Kerker mixing scheme for self-consistent muffin-tin based all-electron electronic structure calculations

Miriam Winkelmann [1,2,3,4,*] Edoardo Di Napoli [1,2] Daniel Wortmann [1,3] and Stefan Blügel [1,3]

[1] *Institute for Advanced Simulation, Forschungszentrum Jülich, and JARA-CSD, D-52425 Jülich, Germany*

[2] *Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany*

[3] *Peter Grünberg Institute, Forschungszentrum Jülich, D-52425 Jülich, Germany*

[4] *Physics Department, RWTH-Aachen University, D-52062 Aachen, Germany*

(Received 19 March 2020; revised 11 October 2020; accepted 20 October 2020; published 20 November 2020)

We propose a computationally efficient Kerker mixing scheme for robust and rapidly converging self-consistent-field calculations using all-electron first-principles electronic structure methods based on the muffin-tin partitioning of space. The mixing scheme is composed of the Kerker preconditioner in combination with quasi-Newton methods. We construct the Kerker preconditioner in the muffin-tin sphere by determining the screened Coulomb potential in real space, solving a modified Helmholtz equation by adopting Weinert's pseudocharge method for calculating the Poisson equation for periodic charge densities without shape approximation to the solution of the modified Helmholtz equation. Implemented in a full-potential linearized augmented plane-wave (FLAPW) method, we found that the Kerker preconditioning scheme (i) leads to a convergence to self-consistency that is independent of system size, (ii) is extremely robust in the choice of the mixing and preconditioning parameters, (iii) scales linearly with system size in computational cost, and (iv) conserves the total charge. We have related the preconditioning parameter to the density of states of the delocalized electrons at the Fermi energy and developed a model to choose the preconditioning parameter either prior to the calculation or on the fly. Our computationally validated model supports the hypothesis that, in the absence of Kerker preconditioning, the delocalized s and p electrons of simple and transition metals are the primary cause for the slowing of the convergence speed and that the stronger, localized d and f electrons account for only a small fraction of the charge sloshing problem. The presented formulation of the Kerker preconditioning scheme establishes an efficient methodology for the simulation of magnetic and nonmagnetic metallic large-scale material systems by means of muffin-tin-based all-electron methods.

DOI: 10.1103/PhysRevB.102.195138

I. INTRODUCTION

Density functional theory (DFT) is considered the standard model for the calculation of electronic and structural properties of materials. In this formalism, the quantum mechanical ground state of a many-electron system is the exact solution ρ_* to the Kohn-Sham (KS) equations. Due to the nonlinearity of the Kohn-Sham equations, a direct computation of the ground-state charge density ρ_* is not possible and one has to resort to a fixed-point iteration scheme—the "self-consistent-field" (SCF) iteration—where a new, better approximation of ρ_* is calculated at each iteration step. In practical applications, an iteration step, or "cycle," of the SCF iteration includes solving the Poisson equation, evaluating an approximation to the unknown exchange-correlation potential for a given charge density, and subsequently solving the Schrödinger-like Kohn-Sham equation for each occupied state, from which a new charge density is synthesized. In this paper, the abstract functional f describes a cycle, and we aim at improving the convergence of the whole SCF iteration.

In the last 35 years, advanced mixing methods such as Tchebycheff [1], Broyden [2], Anderson [3], or Pulay mixing [4,5], and variants and adaptations [6–14] of these, have been introduced in a wide range of electronic structure methods, so as to ensure SCF convergence with only a few shortcomings in practical simulations. Their basic principle is to assemble a new charge density $\rho^{(m+1)}$ for the next iteration $m+1$ by mixing input charge densities $\rho^{(i)}$ and output charge densities $f(\rho^{(i)})$ of previous iterations $i \leqslant m$.

One such shortcoming is "charge sloshing," which often occurs in large metallic systems. In each cycle, the charge moves too far in the physical sample simulated, resulting in a too strongly restoring electrostatic potential, which eventually leads to oscillations of the charge that prevents the convergence of the SCF iteration. The response strength associated with the misplaced charge density is proportional to the Coulomb interaction, which scales as $4\pi/K^2$ with the magnitude of the sampled wave vector $\boldsymbol{K}$, the smallest of which has length $2\pi/L$, where L is the characteristic size of the unit cell. Because of such dependence, we expect a decrease in the amount of output charge density that is mixed into the next input charge density and an increase of the number of iterations to achieve SCF convergence scaling as the square of the system size (L^2) [15]. In principle,

*Corresponding author: m.winkelmann@fz-juelich.de

the computation of systems of increasing size and chemical or structural complexity could be carried out by increasing the amount of computational resources. On the other hand, the slowing down of the convergence speed forces the use of an ever-increasingly disproportionate amount of resources to the point that it can develop into a roadblock for large-scale *ab initio* calculations. Charge sloshing also impedes a reliable and stable path to self-consistency, an important requirement for high-throughput calculations in the context of autonomous computing and materials discovery. Translated in mathematical terms, the problem of finding the fixed point is ill-conditioned. A common approach to remedy such a problem is designing and using a preconditioner.

Sophisticated SCF preconditioners [16–20] have been developed using a linear approximation of the functional f. This linearization operates on two fronts: first on the susceptibility describing the response of the charge density with respect to the change of the effective potential, and second on the Coulomb and exchange-correlation kernel, describing the response of the effective potential with respect to changes of the charge density. While most SCF preconditioners for metallic systems show excellent convergence results, they are too expensive in practical computations.

For plane-wave-based methods, the Kerker preconditioner [10,11,21] provides a computationally light-weight alternative, while still significantly improving SCF convergence. In contrast to the more expensive methods, the Kerker preconditioner discussed in Sec. II A makes a large simplifying assumption: It replaces the dielectric function of the Kohn-Sham system, by one of uniform distribution of electrons, yielding a simple mathematical form of the preconditioner in reciprocal space. Such a setup effectively enables the screening of long-range electron interactions and thereby suppresses charge sloshing. Thanks to this simple scheme, plane-wave-based pseudopotential [22] or projector augmented plane-wave (PAW) [23,24] methods typically make use of a Kerker preconditioning scheme.

In contrast, "all-electron" methods, such as the augmented spherical wave (ASW) [25], linear muffin-tin orbital techniques (LMTO) [26], Korringa-Kohn-Rostoker (KKR) Green function [27], and full-potential linearized augmented plane-wave (FLAPW) method [28], treat core electrons on the same footing as valence electrons. In order to deal with the Coulomb singularity produced by the nuclear charge and the associated rapid variation of the charge densities in the vicinity of the nucleus, all-electron methods partition the space of the unit cell into muffin-tin spheres in which wave functions and charge densities are represented in real space. A Fourier representation of these quantities would hardly converge. Thus, in order to be able to use the Kerker preconditioner efficiently in all-electron methods, a real-space formulation is advantageous and conceptually consistent with the real-space representation of the charge density. A first implementation and results were reported in Ref. [29]. A real-space Kerker method has also been developed for pseudopotential approaches [30], but the concept is quite different to what is presented in this work due to the replacement of true charge densities in these methods.

Here we present a formulation of the Kerker mixing scheme for full-potential all-electron methods employing the muffin-tin decomposition of charge density and potential for three-dimensional periodic crystalline solids. We focus on the development of a computationally efficient Kerker preconditioner for stable and fast-converging SCF computations, the verification of convergence properties, and the setup of a material-specific rule for the choice of optimal parameters in the mixing and preconditioning parameter space. The scheme is implemented in the FLEUR [31] code, a realization of the FLAPW method [28].

We have demonstrated the performance and efficiency of the Kerker mixing scheme using the example of metal surfaces, which we realize with the supercell approach. The supercell consists of atoms representing the substrate, atoms representing the surface and missing atoms representing the vacuum. This way we work with very unfortunate unit cells, which have very disproportionate c/a ratios, i.e., very thin and long wire-shaped unit cells with many atoms in one direction and a minimum number of atoms in the plane of the unit cell. Such unit cells of metallic systems are prone to charge sloshing. We study 17 different metals—alkali, transition, noble, post-noble, nonmagnetic, as well as magnetic metals. We have chosen both symmetrical and asymmetrical surfaces, varying the number of atoms in the unit cell between 19 and 79, and we treated systems with more than 1200 valence electrons.

Our tests show that the convergence provided by the Kerker mixing scheme is rather robust with respect to the choice of preconditioning and mixing parameter. Most importantly, the SCF convergence is shown to become system-size independent, as it is typical for the Kerker preconditioner in its original reciprocal-space formulation and, at the same time, already effective for relatively small systems. These results make this mixing scheme applicable to a huge span of systems all the way up to huge ones investigated using large parallel compute clusters. The results of the material-specific model supports the hypothesis that the charge sloshing problem arises mainly from the response of the s and p electrons of metals and only in part from the response of the more localized d and f electrons. On the other hand, our model can be used to automatize the parameter selection by direct calculation on the fly, a valuable asset for high-throughput computing.

