

**Atomic forces by quantum Monte Carlo: Application to phonon dispersion calculations**Kousuke Nakano <sup>1,2,\*</sup>, Tommaso Morresi,<sup>3</sup> Michele Casula,<sup>3</sup> Ryo Maezono,<sup>2</sup> and Sandro Sorella<sup>1,†</sup><sup>1</sup>*International School for Advanced Studies (SISSA), Via Bonomea 265, 34136 Trieste, Italy*<sup>2</sup>*Japan Advanced Institute of Science and Technology (JAIST), Asahidai 1-1, Nomi, Ishikawa 923-1292, Japan*<sup>3</sup>*Institut de Minéralogie, de Physique des Matériaux et de Cosmochimie (IMPMC), Sorbonne Université, CNRS UMR 7590, IRD UMR 206, MNHN, 4 Place Jussieu, 75252 Paris, France*

(Received 2 December 2020; revised 16 February 2021; accepted 16 February 2021; published 15 March 2021)

We report a successful application of the *ab initio* quantum Monte Carlo (QMC) framework to a phonon dispersion calculation. A full phonon dispersion of diamond is successfully calculated at the variational Monte Carlo (VMC) level, based on the frozen-phonon technique. The VMC-phonon dispersion is in good agreement with the experimental results, giving renormalized harmonic optical frequencies very close to the experimental values, and improving upon previous density functional theory estimates. Key to success for the QMC approach is the statistical error reduction in the atomic force evaluation. We show that this can be achieved by using well conditioned atomic basis sets and by explicitly removing the basis-set redundancy, which reduces the statistical error of forces by up to two orders of magnitude by combining it with the so-called space-warp transformation algorithm. This leads to affordable and accurate QMC-phonons calculations, which are up to  $10^4$  times more efficient than a bare force treatment, and paves the way to new applications, particularly in correlated materials, where phonons have been poorly reproduced so far.

DOI: [10.1103/PhysRevB.103.L121110](https://doi.org/10.1103/PhysRevB.103.L121110)

The accurate description of phonons in a solid is one of the central research topics in the field of condensed matter physics and materials science for discussing phase stability (i.e., Gibbs-free energy), electron-phonon interaction, and structural phase transitions of materials [1,2]. *Ab initio* phonon calculations based on the Density Functional Theory (DFT) [3,4] have been successful for many compounds, but they often fail in strongly correlated materials. For example, DFT calculations severely underestimate the highest frequency of the optical phonons of graphene at the  $K$  point [5,6], because the electron-electron correlation is not taken into account with sufficient accuracy. Another example is the elemental cerium, whose phonon dispersions measured by neutron scattering strongly mismatch with the calculated DFT Perdew-Burke-Ernzerhof (DFT-PBE) ones [7]. Interestingly, such failure was also seen in a high  $T_c$  cuprate superconductor [8]. Some effort has been made to include correlation in phonon calculations in the DFT + Hubbard  $U$  (DFT+ $U$ ) framework [9] and also within the Dynamical Mean-Field Theory (DMFT) framework [10,11]. In both cases, this requires modeling correlations by an empirical parameter, though physically motivated (i.e., the Hubbard  $U$ ). Indeed, a genuine *ab initio* framework applicable to strongly correlated materials without any empirical parameters remains, so far, a very important theoretical challenge.

The *ab initio* quantum Monte Carlo (QMC) framework, including variational quantum Monte Carlo (VMC) and the diffusion quantum Monte Carlo (DMC) schemes, is among the state-of-art numerical methods to obtain highly accurate

many-body wave functions [12] and cope with the electron correlation more rigorously than DFT. It has been successfully applied to challenging materials that DFT cannot tackle, such as cuprates [13], iron arsenides [14], and graphene [13,15]. So far, unfortunately, almost all QMC applications are mainly based on energy and its first derivative (i.e., atomic force) calculations. Indeed, it is at present an open problem how to evaluate, with an affordable computational effort, second and higher-order derivatives, which are essential for computing various physical properties.

There are three routes to compute the second derivatives (i.e.,  $\frac{\partial^2 E}{\partial R_\alpha \partial R_\beta}$ ), which are needed for evaluating harmonic phonon properties by *ab initio* calculations, i.e., (i) potential energy surface (PES) fitting, (ii) finite difference expression based on atomic force evaluations, and (iii) direct evaluation of second derivatives. All of the above attempts have been successful for isolated molecular systems [16–19]. On the other hand, for solids, only strategy (i) has been successful so far within the QMC framework. For example, Maezono *et al.* calculated Raman frequencies of diamond (phonons at  $q = \Gamma$  point) [20]. However, QMC-phonon calculations of solids have been limited to a single high-symmetry  $q$  point. To the best of our knowledge, the full ( $q$ -resolved) phonon dispersion has been unaccessible so far at both VMC and DMC levels.

In this paper, we report a successful phonon dispersion calculation of diamond at the VMC level by adopting strategy (ii), the so-called frozen phonon technique [4]. The key to success is to reduce the statistical error of atomic forces. We found that removing the nearly linear dependency of the basis set used for the trial wave function parametrization [21] dramatically lowers the statistical error of forces. Its decrease reaches the order of  $\sim \times 10^{-2}$ , which corresponds to a speed

\*kousuke\_1123@icloud.com

†sorella@sissa.it

up of  $\sim 10^4$  times in a VMC computation. This drastic reduction enables us to construct a dynamical matrix within an affordable computational time, and to eventually apply VMC-phonon calculations to new, interesting materials from first principles.

