Rapid Communications

Quasiparticle spectra from molecules to bulk

Vojtěch Vlček,[1,*] Eran Rabani,[2,3,†] and Daniel Neuhauser[1,‡]

[1]*Department of Chemistry and Biochemistry, University of California, Los Angeles California 90095, USA*

[2]*Department of Chemistry, University of California and Materials Science Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, USA*

[3]*The Raymond and Beverly Sackler Center for Computational Molecular and Materials Science, Tel Aviv University, Tel Aviv 69978, Israel*

(Received 13 August 2017; published 16 March 2018)

A stochastic cumulant GW method is presented, allowing us to map the evolution of photoemission spectra, quasiparticle energies, lifetimes, and emergence of collective excitations from molecules to bulklike systems with up to thousands of valence electrons, including Si nanocrystals and nanoplatelets. The quasiparticle energies rise due to their coupling with collective shake-up (plasmon) excitations, and this coupling leads to significant spectral weight loss (up to 50% for the low-energy states), shortening the lifetimes and shifting the spectral features to lower energies by as much as 0.6 eV. Such features are common to all the systems studied irrespective of their sizes and shapes. For small and low-dimensional systems the surface plasmon resonances affect the frequency of the collective excitation and position of the satellites.

DOI: 10.1103/PhysRevMaterials.2.030801

Recent developments in Green's function (GF) techniques have allowed for the description of charge excitations, i.e., quasiparticles (QPs) [1,2], in the bulk over a wide range of QP energies. Band edge excitations are well described by the so-called G_0W_0 approximation [3–5], whereas at higher QP energies corrections are required to account for charge-density fluctuations and hole-plasmon coupling [5–8]. Photoemission experiments on solids reveal significant QP lifetime shortening and coupling to other collective excitations, manifested by satellite structures in the photoemission spectra [6,9,10]. The satellite structure and the QP lifetime shortening is often captured by the cumulant expansion (CE) ansatz to G_0W_0 [6–8,10–14].

In confined systems, the QP spectrum near the band edge is governed by the quantum confinement of electrons and holes. Higher-energy satellite excitations are attributed to simultaneous ionization and excitation of the valence electrons ("shake-up" excitations) [13,15–18]. Transition and differences between the satellite spectral features of molecules and nanostructures with shake-up signatures and bulk with collective plasmon resonances have been difficult to assess as they require many-body treatment of systems with hundreds and thousands of electrons. In fact, the quantum confinement effect on the satellite transitions has received little attention if any.

In this Rapid Communication, we address this challenge by combining the well-known CE ansatz [5,7–9,11,19,20] with the recent stochastic GW approach (sGW [21,22]) to obtain a nearly linear-scaling algorithm that reveals the changes in the QP spectra from a single molecule to covalently bonded nanocrystals (NCs) of unprecedented size (here up to 5288 valence electrons). The formalism is presented and assessed for the two size extremes (molecules and large bulklike systems), followed by the study of the effects of quantum confinement on the satellite structure in silicon NCs of different sizes and shapes. In small NCs the satellite features are affected by the changes in the plasmon energy. For large NCs, we find observable quantum confinement effects on the satellite features below the exciton Bohr radius where the position of the satellite peak and the QP lifetime show small dependence on the size of the system.

The central theoretical quantity for quasiparticles is the spectral function, which in the sudden approximation is directly linked to the photoemission current [5,23,24]. The spectral function of the ith QP state is $A_i(\omega) = \frac{1}{\pi}\mathrm{Im}\, G_i(\omega)$ where the GF fulfills the Dyson equation $G_i(\omega) = G_i^{(0)}(\omega) + G_i^{(0)}(\omega)\Sigma_i(\omega)G_i^{(0)}(\omega) + \cdots$, where $G_i^{(0)}(\omega)$ is the noninteracting GF and Σ_i is the self-energy. All the quantities are nonlocal in space, and all the higher terms in the equation represent a convolution integral, but for brevity we omit the spatial dependence in the notation.

