channel, and Keldysh component $\boldsymbol{k}$.

Finally, we discuss how compressing each Keldysh component of the vertex separately, as was done in this work, compares to running TCI on the entire Keldysh core or full vertex, where the Keldysh components are encoded in an additional leg of a single tensor train. In both cases, spin components are compressed separately. For a fair comparison of these two approaches, recall the following: (1) TCI measures the interpolation error relative to the supremum norm of the target function [see Eq. (19)]. When compressing the entire vertex, any given Keldysh component $\Gamma^{\boldsymbol{k}}$ (or $\Gamma_{\text{core}}^{\boldsymbol{k}}$) should therefore be normalized by $||\Gamma^{\boldsymbol{k}}||_\infty$ (or $||\Gamma_{\text{core}}^{\boldsymbol{k}}||_\infty$), i.e., with the supremum norm of the same Keldysh component $\boldsymbol{k}$. Only then does one achieve the same accuracy as in separate compressions of Keldysh components. (2) Matrix product operator contractions (MPO-MPO contractions), the most expensive operations in QTT-based diagrammatic calculations, scale as $\mathcal{O}(R\chi^4)$ in run-time. As a preliminary investigation, we compressed the entire core vertex at $u = 0.5$, $\beta = 2000$, and $\omega_{\max} = 0.65$ with a tolerance of $\tau = 10^{-2}$ and $R = 8$ quantics bits. The tensor leg for the Keldysh component was placed to the very left and only included the five Keldysh components not related by crossing symmetry or complex conjugation (cf. Appendix C).

The resulting rank χ is compared with the ranks $\chi_{\boldsymbol{k}}$ of individual Keldysh components in Table II. We observe the ratio $\chi^4/\max_{\boldsymbol{k}} \chi_{\boldsymbol{k}}^4 \approx 21.77$; thus, MPO-MPO contractions take about 22 times longer for two entire vertices than for two individual components. On the other hand, the latter type of contraction would have to be performed $5^2 = 25$ times. However, Table II shows that some Keldysh components have a significantly lower bond dimension than $\max_{\boldsymbol{k}} \chi_{\boldsymbol{k}}$. In summary, both approaches are worth investigating, and a conclusive comparison is only possible in the context of a specific QTT-based diagrammatic code.

IV. SUMMARY AND OUTLOOK

We have presented a QTCI-based method for representing imaginary- and real-frequency mpNRG vertex functions on large grids that are far beyond the reach of previous implementations. The QTCI algorithm allows us to automatically capture all relevant features of the vertex up to a prescribed accuracy and represents the result in a highly compressed format. Repeated sampling during TCI sweeps necessitates optimizations of the vertex evaluation, which we described in detail. We studied the compressibility of the vertex in a systematic fashion: Imaginary- and even real-frequency vertices are representable as QTTs with maximum bond dimensions sufficiently small ($\chi \approx$ a few hundred) to allow for diagrammatic computations with these objects. This holds true across a broad range of temperatures and interaction strengths, both for the full vertex and for its asymptotic decomposition. Our work thus constitutes an important step toward QTCI-based diagrammatic calculations that use a nonperturbative DMFT vertex as input and suggests that these will be feasible. The next step will be to implement a QTCI-based diagrammatic extension of DMFT that augments the local vertex with momentum dependence. An analogous program can be envisioned in the real-frequency setting. Though this is a challenging, computationally demanding endeavor, it would achieve a long-sought goal: a method to obtain nonlocal, real-frequency dynamical response functions of strongly correlated systems.

ACKNOWLEDGMENTS

We thank Seung-Sup Lee and Jae-Mo Lihm for helpful discussions. We thank Seung-Sup Lee for providing the mpNRG code used for computing the PSFs. This work was funded in part by the Deutsche Forschungsgemeinschaft under Germany's Excellence Strategy EXC-2111 (Project No. 390814868). It is part of the Munich Quantum Valley, supported by the Bavarian state government with funds from the Hightech Agenda Bayern Plus. We further acknowledge support from DFG Grant No. LE 3883/2-2. We gratefully acknowledge computational resources from Grant No. INST 86/1885-1 FUGG of the German Research Foundation (DFG) and from the GCS Supercomputer SuperMUC-NG at the Leibniz Supercomputing Centre in Munich provided by the Gauss Centre for Supercomputing e.V. M.F., A.G., M.R., and N.R. acknowledge support from the International Max-Planck Research School for Quantum Science and Technology (IMPRS-QST). N.R. acknowledges funding from a graduate scholarship from the German Academic Scholarship Foundation ("Studienstiftung des deutschen Volkes") and additional support from the "Marianne-Plehn-Programm" of the state of Bavaria.

DATA AVAILABILITY

The mpNRG computations were performed with the MUNRG package [12,54,55], which is based on the QSPACE tensor network library [56–59]. The latest version of QSPACE is available [60] and a public release of MUNRG is intended. The code used in this work to compute and compress vertices is available on GitHub [61]. The partial spectral functions required as input for that code can be found in Ref. [62].

TABLE III. Broadening settings for different NRG datasets. $T = 1/\beta$ denotes the temperature. See main text for definitions of the parameters. We set $e_{\min} = 10^{-6}$, $e_{\max} = 10^4$, and $e_{\text{step}} = 50$. For TCI tolerances $\tau \leqslant 3.4 \times 10^{-3}$, the integration grid was refined to $e_{\text{step}} = 200$ to avoid fitting of numerical noise by the TCI algorithm. We also set $e_{\text{step}} = 200$ to broaden 2p functions.

u	β	σ_{sLG}	γ_{L}
0.5	20/200	0.693	T
0.5/1.0/1.5	2000	0.4	T
3.0/5.0	2000	0.4	T

APPENDIX A: COMPUTING BROADENED KELDYSH KERNELS

In this Appendix, we provide details on the numerical computation of the broadened Keldysh kernels $K_b^{[\lambda]}$ appearing in Eq. (10b). It consists of two steps: broadening the Dirac-δ functions in Eq. (3) to $\delta_b(\omega', \epsilon)$ and convolution of $\delta_b(\omega', \epsilon)$ with the Keldysh kernel $(\omega' + \mathrm{i}\gamma_{0,i}^{\lambda})^{-1}$. The numerical details of this procedure are taken from the MUNRG code of Ref. [12].