The paper is organized as follows: In Sec. II A we introduce the Jacobian to the residual charge density functional, and discuss the Thomas-Fermi approximation to the Kohn-Sham response function and the generalization to a local response function with a free preconditioning parameter. This is followed by a description of the origin of the charge sloshing and a representation-free formulation of the Kerker preconditioner. The solution of the modified Helmholtz equation for a residual charge density is introduced as an elementary step relating the residual charge density to the preconditioned residual density. We conclude with the introduction of the Kerker mixing scheme, a combination of preconditioner and quasi-Newton method. In Sec. II B, we briefly illustrate the representation of the charge density and the potential in the FLAPW method. We then discuss an extension of Weinert's pseudocharge method [32] for the solution of the Poisson equation applied to the modified Helmholtz equation described in Ref. [33]. Section II also discusses aspects of charge neutrality, the treatment of magnetic systems, and

describes the preconditioning and mixing scheme algorithms. In Sec. III, we examine the performance of our implementation of the Kerker preconditioner in the FLAPW code FLEUR [31], showing improvements of convergence even for bulk systems of only 19 atoms. In Sec. IV, we develop and analyze a model that relates the preconditioning parameter to the density of states of delocalized electrons, showing good agreement of our model and computationally determined optimal parameters. We conclude with a summary of the main points and a discussion of some of the open issues.

II. PRINCIPLE AND GENERAL CONSIDERATIONS

A. Charge sloshing phenomenon and Kerker preconditioner

Charge sloshing is a problem with SCF convergence. The SCF convergence is mainly determined by the derivative of the residual functional $\delta\rho = f(\rho) - \rho$ close to the sought-after fixed point ρ^*. Let $\rho^{(m)}$ and $f(\rho^{(m)})$ at some iteration m be the input and output charge density, respectively. It is well known that taking the output charge density as new input for the next iteration, i.e., $\rho^{(m+1)} = f(\rho^{(m)})$, only converges if the eigenvalues of $\frac{df}{d\rho}$ in the vicinity of the fixed point are sufficiently smaller than 1 in absolute value [34]. Instead, one could take a mixture of output and input,

$$\rho^{(m+1)} = \rho^{(m)} + C\delta\rho^{(m)}, \tag{1}$$

where, after discretization, $\delta\rho$ is a vector quantity and C is, in general, a square matrix. This way, one has control over the convergence, since the absolute error of iteration $m+1$ is

$$\begin{aligned}\rho^{(m+1)} - \rho^* &= (\rho^{(m)} - \rho^*) + C\frac{\delta\rho^{(m)} - \delta\rho^*}{\rho^{(m)} - \rho^*}(\rho^{(m)} - \rho^*) \\ &\approx (\mathbb{1} + CJ)(\rho^{(m)} - \rho^*),\end{aligned} \tag{2}$$

where $\mathbb{1}$ is the identity matrix and $J = \frac{d\delta\rho}{d\rho}$ is the Jacobian of the residual close to the fixed point. Choosing $C \approx -J^{-1}$ brings the right-hand side close to 0 and thus $\rho^{(m+1)}$ close to the fixed point. So our aim is to approximate $-J^{-1}$. In general, C depends on the iteration m.

A cycle consists of several steps: Based on a charge density one first solves the Poisson equation and adds to the solution an exchange-correlation potential and, if available, an external potential. The total or "effective" potential then enters the KS equations, which we solve for the KS orbitals, which again produce a new charge density by summation over the lowest-lying orbitals occupied by electrons. The derivative of the cycle functional f [34],

$$J = \frac{df}{d\rho} - \mathbb{1} = \chi_{\mathrm{KS}}(K_{\mathrm{H}} + K_{\mathrm{xc}}) - \mathbb{1}, \tag{3}$$

is closely related to the dielectric matrix, where K_{H} is the Hartree or Coulomb kernel, respectively, K_{xc} is the exchange-correlation kernel, and χ_{KS} is the static Kohn-Sham response function. The derivative of an external potential cancels out, since V_{ext} is external to the electrons and thus does not depend on the charge density.

In metallic systems, the Kohn-Sham susceptibility of delocalized itinerant electrons with large group velocities may well be approximated by the susceptibility of the homogeneous electron gas or jellium system, respectively, known as the static Lindhard response function [35], $\chi_{\mathrm{KS}}(\boldsymbol{r}, \boldsymbol{r}') \simeq \chi_{\mathrm{L}}(\boldsymbol{r} - \boldsymbol{r}')$. Manninen *et al.* [36] introduced an efficient mixing scheme for inhomogeneous jellium systems in real space extended by Kerker [21] to self-consistent pseudopotential calculations in momentum space, and interpreted by Kresse *et al.* [10,11] as a preconditioner, that employs the Thomas-Fermi (TF) approximation to the KS susceptibility, here presented in real-space representation,

$$\chi_{\mathrm{TF}}(\boldsymbol{r} - \boldsymbol{r}') = -\frac{k_{\mathrm{TF}}^2}{4\pi}\,\delta(\boldsymbol{r} - \boldsymbol{r}') \tag{4}$$

(given in Hartree atomic units), which is today understood as the long wavelength limit of the Lindhard function, $\chi_{\mathrm{L}}(K) \xrightarrow{K\to 0} \chi_{\mathrm{TF}}(K) = -\frac{k_{\mathrm{TF}}^2}{4\pi}$, in momentum space. The Thomas-Fermi approximation witnesses a K-independent response function, which means that the electrons respond in the same way for each of the long wavelength perturbations and respond locally in space to the slowly varying potential changes. The Thomas-Fermi wave number, $k_{\mathrm{TF}} = (4/\pi)k_{\mathrm{F}}$, is related to the magnitude of the Fermi wave vector, $k_{\mathrm{F}} = 3\pi^2 n_\circ$, and thus to the constant electron density $n_\circ$ of the homogeneous electron gas. Thus, it is a constant of the system and therefore independent of the state of iteration m. This makes the Jacobian $J^{(m)} \to J$ completely m independent.

Although at long wavelengths one averages over the atomistic details, the true solid is beyond the picture of the homogeneous electron gas, and a one-to-one system specific relation of the Thomas-Fermi wave number with the electronic properties of the actual system might be difficult to achieve. Therefore, one replaces the TF-wave number by a constant λ, $k_{\mathrm{TF}} \to \lambda$, which represents the inverse of a typical length scale over which an individual charged particle exerts a notable effect. Typically λ is treated as an external preconditioning parameter that is fixed "experimentally" based on the experience of an actual calculation.

So far, little attention has been paid to the relation [37],

$$\chi_{\mathrm{TF}}(\boldsymbol{r} - \boldsymbol{r}') = -z(E_{\mathrm{F}})\,\delta(\boldsymbol{r} - \boldsymbol{r}'), \tag{5}$$

between the TF susceptibility [or through (4) the TF wave vector] and the density of states (DOS), $z(E_{\mathrm{F}})$, of the homogeneous electron gas of electron density $n_\circ$ described in terms of the Fermi energy E_{F}. In fact, this is a special example of a more general relation between local susceptibility and local DOS at the Fermi energy. We show in Sec. IV that this relation offers a much greater potential in optimizing the convergence of the SCF cycle than the relation between the susceptibility and the Thomas-Fermi wave number. In the future it might be interesting to test whether the iteration-step dependent density of states $z^{(m)}(E_{\mathrm{F}})$ can be used to accelerate convergence.

With the TF approximation to the susceptibility, by neglecting the effect of the exchange-correlation operator in (3), the preconditioning Jacobian of the Kerker mixing scheme becomes $J_{\mathrm{K}} = -\mathbb{1} + \chi_{\mathrm{TF}}K_{\mathrm{H}}$. Recalling the Poisson equation, $\Delta K_{\mathrm{H}} = -4\pi\mathbb{1}$ (Δ stands for the Laplace operator), amounting to $(\Delta - \lambda^2)K_{\mathrm{H}} = -4\pi(\mathbb{1} + \frac{\lambda^2}{4\pi}K_{\mathrm{H}})$, we formally write the

negative inverse Jacobian as

$$-J_{\mathrm{K}}^{-1} = \frac{\mathbb{1}}{\mathbb{1} + \frac{\lambda^2}{4\pi} K_{\mathrm{H}}} = \mathbb{1} - \frac{\mathbb{1}}{\mathbb{1} + \frac{\lambda^2}{4\pi} K_{\mathrm{H}}} \frac{\lambda^2}{4\pi} K_{\mathrm{H}}$$
$$= \mathbb{1} - \frac{\lambda^2}{4\pi} \cdot \frac{-4\pi\mathbb{1}}{\Delta - \lambda^2}. \tag{6}$$

We use the Yukawa or screened Coulomb potential δV_λ the solution to the modified Helmholtz equation,

$$(\Delta - \lambda^2)\, \delta V_\lambda = -4\pi \delta\rho, \tag{7}$$

to express the preconditioned charge density residual,

$$\delta\rho_\lambda = -J_{\mathrm{K}}^{-1}\, \delta\rho = \delta\rho - \frac{\lambda^2}{4\pi} \delta V_\lambda, \tag{8}$$

appearing in (1) as $C\delta\rho$.

