

Efficient Quasiparticle Determination beyond the Diagonal Approximation via Random Compression

Annabelle Canestraight,* Xiaohe Lei, Khaled Z. Ibrahim, and Vojtěch Vlček



Cite This: *J. Chem. Theory Comput.* 2024, 20, 551–557



Read Online

ACCESS |



Metrics & More

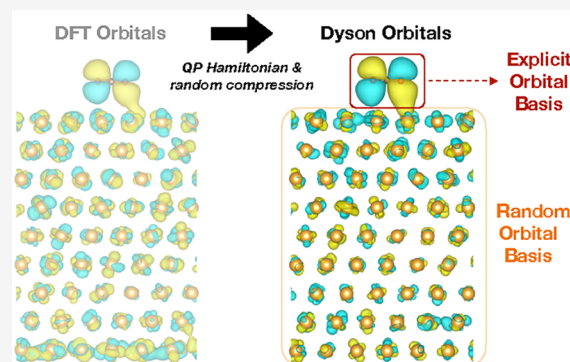


Article Recommendations



Supporting Information

ABSTRACT: Calculations of excited states in the Green's function formalism often invoke the diagonal approximation, in which the quasiparticle states are taken from a mean-field calculation. In this paper, we extend the stochastic approaches applied in the many-body perturbation theory and overcome this limitation for large systems in which we are interested in a small subset of states. We separate the problem into a core subspace whose coupling to the remainder of the system environment is stochastically sampled. This method is exemplified on computing hole injection energies into CO₂ on an extended gold surface with nearly 3000 electrons. We find that in the extended system the size of the problem can be compressed up to 95% using stochastic sampling. This result provides a way forward for self-consistent stochastic methods and determination of Dyson orbitals in large systems.



INTRODUCTION

Single particle states are frequently used in the study of excitation phenomena such as photoionization, electron injection, and generally optical transitions.^{1–9} The physical interpretation of such single particle states often depends on the specific type of observables.^{7,10} In particular, Dyson orbitals, which correspond to the probability amplitude distribution of a specific electron or hole excitation (i.e., quasiparticle state), are directly accessible via orbital tomography and provide insights into the relation between energies and real-space distribution of single particle excitation.^{11,12} This has fundamental implications for chemistry—e.g., hybridization of quasiparticles on surfaces governs the propensity for direct injection of an electron.⁹ These are just a few compelling reasons to account for the physically meaningful orbital distributions, especially for problems concerning (chemical) interfaces.

In practice, however, single-particle states for interfacial systems are typically taken from the Density Functional Theory (DFT),^{13–15} as the cost of higher level theory is too great. While DFT can handle extremely large systems,¹⁶ these calculations can not, even in principle, yield quasiparticle (QP) energies or the Dyson orbitals.^{7,17} A natural and widely applied extension, especially in condensed matter problems, is application of the Many Body Perturbation Theory (MBPT) employing Green's function formalism.^{7,18–20} In particular, the GW approximation, which truncates the correlation expansion to nonlocal charge density fluctuations, has emerged as arguably the most popular approach.^{21,22} Its self-consistent solution yields both QP energies and the Dyson orbitals.^{23–25} However, it is common to apply the GW approach as a one-

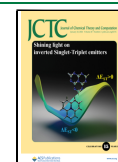
shot correction, G_0W_0 , employing the Kohn–Sham Green's function G_0 and the screened coulomb interaction W_0 derived from the underlying Kohn–Sham DFT solutions. Despite its approximate nature, G_0W_0 often provides good estimates of band gaps.^{25–30} The use of one-shot corrections has been largely motivated by the computational cost,^{31,32} which scales as $O(N^4)$ with the number of electrons in conventional implementations.^{33,34} The computational cost has been significantly decreased by stochastic sampling approaches in GW (and post-GW) to be nearly linear; 1000s of states can thus be studied.^{27,35–37} Such methods allow for the construction of a self-energy operator that acts on a single state and converges to the result of explicit treatment of the entire system. However, even in the stochastic GW, “updating” the single-particle basis (i.e., finding the Dyson orbitals) is difficult³⁸ and, in practice, usually avoided.³⁹ Routine calculations of QP orbitals in realistic systems with thousands of electrons are still elusive. This is true even if one is, in principle, interested in treating a *small subset* of states, as exemplified in this work (see below). In cases where the Dyson orbitals differ greatly from the KS DFT Orbitals (i.e., when there are large off-diagonal terms), there may be large changes in the QP energies. Therefore, for many systems, the G_0W_0

Received: September 26, 2023

Revised: December 4, 2023

Accepted: December 26, 2023

Published: January 4, 2024



approximation may be insufficient, and it becomes necessary to go beyond the diagonal approximation.