The broadening combines symmetric log-Gaussian and linear broadening (cf. Appendix E 2 of Ref. [15]):

$$\delta_b(\omega', \epsilon) = \int_{\mathbb{R}} \mathrm{d}\epsilon' \delta_{\text{F}}(\omega', \epsilon') \delta_{\text{sLG}}(\epsilon', \epsilon), \tag{A1a}$$

$$\delta_{\text{sLG}}(\epsilon', \epsilon) = \frac{\Theta(\epsilon'\epsilon)}{\sqrt{\pi}\sigma_{\text{sLG}}|\epsilon|} \exp\left[-\left(\frac{\ln|\epsilon/\epsilon'|}{2\sigma_{\text{sLG}}} - \frac{\sigma_{\text{sLG}}}{4}\right)^2\right], \tag{A1b}$$

$$\delta_{\text{F}}(\omega', \epsilon') = \frac{1}{2\gamma_{\text{L}}}\left(1 + \cosh\frac{\omega' - \epsilon'}{\gamma_{\text{L}}}\right)^{-1}. \tag{A1c}$$

The broadening parameters γ_{L} and σ_{sLG} along with other numerical settings for all physical parameter sets are listed in Table III. However, the linear broadening γ_L is multiplied with a prefactor that depends on the current permutation p, the fully retarded index λ, and the dimension i. This scheme will be explained further at the end of this Appendix [see Eq. (A4)]. The Integral (A1a) is performed by trapezoidal quadrature, where ω' and ϵ' are discretized on logarithmic grids. These grids contain 0, are symmetric around the origin, and range from $e_{\min}$ to $e_{\max}$ with e_{step} points per decade (see Table III). For the ϵ' grid, $e_{\min}$ is automatically replaced by a lower boundary $x_{\min}$ with $0 < x_{\min} < e_{\min}$ if the low-frequency tail of the log-Gaussian broadening kernel extends below $e_{\min}$. This ensures an accurate integration of $\delta_{\text{sLG}}(\epsilon', \epsilon)$. The energies ϵ specifying the location of the spectral function peaks also reside on a logarithmic grid, which arises from the mpNRG computation.

The numerical integration described above yields $\delta_b(\omega', \epsilon)$ with ω' and ϵ on logarithmic grids. To convolve δ_b with the Keldysh kernel $(\omega' + \mathrm{i}\gamma_{0,i}^{\lambda})^{-1}$, we use the identity

$$\lim_{\gamma_0 \to 0^+} \int_{\mathbb{R}} \mathrm{d}\omega' \frac{\delta_b(\omega', \epsilon)}{\omega - \omega' + \mathrm{i}\gamma_{0,i}^{\lambda}} = \mathcal{P}\int_{\mathbb{R}} \mathrm{d}\omega' \frac{\delta_b(\omega', \epsilon)}{\omega - \omega'} - \mathrm{i}\pi\,\mathrm{sgn}\big(\gamma_{0,i}^{\lambda}\big)\delta_b(\omega, \epsilon), \tag{A2}$$

TABLE IV. Linear broadening of a 3p correlator with doubled broadening on the composite operator q_{ij}. The operator q_{ij} is in the first slot of the operator tuple for the identity permuation $p = [123]$. Left: Permutation $p = [123]$. Right: Permutation $p = [213]$.

	$i = 1$	$i = 2$		$i = 1$	$i = 2$
$\lambda = 1$	$2\gamma_{\text{L}}$	γ_{L}	$\lambda = 1$	$3\gamma_{\text{L}}$	γ_{L}
$\lambda = 2$	$2\gamma_{\text{L}}$	γ_{L}	$\lambda = 2$	γ_{L}	γ_{L}
$\lambda = 3$	$2\gamma_{\text{L}}$	$3\gamma_{\text{L}}$	$\lambda = 3$	γ_{L}	$3\gamma_{\text{L}}$

where $\mathcal{P}$ denotes the Cauchy principal value (PV) integral. Also, recall the Definition (10c) of $\gamma_{0,i}^{\lambda}$. Importantly, $\delta_b(\omega', \epsilon)$ has been computed on a logarithmic grid, while ω in the broadened Keldysh kernel $K_b^{[\lambda]}(\omega, \epsilon)$ defined in Eq. (10b) resides on a linear grid. This is because the external frequency grids on which we compute vertices are also linear. To obtain the imaginary part of Eq. (A2) on the linear grid, we use linear interpolation of $\delta_b(\omega, \epsilon)$ in the argument ω. Computing the real part, i.e., the PV integral, is slightly more involved: By the linear interpolation performed for the imaginary part, $\delta_b(\omega, \epsilon)$ can be viewed as a piecewise linear function. We split the PV integral over $\delta_b(\omega', \epsilon)$ into PV integrals over linear functions $(a_{i,\epsilon}\omega' + b_{i,\epsilon})$ on intervals $[\omega_i', \omega_{i+1}']$. These are evaluated using the formula

$$\mathcal{P}\int_{\omega_i'}^{\omega_{i+1}'} \mathrm{d}\omega' \frac{a_{i,\epsilon}(\omega' - \omega_i) + b_{i,\epsilon}}{\omega - \omega'} = -a_{i,\epsilon}(\omega_{i+1}' - \omega_i') - (a_{i,\epsilon}(\omega - \omega_i') + b_{i,\epsilon}) \ln\left|\frac{\omega - \omega_{i+1}'}{\omega - \omega_i'}\right|. \tag{A3}$$

The sum over all PV integrals of the form (A3) then yields the real part of the broadened Keldysh kernel $K_b(\omega, \epsilon)$. For this scheme to be accurate, the extent $[-e_{\max}, e_{\max}]$ of the logarithmic ω' grid should be significantly larger than the frequency box delimited by $\omega_{\max}$. Comparing the values of $e_{\max}$ given in Table III with our default frequency box size $\omega_{\max} = 0.65$, one verifies that this is the case.

