A non-iterative method for vertex corrections of the Kubo formula for electric conductivity

Yi-Wen Wei, Chao-Kai Li, Yuchuang Cao, Ji Feng

International Center for Quantum Materials, School of Physics, Peking University, Beijing 100871, China
Collaborative Innovation Center of Quantum Matter, Beijing 100871, China
CAS Center for Excellence in Topological Quantum Computation, University of Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 14 May 2020
Received in revised form 8 August 2020
Accepted 16 August 2020
Available online 22 August 2020

Keywords:
Quantum transport
Kubo formula
Vertex corrections
Bethe–Salpeter equation
Rank factorization

ABSTRACT

When computing electric conductivity based on the Kubo formula, vertex corrections describe such effects as anisotropic scattering and quantum interference and are important in determining quantum transport properties. These vertex corrections are obtained by solving the Bethe–Salpeter equation, which can become numerically intractable when a large number of \( k \)-points and multiple bands are involved. We introduce a non-iterative approach to vertex correction based on rank factorization of the impurity vertices, which significantly alleviates the computational burden. We demonstrate that this method can be implemented along with effective Hamiltonians extracted from electronic structure calculations on perfect crystals, thereby enabling the quantitative analysis of quantum effects in electron conduction for real materials.

© 2020 Elsevier B.V. All rights reserved.

1. Introduction

The Kubo formula establishes a fundamental link between near-equilibrium processes and equilibrium correlation functions, enabling a microscopic description of various transport phenomena [1]. When applied to electronic transport in the presence of external fields, the Kubo formula for current response leads to the quantum mechanical formulation of electrical conductivity [1,2], the single most important property of materials. Electron conduction itself displays an extraordinarily wide range of unusual phenomena, which sometimes reveals highly nontrivial physical principles. Of particular interest to quantum transport phenomena is the localization of electrons in the presence of disordered impurities in an otherwise ideal metallic system, where electrons tend to be more localized compared to classical particles. Localization of electrons is crucial to the discussions of various problems, including weak localization, magnetoresistance, the quantum Hall effect, and topological transport properties. Such electronic localization is in essence a consequence of quantum interference, which in the Kubo formula is embodied in a set of vertex corrections, including those from the maximally-crossed diagrams [3–6].

The vertex correction corresponding to a partial sum of the maximally-crossed diagrams is obtained by solving the Bethe–Salpeter equation, which is a 4-point Dyson equation. Successful theories have been developed by considering low-energy effective Hamiltonians and only the leading order contribution to vertex correction to account for, for example, localization or absence of it in novel Dirac and Weyl semimetals [7–9]. However, it is challenging to numerically solve the full vertex correction for a generic band structure of a realistic material, where it becomes unlikely to achieve simplification (based on symmetry or otherwise). This challenge arises plainly from the fact that when a large number of \( k \)-points and multiple bands are involved, the vertex is a non-sparse matrix of enormous size, which practically defies obvious methods for solving the matrix equation. This is true despite the fact that high-quality electronic structure of real materials can routinely be obtained based on state-of-the-art electronic structure methods. In particular, Wannier functions [10] have become a standard technique for down-folding the band structure and Brillouin zone interpolation, allowing us to conveniently compute a wide range of physical properties [11–15].

In this paper, we present a method to compute the vertex corrections arising in the calculation of electric conductivity for a crystalline system with a non-iterative approach, which can greatly alleviate the computational burden and boost efficiency. The remainder of this paper is organized as follows. We begin with a brief recapitulation of the Kubo’s theory for electrical conductivity, leading to the vertex corrections, whereupon the essential formulae and notations are introduced. An analysis of the
vertex in the presence of disordered, uncorrelated impurities in a fairly generic form will be presented. It is revealed that the impurity vertex can have fewer degrees of freedom than its apparent dimension, owing to the short-rangedness of impurity potential. This crucial fact leads to an algorithm based on rank factorization, which reduces the dimension of the Bethe–Salpeter equation. The computational load of the rank factorization is further reduced by employing a projective singular value decomposition method. The stability and efficiency of our method for the vertex correction are examined and compared with direct methods, such as matrix inversion and the biconjugate gradient method. We show that the computational complexity of our method is formally \( \sim O(n^3) \) and indeed scales well in our implementation, whereas those of matrix inversion and the biconjugate gradient method are \( \sim O(n^4) \). Finally, we present two concrete examples by applying this method to 2-dimensional monolayer lead (Pb) with a hexagonal lattice and 3-dimensional face-centered cubic Pb to obtain the vertex corrections to conductivity. These illustrative examples show that our method can be efficiently applied to obtain the aforementioned vertex corrections to conductivity, thereby quantitatively analyzing weak localization for real materials.

2. Vertex corrections in conductivity

In this section, the problem that we aim to solve is introduced along with essential notations. The electronic Hamiltonian and the nature of the impurity potential will be clarified. Starting with the Kubo formula for electric conductivity, vertex corrections are introduced, which lead to the Bethe–Salpeter equation in the diagrammatic expansion of the disorder average. In this section, the vertex corrections in electric conductivity are introduced diagrammatically to highlight the basic structure of the theory, and their algebraic forms are postponed to the next section. A nonexhaustive discussion of available methods for computing electric conductivity and vertex corrections is also presented.