Here, we tackle this problem and present a scheme without diagonal approximation for realistic nanoscale systems. The stochastic GW method^{27,37} enables selective access to individual elements of the QP Hamiltonian without evaluating all the terms; it is unclear *a priori* which states must be used to describe key properties of the system. This stochastic framework is exemplified for a CO₂ molecule on a large Au slab. For this problem, the surface contributions to the orbitals are sampled, drastically reducing the cost of QP calculations. This method divides the system into a set of states in a “core” subspace, treated by standard stochastic MBPT, and a rest space, for which additional sampling is introduced. Sampling of the rest space allows for monotonic convergence toward the full deterministic QP energies and for all rest space states to be represented by a few random vectors. This step is combined with a search over the fixed-point solutions of the frequency-dependent QP Hamiltonian, which is basis representation independent and thus enables the QP Hamiltonian to act on a mixed-basis of both KS DFT and random vectors.

We apply these methods to a prototypical system of a small molecule on a plasmonic surface (CO₂ on Au illustrated in the inset in Figure 1). In the practical demonstration for an

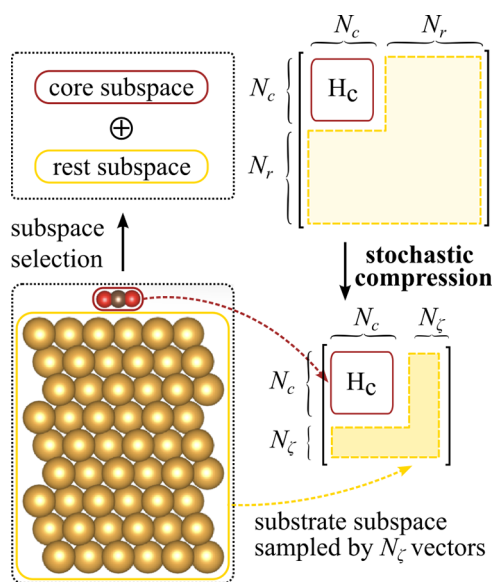


Figure 1. Illustration of the stochastic compression technique, which samples the “rest subspace” using a set of (filtered) random vectors, here spanning the single particle states of the gold substrate.

extended Au(111) surface with 270 atoms (2986 electrons), we found convergence in the hybridized HOMO energy with a 95% rank compression compared to the evaluation on the full canonical orbital basis. This success provides a way to use costly high-level theories to study realistic chemical systems.

■ FORMALISM

The time-ordered Green’s function (GF) contains information about the quasiparticle (QP) energy spectrum and lifetimes, and it corresponds to the probability amplitude of a QP propagation between two space-time points $\mathbf{r}, t$ and $\mathbf{r}', t'$. In the Lehmann representation, the Green’s function for the hole states is expressed as

$$G(\mathbf{r}, \mathbf{r}', \omega) = \sum_n \left[\frac{\psi_n(\mathbf{r})\psi_n(\mathbf{r}')^*}{\omega - \varepsilon_n - i\eta} \right] \quad (1)$$

where the Dyson orbitals are obtained as $\psi_n(\mathbf{r}) = \langle \Psi_{N-1}^n | \hat{\psi}(\mathbf{r}) | \Psi_N^0 \rangle$ from the N -particle ground state and the n^{th} excited state of the $N - 1$ particle system, where $\hat{\psi}(\mathbf{r})$ is the field operator, and n runs over all hole states in the system. The poles of the GF are located at the QP energies, ε_n , here corresponding to the charge removal.⁷ Charge addition is treated analogously. The GF poles are conveniently expressed as solutions to a nonlinear eigenvalue problem for an effective Hamiltonian obtained by downfolding interactions with the system:⁷

$$\hat{H}_{QP}(\omega)|\psi\rangle = \omega|\psi\rangle \quad (2)$$

In practice, the QP Hamiltonian is divided into a static and local term, H_0 , which typically contains all one-body contributions, while a space-time nonlocal portion is represented by the self-energy operator $\hat{\Sigma}$.²² The latter is approximated by the selected types of interaction diagrams (and their resummation). $\hat{v}_{xc}$ represents the mean-field exchange-correlation potential, and $\hat{\Sigma}_{xc}$ represents the nonlocal and dynamical (energy and state-dependent) term comprising the many-body interactions of the given quasiparticle. Therefore, the QP Hamiltonian is practically constructed by replacing $\hat{v}_{xc}$ with the self-energy as a perturbative correction on top of a mean-field starting point

$$\hat{H}_{QP}(\omega) = \hat{H}_{0,KS} - \hat{v}_{xc} + \hat{\Sigma}(\omega) \quad (3)$$

where $\hat{H}_{0,KS}$ is the KS DFT Hamiltonian.

Further, the “one-shot” correction corresponds to

$$\Sigma(\mathbf{r}, \mathbf{r}', \omega) = i \int \frac{d\omega'}{2\pi} G_0(\mathbf{r}, \mathbf{r}', \omega + \omega') W_0(\mathbf{r}, \mathbf{r}', \omega') \quad (4)$$

where G_0 has poles at the DFT Kohn–Sham eigenvalues, ε_0 , and W_0 is the screened coulomb interaction. The self-consistency requires repeated construction of Σ and re-evaluation of eq 2; multiple flavors of self-consistent approaches have been developed.^{23,24} Typically, the convergence pattern is smooth. If the KS DFT single-particle states are close to the Dyson orbitals, the “one-shot” correction provides good estimates of QP energies, yet the quality of the mean-field eigenstates is not *a priori* known.

