

A Defect-Free Model of Amorphous Silicon with Pristine Electronic Structure

Louise A. M. Rosset,¹ Chinonso Ugwumadu,² Stephen R. Elliott,³ David A. Drabold,⁴ and Volker L. Deringer^{1,*}

¹*Inorganic Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford, UK*

²*Quantum & Condensed Matter (T-4) Group, Los Alamos National Laboratory, Los Alamos, NM, USA*

³*Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford, UK*

⁴*Department of Physics and Astronomy, Ohio University, Athens, OH, USA*

Amorphous silicon (*a*-Si) is understood to be the canonical continuous random network material, ideally defined by fully fourfold coordination. Here, we show that a defect-free (‘ideal’) model of *a*-Si from machine-learning-driven molecular-dynamics simulations [L. A. M. Rosset et al., Nat. Commun. **16**, 2360 (2025)], subsequently evaluated with hybrid-level density-functional theory computations, can accurately reproduce the experimentally observed electronic bandgap. We compare this model with one resulting from the Wooten–Winer–Weaire (WWW) bond-switching approach and with other recent approximants to ideal *a*-Si. More broadly, our work provides a platform for studies of band tails, optical properties, and transport in *a*-Si.

Amorphous silicon (*a*-Si) has long been thought of as a continuous random network (CRN) material [1], with fourfold coordination ($N = 4$ for all atoms) but without long-range structural order. The CRN model is an idealized picture that approximates the experimental reality, in which coordination defects ($N \neq 4$) are unavoidable [2], yet it is central to our fundamental understanding of disordered networks. Approaching the defect-free limit is relevant for precision instruments such as LIGO gravitational-wave detectors that rely on hydrogen-free *a*-Si for its low mechanical loss and correspondingly significant reduction in thermal noise [3].

Computationally, the CRN is often approximated using the Wooten–Winer–Weaire (WWW) bond-switching algorithm: starting from a crystalline (*c*-) Si structure, second-neighbor bonds are rearranged to change the network topology, whilst constraining it to $N = 4$ throughout [4]. Other protocols do not enforce fourfold coordination, such that almost all reported defect-free *a*-Si models have been generated with the WWW protocol [5].

Despite the existence of several defect-free structural models, to the authors’ knowledge, none has succeeded in demonstrating quantitative agreement with the experimental bandgap of *a*-Si, which varies between 1.5 and 1.8 eV, depending on the preparation method [6]. This is often due to localized states at tails of the valence- and conduction-band edges, thus reducing the electronic gap. Previous studies have focused on the effect of hydrogenation on cleaning up the mid-gap states by saturating dangling bonds and reducing network strain, thereby modifying the band tails and opening up the bandgap [7–9]. An ideal, defect-free network would not require hydrogenation, instead presenting a ‘clean’ electronic gap.

We recently created a large dataset of *a*-Si structures from a series of melt-quench molecular-dynamics simulations [10], driven by an efficient machine-learning-based interatomic potential (MLIP) [11, 12], systematically sampling the configurational space of *a*-Si by varying the mass density and quench rate. From this library of thousands of disordered structures, we discov-

ered a defect-free 216-atom model quenched at a rate of 10^{10} K/s, denoted as ‘Rosset-10’ in the following. We note that no coordination criterion was imposed during the simulation, so the system was not constrained to be defect-free, unlike in the WWW protocol. The generation of a defect-free *a*-Si network by melt-quenching is therefore a rare and stochastic event, and requires a large number of attempts, which can be accelerated using MLIPs [10, 13].

In this Letter, we characterize the electronic structure of our ‘Rosset-10’ structural model of *a*-Si, and show that it faithfully reproduces the value of the electronic bandgap from experiment. We compare it to three existing high-quality models of *a*-Si (Table I): (i) ‘WWW-216’, obtained from the WWW protocol [4] paired with the Keating potential [18]. Its electronic structure has been characterized previously [15, 19]; (ii) ‘Pedersen-Y’, obtained from a set of DFT annealing, relaxation, and rescaling procedures [16]. While this structure is not entirely defect-free (Table I), it agrees very well with various experimental measurements of short- and medium-range structural order; (iii) ‘Zongo-R2’, generated by the Activation Relaxation Technique nouveau (ARTn), which moves the system along the potential-energy surface by searching for first-order saddle points, coupled with an MLIP [5]. It is the only other defect-free 216-atom model that was not generated with the WWW protocol. We note that Table I is not an exhaustive survey of simulation protocols for *a*-Si: others exist, such as Monte-Carlo [20] or hybrid approaches [21–23].

We employed hybrid-level DFT computations, known to reproduce the experimental bandgap of *c*-Si [24, 25], and previously used for *a*-Si [26, 27]. Computations were carried out using a plane-wave energy cutoff of 540 eV and the projector augmented-wave formalism [28], as implemented in VASP [29, 30]. To evaluate all models on an equal footing, we first relaxed each model using the HSE06 functional [31, 32] at the Γ point only. We then computed electronic properties using a Γ -centered k -point grid of $2 \times 2 \times 2$ with the HSE06 functional. Further

TABLE I. Summary of the relaxed *a*-Si models discussed in the present study, detailing: generation protocol; number of atoms and percentage of 4-fold coordinated ones; mass density; porosity; percentage of locally crystal-like environments by Polyhedral Template Matching (PTM) and Common Neighbor Analysis (CNA); structural similarity to cubic diamond-type Si (*c*-Si) calculated using the SOAP kernel [14]; and energy relative to *c*-Si as predicted by HSE06-level DFT computations.

Sample	Protocol	No. of atoms	4-fold coord. (%)	Density (g/cm ³)	Porosity (%)	PTM (%)	CNA (%)	SOAP sim.	ΔE (eV/at.)
WWW-216 [15]	WWW bond-switching [4]	216	100	2.304	0.31	0	7.9	0.895	0.196
Pedersen-Y [16]	DFT relaxations	215	98.1	2.267	0.22	0	0	0.911	0.189
Zongo-R2 [5]	ARTn runs [17]	216	100	2.256	0.24	0.5	19.9	0.919	0.160
Rosset-10 [10]	Melt-quench, 10 ¹⁰ K/s rate	216	100	2.223	0.28	0	0	0.910	0.138

details are available in the Supplemental Material [13].

A common first criterion in validating a candidate amorphous structure is the predicted mass density. The density of *a*-Si thin films is estimated around 2.28 g/cm³ from ion-implanted samples [2, 33, 34]. The ‘Pedersen-Y’ model is the closest to this value, likely since volume scaling was included in the generation protocol, but the other models fall within 1–3% of the experimental value. The porosity is estimated using a fine 3D mesh and counting empty grid elements with a script developed in the context of Ref. 35. All models present low porosities, agreeing with reports that pores in *a*-Si are deposition artefacts [36] and should not occur in simulated quenched structures [37]. The higher porosity of the ‘WWW-216’ model arises from $m > 8$ rings [13].

To ascertain the CRN nature of the models, we probe the existence of local ordering and compare three common indicators, viz. Polyhedral Template Matching (PTM) [38], Common Neighbor Analysis (CNA) [39–42], and the averaged SOAP kernel similarity [14] to diamond-type (**dia**) Si, where 1 is identical. CNA appears more sensitive than PTM, indicating crystallinity in both the ‘WWW-216’ and ‘Zongo-R2’ models, while PTM only detects ordering in the latter. This crystallinity is consistent with reports of ordering in the original study [5]. In SOAP terms, the ‘Zongo-R2’ model is closest to **dia**, whereas ‘WWW-216’ is most dissimilar. ‘Rosset-10’ shows no apparent crystallinity and an average SOAP similarity to **dia** relative to the other models and the literature [10, 43]. Beyond global structural characteristics, the Supplemental Material includes metrics of short- and medium-range order [13].