We work primarily in the Bloch representation, and operators are expanded in the eigenstates of the self-consistent mean-field single-particle electronic Hamiltonian, \( H \), for instance, from various implementations of the density functional theory (DFT). Thus, the Hamiltonian is only effectively non-interacting because part of the interactions and correlations are already captured by the mean-field theory. Within the static-lattice approximation, the Bloch theorem applies, and we shall assume knowledge of the solution to the Schrödinger equation

\[
H \psi_{nk} = E_{nk} \psi_{nk},
\]

where \( \psi_{nk} \) is the Bloch eigenstate of the nth band at quasimomentum \( \mathbf{k} \). From the Bloch functions, Wannier functions can be obtained to afford a lattice representation in which the impurity potential can be conveniently written. Associating the Bloch function with the Fermion operator \( c_{nk} \), the Fermion operator for destroying an electron in orbital \( a \) at lattice point \( \mathbf{R} \) is

\[
c_{a\mathbf{R}} = \frac{1}{\sqrt{n_k}} \sum_{nk} e^{-i \mathbf{k} \cdot \mathbf{R}} U_{a\mathbf{k}}(\mathbf{k}) c_{nk},
\]

where \( U(\mathbf{k}) \) is a unitary matrix obtained from various approaches to the construction of Wannier-type functions [10]. Here, \( n_k \) represents the number of \( \mathbf{k} \)-points, which is equivalent to the number of lattice sites under the Born–von Kármán boundary condition.

We shall confine ourselves to the problem of electric conductivity in the presence of randomly distributed elastic scatterers. The general form of one-body impurity potential written for a Wannier-type basis on a lattice is

\[
H' = \sum_{\mathbf{R}} \sum_{\mathbf{R}'} \sum_{\mathbf{R}''} V_{ab}(\mathbf{R} - \mathbf{R}'', \mathbf{R}' - \mathbf{R}) c_{a\mathbf{R}'} c_{b\mathbf{R}''} + H.c.,
\]

where \( \mathbf{R}(\mathbf{R}') \) is the lattice vector, and \( a, b \) label the Wannier-type orbitals. \( \mathbf{R} \) is a lattice site where an impurity resides. This form of impurity potential allows hopping between orbitals on the same and different lattice sites. We will focus on short-ranged impurities, which have been widely employed to model scattering in electronic transport, especially for vertex corrections to conductivity [8,9,16,17]. It is assumed that the impurity potential has a cutoff interaction range \( R_c \) such that \( V_{ab}(\mathbf{R} - \mathbf{R}'', \mathbf{R}' - \mathbf{R}) = 0 \), if \( |\mathbf{R} - \mathbf{R}'| > R_c \) or \( |\mathbf{R}' - \mathbf{R}| > R_c \). Then, in the Bloch representation, the general impurity potential has the following matrix form:

\[
H'_{kk'} = S_{k-k'} U_k^\dagger V_{kk'} U_{k'} = S_{k-k'} W_{kk'},
\]

where \( S_k = \eta_i \sum_i \epsilon_{ik} R_i \), and \( V_{kk'} = \sum_R V_{ab}(R, R') \).

We now introduce the Kubo formula for electric conductivity in the presence of disordered impurities. The linear response approach is naturally couched in the machinery of Green’s functions owing to its perturbative nature. An advantage of this method is that for impurity scattering, its result contains quantum corrections to conductivity that are responsible for the weak localization effect, which is not included in the usual Boltzmann conductivity. The longitudinal electrical conductivity in the long-wavelength limit derived from the general Kubo formula for a non-interacting electronic system in the presence of an ensemble of disordered scatterers is [1,7,18–20]

\[
\sigma^{\text{as}}(\omega) = \frac{e^2}{\Omega} \int \frac{d\epsilon}{2\pi} \frac{f(\epsilon) - f(\epsilon + \hbar \omega)}{\omega} \times tr\left(\nu^a \Lambda^b(\epsilon) \nu^b \Lambda^a(\epsilon + \hbar \omega)\right) d\epsilon,
\]

where \( \omega \) is frequency, \( v \) velocity operator, \( f(\epsilon) \) the Fermi–Dirac distribution, and \( \Omega \) the volume of the system. \( \Gamma^{R/A}(\omega) = (\epsilon - H - \Sigma^{R/A}(\omega))^{-1} \) is the retarded/advanced single-particle Green’s function evaluated in the presence of a given disorder configuration, which results in corresponding self energy \( \Sigma^{R/A}(\omega) \). \( \langle \ldots \rangle_A \) indicates disorder averaging. Upon averaging over an ensemble of scatterer configurations, the current response is expanded in terms of scattering vertices into an infinite series of distinct current bubbles. The bare (zeroth order) current bubble yields the Drude conductivity [19,20], whereas other current bubbles contain vertex corrections arising from impurity-induced scattering between electron and hole excitations [21]. In the weak-scattering regime, the leading corrections to bare current bubble arise from simultaneous scattering of propagating quasiparticles by the same impurity site. Neglecting multiple scattering, two classes of diagrams are usually calculated: ladder diagram and maximally-crossed diagram, as shown in Fig. 1(a) and (b), respectively. The ladder diagram shown in Fig. 1(a) is composed of a pair of bare current vertices connected by a 4-point diffusion vertex, \( \Gamma_d \), with countercurrent electron and hole propagators. The diffusion vertex displays diffusive poles with respect to momentum exchange from successive elastic scattering by impurities. The partial sum of the ladder diagrams amounts to the usual vertex correction for anisotropic scattering, which replaces apparent scattering time with transport relaxation time and leads to the equivalent of the Boltzmann conductivity. With the implementation of semiclassical Boltzmann transport formalism, it is already possible to evaluate this part of the bulk conductivity for real materials via maximally-localized Wannier orbitals obtained from DFT calculations without further vertex corrections [14] and the ballistic conductance of a finite structure within the Landauer’s formulation [22].

