PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

View Article Online
View Journal | View Issue

Cite this: *Phys. Chem. Chem. Phys.*, 2022, **24**, 5658

Simple renormalization schemes for multiple scattering series expansions

Aika Takatsu,[a] Sylvain Tricot,[b] Philippe Schieffer,[b] Kevin Dunseath,[b] Mariko Terao-Dunseath,[b] Keisuke Hatada[a] and Didier Sébilleau[b]

Received 3rd December 2021,
Accepted 4th February 2022

DOI: 10.1039/d1cp05530e

rsc.li/pccp

A number of renormalization schemes for improving the convergence of multiple scattering series expansions are investigated. Numerical tests on a small Cu(111) cluster demonstrate their effectiveness, for example increasing the rate of convergence by up to a factor 2 or by transforming a divergent series into a convergent one. These techniques can greatly facilitate multiple scattering calculations, especially for spectroscopies such as photoelectron diffraction, Auger electron diffraction, low energy electron diffraction *etc.*, where an electron propagates with a kinetic energy of hundreds of eV in a cluster of hundreds of atoms.

1 Introduction

Multiple Scattering Theory (MST) is one of the methods favored to model spectroscopies such as photoemission, X-ray absorption or electron diffraction, and more generally, most of the core-level spectroscopies. In particular, as far as condensed matter is concerned, electron diffraction based techniques such as photoelectron diffraction, Auger electron diffraction or low energy electron diffraction are quite simple experimental setups yet powerful enough to infer crystallographic and structural characteristics of thin layers of materials. The flexibility of MST also makes it particularly suited to the study of the electronic structure of materials as it can describe within the same formalism both bound and continuum states.[1,2] In the former case, MST is generally known as KKR theory. Far from being limited to surface or material science, MST also enjoys widespread use in several areas of physics (nuclear physics, condensed matter, acoustics, geophysics, *etc.*), since describing the interaction of a wave with a media composed of distinguishable objects/obstacles acting as scatterers is a more general problem.

Among the different formulations of MST, the so-called scattering path operator approach has been shown to be particularly convenient for modelling both spectroscopies and band structures. Within this approach, the scattering path operator $\bar{\tau}^{ji}$ can be identified as the operator describing all the pathways an electron can travel in order to move from atom i to atom j. As such, it contains all the information about the structural, electronic and magnetic properties of the material under consideration.

In practice, if we denote the potential describing the interaction of the probe particle with atom i as V_i and the propagator describing the behaviour of this probe within the material by G, we can define the scattering path operator as

$$\bar{\tau}^{ji} = \overline{V}_i\delta_{ij} + \overline{V}_j G \overline{V}_i. \tag{1}$$

Here, the overbar is simply a reminder that, because atom i and atom j do not generally coincide with the arbitrary origin we choose, the corresponding operator contains translation operators referring to this origin.[3]

The differential cross sections for many spectroscopies can then be expressed as functions of matrix elements of this scattering path operator.[4] Its evaluation is thus critical in the MST approach, however the definition (1) is not very useful for practical purposes, as it contains the total Green's function or propagator for the full system. Instead, we can express the scattering path operator in terms of the free propagator G_0 and transition operators $\overline{T}_i$ for the individual scattering by the potential $\overline{V}_i$ centered on the atom i. The eqn (1) can then be written as

$$\begin{aligned}\bar{\tau}^{ji} &= \overline{T}_j\delta_{ij} + \sum_{k\neq j}\overline{T}_j G_0 \bar{\tau}^{ki} \\ &= \overline{T}_j\delta_{ij} + \sum_{k\neq j}\bar{\tau}^{jk} G_0 \overline{T}_i,\end{aligned} \tag{2}$$

known as the equation of motion of the scattering path operator, which can be expressed in matrix form as

$$\begin{aligned}\boldsymbol{\tau} &= (\mathbf{I} - \mathbf{T}\mathbf{G}_0)^{-1}\mathbf{T} \\ &= \mathbf{T}(\mathbf{I} - \mathbf{G}_0\mathbf{T})^{-1}.\end{aligned} \tag{3}$$

The elements of the matrix $\boldsymbol{\tau}$ are the individual scattering path operators between the atoms i and j. For spherical potentials, the

[a] *Faculty of Science, University of Toyama, Gofuku 3190, Toyama 930-8555, Japan*
[b] *Univ Rennes, CNRS, IPR (Institut de Physique de Rennes) – UMR 6251, F-35000 Rennes, France*

matrix $\mathbf{T}$ is diagonal with elements $\overline{T}_i$, while the elements of $\mathbf{G}_0$ are $\mathscr{T}(-\vec{\mathbf{R}}_i)G_0\mathscr{T}(\vec{\mathbf{R}}_j)$. $\mathscr{T}(\vec{\mathbf{R}})$ is the translation operator which shifts position vectors by $\vec{\mathbf{R}}$; these translations are necessary to be able to use the transition operators for single scattering, which are determined for a potential centered on the origin. The matrix elements of $\mathscr{T}(-\vec{\mathbf{R}}_i)G_0\mathscr{T}(\vec{\mathbf{R}}_j)$ are also known as the KKR structure constants.

The elements of $\mathbf{G}_0$ can be calculated using a suitable basis (plane waves or spherical waves for example), while the individual, single scattering transition matrix elements can be obtained by matching the logarithmic derivative of the inner and outer solutions of the Schrödinger equation at the surface of the potential, using standard partial wave techniques. It is then straightforward to construct the kernel matrix $\mathbf{G}_0\mathbf{T}$ and compute the inverse of $(\mathbf{I} - \mathbf{G}_0\mathbf{T})$ to obtain $\boldsymbol{\tau}$. This approach, which we shall call the matrix inversion (MI) method, is in principle exact, but is limited by the size of the matrix that can be stored; in a cluster of N_{at} atoms, if for a given kinetic energy angular momenta up to $\ell_{\max}$ are required to adequately represent the propagating electron, the dimension of the matrix to be inverted will be of the order of $N = N_{\mathrm{at}}(\ell_{\max} + 1)^2$. Since the matrix elements are complex, the amount of storage required will be $16N^2$ bytes; the extra work space required in matrix inversion routines for example in the LAPACK library[5] can double this. Furthermore, the time required to invert the matrix increases as N^3, so that calculations can quickly become very lengthy and greedy for memory, particularly for larger systems at higher electron kinetic energies, since $\ell_{\max}$ increases with increasing energy. In such cases, other approaches such as the path operator expansion[6] or Lanczos techniques[7] could be useful.

An alternative procedure is to solve the eqn (2) iteratively by replacing $\overline{\tau}^{ki}$ on the right-hand side by its equation of motion, yielding

$$\overline{\tau}^{ji} = \overline{T}_j\delta_{ij} + \overline{T}_jG_0\overline{T}_i + \sum_{k\neq i,j}\overline{T}_jG_0\overline{T}_kG_0\overline{T}_i + \ldots \tag{4}$$

which is often referred to as the Watson series expansion for the scattering path operator. It is equivalent to a Taylor matrix series (MS) expansion of the inverse:

$$\begin{aligned}\boldsymbol{\tau} &= (\mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \mathbf{K}^3 + \ldots)\mathbf{T} \\ &= \mathbf{T}(\mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \mathbf{K}^3 + \ldots),\end{aligned} \tag{5}$$

where the non-Hermitian kernel matrix ($\mathbf{TG}_0$ or $\mathbf{G}_0\mathbf{T}$ depending on the form used) is now denoted by $\mathbf{K}$. When it comes to model complex phenomena in condensed matter physics arising from the propagation of an electron in a finite collection of hundreds of atoms at kinetic energies of hundreds of eV, the series expansion is the preferred choice. The advantage of this method is that it can be implemented without storing the full matrix, for example by recalculating the elements required when evaluating individual scattering path operators using eqn (2). This will greatly reduce the memory requirements of a simulation to the point of rendering it feasible, at the price of increasing the computation time, especially if the series converges slowly. Even worse, the simulation may fail if the series does not converge at all.

The convergence of series expansions is an old problem in mathematical physics that has been greatly studied over several centuries, leading to the development of a number of interesting approaches (see ref. 8 for an overview of several of these methods).