The Kerker iteration scheme then takes the form,

$$\rho^{(m+1)} = \rho^{(m)} + \delta\rho_\lambda^{(m)}. \tag{9}$$

The formulation of the Kerker preconditioner in terms of the last three equations (7), (8), and (9) has the advantage of being basis independent and can thus be used in any formulation of an electronic structure method.

Since the Kerker preconditioner is a rough approximation, Kresse *et al.* [10,11] have shown that it is advisable not to rely on the preconditioner alone, i.e., using (9), but to use it additionally to advanced mixing methods such as Broyden [2,6,13], Anderson [3], or Pulay mixing [4,5]. These are quasi-Newton methods and the inverse of the Jacobian, $J^{-1(m)} = -\alpha\mathbb{1} + \sum_{j=2}^{m} u_j \otimes v_j$ [38], is typically expressed in terms of dyadic products of vectors u_j and v_j that are recursively synthesized at each iteration step m, where $J^{-1(1)} = -\alpha\mathbb{1}$ is the starting Jacobian and α is a scalar mixing parameter. The preconditioner is applied to the residual charge density $\delta\rho^{(m)}$ [see (1)] and the resulting preconditioned residual charge density $\delta\rho_\lambda^{(m)}$ then substitutes $\delta\rho^{(m)}$ in the entry of the mixing algorithm of choice implemented in the respective electronic structure methods, i.e., $\rho^{(m+1)} = \textsc{Mix}(\rho^{(m)}, \delta\rho^{(m)}, \alpha)$ changes to

$$\rho^{(m+1)} = \textsc{Mix}(\rho^{(m)}, \delta\rho_\lambda^{(m)}, \alpha). \tag{10}$$

When applied in this way, the mixing method does not approximate the inverse Jacobian of the residual functional, but the inverse Jacobian of the preconditioned residual functional. Note that by choosing Kerker preconditioning ($C = -J_{\mathrm{K}}^{-1}$), the matrix $\mathbb{1} + CJ$ in (2) becomes approximately 0 for all $\boldsymbol{r}$, and thus $\rho^{(m+1)}$ is brought close to the fixed point. Therefore, we can apply a stronger mixing than usual with mixing parameter α much closer to 1. In the absence of Kerker preconditioning ($\lambda = 0$) and large materials systems, convergence is reached only for small mixing parameters α. We show in Sec. III, that including Kerker preconditioning ($\lambda > 0$), the convergence radius with α changes fundamentally to the point that it becomes nearly independent of α.

B. Solution in the FLAPW method

The basis-set independent formulation of the Kerker preconditioner presented above provides the ground work for applications in all-electron methods. Typical of all-electron methods is the Coulomb singularity at the center of the atoms due to the positively charged nuclei and the rapidly oscillating core and valence electron wave functions in the vicinity of the nuclei. To deal with these circumstances, in full-potential all-electron methods such as the FLAPW method [28], the computational domain is divided into spheres $B_{R_\alpha}(\boldsymbol{\tau}_\alpha)$ around the centers of atoms—the union of all these spheres is called the muffin-tin (MT) region—and an interstitial (I) region. Both charge densities and potentials are represented by functions,

$$g(\boldsymbol{r}) = \begin{cases} \sum_{\boldsymbol{K}} g^{\mathrm{I}}(\boldsymbol{K}) e^{i\boldsymbol{K}\cdot\boldsymbol{r}} & \boldsymbol{r} \in \mathrm{I} \\ \sum_L g_L^\alpha(r_\alpha) Y_L(\hat{\boldsymbol{r}}_\alpha) & \boldsymbol{r} = \boldsymbol{\tau}_\alpha + \boldsymbol{r}_\alpha \in B_{R_\alpha}(\boldsymbol{\tau}_\alpha), \end{cases} \tag{11}$$

where r_α is the radius of the vector $\boldsymbol{r}_\alpha = \boldsymbol{r} - \boldsymbol{\tau}_\alpha$, $\hat{\boldsymbol{r}}_\alpha = \frac{\hat{\boldsymbol{r}} - \boldsymbol{\tau}_\alpha}{|\hat{\boldsymbol{r}} - \boldsymbol{\tau}_\alpha|}$ a unit vector, $\boldsymbol{\tau}_\alpha$ is the location of the sphere α in the unit cell, and Y_L are spherical harmonics. L is defined as $L := (\ell, m)$. Clearly, the representation in the MT region makes a real-space formulation of the preconditioner necessary.

For all-electron methods, the central impediment for the application of the Kerker preconditioning scheme is the low-cost solution of the modified Helmholtz equation (7) needed to obtain the screened Coulomb or Yukawa potential [39] of the residual charge density. It can be obtained efficiently by extending Weinert's pseudocharge method [32] for the solution of the Poisson's equation ($\lambda = 0$) to the case $\lambda > 0$ [33,40].

In Weinert's method, the Coulomb potential in the I region is solved by the Fourier transformation of the Poisson equation replacing the true charge density in the MT region by a Fourier transformable pseudocharge density with multipole moments equal to the true charge density. Thus, in case of the modified Helmholtz equation, the interstitial Yukawa potential $\delta V_\lambda^{\mathrm{I}}$ undergoes changes indirectly through the pseudocharge density, and directly by the prefactor $\frac{4\pi}{K^2+\lambda^2}$ that substitutes $\frac{4\pi}{K^2}$ in relating potential and charge density. The muffin-tin potential is then determined by a subsequent solution of a Dirichlet boundary value problem. This involves a radial Green function, whose radial dependence is slightly different for the Poisson and modified Helmholtz equation. In the case of the Poisson equation, we work with a polynomial radial (r) dependence, i.e., r^ℓ and $1/r^{\ell+1}$. In the modified Helmholtz equation, these quantities are substituted by the two modified spherical Bessel functions $i_\ell(\lambda r)$ and $k_\ell(\lambda r)$ [41,42]. The different radial behavior of the Green function affects the radial dependence of the muffin-tin Yukawa potential due to both the charge density in the muffin-tin sphere and on the sphere boundary. It also leads to changes in the multipole moments of the interstitial and muffin-tin charge densities in each sphere, $q_L^\alpha[\rho^{\mathrm{I}}]$ and $q_L^\alpha[\rho^\alpha]$, respectively, and in the Fourier components of the pseudocharge density $\tilde{\rho}^\alpha(\boldsymbol{K})$. For more details see Ref. [33].

C. Extension to collinear magnetic systems

In the case of a nonmagnetic system, we precondition the residual charge density by subtracting $\frac{\lambda^2}{4\pi}\delta V_\lambda$ from it [see (8)]. The preconditioned residual charge density then substitutes the original residual in the call of the mixing routine. In a system with collinear magnetism, i.e., with spin-up and spin-down charge densities, we first calculate the total charge

ALGORITHM I. Kerker preconditioner.

Input: charge density residual $\delta\rho = f(\rho) - \rho$, or, in case of collinear magnetism, $\delta\rho^{\uparrow}$ and $\delta\rho^{\downarrow}$; preconditioning parameter λ.
Output: preconditioned charge-density residual $\delta\rho_\lambda$, or, in the case of collinear magnetism, $\delta\rho_\lambda^{\uparrow}$ and $\delta\rho_\lambda^{\downarrow}$.
if collinear magnetism **then**
  Convert the pair $(\delta\rho^{\uparrow}, \delta\rho^{\downarrow})$ to $(\delta\rho, \delta m)$. ▷ Eq. (12)
end if
Compute $\delta V_\lambda \leftarrow$ YUKAWA POTENTIAL$(\delta\rho)$. ▷ Alg. 1 in [33]
Assemble preconditioned charge-density residual $\delta\rho_\lambda$. ▷ Eq. (8)
if collinear magnetism **then**
  Convert the pair $(\delta\rho_\lambda, \delta m)$ to $(\delta\rho_\lambda^{\uparrow}, \delta\rho_\lambda^{\downarrow})$. ▷ Eq. (13)
end if

density $\delta\rho$ and the magnetization δm from the spin-up and spin-down charge densities,

$$\delta\rho = \delta\rho^{\uparrow} + \delta\rho^{\downarrow} \quad \text{and} \quad \delta m = \delta\rho^{\uparrow} - \delta\rho^{\downarrow}. \tag{12}$$

We only precondition the total charge density, i.e., we compute $\delta\rho_\lambda$ from $\delta\rho$ by (8) and the screened Coulomb potential as described in the previous section. Finally we transform the above system back to the spin-up and spin-down charge densities as if the preconditioning step had not taken place, i.e.,

$$\delta\rho_\lambda^{\uparrow} = \frac{\delta\rho_\lambda + \delta m}{2} \quad \text{and} \quad \delta\rho_\lambda^{\downarrow} = \frac{\delta\rho_\lambda - \delta m}{2}. \tag{13}$$

The idea behind this is that we do not need to precondition the magnetization, because the origin of the magnetism, the exchange interaction, is essentially short ranged.