As usual, the noninteracting system is described by the Kohn-Sham (KS) density functional theory (DFT) [25,26] (see the details in Ref. [27]) with Troullier-Martins pseudopotentials. The sGW approach is detailed in Refs. [21,22]). The self-energy is then given in the diagonal G_0W_0 approximation as [3]: $\tilde{\Sigma}_i(t) = i\langle\phi_i|\tilde{G}_i^{(0)}(t)\tilde{W}(t^+)|\phi_i\rangle$, where t^+ is infinitesimally after t, ϕ_i is the KS eigenstate, $W(\omega) = \epsilon^{-1}(\omega)v_c$, v_c is the Coulomb kernel, and $\epsilon^{-1}(\omega)$ is the inverse dielectric function. Quantities in frequency and time domains (e.g., G and $\tilde{G}$) are simply related by their Fourier transforms. From the calculated $\Sigma_i(\omega)$ the G_0W_0 spectral function is given by

$$A_i^{GW}(\omega) = \frac{1}{\pi}\frac{|\mathrm{Im}\,\Sigma_i(\omega)|}{[\omega - \varepsilon_i - \mathrm{Re}\,\Sigma_i(\omega) + \overline{v}_{XC}]^2 + [\mathrm{Im}\,\Sigma_i(\omega)]^2}, \quad (1)$$

*vojtech@chem.ucla.edu
†eran.rabani@berkeley.edu
‡dxn@ucla.edu

where ε_i is the KS eigenstate energy and $\overline{v}_{XC}$ is the expectation value of the mean-field exchange-correlation potential. A_i^{GW} has peaks at the quasiparticle energies ε_i^{qp} that fulfill the fixed-point equation,

$$\varepsilon_i^{qp} = \varepsilon_i + \mathrm{Re}\,\Sigma_i\left(\omega = \varepsilon_i^{qp}\right) - \overline{v}_{XC}. \quad (2)$$

In this GW approximation, the inverse lifetime of the QP is given by Im Σ at the peak. However, the actual plasmon-hole coupled excitations are not in general represented by the isolated poles in Eq. (1), and $A^{GW}(\omega)$ thus does not yield a proper description of satellite structures. In addition, spurious secondary peaks arise if Eq. (2) has multiple solutions [6,7,20].

The CE formulation is required to account for the effect of hole-plasmon coupling. For the ith occupied state, the GF in the CE *ansatz* reads [5,9,19]

$$\tilde{G}_i(t) = -ie^{i\varepsilon_i t}e^{C_i(t)}\theta(-t) = -ie^{i\varepsilon_i t + C_i^{qp}(t)}e^{C_i^s(t)}\theta(-t), \quad (3)$$

where C_i is the cumulant, obtained from the Dyson series expansion. Furthermore, following Ref. [6] the cumulant contribution is separated into two components. The first is a QP cumulant C_i^{qp}, derived explicitly in Ref. [6] and associated with a portion of the spectral function describing the main QP peak,

$$A_i^{qp}(\omega) = \frac{\mathbf{Z}_i}{\pi}\frac{\left|\mathrm{Im}\,\Sigma\left(\varepsilon_i^{qp}\right)\right|}{\left(\omega - \varepsilon_i^{qp}\right)^2 + \left[\mathrm{Im}\,\Sigma\left(\varepsilon_i^{qp}\right)\right]^2}, \quad (4)$$

where the renormalization factor due to redistribution of the spectral weight into the satellite peaks is $\mathbf{Z}_i = e^{\alpha_i}$ with $\alpha_i = \frac{\partial\Sigma(\omega)}{\partial\omega}|_{\omega=\varepsilon_i}$. The lifetime of the QP is $1/|\mathrm{Im}\,\Sigma(\varepsilon_i^{qp})|$.