The aim of this article is to explore a way of extending the boundaries of the standard series expansion implementation by improving its convergence. The technique proposed is the so-called renormalization of the MS series, which was introduced a few years ago by Sébilleau and Natoli.[8]

The paper is organized as follows. In the first part (Section 2), we discuss an efficient way of monitoring the convergence properties of the series expansion, namely the application of the power method to compute the spectral radius of the kernel matrix. Next (Section 3), we present three different renormalization schemes aimed at improving the convergence. Finally, in Section 4 we apply these schemes in the calculation of the photoelectron diffraction diagram of a Cu(111) cluster, demonstrating that the renormalized series expansion can reproduce very well the exact matrix inversion solution. These methods are implemented in MsSpec, a multiple scattering package for spectroscopies.[9,10]

2 Monitoring the convergence of the MS expansion

2.1 The spectral radius

The MS series expansion approach is only meaningful if this expansion converges. Therefore, we need a tool that can help us monitor the convergence properties of the expansion. The spectral radius is exactly the tool we need. For a given matrix $\mathbf{K}$, the spectral radius is defined by

$$\rho(\mathbf{K}) = \max|\lambda_i| \quad i \in [1, N], \tag{6}$$

where λ_i is the ith eigenvalue of $\mathbf{K}$ and N is the dimension of the matrix. As was already recognized by Natoli and Benfatto,[11] if $\rho(\mathbf{K}) > 1$, the series expansion diverges; if $\rho(\mathbf{K}) < 1$, the series converges but the number of terms required is larger the closer $\rho(\mathbf{K})$ is to 1. Provided enough terms are included, the series expansion should give the same results as matrix inversion.

There is however more to the story, as the spectral radius can also be used to estimate the number of expansion terms necessary in order to reach a given accuracy. To illustrate this use, let us consider the scalar power series expansion of $1/(1-x)$, where x is a real number with absolute value less than 1. Table 1 gives the largest power of x that needs to be retained in the power series in order to converge to within a given accuracy, for x between 0.4 and 0.9. We see that to converge to within a relative error of 5% with fewer than 10 terms (*i.e.* retaining terms up to x^9) requires $x \lesssim 0.7$, while 20 terms will give convergence to about 1% for values of $x \lesssim 0.8$. This suggests that if we want a series expansion to converge in a reasonable number of terms, for example 10, the corresponding matrix should have a spectral radius of 0.7 or less.

Table 1 Highest power of x that needs to be retained in the power series expansion of $1/(1-x)$ (x a real number) for a given relative accuracy

x	0.4	0.5	0.6	0.7	0.8	0.9
Relative error $<10\%$	2	3	4	6	10	21
Relative error $<5\%$	3	4	5	8	13	28
Relative error $<1\%$	5	6	9	12	20	43

A more detailed analysis for the case of the kernel matrix $\mathbf{K} \equiv \mathbf{G}_0\mathbf{T}$ is given in Appendix 1. In particular, eqn (40) and (42) respectively provide estimates of the expansion order required to ensure that the absolute or relative truncation errors are within a specified accuracy. The two relations provide estimates that agree exactly with the numerical values in Table 1 for the convergence of the power series expansion of $1/(1-x)$.

2.2 Computing the spectral radius

Following the definition (6), the spectral radius $\rho(\mathbf{K})$ may be computed by direct diagonalization of the matrix $\mathbf{K}$. As an example of this, we consider a small Cu(111) cluster of 50 atoms arranged in 4 planes with the emitter of the photoelectron at the bottom (see Fig. 1). In Fig. 2 we plot the spectrum of the kernel matrix $\mathbf{K} \equiv \mathbf{G}_0\mathbf{T}$ at 4 different kinetic energies, 25 eV, 54 eV, 98 eV and 181 eV, as well as the spectral radius as a function of kinetic energy from 10 to 200 eV. The photoelectron wave is usually damped due to inelastic scattering along the electron path and also due to lattice vibrations. These two damping sources are known to strongly affect the convergence of the MS series: a short electron mean free path or large thermal vibration amplitudes of the atoms in the cluster will reduce the spectral radius and thus improve the convergence of the MS series. For this study, we will only consider damping due to the electron mean free path, included through the imaginary part of the complex Hedin–Lundqvist exchange and correlation potential.[12,13] Lattice vibrations can be described using a temperature-dependent Debye–Waller factor or by averaging over T-matrix elements, both of which are available in the MsSpec computer package. In this work, however, we deliberately choose to neglect vibrational damping in order to emphasize possible divergences while treating a relatively small system amenable to direct matrix inversion as well as series expansion calculations. The spectral radius may of course be greater than 1 for larger clusters, even when atomic vibrations are taken into account.

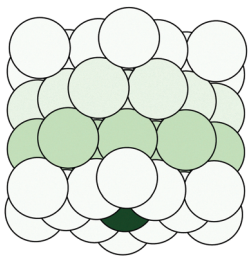

Fig. 1 A cylindrical cluster of 50 copper atoms, organized in 4 planes. The dark atom at the center of the bottom plane is the photo-electron emitter.

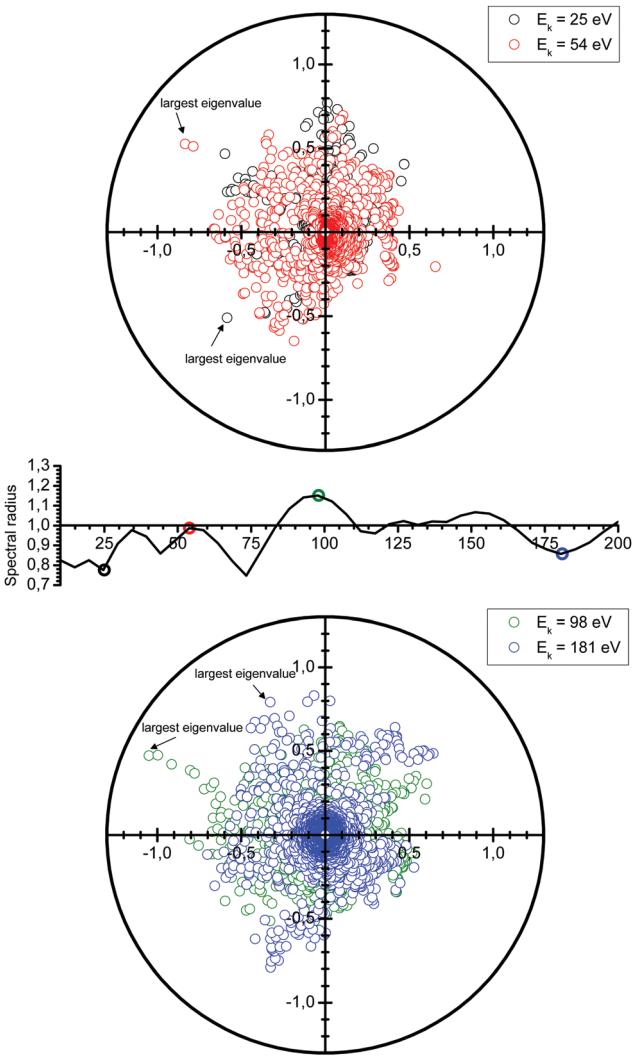

Fig. 2 Spectrum of the kernel matrix $\mathbf{K} \equiv \mathbf{G}_0\mathbf{T}$ for the kinetic energies of 25 eV, 54 eV, 98 eV and 181 eV. The variations of the spectral radius are also shown (middle figure).

We see in Fig. 2 that the spectral radius is greater than 1 for kinetic energies around 100 eV and around 150 eV. We also observe oscillations in the spectral radius. While not fully understood yet, these oscillations are the result of a complex interplay of the electron kinetic energy, the crystal structure, the electronic structure, as well as the number and type of atoms present.

