

This work is on a Creative Commons Attribution 4.0 International (CC BY 4.0) license, <https://creativecommons.org/licenses/by/4.0/>. Access to this work was provided by the University of Maryland, Baltimore County (UMBC) ScholarWorks@UMBC digital repository on the Maryland Shared Open Access (MD-SOAR) platform.

Please provide feedback

Please support the ScholarWorks@UMBC repository by emailing [scholarworks-group@umbc.edu](mailto:scholarworks-group@umbc.edu) and telling us what having access to this work means to you and why it's important to you. Thank you.

# A pathway towards high throughput Quantum Monte Carlo simulations for alloys: A case study of two-dimensional (2D) $\text{GaS}_x\text{Se}_{1-x}$

Daniel Wines,<sup>1</sup> Kayahan Saritas,<sup>2</sup> and Can Ataca<sup>1, a)</sup>

<sup>1)</sup>*Department of Physics, University of Maryland Baltimore County, Baltimore MD 21250*

<sup>2)</sup>*Department of Applied Physics, Yale University, New Haven CT 06520*

(Dated: 14 July 2021)

The study of alloys using computational methods has been a difficult task due to the usually unknown stoichiometry and local atomic ordering of the different structures experimentally. In order to combat this, first-principles methods have been coupled with statistical methods such as the Cluster Expansion formalism in order to construct the energy hull diagram, which helps to determine if an alloyed structure can exist in nature. Traditionally, density functional theory (DFT) has been used in such workflows. In this work we propose to use chemically accurate many-body variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) methods to construct the energy hull diagram of an alloy system, due to the fact that such methods have a weaker dependence on the starting wavefunction and density functional, scale similarly to DFT with the number of electrons, and have had demonstrated success for a variety of materials. To carry out these simulations in a high-throughput manner, we propose a method called *Jastrow sharing*, which involves recycling the optimized Jastrow parameters between alloys with different stoichiometries. We show that this eliminates the need for extra VMC Jastrow optimization calculations and results in a significant computational cost savings (on average 1/4 savings of total computational time). Since it is a novel post-transition metal chalcogenide alloy series that has been synthesized in its few-layer form, we used monolayer  $\text{GaS}_x\text{Se}_{1-x}$  as a case study for our workflow. By extensively testing our *Jastrow sharing* procedure for monolayer  $\text{GaS}_x\text{Se}_{1-x}$  and quantifying the cost savings, we demonstrate how a pathway towards chemically accurate high-throughput simulations of alloys can be achieved using many-body VMC and DMC methods.

## I. INTRODUCTION

The accurate modeling of alloys with first-principles methods is a difficult computational effort in the materials science community<sup>1</sup>. The main issue arises from the usually unknown stoichiometry and local atomic ordering of the alloyed structures. One way to combat this is to manually create supercells with a specific stoichiometry and local atomic ordering determined empirically or intuitively<sup>2-4</sup>. Another, more systematic option is to create a large set of alloyed structures (of varying compositions and stoichiometries) and determine the energy hull diagram of the alloy series from first-principles coupled with statistical methods<sup>1,5,6</sup>. Popular methods usually involve the Special Quasirandom (SQS)<sup>7</sup> generation of several structures, evaluation of energies from first-principles, and then using the first-principles energies as a training set for statistical methods such as the Cluster Expansion<sup>1</sup> formalism.

Constructing an accurate energy hull diagram is essential for predicting the energetic stability and favorability of alloys. The accuracy of an energy hull diagram depends critically on the choice of the method used to generate total energies. Density functional theory (DFT) has been used for this purpose, thanks to its balance between accuracy and computational scaling on modern computer clusters. On the other hand, diffusion Monte Carlo (DMC)<sup>8,9</sup> is another ground state method which can provide total energy differences with significantly better accuracy than DFT, with minimal functional and basis set dependence<sup>10-14</sup>. DMC has a very similar computational scaling to DFT,  $N^{3-4}$ , though with nearly a thousand times

larger prefactor, but its embarrassingly parallel algorithm can utilize parallel computation architectures using larger number of nodes. Recently, DMC has been used to calculate chemically accurate total energies and electronic properties for a wide variety of two-dimensional (2D) and bulk structures, demonstrating the routine applicability of this method<sup>11-34</sup>. In this work, building on our previous DMC study of 2D  $\text{GaSe}$ <sup>10</sup>, we propose to use DMC to reliably calculate the energetics of 2D  $\text{GaS}_x\text{Se}_{1-x}$  as a case study. This will enable accurate simulations of alloy properties for experimental characterization and validation.

DMC calculations require multiple steps to obtain the total energy of a structure and the stochastic nature of the method introduces uncertainty in the observables, in contrast to the deterministic nature of DFT. Constructing an adequate energy hull diagram requires the high-throughput calculation of the energy of several structures. Similarly to the development of the decade old field of high throughput DFT<sup>35</sup>, researchers have attempted to bridge the gap between high throughput computing and Quantum Monte Carlo (QMC) and increase the computational efficiency of the QMC workflow<sup>36,37</sup>. For example, Saritas et al.<sup>37</sup> developed a high throughput procedure to determine the DMC formation energy of certain materials directly from the ICSD database. In this work, we propose a new high throughput workflow to determine the alloy formation energy of 2D  $\text{GaS}_x\text{Se}_{1-x}$  with near chemical accuracy. We went on to test various wavefunction optimization methods, which included recycling Jastrow parameters (which we call *Jastrow sharing*) among various alloy configurations to reduce the overall computational cost and localization errors.

In addition to 2D  $\text{GaS}_x\text{Se}_{1-x}$  being a suitable case study that builds off our previous QMC work<sup>10</sup>, it has been reported

<sup>a)</sup>Electronic mail: [ataca@umbc.edu](mailto:ataca@umbc.edu)

that alloying is a promising route to control the properties of several monolayer or few layer materials<sup>3,4,38-43</sup> (including post transition metal chalcogenides, or PTMCs<sup>5,6,44-46</sup>), which makes the study of 2D alloys with accurate QMC methods useful for theorists and guiding experimentalists in synthesis and characterization. For example, it has experimentally been reported that when GaSe nanostructures are alloyed with Te, the material undergoes a hexagonal to monoclinic transition and there exists an instability region where the phases compete and two different band gaps can be found at the same composition<sup>6</sup>. Janus monolayers<sup>45,47</sup> (a class of two-dimensional 2D alloyed structures) have also gained attention due to their applications for Schottky contacts, as demonstrated theoretically with Janus GaS<sub>x</sub>Se<sub>1-x</sub> on top of graphene<sup>45</sup>. In addition, PTMCs such as GaSe and GaS can be engineered for specific applications by chemical functionalization<sup>48-50</sup>, creating heterostructures<sup>45,51-54</sup>, and applying strain<sup>55-57</sup>. PTMCs such as GaSe have also been reported to be suitable substrate material for other 2D structures, as well<sup>58,59</sup>.

In section II we outline our DFT and QMC approaches and convergence criteria. In section III A we present our DFT results in order to initially screen 2D GaS<sub>x</sub>Se<sub>1-x</sub> alloys. In section III B we describe our *Jastrow sharing* methodology and present a detailed analysis of this methodology at the variational Monte Carlo (VMC) and DMC level. In section III C we present DMC results using various supercell sizes for alloys and compare extrapolated results using different wavefunction optimization methods. In section III D we give an analysis of the computational cost savings of our approach. Finally, we provide concluding remarks and future perspectives in section IV.

## II. COMPUTATIONAL METHODS AND THEORY

Benchmarking DFT calculations were performed using the VASP code with projector augmented wave (PAW) potentials<sup>60,61</sup>. For these VASP benchmarking calculations, the Perdew-Burke-Ernzerhof (PBE)<sup>62</sup> and strongly constrained and appropriately normed (SCAN)<sup>63</sup> meta-GGA functionals were used. In addition, the PBE+D2<sup>64</sup> and PBE+D3<sup>65</sup> methods of Grimme and the SCAN+rvv10<sup>66</sup> functionals were used to investigate vdW effects at the DFT level. At least 20 Å of vacuum was given between periodic layers of GaS<sub>x</sub>Se<sub>1-x</sub> in the *c*-direction, a kinetic energy cutoff of 350 eV was used, and a 6x6x1 reciprocal grid was used for the primitive cells of the alloys and the number of k-points were scaled appropriately with supercell size.

DMC calculations use wavefunctions from other ab-initio methods, in our case DFT-PBE, as input (trial wavefunction) to obtain an equilibrium that is computationally tractable. For DFT calculations within the QMC workflow, the Quantum Espresso (QE)<sup>67</sup> code was used. For Ga, S, and Se we used norm-conserving Burkatzki-Filippi-Dolg (BFD)<sup>68,69</sup> pseudopotentials. DMC calculations require norm-conserving pseudopotentials and BFD pseudopotentials have been thoroughly tested in DMC<sup>10,70</sup>. For details of pseudopotential

testing, conversion and validation, refer to the discussion and Table S1 in the Supplementary Information (SI). For the pseudopotentials, we used a kinetic energy cutoff of 160 Ry (see Fig. S1), which gives a convergence of less than 1 meV at the DFT level. We used the same supercell reciprocal twist grid in QE as our benchmarking VASP calculations which is tested in Fig. S2.

After the trial wavefunction is generated using DFT, VMC and DMC<sup>8,9</sup> calculations were carried out using the QMCPACK<sup>71,72</sup> code. VMC calculations are the intermediate steps between the DMC and DFT calculation, where the single determinant DFT wavefunction is converted into a many-body wavefunction, through the Jastrow parameters<sup>73,74</sup>. Jastrow parameters help model the electron correlation and reduce the uncertainty in the DMC calculations<sup>75,76</sup>. The automated DFT-VMC-DMC workflows were generated using the Nexus<sup>77</sup> software suite. Up to three-body Jastrow<sup>78</sup> correlation functions were included. The linear method<sup>79</sup> was used to minimize the variance and energy respectively of the VMC total energies. The cost function of the variance optimization is 100% variance minimization and the cost function of the energy optimization is split as 95% energy minimization and 5% variance minimization, which has been shown to reduce uncertainty for DMC results<sup>75</sup>. The explicit details of how we modified this optimization procedure for our high-throughput QMC workflow will be given in section III B.

Being a real-space wavefunction method, DMC calculations need to be performed at increasing supercell sizes to eliminate finite-size errors. We used supercell sizes up to 72 atoms and extrapolated to the infinite cell size. In order to smooth the image interactions at each supercell used in the extrapolation, we used the `optimal_tilematrix` function in Nexus, which allows constructing appropriate supercells with the largest Wigner-Seitz (WS) radius for a given size. We used Jackknife fitting to obtain a linear fit of DMC data and extrapolate to the infinite-size limit for total ground state energy of each alloyed structure. In addition, the locality approximation<sup>76</sup> was used to evaluate the nonlocal pseudopotentials in DMC. A timestep of 0.01 Ha<sup>-1</sup> (convergence tests are shown in Fig. S3) was used for all DMC simulations.

To initially create and screen these GaS<sub>x</sub>Se<sub>1-x</sub> alloyed structures with varying concentrations, the Special Quasirandom Structure (SQS) method<sup>7</sup> was employed as implemented in the ATAT code<sup>1</sup>. The alloy formation energy in units of eV/formula unit (otherwise known as the mixing enthalpy) is defined as:

$$E_{\text{form}} = E_{\text{GaS}_{1-x}\text{Se}_x} - (1-x)E_{\text{GaS}} - xE_{\text{GaSe}} \quad (1)$$

where  $E_{\text{GaS}_{1-x}\text{Se}_x}$  is the energy of a GaS<sub>x</sub>Se<sub>1-x</sub> cell,  $E_{\text{GaS}}$  is the energy of a GaS cell, and  $E_{\text{GaSe}}$  is the energy of a GaSe cell, all per formula unit (f.u.). After the SQS generation of random alloys, first-principles calculations are done to calculate the formation energy of each structure. These energies are then used to obtain fitted energies via the Cluster Expansion formalism<sup>1</sup>. Further details and convergence criteria can be found in the Supplementary Information (SI).

### III. RESULTS AND DISCUSSION

#### A. DFT Calculations

In contrast to a previous DFT study where the 2D  $\text{GaS}_x\text{Se}_{1-x}$  alloys were constructed manually<sup>2</sup>, we used the SQS method to generate alloys and DFT to initially optimize the geometries. Lattice parameters of 2D GaS and GaSe have been determined experimentally<sup>2</sup>, therefore accurate determination of the geometry is an important step for the respective 2D alloys, which have unknown experimental lattice constants. Due to this, we performed geometry and lattice optimization calculations with PBE, PBE-D2, PBE-D3, SCAN and SCAN+rvv10 DFT functionals to benchmark. This process also allows us to construct the energy hull diagrams using each functional (see Fig. 1 and Fig. S4). We find that all the tested functionals are in good agreement with each other in their formation energies, but lattice parameters differ slightly for each structure. However, we find that the SCAN+rvv10 functional gives the overall best agreement with the experimental lattice parameters of 2D GaS and GaSe<sup>2,80</sup> (see Table S2 in SI). This is expected since it has been previously reported that SCAN yields lattice constants and optimal geometry in closer agreement to experiment for several 2D materials<sup>10,81</sup>. Therefore, we use the geometries from the SCAN+rvv10 functional as a starting point for DMC calculations. To ensure that the optimal geometry of these structures are closest to exact and do not bias the energy, we isotropically scaled the lattice (starting from the geometry and lattice constants obtained from SCAN+rvv10) and confirmed what the location of the minimum energy was with respect to lattice constant for the equation of state (see Fig S5-9). The benchmarking of lattice parameters with different functionals are done using VASP because the plane wave cutoff of PAW potentials are more than an order of magnitude smaller than BFD pseudopotentials.