We now turn to the prefactors of the linear broadening γ_{L} mentioned above. In the linear broadening kernel $\delta_{\text{F}}(\omega_{\bar{1}\ldots\bar{i}}', \epsilon_i')$ to be convolved with the Keldysh kernel $(\omega_{\bar{1}\ldots\bar{i}}' + \gamma_{0,i}^{\lambda})^{-1}$, the broadening with γ_{L} is replaced by

$$\gamma_{\text{L},i}^{\lambda} = \begin{cases} \gamma_{\text{L}} \cdot (\ell - i), & \text{for } i \geqslant \lambda, \\ \gamma_{\text{L}} \cdot i, & \text{for } i < \lambda. \end{cases} \tag{A4}$$

This choice was found to reduce broadening artifacts in an mpNRG treatment of the Hubbard atom in Ref. [63]. Moreover, composite operators q_{ij} in 3p correlators [cf. Eq. (96) of Ref. [15]] receive a doubled linear broadening. This was found to cancel discretization and broadening artifacts when computing $\mathcal{K}_{2^{(\prime)}}^r$ by multiplication with self-energies; see Eq. (B3) and Ref. [63]. The broadening of 3p correlators is exemplified in Table IV. Finally, the 2p correlators required for self-energies in the symmetric estimators for Γ_{core} and $K_{2^{(\prime)}}^r$ (cf. Sec. II D and Appendix B.) are broadened according to Table V.

TABLE V. Linear broadening of a 2p function used for the aIE self-energy. The two rightmost columns correspond to the two possible permutations.

	$i = 1, p = [12]$	$i = 1, p = [21]$
$\lambda = 1$	$3\gamma_L$	γ_L
$\lambda = 2$	γ_L	$3\gamma_L$

APPENDIX B: COMPRESSION OF 1D AND 2D VERTEX CONTRIBUTIONS

A QTCI-based parquet calculation exploiting the asymptotic decomposition (11) requires not only Γ_{core}, but also $\mathcal{K}_1^r$ and $\mathcal{K}_{2^{(\prime)}}^r$ represented as QTTs. Recall that $r = a, p, t$ labels the three frequency channels. In this Appendix, we verify that these asymptotic contributions indeed have a significantly lower rank than Γ_{core}, as expected from their simpler structure.

The $\mathcal{K}_1^r$ contributions are simply given by 2p correlators of composite operators (see Sec. IV F of Ref. [15]). We illustrate the evaluation of $\mathcal{K}_{2^{(\prime)}}^r$ using $\mathcal{K}_2^t$ as an example. For derivations and the remaining $\mathcal{K}_{2^{(\prime)}}^r$ components, we refer to Secs. IV C and IV F of Ref. [15]. Since the self-energy is spin-diagonal, we again omit spin indices. First, the operator $q = [d, H_{\text{int}}]$ introduced in Sec. II D is used to define the operator $q_{34} = \{q, d^\dagger\} = qd^\dagger + d^\dagger q$. One further introduces 3p correlators $G[q_{34}, a_1, a_2^\dagger]$, where $a_1, a_2 \in \{d, q\}$. They are defined in terms of connected correlators as

$$G^{\boldsymbol{k}}[q_{34}, a_1, a_2^\dagger] = P^{k_1 k_2 (k_3+k_4)} G_{\text{con}}^{\boldsymbol{k}}[q_{34}, a_1, a_2^\dagger]. \tag{B1}$$

In the Keldysh formalism, the tensor P reads

$$P^{k_1 k_2 (k_3+k_4)} = \begin{cases} \frac{1}{\sqrt{2}}, & \text{if } \sum_i k_i \text{ is odd,} \\ 0, & \text{else,} \end{cases} \tag{B2}$$

while it is set to unity in Matsubara. Using the symbol Y_{x_i} introduced in Eq. (13), K_2^t can then be expressed as

$$\mathcal{K}_2^t(\omega_t, \nu_t) = \sum_{a_1, a_2 \in \{d,q\}} Y_{a_1} G[q_{34}, a_1, a_2^\dagger](-\omega_{12}, \omega_1, \omega_2) Y_{a_2}. \tag{B3}$$

To evaluate $\mathcal{K}_2^t$, the external frequencies ω_1, ω_2 appearing on the right-hand side of Eq. (B3) are expressed in the t-channel parametrization, i.e., in terms of ω_t and ν_t.

The ranks of Matsubara and Keldysh asymptotic contributions $\mathcal{K}_1^{t,\uparrow\uparrow}$ and $\mathcal{K}_2^{t,\uparrow\uparrow}$ for different parameters are shown in Figs. 10(a)–10(c) and 10(d)–10(f), respectively. We use the t-channel frequency parametrization, thus viewing $\mathcal{K}_1^t(\omega_t)$ as a 1D and $\mathcal{K}_2^t(\omega_t, \nu_t)$ as a 2D function. A comparison of Fig. 10 with Figs. 5 and 9 confirms that the three-dimensional vertex functions will dominate the cost of a diagrammatic calculation: For a tolerance of $\tau = 10^{-3}$, the ranks of the Matsubara $\mathcal{K}_2^t$ component are no larger than 20. The variation of the rank with β and u does therefore not bare much significance. For smaller tolerances, the rank of $\mathcal{K}_2^t$ remains much lower than that of the core vertex. Analogous observations hold for the Keldysh $\mathcal{K}_2^t$ component with a target tolerance of $\tau = 10^{-2}$. Like the Keldysh core vertex, its rank increases slowly with β and decreases for strong coupling u, but the changes are small compared to the core vertex.

