I. INTRODUCTION

The low efficiency ($\eta$) of optical absorption in crystalline Si (c-Si) photovoltaics is one of the fundamental challenges impeding improvement in today’s most prolific1 solar-cell substrate. Since Carlson and Wronski2 reported the first feasible amorphous Si (a-Si) p-i-n solar cell in 1976, relatively inexpensive Si-based thin-film materials [including hydrogenated a-Si (a-Si:H) and derivative mixed-phase materials containing microcrystalline and/or nanocrystalline Si phases dispersed in a-Si:H] (Ref. 3) have been expected to revolutionize the photovoltaics industry. However, the highest confirmed $\eta$ for a single-junction a-Si:H cell is still less than 10% (Ref. 4), which is inadequate to provide serious competition to reigning c-Si solar-cell technologies. The redshifted absorption coefficient [$\alpha(\omega)$] and enlarged band gap ($E_g \approx 1.8$ eV) (Refs. 1 and 5) of a-Si:H are particularly advantageous relative to c-Si with respect to the spectral irradiance6 from the sun, but the low $\eta$ ceiling of a-Si:H materials will remain until the light-induced degradation of the reversible Staebler-Wronski effect7,8 is circumvented. Significant effort is required to improve c-Si absorption through techniques including transition metal doping to create intermediate bands9 and growth of Si nanowire arrays that function as broadband absorption layers.10

It is well established that variations in phase,11 defect-content,12,13 and morphology14–18 (porous Si morphologies vary greatly) of Si produce appreciable changes in the absorption spectra. Knief and von Niessen12 succinctly classified the absorption spectra of a-Si materials into three categories: (1) high-energy absorption (like c-Si) in which transitions occur between extended electronic states (bands); (2) Urbach absorption at lower photon energies in which optical transitions occur between localized states in one band (valence or conduction) and extended states in the opposite band; and (3) tail or defect absorption at the lowest photon energies in which optical transitions are assisted by localized states near the Fermi energy. Localized defects include native defect clusters, which are often generated during ion implantation and crystal growth18 processes. The recent theoretical work of Pan et al.19 proposed an atomistic topological relationship connecting short (long) Si-Si bonds with valence- (conduction- ) band character in the electronic structure of a-Si, which suggests that the strain distribution associated with the bond topology of an arbitrary structure can modify the joint density of states, and consequently optical absorption, in Si.

The ubiquitous presence of vacancies in all crystalline materials and their further proliferation through a variety of semiconductor processes has made them extensively studied using both experimental and computational techniques. Early theoretical studies20–23 focused on the ring hexavacancy ($V_6$) because of its simple structure and exceptionally stable fourfold-coordinated (FC) configuration. Makrov and Lewis24 later used the results of density-functional theory for small vacancy clusters ($V_n$, $n=3 \leq n \leq 5$). Lee and Hwang25,26 recently generalized the significance of FC vacancy clusters by detailing their thermodynamic preference through systematic application of an integrated atomistic modeling procedure25–28 for small $n$ ($V_n$, $n=48$). Despite these computational achievements, small FC clusters have remained elusive to physical verification because of their expected transparency to common experimental techniques such as deep-level transient spectroscopy (DLTS) and photoluminescence.21 Similarly, Davies et al.29 report classes of vacancy-cluster defects which are invisible to infrared spectroscopy and DLTS, yet detectable through positron annihilation and broadband luminescence; however, the structures of their reported cluster defects are not identified. A disconnect remains between theory and experiment because spectroscopic techniques, in general, cannot provide the structural configurations for small $V_n$, and microscopy techniques, likewise, cannot provide the subnanometer resolution necessary to resolve small $V_n$.

In this paper, we apply a DFT simulation approach to explore the contribution of structural disorder to optical absorption in Si using charge-neutral, ground-state Si vacancy...
clusters \(V_n, n = 6, n = 12 \) and 32). In particular, we present calculated optical spectra for the dielectric function \(\varepsilon(\omega)\) and absorption coefficient \(\alpha(\omega)\) for previously identified FC vacancy clusters and discuss the results in context with the archetypal spectra of the a-Si and c-Si phases. In addition, we extend our computational approach to assess the effect of defect concentration, using incremental dilutions of monovacancy optical response, and the effect of density, using amorphous samples, on optical absorption in Si. We intend our simulation results to provide improved understanding of bulk Si optical absorption by exploring the influence of subnanometer-scale structural disorder in c-Si via native defect cluster incorporation.

II. COMPUTATIONAL FRAMEWORK

All atomic configurations and optical spectra reported herein were computed using a plane-wave pseudopotential method within the generalized gradient approximation of Perdew and Wang (GGA-PW91) (Ref. 30) to DFT,\(^{31}\) as implemented in the Vienna ab initio simulation package (VASP)\(^{32}\). We employed Vanderbilt-type ultrasoft pseudopotentials\(^{33}\) (US-PPs) to represent the interaction between ion cores and valence electrons for initial structural optimization and during volume relaxation of amorphous reference samples, while projector-augmented wave (PAW) pseudopotentials\(^{34}\) were employed throughout the entire optical calculation sequence. In principle, the PAW methodology comprises an all-electron frozen-core approach that represents exact valence wave functions; therefore, we employed PAW pseudopotentials to help ensure enhanced representation of electronic transitions.