The maximally-crossed diagrams [3–6] shown in Fig. 1(b) are important to the discussion of quantum correction to conductivity, arising from interference owing to the wave nature of electrons. The maximally-crossed diagrams can be depicted with
Fig. 1. Current bubbles with vertex correction from (a) the Ladder diagram and (b) the maximally-crossed diagram. The wiggly lines are bare velocity vertices. \(\alpha, \alpha, \beta, \beta\) are indices that combine band and momentum. The arrowed solid lines are disorder averaged retarded or advanced single-particle propagators, as labeled by \(R\) and \(A\), respectively. An impurity vertex is represented by a dashed line with a cross in the center. The vertices shown as shade blocks are the (a) diffuson and (b) cooperon. The corresponding Bethe–Salpeter equations are depicted diagrammatically in (c) and (d).

an exact diagrammatic twist in a time-reversal invariant system by reversing one of the single-particle Green’s function line. This leads to a pair of current vertices connected to the 4-point cooperon vertex, \(\Gamma_c\), a propagator in the particle–particle channel. \(\Gamma_c\) is also a ladder diagram but with concurrent single-particle lines; therefore, it displays diffusive poles with respect to the total momentum of scattering. The cooperon diagram is especially relevant to the phenomena of weak localization and weak antilocalization in the presence of disorder when time-reversal symmetry is preserved, which can be suppressed upon the removal of time-reversal symmetry by magnetic impurity or external magnetic field. Weak localization and weak antilocalization have a dramatic impact on the transport properties at low temperatures, particularly in low-dimensional quantum systems. Quantum vertex correction has been widely used to explain the experimentally observed magnetooconductivity induced by the weak localization or weak antilocalization effect in metallic systems, such as metal films [23], graphene [24], topological insulator thin films [25], and Weyl semimetals [26,27]. Clearly, the electronic structure of these materials may not always be simple enough to justify approximations that make analytical results viable, in which case an efficient numerical approach to the quantum correction to longitudinal conductivity is desirable.

Given access to accurate electronic structures for real materials from various methods, it is highly desirable to be able to evaluate electrical conductivity, including vertex corrections, for real materials, which requires the evaluation of the diffusion vertex \(\Gamma_d\) and cooperon vertex \(\Gamma_c\). Although useful (and sometimes analytical) results can be obtained for model Hamiltonians with high symmetry in conjunction with simple scattering potential [8,16], the task is computationally challenging for generic materials with more complex scattering potentials, for which multiple bands and a large \(k\)-point set are typically involved in the calculations and limited simplifications, if any, can be made from symmetry.

With this outline of the problem, a non-iterative approach to the Bethe–Salpeter equation for vertex correction to electric conductivity is proposed and shown to alleviate the burden of computation. In the subsequent section, an analysis reveals that for scattering potential of a finite range, the rank of scattering vertex in the Bethe–Salpeter equation can be significantly lower than its dimension. This fact will then be systematically exploited to simplify the solution of the Bethe–Salpeter equation.

3. Analysis and algorithm

As introduced in the previous section, the diffusion and cooperon vertices can both be expressed as infinite series. The series can be folded into integral equations, referred to as the Bethe–Salpeter equation, as diagrammatically shown in Fig. 1(c) and (d). Expanded in a complete set of Bloch functions \(\psi_{nk}\) as in Eq. (1) with the abbreviations \(\alpha, \beta, \alpha', \beta', \alpha_i, \beta_i \rightarrow nk\), the Bethe–Salpeter equation can be written as a matrix equation [21]:

\[
\Gamma_{\alpha'\beta'\rho}(\varepsilon, \omega) = \Gamma_{\alpha'\beta'\rho}^0 + \sum_{i,j} \Gamma_{\alpha'\beta'\rho}^0 G^{\alpha_i\beta_i}_\rho(\varepsilon \pm \hbar \omega) \times G^{\alpha_j\beta_j}_\rho(\varepsilon \pm \hbar \omega),
\]

where the bare impurity vertex \(\Gamma^0\) is a matrix composed of elements corresponding to simultaneous scattering off of a single impurity by electron and hole:

\[
\Gamma_{\alpha'\beta'\rho}^0(\varepsilon, \omega) = \langle H_{\alpha'\rho} H_{\beta'\rho} \rangle_d.
\]

where \(H_d\) is the static impurity potential given in Eq. (3). The diffusion and cooperon’s Bethe–Salpeter equations differ in quasi-momentum conservation, dictated by the momentum flows of the single-particle propagators in the conductivity bubbles, as

\[
k_a \mp k_b = k_a' \mp k_b' = q.
\]

Note that the disorder averaged single-particle propagators, \(G^{\alpha_i\beta_i}_\rho(\varepsilon)\), are diagonal in \(k\), i.e., \(k_a = k_b\), but are in general non-diagonal in band indices due to self-energy insertion. Therefore, the Bethe–Salpeter equation, in consideration of the quasi-momentum conservation, is decoupled into separate equations for each \(q\) as

\[
\Gamma_{mnk,n'\alpha'\beta'}^{\alpha_\rho \beta_\rho} = \Gamma_{mnk,n'\alpha'\beta'}^{\alpha_\rho \beta_\rho} + \sum_{m_1n_1k_1,m_2n_2k_2} \Gamma_{mnk,m_1n_1k_1}^{0} \times K_{m_1n_1\alpha_1,m_2n_2\alpha_2} \Gamma_{m_2n_2\beta_1,k_1}^{r}(q),
\]

where \(K_{m_1n_1\alpha_1,m_2n_2\alpha_2} = \delta_{\alpha_1\alpha_2}\delta_{\beta_1\beta_2} G^{\alpha_\rho}_\rho(m_1n_1k_1) G^{\beta_\rho}_\rho(m_2n_2k_2)\). Here, \(\alpha, \beta, \) etc. have been replaced by \(mn, nk\), etc. The energy dependencies of the vertices and propagators are omitted for brevity as they can be unambiguously inferred.