The four models show large variations in their energies relative to **dia**, as computed with HSE06. In particular, the ‘Rosset-10’ model has the lowest energy among the four. We have previously shown that paracrystalline models have lower energies on average than fully disordered networks [10], hence this lower energy could reflect crystalline ordering that was not captured by PTM, CNA or SOAP. Yet, the ‘Rosset-10’ model provides excellent agreement with the calorimetrically-determined en-

thalpy of crystallization of $\Delta H = 0.142 \pm 0.003$ eV/atom obtained from ion-implanted samples [44]. Importantly, this means that only the ‘Rosset-10’ model fulfills the energy criterion set out in Ref. 45, that requires the energy of a candidate ideal *a*-Si model to fall within the range of 0.07 to 0.15 eV/atom above that of **dia**.

We now turn to the electronic densities of states (DOS), plotted in Fig. 1(a). The shape of the DOS is well reproduced throughout, with the 3s orbitals mainly contributing in the –15 to –5 eV range, and the 3p orbitals between –5 and 0 eV. The relative heights of the valence-band (VB) and conduction-band (CB) maxima are similar and agree well with previous literature [47, 48]. All three defect-free structures have clean bandgaps free of mid-gap states, while the DOS of the ‘Pedersen-Y’ model does have such states, associated with the aforementioned coordination defects (Table I).

Other important features of the DOS are the exponential tails at the VB and CB edges, known as Urbach tails [49], which have been resolved experimentally [50]. These tails have been linked to atomic filaments, whose electron eigenstates form interconnected clusters of charge [48], and also to topological filaments associated with bond stretching and bending [19, 51]. The ‘WWW-216’ and ‘Zongo-R2’ models show large tails at their CB and VB edges, respectively, while the ‘Rosset-10’ model shows small and narrower VB and CB tails.

We report the estimated values of the electronic bandgap, E_{gap} , in Fig. 1(b), defined as the difference in energy between the lowest occupied CB (LUMO) state and the highest occupied VB (HOMO) state. For the ‘Pedersen-Y’ model, we exclude the localized mid-gap states when defining the bandgap. There is a large variation in the estimated E_{gap} values between models, directly reflecting the differences in the Urbach tails. Previously reported values for *a*-Si in the computational literature range between 0.8 and 1.35 eV [53, 54]; the ‘WWW-216’, ‘Pedersen-Y’ and ‘Zongo-R2’ models are consistent with these. The ‘Rosset-10’ model has a substantially larger E_{gap} , which provides excellent and unprecedented agreement with the experimentally reported range of op-

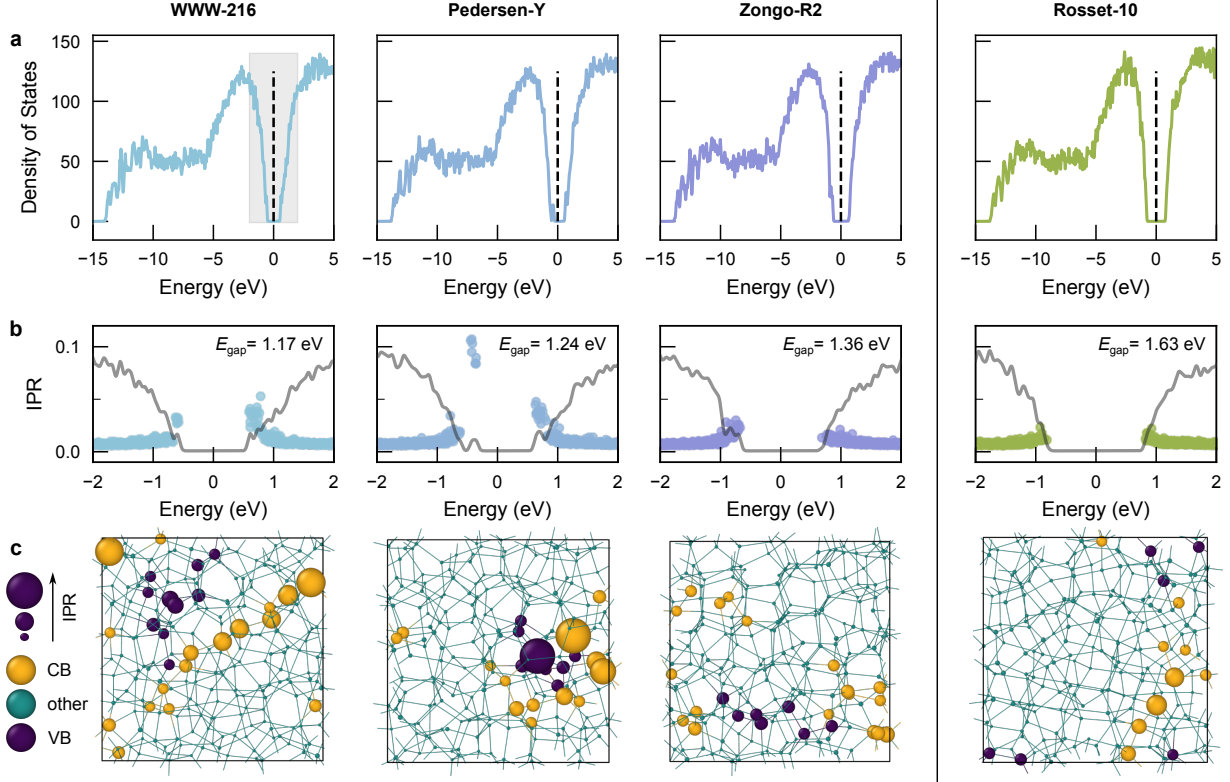


FIG. 1. Electronic structure. (a) Electronic densities of states of the four structural model considered in this study, where a dashed vertical line indicates the Fermi level (E_F) defined as the middle of the gap and centered at 0 eV. Enlarged plots are available in the Supplemental Material [13]. (b) Inverse Participation Ratio (IPR) plots in the region of the bandgap for each model, as shaded in gray for the ‘WWW-216’ model. The IPR is evaluated for each band and k -point pair; thick gray line overlays show the densities of states. The quoted E_{gap} value refers to the energy difference between the lowest occupied conduction-band (LUMO) state and the highest occupied valence-band (HOMO) state, omitting the mid-gap states for the ‘Pedersen-Y’ model. (c) Visualizations of the spatial localization of the highest IPR sites for each structure using OVITO [46], where atoms are color-coded by their IPR contribution to the VB, CB, or other as purple, yellow, or teal, respectively. The atomic radii are scaled proportionally to the relative IPR magnitudes, normalized against the largest IPR contribution across all structural models.

tical bandgaps of a -Si, typically between 1.5 and 1.8 eV [6, 13, 55–57]. We note that our results are obtained from computations with a $2 \times 2 \times 2$ k -point grid, and that the estimated E_{gap} varies with k -sampling and the associated dispersion in the supercell [58, 59]. Identical calculations at the Γ -point result in gaps that are systematically larger, by 0.02 to 0.12 eV, giving $E_{\text{gap}} = 1.71$ eV for ‘Rosset-10’.