By itself, $A_i^{qp}(\omega)$ does not include any satellite contributions—it is a single Lorentzian-shaped peak around the QP energy. The satellite peaks stem from resonances identified as poles in $W(\omega)$ [i.e., zeros of $\epsilon(\omega)$] and appear as strong maxima in the imaginary part of the self-energy; they are introduced by $e^{C_i^s(t)}$ [Eq. (3)] that derives from the spectral representation of Σ_P [6,9,19],

$$C_i^s(t) = \frac{1}{\pi}\lim_{\eta\to 0}\int_{-\infty}^{\mu}\frac{\mathrm{Im}\,\Sigma_P(\omega)e^{-i(\omega-\varepsilon_i+i\eta)t}}{(\omega-\varepsilon_i+i\eta)^2}d\omega. \quad (5)$$

We solve for $C_i^s(t)$ within the stochastic framework using Σ_P, which is obtained from the stochastic G_0W_0 calculation in the time domain. This is distinct from the previous calculations that either employed time-dependent DFT for core electrons [28] or employed the G_0W_0 approximation for the cumulant in the frequency domain. Due to computational cost, the latter is limited to small periodic systems, and $C_i^s(t)$ was in many cases further approximated by a single excitation pole in Σ_P. Here, the self-energy is obtained in a high-frequency resolution ($<0.01E_h$) on a wide range of frequencies $\pm 60E_h$; no further approximation is introduced.

The computed satellite cumulant $C_i^s(t)$ is inserted into Eq. (3), which is Fourier transformed to yield $G_i(\omega)$ and thereby $A_i(\omega) = \mathrm{Im}[G_i(\omega)]$ [29].

We next verify our approach using a large NC $Si_{705}H_{300}$, that is close to the bulk limit. Figure 1 shows the spectral function

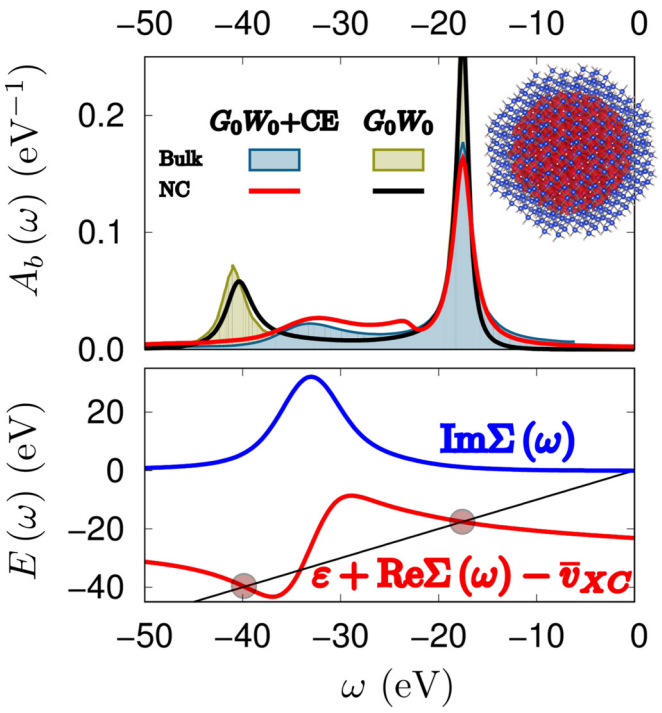

FIG. 1. Top: Spectral function for the bottom valence-band (VB) $A_b(\omega)$ for a bulk solid from Ref. [10] and for the $Si_{705}H_{300}$ NC. The G_0W_0 + cumulant spectral function (red) has an asymmetric satellite peak at the maximum of Im $\Sigma(\omega)$. The G_0W_0 prediction (black) has an artificial second maximum at low energies due to a spurious additional solution of Eq. (2). The inset plots the structure of the nanocrystal and orbital density isosurface (red; Si and H are the blue and white circles). Bottom: Graphical solutions to the QP equation, marked with red circles, are found at the intersection of the red curve [ε + Re $\Sigma(\omega)$ − $\bar{v}_{XC}$] with the diagonal ω line.

of the bottom of the valence band (VB—denoted A_b^{GW}) with a pronounced QP peak at −17.5 eV. If a cumulant expansion is not used, A_b^{GW} shows an additional maximum at −39.8 eV. This is in excellent agreement with previous GW calculations for bulk systems but is not observed experimentally and is attributed to spurious secondary solutions to Eq. (2) [6–8,10,11].