As seen in Fig. 1 f), all the monolayer  $\text{GaS}_x\text{Se}_{1-x}$  structures have positive formation energy, which implies that these structures are potentially not energetically favorable at zero temperature and cannot form spontaneously in vacuum (metastable). In addition, Janus  $\text{GaS}_{0.5}\text{Se}_{0.5}$  (where S and Se atoms are on the opposing faces of the 2D layer), has the highest formation energy, which means that this structure is the least energetically favorable in free-standing form. This is a similar finding to Ersan *et al.*<sup>47</sup>, where it was found that monolayer Janus PtXY (X = S, Se, Te) is the least favorable alloyed structure in the 2D PtXY alloy series. Even though these reported Janus structures are thermodynamically stable (as a result of phonon and molecular dynamics simulations) in vacuum<sup>47</sup>, they can perhaps be synthesized on a substrate material<sup>82,83</sup>. In free-standing form, molecular dynamics simulations showed that certain Janus structures can spontaneously curl<sup>84</sup>. Figure S4 and 1 f) also reflects the band gap tunability (color axis) of  $\text{GaS}_x\text{Se}_{1-x}$  monolayers at the DFT level, opening the door to various electronic applications. Following the initial geometric optimization and construction of the energy hull diagram with DFT, we selected certain structures to further run QMC on to test our *Jastrow sharing* pro-

cedure which we will discuss in the next subsection.

#### B. Jastrow Parameter Sharing

In attempt to reduce the number of wavefunction optimization simulations required to perform DMC for all the 2D  $\text{GaS}_x\text{Se}_{1-x}$  structures involved in the Cluster Expansion formalism, we propose a systematic method we call *Jastrow sharing*. Wavefunction optimization using Jastrow parameters is key to reducing variance and the localization error of subsequent DMC simulations<sup>32,70,85</sup>. Jastrow parameters are typically optimized using the cheaper VMC calculations, but these costs can be important considering the large number of materials simulated in the Cluster Expansion formalism. Computational cost analysis will further be discussed in Section III D. Our *Jastrow sharing* approach involves a judicious selection of optimized Jastrow parameters from a single compound, and reusing it in all the rest of the materials studied. In this section we discuss the various methods we used to test and validate this method for  $\text{GaS}_x\text{Se}_{1-x}$  and the limitations of this method with respect to other material classes.

In our method, we use a supercell that has the smallest Wigner-Seitz (WS) radius to obtain the optimized Jastrow parameters that will later be used in all DMC calculations. The maximum Jastrow cutoff radius possible in a structure is equal to its WS radius, which is equal to the radius of the largest inscribing sphere that can fit into the simulation cell. Even though the Jastrow cutoff can be optimized, this increases the computational scaling of the VMC optimization<sup>78</sup>, hence it is more practical to use the maximum Jastrow cutoff possible for a periodic system (WS radius), and update each Jastrow parameter at a fixed cost. Therefore, it is imperative that the smallest WS radius among all the structures is determined in advance in order to reuse the same set of Jastrow parameters for all the structures. For three-body Jastrows, a cutoff radius of 2 Å (on the order of  $\text{GaS}_x\text{Se}_{1-x}$  bond lengths) was used. Due to the fact that three-body Jastrows (energy optimization) usually capture shorter-range correlations, we made this arbitrary choice to minimize computational cost while attempting to capture such interactions. The full workflow of our procedure is given in Fig. 5.

We performed DMC calculations on four structures:  $x = 0.333$  (12 atom cell, Fig. 1 c)),  $x = 0.5$  (16 atom cell, Fig. 1 b)) and  $x = 0.667$  (12 atom cell, Fig. 1 d)) and the Janus  $\text{GaS}_{0.5}\text{Se}_{0.5}$  structure (highest formation energy, 16 atom cell, Fig. 1 a)). Except for the Janus structure, all these compounds lie on the bottom of the DFT hull diagram. However, this selection can be made on a finite threshold above the DFT hull or more ambitious plans can involve all the SQS generated structures. The reason we decided to investigate materials with low and high formation energy is to observe if QMC can correctly capture these energy differences and to demonstrate that our method can work for the entire energy hull.

Among the unit cells of the alloyed structures we selected to study,  $x = 0.875$  has the smallest WS radius (see Fig. 1 e)). Therefore, according to our procedure, we would use  $\text{GaS}_{0.875}\text{Se}_{0.125}$  to generate the Jastrow parameters to be used

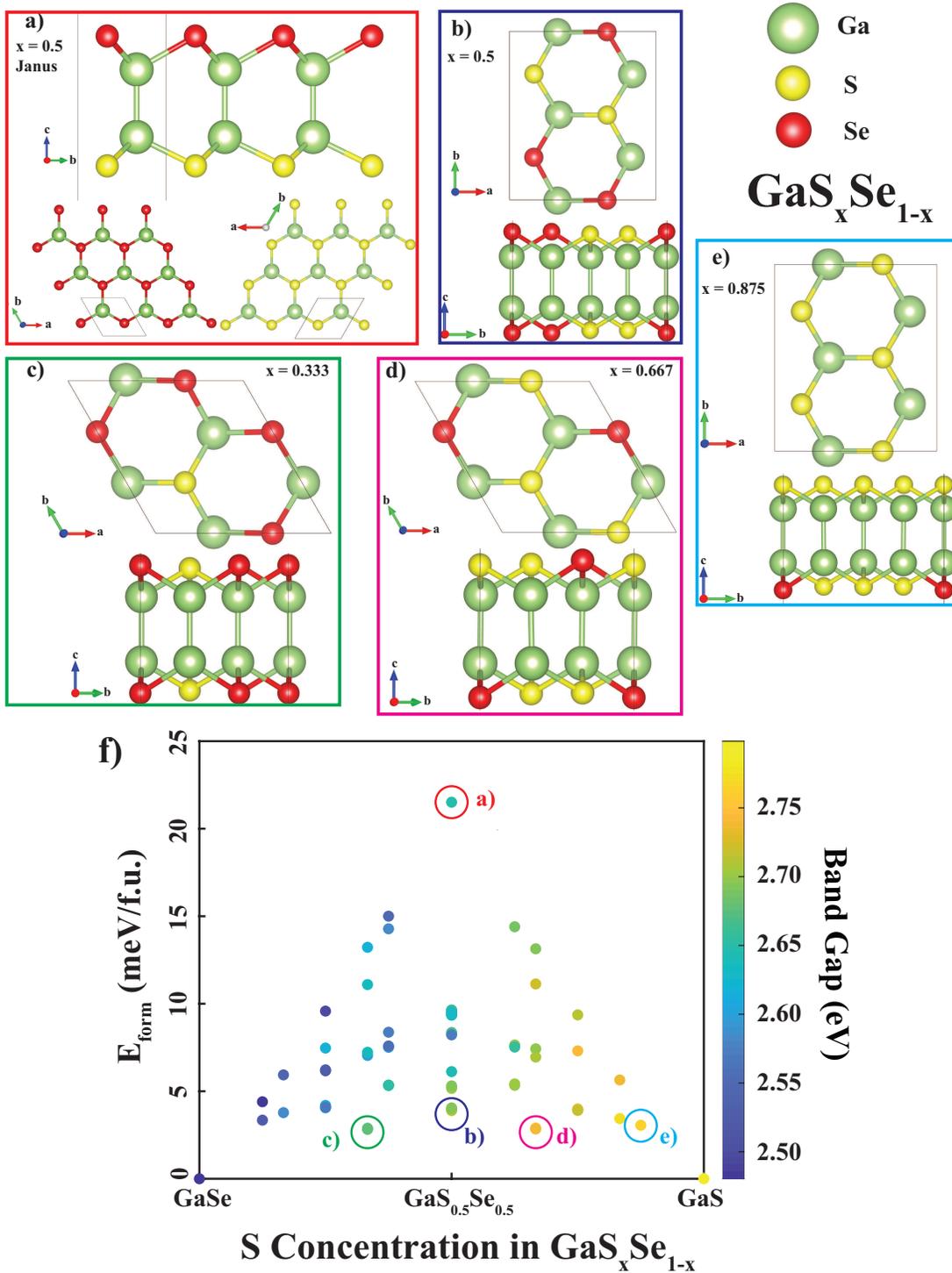


FIG. 1. The atomic structure (top and side view) of selected SQS GaS<sub>x</sub>Se<sub>1-x</sub> alloys from the energy hull diagram: a) Janus GaS<sub>0.5</sub>Se<sub>0.5</sub>, b) GaS<sub>0.5</sub>Se<sub>0.5</sub> (lowest  $E_{\text{form}}$ ), c) GaS<sub>0.333</sub>Se<sub>0.667</sub> (lowest  $E_{\text{form}}$ ), d) GaS<sub>0.667</sub>Se<sub>0.333</sub> (lowest  $E_{\text{form}}$ ), e) GaS<sub>0.875</sub>Se<sub>0.125</sub> (smallest Wigner-Seitz radius) and f) the energy hull diagram obtained using DFT (SCAN+rvv10, PAW potentials), where the structures depicted (a - e) are circled with the corresponding color outline and the band gap is given in the color axis for each structure.

DMC-J3	Janus	x = 0.5	x = 0.333	x = 0.667
Janus Jastrow	0	-6(6)	2(6)	3(10)
x = 0.5 Jastrow	0(4)	0	4(7)	15(11)
x = 0.333 Jastrow	-1(4)	-2(4)	0	17(10)
x = 0.667 Jastrow	-6(5)	-10(6)	-1(6)	0
x = 0.875 Jastrow	-2(5)	-15(7)	-2(5)	9(9)
DMC-J2	Janus	x = 0.5	x = 0.333	x = 0.667
Janus Jastrow	0	-3(7)	-9(7)	-6(7)
x = 0.5 Jastrow	3(4)	0	5(7)	-9(8)
x = 0.333 Jastrow	-7(5)	4(9)	0	6(9)
x = 0.667 Jastrow	-9(7)	-2(8)	-2(6)	0
x = 0.875 Jastrow	-8(4)	-5(8)	-2(6)	-3(7)

TABLE I. The differences (in meV/f.u.) between total energies (DMC) calculated with *self-Jastrows* and *shared-Jastrows* including up to two-body and up to three-body terms with the associated error bars in parenthesis. The rows represent which Jastrows are used and the columns represent each alloyed structure for which the DMC energies are calculated.

in the remainder of the calculations. However, in order to understand the sensitivity of our method to this decision, we make a benchmark study where we optimize Jastrow parameters of all the structures separately as it is typically done in DMC calculations. For testing, a 16 atom cell was used for  $x = 0.5$ , Janus, 0.875 and a 12 atom cell was used for  $x = 0.333$ , 667. To be able to use all 5 of these sets of Jastrow parameters across all the structures we study, we still use the smallest WS radius as the cutoff parameter for two-body Jastrow parameters. The smallest WS radius we have in this case is 3.12 Å. This is longer than 2 Å, hence no modifications are required for 3-body Jastrows. The testing of our *Jastrow sharing* procedure is tabulated in Table I, where we show the energy difference (in meV/f.u.) between the total energy of each structure using its own optimized Jastrow parameters (we will define this as a *self-Jastrows*) and the total energy of each structure using the optimized Jastrow parameters of a different structure (we will define these as *shared-Jastrows*), with the associated error. In order to accept the *Jastrow-sharing* procedure as valid, we expect the localization error obtained with *self-Jastrows* and *shared-Jastrows* to be comparable and the total energy difference (per f.u.) between *self-Jastrows* and *shared-Jastrows* to be as minimal as possible.

In Table I, the columns represent the different selected structures and the rows represent the various optimized Jastrow parameters. The upper quadrant contains the energy difference at the DMC level using up to three-body Jastrows and the lower quadrant contains the energy difference at the DMC level using up to two-body Jastrows. A more detailed table with VMC energies, including two- and three-body Jastrows are given in Table S3 in the SI. From the data in Table I, we observe that at the DMC level, the total energy differences between *shared-Jastrows* and *self-Jastrows* are overall smaller, with differences ranging from 0 - 17 meV/f.u. We also observe that the energy differences calculated with DMC are nearly identical whether or not two-body or three-body interactions are included in the Jastrow factor due to the fact

that we don't have enough resolution to differentiate between the results. The lack of significant variations in the DMC energies here also suggest that the localization errors of these pseudopotentials are very small or negligible since the quality of the Jastrow parameters, with regards to the Jastrow cutoff, do not affect the DMC calculations.