To probe these variations in the electronic structure, we focus on the spatial localization of states around the Fermi level, E_F , in Fig. 1(b), plotting the DOS together with inverse participation ratio (IPR) values [60]:

$$\text{IPR}_{nk} = \frac{\sum_a |c_{nka}|^4}{(\sum_a |c_{nka}|^2)^2}$$

where n is the band index, k is the k -point index, a is the atomic orbital and c is the wavefunction coefficient, as calculated with PYPROCAR [61, 62]. The IPR mea-

sures how localized or extended the electron states are [48, 53]: low values ($\rightarrow 0$) correspond to delocalized or extended states that are distributed across the structure, while high values ($\rightarrow 1$) correspond to spatially localized states. However, it should be noted that it is not possible to distinguish unambiguously between delocalized and weakly localized states when the localization length is comparable to the simulation-box size, here ≈ 15 Å.

The ‘Pedersen-Y’ model has highly localized states adjacent to the VB edge, arising from the dangling-bond defect. It also displays some localization at the CB edge, associated with a broader conduction-band tail. The ‘WWW-216’ and ‘Zongo-R2’ models reveal appreciable localization at the CB edge and the VB edge, respectively. The ‘Rosset-10’ model exhibits no significant electron-localization, instead seemingly having fully extended (delocalized) states throughout both the VB and CB edges. It appears that the degree of localiza-

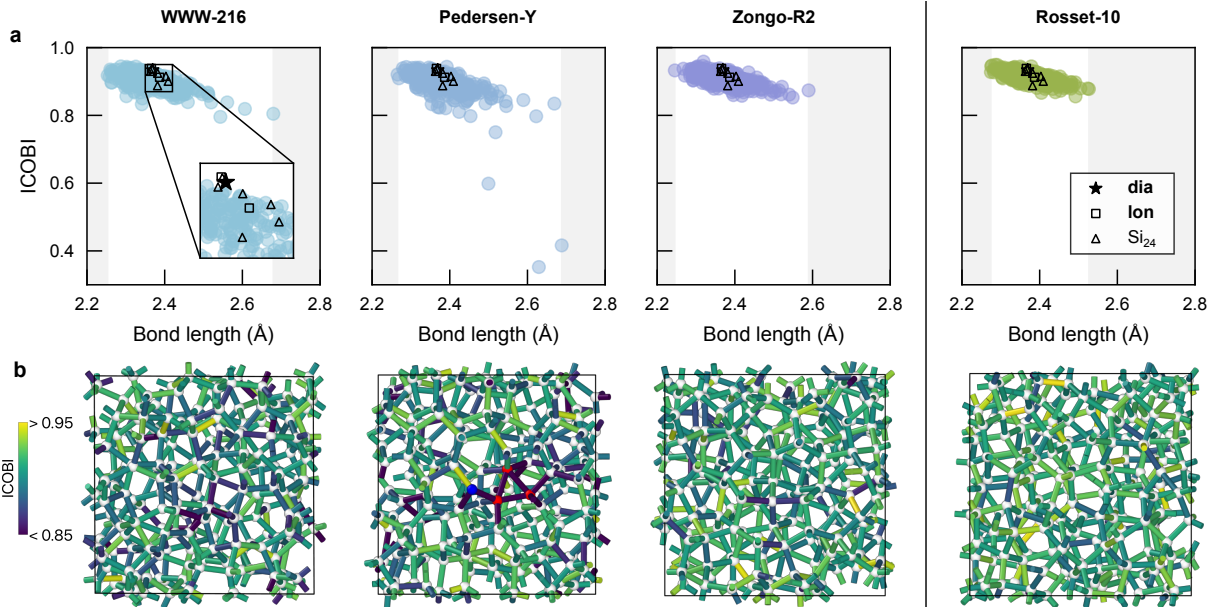


FIG. 2. Chemical bonding. (a) Scatter plots of integrated pairwise crystal orbital bond index ($\text{ICObI}^{(2)}$) values against bond lengths, where gray shading shows the edges of the bond-length distributions. An inset highlights relevant *c*-Si modifications: cubic diamond-type (“**dia**”; star), hexagonal diamond-type (“**lon**”; square) and the open-framework structure of Ref. 52 (“**Si₂₄**”; triangle). (b) Structure visualizations of the models using OVITO [46], where atoms are color-coded by coordination (blue/white/red for 3/4/5-fold coordination, respectively), and bonds are color-coded by ICObI value.

tion around the bandgap is inversely correlated with the width of the gap: the models with the smallest E_{gap} values, i.e., ‘WWW-216’ and ‘Pedersen-Y’, show the most tail localization, while the ‘Rosset-10’ model shows no localization and has the largest bandgap.

The width of the bandgap and the degree of localization are associated with the degree of structural disorder and the presence of coordination defects, respectively [63]. To examine the spatial character of the localized states, Fig. 1(c) shows the atomic regions contributing most strongly to the IPR for states near the band edges, where the atomic radii are scaled proportionally to the largest IPR contribution in the structure. The ‘WWW-216’ and ‘Pedersen-Y’ models exhibit localization on relatively small groups of atoms, arranged in compact clusters. This is consistent with the larger IPR values observed for these models, as higher IPR indicates stronger confinement of the electronic state to a small number of sites [63]. By contrast, the ‘Zongo-R2’ and ‘Rosset-10’ models show lower IPR values at the band edges and correspondingly more spatially extended localization patterns, which appear as interconnected chains within the network. This behavior is consistent with earlier studies showing that band-tail states in *a*-Si arise from chains of distorted bonds rather than isolated point-like defects [51, 63–65]. In particular, mildly short bonds contribute predominantly to VB tail states, while mildly long bonds contribute predominantly to CB tail states, a trend that

is reproduced by the ‘Pedersen-Y’ and ‘Rosset-10’ models, as shown in the Supplemental Material [13].

We turn to the chemical bonding in these structures to understand the large variation in E_{gap} and the exceptionally extended states in the ‘Rosset-10’ model. In Fig. 2(a), we plot the integrated crystal orbital bond index ($\text{ICObI}^{(2)}$), obtained with LOBSTER [13, 66–68]. $\text{ICObI}^{(2)}$ quantifies the pairwise bond order, such that strong covalent bonds have a value close to 1 [68]. Stretched bonds in the ‘WWW-216’ and ‘Pedersen-Y’ models result in lower ICObI values, but there are strong outliers in ‘Pedersen-Y’ falling beyond the exponential bond-length–bond-strength relationship [69]. The distributions for the ‘Zongo-R2’ and ‘Rosset-10’ models are qualitatively similar, but the bond-length distribution of the latter is narrower, resulting in a more compact scatter. For comparison, an inset shows ICObI values for crystalline allotropes of Si: cubic and hexagonal diamond-type, and an open-framework structure [52].

To resolve the spatial character of bonding, we visualize the structural models in Fig. 2(b), coloring bonds by their ICObI values. The low values in the ‘Pedersen-Y’ model arise from a 3-membered ring formed by the overcoordinated Si atoms (red). The three long bonds strain their local environment into a larger cluster with disrupted bonding, in turn leading to the localized states in the gap and in the band tails. The bonds around the undercoordinated Si atom (blue) also leads to local

strain and higher ICOBI values. In the ‘WWW-216’ model, some bonds with low ICOBI values belong to 4-membered rings, while others form linear chains, or isolated stretched bonds. Conversely, the ICOBI values are more uniformly distributed in the ‘Zongo-R2’ and ‘Rosset-10’ models, and the bonds with low ICOBI values do not cluster.