During structural optimization with US-PPs, we employed a plane-wave basis set with a kinetic-energy cutoff \((E_{\text{cut}})\) of 160 eV, sampled the Brillouin zone (BZ) at the \(\Gamma\) point, and allowed all atoms to relax until residual forces between atoms had converged within \(5 \times 10^{-2} \) eV/Å. For all c-Si supercells, we used a fixed, DFT-optimized lattice constant of 5.457 Å along (100) or 3.859 Å along (110). For all a-Si structures, initial structural optimization was followed by DFT-based volume relaxation on each cubic supercell to determine the equilibrium volume and density of each structure. In our previous work, we justifies these alternative modeling philosophies through discussion of the trade-offs encountered between the fixed-cell and volume-relaxed approaches to modeling both embedded clusters and amorphous phases of Si. Note that the application of US-PPs throughout this work is primarily a means of initial structural relaxation at reduced computational cost; in some cases, where well-relaxed structures were already available, structural optimization with US-PPs was deemed unnecessary. This situation typically occurred in the generation of derivative structures, such as the process of embedding a previously relaxed cluster in a supercell of larger size.

A. Optical calculations

For optical calculations, we utilized a four-step procedure to compute the complex frequency-dependent dielectric function with VASP following the general theoretical framework described by Adolph et al.\(^{37}\) Since this calculation procedure is based on the independent-particle approximation, quasiparticle self-energy corrections, local-field effects, and excitonic contributions are neglected. Table I is provided to recapitulate the DFT conditions for the optical calculation sequence for the essential supercell sizes used throughout this work.

In the first optical calculation step, the configuration of the relaxed structure (US-PPs) was optimized with PAW pseudopotentials using the same force convergence criterion with cut-off energies and BZ sampling as indicated in Table I. Second, the charge-density distribution was obtained using increased cutoff energies and BZ sampling in accordance with Table I. Third, the frequency-dependent dielectric matrix was determined from the optimized charge-density distribution using the tetrahedron method\(^{32}\) and an adequate number of conduction bands that we carefully scaled with supercell size. Since the optical spectra is sensitive to the number of conduction bands, especially at high energies, we performed preliminary calculations to ensure saturation of both the optical spectra and low-frequency dielectric constant \(\varepsilon_{\text{r}}\) as a function of the number of bands simulated. Performing DFT calculations on large 980-atom supercells also presented an unusual challenge. Bound by the constraint of the tetrahedron method (minimum of four \(k\)-points), yet also seeking to minimize \(k\)-point quantity for the conservation of computational resources, we used a \(2 \times 2 \times 2\) \(\Gamma\)-centered mesh in the full BZ (\(k\)-point reduction via symmetry suppressed), instead of the VASP (Ref. 32) implementation of the standard Monkhorst-Pack scheme used to generate the other meshes. Inclusion of the \(\Gamma\) point on even meshes often results in more \(k\)-points, while identical meshes are observed for both schemes for odd meshes. This detail was necessary to ensure a mesh with eight \(k\)-points for all 980-atom supercells so the tetrahedron method could be applied in all cases, including c-Si. Finally, the OPTICS code of Furthmüller\(^{38}\) was used to obtain both the imaginary \(\varepsilon_2\) and real \(\varepsilon_1\) parts of \(\varepsilon(\omega)\). The imaginary part of \(\varepsilon(\omega)\) (in atomic units) is evaluated by summation over the frequency-dependent dielectric matrix elements using

<table>
<thead>
<tr>
<th>Base supercell</th>
<th>(k)-point mesh(^{a})</th>
<th>(E_{\text{cut}}) (^{b}) (eV)</th>
<th>Conduction bands</th>
</tr>
</thead>
<tbody>
<tr>
<td>72</td>
<td>(\alpha(\gamma))</td>
<td>(\delta(\gamma))</td>
<td>200</td>
</tr>
<tr>
<td>480</td>
<td>(\Gamma(\beta))</td>
<td>(\delta(\gamma))</td>
<td>750</td>
</tr>
<tr>
<td>980</td>
<td>(\Gamma(\alpha))</td>
<td>(\delta(\gamma))</td>
<td>1300</td>
</tr>
</tbody>
</table>

\(\Gamma = \Gamma\) point, \(\alpha = (2 \times 2 \times 2), \beta = (3 \times 3 \times 3), \) and \(\chi = (9 \times 9 \times 9)\).  
\(^{a}\)\(\delta = 250\) eV and \(\gamma = 300\) eV.
ROLE OF STRUCTURAL DISORDER IN OPTICAL…

\[ \text{Im}[\varepsilon(\omega)] = \varepsilon_2(\omega) = \frac{4 \pi^2}{\Omega \omega^2} \sum_{i \in \text{VB}, j \in \text{CB}} \sum_k w_k |p_{ij}|^2 \times \delta (\varepsilon_{kj} - \varepsilon_{ki} - \omega), \]

where \( \Omega \) is the unit-cell volume, VB represents the valence band, CB represents the conduction band, \( w_k \) is a weighting associated with a \( k \)-point \( k \), and \( \varepsilon_{kn} \) represent energy levels in their respective bands.\(^{39}\) Transition matrix elements in Eq. (1) are represented by \( p_{ij} = (\mathbf{k}_i | \mathbf{\hat{p}} | \mathbf{k}_j) \), where \( \{ \mathbf{k}_i \} \) are Bloch-state wave functions with momentum \( \mathbf{k} \). We used the longitudinal form\(^{32,39,40}\) of the dielectric matrix in our VASP calculations throughout this work. When the imaginary part \( \varepsilon(\omega) \) is known for all photon energies, application of the Kramers-Kronig transformation (in atomic units),

\[ \text{Re}[\varepsilon(\omega)] = \varepsilon_1(\omega) = 1 + \frac{2}{\pi} \mathbf{P} \int_0^\infty \frac{\omega' \varepsilon_2(\omega')}{\omega^2 - \omega'^2} d\omega', \]

where \( \mathbf{P} \) represents the principal integral value, allows subsequent evaluation of the real component of \( \varepsilon(\omega) \). We found that sampling optical transitions up to 20 eV was adequate to obtain the key features in the dielectric function spectra. With \( \varepsilon_1 \) and \( \varepsilon_2 \) determined at each \( \omega \), the absorption coefficient is computed as an algebraic expression,

\[ \alpha(\omega) = \frac{\rho_\mathbf{e}}{\hbar} \left[ \sqrt{\varepsilon_1^2(\omega) + \varepsilon_2^2(\omega)} - \varepsilon_1(\omega) \right]^{1/2}, \]

where \( \hbar \) represents photon energy, \( \hbar \) represents the reduced Planck’s constant, and energy and frequency are essentially interchangeable via \( E = \hbar \omega \).

