Gregory Beylkin,[1,a)] Martin J. Mohlenkamp,[2,b)] and Fernando Pérez[1,c)]
[1] *A , C , B , 256 CB, B , C 80309-0526, A*
[2] *, , 321 .., A , 45701, A*

(Received 25 July 2007; accepted 25 January 2008; published online 14 March 2008)

The wavefunction for the multiparticle Schrödinger equation is a function of many variables and satisfies an antisymmetry condition, so it is natural to approximate it as a sum of Slater determinants. Many current methods do so, but they impose additional structural constraints on the determinants, such as orthogonality between

that approximates the wavefunction as a linear combination of a reference state and excited states, you could not learn that the wavefunction is better approximated as a linear combination of several nonorthogonal, near-reference states. Thus, the choice of numerical method is not just a computational issue; it can help or hinder our understanding of the wavefunction.

For these reasons, our goal is to construct an adaptive numerical method without imposing

$$\mathcal{H} = \mathcal{T} + \mathcal{V} + \mathcal{W} = -\frac{1}{2}\sum_{=1} \Delta \; + \sum_{=1} v(\mathbf{r}) + \frac{1}{2}\sum_{=1} \sum_{\neq} \frac{1}{\|\mathbf{r} - \mathbf{r}\|}, \tag{1}$$

where Δ is the three-dimensional Laplacian acting in0.5526 0 TD(−)Tj0.8733 0.7398 TD(1)T8uariabler

1. W ar

In this work, we construct and demonstrate a method that also uses a wavefunction of the form (5) but without constraints on the ϕ. We remove both structural constraints, such as an excitation pattern or orthogonality between single-electron functions, and representation constraints, such as those imposed by using a predetermined basis set.

Many methods (e.g., Refs. 56, 48, 40, 2, 21, 18, 20, 3, 17, 61, and 42) have loosened the constraints on the Slater determinants in one way or another, often with encouraging results. These works, however, only partially removed the constraints, and so, we claim, did not achieve the full

importantly, in Sec. II B 2, we modify the iteration to preserve our wavefunction representation (5). Second, we approximate the Green's function using Gaussians to some fixed but arbitrary accuracy in operator norm (Sec. V A). Third, we delegate the task of representing and operating on functions of the single-electron variable γ to an adaptive numerical method (see Ref. 5) with its own accuracy control. We note that an alternative to Ref. 5 may be used for this purpose as long as it is adaptive and has controlled accuracy. In this paper, we do not address some technical issues, such as finding a good initial guess or determining stopping criteria.

Define the Green's function

$$\mathcal{G}_\mu = (\mathcal{T} - \mu\mathcal{I})^{-1} \tag{7}$$

for $\mu < 0$ and consider the Lippmann–Schwinger integral equation

$$\lambda_\mu \psi_\mu = -\mathcal{G}_\mu[(\mathcal{V}+\mathcal{W})\psi_\mu]. \tag{8}$$

The subscript μ on λ_μ and ψ_μ are to emphasize the dependence of the eigenvalues and eigenfunctions on μ. The operator $\mathcal{G}_\mu[(\mathcal{V}+\mathcal{W})]$ is bounded (see Refs. 35, 55, and 53) but without additional assumptions, it is not compact (see Ref. 54 Sec. XIII 5). Since we are interested in a bound state, we assume that the operators act on a bounded domain, as is justified by the exponential decay of the wavefunction.[1] Under this assumption, $\mathcal{G}_\mu[(\mathcal{V}+\mathcal{W})]$ is compact, so (8) has only a discrete spectrum.

If $\mu =$, then there is an eigenvalue $\lambda_\mu = 1$ and the corresponding eigenfunction ψ_μ of (8) is the desired ground-state eigenfunction of (2), as one can see by rearranging (8) into (2). One can show that $\lambda_\mu = 1$ is the largest eigenvalue (see Ref. 45), so a simple iteration such as the power method yields the desired ground-state eigenfunction.

The eigenvalues λ_μ depend analytically on μ, so when μ is sufficiently close to , the power method will still yield an eigenfunction of (8) with energy near the minimum of (4). The convergence rate of the power method to produce ψ_μ and λ_μ is linear, and depends, as usual, on the gap between the eigenvalues in (8). From ψ_μ and λ_μ, one can then compute an improved estimate μ for . In the practical use of this approach, one does not wait for the power method to converge at each step but instead intertwines it with the update of μ. Beginning with an approximation to the energy $\mu_0 \approx$ and an approximate wavefunction ψ_0, one converts (8) to an iteration

$$\hat{\psi} = -\mathcal{G}_\mu[(\mathcal{V}+\mathcal{W})\psi]. \tag{9}$$

After each iteration, one normalizes by setting

$$\psi_{+1} = \hat{\psi} / \|\hat{\psi}\|. \tag{10}$$

Following the approach of Ref. 29, we can use the update rule

$$\mu_{+1} = \mu - \langle (\mathcal{V}+\mathcal{W})\psi , \psi - \hat{\psi} \rangle / \|\hat{\psi}\|^2. \tag{11}$$

1: An eigenfunction of (8) is an eigenfunction of (2) only when μ is an eigenvalue of (2) and the eigenfunction of (8) has eigenvalue $\lambda = 1$. To obtain the next eigenfunction of (2) above the ground state, it appears that one needs to compute the first two eigenfunctions of (8), the first with eigenvalue $\lambda > 1$ and the second with eigenvalue $\lambda = 1$. We will develop such deflation procedure elsewhere.