A model of *a*-Si with no localized states, a narrow bond-angle distribution, and low excess energy with respect to **dia** has been hypothesized to represent an *ideal a*-Si network [53]. All three criteria are fulfilled with the ‘Rosset-10’ model, as shown in Fig. 1, Fig. S2 [13], and Table I. We further propose that an ‘ideal’ *a*-Si model must have a narrow distribution of bond lengths without excessively short nor long bonds, a criterion that appears necessary to achieve an absence of localized states at the VB and CB band edges, respectively.

Looking forward, our model could serve as a platform to further understand the effect of defects and dopants in *a*-Si [70], or to make accurate predictions of carrier mobility [71]. These effects are of interest since *a*-Si is used in both intrinsic and doped forms in solar-cell heterojunctions [72]. The model’s large bandgap may be of interest in studying the optical absorption of *a*-Si, relevant for LIGO applications [3]. Our model could also be used in the future to investigate why *a*-Si has a larger bandgap than the indirect gap in **dia**, in an analogous study to the one recently conducted on *c*-Si in Ref. 73. More broadly, this model of an ideal random network opens up new avenues for the exploration of the electronic structure of CRN-like disordered materials.

This work was supported by UK Research and Innovation [grant number EP/X016188/1]. C.U. acknowledges support from the laboratory directed Research and Development (LDRD) program at Los Alamos National Laboratory (LANL) through a Director’s Postdoctoral Fellowship (Project No. 20240877PRD4). S.R.E. is grateful to the Leverhulme Trust (UK) for a Fellowship. D.A.D. acknowledges support from the U.S. National Science Foundation (Project No. MRI-2320493). We are grateful for computational support from the UK national high performance computing service, ARCHER2, for which access was obtained via the UKCP consortium and funded by EPSRC [grant number EP/X035891/1]. Data supporting this work are provided at <https://github.com/vldgroup/papers-defect-free-a-si>.

* volker.deringer@chem.ox.ac.uk

[1] W. H. Zachariasen, *J. Am. Chem. Soc.* **54**, 3841 (1932).
 [2] K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J. L. Robertson, J. Wang, and S. C. Moss, *Phys. Rev. B* **60**, 13520 (1999).
 [3] R. Birney, J. Steinlechner, Z. Tornasi, S. MacFoy, D. Vine, A. S. Bell, D. Gibson, J. Hough, S. Rowan,

P. Sortais, S. Sproules, S. Tait, I. W. Martin, and S. Reid, *Phys. Rev. Lett.* **121**, 191101 (2018).
 [4] F. Wooten, K. Winer, and D. Weaire, *Phys. Rev. Lett.* **54**, 1392 (1985).
 [5] K. Zongo, H. Sun, C. Ouellet-Plamondon, N. Mousseau, and L. K. Béland, *Phys. Rev. B* **111**, 214209 (2025).
 [6] R. H. Klazes, M. H. L. M. van den Broek, J. Bezemer, and S. Radelaar, *Philos. Mag. B* **45**, 377 (1982).
 [7] W. Y. Ching, D. J. Lam, and C. C. Lin, *Phys. Rev. B* **21**, 2378 (1980).
 [8] G. Allan, C. Delerue, and M. Lannoo, *Phys. Rev. B* **57**, 6933 (1998).
 [9] R. V. Meidanshahi, S. Bowden, and S. M. Goodnick, *Phys. Chem. Chem. Phys.* **21**, 13248 (2019).
 [10] L. A. M. Rosset, D. A. Drabold, and V. L. Deringer, *Nat. Commun.* **16**, 2360 (2025).
 [11] J. D. Morrow and V. L. Deringer, *J. Chem. Phys.* **157**, 104105 (2022).
 [12] The MLIP model of Ref. 11 was distilled from the Si-GAP-18 potential introduced in A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Phys. Rev. X* **8**, 041048 (2018).
 [13] Further details are given in the Supplementary Material.
 [14] A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
 [15] B. R. Djordjević, M. F. Thorpe, and F. Wooten, *Phys. Rev. B* **52**, 5685 (1995).
 [16] A. Pedersen, L. Pizzagalli, and H. Jónsson, *New J. Phys.* **19**, 063018 (2017).
 [17] N. Mousseau, L. K. Béland, P. Brommer, J.-F. Joly, F. El-Mellouhi, E. Machado-Charry, M.-C. Marinica, and P. Pochet, *J. Phys. B: At. Mol. Phys.* **2012**, 925278 (2012).
 [18] P. N. Keating, *Phys. Rev.* **145**, 637 (1966).
 [19] Y. Pan, M. Zhang, and D. Drabold, *J. Non-Cryst. Solids* **354**, 3480 (2008).
 [20] G. Opletal, T. C. Petersen, I. K. Snook, and D. G. McCulloch, *J. Chem. Phys.* **126**, 214705 (2007).
 [21] M. J. Cliffe, M. T. Dove, D. A. Drabold, and A. L. Goodwin, *Phys. Rev. Lett.* **104**, 125501 (2010).
 [22] A. Pandey, P. Biswas, and D. A. Drabold, *Sci. Rep.* **6**, 33731 (2016).
 [23] D. Igram, B. Bhattarai, P. Biswas, and D. Drabold, *J. Non-Cryst. Solids* **492**, 27 (2018).
 [24] J. Heyd, J. E. Peralta, G. E. Scuseria, and R. L. Martin, *J. Chem. Phys.* **123**, 174101 (2005).
 [25] H. J. Xiang, B. Huang, E. Kan, S.-H. Wei, and X. G. Gong, *Phys. Rev. Lett.* **110**, 118702 (2013).
 [26] K. Jarolimek, E. Hazrati, R. De Groot, and G. De Wijs, *Phys. Rev. Appl.* **8**, 014026 (2017).
 [27] J. D. Morrow, C. Ugwumadu, D. A. Drabold, S. R. Elliott, A. L. Goodwin, and V. L. Deringer, *Angew. Chem. Int. Ed.* **63**, e202403842 (2024).
 [28] P. E. Blöchl, *Phys. Rev. B* **50**, 17953 (1994).
 [29] G. Kresse and J. Furthmüller, *Phys. Rev. B* **54**, 11169 (1996).
 [30] G. Kresse and D. Joubert, *Phys. Rev. B* **59**, 1758 (1999).
 [31] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **118**, 8207 (2003).
 [32] J. Heyd, G. E. Scuseria, and M. Ernzerhof, *J. Chem. Phys.* **124**, 219906 (2006).
 [33] J. S. Custer, M. O. Thompson, D. C. Jacobson, J. M. Poate, S. Roorda, W. C. Sinke, and F. Spaepen, *Appl. Phys. Lett.* **64**, 437 (1994).

