Low-scaling GW algorithm applied to twisted transition-metal dichalcogenide heterobilayers

Maximilian Graml, 1,2 Klaus Zollner, 1 Daniel Hernangómez-Pérez, 2 Paulo E. Faria Junior, 1 and Jan Wilhelm 1,2,*

¹Institute of Theoretical Physics, University of Regensburg, 93053 Regensburg, Germany
²Regensburg Center for Ultrafast Nanoscopy (RUN), University of Regensburg, 93053 Regensburg, Germany
³Department of Molecular Chemistry and Materials Science,
Weizmann Institute of Science, Rehovot 7610001, Israel

The GW method is widely used for calculating the electronic band structure of materials. The high computational cost of GW algorithms prohibits their application to many systems of interest. We present a periodic, low-scaling and highly efficient GW algorithm that benefits from the locality of the Gaussian basis and the polarizability. The algorithm enables G_0W_0 calculations on a $MoSe_2/WS_2$ bilayer with 984 atoms per unit cell, in 42 hours using 1536 cores. This is four orders of magnitude faster than a plane-wave G_0W_0 algorithm, allowing for unprecedented computational studies of electronic excitations at the nanoscale.

Electronic excitations in matter play a pivotal role in various physical phenomena, including light absorption and transport. The characteristics of these excitations are strongly influenced by the host material. Excitons, which are bound electron-hole pairs, exhibit a remarkable and unusually strong electron-hole binding in low-dimensional semiconductors that have emerged in the last decade [1]. When stacking two atomically thin semiconductors on top of each other, the atomic alignment between the layers can exhibit periodic variations, leading to a new type of in-plane superlattice known as the moiré superlattice. Excitons in moiré structures have gained enormous attention recently [2–12] thanks to their highly unusual exciton properties which include spatial confinement due to the moiré potential [2], interlayer [5, 6], and intralayer charge transfer [9]. Furthermore, electronic properties of moiré lattices can be tuned by the band alignment and the twist angle between the layers such that moiré structures hold great promise as an exciting platform for probing electronic and photonic quantum phenomena over the next decade [12].

Gaining insights into excitons in moiré structures can be achieved through a combination of experiments, theoretical models, and computations. As an example, low-angle MoSe₂/WS₂ moiré structures have shown an interesting interplay of intra- and interlayer exciton hybridization because of the nearly degenerate conduction bands of the MoSe₂ and WS₂ layers. The conduction band offset and the wavefunction hybridization between layers, however, is still under debate [3, 7, 13–15]. Detailed knowledge about the electronic band structure of the MoSe₂/WS₂ moiré bilayer and the implication on exciton formation and binding is thus crucial to resolve this controversy.

In this work, we focus on the GW method from many-body-perturbation theory [16–18] which is an approximation for the electronic self-energy that allows for computing the electronic band structure of a given material. Importantly, GW accounts for the nonlocal, frequency-dependent screening of the interaction between electrons which is crucial in moiré bilayers.

The GW band structure is then the basis for the description of excitons via the Bethe-Salpeter equation [17, 19]. Currently available plane-wave-based GW algorithms are however incapable of treating low-angle moiré cells that contain thousands of atoms [20], despite their computational scalability to the largest supercomputers [21–25]. Stochastic GW methods may enable large-scale GW calculations [26, 27], but it is not clear whether the numerical precision of this approach is sufficient for its application across the whole chemical space [28]. For computing the GW band structure in large moiré cells, pristine unit-cell matrix projection (PUMP) has been suggested [9, 20]. PUMP is based on expanding the moiré cell wavefunctions in terms of the pristine unit-cell wavefunctions. By construction, PUMP cannot capture nanometer-scale atomic reconstruction of moiré structures which can dramatically influence their electronic band structure [4].

The GW space-time method [29] offers a promising route towards large-scale GW calculations. This is because the computational scaling is reduced from $O(N_{\rm at}^4 N_k^2)$ for standard GW algorithms to $O(N_{\rm at}^3 N_k)$ in the GW space-time method, where $N_{\rm at}$ is the number of atoms in the unit cell and N_k the number k-points used to discretize the Brillouin zone. For achieving the scaling reduction, it is required to use a spatially local basis instead of plane waves. The local basis can be chosen as real-space grid where studies of unit cells up to one hundred atoms have been reported [29, 30]. Another choice of the spatially local basis is an atomic-orbital-like basis [31]. This choice is highly efficient in the GW space-time method enabling GW calculations on molecules with more than 1000 atoms [32–36].

Periodic boundary conditions in the GW space-time method with atomic-orbital-like basis functions have not been reported yet. The main inhibiting factor has been the inclusion of k-dependent Coulomb interactions which represent a major challenge regarding computational efficiency and numerical precision [37–39]. In this work, we overcome this challenge by employing real space representations of the polarizability, the screened Coulomb interaction and the self-energy. The

real-space representation allows us to use the minimum image convention (MIC) [40, 41], i.e., each atomic orbital in the simulation interacts only with the closest image of another atomic orbital. We benchmark the algorithm on G_0W_0 bandgaps of monolayer MoS_2 , $MoSe_2$, WS_2 , and WSe_2 finding an average deviation of only 0.06 eV from reference calculations [42, 43]. We also apply the GW algorithm to a $MoSe_2/WS_2$ bilayer with an unprecedented cell size of 984 atoms which has an order of magnitude more atoms than previous state-of-the-art large-scale GW calculations [25].