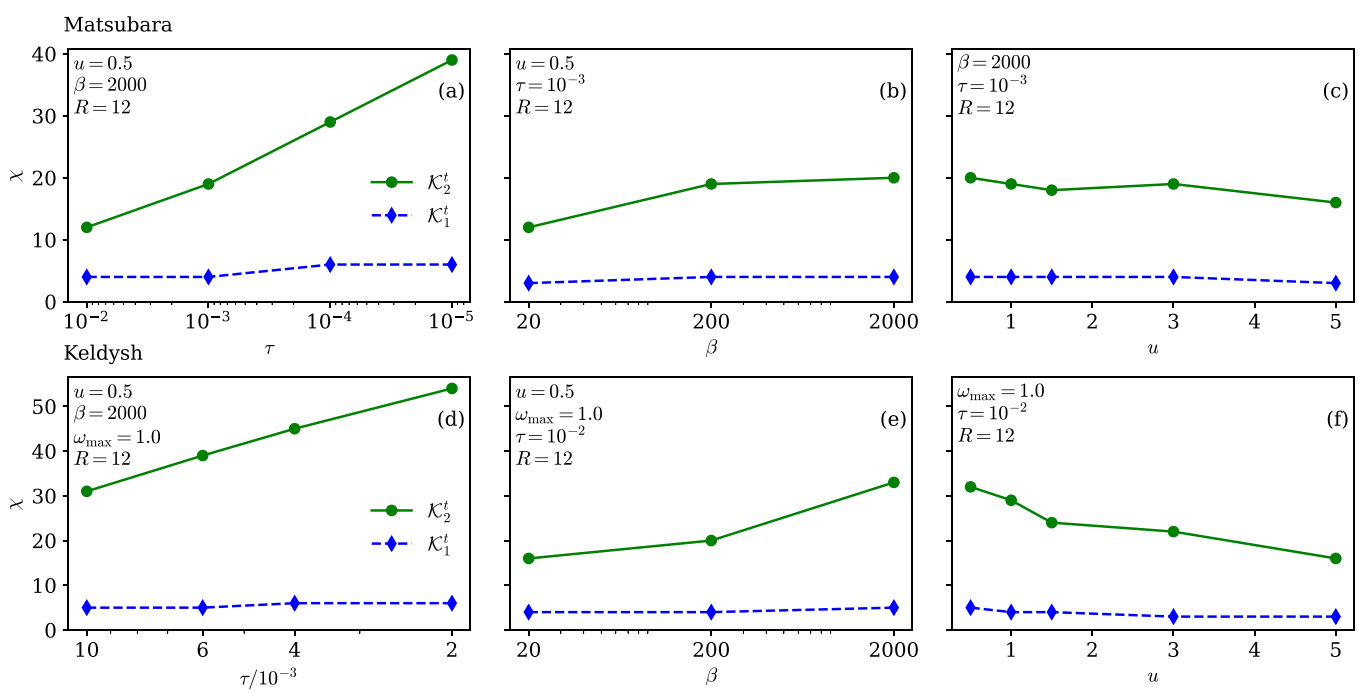

FIG. 10. QTT ranks of Matsubara [top row, panels (a)–(c)] and Keldysh [bottom row, panels (d)–(f)] $\mathcal{K}_1^{t,\uparrow\uparrow}$ and $\mathcal{K}_2^{t,\uparrow\uparrow}$ contributions to the full vertex vs: (a), (d) tolerance τ, (b), (e) inverse temperature β, and (c), (f) interaction strength u. In Keldysh we chose, of all components, the component $\boldsymbol{k} = (22)$ for $\mathcal{K}_1^{t,\uparrow\uparrow}$ and $\boldsymbol{k} = (112)$ for $\mathcal{K}_2^{t,\uparrow\uparrow}$. These components were found to have the highest rank, respectively.

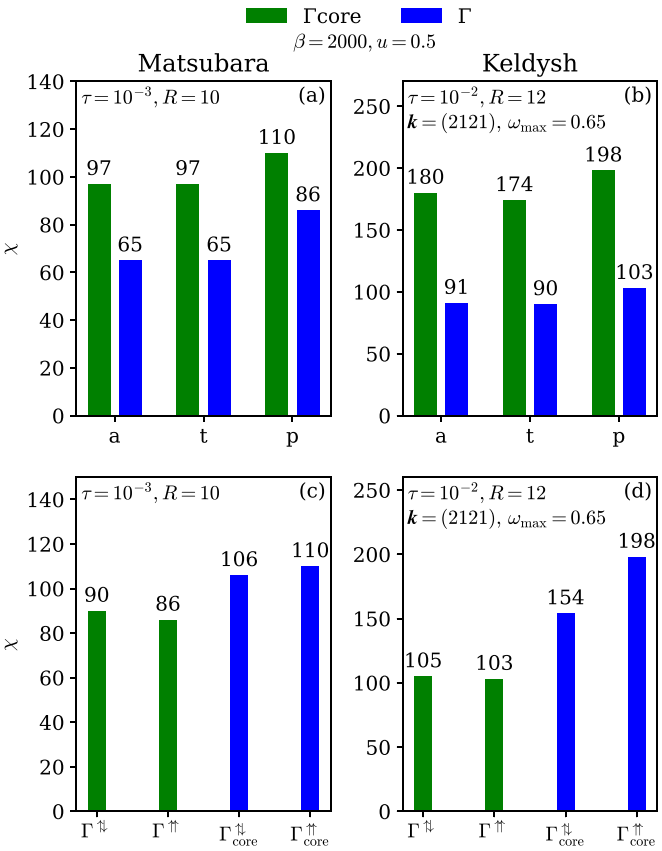

FIG. 11. Top row [panels (a) and (b)]: TCI ranks of $\Gamma^{\uparrow\uparrow}$ and $\Gamma^{\uparrow\uparrow}_{\text{core}}$ in the three channels a, p, t. Bottom row [panels (c) and (d)]: TCI ranks of Γ and Γ_{core} in the p channel for the two flavors $\uparrow\uparrow$ and $\uparrow\downarrow$. Matsubara vertices [panels (a) and (c)] were compressed with $\tau = 10^{-3}$ and $R = 10$, and Keldysh vertices [panels (b) and (d)] with $\tau = 10^{-2}$ and $R = 12$.