- [34] K. Laaziri, S. Roorda, and J. M. Baribeau, *J. Non-Cryst. Solids* **191**, 193 (1995).
- [35] Y. Liu, Y. Zhou, R. Ademuwagun, L. Walterbos, J. George, S. R. Elliott, and V. L. Deringer, *J. Am. Chem. Soc.* **148**, 9400 (2026).
- [36] S. C. Moss and J. F. Graczyk, *Phys. Rev. Lett.* **23**, 1167 (1969).
- [37] L. J. Lewis, *J. Non-Cryst. Solids* **580**, 121383 (2022).
- [38] P. M. Larsen, S. Schmidt, and J. Schiøtz, *Model. Simul. Mater. Sci. Eng.* **24**, 055007 (2016).
- [39] J. D. Honeycutt and H. C. Andersen, *J. Phys. Chem.* **91**, 4950 (1987).
- [40] A. S. Clarke and H. Jónsson, *Phys. Rev. E* **47**, 3975 (1993).
- [41] A. Stukowski, *Model. Simul. Mater. Sci. Eng.* **20**, 045021 (2012).
- [42] E. Maras, O. Trushin, A. Stukowski, T. Ala-Nissila, and H. Jónsson, *Comput. Phys. Commun.* **205**, 13 (2016).
- [43] N. Bernstein, B. Bhattacharai, G. Csányi, D. A. Drabold, S. R. Elliott, and V. L. Deringer, *Angew. Chem. Int. Ed.* **58**, 7057 (2019).
- [44] S. Roorda and L. J. Lewis, *Science* **338**, 1539 (2012).
- [45] D. A. Drabold, *Phys. Status Solidi Rapid Res. Lett.* **5**, 359 (2011).
- [46] A. Stukowski, *Model. Simul. Mater. Sci. Eng.* **18**, 015012 (2010).
- [47] M. F. Thorpe and D. Weaire, *Phys. Rev. Lett.* **27**, 1581 (1971).
- [48] J. Dong and D. A. Drabold, *Phys. Rev. Lett.* **80**, 1928 (1998).
- [49] F. Urbach, *Phys. Rev.* **92**, 1324 (1953).
- [50] S. Aljishi, J. D. Cohen, S. Jin, and L. Ley, *Phys. Rev. Lett.* **64**, 2811 (1990).
- [51] D. A. Drabold, Y. Li, B. Cai, and M. Zhang, *Phys. Rev. B* **83**, 045201 (2011).
- [52] D. Y. Kim, S. Stefanoski, O. O. Kurakevych, and T. A. Strobel, *Nat. Mater.* **14**, 169 (2015).
- [53] M. Durandurdu, D. A. Drabold, and N. Mousseau, *Phys. Rev. B* **62**, 15307 (2000).
- [54] G. T. Barkema and N. Mousseau, *Phys. Rev. B* **62**, 4985 (2000).
- [55] G. K. M. Thutupalli and S. G. Tomlin, *J. Phys. C: Solid State Phys.* **10**, 467 (1977).
- [56] N.-M. Park, C.-J. Choi, T.-Y. Seong, and S.-J. Park, *Phys. Rev. Lett.* **86**, 1355 (2001).
- [57] C.-C. Huang, Y. Tang, M. van der Laan, J. van de Groep, A. F. Koenderink, and K. Dohnalová, *ACS Appl. Nano Mater.* **4**, 288 (2021).
- [58] D. A. Drabold, P. A. Fedders, O. F. Sankey, and J. D. Dow, *Phys. Rev. B* **42**, 5135 (1990).
- [59] K. Prasai, P. Biswas, and D. A. Drabold, *Semicond. Sci. Technol.* **31**, 073002 (2016).
- [60] D. J. Thouless, *Phys. Rep.* **13**, 93 (1974).
- [61] U. Herath, P. Tavadze, X. He, E. Bousquet, S. Singh, F. Muñoz, and A. H. Romero, *Comput. Phys. Commun.* **251**, 107080 (2020).
- [62] L. Lang, P. Tavadze, A. Tellez, E. Bousquet, H. Xu, F. Muñoz, N. Vasquez, U. Herath, and A. H. Romero, *Comput. Phys. Commun.* **297**, 109063 (2024).
- [63] Y. Pan, F. Inam, M. Zhang, and D. A. Drabold, *Phys. Rev. Lett.* **100**, 206403 (2008).
- [64] P. A. Fedders, D. A. Drabold, and S. Nakhmanson, *Phys. Rev. B* **58**, 15624 (1998).
- [65] F. Inam, J. P. Lewis, and D. A. Drabold, *Phys. Status Solidi A* **207**, 599 (2010).
- [66] S. Maintz, V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, *J. Comput. Chem.* **34**, 2557 (2013).
- [67] R. Nelson, C. Ertural, J. George, V. L. Deringer, G. Hautier, and R. Dronskowski, *J. Comput. Chem.* **41**, 1931 (2020).
- [68] P. C. Müller, C. Ertural, J. Hempelmann, and R. Dronskowski, *J. Phys. Chem. C* **125**, 7959 (2021).
- [69] L. Pauling, *J. Am. Chem. Soc.* **69**, 542 (1947).
- [70] K. Winer, I. Hirabayashi, and L. Ley, *Phys. Rev. Lett.* **60**, 2697 (1988).
- [71] Y. Lee, Y. Hu, D. Kim, S. Datta, and K. Cho, *Phys. Rev. B* **105**, 085201 (2022).
- [72] H. Lin, M. Yang, X. Ru, G. Wang, S. Yin, F. Peng, C. Hong, M. Qu, J. Lu, L. Fang, C. Han, P. Procel, O. Isabella, P. Gao, Z. Li, and X. Xu, *Nat. Energy* **8**, 789 (2023).
- [73] E. Oliphant, V. Mantena, M. Brod, G. J. Snyder, and W. Sun, *Mater. Horiz.* **12**, 3073 (2025).

Supplemental Material for
**‘A Defect-Free Model of Amorphous Silicon with Pristine
Electronic Structure’**

Louise A. M. Rosset¹, Chinonso Ugwumadu², Stephen R. Elliott³,
David A. Drabold⁴, and Volker L. Deringer*¹

¹Inorganic Chemistry Laboratory, Department of Chemistry, University of Oxford, Oxford,
UK

²Quantum & Condensed Matter (T-4) Group, Los Alamos National Laboratory, Los
Alamos, NM, USA

³Physical and Theoretical Chemistry Laboratory, Department of Chemistry, University of
Oxford, Oxford, UK

⁴Department of Physics and Astronomy, Ohio University, Athens, OH, USA

*volker.deringer@chem.ox.ac.uk

S1 Methods

S1.1 Structural models

The ‘Rosset-10’ model originates from a dataset generated in Ref. [S1](#) comprised of structures of 64, 216, 512, and 1,000 atoms. None of the structures of 512 or 1,000 atoms were defect-free. The structures of 64 atoms were omitted due to their very small size. Within the subset of 782 structural models of 216 atoms, only four were defect-free. From these four, two showed substantial paracrystalline content according to CNA and PTM analysis (see main text). This would give a rate of generating continuous random networks that are defect-free of $\approx 0.26\%$. However, if we consider that only 381 structures were generated with a slow quench rate, that is, either 10^{10} or 10^{11} K/s, then this rate increases to $\approx 0.52\%$.

Of the two defect-free CRN-like models, we focused on the one with a larger bandgap and narrower bond-length distribution, denoted ‘Rosset-10’ as it was generated with a quench rate of 10^{10} K/s. The other model, generated with a quench rate of 10^{11} K/s, showed a smaller bandgap of 1.42 eV, with localization on the edge of the valence band. It has a wider bond-length distribution, with $\approx 0.4\%$ of bonds with $l > 2.6$ Å.

S1.2 DFT computations

DFT computations were carried out in VASP v6.5.1, [S2](#)[S3](#) using the projector augmented-wave formalism. [S4](#)[S5](#) All structural models were relaxed with HSE06 using a plane-wave energy cutoff of 540 eV at the Γ point, with a convergence criterion of 10^{-6} eV.