The DFT-GGA approach employed in this paper characteristically underestimates \( E_g \), but the resulting qualitative trends we show are considered reliable. Numerous correction methods have been reported to address the electronic band-structure deficiencies of DFT, but each method necessitates additional trade-offs and the most appropriate correction scheme is dictated by the material system. Hybrid functionals use a mixture of Hartree-Fock (HF) exchange correlation with local or semilocal density functionals (GGA) to improve semiconductor \( E_g \) representations, but HF evaluation in periodic systems can be computationally expensive and the ratio of HF/DFT representation is an additional variable.\(^{41}\) The quasiparticle GW method demands significantly more computational expense because the solution of single-particle Kohn-Sham equations in DFT is supplanted by calculation of both the dielectric function and single-particle Green’s function via rigorous solution of the Dyson equation.\(^{42}\) Application of either correction method to our current work would be cost prohibitive because of the large supercells required to represent large defect clusters.

B. Structural models

Lee and Hwang\(^{25,26}\) recently reported configurations of FC vacancy clusters (\( V_n \), \( n \leq 48 \)) using an integrated atomistic modeling procedure\(^{25-28}\) combining Metropolis Monte Carlo (MMC), tight-binding molecular dynamics, and DFT calculations. We use these same structures as the foundation for the current investigation presented here. Figure 1 presents the neutral, ground-state FC configurations of \( V_1 \) through \( V_6 \) along with the monovacancy (\( D_{2h} \) symmetry) and divacancy (\( C_{2v} \) symmetry) configurations. We have chosen to illustrate the intricate structures of FC vacancy clusters through identification of the highly strained atoms neighboring the vacancies, rather than explicitly depicting the original vacancy sites with visual identifiers as commonly seen in the literature\(^{24}\) for PHR vacancy clusters. Unless noted otherwise, \( V_1 \) through \( V_6 \) calculations were conducted in 480-atom basis supercells. Figure 2 provides illustrations of the larger structures investigated that required the use of 980-atom supercells. Unlike Fig. 1, we have isolated the previously reported structural cores for \( V_{12} \) [Fig. 2(a)] (Ref. 25) and \( V_{32} \) [Fig. 2(b)] (Ref. 26) from the Si lattice to enhance visual clarity for the reader. Figure 2(c) shows the 980-atom a-Si supercell and its radial distribution function (RDF), which we will use extensively as a reference case throughout this work.

We constructed a-Si samples free of coordination defects using a derivative approach of the method described by Yu et al.\(^{44}\) that employs continuous random network (CRN) model-based MMC sampling with Keating-type potentials. This CRN-MMC method was used to generate at least five samples of each unique size/density of a-Si supercell studied.
including five different densities (vide infra Table IV) of a-Si in 480-atom basis supercells as well as the 980-atom a-Si supercell. The variable-density a-Si samples were constructed using the experimental density of 2.28 g/cm$^3$ \textit{Ref}.$^{45}$ as the basis for the aSi$_{480}$ reference structure containing 480 atoms. Lower density structures aSi$_{468}$, aSi$_{456}$, aSi$_{444}$, and aSi$_{432}$ were subsequently generated by seeding the same volume supercell with incrementally fewer Si atoms. We used cost-effective US-PPs to select the lowest energy amorphous samples at each density, then applied sample-specific volume relaxation on the remaining candidate(s) to identify the lowest energy representative structure for each cell size and/or density.

III. RESULTS AND DISCUSSION

Our investigation originated as a systematic evaluation of optical spectra based on the FC vacancy clusters identified by Lee and Hwang.$^{25,26}$ The effects of Si divacancies (dangling bonds) on optical absorption were previously measured with spectrophotometry by Stein \textit{et al}.$^{13}$ and Knief and von Niessen$^{12}$ modeled the effect of both vacancy concentration and coordination defects on optical absorption. Based on these previous results, we were motivated to determine if optical signatures would be readily observable for FC vacancy clusters despite their expected transparency to common experimental characterization techniques.$^{20,21,29}$

A. Structural disorder effects in c-Si

1. Small cluster absorption \( \left( V_n, \; n \leq 6 \right) \)

Figure 3 provides absorption coefficient spectra as a function of both photon energy and wavelength for all $V_n$ shown in Fig. 1 along with c-Si as a reference for comparison. While an optical signature was anticipated based on the results of Stein \textit{et al}.$^{13}$ for clusters with dangling bonds ($V_1$ and $V_2$), the substantial magnitude and bandwidth of optical absorption enhancement over c-Si was unexpected. The unsaturated bonds in $V_1$ and $V_2$ correspond to typical defect
absorption\textsuperscript{12} in which sharp absorption peaks are observed at low energies. Our results reveal that similar increases in absorption are also observed for all FC clusters ($V_n$, $3 \leq n \leq 6$). For these clusters, the onset of incipient absorption occurs at relatively higher photon energies ($E > 0.5$ eV) and the general shape of the absorption envelope tends to be much smoother than that seen for the $V_1$ and $V_2$ cases.