All VMC and lattice regularized diffusion Monte Carlo (LRDMC) [29,30] calculations in this study were performed by the TURBORVB [31,32] SISSA quantum Monte Carlo package. We employed the Jastrow-Slater determinant (JSD) ansatz, i.e.,  $\Psi = \Psi_{\text{SD}} \times \exp J$ , where  $\Psi_{\text{SD}}$  and  $J$  are the Slater determinant and Jastrow terms, respectively. The Slater determinant part is expressed in terms of molecular orbitals  $\phi_k(\mathbf{r}) = \sum_{i=1}^L c_{i,k} \psi_i(\mathbf{r})$  expanded in a periodized Gaussian basis set  $\{\psi_i\}$ . The valence triple zeta (VTZ) basis set accompanying an energy-consistent pseudopotential developed by Burkatzki *et al.* [33] was employed for the primitive Gaussian atomic orbitals (Table S-I [34]). The coefficients of atomic orbitals (i.e.,  $c_{i,k}$ ) were obtained by a DFT calculation with the LDA-PZ exchange-correlation functional [35] and were left unchanged during the VMC optimization. The Jastrow factor was composed of inhomogeneous one-, two-, and three-body contributions ( $J = J_1^{\text{inh}} + J_2 + J_3$ ) [34]. The variational parameters in the Jastrow terms were optimized by the stochastic reconfiguration [36] and/or the modified linear method [32,37] implemented in TURBORVB. LRDMC calculations were performed by the single-grid scheme [29] with a lattice space,  $a = 0.2$  bohr.

In this paper, we focus on diamond (Space group:  $Fd\bar{3}m$ ) as a proof of concept for the first example of a VMC-based phonon dispersion calculation.  $2 \times 2 \times 2$  conventional supercells (256 electrons / 64 carbon atoms in a simulation cell) were used for most calculations, while  $3 \times 3 \times 3$  conventional supercells (864 electrons / 216 carbon atoms in a simulation cell) were also used for several calculations to investigate the finite-size errors.  $L$ -twist (i.e.,  $k = \pi, \pi, \pi$ ) was employed for alleviating the so-called one-body finite-size effects [20,38,39]. Dynamical matrices and the corresponding phonon dispersions were calculated based on the frozen-phonon technique implemented in the Phonopy module [4], where a 0.15 bohr displacement of carbon atoms was large enough to work with a reasonable signal/noise ratio in QMC forces. This displacement underestimates harmonic frequencies only by  $\sim 0.1$  THz, as shown in Fig. S-3 [34]. Error bars in a phonon dispersion were estimated by the jackknife method [40]. Phonon calculations based on DFT were performed using the QUANTUM ESPRESSO package [41] with LDA-PZ [35] and GGA-PBE [42] exchange-correlation functionals at the experimental lattice parameter [34]. The phonon dispersion of diamond has already been studied using DFT calculations by many groups so far at the theoretical [43,44] and experimental [20] lattice parameters. That makes this system a very good testbed for any new methodological implementation of phonons calculations. As shown later, our DFT calculations are consistent with the previous study.

Figure 1 shows the phonon dispersion obtained by our VMC calculations using the conventional  $2 \times 2 \times 2$  diamond supercell [49]. Observed experimental frequencies [22,24,27] are also plotted for comparison. A phonon dispersion obtained by the finite displacement method does not include anharmonic effects, which can decrease harmonic frequencies by

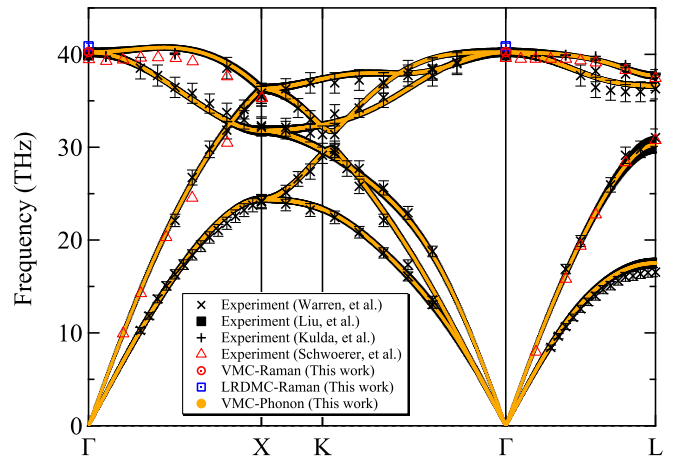


FIG. 1. The phonon dispersion and Raman frequencies of diamond calculated using a  $2 \times 2 \times 2$  conventional supercell at the VMC level, where the anharmonic and one-body finite size corrections are included. The experimental lattice parameter (i.e., 6.741 bohr [26]) was employed. The error bars were estimated by the jackknife method. Observed experimental results are also plotted for comparison. Experimental data are taken from Refs. [22–24,27]. The data in Ref. [23] were digitalized using WEBPLOTDIGITIZER [28].