APPENDIX C: COMPRESSION FOR DIFFERENT CHANNELS, FLAVORS, AND KELDYSH COMPONENTS

In this Appendix, we investigate the compressibility of core and full vertices for different flavors ($\uparrow\uparrow, \uparrow\downarrow$), frequency channels ($r = a, p, t$), and Keldysh components $\boldsymbol{k} = (k_1k_2k_3k_4)$.

Figure 11 shows the QTT ranks of the Matsubara and Keldysh full and core vertices for different channels and flavors, at $\beta = 2000$ and $u = 0.5$. The tolerances are $\tau = 10^{-3}$ and $\tau = 10^{-2}$ for Matsubara and Keldysh vertices, respectively. We observe that the p channel exhibits the highest ranks throughout, which is why we used this frequency parametrization in the main text. The QTT ranks of the Matsubara vertices shown in Fig. 11(c) barely differ between the two flavors. In contrast, $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}$ has a significantly higher rank than $\Gamma^{2121,\uparrow\downarrow}_{\text{core}}$ ($\chi = 198$ versus $\chi = 154$).

The rank of $\Gamma^{\uparrow\uparrow}_{\text{core}}$ and $\Gamma^{\uparrow\uparrow}$ depending on the Keldysh component is shown in Fig. 12. Only components that are not related by crossing or complex conjugation symmetry [50] are considered. We show data for $\beta = 2000$, $u = 0.5$, and $\tau = 10^{-3}$. The (2121) component of Γ_{core} is found to have the highest rank. We therefore selected $\Gamma^{2121,\uparrow\uparrow}_{\text{core}}$ for our analysis of rank saturation and parameter dependence in Figs. 8 and 9.

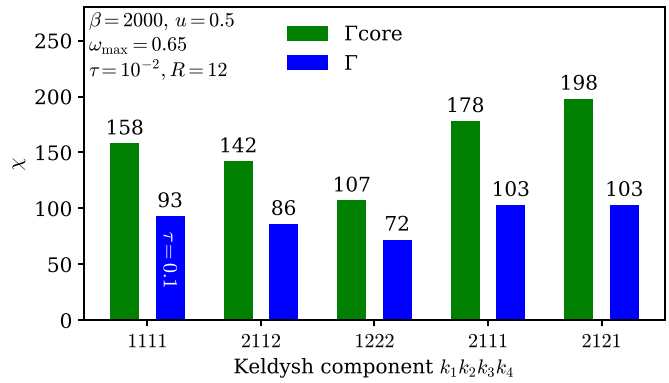

FIG. 12. QTT rank of Keldysh core vertex $\Gamma^{\boldsymbol{k},\uparrow\uparrow}_{\text{core}}$ in the p channel vs Keldysh component. From the 16 Keldysh components, only those are shown that are not related by crossing symmetry or complex conjugation. The $\Gamma^{1111,\uparrow\uparrow}$ component of the full vertex was compressed with tolerance $\tau = 0.1$, since it is about a factor of 10 smaller than the other Keldysh components of the full vertex. This is because the $\mathcal{K}^r_1$ contributions, which dominate other components $\Gamma^{\boldsymbol{k}\neq 1111}$, vanish in the $\boldsymbol{k} = (1111)$ component.

APPENDIX D: FREQUENCY CONVENTIONS

We use the following parametrizations for the t (particle-hole), p (particle-particle), and a (transverse particle-hole) channels:

$$\boldsymbol{\omega} = \begin{cases} (-\nu_r, \omega_r + \nu_r, -\omega_r - \nu'_r, \nu'_r), & \text{for } r = t \text{ (ph)}, \\ (-\nu_r, \omega_r - \nu'_r, -\omega_r + \nu_r, \nu'_r), & \text{for } r = p \text{ (pp)}, \\ (-\nu_r, \nu'_r, -\omega_r - \nu'_r, \omega_r + \nu_r), & \text{for } r = a \text{ (}\overline{\text{ph}}\text{)}. \end{cases} \tag{D1}$$

These are the same as in Ref. [15], up to a global minus sign. Spin and, if present, Keldysh indices of the vertex $\Gamma^{\boldsymbol{k},\sigma_1\sigma_2\sigma_3\sigma_4}(\boldsymbol{\omega})$ are ordered according to the underlying impurity Green's function $G^{\boldsymbol{k}}_{\text{con}}[d_{\sigma_1}d^\dagger_{\sigma_2}d_{\sigma_3}d^\dagger_{\sigma_4}](\boldsymbol{\omega})$. Finally, the evaluation of a 2p correlator G at a frequency ν is defined as $G(\nu, -\nu)$ in our convention. This is relevant for evaluating the self-energy Σ in Eq. (13), because computing the self-energy comes down to evaluating 2p correlators according to the asymmetric estimators we employed; see Eq. (27) in Ref. [15].