A step beyond this practice is to diagonalize H_{QP} in eq 2 in the orbital basis, yielding Dyson orbitals (in the first iteration) and updated one-shot QP energies in the GW approximation.⁷ Note that, in principle, the nonlinear problem in eq 2 holds for multiple values of ω associated with satellite features.^{40–42} In this Letter, we will focus only on the primary QP peaks, i.e., we seek a single solution to the QP Hamiltonian in the vicinity of ε_0 and look for the fixed point solutions to $\omega_i = \langle \psi_i | \hat{H}_{QP}(\omega_i) | \psi_i \rangle$. Note that H_{QP} is non-Hermitian, and each QP state, in general, corresponds to H_{QP} computed at a different frequency.

In practical schemes,^{25,39,43,44} it is common to construct a single “static” effective Hamiltonian (yielding orthogonal eigenstates). However, due to the nonlinearity of this problem, it is not entirely clear at what frequency the self-energy should be evaluated. For strongly diagonally dominant H_{QP} , i.e., those where KS DFT orbitals are, in fact, close to the Dyson orbitals, one may evaluate ω_i as the fixed point solution for the diagonal

entries. The remaining off-diagonal self-energy is e.g., $\Sigma_{ij} = \frac{1}{4}[\Sigma_{ij}(\omega_i) + \Sigma_{ji}(\omega_i) + \Sigma_{ij}(\omega_j) + \Sigma_{ji}(\omega_j)]$. In this form, it is possible to construct a static and hermitized QP Hamiltonian. By enforcing the hermicity of H_{QP} , we impose that the resulting QP states are orthonormal. The QP energies are purely real, corresponding to an infinite lifetime QP. Alternatively, one can therefore relax the latter step by taking $\Sigma_{ij} = \frac{1}{2}[\Sigma_{ij}(\omega_i) + \Sigma_{ji}(\omega_j)]$.

Note that both approaches strongly depend on the basis choice. We illustrate this in detail in the [Supporting Information \(SI\)](#) for the acrolein molecule, for which the magnitudes of the off-diagonal terms are 98.5% smaller than the diagonal terms for the canonical KS DFT basis. The situation changes dramatically when maximally localized (Pipek-Mezek-Wannier) orbitals are employed. These differ from the canonical states by a unitary transformation. Depending on the construction of a single H_{QP} , the resulting QP energies change by as much as 10% and translate to changes of 0.77 eV on average for acrolein.

Since our goal is to determine Dyson orbitals for a selected subspace of interest (which will be constructed from localized basis states), we avoid any approximation to the fixed point solution. In this method, the whole QP Hamiltonian is evaluated at multiple frequencies, and the QP eigenvalues are found as the fixed point solutions in [eq 2](#). No assumptions are made about the hermicity of the Hamiltonian matrix; a graphical example of such a fixed point solution for H_{QP} is also illustrated in the [SI](#).

STOCHASTIC COMPRESSION OF QP STATES

When a large system with a subspace of particular interest is studied, it is prohibitively expensive to employ all M electronic states. It is also insufficient to assume that the Hamiltonian matrix takes a block-diagonal form due to the coupling between the subspace and its orthogonal complement. To handle such a case, we propose a method of stochastic matrix compression, where a portion of the H_{QP} matrix is represented by a set of random vectors. These vectors sample a large portion of the Hilbert space, which contributes to the overall QP shift and affects the Dyson orbitals, but for which each individual single particle state has only a limited contribution.

As illustrated in [Figure 1](#), we separate the “core subspace” spanned by N_c deterministic states, $\{\phi^c\}$ (e.g., the original KS DFT eigenstates), and the remainder spanned by a N_s stochastic state $\{\zeta\}$, constructed as a random linear combination of the KS states that are orthogonal to the $\{\phi^c\}$ set: $|\zeta\rangle = \sum_{i=1 \perp \phi^c}^{N_s} c_i |\phi_i\rangle$. In the final step, the individual random states are orthogonalized via the Gram-Schmidt process. Because this change of basis is guaranteed to be a unitary transformation of the Hamiltonian matrix, when the whole system is diagonalized, the resulting eigenstates will be the same. When the Hamiltonian matrix is truncated on this new stochastic basis, the coupling of each stochastic state to the core subspace will represent the subspace interaction with the full environment. In this way, we have “compressed” the information on the whole system environment into a single state. Given that the fixed point solution is basis independent (as illustrated in [Figure S.3](#)), a total number of states $N_c + N_s$ is the same as the dimension of H_{QP} , M , we necessarily obtain the same QP energies. For fewer random states, $N_c + N_s < M$, the computation is less expensive. The accuracy of the eigenvalues

of the truncated Hamiltonian can, in theory, be systematically improved by the expansion of the core or stochastic subspaces. The conceptual advantage to increasing N_s rather than N_c lies in the removal of bias. We can not presume to know the optimal metric for identifying N_c for a given system, but through the use of stochastic states, we can achieve an unbiased, systematic improvement of the result. Note that the QP energy has a finite statistical error, which decreases as $1/\sqrt{N_s}$ with the number of states sampling the off-diagonal self-energy contributions. As we show below, the convergence of the QP energies is smooth. Further, note that instead of the canonical single particle states in the above equation, we achieve further speedup if a preselected (filtered) subset of states (orthogonal to the $\{\phi^c\}$) is used in the construction of $|\zeta\rangle$.