D. Scheme summary

Algorithm I summarizes the preconditioning scheme for the simulation of nonmagnetic and collinear-magnetic systems. The main steps are the computation of the Yukawa potential and assembly of the preconditioned residual charge density.

Algorithm II displays the embedding of the preconditioning scheme into the SCF iteration. The main difference to a nonpreconditioned SCF iteration is the additional preconditioning step prior to the mixing. The mixing step only changes insofar as the set of preconditioned residual charge densities obtained in the current and previous cycles substitutes the set of residual charge densities in the call of the mixing function.

E. Adjusting the charge

We would like to point out that the preconditioning of the SCF iteration by (7) and (8) in general prohibits any adjustment of the total charge that may have been calculated during a cycle, i.e.,

$$\int_\Omega \delta\rho_\lambda = 0, \tag{14}$$

even if

$$\int_\Omega \delta\rho \neq 0. \tag{15}$$

In the plane-wave formulation the change in charge is given by the volume of the unit cell times $\delta\rho_\lambda(\boldsymbol{K} = \boldsymbol{0})$, which is zero by construction. Alternatively, it can be proven directly for the basis-independent formulation by applying the divergence theorem and the periodic boundary conditions,

$$\int_\Omega \Delta\delta V_\lambda d\boldsymbol{r} = \int_\Omega \mathrm{div}(\nabla\delta V_\lambda) d\boldsymbol{r} = \int_{\partial\Omega} \nabla\delta V_\lambda \cdot \mathbf{n} d\omega = 0. \tag{16}$$

Then integration of the differential equation (7) leads to

$$\int_\Omega \delta V_\lambda = \frac{4\pi}{\lambda^2} \int_\Omega \delta\rho, \tag{17}$$

which, inserted into (8), yields

$$\int_\Omega \delta\rho_\lambda = \int_\Omega \delta\rho - \int_\Omega \delta\rho = 0. \tag{18}$$

ALGORITHM II. Preconditioned self-consistent-field iteration.

Input: starting charge density $\rho^{(0)}$, preconditioning parameter λ and mixing parameter α.
Output: ground-state charge density ρ^*.
Set input charge density $\rho = \rho^{(0)}$.
Compute output charge density $f(\rho)$. This is a complex step including the construction of an exchange-correlation potential and the solution of the Poisson and Kohn-Sham equations.
while distance$(f(\rho), \rho) > 10^{-6}$ **do**
  Compute the residual $\delta\rho = f(\rho) - \rho$.
  Compute the preconditioned residual $\delta\rho_\lambda \leftarrow$ Kerker Preconditioner$(\delta\rho)$. ▷ Alg. I
  Mix new input charge density $\rho \leftarrow \mathrm{Mix}(\rho, \delta\rho_\lambda, \alpha)$.
  Compute output charge density $f(\rho)$.
end while
The ground-state charge density is $\rho^* = f(\rho)$.

As a consequence, if the starting density does not already integrate to the fully correct total charge, an additional charge correction treatment is necessary.

III. PERFORMANCE OF THE PRECONDITIONER

In this section we address the following three questions: (1) Does our implementation improve the SCF convergence at all? (2) How stable is the SCF convergence with respect to variations of our parameters? (3) Is the SCF convergence, as expected, independent of the system size?

In order to answer the questions we run a number of computational tests on the JURECA [43] commodity cluster located at the Jülich Supercomputing Centre. Each node of this cluster is equipped with two Xeon E5-2680 v3 Haswell CPUs (24 cores) over a Fat-Tree EDR Infiniband interconnect, and 128 GB of memory. Our implementation of the preconditioner in the FLEUR [31] code features MPI and OpenMP parallelism over the reciprocal lattice vectors $\boldsymbol{K}$ of the modified multiple moments synthesized from the interstitial charge density $q_L^\alpha[\rho^{\mathrm{I}}]$, the Fourier components of modified pseudocharge density $\bar{\rho}^\alpha(\boldsymbol{K})$, and the synthesis of the terms of the interstitial potential at sphere boundary $V_L^{\mathrm{I}}(R_\alpha;\boldsymbol{\tau}_\alpha)$, which are all calculated within the scope of Alg. 1 from Ref. [33]. As elucidated below, the Kerker mixing scheme is very efficient as it demands only a tiny fraction of the time needed for a self-consistency step.

As a measure of SCF convergence we use the L_2-norm-induced metric,

$$\mathrm{distance}(\rho_{\mathrm{out}},\rho_{\mathrm{in}}) = \left(\frac{1}{|\Omega|}\int_\Omega \|\delta\rho\| d\boldsymbol{r}\right)^{\frac{1}{2}}. \qquad (19)$$

We say that the SCF iteration is converged when the distance has decreased below a tolerance of 10^{-6} a.u.$^{-3}$. We use Anderson mixing with mixing parameter α in all cases. A preconditioning parameter of $\lambda = 0$ stands for no preconditioned charge density.

In order to select test cases that are prone to charge sloshing, we work with metallic systems that have strongly elongated unit cells with large aspect ratios of length to width and that simultaneously exhibit large charge density inhomogeneities. Such unit cells typically occur in the simulation of metal surfaces using the supercell approach. Except for the test cases answering question 3, where system sizes vary, our test cases are supercells of aspect ratio 20:1 consisting of 19 atoms simulating the substrate and surface atoms and one layer of vacant atoms simulating the vacuum. In the Cu unit cell shown in Fig. 1(a) the vacant atom layer simulating a vacuum is in the center. Despite the limited number of atoms, such a system is large enough to already exhibit charge sloshing while the stretch in one dimension reduces computational cost. The noble metal Cu was chosen due to its delocalized s and p electrons besides its localized d electrons. We use local orbitals as chosen automatically (and conservatively) by FLEUR's input generator, which for most of the 19-atom test cases translates to a number of 200 to 300 electrons treated as valence electrons. Switching to the larger systems with 79 atoms this number increases to 869 electrons for Cu and 1264 electrons for Fe.

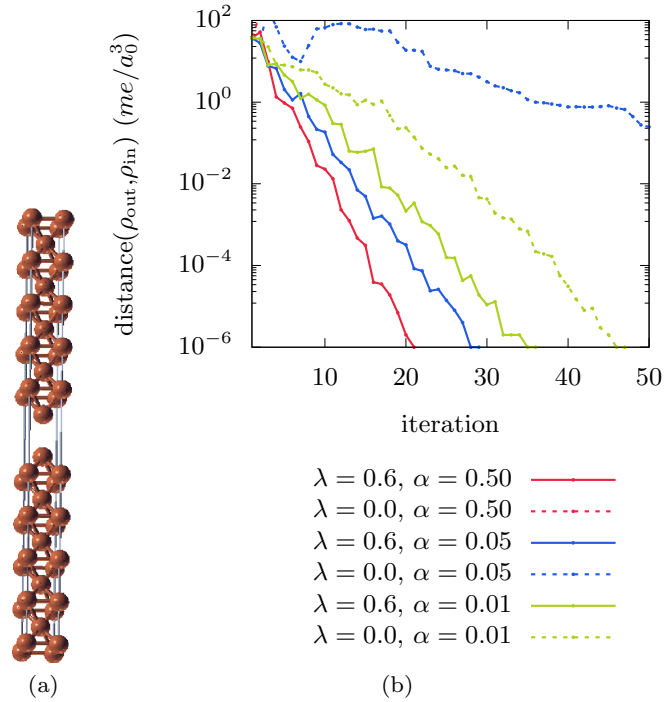

FIG. 1. (a) 19-atom Cu unit cell with empty atom position in the center. (b) The convergence behavior for the 19-atom Cu test system for some parameter combinations.