APPENDIX E: EVALUATING THE KELDYSH VERTEX ON NONLINEAR GRIDS

As mentioned in Sec. II A, the PSFs consist of spectral peaks residing on a logarithmic energy grid. This raises the question whether the vertices could also be computed on logarithmic (or arbitrary nonlinear) grids. Since the Matsubara vertex is inherently defined on an equidistant grid, we consider this question to be relevant only for the Keldysh vertex. While the spectral representation and symmetric estimators yield the Keldysh vertex at arbitrary frequency points, evaluating the vertex on a nonlinear grid turns out to be much more expensive. This Appendix explains why that is the case.

Consider a grid $W \subset \mathbb{R}$, where the frequencies $\omega \in W$ may be spaced, e.g., logarithmically. Suppose we want to evaluate the Keldysh vertex on points $\boldsymbol{\omega} \in W^3 \subset \mathbb{R}^3$. As detailed in Secs. II C and II D, this entails the evaluation of full correlators $G^{\boldsymbol{k}}(\boldsymbol{\omega})$ with $\boldsymbol{\omega} \in W^3$, which in turn requires $G^{[\lambda]}_p(\boldsymbol{\omega}_p)$

[see Eq. (29a)]. Crucially, the broadened kernels $k_b^{[\lambda,i]}$ are evaluated at the transformed frequency $\omega_{\bar{1}\ldots\bar{i}}$ in Eq. (29a). Let us denote this linear transformation $\boldsymbol{\omega} \mapsto (\omega_{\bar{1}}, \omega_{\bar{1}\bar{2}}, \omega_{\bar{1}\bar{2}\bar{3}})$ for 4p correlators by T_p. The broadened kernels $k_b^{[\lambda,i]}$ must be evaluated on points in the image of the chosen grid W^3 under T_p, i.e., $T_p(W^3) \ni \boldsymbol{\omega}_p$. In order to save computation time by precomputing the kernels, one can embed the image grid in a Cartesian product of one-dimensional grids $W_{p,1}, W_{p,2}, W_{p,3} \subset \mathbb{R}$, i.e., $T_p(W^3) \subset W_{p,1} \times W_{p,2} \times W_{p,3}$. It is the grids $W_{p,i}$ on which the kernels are broadened and, if desired, SVD compressed as explained in Sec. II F 2. In the case of an *equidistant* grid, the $W_{p,i}$ have at most $3 \cdot |W| = \mathcal{O}(|W|)$ points.

In constrast, if W is nonlinear, one has $|W_{p,i}| = \mathcal{O}(|W|^3)$ in general. This means that, for large grid sizes, precomputing and storing the broadened kernels $k_b^{[\lambda,i]}$ on W_p is no longer possible. The same holds for the SVD compression of the kernels. This issue leaves two alternatives for evaluating the Keldysh vertex $\Gamma^{\boldsymbol{k}}$ on nonlinear grids:

(1) For each individual point $\boldsymbol{\omega} \in W^3$, compute the broadened kernels upon evaluating $\Gamma^{\boldsymbol{k}}(\boldsymbol{\omega})$. This is not implemented in our code [61], but we expect it to be prohibitively expensive even if caching of previously computed kernel values was introduced.

(2) Precompute the kernels on a very fine, equidistant grid and evaluate the vertex by interpolation from linear grids. This approach is implemented in our code [61] and has been employed in Ref. [16] when investigating the fulfillment of diagrammatic identities by mpNRG data.

TABLE VI. QTT ranks of different components of the t-channel Keldysh core vertex $\Gamma^{\uparrow\uparrow}_{\text{core}}$ computed on a logarithmic grid and, for comparison, on an $R = 15$ linear quantics grid. The tolerance is $\tau = 10^{-2}$.

$\boldsymbol{k}$	1111	2111	2121	2112	1222
$\chi_{\log}$	258	475	718	379	445
χ_{lin}	129	150	177	127	101

In conclusion, we do not see an affordable method of evaluating the Keldysh vertex "directly" on a large, nonlinear grid. However, interpolation from a linear grid provides a viable alternative: Using QTCI, one can determine at which grid size the vertex is resolved to the desired accuracy and then use the resulting quantics tensor train for interpolation.

Finally, we report a preliminary investigation of the QTCI compressibility of Keldysh vertex components stored on logarithmic grids: We compressed different Keldysh components of the t-channel Keldysh core vertex at $\beta = 2000$ and $u = 0.50$ computed on a $243 \times 243 \times 243$ logarithmic grid. The TCI tolerance was $\tau = 10^{-2}$. The grid spans a large frequency range from -3.18 to 3.18. The vertex was interpolated trilinearly from a quantics grid with $R = 15$ bits (see Ref. [16]). The results are summarized in Table VI, showing large QTT ranks $\chi_{\log}$ compared to the QTT ranks on a linear grid, χ_{lin}. This indicates that a QTCI compression of Keldysh vertex data on a logarithmic grid is not effective.

[1] W. Metzner, M. Salmhofer, C. Honerkamp, V. Meden, and K. Schönhammer, Functional renormalization group approach to correlated fermion systems, Rev. Mod. Phys. **84**, 299 (2012).

[2] F. B. Kugler and J. von Delft, Multiloop functional renormalization group that sums up all parquet diagrams, Phys. Rev. Lett. **120**, 057403 (2018).

[3] F. B. Kugler and J. von Delft, Multiloop functional renormalization group for general models, Phys. Rev. B **97**, 035162 (2018).

[4] F. B. Kugler and J. von Delft, Derivation of exact flow equations from the self-consistent parquet relations, New J. Phys. **20**, 123029 (2018).

[5] N. E. Bickers, Self-Consistent Many-Body Theory for Condensed Matter Systems, in *Theoretical Methods for Strongly Correlated Electrons, CRM Series in Mathematical Physics*, edited by D. Sénéchal, A.-M. Tremblay, and C. Bourbonnais (Springer, New York, 2004), pp. 237–296.