The memory required to store $\mathbf{K}$ increases rapidly for larger clusters and higher kinetic energies, so that it becomes unfeasible to store and directly diagonalize the full matrix. An alternative that avoids storing the whole matrix is based on the power method, an iterative approach that gives a direct approximation to the spectral radius. The method is based on the assumption that the matrix $\mathbf{K}$ has one dominant eigenvalue, whose magnitude is much larger than the magnitudes of the other eigenvalues. It starts by choosing an initial approximation $\boldsymbol{x}_0$ to the eigenvector associated with this dominant eigenvalue, then forms the sequence of products

$$\boldsymbol{x}_1 = \mathbf{K}\boldsymbol{x}_0$$

$$\boldsymbol{x}_2 = \mathbf{K}\boldsymbol{x}_1 = \mathbf{K}^2\boldsymbol{x}_0$$

$$\boldsymbol{x}_3 = \mathbf{K}\boldsymbol{x}_2 = \mathbf{K}^3\boldsymbol{x}_0$$

$$\ldots$$

$$\boldsymbol{x}_n = \mathbf{K}\boldsymbol{x}_{n-1} = \mathbf{K}^n\boldsymbol{x}_0.$$

These values can give a good approximation to the dominant eigenvalue and eigenvector, and so of the spectral radius. We assume that none of the N eigenvalues of $\mathbf{K}$ are zero. Then there are N linearly independent normalized eigenvectors $\{\boldsymbol{v}_i\}$ with eigenvalues $\{\lambda_i\}$, with $i \in [1, N]$. We also assume that the eigenvalues are ordered so that $|\lambda_1| > |\lambda_i|$ for $i > 1$. The eigenvectors form a basis and the initial approximation $\boldsymbol{x}_0$ can be written as a linear combination of the eigenvectors:

$$\boldsymbol{x}_0 = \sum_{i=1}^{N} \alpha_i \boldsymbol{v}_i.$$

Multiplying both sides by $\mathbf{K}$ gives

$$\mathbf{K}\boldsymbol{x}_0 = \sum_{i=1}^{N} \alpha_i \mathbf{K}\boldsymbol{v}_i = \sum_{i=1}^{N} \alpha_i \lambda_i \boldsymbol{v}_i.$$

Repeating n times gives

$$\mathbf{K}^n\boldsymbol{x}_0 = \sum_{i=1}^{N} \alpha_i \lambda_i^n \boldsymbol{v}_i = \lambda_1^n \sum_{i=1}^{N} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^n \boldsymbol{v}_i. \tag{7}$$

Introducing an appropriate vector norm $\|\boldsymbol{x}\|$, such as the Euclidian norm

$$\|\boldsymbol{x}\|_2 = \left[\sum_i |x_i|^2\right]^{\frac{1}{2}},$$

and comparing successive approximations, we obtain an estimate for the spectral radius of $\mathbf{K}$:

$$\frac{\|\mathbf{K}^n\boldsymbol{x}_0\|_2}{\|\mathbf{K}^{n-1}\boldsymbol{x}_0\|_2} = |\lambda_1| \sqrt{\frac{\sum_i |\alpha_i|^2 \left|\frac{\lambda_i}{\lambda_1}\right|^{2n}}{\sum_i |\alpha_i|^2 \left|\frac{\lambda_i}{\lambda_1}\right|^{2n-2}}} \approx \rho(\mathbf{K}). \tag{8}$$

Another estimate of the spectral radius can be obtained from the Rayleigh quotient

$$R_{\mathbf{K}}(\boldsymbol{x}) = \frac{\boldsymbol{x}^\dagger \mathbf{K}\boldsymbol{x}}{\boldsymbol{x}^\dagger \boldsymbol{x}}. \tag{9}$$

Clearly if $\boldsymbol{x}$ is an eigenvector of $\mathbf{K}$ with eigenvalue λ, then $R_{\mathbf{K}}(\boldsymbol{x}) = \lambda$. Following the same procedure outlined above, an approximation to the spectral radius after n iterations is then given by

$$\frac{|(\mathbf{K}^n\boldsymbol{x}_0)^\dagger \mathbf{K}(\mathbf{K}^n\boldsymbol{x}_0)|}{|(\mathbf{K}^n\boldsymbol{x}_0)^\dagger (\mathbf{K}^n\boldsymbol{x}_0)|} = |\lambda_1| \sqrt{\frac{\left|\sum_i |\alpha_i|^2 \left|\frac{\lambda_i}{\lambda_1}\right|^{2n} \left(\frac{\lambda_i}{\lambda_1}\right)\right|}{\sum_i |\alpha_i|^2 \left|\frac{\lambda_i}{\lambda_1}\right|^{2n}}} \approx \rho(\mathbf{K}). \tag{10}$$

Note that there is a complex term λ_i/λ_1 in the numerator, in contrast to the expression (8).

As n increases, the sequence of approximations (8) and (10) converges to the spectral radius $\rho(\mathbf{K})$. The rate of convergence depends on the initial guess $\boldsymbol{x}_0$ and on the ratio $|\lambda_i/\lambda_1|$ for $I > 1$; if this is close to 1, the convergence will be slower. Acceleration techniques such as the epsilon algorithm[14] can also be used to extrapolate the sequence of approximations to obtain a better value. Several such acceleration methods are implemented in the calculation of the spectral radius in the MsSpec computer code.[9]

In order to implement these two approximations, we need to define a starting vector $\boldsymbol{x}_0$. Several choices have been tested,[8] and the value $\boldsymbol{x}_0 = \left(1, j^{\frac{1}{2}}, j, j^{\frac{3}{2}}, \ldots\right)$ was found to perform particularly well in terms of the number of iterations needed in order to achieve convergence. Here, we have taken j as the cube root of unity, $e^{2\pi i/3}$.

To better understand the behaviour of the iterative process, we present in Fig. 3 the evolution of the ratio $|\lambda_2|/|\lambda_1|$ as a function of energy (upper panel), and of the approximation to the spectral radius given by eqn (8) (lower panel). We note that when the two largest eigenvalues are close, the power method as well as the Rayleigh quotient can be misled and jump to the energy trajectory of the second largest eigenvalue. More precisely, even though the power method's approximation to the spectral radius may not follow exactly the largest eigenvalue, we see that in practice this happens when the second largest eigenvalue lies within approximately 5% of $|\lambda_1|$; the resulting approximate spectral radius is however still a good measure of the convergence of the matrix series expansion.

3 Simple renormalization schemes

Now that we have developed an efficient tool to monitor the behaviour of a matrix series expansion, we can apply it to the

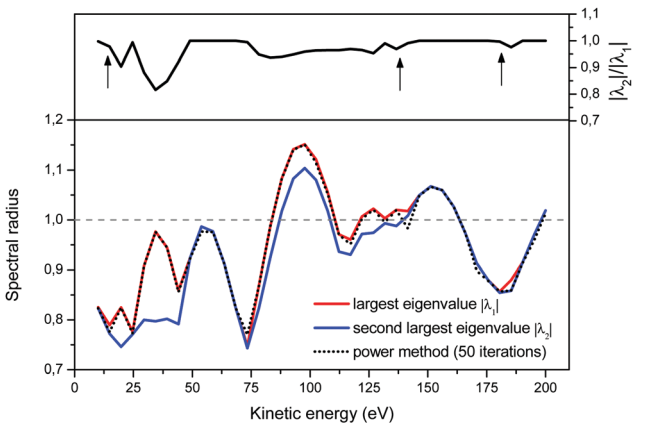

Fig. 3 The Euclidian 2-norm approximation to the spectral radius, eqn (8), after 50 iterations in the power method. Note that the spectral algorithm can jump from following the largest eigenvalue to following the second-largest eigenvalue when the ratio of their magnitudes is close to 1 (upper panel).

MS problem, and more particularly to the Watson expansion. Our aim in this section is to improve the convergence properties of matrix series expansions of the form

$$(\mathbf{I} - \mathbf{K})^{-1} = \mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \mathbf{K}^3 + \ldots \tag{11}$$

For this, we follow ideas developed by Janiszowski[15] for deriving an algorithm for matrix inversion in processors with limited calculation abilities. The main ideas have already been outlined by Sébilleau *et al.*,[16] but we go here into more details.

We seek a general transformation of the form

$$(\mathbf{I} - \mathbf{K})^{-1} = (\mathbf{I} - \mathbf{M}(\omega, \mathbf{K}))^{-1}\mathbf{N}(\omega, \mathbf{K}), \tag{12}$$

implying

$$\mathbf{M}(\omega, \mathbf{K}) = \mathbf{I} - \boldsymbol{N}(\omega, \mathbf{K})(\mathbf{I} - \mathbf{K}), \tag{13}$$

where $\boldsymbol{\omega}$ is a complex scalar chosen to minimize the spectral radius of $\boldsymbol{M}$ so that the series expansion of $(\mathbf{I} - \mathbf{M})^{-1}$ converges rapidly. We call $\boldsymbol{N}$ the renormalization matrix. Several renormalization schemes were investigated under the constraint provided by eqn (12), including the Löwdin second-order iteration method,[17] or the Euler expansion,[17] but in the end, we focus on the three transformations described in detail in the following subsections.