up to  $\sim 5$ – $10\%$  for the lightest elements [50]. Therefore, for comparison with the experimental results, anharmonic corrections were added to the VMC phonon dispersion in this study. The phonon dispersion before the correction is shown in Fig. S-1 [34]. The anharmonic renormalizations were estimated in this study using path integral molecular dynamics simulations [51,52] at the DFT level with the PBE exchange-correlation function (Fig. S-5 [34]), giving  $-0.411$  THz,  $-0.177$  THz, and  $-0.280$  THz for  $\Gamma$ ,  $X$ , and  $L$ , respectively (see the Supplementary Material for details [34]). The value at  $\Gamma$  is consistent with a reported estimate of  $-17.4 \text{ cm}^{-1} = -0.522$  THz [25]. Notice that other possible sources of error were also considered. The phonon dispersion in Fig. 1 also includes the one-body finite size corrections that were estimated by DFT calculations (see Fig. S-2 [34]). The two-body finite-size error is negligible because the average density and the volume do not change in phonon calculations. Table I shows a detailed comparison of the highest harmonic phonon frequencies at three  $q$  points, i.e.,  $\Gamma = (0, 0, 0)$ ,  $X = (2\pi, 0, 0)$ , and  $L = (\pi, \pi, \pi)$ . Raman frequencies at  $\Gamma$  obtained by a direct fit of VMC energies, VMC forces, and LRDMC energies of the structures displaced along the corresponding eigenmode are also plotted in Fig. 1 [53]. In Table I and hereafter, (P) denotes the interpolated frequencies obtained by Phonopy, (F) denotes the phonon frequency at  $\Gamma$  obtained by force fitting, and (E) denotes the phonon frequency at  $\Gamma$  obtained by energy fitting. The corresponding Raman frequencies are consistent within the error bars, indicating that the Slater determinant obtained by DFT is almost optimal also in the presence of the Jastrow factor, thus explaining why this consistency is satisfied quite accurately [54]. In other words, if the wavefunction is at its minimum, the consistency is a consequence. If it is not at its minimum, the consistency may also be satisfied by chance or good behavior of the used basis set.

TABLE I. The highest harmonic phonon frequencies (THz) of diamond at high-symmetry  $q$  points. All phonons are calculated using the experimental lattice parameter. In this table, (P) denotes the interpolated frequencies obtained by Phonopy, (F) denotes the phonon frequency at  $\Gamma$  obtained by fitting forces of distorted structures along the Raman mode with a linear function at the VMC level, and (E) denotes the phonon frequency at  $\Gamma$  obtained by fitting energies of undistorted and distorted structures along the Raman mode with a quadratic function at the VMC and LRDMC levels. The last column indicates the harmonic frequencies estimated by the raw experimental values: 39.938 THz [22], 35.299 THz [23], and 37.962 THz [24] at the  $\Gamma$ ,  $X$ , and  $L$  points, respectively.

$q$	DFT		Previous work		This work <sup>d</sup>				Experiment
	LDA-PZ	GGA-PBE	VMC <sup>c</sup>	DMC <sup>c</sup>	VMC(P)	VMC(F)	VMC(E)	LRDMC(E)	Harmonic (Estimated)
$\Gamma$	38.55 <sup>a</sup>	38.82 <sup>b</sup>	41.64(9)	41.22(12)	40.65(38)	40.49(4)	40.68(29)	41.52(22)	40.460 <sup>e</sup> , 40.349 <sup>f</sup>
$X$	35.64	35.87	–	–	36.48(40)	–	–	–	35.476 <sup>f</sup>
$L$	37.31	37.47	–	–	38.01(31)	–	–	–	38.242 <sup>f</sup>

<sup>a</sup>38.40 THz in the previous DFT study employing the LDA functional. See Ref. [20].

<sup>b</sup>38.73 THz in the previous DFT study employing the GGA-PBE functional. See Ref. [20].

<sup>c</sup>These values are taken from Ref. [20].

<sup>d</sup>These values include the one-body finite size corrections, i.e.,  $-0.16$  THz,  $-0.18$  THz, and  $-0.23$  THz for  $q = \Gamma$ ,  $X$ , and  $L$ , respectively.

<sup>e</sup>The anharmonic renormalization,  $-17.4 \text{ cm}^{-1} = -0.522$  THz was employed. See Ref. [25].

<sup>f</sup>The anharmonic renormalizations are  $-0.411$  THz,  $-0.177$  THz, and  $-0.280$  THz for  $\Gamma$ ,  $X$ , and  $L$ , respectively, which were estimated by molecular dynamics simulations performed in this work.

All VMC (P), VMC (F), and VMC (D) calculations give the harmonic phonon frequency at  $\Gamma$  very close to the experimental value, considering the renormalization of the anharmonic effect, i.e., the discrepancy is just  $\sim 0.3$  THz. On the other hand, both DFT calculations with LDA-PZ and GGA-PBE exchange-correlation functionals underestimate the highest frequency at  $\Gamma$  by  $\sim 1.8$  THz and  $\sim 1.5$  THz, respectively. Table I shows that our VMC phonon dispersion calculation also gives accurate harmonic frequencies at other  $q$  points, i.e., at  $X$  and  $L$ . Compared with the previous VMC study [20], our VMC frequency at  $\Gamma$  point is closer to the experimental value, as reported in Table I [55]. The improvement at the VMC level certainly derives from the more flexible Jastrow factor employed in this study, while the one used in the previous VMC study was much simpler [20,56].

It is intriguing that our LRDMC calculation gives a 1.2(2) THz higher Raman frequency than the experimental one, as shown in Table I. The previous DMC study also overestimated the Raman frequency by 0.9(1) THz [20]. To discuss the origin of the discrepancy, we investigated the effect of the lattice-space error in our LRDMC calculations. An extrapolation ( $a \rightarrow 0$ ) with four lattice spaces (i.e., 0.20, 0.30, 0.40, and 0.50 bohr) yields 41.89(44) THz, suggesting that the lattice-space error is not the origin of the discrepancy. We also suspected that the experimental lattice parameter employed in the phonon calculation could be significantly different from the theoretical one, but this is also not the origin, as shown later. Therefore, the discrepancy should arise from the fixed-node error and this should be alleviated by a nodal surface optimization [32], which however is prohibitive in the  $2 \times 2 \times 2$  conventional supercells due to the large number of variational parameters of the distorted structures. A possible future work to study the nodal surface effect is the use of various exchange-correlation functionals suitable for solids such as HSE06 [57] and SCAN [58].