Solving the Bethe–Salpeter equation with a direct method, either by matrix inversion or by an iterative procedure such as the biconjugate gradient method [28], poses a significant computational challenge, especially for 3-dimensional systems with complex Fermi surfaces. Usually a large number of \(k\)-points are necessary to ensure adequate sampling of the Fermi surface, which means the dimension \(n\) of the complex-valued non-sparse matrices \(\Gamma^0\) and \(\Gamma\) is large, where \(n = n_{\text{f}}^2n_k\) and \(n_{\text{f}}\) and \(n_k\) are respectively the numbers of bands and \(k\)-points. This can hinder the use of standard libraries for numerical linear algebra to perform the matrix inversion because of excessive memory load. In an iterative procedure, such as the biconjugate gradient method, the equation is solved in a column-wise fashion to avoid cramping the memory. However, both of these direct methods have a nominal computational complexity of \(O(n^3)\), i.e., the asymptotic complexity of \(\text{(naive) matrix multiplication. Even when adopting the Strassen algorithm with a reduced numerical stability} [29],\) the complexity exponent is \(\log_27 \approx 2.807\). Although coherent-potential approximation has been used to simplify vertex corrections for both equilibrium [30] and non-equilibrium [31] theories, the maximally-crossed diagrams are naturally absent in this approximation [20]. Therefore, it is highly desirable to find an
approach to Eq. (8) that is both computationally efficient and memory-saving.

Our approach is based on the fact that the rank of \( \Gamma^0 \) for the impurity potential described above is no greater than \( n_i^2 \times \min(n_k, n_i^2) \), where \( n_i \) is the number of lattice sites within the range \( R_i \). The claim is proved as follows. The disorder averaging of the impurity vertex only reduces the geometric structure factor of the impurities to a constant factor, \( \langle S(k - k') | S(k' - k) \rangle \rightarrow n_i/n_i^2 \), where \( n_i \) is the number of impurities, for electron–hole scattering by the same impurity site. For a given transfer or total momentum \( \mathbf{q} \), the bare vertex \( \Gamma^0 \) is composed of \( n_i \times n_i \) blocks, which follows from quasimomentum conservation. Each individual block, \( \Gamma_{kk}^0 \in \mathbb{C}^{n_i \times n_i} \), corresponds to a pair of momenta on the \( k \)-mesh used in the computation. Then, for a given \( \mathbf{q} \), we have a single impurity vertex given by

\[
\Gamma_{kk}^0(\mathbf{q}) = \frac{n_i}{n_i^2} W_{kk'} \otimes W_{q+k,q+k'},
\]

where \( \otimes \) stands for tensor multiplication with respect to band indices. The full \( \Gamma^0 \) matrix is a tensor product with respect to two vectors composed of matrix blocks.

Now we construct a tensor product \( \Lambda \) to whom \( \Gamma^0 \) is a submatrix. Explicitly, we define a matrix \( \Phi \in \mathbb{C}^{n_i \times n_i} \), \( \Phi_{mk,nk} = e^{i\mathbf{k} \cdot \mathbf{r}_m} \delta_{nn_k} / \sqrt{n_i} \). The rank of \( \Phi \) is no greater than \( \min(n_i,n_i,n_i^2) \). Also note that the rank of \( V \) is no greater than \( n_i \), for the rank of a matrix cannot exceed the smaller of its dimensions. Then, the rank of \( \Phi \otimes V \otimes \Phi \) is no greater than \( \min(n_i,n_i,n_i^2) \) from the rank inequality of the matrix. Now we define another square matrix \( \Lambda \) with dimensions \( n_i^2 \times n_i^2 \):

\[
\Lambda = \frac{n_i}{n_i^2} (\Phi \otimes V \otimes \Phi) \Lambda(\Phi \otimes V \otimes \Phi),
\]

meaning that \( A_{ijkl} = \frac{n_i}{n_i^2} (\Phi \otimes V \otimes \Phi)_{ij} \delta_{ij} \). From the rank inequality of tensor product, we see that rank(\( \Phi \)) \leq \min(n_i^2,n_i^2,n_i^2). It only remains to recognize that \( \Gamma^0 \) is a momentum-conserving submatrix of \( \Lambda \) of order \( n_i^2 \), since

\[
\Gamma_{mn,m'n'}^0(q) = \lambda_{mn,m'n'}(q) \equiv \delta_{m'm} \delta_{nn'}(q). \]

It is then implied that the rank of \( \Gamma^0 \) is no greater than rank(\( \Lambda \)), for a submatrix cannot have a higher rank. Therefore, let \( r = \text{rank}(\Gamma^0) \), and we have

\[
r \leq \min(n_i^2,n_i^2). \tag{11}
\]

This is the first key result of this paper, showing that given a general form of the scattering potential on the lattice model of Eq. (3), the independent degrees of freedom of the impurity vertex is less than \( \min(n_i^2,n_i^2) \). When \( n_i^2 < n_i \), the rank of \( \Gamma^0 \) is no greater than \( n_i^2 n_i^2 \), regardless of the number of \( k \)-points included in the calculation. In particular, \( r \leq n_i^2 \) if the impurity scattering is completely localized, i.e., \( n_i = 1 \).