Without the Kerker preconditioner, our 19-atom Cu test system does not converge for the standard mixing parameter α of 0.05 [see the blue dotted line in Fig. 1(b)]. Convergence can be ensured by reducing the mixing parameter to 0.01. However, with 48 iterations needed, convergence is rather slow (green dotted line). The preconditioner is much more effective. For both $\alpha = 0.05$ and $\alpha = 0.01$, SCF convergence is improved by using a preconditioning parameter of $\lambda = 0.6$. Even better results can be achieved, if higher mixing parameters are used in combination with the preconditioner: For $\alpha = 0.5$, which would quickly diverge without the preconditioner, the SCF iteration is converged in only 22 iterations.

We observe the same effect for other metals. Figure 2 shows the convergence behavior for Fe, Cr, Ca, and Al with (solid line) and without (dashed line) preconditioner. In comparing SCF iteration with and without preconditioner, we always choose parameters leading to the best scenario: We select α and $\lambda > 0$ (preconditioning case) or a parameter α (nonpreconditioning case) that results in the lowest number of iterations needed to achieve convergence. For these and other metals, Table I lists the number of iterations up to convergence for the preconditioning and nonpreconditioning cases, and, additionally, the speedup we obtain by the usage of the preconditioner. For most metals, the preconditioned SCF convergence is 2–3 times as fast as the nonpreconditioned SCF convergence.

We note that the systems described above inhibit z-reflection symmetry and thus the length over which charge can slosh is actually only half of the unit cell. Switching to a nonsymmetric system of equal size therefore amplifies the efficacy of the Kerker preconditioner, as the following

TABLE I. Comparison of the SCF convergence with and without preconditioner for various metallic systems with 19 atoms, and best parameters.

	Best SCF convergence without preconditioner		Best SCF convergence with preconditioner			
	α	No. it.	λ	α	No. it.	speedup
Cu	0.01	48	0.6	0.50	22	2.2
Ag	0.02	37	0.7	0.90	18	2.1
Au	0.009	42	0.5	0.90	19	2.2
Li	0.40	23	0.9	0.90	12	1.9
Na	0.35	22	0.7	0.85	12	1.8
K	0.06	28	0.6	0.95	13	2.2
Ca	0.02	38	0.8	0.95	18	2.1
Sr	0.04	39	0.8	0.95	17	2.3
Sc	0.007	55	1.0	0.90	28	2.0
Ti	0.006	54	0.8	0.60	30	1.8
V	0.009	49	1.3	0.75	28	1.8
Cr	0.02	146	1.6	0.90	47	3.1
Fe	0.009	70	1.4	0.50	35	2.0
Co	0.004	67	0.7	0.05	34	2.0
Ni	0.007	63	0.8	0.30	35	1.8
Zn	0.003	133	0.6	0.05	44	3.0
Al	0.20	28	0.8	0.65	14	2.0

example shows. To break z-reflection symmetry we replace one of the Cu atoms next to the vacancy in the 19-atom Cu system described above by a V atom. While by using the Kerker preconditioner in the symmetric 19-atom Cu and V systems we gained a speedup of about 2, reducing the number of iterations needed for SCF convergence from 48 to 22, and from 49 to 28, respectively, the effect of Kerker preconditioning in the combined nonsymmetric system with 18 Cu atoms and 1 V atom is even larger, with a speedup of about 3, reducing the number of iterations from 75 (for $\alpha = 0.007$ and $\lambda = 0.0$) to 26 (for $\alpha = 0.5$ and $\lambda = 0.9$). It makes sense that the number of iterations needed does not change so much between the symmetric and nonsymmetric systems when the Kerker preconditioner is used, and that it does change when preconditioning is not used, since the preconditioner makes the convergence largely system-size independent and the transition from the symmetric to the nonsymmetric system essentially doubles the length over which charge sloshing can occur.

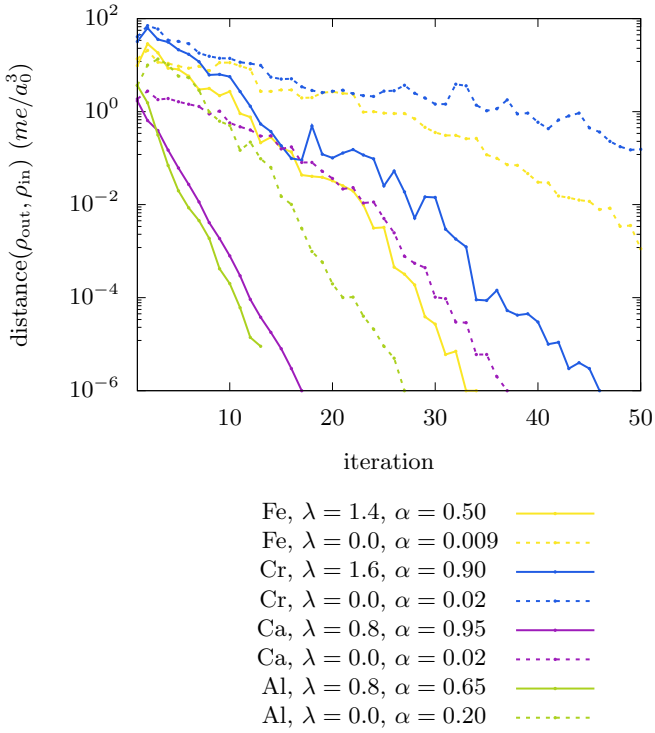

FIG. 2. The best SCF convergence obtained for Fe, Cr, Ca, and Al, with and without preconditioner.

To analyze the dependence of the SCF convergence on the choice of parameters, for the Cu and a similarly constructed 19-atom Fe system, we ran one calculation for each pair of 19 mixing and 13 preconditioning parameters in order to scan the two-dimensional parameter space (α, λ). The heat maps in Figs. 3(a) and 3(b) show the distances after 20 (for Cu) and 35 iterations (for Fe), respectively. Qualitatively the same is shown by the slope of a linear fit of the logarithmically plotted distance in Figs. 3(c) and 3(d). The optimal preconditioning parameter depends on the metal and is $\lambda = 0.6$ for Cu and $\lambda = 1.4$ for Fe. But more importantly, these heat maps reveal a convergence stability with respect to variations in the two parameters α and λ. This is particularly relevant when the optimal parameter combination for a material is not yet known. For instance, in our experience, using $\lambda = 0.8$ improves the SCF convergence of almost any metallic system for almost any mixing parameter, without further knowledge of the system.

If the optimal preconditioning parameter is used, the SCF convergence is expected to be independent of the system size, the reason being that the size dependence cancels out by multiplying J with C in (2). In order to test the validity of this conjecture for real systems, we systematically expanded our