With the cumulant GW [Eq. (3)] the spectrum changes drastically, and an additional peak is obtained at −32.3 eV in excellent agreement with a result for bulk Si [10]. This peak is physically meaningful as it corresponds to the maximum of Im $\Sigma(\omega)$ associated with a collective excitation of the valence electrons (plasmon). The appearance of the satellite structure is accompanied by a reduction of the intensity of the main QP peak so that the renormalization factor is $\mathbf{Z} = 0.61$, i.e., 39% of the intensity is transferred to the satellites. The additional shoulder in the satellite peak is due to the difference between the effective masses of the QP and the plasmon [12]. The pronounced transfer of the spectral weight to the plasmon satellite for the bottom valence excitations is a consequence of their high energy and spatial extent (leading to large overlaps with other states). An isosurface of the bottom valence orbital of $Si_{705}H_{300}$ indeed exhibits spherical symmetry and lacks nodal planes as seen from the inset in Fig. 1.

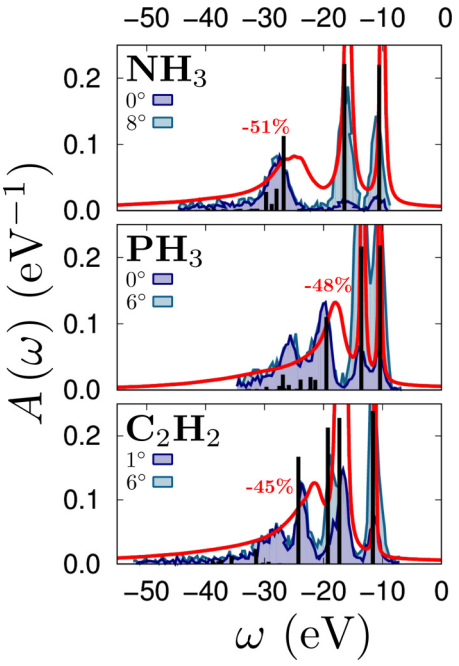

FIG. 2. Spectral functions (the red line) from stochastic G_0W_0 + CE for C_2H_2, PH_3, and NH_3. The spectral weight loss from the bottom valence state to the satellites is shown above the peak. Symmetry-adapted-cluster configuration-interaction (SAC-CI) [30,31] results are shown by the vertical black lines with height proportional to the relative intensities; the colored areas refer to the experimental photoemission spectra for two relative azimuthal angles [31,32].

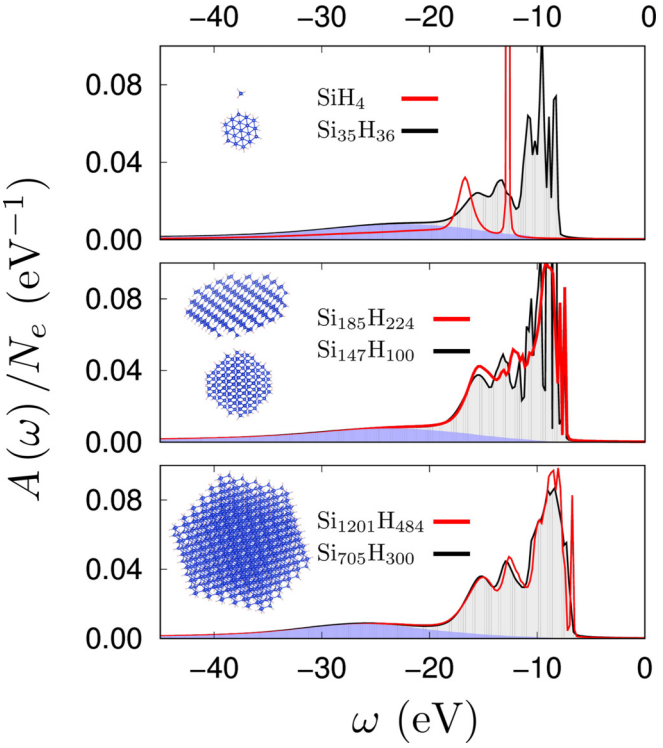

FIG. 3. Spectral functions for silicon nanocrystals, a platelet (middle panel: red) and a silane molecule. Satellite contributions to the spectral function are shown by a blue shaded area. Spectral functions are divided by the number of electrons N_e.