We start with details on the main algorithmic advances for achieving large-scale *GW* calculations on two-dimensional semiconductors. The full *GW* algorithm is given in the Supporting Information.

Following our previous work [35, 44], we compute the irreducible polarizability $\chi_{PQ}(\mathbf{k}=\mathbf{0}, i\tau)$ in imaginary time $i\tau$ at the Γ -point in an auxiliary atomic-orbital-like Gaussian basis set with indices P, Q. The polarizability $\chi_{PQ}(\mathbf{k}, i\tau)$ is however needed on a dense k-point mesh because it is later multiplied with the bare Coulomb interaction that diverges at the Γ -point and thus requires a fine k-point sampling. The atom-centered basis allows us to decompose the Γ -point result, $\chi_{PQ}(\mathbf{k}=\mathbf{0}, i\tau)$, using the identity

$$\chi_{PQ}(\mathbf{k} = \mathbf{0}, i\tau) = \sum_{\mathbf{R}} \chi_{PQ}^{\mathbf{R}}(i\tau) , \ \chi_{PQ}^{\mathbf{R}} = \langle \varphi_P^{\mathbf{0}} | \chi | \varphi_Q^{\mathbf{R}} \rangle , \quad (1)$$

where $\chi_{PQ}^{\mathbf{R}}$ is the real-space representation of the polarizability and $\varphi_P^{\mathbf{R}}$ denotes a Gaussian which is localized in cell \mathbf{R} . For non-metallic systems, the polarizability $\chi(\mathbf{r},\mathbf{r}',i\tau)$ is space-local, i.e. $\chi(\mathbf{r},\mathbf{r}',i\tau)$ exponentially decays with increasing $|\mathbf{r}-\mathbf{r}'|$. [45, 46] The matrix element $\chi_{PQ}^{\mathbf{R}}$ thus vanishes in case of a large distance between the center of $\varphi_P^{\mathbf{0}}$ and the center of $\varphi_Q^{\mathbf{R}}$. We employ MIC, i.e., we assume that $\chi_{PQ}^{\mathbf{R}}(i\tau)$ in Eq. (1) is non-zero only if the atomic center of $\varphi_P^{\mathbf{0}}$ and the atomic center of $\varphi_Q^{\mathbf{R}}$ are closest together among all cells \mathbf{R} . In this way, we extract $\chi_{PQ}^{\mathbf{R}}(i\tau)$ from Eq. (1),

$$\chi_{PQ}^{\mathbf{R}}(i\tau) = \begin{cases} \chi_{PQ}(\mathbf{k}=\mathbf{0}, i\tau) & \text{if } \varphi_{P}^{\mathbf{0}}, \varphi_{Q}^{\mathbf{R}} \text{ closest}, \\ 0 & \text{else}, \end{cases}$$
 (2)

which is exact in the limit of a large, non-metallic unit cell. Using Eq. (2), we obtain the polarizability at any k-point at negligible computational cost,

$$\chi_{PQ}(\mathbf{k}, i\tau) = \sum_{\mathbf{R}} e^{-i\mathbf{k}\cdot\mathbf{R}} \chi_{PQ}^{\mathbf{R}}(i\tau). \tag{3}$$

Following the *GW* space-time method [29], we compute the screened interaction in real space (full algorithm in SI),

$$W_{PQ}^{\mathbf{R}}(i\tau) := \langle \varphi_P^{\mathbf{0}} | W(i\tau) | \varphi_Q^{\mathbf{R}} \rangle , \qquad (4)$$

leading to the self-energy $\Sigma(\mathbf{r}, \mathbf{r}', i\tau) = iG(\mathbf{r}, \mathbf{r}', i\tau)W(\mathbf{r}, \mathbf{r}', i\tau)$ [29]. $\Sigma(\mathbf{r}, \mathbf{r}', i\tau)$ is space-local as $G(\mathbf{r}, \mathbf{r}', i\tau)$ is space-local [45] and only elements of $W(\mathbf{r}, \mathbf{r}', i\tau)$ with small $|\mathbf{r} - \mathbf{r}'|$ contribute to Σ . We thus continue with the minimum image of Eq. (4)

$$W_{PQ}^{\text{MIC}}(i\tau) := W_{PQ}^{\mathbf{R}_{PQ}^{\text{min}}}(i\tau), \qquad (5)$$

where the cell vector

$$\mathbf{R}_{PQ}^{\min} = \underset{\mathbf{R}}{\operatorname{argmin}} \left| \mathbf{R}_{P} - (\mathbf{R}_{Q} + \mathbf{R}) \right| \tag{6}$$

gives the smallest distance between the atomic centers \mathbf{R}_P of φ_P^0 and the atomic center $\mathbf{R}_Q + \mathbf{R}$ of $\varphi_Q^{\mathbf{R}}$. We use $\mathbf{W}^{\text{MIC}}(i\tau)$ to calculate the self-energy $\Sigma_{\mu\nu}(\mathbf{k}=0,i\tau)$ in the atomic-orbital basis $\{\phi_{\mu}(\mathbf{r})\}$ at the Γ -point. We thus avoid k-point sampling in this computationally expensive step. k-points in Σ follow from MIC at negligible computational cost, cf. Eqs. (2), (3),

$$\Sigma_{\mu\nu}(\mathbf{k}, i\tau) = \sum_{\mathbf{R}} e^{i\mathbf{k}\cdot\mathbf{R}} \cdot \begin{cases} \Sigma_{\mu\nu}(\mathbf{k}=\mathbf{0}, i\tau) & \text{if } \phi_{\mu}^{\mathbf{0}}, \phi_{\nu}^{\mathbf{R}} \text{ closest}, \\ 0 & \text{else}. \end{cases}$$
(7)