3.1 $\mathbf{G}_n$ renormalization

We set

$$\mathbf{M} \equiv \mathbf{G}_n = (1 - \omega^n)\mathbf{I} + \omega_n\mathbf{K} \tag{14}$$

so that

$$\begin{aligned}(\mathbf{I} - \mathbf{K})^{-1} &= \omega^n(\mathbf{I} - \mathbf{G}_n)^{-1} \\ &= \omega^n\left(\mathbf{I} + \mathbf{G}_n + \mathbf{G}_n{}^2 + \mathbf{G}_n{}^3 + \ldots\right).\end{aligned} \tag{15}$$

The matrix $\mathbf{N}$ is then just $\omega^n\mathbf{I}$.

This is the simplest renormalization scheme; the renormalized matrix $\mathbf{M}$ is a linear combination of the identity matrix and the kernel matrix. The renormalization parameter appears as ω^n, where n is an integer. This general expression may be useful in formally deriving other renormalization schemes but is less useful for numerical work since by noting $\tilde{\omega} = \omega^n$, we can rewrite $\mathbf{G}_n(\omega) = \mathbf{G}_1(\tilde{\omega})$. In the examples below, we will in general only consider the $\mathbf{G}_1$ renormalization.

Following the definition (14), the eigenvalues λ^G of $\mathbf{G}_n$ are related to those of the kernel matrix $\mathbf{K}$ by

$$\lambda_i^G = (1 - \omega^n) + \omega^n\lambda_i. \tag{16}$$

We then have

$$\begin{aligned}\rho(\mathbf{G}_n) &= \max_i|(1 - \omega^n) + \omega^n\lambda_i| \\ &\leq |1 - \omega^n| + |\omega^n|\rho(\mathbf{K}).\end{aligned} \tag{17}$$

The condition $\rho(\mathbf{G}_n) < \rho(\mathbf{K})$ will be satisfied if

$$|1 - \omega^n| + |\omega^n|\rho(\mathbf{K}) < \rho(\mathbf{K}). \tag{18}$$

Writing $\omega = r\exp(\mathrm{i}\theta)$, the condition (18) becomes

$$|1 - r^n\mathrm{e}^{\mathrm{i}n\theta}| < (1 - r^n)\rho(\mathbf{K}), \tag{19}$$

and since $\rho(\mathbf{K})$ is positive by definition, we must have $r < 1$ so that ω lies within the unit circle. Squaring both sides of (19), rearranging and setting $x = (1 + r^{2n})/2r^n$, with $x > 1$ when $r < 1$, yields

$$\cos(n\theta) > x + (1 - x)\rho^2(\mathbf{K}) \tag{20}$$

as a constraint on the values of the argument of ω to ensure that $\rho(\mathbf{G}_n) < \rho(\mathbf{K})$.

3.2 $\boldsymbol{\Sigma}_n$ renormalization

The different $\mathbf{G}_n$ renormalizations satisfy a recurrence relation

$$\mathbf{G}_n(\omega) = (1 - \omega)\mathbf{I} + \omega\mathbf{G}_{n-1}, \tag{21}$$

with $\mathbf{G}_0 = \mathbf{K}$. We define the $\boldsymbol{\Sigma}_n$ renormalization scheme of order n as the average of the partial sums of $\mathbf{G}_k$ for k from 0 up to n:

$$\begin{aligned}\mathbf{M} \equiv \Sigma_n &= \frac{1}{n+1}\sum_{k=0}^{n}\mathbf{G}_k \\ &= (1 - s_n)\mathbf{I} + s_n\mathbf{K},\end{aligned} \tag{22}$$

where

$$s_n = \frac{1}{n+1}\sum_{k=0}^{n}\omega^k = \frac{1 - \omega^{n+1}}{(n+1)(1-\omega)}. \tag{23}$$

We then have

$$(\mathbf{I} - \mathbf{K})^{-1} = s_n(\mathbf{I} - \boldsymbol{\Sigma}_n)^{-1} \tag{24}$$

with $\mathbf{N} = s_n\mathbf{I}$.

The definitions (14) and (22) for $\mathbf{G}_n$ and $\boldsymbol{\Sigma}_n$ have the same functional form; they are both linear in the kernel matrix $\mathbf{K}$, and will have the same eigenvalue spectrum if the renormalization parameter in the $\boldsymbol{\Sigma}_n$ case, ω_Σ, is chosen so that $s_n(\omega_\Sigma) = \omega_G^n$. In particular, for $n = 1$, we see that $\boldsymbol{\Sigma}_1(2\omega - 1) = \mathbf{G}_1(\omega)$. Mathematically, these two renormalization schemes are equivalent, giving identical smallest spectral radii so that the convergence properties of the renormalized series expansion are the same. It is however interesting to distinguish the two renormalization schemes: as we shall see later, in cases where we cannot find exactly the optimal value for ω that minimizes the spectral radius, the $\boldsymbol{\Sigma}_n$ renormalization scheme can have a wider range of ω giving acceptably small spectral radii than in the $\mathbf{G}_n$ scheme.

The eigenvalues λ^Σ of $\boldsymbol{\Sigma}_n$ are related to those of the kernel matrix $\mathbf{K}$ by

$$\lambda_i^\Sigma = (1 - s_n) + s_n\lambda_i. \tag{25}$$

To ensure that $\rho(\boldsymbol{\Sigma}_n) < \rho(\mathbf{K})$, the condition (20) with $r < 1$ still applies, where now $s_n = r\exp(\mathrm{i}\theta)$.

3.3 The Löwdin $\mathbf{\Pi}_n$ renormalization

We define the $\mathbf{\Pi}_n$ renormalization matrix as the product of $\mathbf{G}_k$ for k from 0 up to n:

$$\mathbf{\Pi}_n = \prod_{k=0}^{n} \mathbf{G}_k. \tag{26}$$

For $n = 1$ this gives

$$\mathbf{M} \equiv \mathbf{\Pi}_1 = (1 - \omega)\mathbf{K} + \omega\mathbf{K}^2 \tag{27}$$

so that

$$(\mathbf{I} - \mathbf{K})^{-1} = (\mathbf{I} - \mathbf{\Pi}_1)^{-1}(\mathbf{I} + \omega\mathbf{K}) \tag{28}$$

where $\mathbf{N} = \mathbf{I} + \omega\mathbf{K}$. The eigenvalues of $\mathbf{\Pi}_1$ are

$$\lambda_i^{\Pi} = (1 - \omega)\lambda_i + \omega\lambda_i^2. \tag{29}$$

More generally, we may define

$$\mathbf{X}_n = \omega^n \mathbf{\Pi}_{n-1} = \omega\mathbf{X}_{n-1}\mathbf{G}_{n-1}$$

with $\mathbf{X}_0 = \mathbf{I}$ and

$$\boldsymbol{\sigma}_n = \sum_{k=0}^{n} \mathbf{X}_k.$$

We then have

$$\omega\mathbf{X}_{n-1} - \mathbf{X}_n = \omega\mathbf{X}_{n_1}(\mathbf{I} - \mathbf{G}_{n-1}) = \omega^n\mathbf{X}_{n-1}(\mathbf{I} - \mathbf{K}),$$

where we have used

$$\mathbf{G}_n = (1 - \omega^n)\mathbf{I} + \omega^n\mathbf{K} \Rightarrow (\mathbf{I} - \mathbf{G}_n) = \omega^n(\mathbf{I} - \mathbf{K}).$$

We then can write

$$\begin{aligned}
\mathbf{X}_n &= \omega\mathbf{X}_{n-1} + \omega^n\mathbf{X}_{n-1}(\mathbf{I} - \mathbf{K}) \\
&= \omega^2\mathbf{X}_{n-2} + \omega^n[\mathbf{X}_{n-1} + \mathbf{X}_{n-2}](\mathbf{I} - \mathbf{K}) \\
&= \dots \\
&= \omega^k\mathbf{X}_{n-k} + \omega^n[\mathbf{X}_{n-1} + \mathbf{X}_{n-2} + \dots + \mathbf{X}_{n-k}](\mathbf{I} - \mathbf{K}) \\
&= \dots \\
&= \omega^n\mathbf{I} - \omega^n\sum_{k=0}^{n-1}\mathbf{X}_k(\mathbf{I} - \mathbf{K}) \\
&= \omega^n\mathbf{I} - \omega^n\sigma_k(\mathbf{I} - \mathbf{K}).
\end{aligned}$$

Using the definition of $\mathbf{X}_n$, we finally obtain

$$(\mathbf{I} - \mathbf{K})^{-1} = (\mathbf{I} - \mathbf{\Pi}_n)^{-1}\boldsymbol{\sigma}_n, \tag{30}$$

which agrees with eqn (28) when $n = 1$ since $\boldsymbol{\sigma}_1 = \mathbf{I} + \omega\mathbf{K}$.