In Figure 2, we additionally show that DMC-J3 total energies are always lower than the DMC-J2 energies. This has been observed for a variety of systems<sup>85</sup>. However, in a practical sense, the variation in a set of Jastrow parameters (J2 or J3) is more important than the separation between the DMC energies calculated with J2 and J3 Jastrow parameters. This is because often the user makes one selection regarding the number of Jastrow parameters and applies it throughout for all similar materials studied in the same calculation set, such as polymorphs or alloys. Ideally, the quality of the Jastrow parameters should only change the effort required to get the target uncertainty in the DMC total energies. However, as observed in here, we often see that the quality of the Jastrow parameters also change the degree of localization error<sup>32,70,85</sup> which can lead to inconsistencies when energy differences are calculated between two structures.

In order to quantify the amount of localization error introduced as a function of the quality of the Jastrow parameter, we can utilize a quantity called Jastrow sensitivity. The Jastrow sensitivity is defined as the amount of energy decrease in the DMC energy per unit decrease in the VMC energy and is a property of the pseudopotential used in the calculation<sup>85</sup>. Ideally, the Jastrow sensitivity should be small for a given pseudopotential for all of the relevant valence environments (i.e. different phases of a given material). We infer that a low enough Jastrow sensitivity could be a reason why we were successfully able to transfer different Jastrow parameters across the 2D GaS<sub>x</sub>Se<sub>1-x</sub> series. In theory, if the sensitivity of the pseudopotentials is low for the material class of study, our *Jastrow sharing* procedure should be valid.

To further quantify Jastrow sensitivity in our work, we define a new quantity called *shared-Jastrow sensitivity* ( $S_{J,S}$ ), which is defined as the DMC energy per unit decrease in the VMC energy calculated with *self-Jastrows* and *shared-Jastrows*. In order to calculate this for each structure, we plotted the DMC total energies vs. the VMC total energies (per f.u.) and obtained the slope for the data (using J2 and J3 parameters, from Table I), which is depicted in Fig. 2. Due to the fact that VMC energies obtained with the Janus Jastrow parameters have a large deviation attributed to cutoff radius, we excluded these points from the linear fitting. A more detailed discussion related to the Jastrow parameters of the Janus structure can be found in the SI. From Fig. 2 we find the *shared-Jastrow sensitivity* (per f.u.) to be 0.094 (0.047 per atom) for the Janus structure, 0.116 (0.058 per atom) for the  $x = 0.5$  structure, 0.084 (0.042 per atom) for the  $x = 0.333$  structure, and 0.115 (0.058 per atom) for the  $x = 0.667$  structure.

These sensitivities are comparable in value to the lower bounds on the Jastrow sensitivities reported by Krogel and Kent<sup>85</sup> (0.05 for Ce 4+ using the locality approximation) and Dzubak, Krogel, and Reboredo<sup>70</sup> (0.07, 0.05, 0.04 for Mn, Fe, Co respectively). In addition to our calculated *shared-Jastrow*

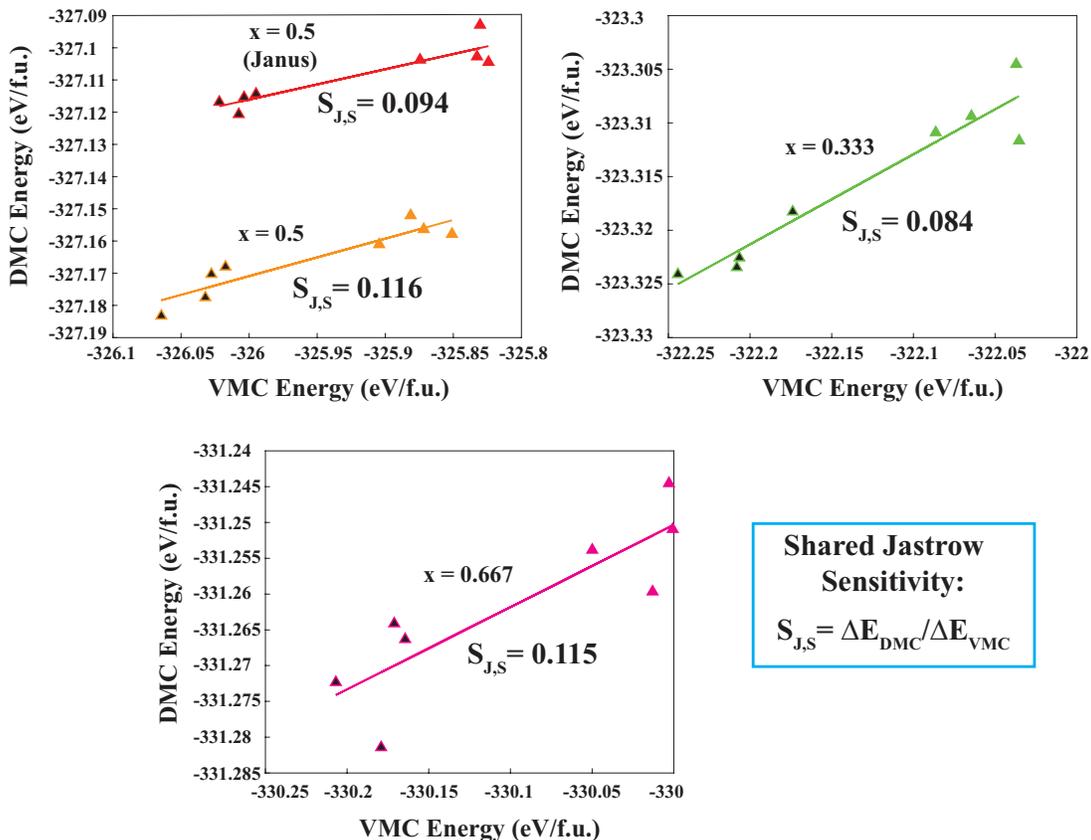


FIG. 2. The total DMC energy (eV/f.u.) as a function of total VMC energy (eV/f.u.) for the Janus,  $x = 0.5$ ,  $x = 0.333$ , and  $x = 0.667$  structures. The formula for *shared-Jastrow sensitivity* (per f.u.), which is defined as the change in DMC energy per change in VMC energy (slope of the graph), is given in each inset. The data points with a black center represent J3 results while the other points represent J2 results. Due to their non linearity, the data points using the Janus Jastrow were excluded (see SI).

*sensitivity*, we calculated the standard Jastrow sensitivity of the individual Ga, S and Se atoms to further prove transferability across different atomic systems. We obtained this by calculating the DMC energy with no Jastrow parameters and the DMC energy using up to J2 and J3 Jastrows, then fitting a line to the three data points. We obtain a Jastrow sensitivity of 0.07, 0.14 and 0.04 for Ga, S and Se respectively. As a result of these sensitivity calculations, we can quantify that the Jastrows can be transferable between various stoichiometries in the  $\text{GaS}_x\text{Se}_{1-x}$  system. If one wished to implement this *Jastrow-sharing* procedure for another system, calculating the *shared-Jastrow sensitivity* using VMC and DMC energies from *self-Jastrows* and *shared-Jastrows* and the standard Jastrow sensitivity of the respective atoms and comparing to tabulated results in literature could be a viable method to check if *Jastrow-sharing* is feasible.

### C. Finite-size Effects

To construct our DMC energy hull diagram and eliminate finite-size effects, we extrapolate all structures to the thermodynamic limit. Therefore, our DMC calculations are per-

formed at different supercell sizes and then extrapolated to the infinite sized limit as the energy scales as  $1/N$ , where  $N$  is the number of particles. In addition to this, we use a converged reciprocal grid obeying periodic boundary conditions for each supercell, which is known as twist averaging. For an additional benchmark to assess the validity of our approach, we calculated the DMC cohesive energy (per f.u.) of each structure by subtracting the appropriate single atom energies of Ga, Se and S (see SI for single atom calculation details) from the extrapolated total energy. We calculated the alloy formation energy ( $E_{\text{form}}$ ) at the DMC level using equation 1 (see Table II). The total energy of the GaSe and GaS supercells with the exact same number of atoms as the specific alloy supercell was subtracted from the total energy of said alloy. This alloy formation energy was calculated at multiple supercell sizes and extrapolated to the thermodynamic limit (see Figure 3).

For the DMC extrapolated cohesive energies and formation energies, *self-Jastrows* were used for each respective supercell. In addition, the optimized Jastrow parameters of the alloy with the smallest WS radius (we will refer to this as *WS-Jastrow*, with parameters taken from the 16 atom  $\text{GaS}_{0.875}\text{Se}_{0.125}$  cell) were used to calculate the total DMC energy for each supercell size. A comparison of the DMC results

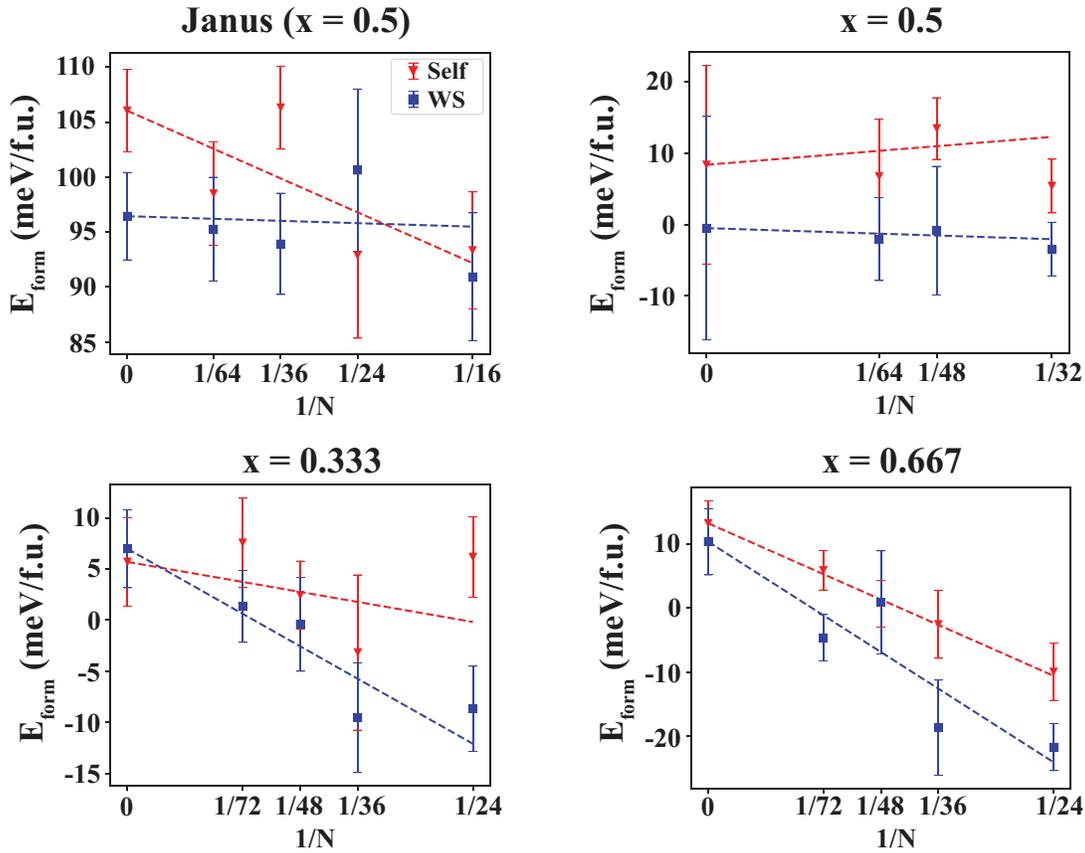


FIG. 3. The DMC calculated formation energies as a function of inverse number of atoms ( $N$ ) in the supercell, extrapolated to the thermodynamic limit ( $N \rightarrow \infty$ ) for the Janus,  $x = 0.5$ ,  $x = 0.333$ , and  $x = 0.667$  structures. Red represents DMC energies calculated with *self-Jastrows* while blue represents DMC energies calculated with *WS-Jastrows*.

(for extrapolated cohesive energies and alloy formation energies) using *self-Jastrows* and *WS-Jastrows* and comparison to DFT is found in Table II. In addition, the partial energy hull diagram for the selected structures using DFT and DMC (with *self-Jastrows* and *WS-Jastrows*) is depicted in Figure 4. From the linear extrapolated results (Fig. 3) depicted in Fig. 4 and tabulated in Table II, we clearly see the energies calculated (and then extrapolated) using *self-Jastrows* and *WS-Jastrows* are nearly indistinguishable.