We transform the self-energy to real energy [18] and the Bloch basis which allows us to compute quasiparticle energies $\varepsilon_{nk}^{G_0W_0}$,

$$\varepsilon_{n\mathbf{k}}^{G_0W_0} = \varepsilon_{n\mathbf{k}} + \operatorname{Re} \Sigma_{n\mathbf{k}} (\varepsilon_{n\mathbf{k}}^{G_0W_0}) - v_{n\mathbf{k}}^{\mathrm{xc}}, \tag{8}$$

where $v_{n\mathbf{k}}^{\text{xc}}$ is the diagonal of the exchange-correlation matrix.

The numerical trick in the presented GW algorithm is the MIC used in Eqs. (2), (5), and (7). MIC is exact in the limit of a large unit cell. We determine the critical cell size for the validity of MIC by computing the G_0W_0 bandgap of monolayer MoS_2 , $MoSe_2$, WS_2 , and WSe_2 , presented in Fig. 1. For the four materials, the bandgap changes on average by only 11 meV between the 10×10 supercell (300 atoms in the unit cell) and

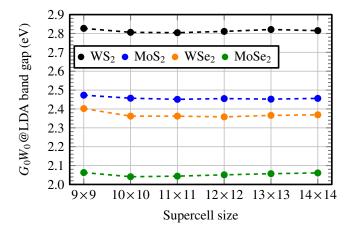


Figure 1. G_0W_0 bandgap of monolayer WS₂, MoS₂, WSe₂ and MoSe₂ calculated from Eq. (8) as function of the supercell size (TZVP-MOLOPT basis set [47], without spin-orbit coupling (SOC)).

Table I. G_0W_0 @PBE bandgap (in eV, without SOC) of monolayer WS₂, MoS₂, WSe₂ and MoSe₂ computed from Eq. (8) (TZV2P-MOLOPT basis [47], 10×10 supercell, detailed convergence test in the SI) and computed from plane-wave codes [39, 42, 43].

Software package	MoS_2	MoSe ₂	WS_2	WSe ₂
This work, (CP2K [48, 49])	2.47	2.07	2.81	2.37
GPAW [42] (SOC removed)	2.53	2.12	2.75	2.30
BerkeleyGW [39] (details in SI)	2.45	2.09	2.61	2.34
VASP [43]	2.50	2.06	2.70	2.34

the 14×14 supercell (588 atoms in the unit cell). We conclude that the GW algorithm from this work can be used to study unit cells which are as large as a 10×10 supercell or larger. In the Supporting Information, we show additional convergence tests on the basis set size, the number of time and frequency points, the k-point mesh size, filter threshold for sparse operations, and the vertical box height.

We compare the G_0W_0 bandgap of monolayer MoS₂, MoSe₂, WS₂, and WSe₂ to the G_0W_0 bandgap computed from three different plane-wave codes [39, 42, 43], see Table I. We find that our G_0W_0 bandgaps deviate on average by only 0.06 eV to the bandgaps from plane-wave based codes. This small discrepancy might be due to the use of different pseudopotentials and the difficulty to reach the complete-basis-set limit.

The presented algorithm has several computational advantages over plane-wave-based *GW* algorithms. The computational bottleneck in plane-wave-based *GW* algorithms is the calculation of the irreducible polarizability [21, 22, 50].

$$\chi_{\mathbf{G}\mathbf{G}'}(\mathbf{q}, i\omega) = \sum_{n}^{\text{occ}} \sum_{n'}^{\text{empty}} \sum_{\mathbf{k}} \frac{1}{\varepsilon_{n\mathbf{k}+\mathbf{q}} - \varepsilon_{n'\mathbf{k}} + i\omega}$$

$$\times \langle n\mathbf{k} + \mathbf{q} | e^{i(\mathbf{q} + \mathbf{G}) \cdot \mathbf{r}} | n'\mathbf{k} \rangle \langle n'\mathbf{k} | e^{-i(\mathbf{q} + \mathbf{G}') \cdot \mathbf{r}} | n\mathbf{k} + \mathbf{q} \rangle ,$$
(9)

where G, G' are reciprocal lattice vectors characterizing the plane wave $e^{i\mathbf{G}\cdot\mathbf{r}}$, \mathbf{q} is a vector in the first Brillouin zone, n, n' refer to occupied and empty bands, respectively, and the brackets in the second line denote integrals of a plane wave and Bloch states. The matrix in Eq. (9) is evaluated up to $|G^2| < |E_{cut}|$ for both ${\bf G}$ and ${\bf G}'$ where $E_{\rm cut}$ is the dielectric energy cutoff. We calculate the number of floating point operations necessary to perform the multiplications in Eq. (9), see gray traces in Fig. 2. The estimate corresponds to the computational effort of a plane-wave based G_0W_0 algorithm for 2D semiconductors and is based on realistic numerical parameters used in large-scale G_0W_0 calculations [25, 39]. We also report the required number of operations of our presented G_0W_0 algorithm in Fig. 2 (black traces). Our G_0W_0 algorithm requires a similar number of floating point operations for a 9×9 supercell as a plane-wave G_0W_0 algorithm for a 2×2 supercell. For a