2. Large cluster absorption ($V_{12}, V_{32}$)

To extend the significance of our observations on small $V_n$ $(n \leq 6)$, we proceeded to examine the optical spectra on the large, FC configurations of $V_{12}$ and $V_{32}$. Since the expanded $\alpha(\omega)$ envelope we observed for small $V_n$ is somewhat reminiscent of the well-known a-Si $\alpha(\omega)$ enhancement,\textsuperscript{2} we also include a second reference for a-Si throughout most of our remaining results for comparison. Figure 4 provides the $\alpha(\omega)$ spectra for $V_{12}$ and $V_{32}$ along with c-Si (Ref. 46) and a-Si references, all modeled using 980-atom basis supercells, while we introduce the corresponding imaginary and real components of $\varepsilon(\omega)$ for the same structures in Fig. 5. From inspection of Fig. 4, it is evident that the enhanced absorption seen for the small $V_n$ clusters in Fig. 3 is indeed a general attribute of a vacancy cluster embedded in c-Si. In Fig. 4, we also observe that the $\alpha(\omega)$ spectra are similar, yet unique, for $V_{12}$ and $V_{32}$ and both are strikingly similar to the $\alpha(\omega)$ spectra of a-Si. $V_{12}$, $V_{32}$, and a-Si all show incipient absorption around 0.5 eV, where the sharp peaks and discontinuities observed exclusively for a-Si are likely attributable to highly strained bonds since our structural analysis\textsuperscript{47} confirms complete fourfold coordination throughout the final configuration.

The additional details provided by $\varepsilon_2(\omega)$ in Fig. 5(a) help exemplify incremental transition across the four sample structures. Using the nomenclature of Yu and Cardona\textsuperscript{5} for labeling optical transition energies in tetrahedrally-bonded semiconductors, the sharpness of the fundamental absorption edge, or $E_0$ transition energy, washes out and concomitantly redshifts as we proceed through this structural sequence: c-Si, $V_{12}$, $V_{32}$, and a-Si. Likewise, the sharpness of the shoulder just above $E_0$, known as the $E_1$ transition, exhibits an incremental softening across the same aforementioned structural sequence. The critical point at $E_1$ arises from optical transitions occurring along (111) directions in the BZ (Refs. 5 and 48) and has been shown as a resilient metric\textsuperscript{48} in the quantification of disorder during amorphization well beyond the disintegration of other features in the optical spectra. Furthermore, the maximum $\varepsilon_2(\omega)$ peak intensity, or $E_3$ transition, decays in magnitude and slightly redshifts across the same structural sequence. Correspondingly, inspection of $\varepsilon_1(\omega)$ in Fig. 5(b) shows both a broadening in anomalous dispersion and an increase in low-frequency dielectric constant ($\varepsilon_\infty = \lim_{\omega \to 0} \text{Re}[\varepsilon(\omega)]$) across the same structural sequence.

3. Absorption in constant-density cluster distributions

To isolate the influence of topological disorder observed in absorption spectra from the density variation that occurs with changes in vacancy concentration, we constructed three different constant-density, 980-atom basis supercells each containing a total of twelve vacancies using different combinations of small, FC $V_n$ clusters in Fig. 1. Each of these derivative supercells, $V_n \times 2$, $V_n \times 3$, and $V_n \times 4$, contains 968 atoms in a c-Si matrix, where the nomenclature $V_n \times q$ indicates the quantity ($q$) of cluster size ($n$) embedded in each supercell. During the construction of each supercell, we attempted to homogenize the cluster spatial distribution in
structures using cumulative distribution functions (CDFs) of bond lengths and angles. Since the strain fields of small FC clusters can be spatially extensive, CDFs were chosen to elucidate subtle variations among the constant-density vacancy structure bond distributions that we expected to see. Figure 7 provides structural distribution signatures for each \( V_n \) cluster that highlight the unique character of each cluster in contrast to the almost Gaussian distributions of a-Si. It is readily apparent that the bond-length distributions for all \( V_n \) are distinctly tensile, while the bond-angle distributions do not exhibit significant deviations in the mean, relative to c-Si. The largest cluster, \( V_{12} \), shows bond-length and bond-angle distributions that are beginning to exhibit some Gaussian character, which is a trend we expect to continue with increasing \( n \).

Figure 8 visually depicts the intermediate nature of the structural bond topologies of the constant-density vacancy structures relative to the extreme cases offered by the c-Si and a-Si distributions. For bond lengths [Fig. 8(a)], we discern a general trend above the 50th percentiles (with some crossover between \( V_4 \times 3 \) and \( V_4 \times 4 \)) of increased spread in bond-length distributions correlating with the same structural sequence (\( V_{12} \), \( V_6 \times 2 \), \( V_4 \times 3 \), and \( V_4 \times 4 \)) of optical absorption enhancement. Interestingly, the CDFs for \( V_4 \times 3 \) and \( V_4 \times 4 \) indicate considerably more compressive bond-length character than the CDFs for \( V_6 \times 2 \) and \( V_{12} \). Reinforcing the distributions of Fig. 7, the plenary bond-angle data provided by Fig. 8(b) shows a more symmetric nature for all structures relative to the bond-length distributions. In addition, increased spread (flattening in CDF) in bond-angle distributions below the 25th and above the 70th percentiles for the constant-density vacancy structures is again observed in this sequence: \( V_{12} \), \( V_6 \times 2 \), \( V_4 \times 3 \), and \( V_4 \times 4 \). Thus far, we consistently observe a pattern for enhanced optical absorption corresponding to both increased spatial dispersion (\( V_n \times q \), \( n \) decreases, \( q \) increases) of a fixed vacancy quantity (\( n \times q = 12 \)) in concert with increased dispersion in bond-topology distributions.