To further test our approach on finite systems, we applied (Fig. 2) the stochastic G_0W_0 approximation with CE to a series of small molecules for which experimental photoemission spectra are available. The results in Fig. 2 were further scaled so that the bottom valence state peak has the same intensity as the G_0W_0 + CE curve. The G_0W_0 + CE description relies on the concept of a plasmon associated with charge excitations in extended systems. Yet, Eq. (5) is applicable to small systems and yields good results. The stochastic GW with damped real-time propagation of the exited state [21,22] is in qualitative agreement with experiment and with high level SAC-CI calculations, computationally feasible for small molecules [33]. We note that:

(i) The QP energies at the top valence band are captured well by G_0W_0. This is the energy region where DFT is a good starting point. But G_0W_0 fails to reproduce the bottom VB where it underestimates the position of the peaks by a significant amount of 2 eV. For these states, DFT is not a good starting point, and the "single-shot" G_0W_0 procedure is inaccurate.

(ii) Most importantly, the G_0W_0 + CE description captures the satellite overall decay although without the fine-structure peaks in the satellite region. The pronounced satellite spectral weight comes at the expense of the QP peaks which transfer up to 51% of their intensity to the satellite tails. The broadening of the satellite peaks observed in G_0W_0 + CE is a consequence of an intrinsic decay of the density-density correlation function in time (τ). The peak width is independent of the maximal time used to simulate the screening (varied between 1 and 24 fs without affecting the lifetime). An infinite propagation time would result in the presence of many sharp poles due to recurrences in the correlation function. As clearly can be seen in Fig. 2, finite τ yields a photoelectron spectrum in good agreement with experiment, likely due to other mechanisms suppressing the recurrences in photoelectron spectroscopy.

Furthermore, the G_0W_0 + CE spectral function has maxima that are shifted with respect to the G_0W_0 QP energies. The shift is large for the bottom VB; e.g., for NH_3 the G_0W_0 peak is at -25.0 eV, whereas the G_0W_0 + CE maximum is at -25.7 eV. The 0.7-eV difference is significant as it is 17% of the GW correction to the LDA energy (-20.8 eV). Thus, the usual practice where G_0W_0 results are directly compared to the photoionization experiment is problematic, especially for low-energy states as it does not include the coupling of these states to the shake-up excitations.

In the next main part of this Rapid Communication we investigate the evolution of the spectral function with system size; the results for a series of Si NCs are shown in Fig. 3. All NCs exhibit a discrete and narrow spectrum near the top of the VB. Due to the quantum confinement effect, the top of the VB shifts to higher energies with increasing size; the highest occupied state has energies of -8.1 and -6.4 eV for $Si_{35}H_{36}$ and for $Si_{1201}H_{484}$, respectively. For deeper hole excitations, the sharp features merge into a semicontinuous spectral response with significant lifetime shortening. This is accompanied by significant spectral weight transfer ($\sim$50%) to the satellites. The bottom of the VB depends weakly on the system size, spanning an energy between -17.3 and -17.7 eV

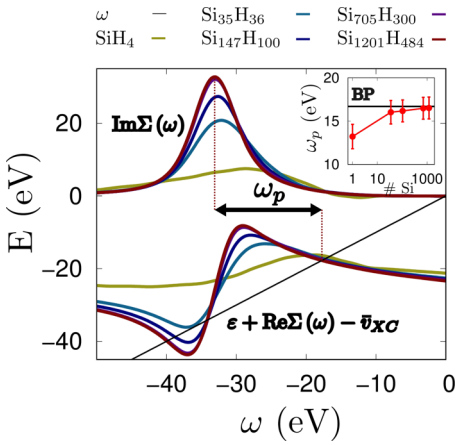

FIG. 4. Results for the bottom VB of Si nanocrystals of different sizes: The upper lines (above $E = 0$) show the imaginary part of the self-energy. The lower curves below $E = 0$ give the shifted real part of the self-energy and represent the graphical solution to Eq. (2) (the QP energy is the intersection with the frequency line in black). The average energy of the plasmon resonance (ω_p) is shown in the inset with error bars indicating a dispersion of the values between bottom and top valence states; the experimental value of the bulk plasmon (Ref. [35]) is indicated by a black horizontal line.