RESULTS

We now demonstrate the method practically for the CO₂ molecule on the Au substrate for which we intend to extract the energies of quasi-hole states on the molecule (i.e., corresponding to the charge removal energies from CO₂ on the surface). We first construct a minimal example that we can solve entirely and illustrate how stochastic sampling smoothes the convergence of the QP energies. Later, we show a realistic example with nearly 3,000 electrons, which cannot be easily solved without the sampling methodology (See the [SI](#) for computational details.).

We will demonstrate the success of our stochastic sampling method on a minimal system of CO₂ on a bilayer of 8 gold atoms. This system contains only 52 occupied states, which we treat explicitly at the GW level using the Stochastic GW methodology and the diagonalization procedure discussed in the [SI](#). Note that, in principle, the hybridization extends beyond merely the occupied manifold, but to illustrate the methodology, we consider only the rotation within the occupied subspace. To see the surface-induced changes, calculate the QP states for a CO₂ molecule in a vacuum ($N = 8$) and for the minimal composite system ($N = 52$). We find, via projection, that the Dyson orbitals of the seven lowest valence states of the molecule change only within the molecular subspace, but they do not hybridize with the surface, and adding the surface states does not impact the deep valence states of the molecule.

In contrast, the HOMO state behaves differently: no single DFT state is in one-to-one correspondence to the isolated molecular HOMO (indeed, this is also true for the Dyson orbitals computed at the G_0W_0 level). Instead, there are multiple *hybridized* states sufficiently localized on the molecule, whose eigenvalues lie within a small range of energies. We aim to characterize them and, consequently, to find a characteristic QP energy for this distribution of HOMO QP for the CO₂ molecule on Au.

We thus define a “core subspace” comprising the states with the most molecular character. In the realistic system, the molecular HOMO has hybridized with hundreds of surface states (see the inset of [Figure 3](#)), preventing us from including all of these hybridized states explicitly, i.e., in the core subspace. In practice, we define the core molecular states based on their projection onto a set of maximally localized molecular orbitals. These are constructed by applying a unitary transformation to the entire system which minimizes the spread of 8 orbitals around the CO₂ molecule per the Pipek-

Mezey Wannierization scheme.^{36,45} The corresponding projection value is

$$P_i = \sum_j |\langle \xi_j | \phi_i \rangle|^2 \quad (5)$$

Here, $\{|\xi\rangle\}$ and $\{|\phi\rangle\}$ are the sets of transformed (localized) and canonical KS DFT states, respectively. Each KS state with P greater than a chosen threshold is included in the core region. This preselection separates the “core” subspace from the rest.

We now track the fixed point HOMO QP solution with the number of states considered in the H_{QP} , i.e., we gradually add more states outside of the core subspace. The molecular HOMO is hybridized with many surface states. We thus define a single energy for this state by taking its mean value, constructed by weighting by the projection onto the HOMO of CO₂ in a vacuum. The results are shown by the green color in Figure 2. The leftmost point represents only the core space,

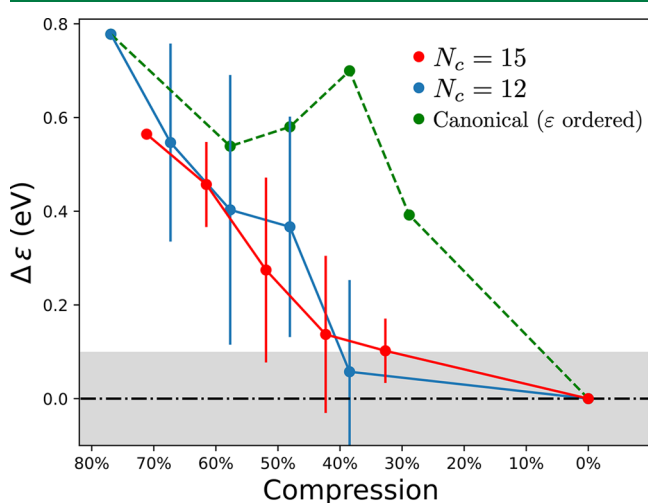


Figure 2. Hybridized HOMO Convergence (Minimal System): Core sizes of $N_c = 12$ and $N_c = 15$ are used, with the remaining states sampled with equal weight. In contrast, adding the states by energy (ϵ ordered) demonstrates the lack of smooth convergence. The gray-shaded region shows where the spectrum converges within 0.1 eV.

containing 12 orbitals corresponding to 23% of the entire H_{QP} . The size of the problem is increased by adding states depending on their distance from the KS DFT HOMO energy. One would expect, from perturbation theory, that the hybridization of states will be small for energetically distant states. However, this does not produce a smooth convergence (green line in Figure 2) as it appears that the first order perturbation correction in orbitals would not suffice and the correlations (accounted for at the GW level) play a larger role.