2. A a d a a

We restrict the method to approximate wavefunctions of the form (5), with fixed, by replacing the definition of $\hat{\psi}$ in (9). We would like to redefine $\hat{\psi}$ to be the function of the form (5) that minimizes the (least-squares) error

$$\left\|\widehat{\psi} - (-\mathcal{G}_{\mu}[(\mathcal{V}+\mathcal{W})\psi])\right\|. \tag{12}$$

In general, such minimization problems may be ill posed in the sense that the infimum may occur at a limit point (see Ref. 16 and the references therein). In Ref. 6, we describe a method to balance the least-squares error with the loss-of-precision error due to a large condition number, so that the problem becomes well posed. However, even then, there are no known algorithms that assure convergence to the global minimum. Instead, we settle for an algorithm that, at each iteration, constructs a $\widehat{\psi}$ with lower value of (12), unless it is already at a minimum.

Since fitting using (12) instead of directly using (9) introduces an error, the update rule (11) may no longer give a quadratic convergence and, in any case, is not expected to converge to the true energy. One may choose to replace the update rule (11) with the more robust but slower converging rule

$$\mu_{+1} = \frac{\langle \mathcal{H}\psi_{+1}, \psi_{+1}\rangle}{\|\psi_{+1}\|^2}, \tag{13}$$

which is based on (4). Other rules may be possible as well. At present, we do not have enough numerical experience to decide which rule to prefer.

The Green's function iteration itself does not enforce the antisymmetry condition. In order to assure convergence to an antisymmetric solution, we use the pseudonorm induced by the pseudo-inner-product $\langle \cdot, \cdot \rangle_{\mathcal{A}} = \langle \mathcal{A}(\cdot$

$$A(\ ,\ ) = \langle\ ,\ \rangle \text{ and } \quad (\ ) = \langle\ ,\ \rangle, \tag{16}$$

- r, the separation rank used in (5).
- , the cost to represent a function of γ, as discussed in Sec. II B 2.
- $_{\mathcal{P}}$, the cost to convolve a function of γ with the Poisson kernel $1/\|\mathbf{r}\|$. A Fourier-based Poisson solver on a uniform grid would achieve log . For adaptive methods such as that we used, $_{\mathcal{P}}$ may be considered proportional to log as well, but since the complexity depends on the type of function to which the operator is applied, the estimate of the cost is complicated (see the discussion in Refs. 5 and 19).
- , the number of Green's function iterations, as discussed in Sec. II B 1. If we used more than one alternating least-squares pass (Sec. II B 2) per iteration, then would include a factor of the number of passes.
- , the number of conjugate gradient iterations used to solve the system in Sec. II B 2. Although in theory could be as many as the number of degrees of freedom , we generally have a very good starting point and so expect only a very small number to be needed.
- , the number of terms used to approximate the Green's function to relative error ϵ with Gaussians. In Sec. V A, we prove that $=\mathcal{O}((\ln \epsilon)^2)$ independent of μ and .

In terms of these parameters, the cost to store the representation (5) is

$$\mathcal{O}(\quad) \tag{19}$$

and the computational cost to perform the algorithm is

$$\mathcal{O}(\quad ^2 \quad ^2[\quad (\quad + \quad _{\mathcal{P}}) + \quad (\quad + \quad)]). \tag{20}$$

For comparison, the cost to evaluate a single antisymmetric inner product via Löwdin's rules is $\mathcal{O}(\quad ^2(\quad + \quad))$.

C. F

We have implemented the method developed here and tested it sufficiently to verify the correctness of the algorithm as presented. The numerical results are too preliminary to allow us to make any particular claims at this point, however, so we will present them separately. The linear algebra accelerations based on Appendix B have not yet been implemented.

We develop the method in terms of the total variable γ without specifying the spin states. If a specific spin state is imposed on our initial trial wavefunction ψ_0, the iteration will preserve this state.

The representation (5) does not account for the interelectron cusp (see, e.g., Refs. 57, 47, 37, 50, 51, 36, and 27), and thus we cannot hope to achieve a small error ϵ in the wavefunction with small r. As with CI methods, we may still be able to achieve a small error in the energy difference of two systems, which is often the quantity of interest in chemistry. For the current work, we fix r and adapt $\phi(\gamma)$ and $_{l,}$ to minimize the error ϵ rather than fix ϵ and adaptively determine r. We are developing an extension to (5) that incorporates the cusp and hope to achieve a small error ϵ through it.

Similarly, (5) is not size consistent/extensive. For example, if one applied it to a long line of identical, noninteracting subsystems, then r is expected to grow exponentially in the number of subsystems. We are developing a hierarchical extension to (5) suitable for such extended systems and hope to achieve linear scaling through it.

Although we have focused on the multiparticle Schrödinger equation, the tools that we have developed are another step toward general-purpose, automatically adaptive methods for solving high-dimensional problems.

III. AN I MME C INNE D C

In this section, we develop methods for computing antisymmetric inner products involving $\mathcal{W}$, $\mathcal{V}$, and $\mathcal{T}$. For this purpose, after setting notation, we develop methods for computing with low-rank perturbations of matrices, review the antisymmetry constraint, and define a notion of maximum coincidence. With these, tools we then derive the main formulas.