for the range of NCs studied. The QP peak also overlaps with the emerging satellite, which is already well developed into its bulk shape for $Si_{35}H_{36}$ and found in the range typical for bulk silicon [7]. This result is rather surprising since both the QP spectrum near the band edge and the plasmonic excitations are sensitive to the system size. We further observe that the dimensionality does not strongly affect the main QP peaks: The silicon platelet has ∼60% of the Si atoms on the surface, yet its spectral function is similar to the NCs.

On closer inspection, we observe that the satellite maximum exhibits nonmonotonic shifts: First it shows a strong decrease in energy for systems from $Si_{35}H_{36}$ to $Si_{705}H_{300}$ (from −22.5 to −26.1 eV, respectively), which is followed by a slight move back to higher energies by 0.6 eV. The initial regime stems from the decrease in the plasmon resonance frequency (ω_p) discussed below in detail. Once ω_p converges, the satellite maximum follows the changes in the QP density of states of the valence states governed by quantum confinement, i.e., the spectrum moves to higher energies (cf., Fig. 3).

In Fig. 4 we show Im $\Sigma_P(\omega)$ together with the graphical solution to Eq. (2) (which also depicts spurious secondary solutions found already for $Si_{35}H_{36}$ at −36 eV). The plasmon peak in the Im $\Sigma_P(\omega)$ curve changes until the asymptotic limit is reached; ultimately the curves for NCs with 3120 electrons ($Si_{705}H_{300}$) and 5288 electrons ($Si_{1201}H_{484}$) have practically identical heights, widths, and positions. The distance between the maximum of Im $\Sigma_P(\omega)$ and the QP energy corresponds to ω_p coupled to the bottom valence hole; for the largest system $\omega_p = 15.3$ eV. Convergence of ω_p with system size is shown in the inset for the top and the bottom VBs.

Unlike in solids, the holes in finite systems couple to two types of plasmon resonances: low-energy surface plasmons and high-energy bulk plasmons. For small NCs, both contribute and lead to a broad peak in Im $\Sigma_P(\omega)$. The surface plasmon resonances also strongly contribute in low-dimensional structures—the plasmon satellite of the platelet has a maximum at −22.6 eV, which is almost identical to the smallest NC ($Si_{35}H_{36}$). A similar strong shift of satellites and enhancement of surface plasmon signatures in low-dimensional systems were demonstrated experimentally [34]. For big systems, the hole becomes more localized inside the NC (cf., the inset of Fig. 1), and the coupling to the bulk plasmon dominates, leading to larger ω_p's. The distribution of the resonances becomes narrower, and the peak in Im $\Sigma_P(\omega)$ decreases in width.

To summarize, our work presents first principle theoretical predictions of the photoemission spectra, quasiparticle energies, and lifetimes covering the wide region between molecules and bulklike systems. The calculations show that the QP energies gradually increase with system size, and this is accompanied by changes in the position of the satellite peaks which correspond to a simultaneous ionization of the system and creation of a collective (shake-up or plasmon) excitation. The characteristic frequency of the plasmon has a narrower energy distribution in comparison to the shake-up, but both are similar in nature and significantly alter the spectrum at low energies. Furthermore, we have shown that for small systems the satellite region merges with the QP peak and shifts the apparent photoemission peak maximum to lower energies. The QP energies and photoemission maxima thus differ for the systems studied by as much as 0.6 eV.

The position of the satellite region is dictated by the QP energies and the frequency of the collective excitation. For small and low-dimensional systems, surface and bulk plasmon resonances contribute to the satellites. With increasing size the higher-energy bulk-plasmon coupling dominates. For small and intermediate systems, the maximum of the satellite decreases in energy and is affected by the plasmon resonance energy. For big systems, the maximum shows a slight increase due to changes in the main part of the QP spectrum.