To accelerate our computations with the HSE06 functional, a pre-convergence with PBE was performed to a convergence criterion of 10^{-6} eV. The WAVECAR and CHGCAR were then used to initialize the HSE06 self-consistent field (SCF) computation. The HSE computation was performed with a range-separation parameter of 0.2 Å $^{-1}$, a fraction of exact exchange of 0.25, a plane wave energy cutoff of 540 eV, a Gaussian smearing width of 0.05 eV, a Γ -centered k -point grid of $2 \times 2 \times 2$, and a convergence criterion of 10^{-6} eV/cell.

We did not include non-spherical contributions related to the gradient of the density in the PAW spheres (LASPH tag). A test in which we set `LASPH=.TRUE.` for the HSE06 computation resulted in a change of 0.1 meV in the predicted E_{gap} of the ‘Rosset-10’ model, much smaller than the precision to which we report E_{gap} (0.01 eV or 10 meV).

We used LOBSTER v5.1.1 for chemical-bonding analysis on the HSE06 SCF outputs, using 3s 3p basis functions, with a Gaussian smearing width of 0.05 eV. We consider chemical bonding interactions for pairs within a cutoff of 2.85 Å, and an energetic window from –15 eV to 10 eV relative to the Fermi level.

S1.3 Experimental bandgaps

Table [S1](#) summarizes experimental literature relating to measurements of the bandgap of amorphous silicon.

Table S1: Summary of reported bandgaps from experimental literature, detailing the sample preparation protocol, the sample type, the bandgap measurement method, the bandgap (E_{gap}), and the year and number of the literature reference.

Protocol	Sample type	Method	E_{gap} (eV)	Year	Reference
Evaporation	Thin film	Tauc with $r = 2$	1.82–1.83	1977	S6
Evaporation	Thin film	Tauc with $r = 2$	1.5	1993	S7
Sol–gel synthesis	Nanoparticle	Photoluminescence	1.53 – 1.82	2021	S8
Sputtering	Thin film	Tauc with $r = 3$	1.1–1.4	1982	S9
PECVD ¹	Thin film	Tauc with $r = 3$	1.6–1.8	1982	S9
PECVD ²	Thin film	Tauc with $r = 2$	1.49	1993	S10
PECVD ³	Quantum dot	Photoluminescence	1.56	2001	S11

¹ Hydrogen content in these samples is not discussed in the reference work.

² Samples contain less than 0.3 at.-% H.

³ Samples were broadly described as having ‘very little’ hydrogen. [S11](#)

The entries in Table [S1](#) vary in preparation protocol, sample type, and in the approximation used to calculate the E_{gap} value from experimental findings. Broadly, the reported bandgaps range from 1.5 to 1.8 eV. We omit the sputtered samples from Ref. [S9](#) from this range as they appear untypically low compared to other reports for *a*-Si, and as the original paper reports that sputtering has yielded a ‘different material’ to the samples obtained in the same study

from glow discharge deposition^{S9} (denoted ‘PECVD’, plasma enhanced chemical vapor deposition, in Table S1).

We note that samples synthesized by PECVD from SiH₄ or related precursors may contain some amount of hydrogen, which may artificially increase the reported bandgap value. Hydrogen contents for each of these samples are discussed as footnotes in Table S1.

We also note that most of these experiments were performed over 30 years ago, and that new, high quality, experimental data would substantially strengthen the comparison between experimental results and computational models.

S1.4 Density of states

We assess whether the resolution of the density of states computation (Fig. S1(a)) is sufficient by plotting the log of the density of states (Fig. S1(b)), here shown for the ‘Rosset-10’ structure as an example. The smooth behavior captured in the logarithmic plot suggests that the resolution is sufficient.

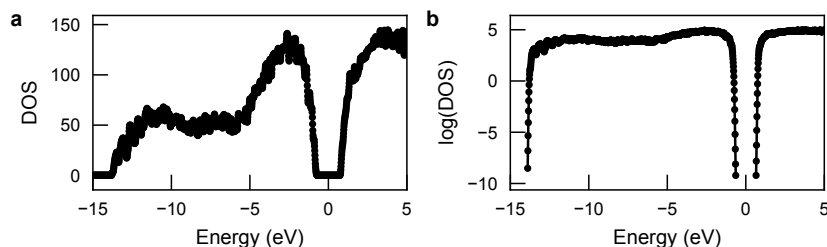


Figure S1: Plots of (a) the density of states and (b) the logarithm of the density of states for the ‘Rosset-10’ model.

S1.5 Molecular-dynamics annealing

We performed a 1 ns annealing run at 300 K and 1 bar in the NpT ensemble using the Si-GAP-18 model^{S12} to obtain well-equilibrated structural models. The Si-GAP-18 model has been extensively validated against experimental data for a -Si^{S13 S14}. The last 100 ps were

used for structural analysis. In Table [S2](#), we investigate the effect of the annealing treatment on some global characteristics, comparing to the structures before annealing (Table 1).

Table S2: Global characteristics of the HSE06-relaxed structural models (‘Static’) compared to the MD-annealed models (‘MD’): mass density, coordination, percentage of locally crystalline-like environments by PTM and CNA, SOAP similarity to **dia**, and the Si-GAP-18 predicted energy of the model relative to **dia**. Results for the MD-annealed models are obtained from averaging over the last 100 ps.

Sample	ρ (g/cm ³)		4-fold coord.		PTM (%)		CNA (%)		SOAP sim. dia		ΔE (eV/at.)	
	Static	MD	Static	MD	Static	MD	Static	MD	Static	MD	Static	MD
WWW-216	2.304	2.251	100	98.9	0	0	7.9	3.3	0.895	0.890	0.193	0.226
Pedersen-Y	2.267	2.276	98.1	97.4	0	0	0	4.9	0.911	0.904	0.186	0.217
Zongo-R2	2.256	2.260	100	99.9	0.5	0.16	19.9	15.7	0.919	0.912	0.155	0.191
Rosset-10	2.223	2.239	100	100	0	0	0	3.1	0.910	0.904	0.138	0.175

After a long 1 ns annealing run, both the ‘Zongo-R2’ and ‘Rosset-10’ models retained ideal fourfold connectivity (to within $\approx 0.1\%$), whereas the other two models contained a small number of coordination defects.

S2 Supplemental results

S2.1 Structural analysis

We report metrics of short- and medium range order in Fig. S2, first plotting bond-length distributions. All four structural models show maxima around 2.38 Å (Table S3), which compares well to the experimental average of 2.35 Å (ion-implanted samples; Ref. S15). An experimental standard deviation of ≈ 0.065 Å was reported, wider than that of the models obtained by peak fitting (≈ 0.04 Å), but the authors of Ref. S15 warn that some broadening may arise from vibrational disorder and insufficient experimental resolution. The tails of the bond-length distributions are where the models differ structurally: while ‘Rosset-10’ and ‘Zongo-R2’ show fewer than 0.5% of short bonds with $l < 2.2$ Å or of long bonds with $l > 2.6$ Å, the ‘WWW-216’ and ‘Pedersen-Y’ models show significant tails from overstretched bonds, 1.0% and 1.4% respectively, as indicated by arrows in Fig. S2. We note that the bond-length distribution presented in Fig. S2(a) is broader than the one shown in Fig. 3 of the main text – Fig. S2(a) shows the distributions obtained from a molecular-dynamics annealing simulation, while the distribution in Fig. 3 is the static HSE06-relaxed structure. The annealing simulations induce thermal vibrations which broaden the distribution of bond lengths.