Table II provides quantification of important statistics that characterize the bond-topology distributions of the 980-atom supercells as well as sample densities and internal strain energies. We computed strain energies by calculating formation energies on a per atom basis of charge-neutral, Si vacancy clusters as

\[
E_f(n) = E_{tot}(n) - \frac{N-n}{N}E_{bulk},
\]

where \( E_{tot}(n) \) is the total energy of each \( N-n \) atom supercell, \( n \) is the number of vacancies, \( N \) is the basis number of atoms in the c-Si supercell, and \( E_{bulk} \) is the total energy of the \( N \)-atom c-Si supercell. Equation (5) is similarly applicable to variable-density amorphous structures, where \( E_f(n) \) instead represents the total energy of the \( N-n \) atom a-Si sample and \( n \) represents the number of removed Si atoms, if any, relative to the experimental reference density. Based on the work of Lee and Hwang,\(^{25}\) it is not surprising that a trend of increased strain energy is observed as the twelve constituent vacancies are dispersed across an increasing number of clusters since the smaller vacancy clusters have higher \( E_f(n) \) on a
per vacancy basis. Statistical trends in bond lengths for the constant-density $V_n \times q$ structures are not readily apparent, but the standard deviations ($\sigma$) and interquartile ranges (IQRs) for the bond angles quantify the observations described in the CDFs of Fig. 8. Perhaps most importantly, the maximum bond-angle ($\theta_{\text{max}}$) data for the constant-density $V_n \times q$ structures reveals that both $V_4 \times 3$ ($\theta_{\text{max}}=153.86^\circ$) and $V_3 \times 4$ ($\theta_{\text{max}}=159.09^\circ$) contain larger bond angles than a-Si ($\theta_{\text{max}}=150.58^\circ$), which offers a plausible structural explanation for the enhanced redshift in incipient absorption observed for $V_4 \times 3$ and $V_3 \times 4$ relative to a-Si in the $\Xi(\omega)$ spectra of Fig. 6.

5. Bond-length contribution

Expecting to identify a relationship between bond length and optical absorption, we continued our investigation by exploiting the unique structural manipulation of hydrostatic strain which provides a method to decouple the contributions of bond length and bond angle. Application of hydrostatic strain permits uniform variation in bond length throughout the structure while the bond angles remain constant. The conclusions of Idrobo et al. show the Si-Si bond exhibits a universal optical response in the dominant spectral peak, regardless of bonding configuration. We extend their results in Table III with our $\varepsilon_2(\omega)$ calculations obtained via application of hydrostatic strain to 72-atom c-Si supercells. We observe that increasing compression of the Si-Si bond corresponds to an increasing redshift in the $E_2$ transition energy, while the $E_2$ transition is simultaneously blueshifting and increasing in intensity. Since Si-Si bond elongation conversely modulates $\varepsilon_2(\omega)$ in ways that contribute to reduced absorption, it is possible that the enhanced absorption associated with the
largely tensile nature of the vacancy clusters is either driven by the minority compressive bonds or through the more elusive effect of bond-angle variation. It is unfortunate that an analogous method of equivalent simplicity to strictly control bond-angle distribution is not readily available for a tetrahedrally-bonded network. We have compared the intermediate nature of topological disorder in constant-density \( V_{\text{12}} / H_{11003} \) structures to the broad, nearly Gaussian distributions that characterize the bond topology of a-Si and the sharp distributions that define c-Si. Conceptually, we suggest that the optical absorption spectra in a Si material of arbitrary morphology can be interpreted as the weighted superposition, or perhaps convolution, of all individual spectra representing the constituent bond lengths and angles in the system.

### B. Defect concentration effect in c-Si

In light of numerous recent endeavors\textsuperscript{20,21,29} that have discussed expected transparency of FC Si vacancy clusters, we

![Graph](image)

**FIG. 8.** (Color online) Normalized cumulative distribution functions representing all (a) bond lengths and (b) bond angles in the constant-density, 968-atom supercells \( V_{12}, V_{6} \times 2, V_{4} \times 3, \) and \( V_{3} \times 4 \) along with their 980-atom c-Si and a-Si references. For both (a) and (b), the insets illustrate the fine details differentiating the constant-density structural distributions, while their respective parent plots provide perspective.

Table II. Summary of selected bond topology statistics compiled for various structures represented in 980-atom basis supercells. Strain energies are computed as formation energies referenced to c-Si and mass densities are given as \( \rho_{\text{m}} \). Bond-angle and bond-length distribution statistics are computed comprehensively over the entire configuration represented in each supercell. Standard deviations and interquartile ranges are represented by \( \sigma \) and IQR, respectively.

<table>
<thead>
<tr>
<th>Structure</th>
<th>N-n</th>
<th>Strain energy (eV/atom)</th>
<th>( \rho_{\text{m}} ) (g/cm(^3))</th>
<th>Bond angle (deg)</th>
<th>Bond length (Å)</th>
</tr>
</thead>
<tbody>
<tr>
<td>c-Si</td>
<td>980</td>
<td>2.296</td>
<td>199.47, 0.06, Maximum, 109.69, 0.06</td>
<td>2.363, 0.001</td>
<td>2.373, 0.048</td>
</tr>
<tr>
<td>a-Si</td>
<td>980</td>
<td>0.203</td>
<td>2.252</td>
<td>109.18, 10.31, 150.58, 14.15</td>
<td>2.373, 0.094</td>
</tr>
<tr>
<td>( V_{12} )</td>
<td>968</td>
<td>0.014</td>
<td>2.268</td>
<td>109.45, 2.48, 139.06, 0.56</td>
<td>2.373, 0.029</td>
</tr>
<tr>
<td>( V_{6} \times 2 )</td>
<td>968</td>
<td>0.020</td>
<td>2.268</td>
<td>109.45, 2.92, 145.59, 0.74</td>
<td>2.373, 0.044</td>
</tr>
<tr>
<td>( V_{4} \times 3 )</td>
<td>968</td>
<td>0.024</td>
<td>2.268</td>
<td>109.45, 3.47, 153.86, 1.21</td>
<td>2.373, 0.036</td>
</tr>
<tr>
<td>( V_{3} \times 4 )</td>
<td>968</td>
<td>0.029</td>
<td>2.268</td>
<td>109.44, 3.98, 159.09, 1.29</td>
<td>2.375, 0.039</td>
</tr>
<tr>
<td>( V_{32} )</td>
<td>948</td>
<td>0.029</td>
<td>2.221</td>
<td>109.43, 3.57, 137.99, 1.85</td>
<td>2.386, 0.037</td>
</tr>
</tbody>
</table>