A. N

We denote a column with suppressed indices by $\mathbf{F}$ and with explicit indices by ( ). We denote its conjugate transpose by $\mathbf{F}^*$. We use $\mathbf{e}$ to denote the column vector that is 1 in coordinate and zero otherwise. A linear is written $\mathcal{L}$. We denote a with suppressed indices by $\mathbb{L}$ and with explicit indices by (,). Recalling that $\mathbf{r}=(, ,) \in \mathbf{R}^3$, we combine spatial integration with summation over spins and define the integral

$$\int ( ) \quad = \sum_{\sigma \in}$$

$$\begin{vmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{vmatrix} = |\mathbb{A}||\mathbb{D} - \mathbb{C}\mathbb{A}^{-1}\mathbb{B}|. \tag{25}$$

Proof: (See, e.g., Ref. 52) It is easy to verify directly that

$$\begin{bmatrix} \mathbb{A} & \mathbb{B} \\ \mathbb{C} & \mathbb{D} \end{bmatrix} = \begin{bmatrix} \mathbb{I} & 0 \\ \mathbb{C}\mathbb{A}^{-1} & \mathbb{I} \end{bmatrix} \begin{bmatrix} \mathbb{A} & 0 \\ 0 & \mathbb{D} - \mathbb{C}\mathbb{A}^{-1}\mathbb{B} \end{bmatrix} \begin{bmatrix} \mathbb{I} & \mathbb{A}^{-1}\mathbb{B} \\ 0 & \mathbb{I} \end{bmatrix}. \tag{26}$$

Since the determinants of the first and third matrices are equal to 1, the determinant of the middle matrix gives the desired result. □

Lemma 3 (*):* $\{\mathbf{u}_\ell\}_{\ell=1}^{L}$ $\{\mathbf{v}_\ell\}_{\ell=1}^{L}$ $\mathbf{u}_\ell \mathbf{v}_\ell^*$ $\mathbf{u}_\ell$ $\mathbf{v}_\ell$,

$$\left| \mathbb{I} + \sum_{\ell=1}^{L} \mathbf{u}_\ell \mathbf{v}_\ell^* \right| = \begin{vmatrix} 1 + \mathbf{v}_1^*\mathbf{u}_1 & \mathbf{v}_1^*\mathbf{u}_2 & \cdots & \mathbf{v}_1^*\mathbf{u}_L \\ \mathbf{v}_2^*\mathbf{u}_1 & 1 + \mathbf{v}_2^*\mathbf{u}_2 & \cdots & \mathbf{v}_2^*\mathbf{u}_L \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{v}_L^*\mathbf{u}_1 & \mathbf{v}_L^*\mathbf{u}_2 & \cdots & 1 + \mathbf{v}_L^*\mathbf{u}_L \end{vmatrix}. \tag{27}$$

$$\mathbb{A}^{\perp} = \sum_{=\ -\ell+1} \mathbf{v}\,\mathbf{u}^{*}. \tag{32}$$

For a modified pseudoinverse, we have the following.

5 (

$$\mathbb{A}^{\ddagger} = \mathbb{A}^{\dagger} + \mathbb{A}^{\perp}. \tag{33}$$

Note that $\mathbb{A}^{\perp}$, and thus $\mathbb{A}^{\ddagger}$

$$\mathbf{\Theta} = \mathbb{L}^{-1}\tilde{\mathbf{\Phi}}, \tag{39}$$

we have

$$\mathcal{A}\Theta = \frac{1}{!}|[(\mathbb{L}^{-1}\tilde{\mathbf{\Phi}})(\gamma_1)\ \cdots\ (\mathbb{L}^{-1}\tilde{\mathbf{\Phi}})(\gamma\ )]| = |\mathbb{L}^{-1}|\frac{1}{!}|[\tilde{\mathbf{\Phi}}(\gamma_1)\ \cdots\ \tilde{\mathbf{\Phi}}(\gamma\ )]| = |\mathbb{L}^{-1}|\mathcal{A}\hat{\Phi}. \tag{40}$$

Thus, the antisymmetrizations of $\hat{\Phi}$ and Θ are the same up to a constant, and we can use Θ instead of $\hat{\Phi}$ in calculations. The advantage of using Θ is that the resulting matrix of inner products $\mathbb{L} = \mathbb{L}(\Theta, \Phi) = \mathbb{I}$; in other words, we have the biorthogonality property $\langle \theta, \phi \rangle =$

$$\frac{1}{2}\frac{|\mathbb{L}|}{!}\sum_{\neq}\int\frac{1}{\|\mathbf{r}-\mathbf{r}'\|}\begin{vmatrix} 1 & \cdots & \phi(\gamma)\theta_1(\gamma) & \cdots & \phi(\gamma')\theta_1(\gamma') & \cdots & 0 \\ & & & & & & \end{vmatrix}$$

$$\frac{1}{2}\frac{|\mathbb{U}||\mathbb{V}^*|\prod_{,}}{!}\int \mathbf{\Phi}^*\sum_{=1}{}_{,}{}^{-1}\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}\mathcal{W}_{\mathcal{P}}\left[\mathbf{\Phi}^*\sum_{=1}{}_{,}{}^{-1}\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}\right] - \mathbf{\Phi}^*\mathcal{W}_{\mathcal{P}}\left[\sum_{=1}{}_{,}{}^{-1}\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}\mathbf{\Phi}^*\right]\sum_{=1}{}_{,}{}^{-1}\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}\ \ \gamma$$
$$= \frac{1}{2}\frac{|\mathbb{U}||\mathbb{V}^*|}{!}\sum_{=1}\sum_{=1}\prod_{\neq,}{}_{,}\int \mathbf{\Phi}^*\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^*\mathbf{v}\ \mathbf{u}^*\tilde{\mathbf{\Phi}}] - \mathbf{\Phi}^*\mathbf{v}\ \mathcal{W}_{\mathcal{P}}[\mathbf{u}^*\tilde{\mathbf{\Phi}}\mathbf{\Phi}^*\mathbf{v}\ ]\mathbf{u}^*\tilde{\mathbf{\Phi}}\ \ \gamma. \tag{49}$$

If $\mathbb{L}$ is singular, then at least one $_{,}$ is zero, and only terms that exclude those from the product in (49) are nonzero. Since we exclude two indices in the product, if more than two $_{,}$ are zero, then the entire inner product is zero. If exactly two are zero, then only one term in the sum survives. If exactly one is zero, then we can simplify from a double to a single sum using symmetry. Recalling the modified pseudoinverse from Definition 5 and sorting the zero $_{,}$ to the beginning for notational convenience, we obtain the following propositions.