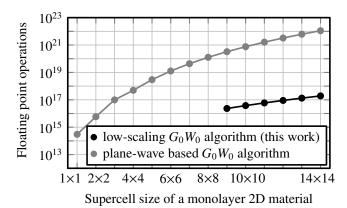


Figure 2. Number of floating point operations (real double precision) needed for executing G_0W_0 algorithms. Black: low-scaling G_0W_0 algorithm from this work using a TZVP-MOLOPT basis set [47], gray: plane-wave-based G_0W_0 algorithm, Eq. (9). Underlying computational parameters are typical for monolayer MoS_2 , $MoSe_2$, $MoSe_2$, and $MoSe_2$, see detailed raw data available in the SI.

 14×14 supercell, our G_0W_0 algorithm requires 10^5 times less operations compared to a plane-wave based algorithm. This large factor has several origins, most important are the following: The plane-wave basis $\{e^{i\mathbf{G}\cdot\mathbf{r}}\}$ resolves large vacuum regions [25, 39] for two-dimensional materials and is thus a factor 10 larger than the Gaussian auxiliary basis $\{\varphi_P\}$. We thus need to calculate 100 times less matrix elements of χ in a Gaussian basis compared to Eq. (9). Integrals over Gaussians, similar to the second line of Eq. (9), are sparse due to the spatial locality of Gaussians [35]. Only 3 % of the integrals need to be considered for a 14×14 supercell reducing the number of operations by another factor 30. In the present algorithm, χ is evaluated at the Γ -point using real-valued matrix algebra [35] which makes another factor 4 compared to the complex matrix algebra in Eq. (9). In Eq. (9), at least a 3×3 mesh for **q** is necessary [25] which is responsible for another factor of 5 [51]. These numerical parameters thus explain a factor of 60.000 between the required operations of a plane-wave G_0W_0 algorithm and the G_0W_0 algorithm from this work.

Further advantages compared to plane-wave based algorithms include the cheap diagonalization of the Kohn-Sham matrix to obtain Bloch states thanks to the small Gaussian basis. Also, non-periodic directions are easily dealt with in our GW algorithm by restricting the sum over cells \mathbf{R} to periodic directions. It is not necessary to truncate the Coulomb operator [39, 52] in non-periodic directions as in plane-wave algorithms. Moreover, the self-energy (7) is available in the Gaussian basis set which allows to compute the G_0W_0 correction for all Bloch states at low computational cost.

We measure the computation time of the algorithm, shown in Fig. 3. The computation time is moderate; as an example, a G_0W_0 calculation on the 10×10 MoSe₂ supercell (300 atoms)

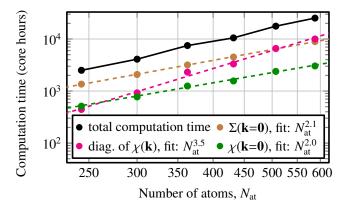


Figure 3. Execution time of a G_0W_0 calculation for MoSe₂ $9\times 9-14\times 14$ supercells (TZVP-MOLOPT basis set) on Supermuc-NG (Intel Skylake Xeon Platinum 8174). Magenta points show the computational cost to diagonalize the polarizability $\chi(\mathbf{k})$ which allows us to remove all spurious negative eigenvalues of $\chi(\mathbf{k})$ to ensure numerical stability. Dashed lines show a fit $\alpha N_{\rm at}^{\beta}$ to the execution time, where α and β are fit parameters. Raw data is available in the SI.

takes only 7 hours on 576 cores. Assuming ideal scalability starting from the 9×9 cell, we estimate that a G_0W_0 calculation on 4500 atoms is in reach [53]. Scalability improvements are subject of ongoing work to achieve this system size in practice.

We now focus on an application of the G_0W_0 algorithm to transition-metal dichalcogenide heterobilayers which recently gained increased attention due to twist-angle dependent moiré potentials and interlayer excitons [3, 4, 7–9, 12–15]. Recent large-scale plane-wave-based GW calculations on twisted heterostructures were limited to 75 atoms in the unit cell [25]. This GW computation [25] has been described to be highly cumbersome and it was only achieved owing to an advanced accelerated large-scale version of the BerkeleyGW code which scales to entire leadership high-performance computers with more than half a million cores [22, 23]. Small unit cells with 75 atoms only allow for the study of heterobilayers with selected, large twist angles and absent atomic reconstruction.

In order to illustrate the large-scale capabilities of our G_0W_0 algorithm beyond monolayers, we focus on the prototypical MoSe₂/WS₂ twisted heterostructures. On one hand, the different lattice parameters of MoSe₂ and WS₂ gives rise to a considerably large moiré periodicity at zero twist angle (\sim 8 nm), thus requiring a large number of atoms in the structure. On the other hand, low-angle MoSe₂/WS₂ have shown an interesting interplay of intra- and inter-layer exciton hybridization because of the nearly degenerate conduction bands. This feature, however, is still under debate in the literature [3, 7, 13–15]. The underlying electronic structure is thus crucial to resolve this controversy and is exactly the kind of problem that require large-scale GW calculations. Here we considered MoSe₂/WS₂ moiré superstructures with twist angles between

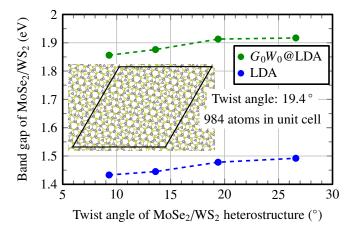


Figure 4. Band gap of a $MoSe_2/WS_2$ heterostructure as function of the twist angle. Inset: Unit cell (black rhomboid) for 19.4° twist angle contains 984 atoms.