Table III. Summary of Si-Si bond length influence on selected critical features of \( E_{2} \) using hydrostatically-strained 72-atom c-Si supercells. The bond lengths shown represent averages at each strain condition, where the \( \sigma \) for each bond-length distribution is \( -3 \times 10^{-4} \) Å and the average bond angle for each distribution is 109.47° ± 0.01°. \( E_{2} \) peak intensities are normalized to the strain-free case.

<table>
<thead>
<tr>
<th>Hydrostatic strain (%)</th>
<th>Si-Si bond length (Å)</th>
<th>( E_{0} ) transition (eV)</th>
<th>( E_{2} ) transition (eV)</th>
<th>( E_{2} ) intensity (normalized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>-5</td>
<td>2.245</td>
<td>0.42</td>
<td>4.04</td>
<td>1.102</td>
</tr>
<tr>
<td>-4</td>
<td>2.268</td>
<td>0.49</td>
<td>4.08</td>
<td>1.080</td>
</tr>
<tr>
<td>-3</td>
<td>2.292</td>
<td>0.56</td>
<td>3.92</td>
<td>1.085</td>
</tr>
<tr>
<td>-2</td>
<td>2.316</td>
<td>0.62</td>
<td>3.87</td>
<td>1.018</td>
</tr>
<tr>
<td>-1</td>
<td>2.339</td>
<td>0.68</td>
<td>3.80</td>
<td>1.004</td>
</tr>
<tr>
<td>0</td>
<td>2.363</td>
<td>0.74</td>
<td>3.75</td>
<td>1.000</td>
</tr>
<tr>
<td>1</td>
<td>2.387</td>
<td>0.79</td>
<td>3.72</td>
<td>0.992</td>
</tr>
<tr>
<td>2</td>
<td>2.410</td>
<td>0.83</td>
<td>3.67</td>
<td>0.987</td>
</tr>
<tr>
<td>3</td>
<td>2.434</td>
<td>0.87</td>
<td>3.63</td>
<td>0.938</td>
</tr>
<tr>
<td>4</td>
<td>2.457</td>
<td>0.90</td>
<td>3.58</td>
<td>0.907</td>
</tr>
<tr>
<td>5</td>
<td>2.481</td>
<td>0.93</td>
<td>3.52</td>
<td>0.892</td>
</tr>
</tbody>
</table>

In light of numerous recent endeavors\textsuperscript{20,21,29} that have discussed expected transparency of FC Si vacancy clusters, we
were compelled to offer a plausible explanation for the distinguishing optical signatures of FC vacancy clusters observed from DFT-GGA simulation results. Since the computational demands of our first-principles approach essentially restricts a feasible supercell size to approximately \( N \leq 1 \times 10^3 \), it is usually prudent to estimate how DFT calculation results at the subnanometer scale will translate with system size as guidance for subsequent studies.

In Fig. 9, we provide both \( \alpha(\omega) \) spectra and its cumulative counterpart for a monovacency embedded in a variety of supercell sizes (192, 256, 480, 576, and 980 atoms) to gauge the effect of concentration on our DFT results. In this case, we chose \( V_1 \), rather than a larger FC structure, to maximize the range of feasible basis supercell sizes. To accommodate the additional basis supercells, BZ sampling was increased to \( 4 \times 4 \times 4 \) meshes for the 192- and 256-atom basis supercells, while the 576-atom cell retained the \( 3 \times 3 \times 3 \) \( k \)-point mesh. Except for the 980-atom supercell, all cells used 750 conductor bands. No other changes were made to the previously detailed optical approach.

The utility of expressing \( \alpha(\omega) \) in its cumulative form is immediately evident in comparison of Figs. 9(a) and 9(b). Inspection of \( \alpha(\omega) \) spectra provides an impression that the absorption signal of \( V_1 \) is diluted with increasing supercell size, while \( \Xi(\omega) \) is visually more effective at establishing the same trend. Since the dilution of the \( V_1 \) signal is sufficiently established over less than an order of magnitude change in basis supercell size (192 \( \leq N \leq 980 \)), it is reasonable to conclude that concentration has a significant effect on the practical measurement of optical absorption. Consequently, we can reasonably speculate from our simulated \( \Xi(\omega) \) that a critical threshold \( V_n \) concentration is required to optically validate the presence of \( V_n \) clusters with experimental techniques and insufficient \( V_n \) concentration is a primary reason for the reported transparency of FC vacancy clusters.

### C. Density effect in a-Si

In order to assess the contribution of atomic/mass density variation in optical absorption, we applied our optical computational procedure to a set of five variable-density a-Si structures. Not surprisingly, application of volume relaxation to each structure reduces the total range and incremental change in density covered by the resultant five-sample set, but our previous simulation work has shown this method to be most effective in reproduction of expected material properties.