These extrapolated formation energies demonstrate that an adequate energy hull diagram can be constructed from DMC using just one set of Jastrow parameters from the smallest WS cell, in contrast to optimizing the Jastrows for every structure at every supercell size. Additionally, using a many-body approach such as QMC for alloys can provide near-chemically accurate confirmation of stability and other important electronic and energetic properties. 2D  $\text{GaS}_x\text{Se}_{1-x}$  was chosen because it is a convenient system to test this methodology on. This is convenient because it builds on our previous 2D  $\text{GaSe}$ <sup>10</sup> work and since there is relatively good agreement between DFT functionals for the formation energy hull diagram (Fig. S4), it allows us to benchmark the resolution of our DMC extrapolated (to meV resolution) energies with reliable

DFT energies.

A full schematic of this workflow is given in Fig. 5. This workflow starts with an input of an initial lattice, followed by the SQS generation of random alloys and then the construction of the energy hull diagram with DFT. After analyzing the DFT hull diagram, the user can choose to run QMC for structures below a certain energy threshold or discard them. From the SQS structures, the user must determine which has the smallest WS radius, and then perform J2 and J3 Jastrow optimization for such. These J2 and J3 parameters are then reused for DMC calculations of the other selected structures. From this, the DMC energy diagram can be constructed. A more detailed analysis of the cost savings of this workflow is given in the following section.

#### D. Computational Cost Analysis

We claim that reusing a fixed single set of Jastrow parameters across all structures involved in creating a DMC energy hull diagram is a valid and computationally efficient approach. Therefore, we compare the efficiency of our procedure to another more conventional way of performing these

Cohesive Energy (eV/f.u.)			
Structure	DFT (PBE)	DMC self-Jastrow	DMC WS-Jastrow
GaSe	-7.135	-7.028(3) <sup>10</sup>	-
GaS	-7.789	-7.659(4)	-
Janus	-7.360	-7.231(6)	-7.240(6)
x = 0.5	-7.456	-7.35(1)	-7.35(2)
x = 0.333	-7.347	-7.259(6)	-7.258(6)
x = 0.667	-7.564	-7.461(6)	-7.464(7)

Formation Energy (eV/f.u.)			
Structure	DFT (PBE)	DMC self-Jastrow	DMC WS-Jastrow
Janus	0.1019	0.106(4)	0.096(4)
x = 0.5	0.0059	0.01(1)	0.00(2)
x = 0.333	0.0057	0.006(4)	0.007(4)
x = 0.667	0.0064	0.013(3)	0.010(5)

TABLE II. The calculated cohesive energy (top half) and alloy formation energy (bottom half) all in eV/f.u. obtained from DFT (PBE, BFD pseudopotentials) and DMC (using *self-Jastrows* and *WS-Jastrows*) for the Janus,  $x = 0.5$ ,  $x = 0.333$ , and  $x = 0.667$  structures.

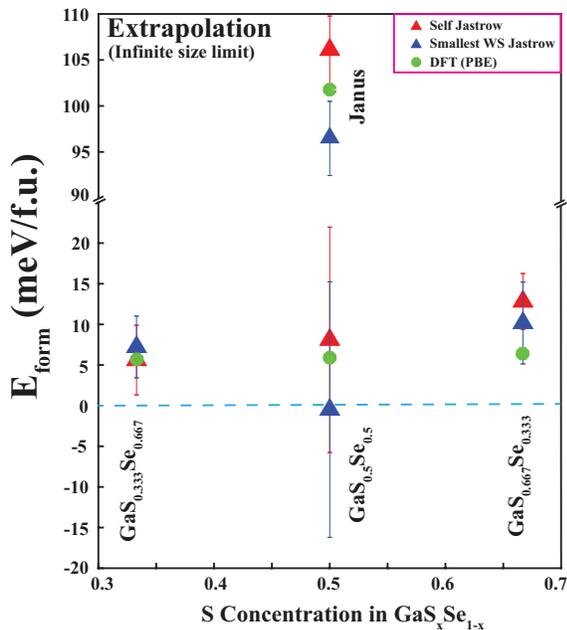


FIG. 4. The partially constructed energy hull diagram for  $\text{GaS}_x\text{Se}_{1-x}$  calculated with DFT (PBE) in green and DMC (using *self-Jastrows* in red and *WS-Jastrows* in blue), with associated error bars.

DMC calculations, where the Jastrow parameters are optimized separately for each structure including the different sized supercells used in the finite size extrapolations. Otherwise, all the settings in the DMC and VMC calculations are uniform across all the calculations, including the number of cores and number of walkers used. In this section we analyze

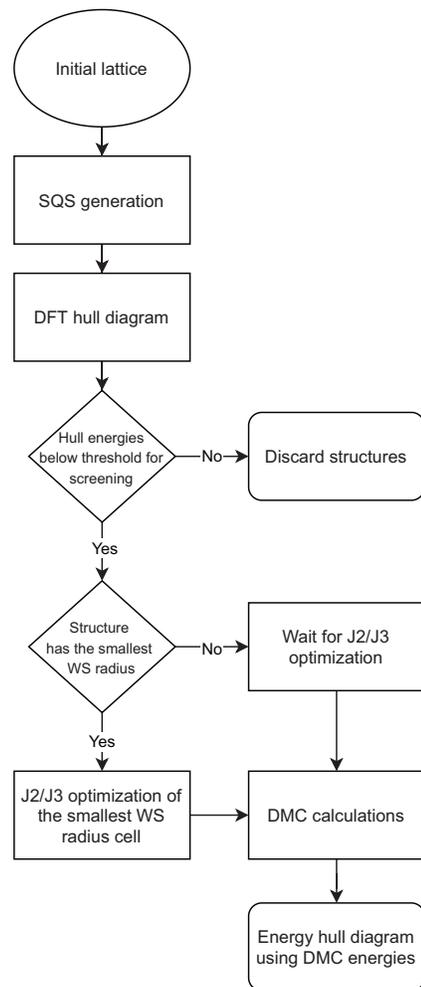


FIG. 5. The full high throughput workflow proposed in this work to obtain the energetics of an alloy system with QMC methods.

the computational cost associated with *Jastrow sharing*. We can see from our results in Table I that the VMC and DMC energies are converging to the appropriate ground states with comparable localization error using the same VMC and DMC parameters on the same number of nodes/cores, which demonstrates that *Jastrow sharing* for  $\text{GaS}_x\text{Se}_{1-x}$  is a valid approach. Therefore, we report the computational time for these two procedures.

Before proceeding any further, we provide details of our computational environment. All of the VMC and DMC simulations are conducted on 8 nodes of our local cluster. Each node is connected via OmniPath networking and contains 2 Intel 20-Core Xeon Gold 6138 CPUs clocked at 2GHz and 384 GB of 2666MHz DDR4 memory. QMCPack version 3.9.0 is compiled using OpenMPI version 3.0.1 and Intel Fortran, C and C++ (version 19.0.4.243) compiler.

To quantitatively investigate the computational cost savings of *Jastrow sharing*, we report the computational time of VMC and DMC (total time including every twist) in Table III. We report the computational time obtained with parameters we

VMC Computational Time			
	1 Supercell	Extrapolation	Energy Hull
Self Jastrow	$755 \pm 180$	$7,124 \pm 1,816$	$24,043 \pm 5,881$
WS Jastrow	$134 \pm 40$	$134 \pm 40$	$134 \pm 40$

DMC Computational Time			
	1 Supercell	Extrapolation	Energy Hull
Self Jastrow	$2,143 \pm 1,071$	$32,899 \pm 16,450$	$93,066 \pm 46,533$
WS Jastrow	$2,109 \pm 1,054$	$33,259 \pm 16,630$	$93,170 \pm 46,616$

TABLE III. A measure of the computational cost (time in seconds) using Self Jastrows and WS Jastrows for 1 supercell, one extrapolated result, and the entire energy hull diagram for VMC (upper half) and DMC (lower half) simulations. The quantity in the  $\pm$  gives the upper and lower bound on cost (discarding unnecessary QMC steps after target error is reached).

used in our QMC simulations (upper bound) and calculate a lower bound, which is obtained by removing the number of extra unnecessary steps after a target uncertainty was reached ( $\sim 10$  meV/f.u. for VMC,  $\sim 5$  meV/f.u. for DMC). We used the 36 atom supercell of  $\text{GaS}_{0.333}\text{Se}_{0.667}$  as an example of a single supercell (second column of Table III). As seen from Table III (second column), the VMC time associated with optimizing the Jastrow parameters of the smallest WS structure is much lower due to the fact that the WS structure only has 16 atoms (while the other supercell has 36 atoms).

To report total cost savings for all of the calculations performed in this paper, we give the VMC and DMC time needed to achieve the extrapolated results (using  $x = 0.333$  as an example, see third column of Table III). We also report the computational time to achieve the extrapolated results for the four considered structures (Janus,  $x = 0.5$ ,  $x = 0.333$ , and  $x = 0.667$ ) to get an estimate of the cost savings for constructing the energy hull diagram (see Fig. 4). Since the DMC extrapolated results using the *WS-Jastrow* only requires one VMC optimization simulation (extra VMC calculations are avoided), the VMC time for the extrapolation is the same as the single supercell (third column of Table III). The substantial savings of constructing the hull diagram (at the VMC level) is indicated in the fourth column of Table III, where the VMC computational time of using *WS-Jastrows* is two orders of magnitude smaller than using *self-Jastrows*.

From Table III, we observe that using either the *self-Jastrows* or the *WS-Jastrow* does not change the DMC computational time significantly, which means that *Jastrow-sharing* does not increase cost at the DMC level in our case. Specifically, each set of Jastrow parameters are of comparable quality and do not increase the uncertainty or amount of DMC steps needed to reach the target uncertainty for each calculation. It is clear from this cost analysis that by sharing Jastrow parameters, it is possible to achieve substantial savings by reducing the VMC computational time. The cost of VMC calculations is often disregarded when compared to DMC cost, but in our work we observe that when *self-Jastrows* are used, VMC accounts for roughly 1/4 (on average) of the total cost. Therefore by using the *WS-Jastrow* approach, the high-throughput QMC calculations of alloy systems can be significantly less costly.

## IV. CONCLUSION

We have outlined a high-throughput procedure to calculate the energy hull diagram for an alloy system using QMC methods. This involves optimizing the Jastrow parameters of the alloy with the smallest WS radius, and using these optimized parameters for all of the other subsequent structures. We tested the validity of this *Jastrow-sharing* procedure on monolayer  $\text{GaS}_x\text{Se}_{1-x}$  by swapping the Jastrow parameters of structures with different stoichiometries and tabulated the Jastrow sensitivities. Finally, we quantified the substantial computational cost savings obtained from avoiding extra VMC optimization simulations. We hope this method can be implemented for other systems where the Jastrow sensitivities of the subsequent pseudopotentials are low, which can aid in future accurate and high-throughput studies for alloys.

## SUPPLEMENTARY MATERIAL

See the supplementary information for additional details and convergence tests of DFT and DMC calculations, convergence criteria for the SQS and Cluster Expansion, DFT and DMC details about the optimal geometries of each structure, testing of *Jastrow-sharing* at the VMC level, and details about finite-size extrapolation.

## ACKNOWLEDGMENTS

This work was supported by the National Science Foundation through the Division of Materials Research under NSF DMR-1726213.

## DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request.

## REFERENCES

- <sup>1</sup>A. van de Walle, "Multicomponent multisublattice alloys, nonconfigurational entropy and other additions to the alloy theoretic automated toolkit," *Calphad* **33**, 266 – 278 (2009).
- <sup>2</sup>C. S. Jung, F. Shojaei, K. Park, J. Y. Oh, H. S. Im, D. M. Jang, J. Park, and H. S. Kang, "Red-to-ultraviolet emission tuning of two-dimensional gallium sulfide/selenide," *ACS Nano* **9**, 9585–9593 (2015).
- <sup>3</sup>F. Ersan, H. D. Ozaydin, and E. Aktürk, "Influence of chalcogen composition on the structural transition and on the electronic and optical properties of the monolayer titanium trichalcogenide ordered alloys," *Phys. Chem. Chem. Phys.* **20**, 1431–1439 (2018).
- <sup>4</sup>M. Kanli, A. Onen, A. Mogulkoc, and E. Durgun, "Characterization of two-dimensional  $\text{Ga}_{1-x}\text{Al}_x\text{N}$  ordered alloys with varying chemical composition," *Computational Materials Science* **167**, 13 – 18 (2019).
- <sup>5</sup>D. Wines, F. Ersan, and C. Ataca, "Engineering the electronic, thermoelectric, and excitonic properties of two-dimensional group-III nitrides through alloying for optoelec-