8: [illegible] $\mathbb{L}$ [illegible] *2,* [illegible] *(41)* [illegible].

9: [illegible] $\mathbb{L}$ [illegible] *2,* [illegible] *(41)* [illegible]

$$\frac{1}{|\mathbb{L}^{\ddagger}|\ !}\int \mathbf{\Phi}^*\mathbf{v}_1\mathbf{u}_1^*\tilde{\mathbf{\Phi}}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^*\mathbf{v}_2\mathbf{u}_2^*\tilde{\mathbf{\Phi}}] - \mathbf{\Phi}^*\mathbf{v}_1\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^*\mathbf{v}_2\mathbf{u}_1^*\tilde{\mathbf{\Phi}}]\mathbf{u}_2^*\tilde{\mathbf{\Phi}}\ \ \gamma. \tag{50}$$

10: [illegible] $\mathbb{L}$ [illegible] *1,* [illegible] $\mathbf{\Theta} = \mathbb{L}^{\dagger}\tilde{\mathbf{\Phi}}$ [illegible] $\mathbf{\Theta} = \mathbb{L}^{\ddagger}\tilde{\mathbf{\Phi}}$, [illegible] *(41)* [illegible]

$$\frac{1}{|\mathbb{L}^{\ddagger}|\ !}\int \mathbf{\Phi}^*\mathbf{v}_1\mathbf{u}_1^*\tilde{\mathbf{\Phi}}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^*\mathbf{\Theta}] - \mathbf{\Phi}^*\mathbf{v}_1\mathcal{W}_{\mathcal{P}}[\mathbf{u}_1^*\tilde{\mathbf{\Phi}}\mathbf{\Phi}^*]\mathbf{\Theta}\ \ \gamma. \tag{51}$$

In computing (50), constructing $\mathbf{\Phi}^*\mathbf{v}_1$, $\mathbf{\Phi}^*\mathbf{v}_2$, $\mathbf{u}_1^*\tilde{\mathbf{\Phi}}$, and $\mathbf{u}_2^*\tilde{\mathbf{\Phi}}$ costs $\mathcal{O}(\ \ )$, applying $\mathcal{W}_{\mathcal{P}}[\cdot]$ costs $\mathcal{O}(\ _{\mathcal{P}})$ and, finally, the integral in γ costs $\mathcal{O}(\ \ )$. In computing (51), the first term costs $\mathcal{O}(\ \ )$ to form $\mathbf{\Phi}^*\mathbf{\Theta}$, $\mathcal{O}(\ _{\mathcal{P}})$ to apply $\mathcal{W}_{\mathcal{P}}[\cdot]$, and $\mathcal{O}(\ \ )$ to integrate in γ. The second term costs $\mathcal{O}(\ \ )$ to form $\mathbf{u}_1^*\tilde{\mathbf{\Phi}}\mathbf{\Phi}$, $\mathcal{O}(\ \ _{\mathcal{P}})$ to apply $\mathcal{W}_{\mathcal{P}}[\cdot]$, $\mathcal{O}(\ \ )$ to apply $\mathbf{\Theta}$, and $\mathcal{O}(\ \ )$ to integrate in γ. In total, the computational cost for the singular cases is less than the cost of the nonsingular case.

11: [illegible] C [illegible] *2*[illegible] $\mathbf{u}$ $\mathbf{v}$ [illegible] $\mathbf{\Phi}$ [illegible] $\tilde{\mathbf{\Phi}}$.

Proof: We follow the same procedure as we used for the electron-electron operator $\mathcal{W}$ in Sec. III F. Instead of (44), we have the simpler expression

$$\frac{|\mathbb{L}|}{!}\sum \int |\mathbb{I} + ((\mathcal{T}_* + \mathcal{V}_*)[\phi\,](\gamma)\boldsymbol{\Theta}(\gamma) - \mathbf{e}\,)\mathbf{e}^*|\;\; \gamma. \tag{54}$$

Applying Proposition 3, we obtain (53). □

To analyze the computational cost to compute (53), we note that it costs $\mathcal{O}(\quad)$ to apply $(\mathcal{T}_*+\mathcal{V}_*)[\cdot]$. Including the cost for the maximum coincidence transformation, our total cost is thus $\mathcal{O}(\ ^2(\ +\ ))$.

1. T S J I a Ca

We now state the formula when $\mathbb{L}$ is singular. The analysis is similar to that for $\mathcal{W}$ in Sec. III F 1.

13: L , 1, (52) , 1,

$$\frac{1}{|\mathbb{L}^{\ddagger}|\;\;!}\int (\mathcal{T}_* + \mathcal{V}_*)[\boldsymbol{\Phi}^*\mathbf{v}_1]\mathbf{u}_1^*\tilde{\boldsymbol{\Phi}}\;\; \gamma. \tag{55}$$

To compute (55), it costs $\mathcal{O}(\quad)$ to form $\boldsymbol{\Phi}^*\mathbf{v}_1$ and $\mathbf{u}_1^*\tilde{\boldsymbol{\Phi}}$, and $\mathcal{O}(\quad)$ to apply $(\mathcal{T}_*+\mathcal{V}_*)[\cdot]$.

I . INC A ING DEL A F NC I N IN HE AN I MME C INNE D C

In this section, we show how to compute antisymmetric inner products when one of the component functions is replaced by a delta function. For concreteness, we will replace $\hat{\phi}_1(\gamma_1)$ by $\delta(\gamma-\gamma_1)$.