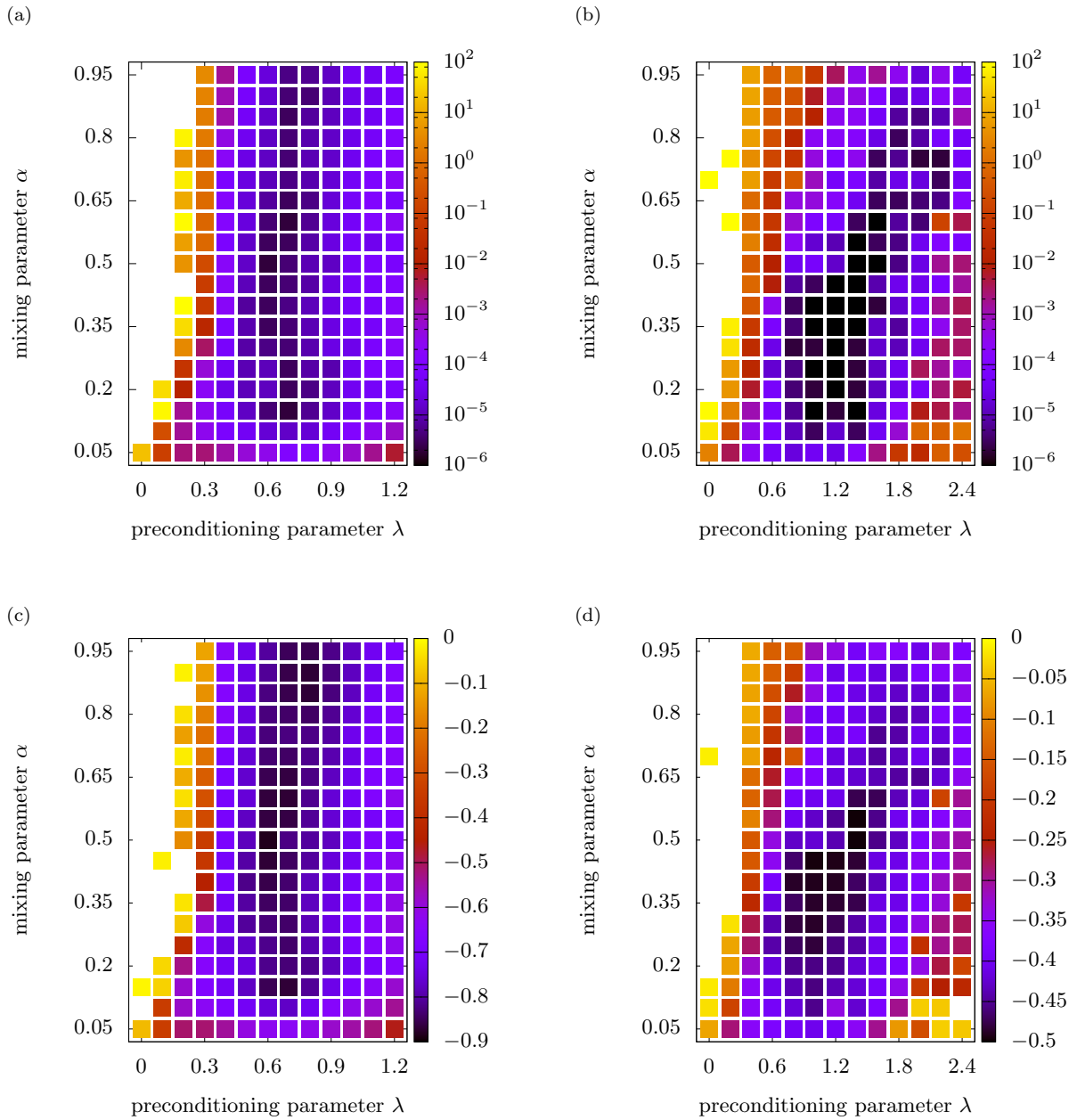

FIG. 3. The distance of ρ_{out} and ρ_{in} in me/a_0^3 for a number of mixing parameters α and a number of Kerker parameters λ for the 19-atom system of (a) Cu after 20 and (b) Fe after 35 iterations. The slope of a linear fit of the logarithmically plotted distance for the same systems of (c) Cu and (d) Fe. The color codes in the slope plots are restricted to values $\leqslant 0$. White spots mean that the iteration did not converge within (a) 20, (b) 35, (c) 30, and (d) 50 iterations, respectively.

19-atom Cu and Fe test systems to 39-, 59- and 79-atom systems. As for the 19-atom system, they all have only one vacant atom defect in the center of the supercell. Size independence is clearly depicted in Fig. 4 for both the Cu and Fe systems: The lines representing the convergence of the systems of varying size almost indistinguishably end in the same number of SCF iterations.

In terms of computational cost, the Kerker preconditioner is quite inexpensive compared to the benefit it delivers. For the 19-atom Cu system in a calculation on 1 MPI rank with 24 threads and with 30 k points, the overhead of the preconditioner over the entire computation is 0.2s per cycle, or equivalently 1.5% of the total time. For a 39-atom system the overhead reduces to 0.33% of the total time. This is mainly due to the linear growth of time spent in the preconditioner compared to the cubic growth of time needed for the rest of the cycle. Thus, the price paid for introducing the preconditioner in the SCF iteration matters less and less the larger the system is, while it delivers a remarkable advantage in terms of the total execution time of the simulation.

IV. A MODEL FOR THE PRECONDITIONING PARAMETER UTILIZING THE DENSITY OF STATES

For a homogeneous electron gas, the Thomas-Fermi approximation to the Kohn-Sham response function [37] [see (4)

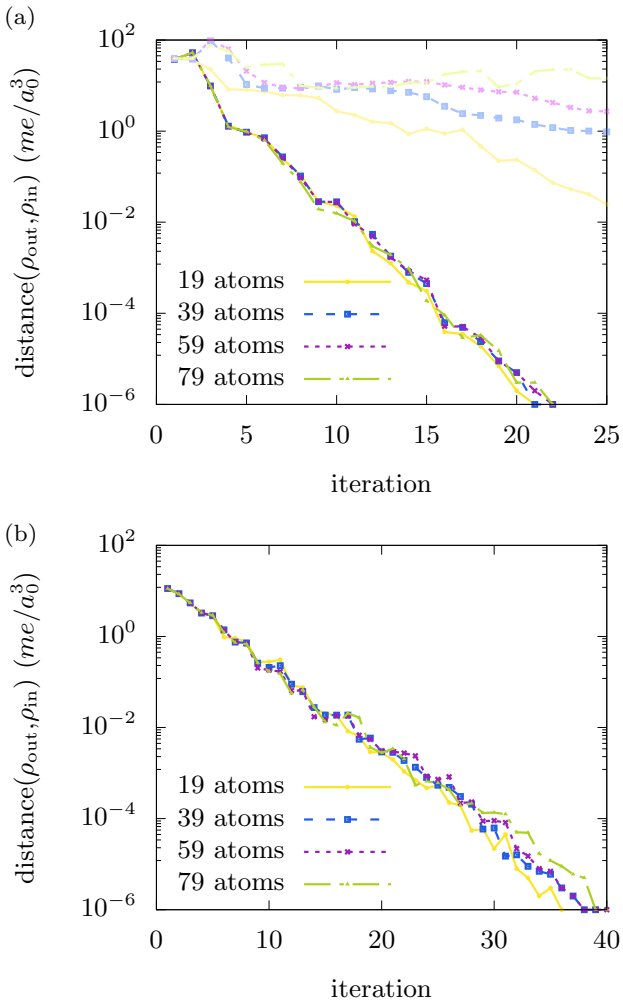

FIG. 4. System-size independence of the SCF convergence for Cu and Fe. (a) Cu for $\alpha = 0.5$ and $\lambda = 0.6$ (dark colors) and for comparison without preconditioner for α values resulting in best SCF convergence (pale colors). (b) Fe for $\alpha = 0.1$ and $\lambda = 1.0$.

and (5)] provides a connection between the preconditioning parameter λ and the density of states z at the Fermi energy E_F,

$$\lambda^2 = 4\pi z(E_F), \tag{20}$$

in Hartree atomic units. The density of states in Eq. (20) is a volume-averaged quantity, making the preconditioning parameter system-size independent. In the following we investigate to which extent this relationship also holds for more complex metallic systems.

To gain a deeper understanding of the relation of the preconditioning parameter λ and the density of states in the context of the homogeneous electron gas, it is worth resorting to an atomiclike angular-momentum-resolved characterization of the valence electron states, e.g., denoted by s, p, etc. Apart from s and p electrons of metals, which due to their delocalization behave similarly to the electron gas, a metallic system may also comprise strongly localized d and f electrons. Since these participate to a lesser extent in the long-range interactions that cause the charge sloshing, we expect that the contribution of the d and f density of states to the preconditioning parameter will be much smaller. We approximate the proportion of the density of states relevant for the preconditioning by the full s and p parts plus strongly reduced d and f fractions. More precisely, we model the preconditioning parameter by

$$\lambda^2 = \frac{4\pi}{V^\Omega}\left(Z_{sp}(E_F) + Z^I_{sp}(E_F) + \alpha Z_{df}(E_F) + \beta Z^I_{df}(E_F)\right), \tag{21}$$

where we now average over the unit-cell volume V^Ω explicitly and where $Z_{sp}(E_F)$, $Z_{df}(E_F)$, $Z^I_{sp}(E_F)$, and $Z^I_{df}(E_F)$ refer to the contributions from the s and p electrons and d and f electrons, respectively, to the local density of states integrated over the muffin-tin and interstitial (I) regions, respectively, having the space partitioning of the FLAPW method in mind. Since the notion of s, p, d, and f electrons as well as the corresponding density of states are properly defined only in the muffin-tin sphere, we estimate their contributions, $Z^I_{sp}(E_F)$ and $Z^I_{df}(E_F)$, to the interstitial density of states based on the easily obtainable total density of states of the unit cell, $Z_{total}(E_F)$, the orbital-projected local densities of states, $Z_{sp}(E_F)$ and $Z_{df}(E_F)$, as well as the volumes of the unit cell V^Ω and the MT region V^{MT}. We estimate the contribution of the delocalized s and p electrons to the density of states in the interstitial volume, $V^I = V^\Omega - V^{MT}$, by making the reasonable assumption,

$$\frac{Z^I_{sp}(E_F)}{V^I} = \frac{Z_{sp}(E_F)}{V^{MT}}, \tag{22}$$

of a homogeneous density of states across the unit cell and thus, their total contribution to the parameter by

$$\frac{Z^I_{sp}(E_F) + Z_{sp}(E_F)}{V^\Omega} = \frac{Z_{sp}(E_F)}{V^{MT}}. \tag{23}$$

In the model proposed, the strongly localized d and f electrons contribute with $\alpha Z_{df}(E_F) + \beta Z^I_{df}(E_F)$, where we expect the interstitial contribution to have a stronger weight than the muffin-tin contribution, i.e., $0 \leqslant \alpha \leqslant \beta \leqslant 1$, and where the interstitial contribution is obtained by subtracting all other contributions from the total density of states $Z_{total}(E_F)$,

$$Z^I_{df}(E_F) = Z_{total}(E_F) - Z_{df}(E_F) - Z_{sp}(E_F) - Z^I_{sp}(E_F). \tag{24}$$

Inserting (24) and (22) into (21), we finally obtain as a model for the preconditioning parameter,

$$\lambda^2 = \frac{4\pi}{V^\Omega}\Bigg((1-\beta)\frac{V^\Omega}{V^{MT}}Z_{sp}(E_F) + (\alpha-\beta)Z_{df}(E_F) + \beta Z_{total}(E_F)\Bigg), \tag{25}$$

containing solely quantities easy to read off during the SCF cycle, whereas α and β are adjustable parameters. In principle, optimal material-class specific parameters can be determined.