One limitation of our approach is that the density effect we report is specific to one type of morphology. Using the method described in Sec. II B for supercell sizes where \( N \leq 480 \), the reduction in density generated is approximately uniform if the structural topology is extrapolated to larger size scales; however, we acknowledge that experimentally produced, low-density Si, such as a-Si (Refs. 50 and 51) or porous Si (p-Si), typically exhibits more heterogeneous morphologies in which lower density is often achieved from the contribution of large, open-volume voids.

In Table IV, we present the effect of density variation on \( \varepsilon_2(\omega) \), which is rather minor for the density range investigated. We provide the critical features of \( \varepsilon_2(\omega) \), rather than the spectra, because the effect of density is sufficiently small such that all spectra appear nearly identical. A consistent reduction in \( E_2 \) peak intensity is observed as density decreases, as well as a less-distinguished trend for \( E_0 \) to redshift. We use the utility of the logarithmic scale for \( \alpha(\omega) \) to resolve the small changes in optical response among the five variable-density a-Si structures in Fig. 10. Relative to the structural effect on optical absorption discussed in Sec. III A and observed in Figs. 3, 4, and 6, the effect of density on optical absorption is secondary for the amorphous morphologies investigated. The main feature distinguishing the spectra in Fig. 10 is the transition energy for incipient absorption. Even this feature is actually structural in origin since the discontinuities observed in incipient absorption (similar to Fig. 4) are typical signatures of highly localized states (highly strained or dangling bonds) contributing to tail absorption.
The insignificant effect of density in Fig. 10 is counterintuitive to classical optics principles, where application of the effective medium approximation (EMA) to the large difference in dielectric constants between vacuum ($\varepsilon_0 = 1$) and either a-Si ($\varepsilon_0 \approx 16$) (Ref. 54) or c-Si ($\varepsilon_0 = 11.9$) (Ref. 5) suggests an appreciable effect on optical absorption. However, as discussed by Banerjee et al., 85 voids (or vacancy clusters) that are significantly smaller than the wavelength of light, as is the case for all structures presented in this work, invalidate EMA and require computation of an effective $\alpha(\omega)$ comprehending both diffraction of light by nanocrystallites between the voids as well as Rayleigh scattering by the voids. Similar to our work, this mathematical optical theory described by Banerjee et al. 55 predicts significant enhancement of the effective $\alpha(\omega)$ in thin c-Si layers by optimization of both nanovoid dimensions and distributions.

Despite the insignificant effect of density observed on the a-Si morphologies we studied, further investigation is warranted on characteristic multiphase (voids interspersed in a-Si) a-Si morphologies. We suspect the density effect on absorption to be significant, but we believe its importance is probably only relevant on larger length scales. Using simulation methods capable of handling tens of thousands of atoms would allow representation of large, open-volume voids. For void dimensions comparable to the wavelength of light, we suspect a more significant effect of density would be evident using EMA.

The results we have presented suggest that incorporation of modest amounts of localized structural disorder in c-Si on a subnanometer scale is sufficient to provide optical absorption enhancement similar to that characteristic of a-Si. Based on our results, a more complete picture of optical absorption could be obtained with a multiscale study to better understand how both disorder and density contribute to optical absorption in Si. Furthermore, if the ability to precisely control vacancy cluster distributions and concentrations becomes feasible to an extent similar to modern control of Si dopant profiles, manipulation of vacancy cluster profiles could analogously provide a method to dictate Si absorption in a myriad of integrated optoelectronic and photovoltaic applications.

### IV. SUMMARY

To recapitulate, we use a DFT-based methodology to explore the structural contribution to optical absorption using charge-neutral, ground-state Si vacancy clusters ($V_n$, $n = 5$, 6, 7, and 32). We utilize an established PAW methodology based on computation of the longitudinal form of the dielectric matrix to calculate the imaginary and real parts of the complex dielectric function, $\varepsilon(\omega)$, and absorption coefficient, $\alpha(\omega)$. For small clusters ($V_n$, $n = 5$), we observe a substantial redshift of $\sim 2$ eV in the onset of incipient absorption as well as a magnitude of absorbance covering several orders of magnitude. We validated the generality of these observations by examining both $\varepsilon(\omega)$ and $\alpha(\omega)$ on larger, fourfold-coordinated (FC) clusters ($V_{12}$ and $V_{32}$) and compared the resulting spectra to the well-known cases of both a-Si and c-Si. The optical spectra of $V_{12}$ and $V_{32}$ not only substantiated generality of optical absorption enhancement by smaller Si vacancy clusters, but we also observed remarkable similarities in the broad absorption envelopes of FC vacancy clusters when compared to a-Si.

To isolate the effects of topological disorder observed in absorption spectra from the density variation that accompa-

### TABLE IV. Summary of variable-density a-Si sample structures. Structure names indicate the constituent number of Si atoms in each supercell. Strain energies are computed as formation energies referenced to c-Si, mass densities are given as $\rho_m$, and selected critical features of $\varepsilon_2(\omega)$ characterize the optical response of each structure. $E_2$ peak intensities are normalized to the basis a-Si480 structure.