A. L ' $\delta(\gamma\ \gamma_1)$,

The matrix $\mathbb{L}$ from (37) is defined by $(\ ,\ )=\langle\phi\ ,\phi\ \rangle$. If we replace $\hat{\phi}_1(\gamma_1)$ by $\delta(\gamma-\gamma_1)$, then the first row depends on γ and is given by $(1,\ )=\langle\delta(\gamma-\cdot),\phi\ \rangle=\phi\ (\gamma)$. We thus have a matrix that depends on γ,

$$\mathbb{L}(\gamma)=\begin{bmatrix} \phi_1(\gamma) & \phi_2(\gamma) & \cdots & \phi\ (\gamma) \\ \langle\phi_2,\phi_1\rangle & \langle\phi_2,\phi_2\rangle & \cdots & \langle\phi_2,\phi\ \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi\ ,\phi_1\rangle & \langle\phi\ ,\phi_2\rangle & \cdots & \langle\phi\ ,\phi\ \rangle \end{bmatrix}. \tag{56}$$

To compute with $\mathbb{L}(\gamma)$ without resorting to cofactor expansions, we express $\mathbb{L}(\gamma)$ as a rank-1 perturbation of a matrix of numbers. Define

$$\mathbb{E}=\begin{bmatrix} \cdot(1) & \cdot(2) & \cdots & \cdot(\ ) \\ \langle\phi_2,\phi_1\rangle & \langle\phi_2,\phi_2\rangle & \cdots & \langle\phi_2,\phi\ \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle\phi\ ,\phi_1\rangle & \langle\phi\ ,\phi_2\rangle & \cdots & \langle\phi\ ,\phi\ \rangle \end{bmatrix}, \tag{57}$$

where the vector $\mathbf{d}^*$ is chosen to be a unit vector orthogonal to the remaining rows of $\mathbb{E}$. This choice assures that the rank deficiency of $\mathbb{E}$ will be smaller than or equal to the rank deficiency of the matrix with any other first row. It also gives us some convenient properties, namely, $\mathbb{E}\mathbf{d}=\mathbf{e}_1$, $\mathbf{d}^*\mathbb{E}^{\ddagger}=\mathbf{e}_1^*$, $\mathbb{E}^{\ddagger}\mathbf{e}_1=\mathbf{d}$, and $\mathbf{e}_1^*\mathbb{E}=\mathbf{d}^*$, where $\mathbb{E}^{\ddagger}$ is the modified pseudoinverse of Definition 5. It costs $\mathcal{O}(\ ^2\quad)$ to construct $\mathbb{E}$ and $\mathcal{O}(\ ^3)$ to compute $\mathbb{E}^{\ddagger}$ and $|\mathbb{E}|$.

We then have

$$\mathbb{L}(\gamma) = \mathbb{E} + \mathbf{e}_1(\mathbf{\Phi}(\gamma) - \mathbf{d})^* \tag{58}$$

and, with the help of Proposition 3, compute

$$|\mathbb{L}(\gamma)| = |\mathbb{E}||\mathbb{I} + \mathbf{d}(\mathbf{\Phi}(\gamma) - \mathbf{d})^*| = |\mathbb{E}|(1 + (\mathbf{\Phi}(\gamma) - \mathbf{d})^*\mathbf{d}) = |\mathbb{E}|\mathbf{\Phi}(\gamma)^*\mathbf{d}, \tag{59}$$

which yields the following

14:

$$\left\langle \delta(\gamma - \gamma_1)\prod_{=2}\hat{\phi}(\gamma), \prod_{=1}\phi(\gamma) \right\rangle_{\mathcal{A}} = |\mathbb{E}|\mathbf{\Phi}(\gamma)^*\mathbf{d}, \tag{60}$$

$\mathbb{E}$ **d**

15: >1,

$$\langle |\mathbb{E}|\mathbf{\Phi}^*\mathbf{d}, \hat{\phi}\rangle = |\mathbb{E}|\langle \mathbf{\Phi}, \hat{\phi}\rangle^*\mathbf{d} = |\mathbb{E}|\ (\ ,\cdot)^*\mathbf{d} = 0, \tag{61}$$

d $(\ ,\cdot)$, $\mathbb{E}$. *(60)* $\hat{\phi}$ >1. $\mathcal{T}$, $\mathcal{V}$, $\mathcal{W}$

B. A $\delta(\gamma\ \gamma_1)$ $(\mathcal{T}\ /\ \mathcal{V})$,

To compute antisymmetric inner products involving operators, we will modify formulas from Sec. III. The first (trivial) modification is to denote the variable of integration in those formulas by γ', so as not to confuse it with the variable γ in $\delta(\gamma-\gamma_1)$. Next, we replace $|\mathbb{L}|$ with $|\mathbb{L}(\gamma)|$ given by (59). Using (58), we can express

$$\mathbb{L}(\gamma)^{-1} = (\mathbb{E} + \mathbf{e}_1(\mathbf{\Phi}(\gamma) - \mathbf{d})^*)^{-1} = (\ \mathbb{E}(\mathbb{I} + \mathbf{d}(\mathbf{\Phi}(\gamma$$

$$\left\langle \delta(\gamma-\gamma_1)\prod_{=2}\hat{\phi}(\gamma),(\mathcal{T}+\mathcal{V})\prod_{=1}\phi(\gamma)\right\rangle_{\mathcal{A}} \tag{66}$$

$$\frac{|\mathbb{E}|}{!}\Bigg[\mathbf{\Phi}(\gamma)^{*}\Bigg(\mathbf{d}\int(\mathcal{T}_{*}+\mathcal{V}_{*})[\mathbf{\Phi}]^{*}\tilde{\mathbf{\Theta}}\ \ \gamma'-\int(\mathcal{T}_{*}+\mathcal{W}_{*}$$