Since the equilibrium lattice parameter also affects phonon frequencies, we investigated the equation of states (EOSs) of diamond. Figure 2 shows plots of internal energies vs. vol-

umes fitted by the Vinet EOS [34]. Previous VMC and DMC results [20], and the experimentally observed equilibrium lattice parameter, are also plotted in Fig. 2 and summarized in Table II. In Fig. 2, the zero point energy (ZPE) contribution [20,59] is subtracted, to make the comparison possible with internal energies computed at  $T = 0$  and on static lattice. Table II shows that our VMC calculation reproduces the previous VMC study, while our LRDMC calculation gives a slightly smaller lattice parameter [6.702(1) bohr] than the previous DMC study [6.734(4) bohr]. This discrepancy likely derives from the different nodal surfaces used in the two studies, namely the one originating from the DFT-PBE orbitals in Ref. [20] and the one coming from the LDA orbitals in our work. Notice that both one-body and two-body finite size errors are negligible for the EOS calculations, as shown in Fig. S-6 [34]. Table II indicates that the equilibrium lattice

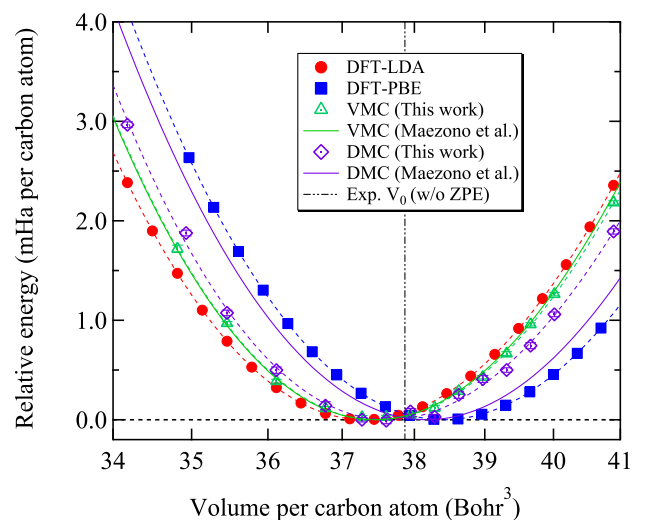


FIG. 2. Equation of states and the curves fitted by the Vinet exponential function [47]. VMC and DMC results (Maezono *et al.*) are taken from Ref. [20].

TABLE II. Vinet EOS parameters. Zero point energy (ZPE) and temperature effects (TE) are not included in these theoretical data.

Parameter	DFT		QMC				Experiment	
	LDA-PZ	GGA-PBE	VMC <sup>a</sup>	DMC <sup>a</sup>	VMC	LRDMC	w/o ZPE, w/o TE	Observed
Lattice (bohr)	6.683	6.748	6.691(4)	6.734(4)	6.693(1)	6.702(1)	6.7193(5) <sup>b</sup>	6.7410(5) <sup>c</sup>
$V_0$ (bohr <sup>3</sup> )	37.32	38.41	37.45(6)	38.17(6)	37.47(2)	37.63(2)	37.920(9) <sup>b</sup>	38.290(9) <sup>c</sup>
$B_0$ (GPa)	465	433	483(4)	448(3)	476(6)	463(5)	457(1) <sup>b</sup> , 453(5) <sup>b</sup>	446(1) <sup>c</sup> , 442(5) <sup>d</sup>
$B'_0$	3.65	3.70	3.8(1)	3.7(1)	4.0(6)	4.9(6)	3.0(1) <sup>b</sup> , 4.0(7) <sup>b</sup>	3.0(1) <sup>c</sup> , 4.0(7) <sup>d</sup>

<sup>a</sup>These values are taken from Ref. [20]. Here, ZPE and TE are subtracted.

<sup>b</sup>ZPE and TE are corrected,  $-0.37$  bohr<sup>3</sup>,  $+11$  GPa, and  $-0.03$  for  $V_0$ ,  $B_0$ , and  $B'_0$ , respectively. See Ref. [20].

<sup>c</sup>These values are taken from Ref. [45].

<sup>d</sup>These values are taken from Refs. [45,46].

parameter of diamond is  $\sim 0.03$  bohr ( $\sim 0.02$  bohr) smaller than the experimental one at the VMC (LRDMC) level. Our DFT calculations (Fig. S-4 [34]) suggest that the Raman frequency is implicitly proportional to the lattice parameter, with a gradient of  $-0.18$  THz /  $0.01$  bohr. Therefore, if the equilibrium lattice parameter (with ZPE) were employed instead of the experimental one, VMC (LRDMC) calculations would give  $\sim 0.5$  THz ( $\sim 0.4$  THz) higher Raman frequencies while still staying close to the experimental values.