9.3° and 26.6° (Fig. 4) that have corresponding unit cells of up to 984 atoms. We emphasize that in all cases, the strain of the individual monolayers is < 0.01% compared to the experimentally determined lattice constants [54, 55], which is important because the bandgap is very sensitive to strain [56, 57]. The G_0W_0 bandgap of the MoSe₂/WS₂ bilayer depends on the twist angle changing from 1.86 eV (9.3°) to 1.92 eV (26.8°), in line with experimental observations of the exciton emission energy [13]. Our GW calculation on the 984-atom heterostructure takes 42 hours on only 1536 cores which is a factor 30.000 faster than with a plane-wave algorithm, see estimate in the SI. Such large-scale GW calculations are an ideal starting point for further analyzing the electronic structure of these materials. For example, with our GW algorithm, the calculation of deep moiré potentials [4] are within reach, which are caused by atomic reconstruction and height variations. Both crucially influence the interlayer screening that is captured by the GW method. On top of a GW calculation, the Bethe-Salpeter equation [17, 19] will enable the study of excitons in large-scale moiré structures. Our computationally efficient scheme also holds great promise for nanoscale excited-state dynamics in low-dimensional materials. Current state-of-the-art studies only report the dynamics in clean monolayers [58–60] and models [61, 62].

Summarizing, we have presented a low-scaling GW algorithm with periodic boundary conditions employing localized basis functions and the minimum image convention. The GW algorithm is numerically precise and requires up to five orders of magnitude less floating point operations compared to plane-wave codes. We carried out a G_0W_0 calculation on a $MoSe_2/WS_2$ heterostructure with 984 atoms in the unit cell which is an order of magnitude more than the state of the art [25]. We are fully convinced that our GW algorithm will

enable routine applications of GW and its time-dependent variants to low-dimensional, nanostructured materials that were previously computationally highly challenging.

CODE AND DATA AVAILABILITY

The low-scaling GW algorithm is implemented in the open-source CP2K package [48] which is freely available from github [49]. Inputs and outputs of the calculations are also available on github [63].

ACKNOWLEDGMENT

We thank Mauro Del Ben, Maria Camarasa-Gomez, Ferdinand Evers, Dorothea Golze, Jürg Hutter, Ole Schütt and Shridhar Shanbhag for helpful discussions. K.Z. and P.E.F.J. acknowledge funding by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) SFB 1277 (Project No. 314695032, projects B07 and B11), SPP 2244 (Project No. 443416183), and the European Union Horizon 2020 Research and Innovation Program under contract number 881603 (Graphene Flagship). D. H.-P. acknowledges support from a Minerva Foundation grant 7135421, ERC Starting grant 101041159, and DFG through the SFB 1277 (No. 314695032, project B10). J.W. acknowledges funding by the DFG via the Emmy Noether Programme (Project No. 503985532). The Gauss Centre for Supercomputing is acknowledged for providing computational resources on SuperMUC-NG at the Leibniz Supercomputing Centre under the project ID pn72pa. The QUANTUM Espresso and BerkeleyGW computations were carried out in the Max Planck Computing and Data Facility cluster.

- * jan.wilhelm@physik.uni-regensburg.de
- [1] K. F. Mak, C. Lee, J. Hone, J. Shan, and T. F. Heinz, *Atomically Thin MoS*₂: *A New Direct-Gap Semiconductor*, Phys. Rev. Lett. **105**, 136805 (2010).
- [2] K. L. Seyler, P. Rivera, H. Yu, N. P. Wilson, E. L. Ray, D. G. Mandrus, J. Yan, W. Yao, and X. Xu, Signatures of moiré-trapped valley excitons in MoSe₂/WSe₂ heterobilayers, Nature 567, 66 (2019).
- [3] L. Zhang, Z. Zhang, F. Wu, D. Wang, R. Gogna, S. Hou, K. Watanabe, T. Taniguchi, K. Kulkarni, T. Kuo, S. R. Forrest, and H. Deng, *Twist-angle dependence of moiré excitons in* WS₂/MoSe₂ heterobilayers, Nat. Commun. 11, 5888 (2020).
- [4] S. Shabani, D. Halbertal, W. Wu, M. Chen, S. Liu, J. Hone, W. Yao, D. N. Basov, X. Zhu, and A. N. Pasupathy, *Deep moiré* potentials in twisted transition metal dichalcogenide bilayers, Nat. Phys. 17, 720 (2021).