It is interesting to note from the definition (26) that[18]

$$\rho(\mathbf{\Pi}_n) \le \prod_{k=0}^{n} \rho(\mathbf{G}_k)$$

and in particular

$$\rho(\mathbf{\Pi}_1) \le \rho(\mathbf{K})\rho(\mathbf{G}_1). \tag{31}$$

This shows that if the regular Taylor and $\mathbf{G}_1$ series expansions both converge, then the $\mathbf{\Pi}_1$ series will converge faster than either of them. Furthermore, the condition $\rho(\mathbf{\Pi}_1) < \rho(\mathbf{K})$ implies that the spectral radius of $\mathbf{G}_1$ must be less than 1; replacing $\rho(\mathbf{K})$ by 1 on the right-hand side of the inequality (18) and rearranging yields

$$\cos(\theta) > \rho(\mathbf{K}) + \frac{r}{2}(1 - \rho^2(\mathbf{K}))$$

with $r < 1/\rho(\mathbf{K})$ as the conditions to be respected.

3.4 Expressing the renormalization schemes as power series in K

The effect of renormalizing the Taylor expansion (11) can be seen by re-expressing the expansions of the transformed inverse as power series in the kernel matrix $\mathbf{K}$:

$$(\mathbf{I} - \mathbf{K})^{-1} = \sum_{k=0}^{+\infty} R_k(\omega, \infty)\mathbf{K}^k \approx \sum_{k=0}^{N_s} R_k(\omega, N_s)\mathbf{K}^k, \tag{32}$$

where N_s is the truncation order of the expansion.

For the $\mathbf{G}_n$ renormalization scheme, writing $g_n \equiv \omega^n$, the first few terms in the series expansion (15) are

$$\begin{aligned}
\mathbf{G}_n^0 &= \mathbf{I} \\
\mathbf{G}_n^1 &= (1 - g_n)\mathbf{I} + g_n\mathbf{K} \\
\mathbf{G}_n^2 &= (1 - g_n)^2\mathbf{I} + 2g_n(1 - g_n)\mathbf{K} + g_n^2\mathbf{K}^2 \\
\mathbf{G}_n^3 &= (1 - g_n)^3\mathbf{I} + 3g_n(1 - g_n)^2\mathbf{K} + 3g_n^2(1 - g_n)\mathbf{K}^2 + g_n^3\mathbf{K}^3 \\
&\dots
\end{aligned}$$

which can be summarized as

$$\mathbf{G}_n^k = \sum_{m=0}^{k}\binom{k}{m}(1 - g_n)^{k-m}g_n^m\mathbf{K}^m \tag{33}$$

where $\binom{k}{m}$ is the standard binomial coefficient. Inspecting the coefficients of $\mathbf{K}^k$ gives

$$\begin{aligned}
R_k(\omega, N_s) &= g_n^{k+1}\sum_{m=k}^{N_s}\binom{m}{k}(1 - g_n)^{m-k} \\
&= g_n\left(\frac{g_n}{1 - g_n}\right)^k\sum_{m=k}^{N_s}\binom{m}{k}(1 - g_n)^m.
\end{aligned} \tag{34}$$

When $\omega = 1$, every coefficient $R_k(\omega, N_s)$ reduces to 1, and we recover the Taylor series expansion. A similar expression holds for the $\mathbf{\Sigma}_n$ renormalization scheme, with s_n replacing g_n in the relation (34).

The corresponding expression for the $\mathbf{\Pi}_1$ renormalization scheme is more complicated. The demonstration is given in the appendix, leading to the final result

$$\begin{aligned}
(\mathbf{I} - \mathbf{K})^{-1} &\approx (\mathbf{I} + \mathbf{\Pi}_1 + \mathbf{\Pi}_1^2 + \cdots \mathbf{\Pi}_1^{N_s})\boldsymbol{\sigma}_1 \\
&= \sum_{k=0}^{2N_s+1} R_k(\omega, N_s)\mathbf{K}^k,
\end{aligned} \tag{35}$$

Table 2 Summary of the different renormalization schemes $(\mathbf{I}-\mathbf{K})^{-1} = (\mathbf{I}-\mathbf{M})^{-1}\mathbf{N}$

Scheme	Equation	$\mathbf{M}$	$\mathbf{N}$	MS expansion $(\mathbf{I}-\mathbf{M})^{-1}$	Renormalized coefficient $R_k(\omega, N_s)$
Taylor	(11)	$\mathbf{K}$	$\mathbf{I}$	$\mathbf{I}+\mathbf{K}+\mathbf{K}^2+\cdots+\mathbf{K}^{N_s}$	1
$\mathbf{G}_n$	(15)	$\mathbf{G}_n$	$g_n\mathbf{I}$	$\mathbf{I}+\mathbf{G}_n+\mathbf{G}_n{}^2+\cdots+\mathbf{G}_n{}^{N_s}$	$g_n^{k+1}\sum_{m=k}^{N_s}(1-g_n)^{m-k}$
$\mathbf{\Sigma}_n$	(24)	$\mathbf{\Sigma}_n$	$s_n\mathbf{I}$	$\mathbf{I}+\mathbf{\Sigma}_n+\mathbf{\Sigma}_n{}^2+\cdots+\mathbf{\Sigma}_n{}^{N_s}$	$s_n^{k+1}\sum_{m=k}^{N_s}\binom{m}{k}(1-s_n)^{m-k}$
$\mathbf{\Pi}_1$	(28)	$\mathbf{\Pi}_1$	$\boldsymbol{\sigma}_1$	$\mathbf{I}+\mathbf{\Pi}_1+\mathbf{\Pi}_1{}^2+\cdots+\mathbf{\Pi}_1^{N_s}$	$\sum_{m=\lceil k/2\rceil}^{\min(k,N)_s}[Y_1^m]_k$

with

$$R_k(\omega, N_s) = \sum_{m=\lceil k/2\rceil}^{\min(k,N_s)} \left[Y_1^m\right]_k \tag{36}$$

where $\lceil k/2\rceil$ represents the integer part of $k/2$, and

$$\left[Y_1^m\right]_k = \omega^{k-m}(1-\omega)^{2m-k}\left[\binom{m}{k-m} + (1-\omega)\binom{m}{k-m-1}\right], \tag{37}$$

with the convention that $\binom{m}{k} = 0$ if $k < 0$ or $k > m$. Since $\mathbf{\Pi}_1$ involves a term in $\mathbf{K}^2$, truncating the expansion of $(\mathbf{I}-\mathbf{\Pi}_1)^{-1}$ at $\mathbf{\Pi}_1{}^{N_s}$ yields a power series in $\mathbf{K}$ of order $2N_s+1$. Due to the fact that the summation over m on the right-hand side of (36) is truncated when $k > N_s$, the expansion (35) coincides with the Taylor expansion (11) up to order $\mathbf{K}^{N_s}$ and differs only for powers of $\mathbf{K}$ between N_s+1 and $2N_s+1$ (see Appendix 2).

The characteristics of the different renormalization schemes are summarized in Table 2.

4 Application to photoelectron diffraction by a copper cluster

As a test case for the renormalization schemes presented above, we consider photoelectron diffraction in the 50 atom Cu(111) cluster illustrated in Fig. 1. This small system can be treated by direct matrix inversion techniques, and is thus suitable for validating our matrix renormalization approaches. In particular, we suppose that photons generated by an X-ray source and incident at 55° with respect to the normal to the material surface, eject a $2p_{1/2}$ electron (ionization energy 967 eV) which is then scattered by atoms in the cluster before eventually escaping the surface. Photoelectrons with a given kinetic energy are detected in different directions (polar and azimuthal angles with respect to the normal at the surface); in the results presented below, we shall only consider polar scans, for which the azimuthal angle is fixed at zero. Inelastic effects are taken into account using a complex Hedin–Lundqvist exchange and correlation potential.[12,13]

The crystal structure is built using the Atomic Simulation Environment (ASE) Python package,[19] while the photoelectron diffraction is treated using MsSpec,[9,10] which we have extended to include the three renormalization schemes presented above. As well as treating several different types of spectrocopy, the MsSpec package can also be used to compute the spectral radius of the kernel and renormalized matrices as a function of the kinetic energy of the photoelectron and the renormalization parameter, either by direct diagonalization or iteratively using the power method combined with various acceleration techniques.