- tronic devices ( $B_{1-x}Al_xN$ ,  $Al_{1-x}Ga_xN$ , and  $Ga_{1-x}In_xN$ )," *ACS Applied Materials & Interfaces* **12**, 46416–46428 (2020).
- <sup>6</sup>H. Cai, B. Chen, M. Blei, S. L. Y. Chang, K. Wu, H. Zhuang, and S. Tongay, "Abnormal band bowing effects in phase instability crossover region of  $GaSe_{1-x}Te_x$  nanomaterials," *Nature Communications* **9**, 1927 (2018).
- <sup>7</sup>A. Zunger, S.-H. Wei, L. G. Ferreira, and J. E. Bernard, "Special quasirandom structures," *Phys. Rev. Lett.* **65**, 353–356 (1990).
- <sup>8</sup>W. M. C. Foulkes, L. Mitás, R. J. Needs, and G. Rajagopal, "Quantum Monte Carlo simulations of solids," *Rev. Mod. Phys.* **73**, 33–83 (2001).
- <sup>9</sup>R. J. Needs, M. D. Towler, N. D. Drummond, and P. L. Ríos, "Continuum Variational and Diffusion Quantum Monte Carlo calculations," *Journal of Physics: Condensed Matter* **22**, 023201 (2009).
- <sup>10</sup>D. Wines, K. Saritas, and C. Ataca, "A first-principles Quantum Monte Carlo study of two-dimensional (2D)  $GaSe$ ," *The Journal of Chemical Physics* **153**, 154704 (2020), <https://doi.org/10.1063/5.0023223>.
- <sup>11</sup>H. Shin, J. T. Krogel, K. Gasperich, P. R. C. Kent, A. Benali, and O. Heinonen, "Optimized structure and electronic band gap of monolayer  $GeSe$  from Quantum Monte Carlo methods," *Phys. Rev. Materials* **5**, 024002 (2021).
- <sup>12</sup>T. Frank, R. Derian, K. Tokár, L. Mitás, J. Fabian, and I. Štich, "Many-body Quantum Monte Carlo study of 2D materials: Cohesion and band gap in single-layer phosphorene," *Phys. Rev. X* **9**, 011018 (2019).
- <sup>13</sup>L. Shulenburg, A. D. Baczewski, Z. Zhu, J. Guan, and D. Tománek, "The nature of the interlayer interaction in bulk and few-layer phosphorus," *Nano Letters* **15**, 8170–8175 (2015).
- <sup>14</sup>A. Annaberdiyev, G. Wang, C. A. Melton, M. C. Bennett, and L. Mitás, "Cohesion and excitations of diamond structure silicon by Quantum Monte Carlo: benchmarks and control of systematic biases," (2021), [arXiv:2102.11998 \[cond-mat.mtrl-sci\]](https://arxiv.org/abs/2102.11998).
- <sup>15</sup>H. Hao, J. Shee, S. Upadhyay, C. Ataca, K. D. Jordan, and B. M. Rubenstein, "Accurate predictions of electron binding energies of dipole-bound anions via Quantum Monte Carlo methods," *The Journal of Physical Chemistry Letters* **9**, 6185–6190 (2018).
- <sup>16</sup>E. Mostaani, N. D. Drummond, and V. I. Fal'ko, "Quantum Monte Carlo calculation of the binding energy of bilayer graphene," *Phys. Rev. Lett.* **115**, 115501 (2015).
- <sup>17</sup>Y. Kadioglu, J. A. Santana, H. D. Ozaydin, F. Ersan, O. U. Akturk, E. Akturk, and F. A. Reboredo, "Diffusion Quantum Monte Carlo and density functional calculations of the structural stability of bilayer arsenene," *The Journal of Chemical Physics* **148**, 214706 (2018).
- <sup>18</sup>M. Szyniszewski, E. Mostaani, N. D. Drummond, and V. I. Fal'ko, "Binding energies of trions and biexcitons in two-dimensional semiconductors from Diffusion Quantum Monte Carlo calculations," *Phys. Rev. B* **95**, 081301 (2017).
- <sup>19</sup>M. Szyniszewski, E. Mostaani, N. D. Drummond, and V. I. Fal'ko, "Erratum: Binding energies of trions and biexcitons in two-dimensional semiconductors from Diffusion Quantum Monte Carlo calculations [Phys. Rev. B **95**, 081301(R) (2017)]," *Phys. Rev. B* **96**, 119902 (2017).
- <sup>20</sup>E. Mostaani, M. Szyniszewski, C. H. Price, R. Maezono, M. Danovich, R. J. Hunt, N. D. Drummond, and V. I. Fal'ko, "Diffusion Quantum Monte Carlo study of excitonic complexes in two-dimensional transition-metal dichalcogenides," *Phys. Rev. B* **96**, 075431 (2017).
- <sup>21</sup>R. J. Hunt, B. Monserrat, V. Zólyomi, and N. D. Drummond, "Diffusion Quantum Monte Carlo and GW study of the electronic properties of monolayer and bulk hexagonal boron nitride," *Phys. Rev. B* **101**, 205115 (2020).
- <sup>22</sup>K. Saritas, J. T. Krogel, S. Okamoto, H. N. Lee, and F. A. Reboredo, "Structural, electronic, and magnetic properties of bulk and epitaxial  $LaCoO_3$  through Diffusion Monte Carlo," *Phys. Rev. Materials* **3**, 124414 (2019).
- <sup>23</sup>K. Saritas, J. T. Krogel, P. R. C. Kent, and F. A. Reboredo, "Diffusion Monte Carlo: A pathway towards an accurate theoretical description of manganese oxides," *Phys. Rev. Materials* **2**, 085801 (2018).
- <sup>24</sup>Y. Luo, A. Benali, L. Shulenburg, J. T. Krogel, O. Heinonen, and P. R. C. Kent, "Phase stability of  $TiO_2$  polymorphs from Diffusion Quantum Monte Carlo," *New Journal of Physics* **18**, 113049 (2016).
- <sup>25</sup>A. Benali, L. Shulenburg, J. T. Krogel, X. Zhong, P. R. C. Kent, and O. Heinonen, "Quantum Monte Carlo analysis of a charge ordered insulating antiferromagnet: the  $Ti_4O_7$  Magnéli phase," *Phys. Chem. Chem. Phys.* **18**, 18323–18335 (2016).
- <sup>26</sup>J. A. Santana, J. T. Krogel, J. Kim, P. R. C. Kent, and F. A. Reboredo, "Structural stability and defect energetics of  $ZnO$  from Diffusion quantum Monte Carlo," *The Journal of Chemical Physics* **142**, 164705 (2015).
- <sup>27</sup>H. Shin, Y. Luo, P. Ganesh, J. Balachandran, J. T. Krogel, P. R. C. Kent, A. Benali, and O. Heinonen, "Electronic properties of doped and defective  $NiO$ : A Quantum Monte Carlo study," *Phys. Rev. Materials* **1**, 073603 (2017).
- <sup>28</sup>H. Shin, A. Benali, Y. Luo, E. Crabb, A. Lopez-Bezanilla, L. E. Ratcliff, A. M. Jokisaari, and O. Heinonen, "Zirconia and hafnia polymorphs: Ground-state structural properties from Diffusion Monte Carlo," *Phys. Rev. Materials* **2**, 075001 (2018).
- <sup>29</sup>K. Saritas, J. T. Krogel, and F. A. Reboredo, "Relative energies and electronic structures of  $CoO$  polymorphs through ab initio Diffusion Quantum Monte Carlo," *Phys. Rev. B* **98**, 155130 (2018).
- <sup>30</sup>K. Saritas, W. Ming, M.-H. Du, and F. A. Reboredo, "Excitation energies of localized correlated defects via Quantum Monte Carlo: A case study of  $Mn^{4+}$ -doped phosphors," *The Journal of Physical Chemistry Letters* **10**, 67–74 (2019).
- <sup>31</sup>K. Saritas, E. R. Fadel, B. Kozinsky, and J. C. Grossman, "Charge density and redox potential of  $LiNiO_2$  using ab initio Diffusion Quantum Monte Carlo," *The Journal of Physical Chemistry C* **124**, 5893–5901 (2020).
- <sup>32</sup>R. Nazarov, L. Shulenburg, M. Morales, and R. Q. Hood, "Benchmarking the pseudopotential and fixed-node approximations in diffusion monte carlo calculations of molecules and solids," *Phys. Rev. B* **93**, 094111 (2016).
- <sup>33</sup>G. Wang, A. Annaberdiyev, and L. Mitás, "Binding and excitations in  $Si_xH_y$  molecular systems using Quantum Monte Carlo," *The Journal of Chemical Physics* **153**, 144303 (2020), <https://doi.org/10.1063/5.0022814>.
- <sup>34</sup>M. C. Bennett, G. Hu, G. Wang, O. Heinonen, P. R. C. Kent, J. T. Krogel, and P. Ganesh, "Origin of metal-insulator transitions in correlated perovskite metals," (2021), [arXiv:2103.09809 \[cond-mat.mtrl-sci\]](https://arxiv.org/abs/2103.09809).
- <sup>35</sup>A. Jain, G. Hautier, C. J. Moore, S. Ping Ong, C. C. Fischer, T. Mueller, K. A. Persson, and G. Ceder, "A high-throughput infrastructure for density functional theory calculations," *Computational Materials Science* **50**, 2295–2310 (2011).
- <sup>36</sup>A. Zen, J. G. Brandenburg, J. Klimeš, A. Tkatchenko, D. Alfè, and A. Michaelides, "Fast and accurate Quantum Monte Carlo for molecular crystals," *Proceedings of the National Academy of Sciences* **115**, 1724–1729 (2018).
- <sup>37</sup>K. Saritas, T. Mueller, L. Wagner, and J. C. Grossman, "Investigation of a Quantum Monte Carlo protocol to achieve high accuracy and high-throughput materials formation energies," *Journal of Chemical Theory and Computation* **13**, 1943–1951 (2017).
- <sup>38</sup>V. Wang, Z. Q. Wu, Y. Kawazoe, and W. T. Geng, "Tunable band gaps of  $In_xGa_{1-x}N$  alloys: From bulk to two-dimensional limit," *The Journal of Physical Chemistry C* **122**, 6930–6942 (2018).
- <sup>39</sup>S. Yang, M. Wu, W. Shen, L. Huang, S. Tongay, K. Wu, B. Wei, Y. Qin, Z. Wang, C. Jiang, and C. Hu, "Highly sensitive polarization photodetection using a pseudo-one-dimensional  $Nb_{1-x}Ti_xS_3$  alloy," *ACS Applied Materials & Interfaces* **11**, 3342–3350 (2019).
- <sup>40</sup>F. Ersan, G. Gökoğlu, and E. Aktürk, "Adsorption and diffusion of lithium on monolayer transition metal dichalcogenides ( $MoS_{2(1-x)}Se_{2x}$ ) alloys," *The Journal of Physical Chemistry C* **119**, 28648–28653 (2015).
- <sup>41</sup>H. Mo, X. Zhang, Y. Liu, P. Kang, H. Nan, X. Gu, K. K. Ostrikov, and S. Xiao, "Two-dimensional alloying molybdenum tin disulfide monolayers with fast photoresponse," *ACS Applied Materials & Interfaces* **11**, 39077–39087 (2019).
- <sup>42</sup>I. Miyazato, S. Sarikurt, K. Takahashi, and F. Ersan, "Controlling electronic structure of single-layered  $HfX_3$  ( $X = S, Se$ ) trichalcogenides through systematic Zr doping," *Journal of Materials Science* **55**, 660–669 (2020).
- <sup>43</sup>F. Raffone, C. Ataca, J. C. Grossman, and G. Cicero, " $MoS_2$  enhanced T-phase stabilization and tunability through alloying," *The Journal of Physical Chemistry Letters* **7**, 2304–2309 (2016).
- <sup>44</sup>R. H. Almadvari, M. Nayeri, and S. Fotoohi, "Engineering of electronic and optical properties of monolayer gallium sulfide/selenide in presence of intrinsic atomic defects," *Materials Research Express* **7**, 015915 (2020).
- <sup>45</sup>H. T. Nguyen, M. M. Obeid, A. Bafekry, M. Idrees, T. V. Vu, H. V. Phuc, N. N. Hieu, L. T. Hoa, B. Amin, and C. V. Nguyen, "Interfacial characteristics, schottky contact, and optical performance of a  $graphene/Ga_2S_6$  van der waals heterostructure: Strain engineering and electric field tunability,"