- [5] O. Karni, E. Barré, V. Pareek, J. D. Georgaras, M. K. L. Man, C. Sahoo, D. R. Bacon, X. Zhu, H. B. Ribeiro, A. L. O'Beirne, J. Hu, A. Al-Mahboob, M. M. M. Abdelrasoul, N. S. Chan, A. Karmakar, A. J. Winchester, B. Kim, K. Watanabe, T. Taniguchi, K. Barmak, J. Madéo, F. H. da Jornada, T. F. Heinz, and K. M. Dani, Structure of the moiré exciton captured by imaging its electron and hole, Nature 603, 247 (2022).
- [6] D. Schmitt, J. P. Bange, W. Bennecke, A. AlMutairi, G. Meneghini, K. Watanabe, T. Taniguchi, D. Steil, D. R. Luke, R. T. Weitz, S. Steil, G. S. M. Jansen, S. Brem, E. Malic, S. Hofmann, M. Reutzel, and S. Mathias, Formation of moiré interlayer excitons in space and time, Nature 608, 499 (2022).
- [7] Y. G. Gobato, C. S. de Brito, A. Chaves, M. A. Prosnikov, T. Woźniak, S. Guo, I. D. Barcelos, M. V. Milošević, F. Withers, and P. C. M. Christianen, *Distinctive g-Factor of Moiré-Confined Excitons in van der Waals Heterostructures*, Nano Lett. 22, 8641 (2022).
- [8] E. Barré, O. Karni, E. Liu, A. L. O'Beirne, X. Chen, H. B. Ribeiro, L. Yu, B. Kim, K. Watanabe, T. Taniguchi, K. Barmak, C. H. Lui, S. Refaely-Abramson, F. H. da Jornada, and T. F. Heinz, Optical absorption of interlayer excitons in transition-metal dichalcogenide heterostructures, Science 376, 406 (2022).
- [9] M. H. Naik, E. C. Regan, Z. Zhang, Y.-H. Chan, Z. Li, D. Wang, Y. Yoon, C. S. Ong, W. Zhao, S. Zhao, M. I. B. Utama, B. Gao, X. Wei, M. Sayyad, K. Yumigeta, K. Watanabe, T. Taniguchi, S. Tongay, F. H. da Jornada, F. Wang, and S. G. Louie, *Intralayer charge-transfer moiré excitons in van der Waals superlattices*, Nature 609, 52 (2022).
- [10] P. Rivera, H. Yu, K. L. Seyler, N. P. Wilson, W. Yao, and X. Xu, Interlayer valley excitons in heterobilayers of transition metal dichalcogenides, Nat. Nanotechnol. 13, 1004 (2018).
- [11] C. Jin, E. Y. Ma, O. Karni, E. C. Regan, F. Wang, and T. F. Heinz, *Ultrafast dynamics in van der Waals heterostructures*, Nat. Nanotechnol. 13, 994 (2018).
- [12] D. Huang, J. Choi, C.-K. Shih, and X. Li, *Excitons in semiconductor moiré superlattices*, Nat. Nanotechnol. **17**, 227 (2022).
- [13] E. M. Alexeev, D. A. Ruiz-Tijerina, M. Danovich, M. J. Hamer, D. J. Terry, P. K. Nayak, S. Ahn, S. Pak, J. Lee, J. I. Sohn, M. R. Molas, M. Koperski, K. Watanabe, T. Taniguchi, K. S. Novoselov, R. V. Gorbachev, H. S. Shin, V. I. Fal'ko, and A. I. Tartakovskii, Resonantly hybridized excitons in moiré superlattices in van der Waals heterostructures, Nature 567, 81 (2019).
- [14] Y. Tang, J. Gu, S. Liu, K. Watanabe, T. Taniguchi, J. Hone, K. F. Mak, and J. Shan, *Tuning layer-hybridized moiré excitons by the quantum-confined Stark effect*, Nat. Nanotechnol. 16, 52 (2020).
- [15] X. Ma, S. Fu, J. Ding, M. Liu, A. Bian, F. Hong, J. Sun, X. Zhang, X. Yu, and D. He, Robust Interlayer Exciton in WS₂/MoSe₂ van der Waals Heterostructure under High Pressure, Nano Lett. 21, 8035 (2021).
- [16] L. Hedin, New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem, Phys. Rev. 139, A796 (1965).
- [17] G. Onida, L. Reining, and A. Rubio, *Electronic excitations: density-functional versus many-body Green's-function approaches*, Rev. Mod. Phys. **74**, 601 (2002).
- [18] D. Golze, M. Dvorak, and P. Rinke, *The GW Compendium:* A Practical Guide to Theoretical Photoemission Spectroscopy, Front. Chem. 7, 377 (2019).