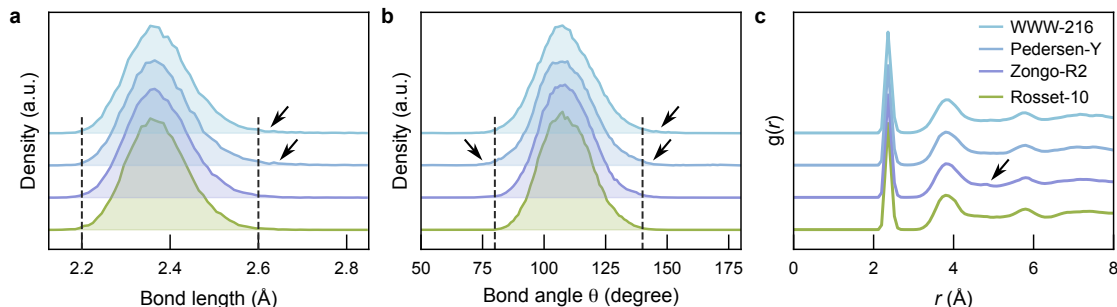


Figure S2: Characteristics of short- and medium-range order. (a) Bond-length distribution, where arrows highlight the distribution tails, and dashed vertical lines indicate short ($r < 2.2$ Å) and long ($r > 2.6$ Å) bonds. (b) Bond-angle distribution, where dashed vertical lines indicate small ($\theta < 80^\circ$) and large ($\theta > 140^\circ$) angles, and arrows highlight the distribution tails. (c) Radial distribution function, $g(r)$, where an arrow indicates an enhancement in the ‘Zongo-R2’ model. Distributions are vertically offset for clarity, and averaged over the last 100 ps of a 1 ns NpT annealing simulation at 300 K and 1 bar, driven by the Si-GAP-18 potential.^{S12}

We plot the bond-angle distributions in Fig. S2(b), with complementary information detailed in Table S3. All four models agree reasonably with the experimental average of $107.83^\circ \pm 0.97^\circ$ (Ref. S15). The ‘Zongo-R2’ model has a slightly narrower bond-angle distribution, likely due to paracrystalline ordering; ‘WWW-216’ and ‘Pedersen-Y’ show broader tails, with distorted small ($\theta < 80^\circ$) and large angles ($\theta > 140^\circ$).

The radial distribution functions [Fig. S2(c)] are similar across all four models, with well-defined peaks for the first- and second-neighbor shells. The ‘Zongo-R2’ model shows a small feature around 4.5 Å: an indicator of paracrystalline ordering observed in experimental S16S17 and computational studies, S1S18 consistent with metrics of crystallinity given in Table 1 of the main text. The ‘Rosset-10’ model does not show such a feature, supporting its CRN character.

Table S3: Mean and standard deviation of the bond-length and bond-angle distributions for each structural model.

Sample	$\langle r \rangle$ (Å)	σ_r (Å)	$\langle \theta \rangle$ (°)	σ_θ (°)
WWW-216	2.382	0.041	109.2	10.48
Pedersen-Y	2.379	0.039	108.9	11.03
Zongo-R2	2.375	0.046	109.2	8.84
Rosset-10	2.376	0.044	109.3	10.05

We plot the ring-size distributions in Fig. S3. The ‘WWW-216’ model has a slightly lower count of $m = 6$ membered rings than the other three structures, and instead has a large count of $m > 8$ rings. Unlike the other three models, the ‘Rosset-10’ has a higher count of $m = 5$ membered rings than $m = 7$ rings.

We show the distributions of dihedral angles in Fig. S4. The dihedral-angle distribution has been linked to the existence of an enhancement in the RDF around ≈ 4.5 Å, S15 and discussed as an indicator of paracrystallinity. S1 The ‘Zongo-R2’ and ‘Rosset-10’ models have similarly sharp distributions, which may be an indicator of paracrystalline ordering in the latter which is not apparent from the other structural metrics presented here, namely PTM, CNA, SOAP,

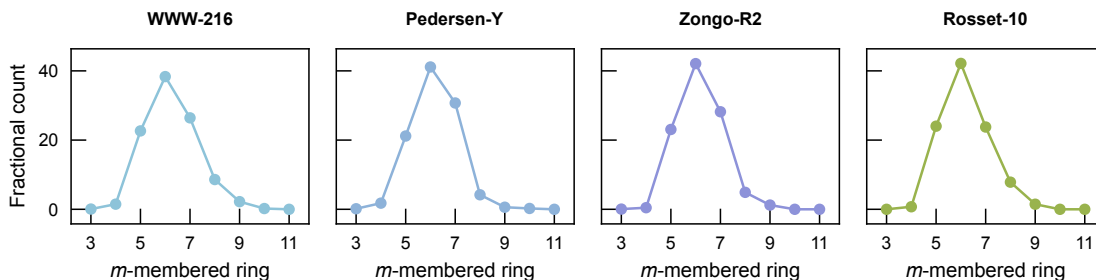


Figure S3: Distributions of the shortest closed-path rings, calculated with MATSCIPY^{S19} obtained from averaging over the last 100 ps of a 1 ns molecular-dynamics simulation at 300 K.

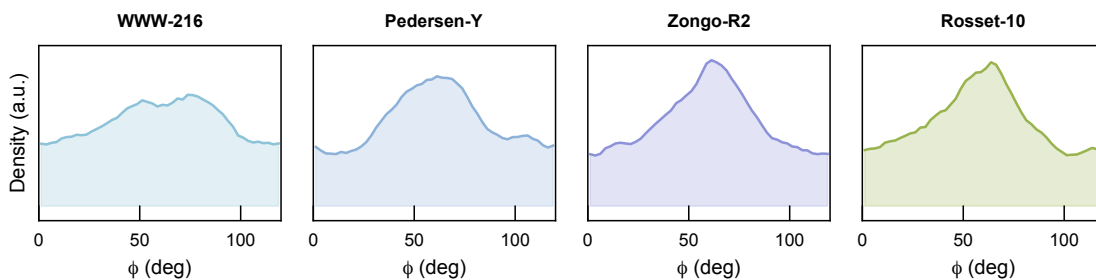


Figure S4: Dihedral-angle distributions, obtained from averaging over the last 100 ps of a 1 ns molecular-dynamics simulation at 300 K.

and the RDF.

We also plot the structure factors of the models in Fig. S5, comparing to the high-quality experimental data from Ref. S15. We note that significant size effects must be considered for models of only 216 atoms. The ‘Rosset-10’ model provides the best match to the high- Q part of the structure factor, but presents a shoulder at the foot of the second sharp diffraction peak. Only the ‘Zongo-R2’ model appears completely devoid of this feature.