γ

$$\frac{1}{2}\frac{|\mathbb{E}|}{!}\Bigg[2(\mathbf{\Phi}(\gamma)^{*}\mathbf{d}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^{*}\tilde{\mathbf{\Theta}}](\gamma)-\mathbf{\Phi}(\gamma)^{*}\mathcal{W}_{\mathcal{P}}[\tilde{\mathbf{\Theta}}\mathbf{\Phi}^{*}\mathbf{d}](\gamma))+\mathbf{\Phi}(\gamma)^{*}\Bigg(\mathbf{d}\int\mathbf{\Phi}^{*}\tilde{\mathbf{\Theta}}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^{*}\tilde{\mathbf{\Theta}}]$$

$$-\mathbf{\Phi}^{*}\mathcal{W}_{\mathcal{P}}[\tilde{\mathbf{\Theta}}\mathbf{\Phi}^{*}]\tilde{\mathbf{\Theta}}\ \ \gamma'-2\int\tilde{\mathbf{\Theta}}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^{*}\tilde{\mathbf{\Theta}}]\mathbf{\Phi}^{*}\mathbf{d}-\tilde{\mathbf{\Theta}}\mathbf{\Phi}^{*}\mathcal{W}_{\mathcal{P}}[\tilde{\mathbf{\Theta}}\mathbf{\Phi}^{*}\mathbf{d}]\ \ \gamma'\Bigg)\Bigg], \tag{72}$$

, $\mathcal{O}(\ ^{3}+\ ^{2}\ _{\mathcal{P}})$.

20: $\mathbb{E}$ *1, (71)*

$$\frac{1}{|\mathbb{E}^{\ddagger}|\ !}\Bigg[(\mathbf{\Phi}(\gamma)^{*}\mathbf{d}\mathcal{W}_{\mathcal{P}}[\mathbf{\Phi}^{*}\tilde{\mathbf{v}}_{1}\tilde{\mathbf{u}}_{1}^{*}\tilde{\mathbf{\Phi}}]($$

Substituting $=_{,}/(-\mu)$ for $\mu<0$ into (75) and dividing by $-\mu$, one has

$$\left|\frac{1}{,}-\sum_{,=1}\frac{,}{-\mu}\exp\left(-\frac{\tau_{,}}{-\mu'}\right)\right|<\frac{\epsilon}{-\mu}, \tag{76}$$

valid on the interval $_{,}\in[-\mu,\infty)$. In Fourier coordinates, we can express

$$\mathcal{G}_{\mu}=\frac{1}{2\pi^2\Sigma\xi^2-\mu}, \tag{77}$$

from which we see that $\|\mathcal{G}_{\mu}\|=1/(-\mu)$. Since the denominator is at least $-\mu>0$, we can substitute into (76) and obtain

$$\left|\mathcal{G}_{\mu}-\sum_{,=1}\frac{,}{-\mu}\ ^{-\tau_{,}}\otimes_{=1}\exp\left(-\frac{2\pi^2\tau_{,}}{-\mu}\xi^2\right)\right|<\frac{\epsilon}{-\mu}=\epsilon\|\mathcal{G}_{\mu}\|. \tag{78}$$

Thus, we obtain an approximation of $\mathcal{G}_{\mu}$ with relative error ϵ in norm using terms, with independent of and μ. To construct $\mathcal{G}_{\mu}$ as an integral operator in spatial coordinates, we apply the inverse Fourier transform to obtain

$$\mathcal{G}_{\mu}\approx\sum_{,=1}\otimes_{=1}\mathcal{F}_{\mathbf{r}},, \tag{79}$$

where the convolution operator $\mathcal{F}_{\mathbf{r}}$, which depends implicitly on μ, is defined by

$$\mathcal{F}_{\mathbf{r}}(\gamma_1,\ldots,\gamma\ )=\left(\frac{,}{-\mu\ \tau_{,}}\right)^{1/}\left(\frac{-\mu}{2\pi\tau_{,}}\right)^{3/2}$$
$$\times\int\exp\left(-\frac{-\mu}{2\tau_{,}}\|\mathbf{r}\ -\mathbf{r}'\|^2\right)(\gamma_1,\ldots,\gamma_{-1},(\mathbf{r}',<),\gamma_{+1},\ldots,\gamma\ )\ \mathbf{r}'. \tag{80}$$

This construction has a theoretical value since it has proved the following theorem.

22: $\epsilon>0$, $\mu<0$, , $\mathcal{G}_{\mu}$ ϵ $=\mathcal{O}((\ln\epsilon)^2$ μ .

B. C - b (18)

In order to do a step in the iteration, we need to construct the right-hand side **b** in the normal equations (15) in Sec. II B 2. Since $\mathcal{A}$ is an orthogonal projection, $\mathcal{A}$ and $\mathcal{G}_{\mu}$ commute, and $\mathcal{G}_{\mu}$ is self-adjoint, the entry (18) is equal to

$$(.)(\gamma)=-\hat{}_{,.}\sum_{,}'\left\langle\mathcal{A}\mathcal{G}_{\mu}\delta(\gamma-\gamma_1)\prod_{=2}\hat{\phi}(\gamma),[\mathcal{V}+\mathcal{W}]\prod_{=1}\phi\ (\gamma)\right\rangle. \tag{81}$$

Substituting (79) in for $\mathcal{G}_{\mu}$ and rearranging, we have

$$(.)(\gamma)=-\hat{}_{,.}\sum_{,}'\sum_{,=1}\left\langle\mathcal{A}\mathcal{F}_{\mathbf{r}_1}\delta(\gamma-\gamma_1)\prod_{=2}\mathcal{F}_{\mathbf{r}}\hat{\phi}(\gamma),[\mathcal{V}+\mathcal{W}]\prod_{=1}\phi\ (\gamma)\right\rangle. \tag{82}$$

The computation is of the same form for each value of the indices $.$, , and $,$, so we can consider a single term and suppress the indices.