<table>
<thead>
<tr>
<th>Structure</th>
<th>Strain energy (eV/atom)</th>
<th>$\rho_m$ (g/cm$^3$)</th>
<th>$E_0$ transition (eV)</th>
<th>$E_2$ transition (eV)</th>
<th>$E_2$ intensity (normalized)</th>
</tr>
</thead>
<tbody>
<tr>
<td>aSi480</td>
<td>0.153</td>
<td>2.251</td>
<td>0.81</td>
<td>3.47</td>
<td>1.000</td>
</tr>
<tr>
<td>aSi468</td>
<td>0.158</td>
<td>2.245</td>
<td>0.75</td>
<td>3.23</td>
<td>0.999</td>
</tr>
<tr>
<td>aSi456</td>
<td>0.149</td>
<td>2.236</td>
<td>0.75</td>
<td>3.47</td>
<td>0.986</td>
</tr>
<tr>
<td>aSi444</td>
<td>0.159</td>
<td>2.226</td>
<td>0.61</td>
<td>3.24</td>
<td>0.974</td>
</tr>
<tr>
<td>aSi432</td>
<td>0.167</td>
<td>2.210</td>
<td>0.68</td>
<td>3.27</td>
<td>0.952</td>
</tr>
</tbody>
</table>
nies vacancy incorporation, we constructed three additional 980-atom supercells \((V_6 \times 2, V_4 \times 3, \text{ and } V_3 \times 4)\), each containing twelve vacancies, by dispersing different combinations of small FC clusters in each supercell. Together with \(V_{12}\), this set of four constant-density \((V_6 \times q, \ n \times q = 12)\) vacancy structures provides a means to evaluate the impact of increasing structural disorder at a fixed density. Evaluation of these cluster dispersion exceeds that of a-Si up to about 1.5 eV. We characterized the structural disorder of each constant-density \(V_6 \times q\) bond topology by comparing CDFs of bond lengths and angles with reference to both a-Si and c-Si, which revealed approximately symmetric distributions of bond angles around the c-Si mean \((109.4\text{ }^\circ)\) as well as the expected tensile nature of bond-length distributions. More importantly, the constant-density CDFs also show increased spread across the same structural sequence \((V_{12} < V_6 \times 2 < V_4 \times 3 < V_3 \times 4)\) which provides evidence for increased structural disorder (strain) as a driving force for optical absorption enhancement. Furthermore, we used hydrostatic strain as a convenient method to uniformly change bond length at constant bond angle to reveal that increasing bond compression correlates to a redshift of the \(E_0\) transition, blueshift of the \(E_2\) transition, and concomitant increase in the \(E_2\) peak intensity.

For completeness, we extended our simulation approach to evaluate the effects of both defect concentration and density on optical absorption. We embedded the monovacancy in increasingly larger supercells \((192, 256, 480, 576, \text{ and } 980 \text{ atoms})\) to show a surprisingly strong concentration effect; extrapolation of these results provides sufficient doubt that our FC clusters can be optically resolved in bulk samples with modern instrumentation, which supports the difficulty reported in the literature in physical verification of small Si FC vacancy clusters. Using variable-density, void-free a-Si samples, we observed the subnanometer-scale effect of density on optical absorption to be secondary to the effect of structural disorder, contrary to expectations grounded in the EMA; however, this is justifiable since EMA is invalidated by subnanometer-scale clusters that are significantly smaller than the range of light wavelengths evaluated. We suspect the density effect to be significant, but practical supercell sizes using first-principles methods are insufficient to describe typical variable-density, multiphase morphologies (voids interspersed in a-Si) that are routinely observed in experimentally grown a-Si samples. Overall, our simulation results provide valuable insight into the fundamentals of subnanometer-scale optical absorption which indicates that increased structural disorder in c-Si via vacancy cluster incorporation constitutes a significant contribution to optical absorption in bulk Si.

**ACKNOWLEDGMENTS**

We acknowledge National Science Foundation (Grant No. CAREER-CTS-0449373) and Robert A. Welch Foundation (Grant No. F-1535) for their financial support. We would also like to thank the Texas Advanced Computing Center for use of their computing resources.

---

*Author to whom correspondence should be addressed; gshwang@che.utexas.edu


Comparison of \( \alpha(\omega) \) spectra for c-Si (480 atoms) in Fig. 3 and c-Si (980 atoms) in Fig. 4 reveals a notable difference in the feet where incipient absorption occurs between approximately 2.0 and 2.5 eV. Since we found it impractical to exactly match optical references across variable supercell sizes using our method, we have made an effort to present our results, whenever possible, in the context of a single basis supercell size.

The algorithm used to identify bond network connectivity searched each Si atom for neighbors within a fixed range, 2.1 \( \leq r_{ij} \leq 3.3 \) Å, where \( r_{ij} \) is the separation distance between two Si atoms. This range effectively sets a minimum and maximum for all bond lengths, so structure classification as fourfold-coordinated is somewhat subjective.

Our initial goal was to construct p-Si structures with the method detailed in Sec. II B because of their well known, but poorly understood, ability to both absorb (Ref. 14) and emit light (Refs. 14 and 16), but the distinctly different morphologies of the resulting structures led us to distinguish the structures generated as variable-density a-Si samples. A large reduction in density can certainly be interpreted as an increase in porosity, but the morphologies of reported (Refs. 14, 16, and 17) p-Si structures are typically categorized as either (1) irregular arrays of quantum wells (pores) or wires on a c-Si surface or (2) sponge-like matrices of embedded Si crystallites in an interconnected network with voids, a-Si, and/or amorphous silicon oxide pockets comprising the remaining volume.

Our initial goal was to construct p-Si structures with the method detailed in Sec. II B because of their well known, but poorly understood, ability to both absorb (Ref. 14) and emit light (Refs. 14 and 16), but the distinctly different morphologies of the resulting structures led us to distinguish the structures generated as variable-density a-Si samples. A large reduction in density can certainly be interpreted as an increase in porosity, but the morphologies of reported (Refs. 14, 16, and 17) p-Si structures are typically categorized as either (1) irregular arrays of quantum wells (pores) or wires on a c-Si surface or (2) sponge-like matrices of embedded Si crystallites in an interconnected network with voids, a-Si, and/or amorphous silicon oxide pockets comprising the remaining volume.

Ref. 14