Reducing the statistical errors of atomic forces is key to a successful phonon calculation. We found that alleviating the linear dependency of a localized atomic basis set drastically decreases the statistical error. In general, a basis set optimized for molecular systems is not suitable for solid state calculations (i.e., strongly linear dependent), due to the presence of orbitals having small exponents (c.f., typically  $< 0.1$ ) [60]. The quality of the basis set is systematically improved by a general and efficient scheme implemented in TURBORVB [21]. Indeed, the linear dependency of a localized atomic basis set  $[\psi_\mu(\vec{r})]$  is characterized by the condition number,  $\kappa(\mathbf{S})$ , where  $S_{\mu,\nu} = \langle \psi_\mu | \psi_\nu \rangle$  is the overlap matrix [61]. In TURBORVB, a redundant basis set is converted into a well-conditioned one by disregarding small eigenvalues and the corresponding eigenvectors of the overlap matrix  $\mathbf{S}$  [21,32]. We note that a well-conditioned basis set can also be constructed by simply removing orbitals having small exponents, while the method employed in this work is more general and systematic. Figure 3 shows the plots of VMC energies, VMC forces, and their statistical errors vs. the inverse of the condition number  $[\kappa(\mathbf{S})^{-1}]$ . Figure 3(a) indicates that the statistical error of the energy is independent of the condition number, at variance with the statistical error of the force, which instead strongly depends on it [Fig. 3(b)]. The error bar in the forces amounts to  $\sim 3.4 \times 10^{-2}$  (Ha/bohr) when the atomic basis set is strongly linear dependent [i.e.,  $\kappa(\mathbf{S})^{-1} = 10^{-15}$ ], a condition that also introduces some bias in the forces because, as we have verified, they are no longer consistent with the finite difference energy derivatives for  $\kappa(\mathbf{S})^{-1} < 10^{-12}$  [see Fig. 3(b)]. On the other hand, the statistical error becomes much smaller,  $\sim 1.7 \times 10^{-3}$  (Ha/bohr), by removing the linear dependency [i.e.,  $\kappa(\mathbf{S})^{-1} = 10^{-7}$ ]. The space warp coordinate transformation (SWCT) [48] is able to further reduce the statistical error. Indeed, Fig. 3(b) shows that the error bar of the force becomes  $\sim 4.1 \times 10^{-4}$  (Ha/bohr) by using the SWCT algorithm combined with a well-conditioned basis set, corresponding

to  $\sim 10^4$  times more efficient computation than a bare force treatment.

We analyze now in detail the reason of this behavior. TURBORVB evaluates atomic forces in a differential form (i.e., by the algorithmic differentiation) [48,62]:

$$F_\alpha = \left\langle \frac{dE}{dR_\alpha} \right\rangle_{|\Psi^{\mathbf{R}}|^2} \simeq \left\langle \frac{E(\mathbf{R} + \Delta\mathbf{R}_\alpha) - E(\mathbf{R})}{\Delta R_\alpha} \right\rangle_{|\Psi^{\mathbf{R}}|^2},$$

where  $\mathbf{R} = (R_1, \dots, R_\alpha, \dots)$  and  $\Delta\mathbf{R}_\alpha = (0, \dots, \Delta R_\alpha, \dots)$ . This equation suggests that the statistical error on forces depends on how much the wavefunction changes after an atom is displaced. In other words, to minimize the stochastic error, the overlap  $\langle \Psi^{\mathbf{R} + \Delta\mathbf{R}_\alpha} | \Psi^{\mathbf{R}} \rangle$  should be close to unity. To investigate the effect of the linear dependency on the overlap, we calculated  $\langle \Psi^{\mathbf{R} + \Delta\mathbf{R}_\alpha} | \Psi^{\mathbf{R}} \rangle$  with linear-dependent and linear-independent basis sets, using correlated sampling techniques [32], where only one carbon atom in the  $1 \times 1 \times 1$  conventional simulation cell was displaced in the  $x$  direction by  $\Delta R_\alpha = 0.005$  bohr. We obtained 0.9999 and 0.9726 for  $\kappa(\mathbf{S})^{-1} = 10^{-7}$  and  $\kappa(\mathbf{S})^{-1} = 10^{-15}$ , respectively. This clearly indicates that the linear dependency of the basis set deteriorates the overlap  $\langle \Psi^{\mathbf{R} + \Delta\mathbf{R}_\alpha} | \Psi^{\mathbf{R}} \rangle$ , thus increasing the statistical error of forces.

The deterioration is explained as follows. Here, a simple Slater wavefunction without Jastrow factor is considered for the sake of clarity. In this case, the overlap  $\langle \Psi^{\mathbf{R}} | \Psi^{\mathbf{R}} \rangle$  reads  $\langle \Psi^{\mathbf{R}} | \Psi^{\mathbf{R}} \rangle = \det(\phi_i^{\mathbf{R}} | \phi_j^{\mathbf{R}})$ , where  $\phi_i^{\mathbf{R}}$  is the  $i$ -th molecular orbital depending on nuclear positions  $\mathbf{R}$ , defining the above  $N \times N$  determinant matrix. The molecular orbital is expanded over localized atomic orbitals, i.e.,  $\phi_i^{\mathbf{R}} = \sum_{a,l} c_{i,(a,l)} \psi_{\{a,l\}}^{R_a}$ , where  $\psi_{\{a,l\}}^{R_a}$  are (periodized) atomic orbitals explicitly dependent on a nuclear position  $R_a$  ( $a = 1, \dots, \alpha, \dots$ ), while  $c_{i,(a,l)}$  are nuclear position independent. We can readily derive  $\langle \Psi^{\mathbf{R}} | \Psi^{\mathbf{R}} \rangle = 1$  when the molecular orbitals are orthonormalized (i.e.,  $\langle \phi_i^{\mathbf{R}} | \phi_j^{\mathbf{R}} \rangle = \delta_{i,j}$ ). What about  $\langle \Psi^{\mathbf{R} + \Delta\mathbf{R}_\alpha} | \Psi^{\mathbf{R}} \rangle$ ? We would like to show how the perturbation  $\frac{d}{dR_\alpha} \phi_i^{\mathbf{R}} \equiv \sum_{a=l} c_{i,(a,l)} \frac{d}{dR_\alpha} \psi_{\{a,l\}}^{R_a}$  affects the overlap. In DFT calculations, it turns out that  $|c_{i,(a,l)}| \gg 1$  when the basis is redundant [e.g.,  $\kappa(\mathbf{S})^{-1} = 10^{-15}$ ], while  $|c_{i,(a,l)}| \ll 1$  when the basis set is well-conditioned [e.g.,  $\kappa(\mathbf{S})^{-1} = 10^{-7}$ ]. When  $|c_{i,(a,l)}| \ll 1$ , the perturbation effect is rather small, and the orthonormalization condition almost holds, while  $|c_{i,(a,l)}| \gg 1$  makes the perturbation effect significant, and by consequence the orthonormalization condition is certainly deteriorated.