To demonstrate the stochastic approach, we now instead sample the remaining KS states using random vectors:

$$|\zeta\rangle = \frac{1}{\sqrt{N_{env}}} \sum_{j=1}^{N_{env}} w_j e^{i\theta_j} |\phi_j\rangle \quad (6)$$

Here, $\theta_j \in [0, 2\pi]$ is randomly chosen, and N_{env} is the number of “rest” states sampled with weight w_j . N_{env} is the upper bound to $M - N_c$, corresponding to the exhaustion of the sampled subspace. Note that we can sample all remaining states evenly ($w_i = 1/\forall i$), but generally, we consider a random selection from

a distribution within the sampled subspace (determined by P_i in eq 5) as we show later.

Once we have obtained the set $\{\zeta\}$, we randomly draw N_s of them, and the fixed point solution is then found for H_{QP} with the dimension of $N_c + N_s$. We plot this result as a function of “percent compression” which is $100 \times (1 - (M/N_{occ}))$. Here, N_{occ} is the number of electrons in the entire system. The results for $N_c = 12$ and variable N_s are in Figure 2 and show a monotonic and smooth convergence toward to the asymptotic value (obtained for the entire 52 occupied states). The stochastic sampling was repeated ten times for each step with a different set of N_s random vectors; the standard deviations are indicated in the plot, and they naturally disappear in the complete basis limit. For instance, for $N_s = 20$, i.e., 38% compression of the system, we see a difference of 0.057 ± 0.19 eV between the mean HOMO QP energies. For an increased core space, $N_c = 15$, we see that the HOMO QP value converges similarly; i.e., the size of the core space is not changing the convergence profile significantly. For $N_s = 20$ (i.e., 32.5% compression of matrix rank), the resulting spectrum mean falls within 100 meV of the value obtained from the diagonalization of the full matrix.

Without any prior knowledge or arbitrary truncation of the KS states, we can capture molecule–surface hybridization effects by employing stochastic states representing the substrate environment. This description is systematically improvable by increasing both N_c and N_s . In general, the cost reduction provided by the stochastic sampling is due to circumventing the explicit treatment of (in principle) all single-particle states that contribute little to the expectation value of an operator.⁴⁶ Here, we sample the QP energies of the selected states of CO₂. For a small system such as the one used here, the amount of compression is less significant, as most of the states contribute to the QP HOMO energy.

We now turn to a realistic large-scale system, for which such a calculation would not be possible with standard methods. Here, we study a CO₂ molecule on an extended Au-111 surface of 270 atoms, containing 2986 electrons. The system is treated analogously to the minimal example: we selected core subspaces of 15 and 25 states. $N_c = 15$ was the core subspace identified when the projection threshold $P_T = 0.04$. Next, the stochastic sampling uses a filtered distribution in eq 6 in which we consider a linear combination of states that are sufficiently localized on the molecules. In practice, this step determines the sampled subspace, which is practically restricted to states with P greater than a selected threshold, P_T . Here, we consider two cases: $P_T = 10^{-3}$ and $P_T = 5 \times 10^{-4}$.

From Figure 3 we can see that the HOMO energy converges with 95% compression (5% of the total number of states used). [In practical execution on HPC machines, namely at the National Energy Research Scientific Computing Center using a Intel Xeon Processor E5-698 v3 at 2.3 GHz, the difference in computational cost for using 96% and 92% compressed matrix amounts to $\sim 77,000$ CPU·hrs. For the maximally compressed converged calculation (96% compression), the entire calculation amounts to $\sim 60,000$ CPU·hrs.] For slightly increased selectivity (i.e., lower projection threshold P), stochastic sampling of the hybridization converges similarly. Further, the size of the core subspace does not significantly impact the convergence rate: when $N_c = 25$ with the filtering threshold of $P_T = 5 \times 10^{-4}$, the curve matches that of the $N_c = 15$ for the same value of P_T . This suggests that the number of states required does not vary significantly with the particular size of

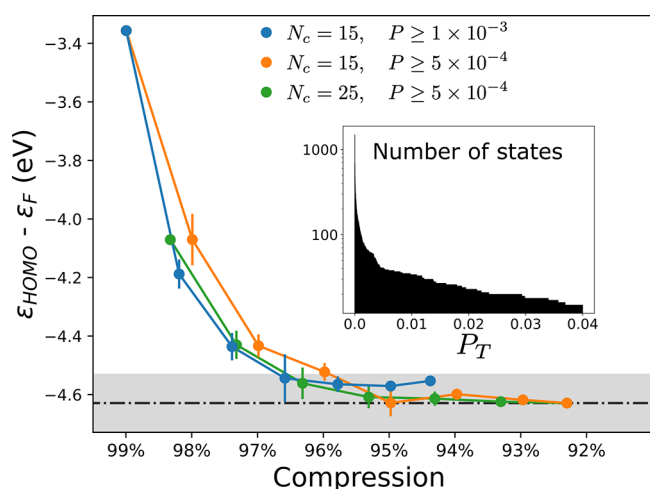


Figure 3. Hybridized HOMO Convergence (Large System with 2986 Electrons): Convergence is demonstrated across multiple core sizes and stochastic sampling filterings. The shaded region indicates that HOMO is within 0.1 eV of the converged value. The inset shows the log-scale histogram of the number of states for which the projection value $P \geq P_T$, where P_T is the filtering cutoff (see text).

the core subspace. Practically, we see that little or no optimization of the core subspace is required for convergence.