To evaluate a single term $\langle\mathcal{A}\mathcal{F}_{\mathbf{r}_1}\delta(\gamma-\gamma_1)\Pi_{=2}\mathcal{F}_{\mathbf{r}}\hat{\phi}(\gamma),[\mathcal{V}+\mathcal{W}]\Pi_{=1}\phi(\gamma)\rangle$, we use the formulas in Propositions 16–21 in Secs. IV B and IV C, with two modifications. The first modification is that $\tilde{\mathbf{\Phi}}$ is replaced with $\mathcal{F}\tilde{\mathbf{\Phi}}$ throughout. This replacement causes no structural change to the

formulas; it just changes the inputs. The second modification is caused by the replacement of $\delta(\gamma-\gamma_1)$ by $\mathcal{F}_{\mathbf{r}_1}\delta(\gamma-\gamma_1)$. The first row of $\mathbb{L}(\gamma)$ in (56) becomes $\mathcal{F}\mathbf{\Phi}(\gamma)^*$, which makes $|\mathbb{L}(\gamma)|$ $=|\mathbb{E}|\mathcal{F}\mathbf{\Phi}(\gamma)^*\mathbf{d}$. Similarly, (65) becomes

$$\mathbf{\Theta}(\gamma,\gamma')=\tilde{\mathbf{\Theta}}(\gamma')-\mathbf{d}\frac{\mathcal{F}\mathbf{\Phi}(\gamma)^*\tilde{\mathbf{\Theta}}(\gamma')-\mathcal{F}\delta(\gamma-\gamma')}{\mathcal{F}\mathbf{\Phi}(\gamma)^*\mathbf{d}}. \tag{83}$$

Tracking $\mathcal{F}$ through the formulas, we find that all we need to do is to modify the formulas in Secs. IV B and IV C by applying $\mathcal{F}$ to the final result.

C. C A (17)

In this section, we construct the kernels in (17) for the normal equations (15) using the same ideas in Sec. IV. We fix $\cdot$ and $\cdot'$ and define

$$(\gamma,\gamma')=\frac{\mathcal{A}(\cdot,\cdot')(\gamma,\gamma')}{\hat{\ }\hat{\ }\cdot'}, \tag{84}$$

$$\mathbf{w}(\gamma'[\phi_2(\gamma')\cdots\hat{\phi}(\gamma')]^*, \tag{85}$$

$$\mathbf{y}(\sigma\sigma[$$

$$\mathcal{O}(\epsilon^{2} \quad {}^{2}(\quad + \quad {}_{\mathcal{P}})). \tag{91}$$

The operation count to solve the normal equations (15) by applying the matrix of integral operators $\mathbb{A}$ times is

$$\mathcal{O}(\epsilon^{2} \quad (\quad + \quad)). \tag{92}$$

As we loop through the directions, we may reuse several quantities, so the total cost of the construction is less than

$$\widetilde{\mathbf{\Theta}}_2 = \mathbb{E}_2^{\ddagger}\mathcal{F}\widetilde{\mathbf{\Phi}}_2 = (\mathbb{E}_1^{\ddagger} + \mathbf{fg}^*)(\mathcal{F}\widetilde{\mathbf{\Phi}}_1 + \mathbf{e}_1(\phi_1 - \hat{\phi}_1)) = \widetilde{\mathbf{\Theta}}_1 + \mathbf{d}_1(\phi_1 - \hat{\phi}_1) + \mathbf{fg}^*\mathcal{F}\widetilde{\mathbf{\Phi}}_1 + \mathbf{fg}^*\mathbf{e}_1(\phi_1 - \hat{\phi}_1) \tag{96}$$

at cost $\mathcal{O}($ $)$. It is insufficient to just update $\widetilde{\mathbf{\Theta}}_2$ in this way, since it would ste

$$\sum_{=1} \prod_{=1} (\phi -). \qquad \text{(A4)}$$

$$|A_1^{\ddagger}| = -(1/\quad)|A^{\ddagger}|. \tag{B4}$$

(2) $\lambda \neq 0$, $=0$, $=0$, $(A_1) = (A)$

$$A_1^{\dagger} = A^{\dagger} - \lambda^{-1}\mathbf{d}\mathbf{e}^*, \tag{B5}$$

$$A_1^{\perp} = A^{\perp}, \tag{B6}$$

$$|A_1^{\ddagger}| = |A^{\ddagger}|\lambda^{-1}. \tag{B7}$$

(3) $=0$ $\neq 0$, $(A_1) = (A)$

$$A_1^{\dagger} = A^{\dagger} - \mu^{-1}\mathbf{d}(\ \mathbf{d}^*A^{\dagger} + \lambda\mathbf{e}^*) + \mu^{-1}\mathbf{g}(-\ \mathbf{e}^* + \lambda\mathbf{d}^*A^{\dagger}), \tag{B8}$$

$$A_1^{\perp} = A^{\perp} - \frac{|\lambda|(\ \bar{\mu} - |\lambda|)\mathbf{g} + \lambda\ \mathbf{d}}{\ |\lambda|\ \bar{\mu}}\mathbf{g}^*A^{\perp}, \tag{B9}$$

$$|A_1^{\ddagger}| = |A^{\ddagger}|\frac{(\lambda - \lambda)|\lambda|^2 + \lambda\mu}{\mu|\lambda|\ \bar{\mu}}. \tag{B10}$$