The fact that the non-sparse \( \Gamma^0 \) matrix has lower rank than its apparent dimensions if \( n_i^2 < n_i \) can be exploited to facilitate the solution of the Bethe–Salpeter equation. The bare vertex \( \Gamma^0 \), which is a square matrix with \( r \) non-zero eigenvalues \( |\gamma_i| = 1, \ldots, r \), is bound to have the compact singular value decomposition

\[
\Gamma^0 = PD^0Q^\dagger, \tag{12}
\]

where \( D^0 \in \mathbb{C}^{r \times r} \) with \( D^0_{ii} = |\gamma_i| \delta_{ii} \), and \( P, Q \in \mathbb{C}^{n_i \times n_i} \) have full column rank. Here, the range of \( P \) \( (\text{ran}(P)) \) and \( Q \) \( (\text{ran}(Q)) \) are each \( r \)-dimensional subspaces of the \( n \)-dimensional linear space. The subspace completeness is ensured by the equality of the number of orthogonal vectors and the rank \( r \). \( P \) and \( Q \) are both column orthogonal:

\[
P^\dagger P = I_{r \times r} = Q^\dagger Q. \tag{13}
\]

It may be noted that \( P \) and \( Q \) are neither row orthogonal nor necessarily unique, because \( |\gamma_i| \) can have degeneracy. If there exists an \( r \times r \) matrix \( D \) (generally not diagonal) that satisfies the equation

\[
\Gamma^0 = D^0 + D^0 K D,
\]

where \( K = Q^\dagger P \). Then \( \Gamma^0 = PDQ^\dagger \) is the solution of the Bethe–Salpeter equation because the solution to Eq. (8), if it exists, is unique. It is also implied rank(\( \Gamma^0 \)) = \( r \), which is a fact not at all obvious at the outset. Solving Eq. (14) by matrix inversion or iteration is much easier since \( r < n \), provided the knowledge of \( P \), \( Q \), and \( D^0 \).

The direct singular value decomposition employing standard libraries [32] requires an iterative procedure, and it is estimated to have time complexity \( O(n^3) \) for a square matrix of order \( n \). Moreover, each iteration revisits a square matrix of order \( n \) from the previous step, adding to the memory burden, which can already be taxing, if not infeasible, for a large system. The fact that \( \Gamma^0 \) has much lower rank than \( n \) can again be exploited, allowing us to determine \( P \) and \( Q \) in a three-step process. A matrix \( \Lambda \in \mathbb{C}^{n_i^2 \times n_i^2} \) is generated randomly. Post-multiplication of \( \Gamma^0 \) by \( \Lambda \) yields a projected matrix composed of \( r \) columns vectors in \( \text{ran}(P) \); and these vectors are subsequently orthogonalized by a matrix \( R_1 \) in a Gram–Schmidt process. Similarly, post-multiplication of \( \Gamma^0 \) by \( \Lambda \) yields a projected matrix composed of \( r \) columns vectors in \( \text{ran}(Q) \), which is subsequently orthogonalized by \( R_2 \). Both \( R_1 \) and \( R_2 \) are in \( \mathbb{C}^{r \times r} \) and upper triangular. Therefore, we have

\[
P_1 = PX; \quad Q_1 = QY.
\]

where \( X = D^0Q^\dagger AR_1 \) and \( Y = D^0P^1AR_2 \), which are yet to be determined as \( P \) and \( Q \) are unknown.

It is clear that \( X \) and \( Y \) are unitary by the column orthogonality of \( P, Q, P_1, Q_1 \). It follows from the unitarity of \( X, Y \) and column orthogonality of \( P, Q \) that

\[
P_1^\dagger \Gamma^0Q_1 = X^\dagger D^0P_1Y.
\]

Evidently, this procedure reduces the original \( n \)-dimensional singular value decomposition to a task in \( r \)-dimensions. Subsequently, \( P \) and \( Q \) can be obtained via Eq. (15). In fact, if \( \Gamma^0 \) and \( V \) are Hermitian, then \( P = Q \) and we have the eigenvalue decomposition \( \Gamma^0 = PDP^\dagger \). The random matrix is projected only once by \( \Gamma^0 \) to get \( P(Q) \), which requires even less computational cost. The method outlined above pivots on the projection from a complete linear space onto its subspaces, and we shall refer to it as the projective method for compact singular value decomposition. As will be illustrated shortly, this projective method is practically stable and computationally efficient for our problem compared to the standard linear algebraic methods.

4. Stability and performance

In this section, we present basic sanity checks on the implementation of the above method for solving Eq. (8) or Eq. (6). We will first show that the projective compact singular value decomposition has reasonable fidelity and stability in dealing with ill-conditioned problems. We will then show, with the example of a 2-dimensional electron gas, that our approach to the Bethe–Salpeter equation compares favorably with two of the direct methods in terms of how the computation time and memory scale with the number of \( k \)-points, \( n_k \), within large \( n_k \) limit.