Finally, note that when the orbital rehybridization is used at the G_0W_0 level, the HOMO QP energy moves down in energy by more than 1 eV. Since approximate semilocal KS DFT is known to suffer from overdelocalization,⁴⁷ it is expected that the physical Dyson orbitals are more localized than the canonical KS DFT eigenstates.

Indeed, if we inspect the orbital shape and distribution for the single most molecule-like state, the Dyson orbital and the corresponding DFT KS eigenstate are distinct, and the former is slightly more localized on the molecule (Figure 4). The

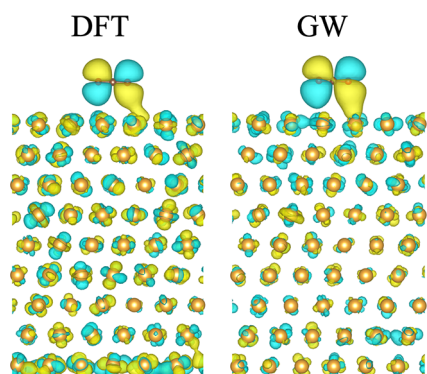


Figure 4. Hybridized HOMO Orbitals: The most molecule-like hybridized HOMO orbitals, computed with DFT and GW, respectively, are plotted. The Dyson orbital (GW), computed with $M = 115$ states (corresponding to the final point of the orange curve of Figure 3), is shown to be more localized on the molecule.

visual changes are, however, not indicative of the magnitude of the QP energy shift accomplished, and hence, these are better inspected in Figure 3. In turn, stronger localization of HOMO is typically associated with its energy decrease.^{9,48} These observations are thus in line with what the MBPT should accomplish and underline the need for a more appropriate treatment of surface phenomena. Without studying the change

in the orbital structure itself, the large downward energy shift, accomplished by accounting for interactions with higher energy states, emphasizes the importance of taking GW calculations beyond the diagonal approximation.

OUTLOOK

The rapid convergence of the QP energies with N_s implies that when we stochastically sample the matrix, aided by preselection and filtering, we can represent the full QP spectrum for a molecule that hybridizes with an extended surface by using less than 5% of the system. The H_{QP} matrix size is thus compressed by 95%. This is largely due to the significant “redundancy” of information encoded in individual single-particle states, and the sampling allows sampling of all (or a large filtered portion of them) simultaneously through random vectors. The approach presented here will enable the treatment of large-scale interfacial problems and opens the door for efficient, self-consistent stochastic MBPT.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge at <https://pubs.acs.org/doi/10.1021/acs.jctc.3c01069>.

Au(111) slab simulation cell parameters, DFT calculations for the minimal system and large system, and self-consistent treatment and basis set independence (PDF)

AUTHOR INFORMATION

Corresponding Author

Annabelle Canestraight – Department of Chemical Engineering, University of California, Santa Barbara, California 93106-9510, United States; orcid.org/0000-0001-6473-3762; Email: acanestraight@ucsb.edu

Authors

Xiaohe Lei – Department of Chemistry and Biochemistry, University of California, Santa Barbara, California 93106-9510, United States

Khaled Z. Ibrahim – Applied Mathematics and Computational Research Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, United States

Vojtěch Vlček – Department of Chemistry and Biochemistry and Department of Materials, University of California, Santa Barbara, California 93106-9510, United States; orcid.org/0000-0002-2836-7619

Complete contact information is available at: <https://pubs.acs.org/doi/10.1021/acs.jctc.3c01069>

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The development (A.C., V.V., and X.L.) of the stochastic compression technique was supported by the NSF CAREER award (DMR-1945098). The implementation and numerical testing (A.C., K.I., and V.V.) is based upon work supported by the U.S. Department of Energy, Office of Science, Office of Advanced Scientific Computing Research and Office of Basic Energy Sciences, Scientific Discovery through Advanced Computing (SciDAC) program under Award Number DE-SC0022198. This research used resources of the National

Energy Research Scientific Computing Center, a DOE Office of Science User Facility supported by the Office of Science of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231 using NERSC award BES-ERCAP0020089.