Figure 5 shows, for the 17 different metals listed in Table I, the comparison of the computationally determined optimal preconditioning parameter (data points indicated as circles) versus the model parameter for $\alpha = 0.05$ and $\beta = 0.5$. We already observed in Sec. III that there is a wide range of preconditioning parameters λ, for which the convergence is substantially improved, meaning that even if the model parameter is far away from the computationally optimal parameter,

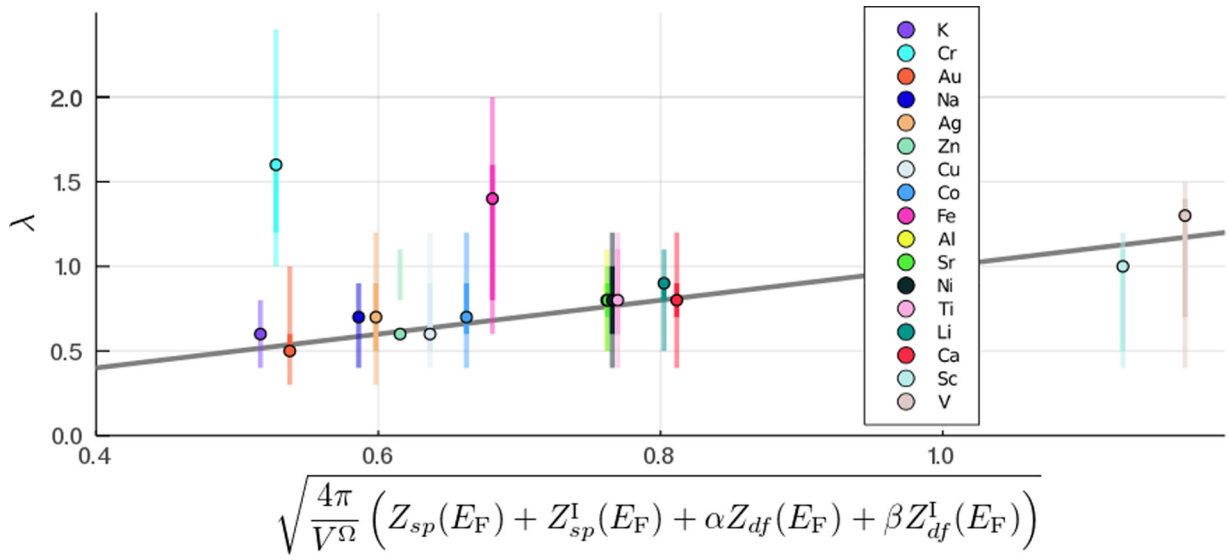

FIG. 5. Comparison of optimal computational preconditioning parameter λ with the parameter determined according to the (density-of-states dependent) model (25) for $\alpha = 0.05$ and $\beta = 0.5$ for a diverse number of metals. The circles indicate the optimal computational λ according to Table I. The short vertical lines show parameter ranges within which the convergence deviates from the optimal convergence by at most 10% (dark lines) or 25% (light lines). The gray diagonal line is the identity and indicates a match of computational and model parameter. The elements in the legend are listed in increasing order of model parameters.

it could be that the model parameter is still sufficiently good in a sense that is still to be specified. Therefore, in addition to the data points, Fig. 5 shows ranges of parameters with acceptable convergence. In particular, the dark (light) lines depict ranges of preconditioning parameters such that, for at least three mixing parameters from the set $\{0.05, 0.10, \ldots, 0.95\}$, the computationally optimal convergence indicated by the circle is at most 10% (25%) faster. For all of the tested metals, except Cr and Zn, the model parameter lies within the 25% range of the computational optimum, and in more than half of the cases even in the 10% range. Nine out of 17 metals find the computationally optimal preconditioning parameter on the identity line (gray line) indicating a match of computational and model parameters.

In order to understand the contribution of the d and f electrons to the accuracy of the model (21) to predict an optimal preconditioning parameter, we present in Fig. 6 a comparison of the computationally determined parameter λ with the model neglecting the contribution of the d and f electrons ($\alpha = \beta = 0$),

$$\lambda^2 = \frac{4\pi}{V^{\mathrm{MT}}} Z_{sp}(E_{\mathrm{F}}). \tag{26}$$

We find only one metal on the identity line, six are in the 10% range, 15 are in the 25% range, and two metals (Cr and Fe) are outside these ranges. In the simpler model, far fewer model parameters coincide with the computationally optimal ones, but since most of them lie in the 10% range, the average increase in the number of iterations amounts to just 13%. Considering that by preconditioning we gain a convergence speed-up of 2–3, we find that the simplified model in which just the local density of states of the s and p electrons in the muffin-tin sphere enters, is still very powerful.

The comparison of the computational experiment and model supports the hypothesis that the charge sloshing problem arises mainly from the response of the s and p electrons and only in part from the response of the stronger localized d and f electrons. Our model not only sheds some light on the origin of the parameter values—why they are material dependent and how they are connected with the charge distribution on the energy scale—it can also be used to automatize the choice of parameter for a metal whose density of states information is readily accessible.

We would like to note that although the Kerker preconditioning scheme is intended for use in metallic systems only, the simple models relating the density of states to the preconditioning parameter is still valid for insulating and semiconducting systems. Not using the preconditioner corresponds to choosing $\lambda = 0$ in (8) and (9), and this is exactly the value we would obtain from the model, since the density of states at the Fermi energy is zero for insulators and semiconductors.

V. CONCLUSION

We developed a Kerker mixing scheme for the SCF iteration of all-electron electronic structure methods based on density functional theory that make use of muffin-tin partitioning of the unit cell applicable to large bulk metallic systems that are either nonmagnetic or collinear magnetic. The Kerker mixing scheme is composed of the Kerker preconditioner and the use of quasi-Newton methods. The scheme requires the determination of the screened Coulomb or Yukawa potential, which involves the adaptation of Weinert's pseudocharge density method for the solution of the Poisson equation such that it provides a solution to the modified Helmholtz equation, as well as provisions enabling minor adjustments of the total

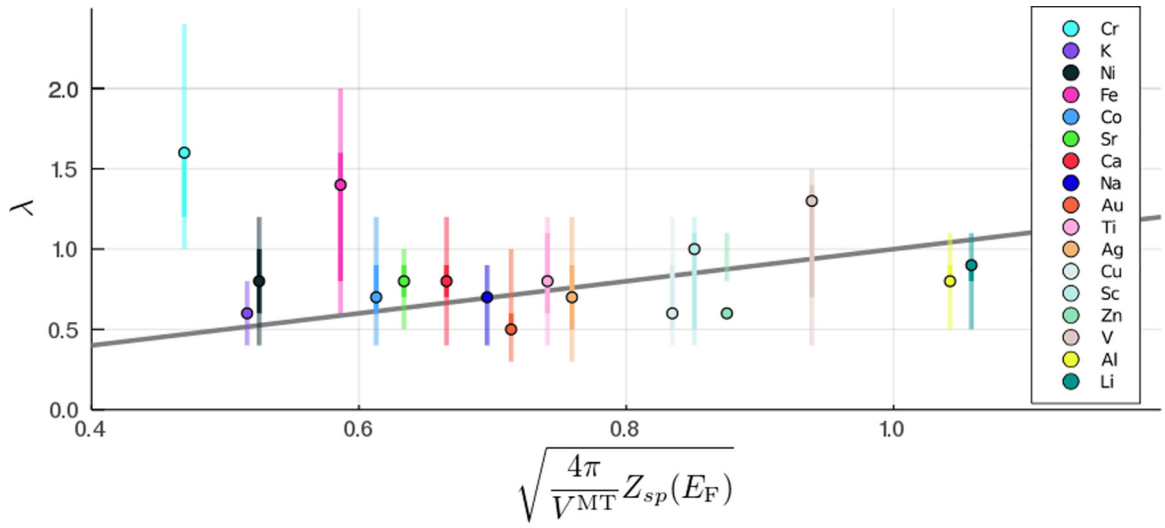

FIG. 6. As Fig. 5, but for parameters $\alpha = 0$ and $\beta = 0$.

charge throughout the SCF iteration. The method is implemented into the FLEUR code, a realization of the FLAPW method.