- [19] X. Blase, I. Duchemin, D. Jacquemin, and P.-F. Loos, The Bethe–Salpeter Equation Formalism: From Physics to Chemistry, J. Phys. Chem. Lett. 11, 7371 (2020).
- [20] H. Li, Z. Xiang, M. H. Naik, W. Kim, Z. Li, R. Sailus, R. Banerjee, T. Taniguchi, K. Watanabe, S. Tongay, A. Zettl, F. H. da Jornada, S. G. Louie, M. F. Crommie, and F. Wang, *Imaging Moiré Excited States with Photocurrent Tunneling Microscopy*, arXiv 2306.00859 (2023).
- [21] D. Sangalli, A. Ferretti, H. Miranda, C. Attaccalite, I. Marri, E. Cannuccia, P. Melo, M. Marsili, F. Paleari, A. Marrazzo, G. Prandini, P. Bonfà, M. O. Atambo, F. Affinito, M. Palummo, A. Molina-Sánchez, C. Hogan, M. Grüning, D. Varsano, and A. Marini, *Many-body perturbation theory calculations using the yambo code*, J. Phys. Condens. Matter 31, 325902 (2019).
- [22] M. Del Ben, F. H. da Jornada, A. Canning, N. Wichmann, K. Raman, R. Sasanka, C. Yang, S. G. Louie, and J. Deslippe, *Large-scale GW calculations on pre-exascale HPC systems*, Comput. Phys. Commun. 235, 187 (2019).
- [23] M. Del Ben, C. Yang, Z. Li, F. H. d. Jornada, S. G. Louie, and J. Deslippe, in SC20: International Conference for High Performance Computing, Networking, Storage and Analysis (2020) pp. 1–11.
- [24] V. W.-Z. Yu and M. Govoni, GPU Acceleration of Large-Scale Full-Frequency GW Calculations, J. Chem. Theory Comput. 18, 4690 (2022).
- [25] S. Kundu, T. Amit, H. R. Krishnamurthy, M. Jain, and S. Refaely-Abramson, *Exciton fine structure in twisted transition metal dichalcogenide heterostructures*, arXiv 2209.05030 (2022).
- [26] 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).
- [27] J. Brooks, G. Weng, S. Taylor, and V. Vlček, *Stochastic many-body perturbation theory for Moiré states in twisted bilayer phosphorene*, J. Phys. Condens. Matter **32**, 234001 (2020).
- [28] V. Vlček, E. Rabani, D. Neuhauser, and R. Baer, Stochastic GW Calculations for Molecules, J. Chem. Theory Comput. 13, 4997 (2017).
- [29] H. N. Rojas, R. W. Godby, and R. J. Needs, *Space-Time Method for Ab Initio Calculations of Self-Energies and Dielectric Response Functions of Solids*, Phys. Rev. Lett. **74**, 1827 (1995).
- [30] P. Liu, M. Kaltak, J. Klimeš, and G. Kresse, *Cubic scaling GW: Towards fast quasiparticle calculations*, Phys. Rev. B 94, 165109 (2016).
- [31] M. Rohlfing, P. Krüger, and J. Pollmann, Efficient scheme for GW quasiparticle band-structure calculations with applications to bulk Si and to the Si(001)-(2×1) surface, Phys. Rev. B 52, 1905 (1995).
- [32] J. Wilhelm, D. Golze, L. Talirz, J. Hutter, and C. A. Pignedoli, Toward GW Calculations on Thousands of Atoms, J. Phys. Chem. Lett. 9, 306 (2018).
- [33] A. Förster and L. Visscher, Low-Order Scaling G_0W_0 by Pair Atomic Density Fitting, J. Chem. Theory Comput. **16**, 7381 (2020).
- [34] I. Duchemin and X. Blase, Cubic-Scaling All-Electron GW Calculations with a Separable Density-Fitting Space-Time Approach, J. Chem. Theory Comput. 17, 2383 (2021).

- [35] J. Wilhelm, P. Seewald, and D. Golze, *Low-Scaling GW with Benchmark Accuracy and Application to Phosphorene Nanosheets*, J. Chem. Theory Comput. **17**, 1662 (2021).
- [36] A. Förster and L. Visscher, *Quasiparticle Self-Consistent GW-Bethe–Salpeter Equation Calculations for Large Chromophoric Systems*, J. Chem. Theory Comput. **18**, 6779 (2022).
- [37] X. Ren, F. Merz, H. Jiang, Y. Yao, M. Rampp, H. Lederer, V. Blum, and M. Scheffler, *All-electron periodic G*₀W₀ *implementation with numerical atomic orbital basis functions: Algorithm and benchmarks*, Phys. Rev. Materials **5**, 013807 (2021).
- [38] T. Zhu and G. K.-L. Chan, *All-Electron Gaussian-Based* G_0W_0 for Valence and Core Excitation Energies of Periodic Systems, J. Chem. Theory Comput. **17**, 727 (2021).
- [39] D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Screening and many-body effects in two-dimensional crystals: Monolayer MoS*₂, Phys. Rev. B **93**, 235435 (2016).
- [40] J. Paier, C. V. Diaconu, G. E. Scuseria, M. Guidon, J. VandeVondele, and J. Hutter, *Accurate Hartree-Fock energy of extended systems using large Gaussian basis sets*, Phys. Rev. B 80, 174114 (2009).
- [41] A. Irmler, A. M. Burow, and F. Pauly, *Robust Periodic Fock Exchange with Atom-Centered Gaussian Basis Sets*, J. Chem. Theory Comput. **14**, 4567 (2018).
- [42] M. N. Gjerding, A. Taghizadeh, A. Rasmussen, S. Ali, F. Bertoldo, T. Deilmann, N. R. Knøsgaard, M. Kruse, A. H. Larsen, S. Manti, T. G. Pedersen, U. Petralanda, T. Skovhus, M. K. Svendsen, J. J. Mortensen, T. Olsen, and K. S. Thygesen, Recent progress of the Computational 2D Materials Database (C2DB), 2D Mater. 8, 044002 (2021).
- [43] M. Camarasa-Gómez, A. Ramasubramaniam, J. B. Neaton, and L. Kronik, Transferable screened range-separated hybrid functionals for electronic and optical properties of van der Waals materials, arXiv 2305.13049 (2023).
- [44] J. Wilhelm, P. Seewald, M. Del Ben, and J. Hutter, *Large-Scale Cubic-Scaling Random Phase Approximation Correlation Energy Calculations Using a Gaussian Basis*, J. Chem. Theory Comput. **12**, 5851 (2016).
- [45] W. Kohn, Density Functional and Density Matrix Method Scaling Linearly with the Number of Atoms, Phys. Rev. Lett. 76, 3168 (1996).
- [46] R. M. Martin, L. Reining, and D. M. Ceperley, *Interacting electrons* (Cambridge University Press, 2016).
- [47] J. VandeVondele and J. Hutter, *Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases*, J. Chem. Phys. **127**, 114105 (2007).
- [48] T. D. Kühne, M. Iannuzzi, M. Del Ben, V. V. Rybkin, P. Seewald, F. Stein, T. Laino, R. Z. Khaliullin, O. Schütt, F. Schiffmann, D. Golze, J. Wilhelm, S. Chulkov, M. H. Bani-Hashemian, V. Weber, U. Borštnik, M. Taillefumier, A. S. Jakobovits, A. Lazzaro, H. Pabst, T. Müller, R. Schade, M. Guidon, S. Andermatt, N. Holmberg, G. K. Schenter, A. Hehn, A. Bussy, F. Belleflamme, G. Tabacchi, A. Glöß, M. Lass, I. Bethune, C. J. Mundy, C. Plessl, M. Watkins, J. Vande Vondele, M. Krack, and J. Hutter, CP2K: An electronic structure and molecular dynamics software package Quickstep: Efficient and accurate electronic structure calculations, J. Chem. Phys. 152, 194103 (2020).
- [49] https://github.com/cp2k/cp2k.