- Phys. Rev. B **102**, 075414 (2020).
- <sup>46</sup>H. Cai, E. Soignard, C. Ataca, B. Chen, C. Ko, T. Aoki, A. Pant, X. Meng, S. Yang, J. Grossman, F. D. Ogletree, and S. Tongay, "Band engineering by controlling vdW epitaxy growth mode in 2D gallium chalcogenides," *Advanced Materials* **28**, 7375–7382 (2016).
- <sup>47</sup>F. Ersan and C. Ataca, "Janus PtX<sub>n</sub>Y<sub>2-n</sub>(X, Y = S, Se, Te; 0 ≤ n ≤ 2) monolayers for enhanced photocatalytic water splitting," *Phys. Rev. Applied* **13**, 064008 (2020).
- <sup>48</sup>S. Zhou, C.-C. Liu, J. Zhao, and Y. Yao, "Monolayer group-III monochalcogenides by oxygen functionalization: a promising class of two-dimensional topological insulators," *npj Quantum Materials* , 16.
- <sup>49</sup>L. Ao, H. Y. Xiao, X. Xiang, S. Li, K. Z. Liu, H. Huang, and X. T. Zu, "Functionalization of a GaSe monolayer by vacancy and chemical element doping," *Phys. Chem. Chem. Phys.* **17**, 10737–10748 (2015).
- <sup>50</sup>X. Meng, A. Pant, H. Cai, J. Kang, H. Sahin, B. Chen, K. Wu, S. Yang, A. Suslu, F. M. Peeters, and S. Tongay, "Engineering excitonic dynamics and environmental stability of post-transition metal chalcogenides by pyridine functionalization technique," *Nanoscale* **7**, 17109–17115 (2015).
- <sup>51</sup>X. Li, M.-W. Lin, J. Lin, B. Huang, A. A. Puzos, C. Ma, K. Wang, W. Zhou, S. T. Pantelides, M. Chi, I. Kravchenko, J. Fowlkes, C. M. Rouleau, D. B. Geohegan, and K. Xiao, "Two-dimensional GaSe/MoSe<sub>2</sub> misfit bilayer heterojunctions by van der Waals epitaxy," *Science Advances* **2** (2016), 10.1126/sciadv.1501882.
- <sup>52</sup>N. Liu, S. Zhou, N. Gao, and J. Zhao, "Tuning Schottky barriers for monolayer GaSe FETs by exploiting a weak Fermi level pinning effect," *Phys. Chem. Chem. Phys.* **20**, 21732–21738 (2018).
- <sup>53</sup>S. Yang, C. Wang, C. Ataca, Y. Li, H. Chen, H. Cai, A. Suslu, J. C. Grossman, C. Jiang, Q. Liu, and S. Tongay, "Self-driven photodetector and ambipolar transistor in atomically thin GaTe-MoS<sub>2</sub> p-n vdW heterostructure," *ACS Applied Materials & Interfaces* **8**, 2533–2539 (2016).
- <sup>54</sup>H. R. Jappor, "Electronic structure of novel GaS/GaSe heterostructures based on GaS and GaSe monolayers," *Physica B: Condensed Matter* **524**, 109 – 117 (2017).
- <sup>55</sup>R. Longuihoss and J. Ribeiro-Soares, "Monitoring the applied strain in monolayer gallium selenide through vibrational spectroscopies: A first-principles investigation," *Phys. Rev. Applied* **11**, 024012 (2019).
- <sup>56</sup>G. Liu, S. Xia, B. Hou, T. Gao, and R. Zhang, "Mechanical stabilities and nonlinear properties of monolayer gallium selenide under tension," *Modern Physics Letters B* **29**, 1550049 (2015).
- <sup>57</sup>C. Wang, S. Yang, H. Cai, C. Ataca, H. Chen, X. Zhang, J. Xu, B. Chen, K. Wu, H. Zhang, L. Liu, J. Li, J. C. Grossman, S. Tongay, and Q. Liu, "Enhancing light emission efficiency without color change in post-transition metal chalcogenides," *Nanoscale* **8**, 5820–5825 (2016).
- <sup>58</sup>D. Wines, J. A. Kropp, G. Chaney, F. Ersan, and C. Ataca, "Electronic properties of bare and functionalized two-dimensional (2D) tellurene structures," *Phys. Chem. Chem. Phys.* **22**, 6727–6737 (2020).
- <sup>59</sup>S. Yang, B. Chen, Y. Qin, Y. Zhou, L. Liu, M. Durso, H. Zhuang, Y. Shen, and S. Tongay, "Highly crystalline synthesis of tellurene sheets on two-dimensional surfaces: Control over helical chain direction of tellurene," *Phys. Rev. Materials* **2**, 104002 (2018).
- <sup>60</sup>G. Kresse and J. Furthmüller, "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set," *Phys. Rev. B* **54**, 11169–11186 (1996).
- <sup>61</sup>G. Kresse and D. Joubert, "From ultrasoft pseudopotentials to the projector augmented-wave method," *Phys. Rev. B* **59**, 1758–1775 (1999).
- <sup>62</sup>J. P. Perdew, K. Burke, and M. Ernzerhof, "Generalized gradient approximation made simple," *Phys. Rev. Lett.* **77**, 3865–3868 (1996).
- <sup>63</sup>J. Sun, A. Ruzsinszky, and J. P. Perdew, "Strongly constrained and appropriately normed semilocal density functional," *Phys. Rev. Lett.* **115**, 036402 (2015).
- <sup>64</sup>S. Grimme, "Semiempirical GGA-type density functional constructed with a long-range dispersion correction," *Journal of Computational Chemistry* **27**, 1787–1799 (2006).
- <sup>65</sup>S. Grimme, J. Antony, S. Ehrlich, and H. Krieg, "A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu," *The Journal of Chemical Physics* **132**, 154104 (2010), <https://doi.org/10.1063/1.3382344>.
- <sup>66</sup>H. Peng, Z.-H. Yang, J. P. Perdew, and J. Sun, "Versatile van der Waals density functional based on a meta-generalized gradient approximation," *Phys. Rev. X* **6**, 041005 (2016).
- <sup>67</sup>P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo, A. D. Corso, S. de Gironcoli, S. Fabris, G. Fratesi, R. Gebauer, U. Gerstmann, C. Gougoussis, A. Kokalj, M. Lazzeri, L. Martin-Samos, N. Marzari, F. Mauri, R. Mazzarello, S. Paolini, A. Pasquarello, L. Paulatto, C. Sbraccia, S. Scandolo, G. Sclauzero, A. P. Seitsonen, A. Smogunov, P. Umari, and R. M. Wentzcovitch, "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials," *Journal of Physics: Condensed Matter* **21**, 395502 (2009).
- <sup>68</sup>M. Burkatzki, C. Filippi, and M. Dolg, "Energy-consistent pseudopotentials for Quantum Monte Carlo calculations," *The Journal of Chemical Physics* **126**, 234105 (2007).
- <sup>69</sup>M. Burkatzki, C. Filippi, and M. Dolg, "Energy-consistent small-core pseudopotentials for 3d-transition metals adapted to Quantum Monte Carlo calculations," *The Journal of Chemical Physics* **129**, 164115 (2008).
- <sup>70</sup>A. L. Dzubak, J. T. Krogel, and F. A. Reboredo, "Quantitative estimation of localization errors of 3d transition metal pseudopotentials in Diffusion Monte Carlo," *The Journal of Chemical Physics* **147**, 024102 (2017).
- <sup>71</sup>J. Kim, A. D. Baczewski, T. D. Beaudet, A. Benali, M. C. Bennett, M. A. Berrill, N. S. Blunt, E. J. L. Borda, M. Casula, D. M. Ceperley, S. Chiesa, B. K. Clark, R. C. Clay, K. T. Delaney, M. Dewing, K. P. Esler, H. Hao, O. Heinonen, P. R. C. Kent, J. T. Krogel, I. Kylänpää, Y. W. Li, M. G. Lopez, Y. Luo, F. D. Malone, R. M. Martin, A. Mathuriya, J. McMinis, C. A. Melton, L. Mitas, M. A. Morales, E. Neuscamman, W. D. Parker, S. D. P. Flores, N. A. Romero, B. M. Rubenstein, J. A. R. Shea, H. Shin, L. Shulenburger, A. F. Tillack, J. P. Townsend, N. M. Tubman, B. V. D. Goetz, J. E. Vincent, D. C. Yang, Y. Yang, S. Zhang, and L. Zhao, "QMCPACK: an open source ab initio quantum monte carlo package for the electronic structure of atoms, molecules and solids," **30**, 195901 (2018).
- <sup>72</sup>P. R. C. Kent, A. Annaberdiyev, A. Benali, M. C. Bennett, E. J. Landinez Borda, P. Doak, H. Hao, K. D. Jordan, J. T. Krogel, I. Kylänpää, J. Lee, Y. Luo, F. D. Malone, C. A. Melton, L. Mitas, M. A. Morales, E. Neuscamman, F. A. Reboredo, B. Rubenstein, K. Saritas, S. Upadhyay, G. Wang, S. Zhang, and L. Zhao, "QMCPACK: Advances in the development, efficiency, and application of auxiliary field and real-space Variational and Diffusion Quantum Monte Carlo," *The Journal of Chemical Physics* **152**, 174105 (2020).
- <sup>73</sup>J. C. Slater, "The theory of complex spectra," *Phys. Rev.* **34**, 1293–1322 (1929).
- <sup>74</sup>R. Jastrow, "Many-body problem with strong forces," *Phys. Rev.* **98**, 1479–1484 (1955).
- <sup>75</sup>C. J. Umrigar and C. Filippi, "Energy and variance optimization of many-body wave functions," *Phys. Rev. Lett.* **94**, 150201 (2005).
- <sup>76</sup>L. Mitas, E. L. Shirley, and D. M. Ceperley, "Non-local pseudopotentials and Diffusion Monte Carlo," *The Journal of Chemical Physics* **95**, 3467–3475 (1991).
- <sup>77</sup>J. T. Krogel, "Nexus: A modular workflow management system for quantum simulation codes," *Computer Physics Communications* **198**, 154 – 168 (2016).
- <sup>78</sup>N. D. Drummond, M. D. Towler, and R. J. Needs, "Jastrow correlation factor for atoms, molecules, and solids," *Phys. Rev. B* **70**, 235119 (2004).
- <sup>79</sup>C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, "Alleviation of the fermion-sign problem by optimization of many-body wave functions," *Phys. Rev. Lett.* **98**, 110201 (2007).
- <sup>80</sup>X. Li, M.-W. Lin, A. A. Puzos, J. C. Idrobo, C. Ma, M. Chi, M. Yoon, C. M. Rouleau, I. I. Kravchenko, D. B. Geohegan, and K. Xiao, "Controlled vapor phase growth of single crystalline, two-dimensional GaSe crystals with high photoresponse," *Scientific Reports* **4**, 5497 (2014).
- <sup>81</sup>I. G. Buda, C. Lane, B. Barbiellini, A. Ruzsinszky, J. Sun, and A. Bansil, "Characterization of thin film materials using SCAN meta-GGA, an accurate nonempirical density functional," *Scientific Reports* **7**, 44766 (2017).
- <sup>82</sup>L. Zhang, J. Yu, M. Yang, Q. Xie, H. Peng, and Z. Liu, "Janus graphene from asymmetric two-dimensional chemistry," *Nature Communications* **4**, 1443 (2013).
- <sup>83</sup>A.-Y. Lu, H. Zhu, J. Xiao, C.-P. Chuu, Y. Han, M.-H. Chiu, C.-C. Cheng, C.-W. Yang, K.-H. Wei, Y. Yang, Y. Yang, D. Sokaras, D. Nordlund, P. Yang, D. A. Muller, M.-Y. Chou, X. Zhang, and L.-J. Li, "Janus monolayers of transition metal dichalcogenides," *Nature Nanotechnology* **12**, 744–749 (2017).

<sup>84</sup>Q.-l. Xiong, J. Zhou, J. Zhang, T. Kitamura, and Z.-h. Li, “Spontaneous curling of freestanding Janus monolayer transition-metal dichalcogenides,” *Phys. Chem. Chem. Phys.* **20**, 20988–20995 (2018).

<sup>85</sup>J. T. Krogel and P. R. C. Kent, “Magnitude of pseudopotential localization errors in fixed node Diffusion Quantum Monte Carlo,” *The Journal of Chemical Physics* **146**, 244101 (2017).

# Supplementary Information: A pathway towards high throughput Quantum Monte Carlo simulations for alloys: A case study of two-dimensional (2D) $\text{GaS}_x\text{Se}_{1-x}$

Daniel Wines,<sup>†</sup> Kayahan Saritas,<sup>‡</sup> and Can Ataca<sup>\*,†</sup>

<sup>†</sup>*Department of Physics, University of Maryland Baltimore County, Baltimore MD 21250*

<sup>‡</sup>*Department of Applied Physics, Yale University, New Haven CT 06520*

E-mail: ataca@umbc.edu

## DFT and DMC convergence testing and additional details

Table S1: To quantify the conversion of pseudopotentials, the calculated ionization potential and electron affinity for Ga, Se and S atoms calculated with the converted plane wave format BFD potentials (using PBE) are given. The experimental values are also given as a reference.

Method	Ionization Potential (eV)	Electron Affinity (eV)
PBE (BFD)	Ga: 5.89, <sup>1</sup> Se: 9.44, <sup>1</sup> S: 10.14	Ga: 0.32, <sup>1</sup> Se: 2.15, <sup>1</sup> S: 2.20
Exp	Ga: 6.00, <sup>2</sup> Se: 9.75, <sup>2</sup> S: 10.36 <sup>2</sup>	Ga: 0.30, <sup>3</sup> Se: 2.02, <sup>4</sup> S: 2.08 <sup>5</sup>

In order to convert these Gaussian potentials to plane wave format, we used the ppconvert<sup>6</sup> tool implemented in QMCPACK. We converted these Gaussian potentials to plane wave format and validated this conversion by calculating the ionization potential (IP) and

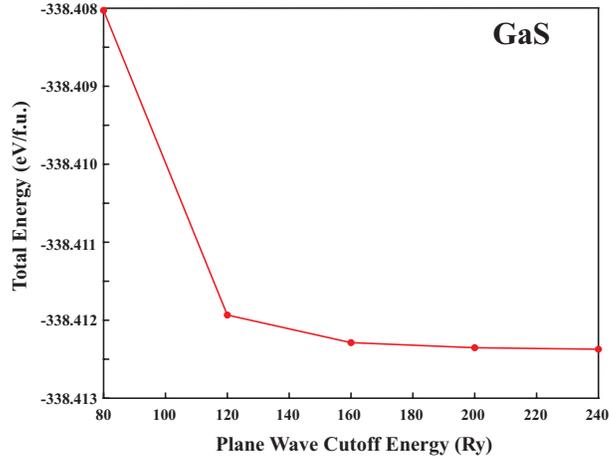


Figure S1: The total energy per formula unit of 2D GaS as a function of plane wave cutoff energy for the BFD pseudopotentials converted to plane wave format (calculated with PBE). The results show a converged value of 160 Ry.