[6] G. Rohringer, H. Hafermann, A. Toschi, A. A. Katanin, A. E. Antipov, M. I. Katsnelson, A. I. Lichtenstein, A. N. Rubtsov, and K. Held, Diagrammatic routes to nonlocal correlations beyond dynamical mean field theory, Rev. Mod. Phys. **90**, 025003 (2018).

[7] A. Georges, G. Kotliar, W. Krauth, and M. J. Rozenberg, Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, Rev. Mod. Phys. **68**, 13 (1996).

[8] C. Taranto, S. Andergassen, J. Bauer, K. Held, A. Katanin, W. Metzner, G. Rohringer, and A. Toschi, From infinite to two dimensions through the functional renormalization group, Phys. Rev. Lett. **112**, 196402 (2014).

[9] A. Toschi, A. A. Katanin, and K. Held, Dynamical vertex approximation: A step beyond dynamical mean-field theory, Phys. Rev. B **75**, 045118 (2007).

[10] K. Held, A. A. Katanin, and A. Toschi, Dynamical vertex approximation: An introduction, Prog. Theor. Phys. Suppl. **176**, 117 (2008).

[11] F. B. Kugler, S.-S. B. Lee, and J. von Delft, Multipoint correlation functions: Spectral representation and numerical evaluation, Phys. Rev. X **11**, 041006 (2021).

[12] S.-S. B. Lee, F. B. Kugler, and J. von Delft, Computing local multipoint correlators using the numerical renormalization group, Phys. Rev. X **11**, 041007 (2021).

[13] K. G. Wilson, The renormalization group: Critical phenomena and the Kondo problem, Rev. Mod. Phys. **47**, 773 (1975).

[14] R. Bulla, T. A. Costi, and T. Pruschke, Numerical renormalization group method for quantum impurity systems, Rev. Mod. Phys. **80**, 395 (2008).

[15] J.-M. Lihm, J. Halbinger, J. Shim, J. von Delft, F. B. Kugler, and S.-S. B. Lee, Symmetric improved estimators for multipoint vertex functions, Phys. Rev. B **109**, 125138 (2024).

[16] N. Ritz, A. Ge, M. Frankenbach, M. Pelz, J. von Delft, and F. B. Kugler, Testing the parquet equations and the U(1) Ward identity for real-frequency correlation functions from the multipoint numerical renormalization group, Phys. Rev. Res. **7**, 033139 (2025).

[17] H. R. Krishna-murthy, J. W. Wilkins, and K. G. Wilson, Renormalization-group approach to the Anderson model of dilute magnetic alloys. I. Static properties for the symmetric case, Phys. Rev. B **21**, 1003 (1980).

[18] R. Bulla, A. C. Hewson, and T. Pruschke, Numerical renormalization group calculations for the self-energy of the impurity Anderson model, J. Phys.: Condens. Matter **10**, 8365 (1998).

[19] K.-M. Tam, H. Fotso, S.-X. Yang, T.-W. Lee, J. Moreno, J. Ramanujam, and M. Jarrell, Solving the parquet equations for the Hubbard model beyond weak coupling, Phys. Rev. E **87**, 013311 (2013).

[20] A. Ge, N. Ritz, E. Walter, S. Aguirre, J. von Delft, and F. B. Kugler, Real-frequency quantum field theory applied to the single-impurity Anderson model, Phys. Rev. B **109**, 115128 (2024).

[21] N. Ritz, A. Ge, E. Walter, S. Aguirre, J. von Delft, and F. B. Kugler, KeldyshQFT: A C++ codebase for real-frequency multiloop functional renormalization group and parquet computations of the single-impurity Anderson model, J. Chem. Phys. **161**, 054118 (2024).

[22] I. V. Oseledets, Approximation of matrices with logarithmic number of parameters, Dokl. Math. **80**, 653 (2009).

[23] B. N. Khoromskij, $O(d \log N)$ quantics approximation of N-d tensors in high-dimensional numerical modeling, Constructive Approximation **34**, 257 (2011).

[24] E. Ye and N. F. G. Loureiro, Quantum-inspired method for solving the Vlasov-Poisson equations, Phys. Rev. E **106**, 035208 (2022).

[25] H. Shinaoka, M. Wallerberger, Y. Murakami, K. Nogaki, R. Sakurai, P. Werner, and A. Kauch, Multiscale space-time ansatz for correlation functions of quantum systems based on quantics tensor trains, Phys. Rev. X **13**, 021015 (2023).

[26] N. Jolly, Y. N. Fernández, and X. Waintal, Tensorized orbitals for computational chemistry, Phys. Rev. B **111**, 245115 (2025).

[27] M. Środa, K. Inayoshi, H. Shinaoka, and P. Werner, High-resolution nonequilibrium GW calculations based on quantics tensor trains, arXiv:2412.14032.

[28] S. Rohshap, M. K. Ritter, H. Shinaoka, J. von Delft, M. Wallerberger, and A. Kauch, Two-particle calculations with quantics tensor trains: Solving the parquet equations, Phys. Rev. Res. **7**, 023087 (2025).

[29] L. Hölscher, P. Rao, L. Müller, J. Klepsch, A. Luckow, T. Stollenwerk, and F. K. Wilhelm, Quantum-inspired fluid simulation of two-dimensional turbulence with GPU acceleration, Phys. Rev. Res. **7**, 013112 (2025).

[30] I. V. Oseledets, Tensor-train decomposition, SIAM J. Sci. Comput. **33**, 2295 (2011).

[31] D. Savostyanov and I. Oseledets, Fast adaptive interpolation of multi-dimensional arrays in tensor train format, in *2011 International Workshop on Multidimensional (nD) Systems, Poitiers, France* (IEEE, New York, 2011), pp. 1–8.

[32] I. Oseledets and E. Tyrtyshnikov, TT-cross approximation for multidimensional arrays, Linear Algebra Appl. **432**, 70 (2010).