REFERENCES

- (1) Ivanov, M. V.; Bangerter, F. H.; Krylov, A. I. Towards a rational design of laser-coolable molecules: insights from equation-of-motion coupled-cluster calculations. *Phys. Chem. Chem. Phys.* **2019**, *21*, 19447–19457.
- (2) Ivanov, M. V.; Jagau, T.-C.; Zhu, G.-Z.; Hudson, E. R.; Krylov, A. I. In search of molecular ions for optical cycling: a difficult road. *Phys. Chem. Chem. Phys.* **2020**, *22*, 17075–17090.
- (3) Gozem, S.; Seidel, R.; Hergenbahn, U.; Lugovoy, E.; Abel, B.; Winter, B.; Krylov, A. I.; Bradforth, S. E. Probing the Electronic Structure of Bulk Water at the Molecular Length Scale with Angle-Resolved Photoelectron Spectroscopy. *J. Phys. Chem. Lett.* **2020**, *11*, 5162–5170.
- (4) Puschnig, P.; Berkebile, S.; Fleming, A. J.; Koller, G.; Emtsev, K.; Seyller, T.; Riley, J. D.; Ambrosch-Draxl, C.; Netzer, F. P.; Ramsey, M. G. Reconstruction of Molecular Orbital Densities from Photoemission Data. *Science* **2009**, *326*, 702–706.
- (5) Tretiak, S.; Mukamel, S. Density Matrix Analysis and Simulation of Electronic Excitations in Conjugated and Aggregated Molecules. *Chem. Rev.* **2002**, *102*, 3171–3212.
- (6) Lüftner, D.; Ules, T.; Reinisch, E. M.; Koller, G.; Soubatch, S.; Tautz, F. S.; Ramsey, M. G.; Puschnig, P. Imaging the wave functions of adsorbed molecules. *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 605–610.
- (7) Martin, R. M.; Reining, L.; Ceperley, D. M. *Interacting Electrons: Theory and Computational Approaches*; Cambridge University Press, 2016.
- (8) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659.
- (9) Lei, X.; Canestraight, A.; Vlcek, V. Exceptional Spatial Variation of Charge Injection Energies on Plasmonic Surfaces. *J. Phys. Chem. Lett.* **2023**, *14* (38), 8470–8476.
- (10) Krylov, A. I. From orbitals to observables and back. *J. Chem. Phys.* **2020**, *153*, 080901.
- (11) Puschnig, P.; Boese, A. D.; Willenbockel, M.; Meyer, M.; Lüftner, D.; Reinisch, E. M.; Ules, T.; Koller, G.; Soubatch, S.; Ramsey, M. G.; Tautz, F. S. Energy Ordering of Molecular Orbitals. *J. Phys. Chem. Lett.* **2017**, *8*, 208–213.
- (12) Martin, R. L. Natural transition orbitals. *J. Chem. Phys.* **2003**, *118*, 4775–4777.
- (13) Wu, Z.-Z.; Zhang, X.-L.; Niu, Z.-Z.; Gao, F.-Y.; Yang, P.-P.; Chi, L.-P.; Shi, L.; Wei, W.-S.; Liu, R.; Chen, Z.; et al. Identification of Cu (100)/Cu (111) interfaces as superior active sites for CO dimerization during CO₂ electroreduction. *J. Am. Chem. Soc.* **2022**, *144*, 259–269.
- (14) Biller, A.; Tamblyn, I.; Neaton, J. B.; Kronik, L. Electronic level alignment at a metal-molecule interface from a short-range hybrid functional. *J. Chem. Phys.* **2011**, *135*, 164706.
- (15) Egger, D. A.; Liu, Z.-F.; Neaton, J. B.; Kronik, L. Reliable energy level alignment at physisorbed molecule–metal interfaces from density functional theory. *Nano Lett.* **2015**, *15*, 2448–2455.
- (16) VandeVondele, J.; Borštnik, U.; Hutter, J. Linear Scaling Self-Consistent Field Calculations with Millions of Atoms in the Condensed Phase. *J. Chem. Theory Comput.* **2012**, *8*, 3565–3573.
- (17) Engel, E.; Dreizler, R. *Density Functional Theory: An Advanced Course; Theoretical and Mathematical Physics*; Springer Berlin Heidelberg, 2011.
- (18) Hedin, L. New method for calculating the one-particle Green's function with application to the electron-gas problem. *Phys. Rev.* **1965**, *139*, A796.
- (19) Hybertsen, M. S.; Louie, S. G. First-Principles Theory of Quasiparticles: Calculation of Band Gaps in Semiconductors and Insulators. *Phys. Rev. Lett.* **1985**, *55*, 1418–1421.
- (20) Godby, R. W.; Schlüter, M.; Sham, L. J. Self-energy operators and exchange-correlation potentials in semiconductors. *Phys. Rev. B* **1988**, *37*, 10159–10175.
- (21) Onida, G.; Reining, L.; Rubio, A. Electronic excitations: density-functional versus many-body Green's-function approaches. *Rev. Mod. Phys.* **2002**, *74*, 601–659.
- (22) Golze, D.; Dvorak, M.; Rinke, P. The GW compendium: A practical guide to theoretical photoemission spectroscopy. *Front. Chem.* **2019**, *7*, 377.
- (23) Rusakov, A. A.; Zgid, D. Self-consistent second-order Green's function perturbation theory for periodic systems. *J. Chem. Phys.* **2016**, *144*, 054106.
- (24) Caruso, F.; Rinke, P.; Ren, X.; Rubio, A.; Scheffler, M. Self-consistent GW: All-electron implementation with localized basis functions. *Phys. Rev. B* **2013**, *88*, 075105.
- (25) Kaplan, F.; Weigend, F.; Evers, F.; van Setten, M. J. Off-Diagonal Self-Energy Terms and Partially Self-Consistency in GW Calculations for Single Molecules: Efficient Implementation and Quantitative Effects on Ionization Potentials. *J. Chem. Theory Comput.* **2015**, *11*, 5152–5160.
- (26) Li, J.; Jin, Y.; Rinke, P.; Yang, W.; Golze, D. Benchmark of GW Methods for Core-Level Binding Energies. *J. Chem. Theory Comput.* **2022**, *18*, 7570–7585.
- (27) Vlcek, V.; Rabani, E.; Neuhauser, D.; Baer, R. Stochastic GW Calculations for Molecules. *J. Chem. Theory Comput.* **2017**, *13*, 4997–5003.
- (28) Rignanese, G.-M.; Blase, X.; Louie, S. Quasiparticle Effects on Tunneling Currents: A Study of C₂H₄ Adsorbed on the Si (001)-(2 × 1) Surface. *Physical review letters* **2001**, *86*, 2110.
- (29) Rohlfing, M.; Krüger, P.; Pollmann, J. Quasiparticle band-structure calculations for C, Si, Ge, GaAs, and SiC using Gaussian-orbital basis sets. *Phys. Rev. B* **1993**, *48*, 17791–17805.
- (30) Shirley, E. L.; Martin, R. M. GW quasiparticle calculations in atoms. *Phys. Rev. B* **1993**, *47*, 15404–15412.
- (31) Bruneval, F.; Rodriguez-Mayorga, M.; Rinke, P.; Dvorak, M. Improved One-Shot Total Energies from the Linearized GW Density Matrix. *J. Chem. Theory Comput.* **2021**, *17*, 2126–2136.
- (32) Bruneval, F. Assessment of the Linearized GW Density Matrix for Molecules. *J. Chem. Theory Comput.* **2019**, *15*, 4069–4078.
- (33) Pham, T. A.; Nguyen, H.-V.; Rocca, D.; Galli, G. G. W calculations using the spectral decomposition of the dielectric matrix: Verification, validation, and comparison of methods. *Phys. Rev. B* **2013**, *87*, 155148.
- (34) Govoni, M.; Galli, G. Large Scale GW Calculations. *J. Chem. Theory Comput.* **2015**, *11*, 2680–2696.
- (35) Romanova, M.; Vlcek, V. Decomposition and embedding in the stochastic GW self-energy. *J. Chem. Phys.* **2020**, *153*, 134103.
- (36) Weng, G.; Vlcek, V. Efficient treatment of molecular excitations in the liquid phase environment via stochastic many-body theory. *J. Chem. Phys.* **2021**, *155*, 054104.
- (37) Neuhauser, D.; Gao, Y.; Arntsen, C.; Karshenas, C.; Rabani, E.; Baer, R. Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic GW Approach. *Phys. Rev. Lett.* **2014**, *113*, 076402.
- (38) Romanova, M.; Vlcek, V. Stochastic many-body calculations of moiré states in twisted bilayer graphene at high pressures. *npj Computational Materials* **2022**, *8*, 11.
- (39) Vlcek, V.; Baer, R.; Rabani, E.; Neuhauser, D. Simple eigenvalue-self-consistent ΔGW. *The Journal of Chemical Physics* **2018**, *149*, 174107.
- (40) Mejuto-Zaera, C.; Weng, G.; Romanova, M.; Cotton, S. J.; Whaley, K. B.; Tubman, N. M.; Vlcek, V. Are multi-quasiparticle interactions important in molecular ionization? *J. Chem. Phys.* **2021**, *154*, 121101.