(4) $\neq 0$ $=0$, $(A_1) = (A)$

$$A_1^{\dagger} = A^{\dagger} - \nu^{-1}(\ A^{\dagger}\mathbf{e} + \lambda\mathbf{d})\mathbf{e}^* + \nu^{-1}(-\ \mathbf{d} + \lambda A^{\dagger}\mathbf{e})\mathbf{f}^*, \tag{B11}$$

$$A_1^{\perp} = A^{\perp} - A^{\perp}\mathbf{f}\frac{(|\lambda|(\ \bar{\nu} - |\lambda|)\mathbf{f} + \lambda\ \mathbf{e})^*}{\ |\lambda|\ \bar{\nu}}, \tag{B12}$$

$$|A_1^{\ddagger}| = |A^{\ddagger}|\frac{(\lambda - \lambda)|\lambda|^2 + \lambda\nu}{\nu|\lambda|\ \bar{\nu}}. \tag{B13}$$

(5) $\neq 0$ $\neq 0$, $(A_1) = (A) + 1$

$$A_1^{\dagger} = A^{\dagger} - \ ^{-1}\mathbf{d}\mathbf{f}^* + \ ^{-1}\mathbf{g}(-\mathbf{e}^* + \lambda\ ^{-1}\mathbf{f}^*), \tag{B14}$$

$$A_1^{\perp} = A^{\perp} - (1/\ \bar{\ })\mathbf{g}\mathbf{f}^*, \tag{B15}$$

$$|A_1^{\ddagger}| = |A^{\ddagger}|[1 + (\ ^{-1}\ ^{-1} - (1/\ \bar{\ }))\mathbf{g}^*A^{\perp}\mathbf{f}], \tag{B16}$$

$A_1^{\dagger}$, $A_1^{\perp}$, $|A_1^{\ddagger}|$ $\mathcal{O}(\ ^2)$.

: The overall method, update rules for (A_1), and update rules for $A_1^{\dagger}$ are taken from Ref. 4, which also lists the useful properties

$$\mathbf{c}^*\mathbf{d} = \mathbf{e}^*\mathbf{b} = \lambda - 1, \quad \mathbf{b}^*\mathbf{f} = , \quad \mathbf{c}^*\mathbf{g} = , \quad \mathbf{d}^*\mathbf{g} = 0, \quad \mathbf{e}^*\mathbf{f} = 0,$$

$$A^{\dagger}A\mathbf{d} = \mathbf{d}, \quad AA^{\dagger}\mathbf{e} = \mathbf{e}, \quad A^*\mathbf{f} = A^{\dagger}\mathbf{f} = 0, \quad A\mathbf{g} = (A^{\dagger})^*\mathbf{g} = 0. \tag{B17}$$

They give update rules for the row and column spans of A_1, which we translate into update rules for $A^{\perp}$. The cases (B3), (B6), and (B15) follow directly. Corresponding to (B9), their update rule is that the row span of $A^{\perp}$ should be extended (orthogonally) by $\mathbf{d}$ and then reduced by projecting orthogonal to $\mathbf{p}$. We translate this into a (Householder) reflection of the vector $\mathbf{g}$ into a vector in

the span of $\mathbf{d}$ and $\mathbf{g}$ perpendicular to $\mathbf{p}$. Adjusting these vectors to have equal norm and real inner product yields the reflection of the vector $\lambda\ \overline{\mu}\mathbf{g}$ to $-|\lambda|(\zeta\ \mathbf{d}-\lambda\mathbf{g})$, resulting in

$$\left(\mathbb{I}-\frac{2(\lambda\ \overline{\mu}\mathbf{g}+|\lambda|(\zeta\ \mathbf{d}-\lambda\mathbf{g}))(\lambda\ \overline{\mu}\mathbf{g}+|\lambda|(\zeta\ \mathbf{d}-\lambda\mathbf{g}))^*}{\|(\lambda\ \overline{\mu}\mathbf{g}+|\lambda|(\zeta\ \mathbf{d}-\lambda\mathbf{g}))\|^2}\right)\mathbb{A}^{\perp}, \tag{B18}$$

which simplifies to (B9). To obtain (B12), we use the same process, extending the column span by $\mathbf{e}$ and then projecting orthogonal to $\mathbf{q}$ by a reflection of $\lambda\ \overline{\nu}\mathbf{f}$ to $-|\lambda|(\cdot\mathbf{e}-\lambda\mathbf{f})$.

To derive the update rules for $|\mathbb{A}_1^{\ddagger}|$, first add the update rules for $\mathbb{A}_1^{\dagger}$ and $\mathbb{A}_1^{\perp}$ and then takTD(A)Tj/F4 a60191 643.1329 Tm

When $\mathbb{A}$ and $\mathbb{A}_1$ are nonsingular, (B5) is the Sherman–Morrisson formula (see, e.g., Ref. 23). For our application, we need the singular vectors in $\mathbb{A}^{\perp}$ rather than $\mathbb{A}^{\perp}$ itself, but then only when $\mathrm{rank}(\mathbb{A}^{\perp}) \leqslant 3$. These singular vectors can be extracted by a simple modification of the power method with deflation.

[1] Agmon, S., [illegible] *(C , 1984)*, Lecture Notes in Mathematics, Vol. 1159 (Springer, Berlin, 1985

[35] Kato, T., “Fundamental properties of Hamiltonian operators of Schrödinger type,” Trans. Am. Math. Soc. **70**, 195 (1951).
[36] Klopper, W., *A* *C* , NIC Series Vol. 1, edited by Grotendorst, J. (John von Neumann Institute for Computing, Jülich, 2000), pp. 153–201.
[37]