[33] M. K. Ritter, Y. Núñez Fernández, M. Wallerberger, J. von Delft, H. Shinaoka, and X. Waintal, Quantics tensor cross interpolation for high-resolution parsimonious representations of multivariate functions, Phys. Rev. Lett. **132**, 056501 (2024).

[34] N. Wentzell, G. Li, A. Tagliavini, C. Taranto, G. Rohringer, K. Held, A. Toschi, and S. Andergassen, High-frequency asymptotics of the vertex function: Diagrammatic parametrization and algorithmic implementation, Phys. Rev. B **102**, 085106 (2020).

[35] L. V. Keldysh, Diagram technique for nonequilibrium processes, in *Selected Papers of Leonid V Keldysh* (World Scientific, Singapore, 2023), pp. 47–55.

[36] J. Schwinger, Brownian motion of a quantum oscillator, J. Math. Phys. **2**, 407 (1961).

[37] L. P. Kadanoff and G. A. Baym, *Quantum Statistical Mechanics* (Benjamin, New York, 1962).

[38] Y. Núñez Fernández, M. K. Ritter, M. Jeannin, J.-W. Li, T. Kloss, T. Louvet, S. Terasaki, O. Parcollet, J. von Delft, H. Shinaoka, and X. Waintal, Learning tensor networks with tensor cross interpolation: New algorithms and libraries, SciPost Phys. **18**, 104 (2025).

[39] C.-W. Ha, Eigenvalues of differentiable positive definite kernels, SIAM J. Math. Anal. **17**, 415 (1986).

[40] N. Chikano, J. Otsuki, and H. Shinaoka, Performance analysis of a physically constructed orthogonal representation of imaginary-time Green's function, Phys. Rev. B **98**, 035104 (2018).

[41] S. Dirnböck, S.-S. B. Lee, F. B. Kugler, S. Huber, J. von Delft, K. Held, and M. Wallerberger, Overcomplete intermediate representation of two-particle Green's functions and its relation to partial spectral functions, Phys. Rev. Res. **6**, 043228 (2024).

[42] H. Shinaoka, J. Otsuki, M. Ohzeki, and K. Yoshimi, Compressing Green's function using intermediate representation between imaginary-time and real-frequency domains, Phys. Rev. B **96**, 035147 (2017).

[43] M. Wallerberger, H. Shinaoka, and A. Kauch, Solving the Bethe-Salpeter equation with exponential convergence, Phys. Rev. Res. **3**, 033168 (2021).

[44] Tensor4all, https://tensor4all.org.

[45] M. Filippone, C. P. Moca, A. Weichselbaum, J. von Delft, and C. Mora, At which magnetic field, exactly, does the Kondo resonance begin to split? A Fermi liquid description of the low-energy properties of the Anderson model, Phys. Rev. B **98**, 075404 (2018).

[46] P. W. Anderson, Localized magnetic states in metals, Phys. Rev. **124**, 41 (1961).

[47] G. Rohringer, A. Valli, and A. Toschi, Local electronic correlation at the two-particle level, Phys. Rev. B **86**, 125114 (2012).

[48] H. Takahashi, R. Sakurai, and H. Shinaoka, Compactness of quantics tensor train representations of local imaginary-time propagators, SciPost Phys. **18**, 007 (2025).

[49] F. D. Picoli, M. Huang, A. Gleis, and J. von Delft, Tangent-space Krylov solver for quantum impurity systems (unpublished).

[50] S. G. Jakobs, M. Pletyukhov, and H. Schoeller, Properties of multi-particle Green's and vertex functions within Keldysh formalism, J. Phys. A: Math. Theor. **43**, 103001 (2010).

[51] C. J. Eckhardt, C. Honerkamp, K. Held, and A. Kauch, Truncated unity parquet solver, Phys. Rev. B **101**, 155104 (2020).

[52] F. Krien and A. Kauch, The plain and simple parquet approximation: Single-and multi-boson exchange in the two-dimensional Hubbard model, Eur. Phys. J. B **95**, 69 (2022).

[53] G. Li, A. Kauch, P. Pudleiner, and K. Held, The victory project v1.0: An efficient parquet equations solver, Comput. Phys. Commun. **241**, 146 (2019).

[54] S.-S. B. Lee and A. Weichselbaum, Adaptive broadening to improve spectral resolution in the numerical renormalization group, Phys. Rev. B **94**, 235127 (2016).
[55] S.-S. B. Lee, J. von Delft, and A. Weichselbaum, Doublon-holon origin of the subpeaks at the Hubbard band edges, Phys. Rev. Lett. **119**, 236402 (2017).
[56] A. Weichselbaum, Non-Abelian symmetries in tensor networks: A quantum symmetry space approach, Ann. Phys. **327**, 2972 (2012).
[57] A. Weichselbaum, Tensor networks and the numerical renormalization group, Phys. Rev. B **86**, 245124 (2012).
[58] A. Weichselbaum, X-symbols for non-Abelian symmetries in tensor networks, Phys. Rev. Res. **2**, 023385 (2020).
[59] A. Weichselbaum, QSpace—An open-source tensor library for Abelian and non-Abelian symmetries, SciPost Phys. Codebases 40 (2024).
[60] A. Weichselbaum, Codebase release 4.0 for QSpace, SciPost Phys. Codebases 40 (2024).
[61] M. Frankenbach and A. Ge, TCI4Keldysh: A Julia code for computing imaginary- and real- frequency local four-point vertices from mpNRG using QTCI (2025), https://github.com/MFrankenbach/TCI4Keldysh.
[62] Data for "Computing and compressing local vertex functions in real and imaginary frequencies from the multipoint renormalization group using quantics tensor cross interpolation", https://doi.org/10.5282/ubm/data.613.
[63] J.-M. Lihm and S.-S. Lee (private communication).