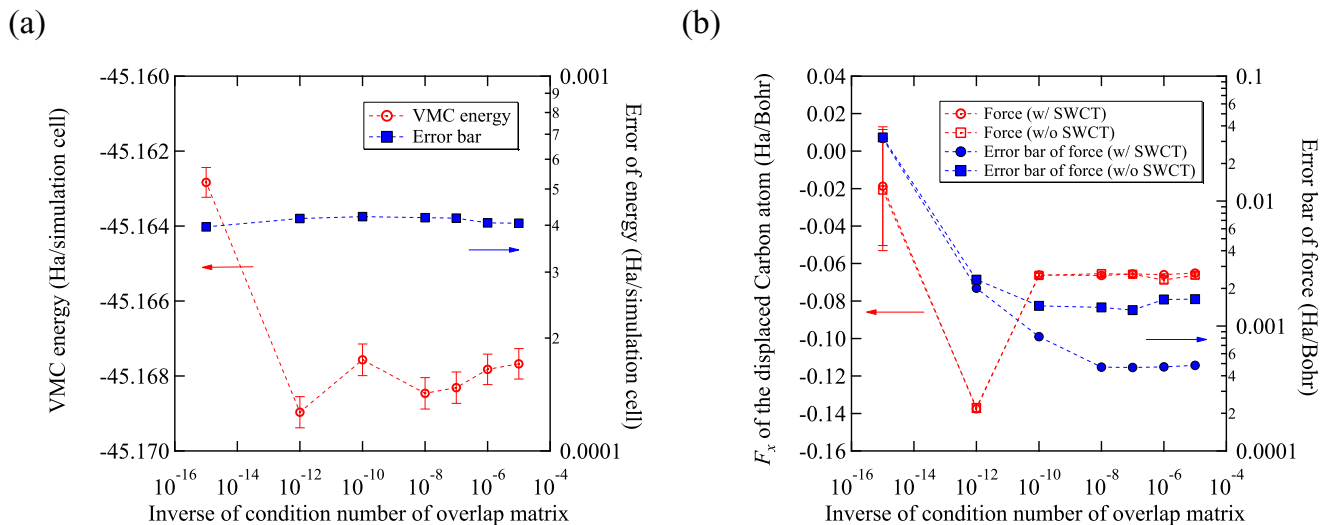


FIG. 3. (a) VMC energies and their statistical error bars and (b) VMC forces and their statistical error bars vs. inverse of the condition number of the overlap matrix. The numerical data is shown in Table S-II [34]. The forces were calculated with and without the space warp coordinate transformations (SWCT) [48]. The Hellmann-Feynman and Pulay contributions to the forces are shown in Fig. S-7 [34]. The same VTZ atomic basis set and the pseudopotential [33] as in the phonon and EOS calculations were adopted, while a  $1 \times 1 \times 1$  conventional cell (8 carbon atoms in a simulation cell) with  $k = \Gamma$  twist was employed. Only one carbon atom was displaced in  $x$  direction by 0.15 bohr, see Table S-III [34].

This is why the linear dependency of an atomic basis set deteriorates the overlap, and thus induces a large error bar in forces. Thus, a complete all-electron and pseudopotential basis set database suitable for QMC solid state calculations will be quite useful for the application of QMC forces to the calculations of phonons in realistic materials.

In summary, we report a VMC determination of the momentum-resolved phonon dispersion of diamond. Our approach combines the *ab initio* quantum Monte Carlo framework with the so-called frozen phonon technique. It gives results in very good agreement with experiments and provides renormalized harmonic optical frequencies consistent with the experimental findings. We estimated the purely harmonic contribution to the phonon spectrum, by evaluating  $q$ -dependent anharmonic corrections by means of a path integral molecular dynamics driven by PBE forces [52]. After including these corrections, the VMC phonon spectrum agrees very well with the experimental phonon dispersion. We found that alleviating the atomic basis-set redundancy of the trial wavefunction is key to reduce the statistical error of atomic forces, and thus to make VMC phonon calculations feasible over the full Brillouin zone. This achievement paves the way to new relevant applications, for instance in correlated materials and Van der Waals crystals (i.e., molecular crystals [63,64]) where sometimes phonons are poorly reproduced within the DFT framework.

K.N. is grateful for computational resources from the facilities of Research Center for Advanced Computing Infrastructure at Japan Advanced Institute of Science and Technology (JAIST). K.N. and S.S. are grateful for computational resources from PRACE Project No. 2019204934. T.M. and M.C. acknowledge that this work was supported by French state funds managed by the ANR within the Investissements d'Avenir programme under reference ANR-11-IDEX-0004-02, and more specifically within the framework of the Cluster of Excellence MATISSE led by Sorbonne University. M.C. is grateful to the French Grand équipement national de calcul intensif (GENCI) for the computational time provided through the Project No. 0906493. K.N., M.C., and S.S. are grateful for computational resources of the supercomputer Fugaku provided by RIKEN through the HPCI System Research Project (Project ID: hp200164). K.N. acknowledges support from the JSPS Overseas Research Fellowships and that from Grant-in-Aid for Scientific Research on Innovative Areas (No. 16H06439). R.M. is grateful for financial supports from MEXT-KAKENHI (19H04692 and 16KK0097), from Toyota Motor Corporation, from the Air Force Office of Scientific Research (AFOSR-AOARD/FA2386-17-1-4049; FA2386-19-1-4015), and from JSPS Bilateral Joint Projects (with India DST). S.S. also acknowledges financial support from PRIN 2017BZPKSZ. The authors appreciate helpful comments by A. Zen.