We now examine the stability of the projective algorithm for singular value decomposition described in the previous section. The stability is measured by the fidelity of the projective method in recovering the singular values for these matrices, which are ill-conditioned in degrees depending on the base \( b \). Such tests are essential, especially in view of the relative instability of the
Gram–Schmidt orthogonalization employed in this approach. In these numerical experiments, we randomly generate test matrices, which are then subject to singular value decomposition by the projective method as well as by standard singular value decomposition routines. The test matrices are all in \( \mathbb{C}^{3000 \times 3000} \) with rank \( r = 25 \), corresponding to 5-band systems created by the following procedure:

1. \( D^0 \) is a diagonal matrix whose diagonal elements are \( \sigma_0^m = b^m, m = 1, 2, \ldots, r \), and the values of the base \( b = \tau, 2, 3, 4 \), where \( \tau \) is the golden ratio;
2. For each \( D^0 \) from step (1), a pair of orthogonal matrices, \( P, Q \in \mathbb{C}^{n \times r} \) are generated. To generate \( P \) or \( Q \), a matrix in \( \mathbb{C}^{n \times r} \) is generated at random with elements selected from a uniform distribution inside a square on the complex plane with corners at \( \pm(1 + i) \) and then orthogonalized. The test matrix is then \( PDP^\dagger \).

The relative errors of the computed singular values are shown in Fig. 2. It can be seen that for \( b = \tau, 2 \), the largest relative errors are less than \( 10^{-10} \) for exact singular values ranging over 7 orders of magnitude. Even in the case of \( b = 3 \), where the singular values range over 12 orders of magnitude, the singular values retrieved have acceptable fidelity with relative errors less than \( 10^{-3} \). Only when the range of singular values is over 17 orders of magnitude for \( b = 4 \), does the method become marginally unstable. These results indicate that this projective method should work as an expedient stratagem for the present application, although it may not offer sufficient stability to be universally applicable.

The computational efficiency of our approach to the Bethe–Salpeter equation is analyzed here in comparison with other standard methods for linear equations, such as direct matrix inversion and the biconjugate gradient method. Once \( K \) and \( F^0 \) are prepared, the Bethe–Salpeter equation in Eq. (8) can be solved using the projective singular value decomposition method, in which matrix multiplications involving \( F^0 \) (\( F^0 A \), \( P^\dagger F^0 P \)) are the leading order operations with a time complexity of \( O(n^2) \). The matrix rank \( r \) is dependent on the interaction range of the impurity potential, and for localized impurity potentials, sufficient accuracy is achieved.

A simple example is now employed to gauge the efficiency of our method in comparison with the matrix inversion and an in-house routine for the biconjugate gradient method. For this example, we use a one band system on a two-dimensional lattice with \( H_k = \cos k_x + \cos k_y \). The impurity is introduced via Eq. (3). It is assumed here that the impurities are completely local, only permitting scattering within each lattice point with an amplitude \( 0 < V < 1 \). The impurity concentration is 0.02, referring to the amount of impurities per lattice. The Bethe–Salpeter equation for the maximally-crossed diagrams is solved for a total quasimomentum \( q = (0.008, 0) \) at zero frequency. A uniform \( k \)-point grid is adopted to sample the Brillouin zone. The disorder averaged single-particle propagator is \( G_k^0 = (\mu - H_k - \Sigma^k)^{-1} \) with chemical \( \mu = 1 \) and self energy \( \Sigma^k = 0.001i \). When using the matrix inversion method, the matrix \( F^0 \) is computed in full and stored in the random-access memory. In our projective singular value decomposition approach and the biconjugate gradient method, columns or blocks of \( F^0 \) split by quasimomentum are computed on-the-flight to reduce the memory requirement. In our tests based on this single-band problem, a range of \( n_s \), which determines the dimension of \( F^0 \) was used. \( n_s \) values of 0, 2, and 4 were used, which lead to \( F^0 \) with different ranks.

Panels (a)–(c) in Fig. 3 show the computation times as functions of \( n_s \) on logarithmic scales for \( R = 0, 2, 4 \), respectively. By counting the theoretical floating-point operations, we find that the computation time for our method is \( O(n_s^1) \), and \( O(n_s^2) \) for both matrix inversion and the biconjugate-gradient method. The computational complexity is also confirmed empirically with the linear least squares of \( \log_{10} t \) versus \( \log_{10} n_s \). While all three methods yielding identical results in all cases, the biconjugate gradient method is the most time consuming. The estimated slopes of \( \log_{10} t \)-vs-\( \log_{10} n_s \) for \( R = 0, 2, 4 \) and using our method.

---

Our code was compiled with the Intel(R) C CompilerXE 19.0 Update 1 for Linux*, and linear algebra routines from Intel Math Kernel Library were used unless otherwise noted. All the measurements were done on 2xIntel Xeon Gold 6130 with 2.10 GHz CPUs, 16 core processors equipped with 256 GB physical RAM.
are about 2.0, whereas it is 2.9 for direct matrix inversion and 3.0 for the biconjugate gradient method. Thus, the empirical computational complexities agree well with theoretical expectations. Although direct matrix inversion method is advantageous for small \( n_h \) and large \( R_c \), it ceases to be feasible for a moderate number of \( k \)-points because of the overwhelming memory requirement for storing the matrices.

Fig. 3(d)–(f) show the peak memory load of the three methods. The amount of working memory required by matrix inversion increases rapidly with \( n_h \), and it ceases to be feasible very quickly for the compute node. Both biconjugate gradient and singular value decomposition methods require much less memory as \( r \ll n_h \). The memory load of our method shows moderate increase with \( R_c \) (or \( r \)). The above results regarding the time and memory costs in Fig. 3 show that the projective singular value decomposition method is both efficient and memory-saving for solving the Bethe–Salpeter equation with a short-ranged impurity potential. The accuracy of our method is also checked by comparing with the exact matrix inversion method. The relative error is less than \( 10^{-10} \), showing sufficient precision for numerical calculations and thus conforming the validity.