This Rapid Communication was supported by the Center for Computational Study of Excited-State Phenomena in Energy Materials (C2SEPEM) at the Lawrence Berkeley National Laboratory, which is funded by the U.S. Department of Energy, Office of Science, Basic Energy Sciences, Materials Sciences and Engineering Division under Contract No. DE-AC02-05CH11231 as part of the Computational Materials Sciences Program. The calculations were performed as part of the XSEDE computational Project No. TG-CHE160092 [36]. 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.

[1] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-Particle Systems* (Dover, Mineola, NY, 2003).

[2] E. K. U. Gross, E. Runge, and O. Heinonen, *Many-Particle Theory* (Hilger, London, 1991).

[3] L. Hedin, New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem, Phys. Rev. **139**, A796 (1965).

[4] M. S. Hybertsen and S. G. Louie, Electron correlation in semiconductors, and, insulators: Band, gaps, and, quasiparticle energies, Phys. Rev. B **34**, 5390 (1986).

[5] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting Electrons* (Cambridge University Press, Cambridge, UK, 2016).

[6] F. Aryasetiawan, L. Hedin, and K. Karlsson, Multiple Plasmon Satellites in Na, and Al Spectral Functions from *Ab Initio* Cumulant Expansion, Phys. Rev. Lett. **77**, 2268 (1996).

[7] M. Guzzo, G. Lani, F. Sottile, P. Romaniello, M. Gatti, J. J. Kas, J. J. Rehr, M. G. Silly, F. Sirotti, and L. Reining, Valence Electron Photoemission Spectrum of Semiconductors: *Ab Initio* Description of Multiple Satellites, Phys. Rev. Lett. **107**, 166401 (2011).

[8] J. J. Kas, J. J. Rehr, and L. Reining, Cumulant expansion of the retarded one-electron Green function, Phys. Rev. B **90**, 085112 (2014).

[9] D. C. Langreth, Singularities in the X-ray spectra of metals, Phys. Rev. B **1**, 471 (1970).

[10] J. Lischner, G. K. Pálsson, D. Vigil-Fowler, S. Nemsak, J. Avila, M. C. Asensio, C. S. Fadley, and S. G. Louie, Satellite band structure in silicon caused by electron-plasmon coupling, Phys. Rev. B **91**, 205113 (2015).

[11] F. Caruso, H. Lambert, and F. Giustino, Band Structures of Plasmonic Polarons, Phys. Rev. Lett. **114**, 146404 (2015).

[12] D. Vigil-Fowler, S. G. Louie, and J. Lischner, Dispersion, and line shape of plasmon satellites in one, two, and three dimensions, Phys. Rev. B **93**, 235446 (2016).

[13] J. McClain, J. Lischner, T. Watson, D. A. Matthews, E. Ronca, S. G. Louie, T. C. Berkelbach, and G. Kin-Lic Chan, Spectral functions of the uniform electron gas via coupled-cluster theory, and comparison to the GW, and related approximations, Phys. Rev. B **93**, 235139 (2016).

[14] M. Z. Mayers, M. S. Hybertsen, and D. R. Reichman, Description of quasiparticle, and satellite properties via cumulant expansions of the retarded one-particle Green's function, Phys. Rev. B **94**, 081109 (2016).

[15] J. Schirmer, L. S. Cederbaum, and O. Walter, New approach to the one-particle Green's function for finite Fermi systems, Phys. Rev. A **28**, 1237 (1983).

[16] L. S. Cederbaum, W. Domcke, J. Schirmer, and W. Von Niessen, Correlation effects in the ionization of molecules: Breakdown of the molecular orbital picture, Adv. Chem. Phys. **65**, 115 (1986).

[17] M. S. Deleuze and L. S. Cederbaum, Correlation effects in the valence x-ray photoionization spectra of ethylene, butadiene, and hexatriene, Int. J. Quantum Chem. **63**, 465 (1997).

[18] A. I. Krylov, in *Reviews in Computational Chemistry*, edited by A. L. Parrill and K. B. Lipkowitz (Wiley, Hoboken, NJ, 2017), Vol. 30.

[19] B. I. Lundqvist, Single-particle spectrum of the degenerate electron gas, Phys. Kondens. Mater. **6**, 193 (1967).