- [1] R. M. Martin, *Electronic Structure: Basic Theory and Practical Methods* (Cambridge University Press, Cambridge, 2004).
- [2] D. Sholl and J. A. Steckel, *Density Functional Theory: A Practical Introduction* (John Wiley & Sons, Hoboken, NJ, 2011).
- [3] S. Baroni, S. de Gironcoli, A. Dal Corso, and P. Giannozzi, *Rev. Mod. Phys.* **73**, 515 (2001).

- [4] A. Togo and I. Tanaka, *Scr. Mater.* **108**, 1 (2015).
- [5] N. Mounet and N. Marzari, *Phys. Rev. B* **71**, 205214 (2005).
- [6] M. Lazzeri, C. Attaccalite, L. Wirtz, and F. Mauri, *Phys. Rev. B* **78**, 081406(R) (2008).
- [7] M. Krisch, D. L. Farber, R. Xu, D. Antonangeli, C. M. Aracne, A. Beraud, T.-C. Chiang, J. Zaretsky, D. Y. Kim, E. I. Isaev,

- R. Ahuja, and B. Johansson, *Proc. Natl. Acad. Sci. U.S.A.* **108**, 9342 (2011).
- [8] D. Reznik, G. Sangiovanni, O. Gunnarsson, and T. Devereaux, *Nature* **455**, E6 (2008).
- [9] A. Floris, S. de Gironcoli, E. K. U. Gross, and M. Cococcioni, *Phys. Rev. B* **84**, 161102(R) (2011).
- [10] I. Leonov, A. I. Poteryaev, V. I. Anisimov, and D. Vollhardt, *Phys. Rev. B* **85**, 020401(R) (2012).
- [11] I. Leonov, V. I. Anisimov, and D. Vollhardt, *Phys. Rev. Lett.* **112**, 146401 (2014).
- [12] W. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, *Rev. Mod. Phys.* **73**, 33 (2001).
- [13] M. Marchi, S. Azadi, and S. Sorella, *Phys. Rev. Lett.* **107**, 086807 (2011).
- [14] M. Casula and S. Sorella, *Phys. Rev. B* **88**, 155125 (2013).
- [15] S. Sorella, K. Seki, O. O. Brovko, T. Shirakawa, S. Miyakoshi, S. Yunoki, and E. Tosatti, *Phys. Rev. Lett.* **121**, 066402 (2018).
- [16] A. Zen, D. Zhelyazov, and L. Guidoni, *J. Chem. Theory Comput.* **8**, 4204 (2012).
- [17] Y. Luo, A. Zen, and S. Sorella, *J. Chem. Phys.* **141**, 194112 (2014).
- [18] Y. Y. F. Liu, B. Andrews, and G. J. Conduit, *J. Chem. Phys.* **150**, 034104 (2019).
- [19] K. Nakano, R. Maezono, and S. Sorella, *J. Chem. Theory Comput.* **15**, 4044 (2019).
- [20] R. Maezono, A. Ma, M. D. Towler, and R. J. Needs, *Phys. Rev. Lett.* **98**, 025701 (2007).
- [21] S. Azadi, C. Cavazzoni, and S. Sorella, *Phys. Rev. B* **82**, 125112 (2010).
- [22] M. S. Liu, L. A. Bursill, S. Prawer, and R. Beserman, *Phys. Rev. B* **61**, 3391 (2000).
- [23] M. Schwoerer-Böhning, A. T. Macrander, and D. A. Arms, *Phys. Rev. Lett.* **80**, 5572 (1998).
- [24] J. Kulda, H. Kainzmaier, D. Strauch, B. Dorner, M. Lorenzen, and M. Krisch, *Phys. Rev. B* **66**, 241202(R) (2002).
- [25] D. Vanderbilt, S. G. Louie, and M. L. Cohen, *Phys. Rev. Lett.* **53**, 1477 (1984).
- [26] W. Parrish, *Acta Cryst.* **13**, 838 (1960).
- [27] J. L. Warren, R. G. Wenzel, and J. L. Yarnell, in *Inelastic Scattering of Neutrons*, Vol. 1 (International Atomic Energy Agency, Vienna, 1965), pp. 361–371; J. L. Warren, J. L. Yarnell, G. Dolling, and R. A. Cowley, *Phys. Rev.* **158**, 805 (1967).
- [28] A. Rohatgi, “Webplotdigitizer: Version 4.3,” <https://automeris.io/WebPlotDigitizer> (2020).
- [29] M. Casula, C. Filippi, and S. Sorella, *Phys. Rev. Lett.* **95**, 100201 (2005).
- [30] K. Nakano, R. Maezono, and S. Sorella, *Phys. Rev. B* **101**, 155106 (2020).
- [31] M. Casula and S. Sorella, *J. Chem. Phys.* **119**, 6500 (2003).
- [32] K. Nakano, C. Attacalite, M. Barborini, L. Capriotti, M. Casula, E. Coccia, M. Dagrada, C. Genovese, Y. Luo, G. Mazzola *et al.*, *J. Chem. Phys.* **152**, 204121 (2020).
- [33] M. Burkatzki, C. Filippi, and M. Dolg, *J. Chem. Phys.* **126**, 234105 (2007).
- [34] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.103.L121110> for details of the DFT and QMC calculations, and for the supplemental Figures S-1-S-7 and Tables S-I-S-III, which includes Refs. [47,65–67].
- [35] J. P. Perdew and A. Zunger, *Phys. Rev. B* **23**, 5048 (1981).
- [36] S. Sorella, M. Casula, and D. Rocca, *J. Chem. Phys.* **127**, 014105 (2007).
- [37] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, *Phys. Rev. Lett.* **98**, 110201 (2007).