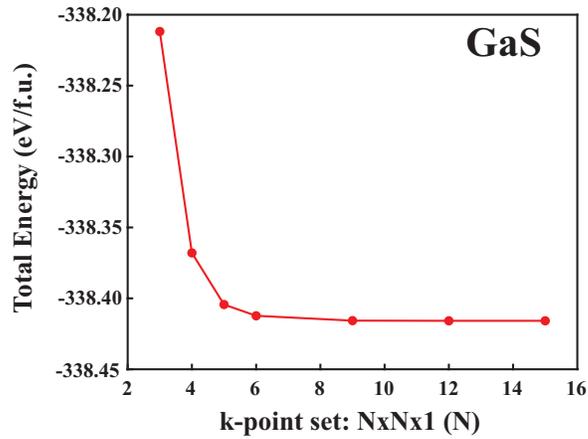


Figure S2: The total energy per formula unit of 2D GaS as a function of k-point grid for the BFD pseudopotentials converted to plane wave format (calculated with PBE). The results show a converged k-point grid of 6x6x1.

electron affinity (EA) at the DFT level. We used the p local channel for these BFD potentials. For our DFT benchmarking calculations in VASP, we used standard PBE PAW (as opposed to the hard or soft versions) potentials from version 5.4.4.

In the ATAT code, what is known as the Cluster Expansion (CE) formalism is performed.<sup>7</sup> In CE, the first-principles formation energies from the SQS structures are used as a training set to calculate fitted energies. CE can be used to validate the training set size of the SQS alloy and even predict the energetics of structures outside of the training set. When the fitted and calculated energies match within a convergence criteria known as the cross validation score, which is designed to estimate the error made in predicting the fitted energy of a structure and is analogous to the root mean square error, we know that we have considered enough random structures in our training set and can verify our predictions. In our CE calculations, we allowed the minimum cross validation score to be lower than 25 meV (thermal energy fluctuations at room temperature).<sup>7</sup> We achieved a smaller cross validation value after 60 simulations of alloys at varying concentrations.

Table S2: The calculated (using various methods) and experimental lattice constants of 2D GaS and GaSe. It is important to note that the DMC lattice constant of GaSe is 3.75(1) Å when LDA and SCAN wavefunctions are used, and 3.74(2) Å when PBE wavefunctions are used.<sup>1</sup>

<b>Functional</b>	<b>GaSe</b> a (Å)	<b>GaS</b> a (Å)
PBE	3.814	3.637
PBE-D2	3.745	3.581
PBE-D3	3.803	3.625
SCAN	3.761	3.599
SCAN+rvv10	3.754	3.589
DMC	3.75(1) <sup>1</sup> 3.74(2) <sup>1</sup>	3.549(3)
Experiment	3.755 <sup>8</sup> 3.74 <sup>9</sup>	3.587 <sup>8</sup>

The testing of our *Jastrow sharing* procedure at the VMC and DMC level is tabulated in Table S3, where we show the energy difference (in meV/f.u.). At the VMC level (first and second quadrant of Table S3), we observe a higher energy difference (up to 306 meV)

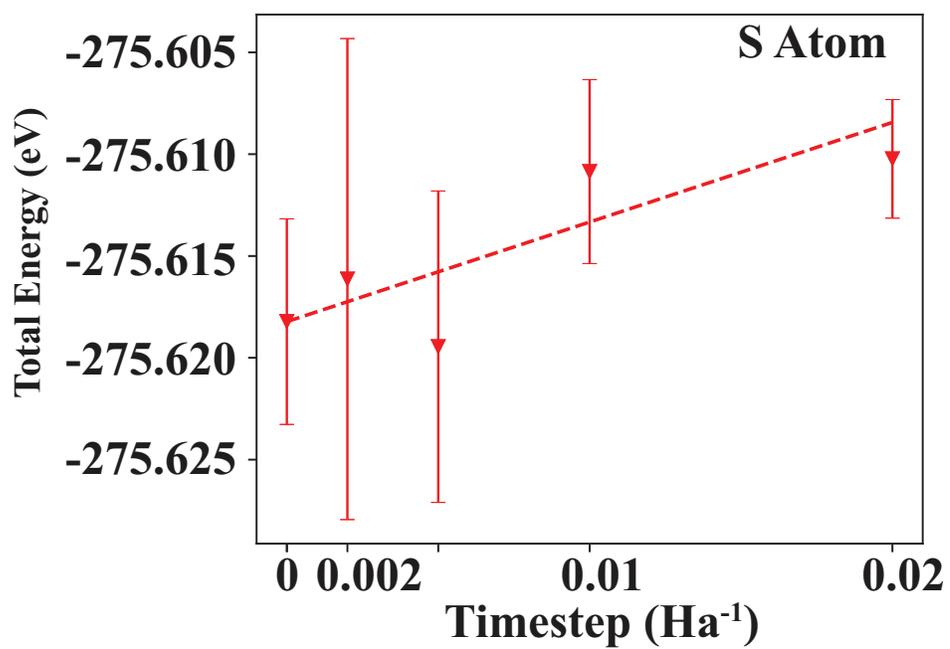


Figure S3: The timestep convergence test for the total energy of a single S atom. A timestep of  $0.01 \text{ Ha}^{-1}$  was used in the calculation of cohesive energy.

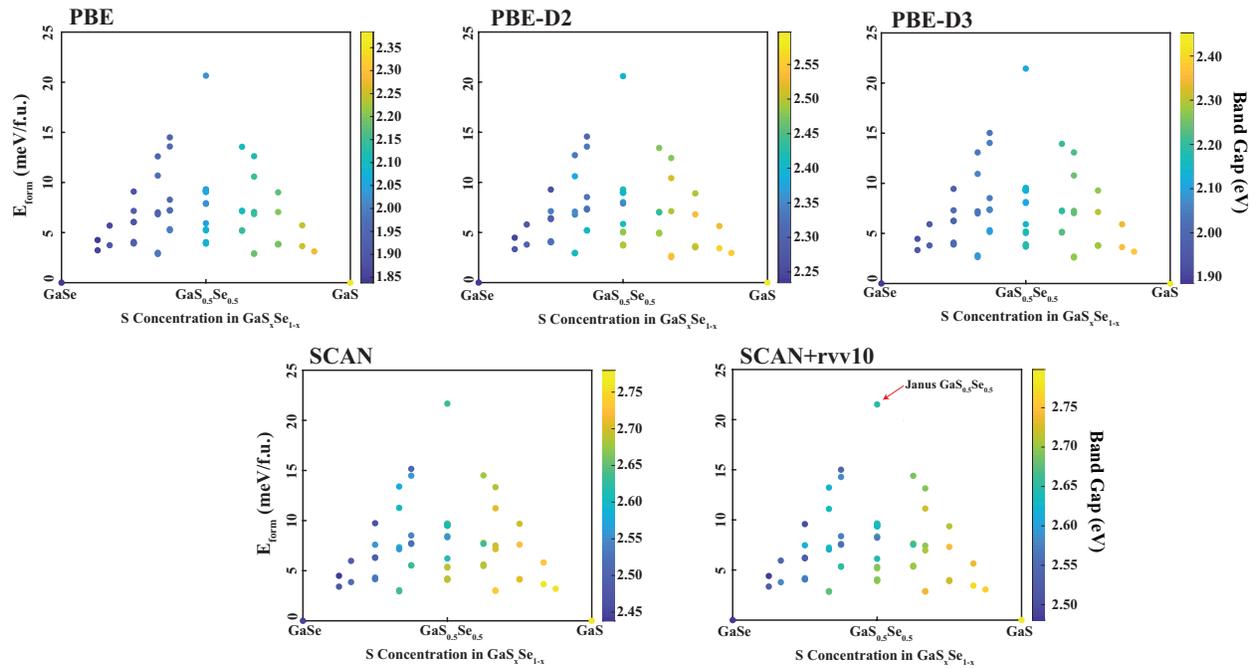


Figure S4: The formation energy per formula unit as a function of S concentration of 2D  $\text{GaS}_x\text{Se}_{1-x}$  alloys created by the SQS method. Each data point represents a different alloyed structure and the structure is optimized and energy is calculated with the PBE, PBE-D2, PBE-D3, SCAN and SCAN+rvv10 functionals. The color axis gives the calculated band gap of each structure.

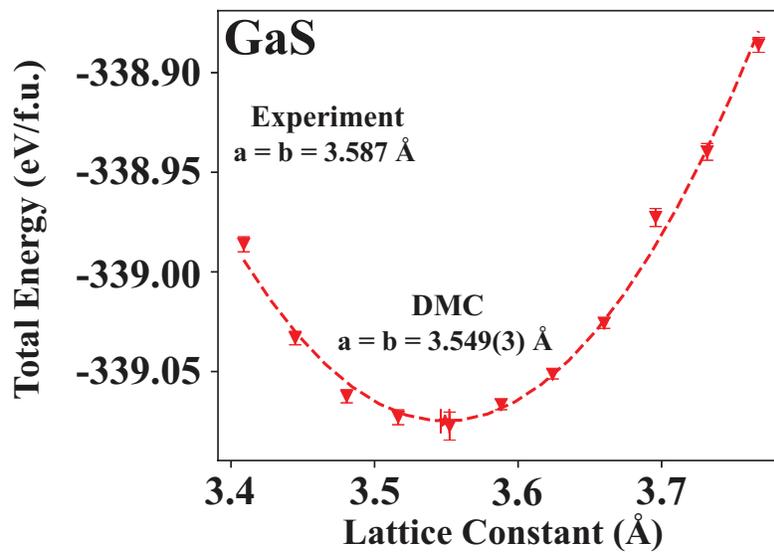


Figure S5: The total energy per formula unit of 2D hexagonal GaS (16 atom cell) calculated with DMC using starting wavefunctions from PBE versus the isotropically scaled lattice constant (in the  $x$  and  $y$  direction). The dotted line represents the fitted curve and the minimum point and respective error bar is marked on the curve (with the value given in the figure inset). The experimental lattice constant from reference<sup>8</sup> is also given in the inset for comparison.

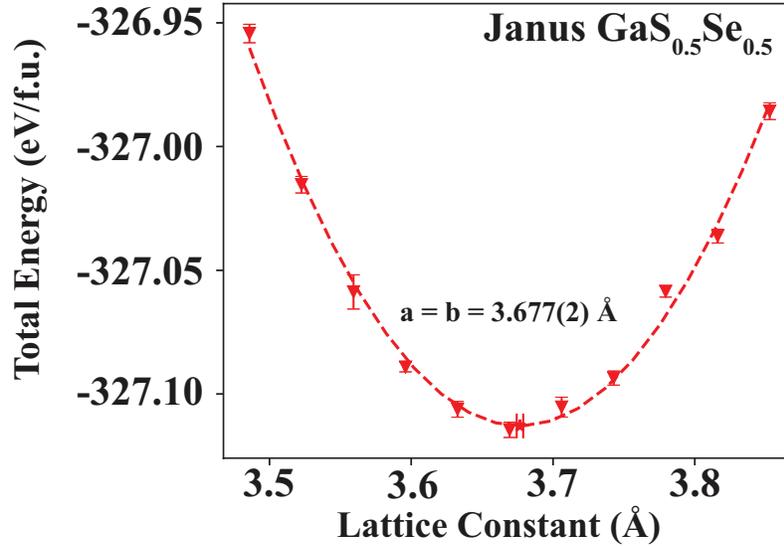


Figure S6: The total energy per formula unit of 2D hexagonal Janus GaS<sub>0.5</sub>Se<sub>0.5</sub> (16 atom cell) calculated with DMC using starting wavefunctions from PBE versus the isotropically scaled lattice constant (in the  $x$  and  $y$  direction). The dotted line represents the fitted curve and the minimum point and respective error bar is marked on the curve (with the value given in the figure inset). Geometric structure is depicted in Fig. 1 a).