[20] J. Lischner, D. Vigil-Fowler, and S. G. Louie, Physical Origin of Satellites in Photoemission of Doped Graphene: An *Ab Initio* GW Plus Cumulant Study, Phys. Rev. Lett. **110**, 146801 (2013).

[21] D. Neuhauser, Y. Gao, C. Arntsen, C. Karshenas, E. Rabani, and R. Baer, Breaking the Theoretical Scaling Limit for Predicting Quasiparticle Energies: The Stochastic GW Approach, Phys. Rev. Lett. **113**, 076402 (2014).

[22] V. Vlček, E. Rabani, D. Neuhauser, and R. Baer, Stochastic GW calculations for molecules, J. Chem. Theo. Comput. **13**, 4997 (2017).

[23] F. Aryasetiawan and O. Gunnarsson, The GW method, Rep. Prog. Phys. **61**, 237 (1998).

[24] G. Onida, L. Reining, and A. Rubio, Electronic excitations: Density-functional versus many-body Green's-function approaches, Rev. Mod. Phys. **74**, 601 (2002).

[25] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, Phys. Rev. **136**, B864 (1964).

[26] W. Kohn and L. J. Sham, Self-consistent equations including exchange, and correlation effects, Phys. Rev. **140**, A1133 (1965).

[27] The local density approximation (LDA) DFT calculation used a real-space grid ($0.6a_0$ for nanocrystals, $0.4a_0$ for PH_3, and NH_3, $0.5a_0$ for the C_2H_2) with Troullier-Martins pseudopotentials. The sGW approach is detailed in Refs. [21,22].

[28] J. J. Kas, F. D. Vila, J. J. Rehr, and S. A. Chambers, Real-time cumulant approach for charge-transfer satellites in x-ray photoemission spectra, Phys. Rev. B **91**, 121112 (2015).

[29] The quasiparticle energies and the spectral functions vary smoothly with frequency [V. Vlcek *et al.*, arXiv:1701.02023] for the extended system studied here. We thus compute the spectral function for several selected states and interpolate the result. The number of states is increased until the interpolation is converged to within 0.1 eV. To evaluate $A(\omega)$ in Fig. 3, a single Lorentzian peak was used to describe the satellites since only one satellite was observed. We found that a third-order polynomial fit to the quasiparticle energies and the parameters of the Lorentzian peaks are sufficient to yield QP energies and satellite positions within 0.4 eV, i.e., better than the resolution of the predicted spectral functions. For systems up to $Si_{705}H_{300}$, we found that calculations for five independent states provide converged results. For $Si_{1201}H_{484}$ three calculations were performed.

[30] H. Wasada and K. Hirao, Computational studies of satellite peaks of the inner-valence ionization of C_2H_4, C_2H_2, and H_2S using the SAC CI method, Chem. Phys. **138**, 277 (1989).

[31] M. Ishida, M. Ehara, and H. Nakatsuji, Outer-and inner-valence ionization spectra of NH_3, PH_3, and AsH_3: Symmetry-adapted cluster configuration interaction general-R study, J. Chem. Phys **116**, 1934 (2002).

[32] E. Weigold, K. Zhao, and W. Von Niessen, Study of the valence electronic structure of ethyne by electron momentum spectroscopy, and Green function methods, J. Chem. Phys. **94**, 3468 (1991).

[33] Similar to Ref. [31], the SAC-CI results were convoluted with a Lorentzian peak with 2-eV broadening.

[34] C. Bostedt *et al.*, Photoemission spectroscopy of germanium nanocrystal films, J. Electron Spectrosc. Relat. Phenom. **126**, 117 (2002).

[35] J. Stiebling, Optische Eigenschaften des einkristallinen Siliziums aus Elektronenenergieverlustmessungen, Z. Phys. **31**, 355 (1978).

[36] J. Towns, T. Cockerill, M. Dahan, I. Foster, K. Gaither, A. Grimshaw, V. Hazlewood, S. Lathrop, D. Lifka, and G. D. Peterson, XSEDE: Accelerating scientific discovery, Comput. Sci. Eng. **16**, 62 (2014).