Fig. 4. Band structures of (a) monolayer Pb with a hexagonal lattice and (b) bulk Pb. Energy dispersion from the tight-binding model (orange dots) is compared with the density functional theory electronic structure (black lines).

5. Applications

To demonstrate its application to real materials, we use our method to evaluate the conductivity correction from maximally-crossed diagrams for 2-dimensional monolayer Pb with a hexagonal lattice and 3-dimensional fcc metal Pb. Here, we focus on the dc conductivity (\( \omega = 0 \)) at zero temperature. Their tight-binding Hamiltonians are constructed with three \( p \) orbitals of Pb using Wannier90 [11], based on the electronic structure calculations from density functional theory [33]. As shown in Fig. 4, the tight-binding models for monolayer Pb with hexagonal lattice and bulk fcc Pb can perfectly capture the band dispersion around the Fermi level. A large number of \( k \)-points are necessary to sufficiently sample the complicated Fermi surfaces of monolayer Pb with hexagonal lattice and fcc Pb, which makes the direct inversion method infeasible for the Bethe–Salpeter equation due to the huge memory requirement, especially in 3-dimensional systems. There are no widely accepted methods for constructing impurity potentials (including its cutoff radius) for real materials from ab initio calculations, and this is a problem deserving further investigation. For the present work, we will use simple model potentials, for the purpose of demonstrating our method.

We start with the tight-binding Hamiltonian of Pb and the impurity scattering potential described by Eq. (3). The disorder averaged Green’s functions \( G^{R/A}_k \) are obtained with an iterative method using the self-consistent Born approximation (SCBA) [17,34] so that the Ward identity [21] is satisfied in the Bethe–Salpeter equation. In SCBA, the disorder averaged one-particle Green’s function is given by

\[
G^{R/A}_k = (\epsilon - H_k - \Sigma^{R/A}_k)^{-1},
\]

where the self energy owing the presence of disorder is

\[
\Sigma^{R/A}_k = \langle H'_{kk} \rangle + \sum_{k'} (H'_{kk} G^{R/A}_{k'k} H'_{k'k}).
\]

In SCBA, the chemical potential is self-consistently determined, in which the Brillouin sum to obtain the total electron number is facilitated by upsampling via Fourier interpolation of the poles of Green’s functions.

With \( G^{R/A}_k \) obtained from SCBA and the impurity potential, our projective singular value decomposition method can be used to solve the Bethe–Salpeter equation of maximally-crossed diagrams. The sum over these diagrams gives correction \( \delta \sigma_{xx} \) as

\[
\delta \sigma_{xx} = \sum_{\mathbf{q}} \delta \sigma_{xx}(\mathbf{q}),
\]

where

\[
\delta \sigma_{xx}(\mathbf{q}) = \frac{e^2}{2\pi \hbar} \sum_{k} v^2_{k} |q - k|^2 |C_{k+q}^R - C_{k-q}^R - C_{k,q}^A| \Delta_{q,k},
\]

The orbital indices are omitted to highlight the momentum dependencies. The summation over \( \mathbf{q} \) is performed for \( 1/l_0 < q < 1/l \). Here, \( l \) is the mean-free path determined by the relaxation time \( \tau \) and diffusion coefficient \( D \), which can be estimated from the calculations of Drude conductivity \( l = \sqrt{D\tau} \). The coherence length characterizing the inelastic scattering processes, \( l_0 \), is introduced as a parameter.

For the 2-dimensional monolayer Pb with hexagonal lattice, we first inspect the case of a fully-localized impurity potential described by Eq. (3) with a concentration of 0.0048, which induces only on-site hopping between \( p \) orbitals with arbitrary numerical values shown in Table 1. The \( q \)-resolved correction \( \delta \sigma_{xx}(\mathbf{q}) \) is calculated to verify convergence with respect to \( n_h \). Fig. 5(a) shows the relative errors of \( \delta \sigma_{xx}(\mathbf{q}) \) versus \( n_h \). The results of \( \delta \sigma_{xx}(\mathbf{q}) \) for the largest \( n_h(10^3) \) are used as the estimate of correct values.
to calculate relative errors. It is seen that a large number of \(k\)-points up to \(10^{4.6}\) are needed to achieve an accuracy within 1%. Fig. 5(b) displays the calculated vertex correction to conductivity as a function of \(l_o\). It is clear that \(\delta\sigma_{xx}\) is negative, indicating weak localization; therefore, it can be fitted well by the anisotropic form of 2-dimensional weak localization [35]:

\[
\delta\sigma_{xx} = -\frac{2e^2}{h\pi} \alpha \log \frac{l_o}{l}, \tag{20}
\]

where \(h\) is the Planck constant, the factor of 2 results from the spin degeneracy, and the coefficient \(\alpha = \sqrt{\frac{6\pi}{\gamma_y}}\) contains the effects of anisotropy with \(\sigma_{xx/yy}\) as the Drude conductivity. This implies that \(\delta\sigma_{xx}(\mathbf{q}) = -e^2/(4\pi^2 \hbar q)^2\) if the momenta is rescaled as \(\mathbf{q} = (q_x, q_y/\alpha)\). Fitting the calculated \(\delta\sigma_{xx}\) versus \(l_o\) shown in Fig. 5(b) using Eq. (20) gives \(\alpha = 1.02\) and \(l = 17.1 a_0\), which are close to \(\sqrt{\frac{2\pi}{\gamma_y}} = 1.01\) and the given mean-free path \(l = 20.66 a_0\). The small mismatch [36] between the fitted and theoretically derived values of \(\alpha\) and \(l\) originates from the deviation of the calculated \(\delta\sigma_{xx}(\mathbf{q})\) with the result based on second-order perturbation theory [35], \(-e^2/(4\pi^2 \hbar q)^2\).