The spectral radius of the kernel matrix $\mathbf{K}$ as a function of the photoelectron kinetic energy from 10 to 200 eV is shown in Fig. 4. For a photoelectron energy of 95 eV, where the spectral radius is the largest, the maximum angular momentum is $\ell_{\max} = 8$ and the size of the matrix $\mathbf{K}$ is $9^2 \times 50 = 4050$. Inverting this relatively small matrix using the Lapack library took a few minutes using a modest Intel Core 8th generation i5 processor, while the power method converges after 50 iterations and took less than a minute. The spectral radius of $\mathbf{K}$ is 1.15, and hence the series expansion of $(\mathbf{I}-\mathbf{K})^{-1}$ diverges.

In what follows, we shall only consider the renormalization schemes $\mathbf{G}_n$, $\mathbf{\Sigma}_n$ and $\mathbf{\Pi}_n$ with $n = 1$; the index n will therefore be omitted, unless otherwise stated.

If all the eigenvalues of the matrix $\mathbf{K}$ are known, the spectral radius of the renormalized matrices can be easily computed using the relations (16), (25) and (29) for any value of ω. Contour plots of the spectral radius as a function of ω for the three renormalized matrices are shown in Fig. 5 at 4 selected energies. The spectral radius of the original matrix $\mathbf{K}$ appears at $\omega = (1, 0)$ in the $\mathbf{G}$ scheme; it is less than 1 at 70 and 75 eV, while it is greater than 1 at 90 and 95 eV. The plots present two interesting features:

- the spectral radius of $\mathbf{\Sigma}$ is smaller than 1 over a relatively wider range of ω than for the other two schemes; in other

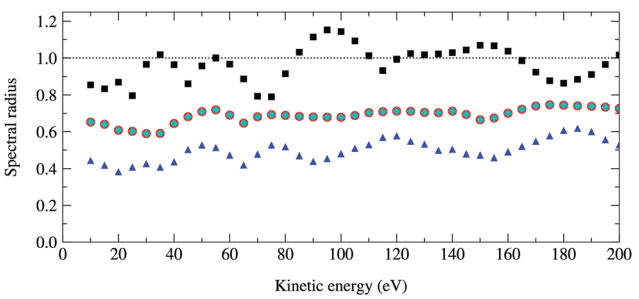

Fig. 4 Optimal spectral radius of the renormalized matrices $\mathbf{G}$, $\mathbf{\Sigma}$ and $\mathbf{\Pi}$ as a function of kinetic energy for a 50 atom Cu(111) system. For each renormalization scheme, these are computed by finding at each energy the value of ω that gives the smallest spectral radius. Squares, circles, diamonds and triangles correspond to the spectral radii of $\mathbf{K}$, $\mathbf{G}_1$, $\mathbf{\Sigma}_1$ and $\mathbf{\Pi}_1$ respectively; the results for $\mathbf{G}_1$ and $\mathbf{\Sigma}_1$ are identical.

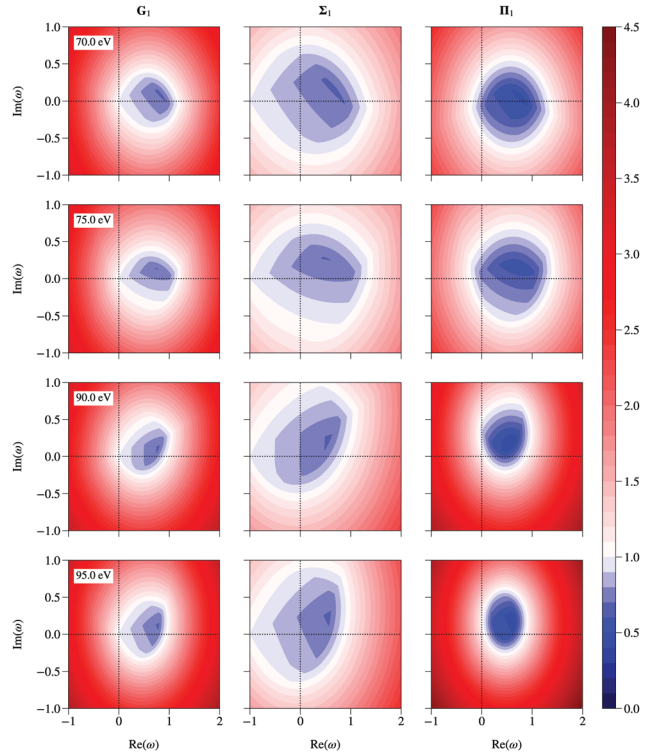

Fig. 5 Contour plots of the spectral radius as a function of the complex parameter ω for the **G**, **Σ** and **Π** renormalized matrices at 70, 75, 90 and 95 eV.

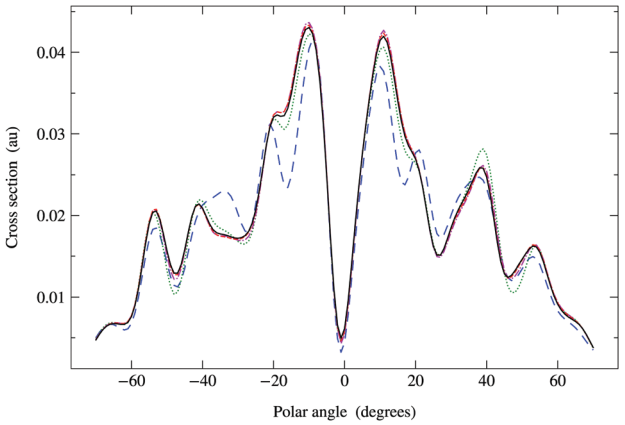

Fig. 6 Comparison of photoelectron diffraction cross sections at 95 eV. Full curve: direct matrix inversion of (**I** − **K**); series expansion with **G** renormalization scheme, optimal ω_G = 0.790 + 0.124i giving a spectral radius $\rho(\mathbf{G}_1)$ = 0.68: long dashed curve, truncation order 3; dotted curve, truncation order 5; dash-dotted curve, truncation order 7; dashed curve, truncation order 9.

words, the spectral radius for **Σ** renormalization is less sensitive to the precise value of ω than in the other two schemes, which may be an advantage in numerical work when searching for an optimal value of ω;

• the strongest reduction occurs for the **Π** renormalization scheme, but over a relatively restricted range of ω.

We also note that the optimal values of ω all lie within the unit circle, as expected following the arguments in section 3 above.

Also shown in Fig. 4 are the optimal spectral radii for the three renormalized matrices **G**, **Σ** and **Π**. The optimal spectral radius at a particular energy is obtained by finding the value of ω that minimizes the spectral radius; each point on a particular curve corresponds to a different ω. The search, performed using the Nelder–Mead simplex algorithm,[20] is usually quick when the eigenvalues of the kernel matrix **K** have been found by direct diagonalization, since the eigenvalues and hence the spectral radii are rapidly obtained from the relations (16), (25) and (29). In contrast, the power method must be re-applied to calculate the spectral radius for each required value of ω, potentially resulting in longer search times which may dominate the overall computing time.

We first note that all spectral radii with renormalization are reduced compared to the spectral radii of the matrix **K**, sometimes by as much as 50%. This implies that the matrix series now converges over the whole energy range considered. As expected, the optimal spectral radii for **G** and **Σ** are identical. The **Π** renormalization scheme gives the smallest optimal spectral radii, and hence potentially the best convergence properties for the matrix series expansion.

In Fig. 6, we compare the cross sections obtained by direct matrix inversion and those obtained by matrix series expansion with **G** renormalization scheme. The optimal value of ω_G is 0.790 + 0.124i, giving a spectral radius $\rho(\mathbf{G}_1)$ = 0.68. Renormalization has thus converted a divergent series into one that converges quite well by order 8 or 9. More precisely, the maximum relative difference (MRD) between the cross sections obtained by direct matrix inversion and by a renormalized series expansion is about 6% after truncating the **G** expansion beyond N_s = 9; terms up to N_s = 13 are required to obtain an MRD of less than 1%. From eqn (42) in Appendix 1, however, we expect the truncation error for the series expansion of the matrix inverse to be about 2% for N_s = 9 and less than 1% for N_s = 11. The cross sections thus converge slightly more slowly with increasing series expansion order, which is understandable since the expression for the cross section involves the square of the scattering path operator. We have also noted that the MRD always seems to occur when the polar angle θ is zero, possibly due to the fact that the cross section has a deep minimum in this direction.