Tests confirmed that the Kerker preconditioner significantly improves the SCF convergence for the above-mentioned class of systems. One of the crucial characteristics of the preconditioner is the system-size independence of the SCF convergence. Moreover, the preconditioner also scales linearly with the system size in all-electron methods, is computationally lightweight, and starts to be effective even in relatively small systems.

Although the application of the preconditioner introduces the additional preconditioning parameter λ, it does not complicate the choice of parameters. In fact, the application of preconditioning increases the radius of convergence with respect to the mixing parameter and reduces the risk of SCF divergence posed by choosing a too large mixing parameter. Furthermore, the SCF convergence is not particularly sensitive to the choice of the preconditioning parameter λ. Our analysis shows that a close-to-optimal model preconditioning parameter can even be computed beforehand or on the fly, given the orbital-projected density of states in the MT spheres and the total density of states in the unit cell at the Fermi level. A similar approach could be pursued in plane-wave methods by calculation of the orbital-projected density of states in a sphere around the atoms. As a consequence, our model can be used to automatize the parameter selection by direct calculation, which can be a very valuable asset for high-throughput computing. Our experimentally validated model supports the hypothesis that the charge sloshing problem arises mainly from the response of the delocalized electrons. These are the *s* and *p* electrons in simple and transition metals, and the *s*, *p*, and *d* electrons in rare earth and actinide systems. The stronger localized *d* and *f* electrons only account for a small portion of the charge sloshing problem and are generally relevant for the preconditioning only due to their large number in the range of the Fermi energy.

The Kerker preconditioner conserves the total charge and thus does not contribute to numerical convergence of the total charge neutrality throughout the SCF cycle. Thus, the starting density should be charge neutral to high numerical precision, or provisions should be in place enabling minor adjustments of the total charge throughout the SCF iteration.

Although the implementation and tests have been performed in the context of the FLAPW method, both readily transfer to all other all-electron methods, making use of the muffin-tin partitioning of space. Since muffin-tin spheres of methods with volume-filling spheres (e.g., KKR-GF, ASW, LMTO) are larger than the ones in the FLAPW method, we speculate that the simplified model (26) for choosing the preconditioning parameter might even be more effective for those methods than for the FLAPW method. The models introduced to choose the preconditioning parameter can be applied directly to all other implementations, making use of the Kerker preconditioning scheme, e.g., pseudopotential or PAW methods.

We conclude that our formulation of the Kerker preconditioner establishes a new state of the art for the simulation of large bulk metallic systems in all-electron methods. A generalization of the Kerker mixing scheme in all-electron methods for systems with strongly varying charge densities such as metallic heterostructures containing interfaces with insulators or systems with surfaces is still pending.

ACKNOWLEDGMENTS

We would like to thank Gregor Michalicek, Jan Winkelmann, and Rudolf Zeller for fruitful discussions and help with conceptual and computational matters. We would like to thank Tatsuya Shishido for fruitful discussions and making us aware of Ref. [29]. This work has been supported by a JARA-HPC seed-fund project and by the program H2020-INFRAEDI-2018-1 of the European Union (Grant No. 824143, project "MaX-materials at the exascale"). The authors gratefully acknowledge the computing time granted by the JARA-HPC Vergabegremium and provided on the JARA-HPC partition's part of the supercomputer JURECA [43] at Forschungszentrum Jülich (DE).

[1] H. Akai and P. H. Dederichs, J. Phys. C **18**, 2455 (1985).
[2] C. G. Broyden, Math. Comput. **19**, 577 (1965).
[3] D. G. Anderson, J. ACM (JACM) **12**, 547 (1965).
[4] P. Pulay, Chem. Phys. Lett. **73**, 393 (1980).
[5] P. Pulay, J. Comput. Chem. **3**, 556 (2004).
[6] G. P. Srivastava, J. Phys. A: Math. Gen. **17**, 2737 (1984).
[7] D. Vanderbilt and S. G. Louie, Phys. Rev. B **30**, 6118 (1984).
[8] D. Singh, H. Krakauer, and C. S. Wang, Phys. Rev. B **34**, 8391 (1986).
[9] D. D. Johnson, Phys. Rev. B **38**, 12807 (1988).
[10] G. Kresse and J. Furthmüller, Comput. Mater. Sci. **6**, 15 (1996).
[11] G. Kresse and J. Furthmüller, Phys. Rev. B **54**, 11169 (1996).
[12] V. Eyert, J. Comput. Phys. **124**, 271 (1996).
[13] L. D. Marks and D. R. Luke, Phys. Rev. B **78**, 075114 (2008).
[14] K. Baarman, T. Eirola, and V. Havu, J. Chem. Phys. **134**, 134109 (2011).
[15] L. Lin and C. Yang, SIAM J. Sci. Comput. **35**, S277 (2013).
[16] K.-M. Ho, J. Ihm, and J. D. Joannopoulos, Phys. Rev. B **25**, 4260 (1982).
[17] J. Auer and E. Krotscheck, Comput. Phys. Commun. **118**, 139 (1999).
[18] J. Auer and E. Krotscheck, Comput. Phys. Commun. **151**, 265 (2003).
[19] A. Sawamura and M. Kohyama, Mater. Trans. **45**, 1422 (2004).
[20] P.-M. Anglade and X. Gonze, Phys. Rev. B **78**, 045126 (2008).
[21] G. P. Kerker, Phys. Rev. B **23**, 3082 (1981).
[22] W. E. Pickett, Comput. Phys. Rep. **9**, 115 (1989).
[23] P. E. Blöchl, Phys. Rev. B **50**, 17953 (1994).
[24] G. Kresse and D. Joubert, Phys. Rev. B **59**, 1758 (1999).
[25] V. Eyert, in *The Augmented Spherical Wave Method: A Comprehensive Treatment*, Lecture Notes in Physics (Springer, Berlin/Heidelberg, 2013), pp. 113–172.
[26] D. Pashov, S. Acharya, W. R. Lambrecht, J. Jackson, K. D. Belashchenko, A. Chantis, F. Jamet, and M. van Schilfgaarde, Comput. Phys. Commun. **249**, 107065 (2020).
[27] N. Papanikolaou, R. Zeller, and P. H. Dederichs, J. Phys.: Condens. Matter **14**, 2799 (2002).
[28] E. Wimmer, H. Krakauer, M. Weinert, and A. J. Freeman, Phys. Rev. B **24**, 864 (1981).
[29] T. Shishido and M. Weinert, http://meetings.aps.org/link/BAPS.2015.MAR.B24.10.
[30] Y. Shiihara, O. Kuwazuru, and N. Yoshikawa, Modell. Simul. Mater. Sci. Eng. **16**, 035004 (2008).
[31] FLEUR: The Jülich FLAPW Code Family, Computer code FLEUR, MaX-Release 3.1, http://www.flapw.de/.
[32] M. Weinert, J. Math. Phys. **22**, 2433 (1981).
[33] M. Hinzen, E. Di Napoli, D. Wortmann, and S. Blügel, arXiv:2003.03370.
[34] P. H. Dederichs and R. Zeller, Phys. Rev. B **28**, 5462 (1983).
[35] J. Lindhard, Kgl. Danske Videnskab. Selskab Mat.-fys. Medd. **28** (1954), http://gymarkiv.sdu.dk/MFM/kdvs/mfm%2020-29/mfm-28-8.pdf.
[36] M. Manninen, R. Nieminen, P. Hautojärvi, and J. Arponen, Phys. Rev. B **12**, 4012 (1975).
[37] N. Ashcroft and N. Mermin, *Solid State Physics* (Saunders College, Philadelphia, 1976).
[38] For an overview of the different quasi-Newton methods expressed in terms of rank-one updates, $u_j \otimes v_j$, and generalizations to higher-rank updates, we refer to Ref. [12].
[39] H. Yukawa, Proc. Physico-Math.l Soc. Japan. 3rd Ser. **17**, 48 (1935).
[40] F. Tran and P. Blaha, Phys. Rev. B **83**, 235118 (2011).
[41] M. Abramowitz and I. A. Stegun, *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*, 9th Dover printing, 10th GPO printing ed. (Dover, New York, 1964).
[42] G. Arfken, H. Weber, and F. Harris, *Mathematical Methods for Physicists: A Comprehensive Guide* (Elsevier Science, New York, 2013).
[43] D. Krause and P. Thörnig, Jülich Supercomputing Centre, Journal of Large-scale Research Facilities **4**, A132 (2018).