- (41) Aryasetiawan, F.; Hedin, L.; Karlsson, K. Multiple Plasmon Satellites in Na and Al Spectral Functions from Ab Initio Cumulant Expansion. *Phys. Rev. Lett.* **1996**, *77*, 2268–2271.
- (42) Lischner, J.; Pálsson, G. K.; Vigil-Fowler, D.; Nemsak, S.; Avila, J.; Asensio, M. C.; Fadley, C. S.; Louie, S. G. Satellite band structure in silicon caused by electron-plasmon coupling. *Phys. Rev. B* **2015**, *91*, 205113.
- (43) Van Schilfgaarde, M.; Kotani, T.; Faleev, S. Quasiparticle Self-Consistent GW Theory. *Phys. Rev. Lett.* **2006**, *96*, 226402.
- (44) Kotani, T.; van Schilfgaarde, M.; Faleev, S. V. Quasiparticle self-consistent GW method: A basis for the independent-particle approximation. *Phys. Rev. B* **2007**, *76*, 165106.
- (45) Weng, G.; Romanova, M.; Apelian, A.; Song, H.; Vlček, V. Reduced Scaling of Optimal Regional Orbital Localization via Sequential Exhaustion of the Single-Particle Space. *J. Chem. Theory Comput.* **2022**, *18*, 4960–4972.
- (46) Baer, R.; Neuhauser, D.; Rabani, E. Stochastic Vector Techniques in Ground-State Electronic Structure. *Annu. Rev. Phys. Chem.* **2022**, *73*, 255–272.
- (47) Mori-Sánchez, P.; Cohen, A. J.; Yang, W. Localization and delocalization errors in density functional theory and implications for band-gap prediction. *Physical review letters* **2008**, *100*, 146401.
- (48) Perdew, J. P.; Parr, R. G.; Levy, M.; Balduz, J. L. Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy. *Phys. Rev. Lett.* **1982**, *49*, 1691–1694.