The **Σ** renormalization scheme yields identical results for optimal $\omega_\Sigma = 2\omega_G - 1$ = 0.581 + 0.247i, although this analytical relation might not be always verified for values determined by numerical search and tiny differences between cross sections might thus appear.

Since the **Π** renormalization scheme yields a smaller spectral radius, $\rho(\mathbf{\Pi}_1)$ = 0.45, the matrix expansion should converge even faster. The results shown in Fig. 7 suggest that a truncation order as low as 4 or 5 is sufficient; the MRD for N_s = 5 is about 1.4% while for N_s = 6 it is only 0.3%. The corresponding truncation errors for the matrix inverse predicted by eqn (42) are respectively 0.8% and 0.4%. This rapid convergence is due to the fact that the **Π** matrix series expansion truncated at order N_s = 5 contains powers of **K** up to order $2N_s + 1$ = 11, the truncation order at which the **G** renormalized series expansion also converged to within 1%.

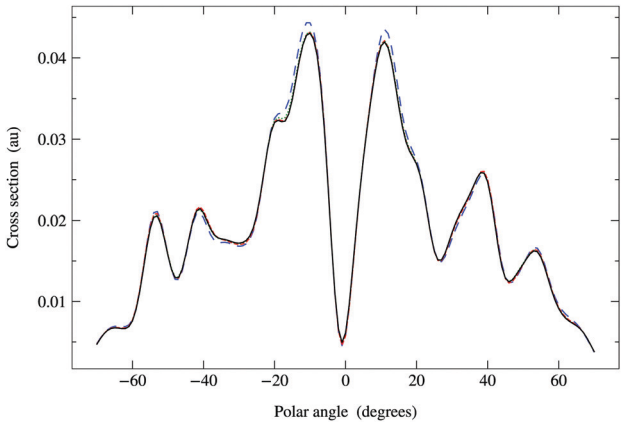

Fig. 7 Comparison of photoelectron diffraction cross sections at 95 eV. Full curve: direct matrix inversion of $(\mathbf{I} - \mathbf{K})$. Series expansion with $\mathbf{\Pi}_1$ renormalization scheme, optimal ω_P = 0.637 + 0.250i giving a spectral radius $\rho(\mathbf{\Pi}_1)$ = 0.45: long dashed curve, truncation order 3; dashed curve, truncation order 4; dotted curve, truncation order 5. The convergence is more rapid than for **G** renormalization, only requiring terms up to order 4 or 5.

Table 3 Comparison of the relative truncation error ε_{rel} in the renormalized series expansion of order N_s of the matrix inverse $(\mathbf{I} - \mathbf{K})$, given by eqn (42), and the maximum relative difference (MRD) between the cross sections obtained by direct matrix inversion and by series expansion. All values are percentages

$\mathbf{G}_1$ renormalization						
N_s	3	5	7	9	11	13
ε_{rel}	21.4	9.9	4.6	2.1	0.98	0.45
MRD	34.8	20.3	13.2	6.01	2.03	0.65
$\mathbf{\Pi}_1$ renormalization						
N_s	3		4	5	6	
ε_{rel}	4.10		1.85	0.83	0.37	
MRD	7.75		6.98	1.35	0.32	

The expected relative truncation errors in the series expansion of the inverse matrix operator $(\mathbf{I} - \mathbf{K})$ and the maximum relative difference (MRD) between the cross sections obtained by direct matrix inversion and by series expansion are summarized in Table 3.

5 Conclusions

The matrix series expansion (5) converges if the spectral radius of the kernel matrix **K** is less than 1. This may be sufficient for further mathematical development, but the convergence may be too slow to be of use in numerical work. The idea of renormalization is to recast the series expansion in another form with faster convergence properties, by introducing a renormalized kernel matrix **M** with a smaller spectral radius. Multiplication by an associated matrix $\boldsymbol{N}$ is then performed to recover the exact result.

We have proposed three simple renormalization schemes involving a single complex parameter ω that can achieve this goal, yielding a convergent series even for cases where the original expansion diverges. Numerical tests on a relatively small Cu(111) cluster show that the $\mathbf{\Pi}_1$ scheme is particularly advantageous, reducing the spectral radius by up to a factor 2 in some cases. The main difficulty is determining an optimal value of the renormalization parameter ω; this can be found using standard optimization approaches such as the Nelder–Mead simplex algorithm[20] combined with an iterative approach for determining the spectral radius.

When used to simulate typical X-ray photoelectron diffraction experiments with kinetic energies in the 100–1500 eV range, the standard Watson series expansion may not converge properly. Computing spectral radii by direct matrix inversion for such cases is not possible because of the large memory requirements. The power method allows the convergence properties to be predicted efficiently for such cases. However, even though the tools and methods presented here succeeded in improving the MS series convergence, the overall procedure of optimizing ω can become quite long. A dedicated study to better understand the origin of the divergence and the variation in the spectral radius is in progress to further improve the renormalization procedure.

Conflicts of interest

There are no conflicts to declare.

Appendix 1: estimate of the truncation order

To simplify the notation, let us write $\mathbf{A} = (\mathbf{I} - \mathbf{K})^{-1}$. We can then approximate **A** by a series expansion of order n:

$$\mathbf{A}_n = \mathbf{I} + \mathbf{K} + \mathbf{K}^2 + \mathbf{K}^3 + \mathbf{K}^n. \quad (38)$$

The absolute truncation error ε_{abs} introduced by this approximation is then

$$\|\mathbf{A} - \mathbf{A}_n\| \leq \varepsilon_{abs}, \quad (39)$$

where $\|\ \|$ represents any matrix norm. Using eqn (38), the previous inequality can be written as

$$\|(\mathbf{K})^{n+1}(\mathbf{I} - \mathbf{K})^{-1}\| \leq \varepsilon_{abs}.$$

We can then use the property

$$\|\mathbf{XY}\| \leq \|\mathbf{X}\| \cdot \|\mathbf{Y}\|,$$

which is valid for any matrix norm, as well as[18]

$$\rho(\mathbf{X}^n) = \rho(\mathbf{X})^n \leq \|\mathbf{X}\|^n.$$

Choosing the minimum norm for which[21]

$$\rho(\mathbf{X}) = \inf_{\|\cdot\|} \|\mathbf{X}\|,$$

the inequality (39) will always be satisfied if

$$\rho(\mathbf{K}^{n+1})(1-\rho(\mathbf{K}))^{-1} \leq \varepsilon_{\text{abs}}.$$

This leads to the condition

$$n \gtrsim \frac{\log[\varepsilon_{\text{abs}}(1-\rho(\mathbf{K}))]}{\log[\rho(\mathbf{K}))]} - 1. \tag{40}$$

For instance, choosing an accuracy $\varepsilon_{\text{abs}} = 0.1$, we find $n \sim 43$ for $\rho(\mathbf{K}) = 0.9$, $n \sim 9$ for $\rho(\mathbf{K}) = 0.7$, while for $\rho(\mathbf{K}) = 0.5$ n drops to ~ 3.

Defining the relative truncation error ε_{rel} by

$$\frac{\|\mathbf{A}-\mathbf{A}_n\|}{\|\mathbf{A}\|} \leq \varepsilon_{\text{rel}}, \tag{41}$$

and using similar arguments, we obtain an estimate for the truncation order required for convergence to within ε_{rel}:

$$n \gtrsim \frac{\log[\varepsilon_{\text{rel}}]}{\log[\rho(\mathbf{K}))]} - 1. \tag{42}$$

Appendix 2: the $\mathbf{\Pi}_1$ renormalization scheme as a power series in K

The $\mathbf{\Pi}_n$ renormalization scheme is (*cf.* eqn (26)):

$$(\mathbf{I}-\mathbf{K})^{-1} = (\mathbf{I}-\mathbf{\Pi}_n)^{-1}\boldsymbol{\sigma}_n,$$

which for $n = 1$ reduces to (*cf.* eqn (28))

$$(\mathbf{I}-\mathbf{K})^{-1} = (\mathbf{I}-\mathbf{\Pi}_1)^{-1}(\mathbf{I}+\omega\mathbf{K}), \tag{43}$$

with $\mathbf{\Pi}_1 = (1-\omega)\mathbf{K} + \omega\mathbf{K}^2$. We now show how to write the renormalized series expansion as a series of powers of $\mathbf{K}$ (*cf.* eqn (35)–(37)):

$$(\mathbf{I}-\mathbf{K})^{-1} = \sum_{k=0}^{\infty} R_k(\omega)\mathbf{K}^k. \tag{44}$$