- [38] R. G. Hennig, A. Wadehra, K. P. Driver, W. D. Parker, C. J. Umrigar, and J. W. Wilkins, *Phys. Rev. B* **82**, 014101 (2010).
- [39] S. Sorella, M. Casula, L. Spanu, and A. Dal Corso, *Phys. Rev. B* **83**, 075119 (2011).
- [40] F. Becca and S. Sorella, *Quantum Monte Carlo Approaches for Correlated Systems* (Cambridge University Press, Cambridge, 2017).
- [41] P. Giannozzi, S. Baroni, N. Bonini, M. Calandra, R. Car, C. Cavazzoni, D. Ceresoli, G. L. Chiarotti, M. Cococcioni, I. Dabo *et al.*, *J. Phys. Condens. Matter.* **21**, 395502 (2009).
- [42] J. P. Perdew, K. Burke, and M. Ernzerhof, *Phys. Rev. Lett.* **77**, 3865 (1996).
- [43] P. Pavone, K. Karch, O. Schütt, W. Windl, D. Strauch, P. Giannozzi, and S. Baroni, *Phys. Rev. B* **48**, 3156 (1993).
- [44] G. Kresse, J. Furthmüller, and J. Hafner, *Europhys. Lett.* **32**, 729 (1995).
- [45] F. Occelli, P. Loubeyre, and R. LeToullec, *Nat. Mater.* **2**, 151 (2003).
- [46] H. McSkimin and P. Andreatch Jr., *J. Appl. Phys.* **43**, 2944 (1972).
- [47] P. Vinet, J. R. Smith, J. Ferrante, and J. H. Rose, *Phys. Rev. B* **35**, 1945 (1987).
- [48] S. Sorella and L. Capriotti, *J. Chem. Phys.* **133**, 234111 (2010).
- [49] A  $2 \times 2 \times 2$  conventional supercell containing 64 atoms in the simulation cell is large enough for obtaining a phonon dispersion almost consistent with experimental ones, as shown in the Supplemental Material [34]. This was confirmed by comparing the phonon dispersion obtained by the linear-response method and that obtained by the linear-response method with very dense  $k$  and  $q$  grids at the DFT level.
- [50] N. J. Harris, *J. Phys. Chem.* **99**, 14689 (1995).
- [51] F. Mouhat, S. Sorella, R. Vuilleumier, A. M. Saitta, and M. Casula, *J. Chem. Theory Comput.* **13**, 2400 (2017).
- [52] T. Morresi, L. Paulatto, R. Vuilleumier, and M. Casula, [arXiv:2103.04094](https://arxiv.org/abs/2103.04094).
- [53] In the Raman mode, two nearest-neighbor carbon atoms are displaced in opposite directions by a distance  $u$  from their high-symmetry positions [20]. The distortion in the  $[1\ 0\ 0]$  direction was employed in this study [25]. We calculated VMC forces for two distorted structures ( $u = 0.03$  bohr and  $u = 0.05$  bohr) and obtained the Raman frequency by fitting the forces with a linear function. We also calculated VMC and LRDMC energies for the undistorted structure and two distorted structures ( $u = 0.03$  bohr and  $u = 0.05$  bohr), then obtained Raman frequencies by fitting the energies with a quadratic function.
- [54] S. Moroni, S. Saccani, and C. Filippi, *J. Chem. Theory Comput.* **10**, 4823 (2014).
- [55] Since smaller displacements were employed in this study (i.e., 0.03 and 0.05 bohr) than the previous one, the error bars in phonon frequencies at  $\Gamma$  point are a bit larger.
- [56] N. D. Drummond, M. D. Towler, and R. J. Needs, *Phys. Rev. B* **70**, 235119 (2004).
- [57] A. V. Krukau, O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria, *J. Chem. Phys.* **125**, 224106 (2006).
- [58] J. Sun, A. Ruzsinszky, and J. P. Perdew, *Phys. Rev. Lett.* **115**, 036402 (2015).

- [59] P. Hao, Y. Fang, J. Sun, G. I. Csonka, P. H. T. Philipsen, and J. P. Perdew, *Phys. Rev. B* **85**, 014111 (2012).
- [60] M. F. Peintinger, D. V. Oliveira, and T. Bredow, *J. Comput. Chem.* **34**, 451 (2013).
- [61]  $\kappa(\mathbf{S})$  = the maximum eigenvalue / the minimum eigenvalue of the overlap matrix ( $S_{\mu,v} = \langle \psi_\mu | \psi_v \rangle$ ).
- [62] C. Attaccalite and S. Sorella, *Phys. Rev. Lett.* **100**, 114501 (2008).
- [63] A. Subedi and L. Boeri, *Phys. Rev. B* **84**, 020508(R) (2011).
- [64] M. Casula, M. Calandra, and F. Mauri, *Phys. Rev. B* **86**, 075445 (2012).
- [65] H. J. Monkhorst and J. D. Pack, *Phys. Rev. B* **13**, 5188 (1976).
- [66] A. Dal Corso, *Comput. Mater. Sci.* **95**, 337 (2014).
- [67] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, Í. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, *Nat. Methods* **17**, 261 (2020).