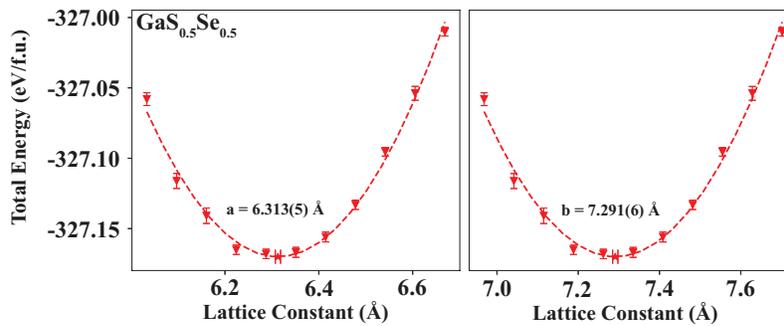


Figure S7: The total energy per formula unit of 2D tetragonal GaS<sub>0.5</sub>Se<sub>0.5</sub> (16 atom cell) calculated with DMC using starting wavefunctions from PBE versus the isotropically scaled lattice constants (in the  $x$  direction  $y$  directions respectively in each plot). The dotted lines represent the fitted curves and the minimum points and respective error bars are marked on the curves (with the value given in each of the figure insets). Geometric structure is depicted in Fig. 1 b).

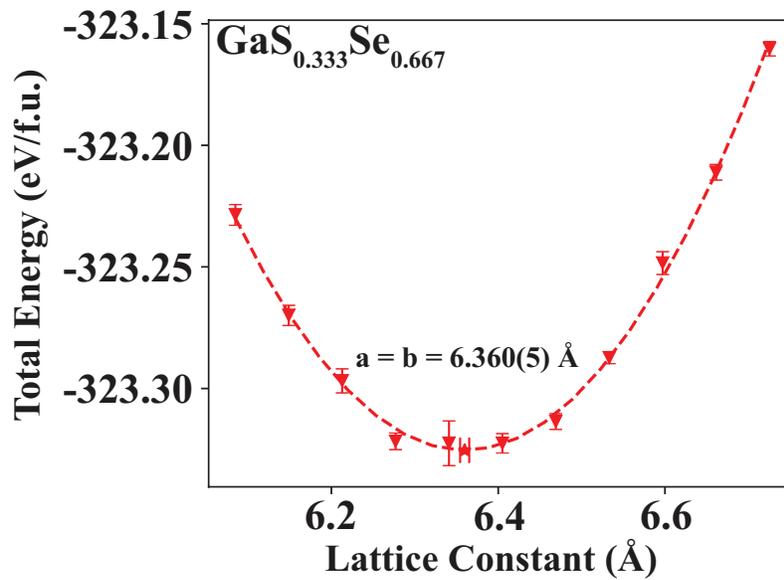


Figure S8: The total energy per formula unit of 2D hexagonal GaS<sub>0.333</sub>Se<sub>0.667</sub> (12 atom cell) calculated with DMC using starting wavefunctions from PBE versus the isotropically scaled lattice constant (in the  $x$  and  $y$  direction). The dotted line represents the fitted curve and the minimum point and respective error bar is marked on the curve (with the value given in the figure inset). Geometric structure is depicted in Fig. 1 c).

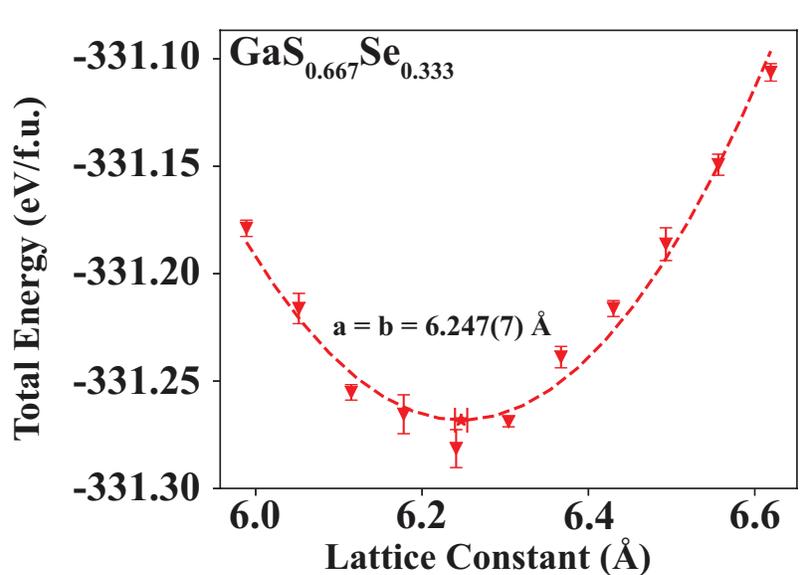


Figure S9: The total energy per formula unit of 2D hexagonal GaS<sub>0.667</sub>Se<sub>0.333</sub> (12 atom cell) calculated with DMC using starting wavefunctions from PBE versus the isotropically scaled lattice constant (in the  $x$  and  $y$  direction). The dotted line represents the fitted curve and the minimum point and respective error bar is marked on the curve (with the value given in the figure inset). Geometric structure is depicted in Fig. 1 d).

Table S3: The differences (in meV/f.u.) between total energies (VMC and DMC) calculated with *self-Jastrows* and *shared-Jastrows* including up to two-body and up to three-body terms with the associated error bars in parenthesis. The rows represent which Jastrows are used and the columns represent each alloyed structure for which the VMC and DMC energies are calculated.

<b>VMC-J3</b>	<b>Janus</b>	<b>x = 0.5</b>	<b>x = 0.333</b>	<b>x = 0.667</b>
Janus Jastrow	0	122(11)	120(13)	156(13)
x = 0.5 Jastrow	257(10)	0	33(15)	15(13)
x = 0.333 Jastrow	249(11)	-10(10)	0	8(12)
x = 0.667 Jastrow	245(10)	-15(11)	-2(13)	0
x = 0.875 Jastrow	230(10)	-47(10)	-38(13)	-28(12)
<b>VMC-J2</b>	<b>Janus</b>	<b>x = 0.5</b>	<b>x = 0.333</b>	<b>x = 0.667</b>
Janus Jastrow	0	28(12)	38(13)	18(14)
x = 0.5 Jastrow	300(12)	0	27(16)	-12(14)
x = 0.333 Jastrow	298(12)	-10(12)	0	-2(13)
x = 0.667 Jastrow	306(12)	21(12)	29(13)	0
x = 0.875 Jastrow	256(12)	-33(12)	-22(13)	-50(13)
<b>DMC-J3</b>	<b>Janus</b>	<b>x = 0.5</b>	<b>x = 0.333</b>	<b>x = 0.667</b>
Janus Jastrow	0	-6(6)	2(6)	3(10)
x = 0.5 Jastrow	0(4)	0	4(7)	15(11)
x = 0.333 Jastrow	-1(4)	-2(4)	0	17(10)
x = 0.667 Jastrow	-6(5)	-10(6)	-1(6)	0
x = 0.875 Jastrow	-2(5)	-15(7)	-2(5)	9(9)
<b>DMC-J2</b>	<b>Janus</b>	<b>x = 0.5</b>	<b>x = 0.333</b>	<b>x = 0.667</b>
Janus Jastrow	0	-3(7)	-9(7)	-6(7)
x = 0.5 Jastrow	3(4)	0	5(7)	-9(8)
x = 0.333 Jastrow	-7(5)	4(9)	0	6(9)
x = 0.667 Jastrow	-9(7)	-2(8)	-2(6)	0
x = 0.875 Jastrow	-8(4)	-5(8)	-2(6)	-3(7)

for the Janus structure using *shared-Jastrows* with two-body and three-body interaction terms. In addition, when the Janus Jastrow parameters are used for the other subsequent structures, the energy differences are on the order of  $\sim 10$  meV when two-body interactions are included and on the order of  $\sim 100$  meV when three-body interactions are included. Due to the fact that Janus  $\text{GaS}_{0.5}\text{Se}_{0.5}$  has an atomic environment that is slightly different than other 2D  $\text{GaS}_x\text{Se}_{1-x}$  structures, with one face of the material being entirely S and the other being entirely Se, *Jastrow sharing* at the VMC level can result in a larger energy difference between *shared-Jastrows* and *self-Jastrows*. A possible explanation for this is that the three-body interactions can be longer range, with electrons of Ga interacting with just S and electrons of Ga interacting with just Se on the respective surfaces (no electron mixing of S and Se).

Due to this, we reoptimized the three-body Jastrow parameters with a larger cutoff radius (3.5 Å) and used these for DMC calculations (see Table S4). We chose 3.5 Å because this is the largest value of cutoff radius that can be used for this cell size (smaller than the WS radius) and we expect to recover more three-body electron mixing interactions of S and Se. Ultimately we observe a smaller energy difference at the DMC and VMC level for the newly optimized set of three-body Jastrows with a larger cutoff radius. Although this has improved the results for the Janus structure, we cannot reuse these Jastrow parameters for the unit cells of the other alloys ( $x = 0.5, 0.333, 0.667$ ), since the Jastrow cutoff for three-body interactions is larger than the WS cell of the remaining structures. However, these newly optimized parameters could be effectively used in larger supercells.

For monolayer GaSe and GaS, the total energy was calculated at supercell sizes of 16, 24, 36, 48, 64 and 72 atoms. For Janus  $\text{GaS}_{0.5}\text{Se}_{0.5}$ , supercells of 16, 24, 36 and 64 atoms were used. For  $\text{GaS}_{0.5}\text{Se}_{0.5}$  supercells of 32, 48 and 64 were used and for the  $\text{GaS}_{0.333}\text{Se}_{0.667}$  and  $\text{GaS}_{0.667}\text{Se}_{0.333}$  structures, supercells of 24, 36, 48 and 72 atoms were used.

Table S4: The differences (in meV/f.u.) between total energies (VMC and DMC) calculated with *self-Jastrows* and *shared-Jastrows* up to three-body terms with the associated error bars in parenthesis for the Janus structure (where Jastrows are optimized using two different J3 cutoff radii). The rows represent which Jastrows are used and the columns represent two different J3 cutoff radii used in the VMC and DMC calculations.

<b>VMC-J3</b>	<b>Janus</b> $r^{J3} = 2 \text{ \AA}$	<b>Janus</b> $r^{J3} = 3.5 \text{ \AA}$
Janus Jastrow	-	-
x = 0.5 Jastrow	257(10)	89(11)
x = 0.333 Jastrow	249(11)	-98(11)
x = 0.667 Jastrow	245(10)	-102(10)
x = 0.875 Jastrow	230(10)	-116(10)
<b>DMC-J3</b>	<b>Janus</b> $r^{J3} = 2 \text{ \AA}$	<b>Janus</b> $r^{J3} = 3.5 \text{ \AA}$
Janus Jastrow	-	-
x = 0.5 Jastrow	0(4)	12(4)
x = 0.333 Jastrow	-1(4)	11(4)
x = 0.667 Jastrow	-6(5)	5(5)
x = 0.875 Jastrow	-2(5)	9(5)

## References

- (1) Wines, D.; Saritas, K.; Ataca, C. A first-principles Quantum Monte Carlo study of two-dimensional (2D) GaSe. *The Journal of Chemical Physics* **2020**, *153*, 154704.
- (2) Lide, D. R. *CRC Handbook of chemistry and physics*; CRC Press, 2008.
- (3) Gibson, N. D.; Walter, C. W.; Crocker, C.; Wang, J.; Nakayama, W.; Yukich, J. N.; Eliav, E.; Kaldor, U. Electron affinity of gallium and fine structure of  $\text{Ga}^-$ : Experiment and theory. *Phys. Rev. A* **2019**, *100*, 052512.
- (4) Vandevraye, M.; Drag, C.; Blondel, C. Electron affinity of selenium measured by photodetachment microscopy. *Phys. Rev. A* **2012**, *85*, 015401.
- (5) Chaibi, W.; Peláez, R. J.; Blondel, C.; Drag, C.; Delsart, C. Effect of a magnetic field in photodetachment microscopy. *The European Physical Journal D* **2010**, *58*, 29–37.

- (6) Kim, J. et al. QMCPACK: an open source ab initio quantum Monte Carlo package for the electronic structure of atoms, molecules and solids. **2018**, *30*, 195901.
- (7) van de Walle, A. Multicomponent Multisublattice Alloys, Nonconfigurational Entropy and other Additions to the Alloy Theoretic Automated Toolkit. *Calphad* **2009**, *33*, 266 – 278.
- (8) Jung, C. S.; Shojaei, F.; Park, K.; Oh, J. Y.; Im, H. S.; Jang, D. M.; Park, J.; Kang, H. S. Red-to-Ultraviolet Emission Tuning of Two-Dimensional Gallium Sulfide/Selenide. *ACS Nano* **2015**, *9*, 9585–9593.
- (9) Li, X.; Lin, M.-W.; Puzos, A. A.; Idrobo, J. C.; Ma, C.; Chi, M.; Yoon, M.; Rouleau, C. M.; Kravchenko, I. I.; Geoghegan, D. B.; Xiao, K. Controlled Vapor Phase Growth of Single Crystalline, Two-Dimensional GaSe Crystals with High Photoresponse. *Scientific Reports* **2014**, *4*, 5497.