We recall that

$$\mathbf{\Pi}_n = \prod_{k=0}^{n} \mathbf{G}_k,$$

with $\mathbf{\Pi}_0 = \mathbf{G}_0 = \mathbf{K}$, so that for $n = 1$,

$$\mathbf{\Pi}_1 = \mathbf{G}_0\mathbf{G}_1 = \mathbf{K}\mathbf{G}_1$$

with $\mathbf{G}_1 = (1-\omega)\mathbf{I} + \omega\mathbf{K}$. The Taylor expansion of the matrix inverse $(\mathbf{I}-\mathbf{\Pi}_1)^{-1}$ in eqn (43) is

$$(\mathbf{I}-\mathbf{K})^{-1} = \left(\sum_{k=0}^{\infty}(\mathbf{K}\mathbf{G}_1)^k\right)(\mathbf{I}+\omega\mathbf{K}).$$

Using the binomial expansion (33) with $n = 1$ and $g_n = \omega$ gives

$$(\mathbf{I}-\mathbf{K})^{-1} = \left(\sum_{k=0}^{\infty}\sum_{m=0}^{k}\binom{k}{m}(1-\omega)^{k-m}\omega^m\mathbf{K}^{k+m}\right)(\mathbf{I}+\omega\mathbf{K}) \tag{45}$$

$$= \sum_{k=0}^{\infty}\sum_{m=0}^{k}\binom{k}{m}(1-\omega)^{k-m}\omega^m\mathbf{K}^{k+m}$$

$$+ \sum_{k=0}^{\infty}\sum_{m=0}^{k}\binom{k}{m}(1-\omega)^{k-m}\omega^{m+1}\mathbf{K}^{k+m+1}. \tag{46}$$

We now examine the coefficients of $\mathbf{K}^n$ on the right-hand side of eqn (46). In the first term, we must have $n = k + m$ or $m = n - k$. Since we must also have $m \leq k$, the values of k can be n, $n-1$, $n-2,\ldots,[(n+1)/2]$, where $[p/2]$ is the integer part of $p/2$. Similarly, in the second term on the right-hand side of eqn (46), we must have $m = n-k-1$ with k taking values between $n-1$ and $[n/2]$. Combining gives the coefficient of $\mathbf{K}^n$ as

$$R_n(\omega) = \sum_{k=[(n+1)/2]}^{n}\binom{k}{n-k}(1-\omega)^{2k-n}\omega^{n-k} + \sum_{k=[n/2]}^{n-1}\binom{k}{n-k-1}(1-\omega)^{2k-n+1}\omega^{n-k},$$

which can in turn be written as

$$R_n(\omega) = \sum_{k=[n/2]}^{n}\omega^{n-k}(1-\omega)^{2k-n}\left[\binom{k}{n-k}+(1-\omega)\binom{k}{n-k-1}\right] \equiv \sum_{k=[n/2]}^{n}\left[Y_1^k\right]_n \tag{47}$$

with the convention that $\binom{k}{m} = 0$ if $m < 0$ or $m > k$.

It should also be noted that if the series expansion for $(\mathbf{I}-\mathbf{\Pi}_1)^{-1}$ is truncated after $N_s + 1$ terms, the largest power of $\mathbf{K}$ in the series (44) will be $2N_s + 1$, while the summations over k in (47) should be truncated at $\min(n, N_s)$. For example, truncating the expansion after $\mathbf{\Pi}_1{}^2$ ($N_s = 2$) gives:

$$\begin{aligned}(\mathbf{I}-\mathbf{K})^{-1} &\approx (\mathbf{I}+\mathbf{\Pi}_1+\mathbf{\Pi}_1{}^2)(\mathbf{I}+\omega\mathbf{K})\\ &= (\mathbf{I}+(1-\omega)\mathbf{K}+(1-\omega+\omega^2)\mathbf{K}^2+2\omega(1-\omega)\mathbf{K}^3\\ &\quad+\omega^2\mathbf{K}^4)(\mathbf{I}+\omega\mathbf{K})\\ &= \mathbf{I}+\mathbf{K}+\mathbf{K}^2+\omega(3-3\omega+\omega^2)\mathbf{K}^3\\ &\quad+\omega^2(3-2\omega)\mathbf{K}^4+\omega^3\mathbf{K}^5.\end{aligned}$$

This also illustrates that the $\mathbf{\Pi}_1$ renormalization scheme truncated after $\mathbf{\Pi}_1{}^{N_s}$ reproduces the Taylor series for $(\mathbf{I}-\mathbf{\Pi}_1)^{-1}$ up to order N_s, followed by correcting terms for powers of $\mathbf{K}$ from $N_s + 1$ to $2N_s + 1$.

Author contributions

D. S. and K. H. contributed to the conceptualization and supervision of this work; A. T., S. T., M. T.-D. and K. D. implemented the ideas in the MsSpec computer package and performed the calculations; all authors participated in the analysis and validation of the results, and in writing the article.

Acknowledgements

K. H. acknowledges funding by JST CREST Grant No. JPMJCR1861 and JSPS KAKENHI under Grant No. 18K05027 and 17K04980. A.T. acknowledges the award of a Tobitate scholarship from the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. The calculations were performed using resources provided by the Institute of Physics of Rennes.

Notes and references

1 H. Ebert, *et al.*, *The Munich SPR-KKR package, version 7.7*, 2017, https://www.ebert.cup.uni-muenchen.de/en/software-en/13-sprkkr.
2 H. Ebert, D. Ködderitzsch and J. Minár, *Rep. Prog. Phys.*, 2011, **74**, 096501.
3 D. Sébilleau, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2000, **61**, 14167–14178.
4 D. Sébilleau, in *Multiple Scattering Theory for Spectroscopies*, ed. D. Sébilleau, K. Hatada and H. Ebert, Springer International Publishing, 2018, ch. 1, pp. 3–34.
5 E. Anderson, Z. Bai, C. Bischof, S. Blackford, J. Demmel, J. Dongarra, J. Du Croz, A. Greenbaum, S. Hammarling, A. McKenney and D. Sorensen, *LAPACK Users' Guide*, 3rd edn, Society for Industrial and Applied Mathematics, Philadelphia, PA, 1999.
6 Y. Chen, F. J. García de Abajo, A. Chassé, R. X. Ynzunza, A. P. Kaduwela, M. A. Van Hove and C. S. Fadley, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 1998, **58**, 13121–13131.
7 A. L. Ankudinov, C. E. Bouldin, J. J. Rehr, J. Sims and H. Hung, *Phys. Rev. B: Condens. Matter Mater. Phys.*, 2002, **65**, 104107.
8 D. Sébilleau and C. R. Natoli, *J. Phys.: Conf. Ser.*, 2009, **190**, 012002.
9 D. Sébilleau, C. Natoli, G. M. Gavaza, H. Zhao, F. Da Pieve and K. Hatada, *Comput. Phys. Commun.*, 2011, **182**, 2567–2579.
10 https://msspec.cnrs.fr.
11 C. R. Natoli and M. Benfatto, *J. Phys. C: Solid State Phys.*, 1986, **8**, 11.
12 L. Hedin and S. Lundqvist, *Solid State Physics*, 1969, p. 1.
13 L. Hedin and B. I. Lundqvist, *J. Phys. C: Solid State Phys.*, 1971, **4**, 2064–2083.
14 P. Wynn, *Math. Comput.*, 1956, **10**, 91–96.
15 K. B. Janiszowski, *Int. J. Appl. Math. Comput. Sci.*, 2003, **13**, 199–204.
16 D. Sébilleau, K. Hatada, H.-F. Zhao and C. R. Natoli, *e-J. Surf. Sci. Nanotechnol.*, 2012, **10**, 599–608.
17 P.-O. Löwdin, *J. Math. Phys.*, 1962, **3**, 969–982.
18 F. Kittaneh, *Proc. Am. Math. Soc.*, 2006, **134**, 385–390.
19 A. Hjorth Larsen, *et al.*, *J. Phys.: Condens. Matter*, 2017, **29**, 273002.
20 J. A. Nelder and R. Mead, *Comput. J.*, 1965, **7**, 308–313.
21 A. Czornik and P. Jurgaś, *Int. J. Appl. Math. Comput. Sci.*, 2006, **16**, 183.