- [50] F. Bruneval and X. Gonze, Accurate GW self-energies in a plane-wave basis using only a few empty states: Towards large systems, Phys. Rev. B 78, 085125 (2008).
- [51] Considering time-reversal symmetry, only 5 irreducible k-points are contained in a 3×3 mesh.
- [52] F. Hüser, T. Olsen, and K. S. Thygesen, Quasiparticle GW calculations for solids, molecules, and two-dimensional materials, Phys. Rev. B 87, 235132 (2013).
- [53] The computation time is

$$T = \left(\frac{N_{\text{at}}}{243}\right)^2 1870 \text{ core hours} + \left(\frac{N_{\text{at}}}{243}\right)^3 443 \text{ core hours}$$
 (10)

where we took the execution time of quadratic and cubic steps as shown in Fig. 3. The maximum job size on Supermuc-NG is 150 000 cores for 24 hours, making 3.6 Mio. core hours, which will allow for a G_0W_0 calculation on 4500 atoms [Eq. (10)]. With restarting, even larger G_0W_0 calculations will become possible.

- [54] W. J. Schutte, J. L. De Boer, and F. Jellinek, Crystal structures of tungsten disulfide and diselenide, J. Solid State Chem. 70, 207 (1987).
- [55] P. B. James and M. T. Lavik, The crystal structure of MoSe₂, Acta Crystallogr. 16, 1183 (1963).
- [56] K. Zollner, P. E. F. Junior, and J. Fabian, *Strain-tunable orbital, spin-orbit, and optical properties of monolayer transition-metal dichalcogenides*, Phys. Rev. B **100**, 195126 (2019).

- [57] A. Chaves, J. G. Azadani, H. Alsalman, D. R. da Costa, R. Frisenda, A. J. Chaves, S. H. Song, Y. D. Kim, D. He, J. Zhou, A. Castellanos-Gomez, F. M. Peeters, Z. Liu, C. L. Hinkle, S.-H. Oh, P. D. Ye, S. J. Koester, Y. H. Lee, P. Avouris, X. Wang, and T. Low, Bandgap engineering of two-dimensional semiconductor materials, npj 2D Mater. Appl. 4, 29 (2020).
- [58] X. Jiang, Q. Zheng, Z. Lan, W. A. Saidi, X. Ren, and J. Zhao, Real-time GW-BSE investigations on spin-valley exciton dynamics in monolayer transition metal dichalcogenide, Sci. Adv. 7, eabf3759 (2021).
- [59] Y.-H. Chan, D. Y. Qiu, F. H. da Jornada, and S. G. Louie, *Giant exciton-enhanced shift currents and direct current conduction with subbandgap photo excitations produced by many-electron interactions*, Proc. Natl. Acad. Sci. **118**, e1906938118 (2021).
- [60] E. Perfetto, Y. Pavlyukh, and G. Stefanucci, Real-Time GW: Toward an Ab Initio Description of the Ultrafast Carrier and Exciton Dynamics in Two-Dimensional Materials, Phys. Rev. Lett. 128, 016801 (2022).
- [61] N. Schlünzen, J.-P. Joost, and M. Bonitz, Achieving the Scaling Limit for Nonequilibrium Green Functions Simulations, Phys. Rev. Lett. 124, 076601 (2020).
- [62] R. Tuovinen, Y. Pavlyukh, E. Perfetto, and G. Stefanucci, Time-Linear Quantum Transport Simulations with Correlated Nonequilibrium Green's Functions, Phys. Rev. Lett. 130, 246301 (2023).
- [63] https://github.com/JWilhelm/Inputs_outputs_low_ scaling_GW_TMDC.