We now turn our attention to two more types of impurity: one with a finite interaction range and another with higher symmetry. We find that in these cases, the theory summarized in Eq. (20) fails. In the former case, the hopping range is extended to the next-nearest neighbors from the localized potential \(V_{ab}(0, 0)\) in Table 1 by \(V_{ab}(\mathbf{R}, \mathbf{R}') = V_{ab}(0, 0) \frac{\sin R \sin R'}{R^2 - R'^2 / \gamma_{xy}}\). Similarly, the results of \(\delta\sigma_{xx}\) shown in Fig. 5(b) can also be described by Eq. (20) with \(\alpha = 0.65\) and \(l = 18.01 a_0\). However, \(\alpha\) differs significantly from the theoretical value \(\frac{2\pi}{\gamma_y} = 1.19\). This discrepancy indicates that the effectiveness of the second-order perturbation is seriously compromised when the impurity potential is not fully localized and has a variation in the \(k\)-space. In the final case, the impurity induced hopping is again fully local, but it is assumed that the impurity potential respects the mirror symmetry of the plane of the hcp lattice. In this case, the hopping matrix elements are the same as in Table 1, except that the hopping between \(p_z\) and \((p_x, p_y)\) orbitals is forbidden. The fitted \(l = 16.47 a_0\) is close to the given value \(18.23 a_0\), while the fitted \(\alpha = 2.00\) is twice the \(\frac{2\pi}{\gamma_y} = 0.99\) shown in Fig. 5. This deviation from Eq. (20) corresponds to adding two decoupled parts derived from the \(p_z\) and \((p_x, p_y)\) orbitals, leading to the simple doubling of \(\alpha\). Therefore, the nature of impurity potential, and in particular its range, orbital, and symmetry, plays a crucial role in quantum correction to conductivity. They could render simpler theories like that of Eq. (20) ineffective. A careful numerical approach becomes necessary in the investigation of quantum correction in realistic materials with complicated impurity scattering.

The quantum correction to conductivity of fcc Pb is shown in Fig. 6 with \(R_i = 0\), \(V_{ab} = 5.0 a_0 eV\), and \(n_i/n_h = 0.02\). The computed weak localization correction to conductivity can be fitted well by the equation for a 3-dimensional metallic system [23]:

\[
\delta\sigma_{xx} = \frac{2e^2}{h^2 \alpha}(1/l_o - 1/l). \tag{21}
\]

Fitting the calculated \(\delta\sigma_{xx}\) versus \(l_o\) shown in Fig. 6 to formula Eq. (21) gives the fitted \(l = 7.86 a_0\). This is slightly larger than the given mean-free path \(11.01 a_0\), which is likely to reflect higher order effects not accounted for by the second-order perturbation theory behind Eq. (21) [35]. These two examples illustrate that the non-iterative approach to the Bethe–Salpeter equation can be successfully applied to the calculation of vertex correction for multi-band Hamiltonian derived from realistic materials.

### 6. Discussions and summary

We have presented a non-iterative method for solving the Bethe–Salpeter equation arising in vertex correction for electric conductivity. This method exploits the fact when the scattering potential is short-ranged, the scattering vertex is typically not full rank, allowing for reduction of matrix dimensions through an expedient projective singular value decomposition. We have shown that this projective singular value decomposition offers sufficient stability and robustness for this problem. The resultant algorithm for the Bethe–Salpeter equation has a theoretical computational complexity of \(O(n^3)\) when the rank of the bare scattering vertex is much smaller than the number of \(k\)-points required for sampling the Fermi surface. It was empirically established, by investigating the case of the single band model in 2 dimensions with short-ranged impurity potentials, that our method compares favorably to conventional linear algebraic approaches with a computational complexity of \(O(n^3)\), namely, matrix inversion and the biconjugate gradient method.

It was then demonstrated that the method can be successfully applied to compute the quantum correction for 2-dimensional and 3-dimensional systems, for which the tight-binding Hamiltonians are obtained from DFT calculations with ab initio accuracy. Our method allows the introduction of various types of impurity potentials. It was found that the calculated quantum correction to conductivity departs qualitatively from the expectation of second-order perturbation theory. This correction is signal modulated by the range, orbital, and symmetry of the impurity potential, indicating that the nature of impurity is essential and requires careful considerations in real materials.

Thus, our method provides an efficient route for evaluating the quantum effects in conductivity of materials based on realistic electronic structures obtained from accurate electronic structure methods. We expect that this newly proposed method could have a profound impact on the study of transport properties by the Kubo formula. Nonetheless, it should be emphasized that at this point not all the inputs of our method are obtained ab initio. First, the Hamiltonian describing how the electrons scatter with an impurity has to be given as a model at this stage. Methods should be developed to evaluate the impurity scattering potentials [37–39]. Second, the coherence length needed for the momentum cutoff is also entered as a given parameter [23].
Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant No. 11725415 and No. 11934001), the Ministry of Science and Technology of the People’s Republic of China (Grant No. 2018YFA0305601 and No. 2016YFA0301004), and by the Strategic Priority Research Program of Chinese Academy of Science (Grant No. XDB28000000).

References