Finally, we probe the information provided by local environments, which has been shown to be valuable in quantifying local disorder.^{S1S21} Here, we plot the local SOAP similarity to **dia** of each atomic environment against its atomic energy above that of **dia** in Fig. S6. To compare the local environments in these models to those from well-annealed disordered samples, we take the subset of disordered ‘CRN’-like networks of 216 atoms generated in Ref. S1. We show all the local environments from this subset in gray. As follows from the chemical-bonding analysis in Fig. 3, the ‘Rosset-10’ model shows a tighter scatter and more

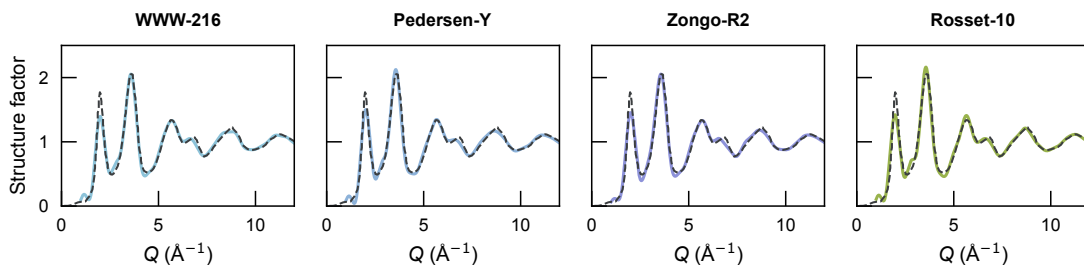


Figure S5: Computed structure factor of each structural model, obtained from the Fourier transform of the radial distribution function from OVITO^[S20]. Data are averaged over the last 100 ps of a 1 ns MD simulation at 300 K. Experimental data from Ref. [S15] are plotted as a dashed black line.

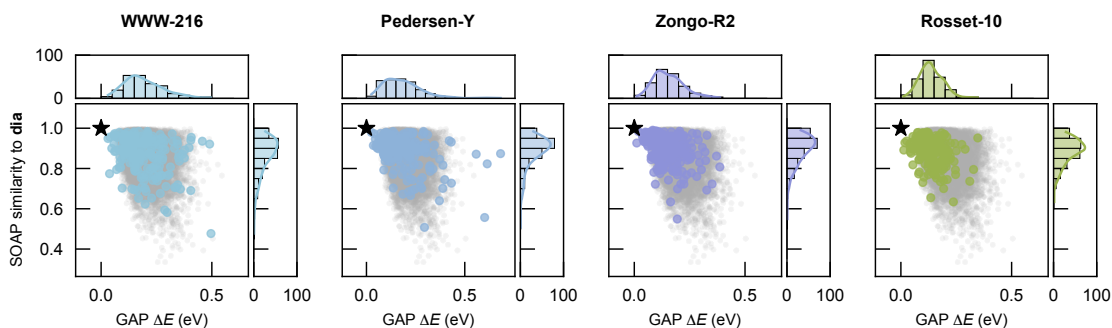


Figure S6: Structure and energetics of individual atomic environments, plotted in the style of Fig. 2 from Ref. [S1]. Scatter plot of the GAP-predicted atomic energies relative to cubic-diamond-type Si (**dia**), against the atomistic SOAP similarity to **dia** for each structural model. A black star indicates the ideal **dia** environment. An underlying set of data points (gray) shows the distribution of local atomic environments in the dataset of disordered *a*-Si structures of 216 atoms from Ref. [S1]. Histograms of the total distribution and kernel density estimates are shown for each axis.

homogeneous disordered environments than the other models, both in terms of local structure and local energy.

We calculate the information entropy of these distribution using: $H = -\sum_i p_i \times \log p_i$, where p_i is the probability of bin i . In relative terms, low entropy is indicative of a more ‘predictable’ distribution, such that the probabilities are concentrated on a small number of bins; high entropy is indicative of a less ‘predictable’ distribution where the probabilities are spread out across many different bins. The SOAP similarity to **dia** distributions have entropy values of 1.84 for ‘WWW-216’, 1.68 for ‘Pedersen-Y’, 1.67 for ‘Zongo-R2’, and 1.64 for ‘Rosset-10’. The ΔE distributions have entropy values of 1.92 for ‘WWW-216’, 1.90 for ‘Pedersen-Y’,

1.71 for ‘Zongo-R2’, and 1.49 for ‘Rosset-10’.

S2.2 Density of states

To further support the discussion of the Urbach tails from Fig. 1 of the main text, we show an enlarged plot of the valence- and conduction-band edges in Fig. S7. The features are somewhat sharp from the low Gaussian broadening of 0.05 eV used in the DFT computations.

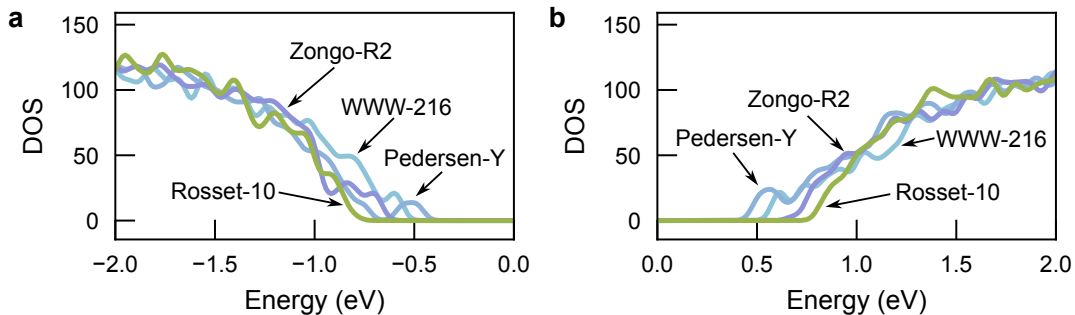


Figure S7: Enlarged plots of the density of states at the: (a) valence-; and (b) conduction-band edges for all four structural models.

The E_{gap} values in Fig. 1 are computed with a k -point grid of $2 \times 2 \times 2$. For a Γ -point only calculation, the computed E_{gap} values are: 1.19 eV for ‘WWW-216’, 1.05 eV for ‘Pedersen-Y’, 1.50 eV for ‘Zongo-R2’, and 1.71 eV for ‘Rosset-10’.

S2.3 Inverse Participation Ratio (IPR)

We examine the mean bond length of atoms contributing most strongly to the high-IPR states near the valence- and conduction-band edges in the four structural models in Fig. S8. The resulting trends are consistent with earlier theoretical work showing that band-tail states in a -Si arise from correlated structural disorder rather than isolated point defects. Specifically, valence-tail states are associated with relatively short-bond networks, whereas conduction-tail states are associated with relatively long-bond networks.^{S22,S24} This trend is reproduced

by the ‘Pedersen-Y’ and ‘Rosset-10’ models. For the ‘WWW-216’ and ‘Zongo-R2’ models, the contributions to both the VB and CB edges arise from longer-than-average bonds, with bonds contributing to the CB being longer on average than those contributing to the VB.

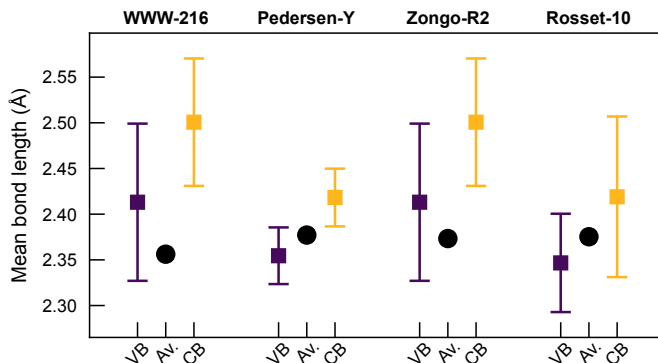


Figure S8: Mean bond length for atoms with a high Inverse Participation Ratio (IPR) at the edge of the valence (VB) and conduction bands (CB) of the four structural models. The mean overall bond length in the model is shown in black. Whiskers show the standard deviation.

S2.4 Chemical bonding

In Fig. S9(a), we show the results of a crystal orbital Hamilton population (COHP) analysis to measure the bond strength, based on orbital interactions. S25S26 The $-COHP$ bonding fingerprints are very similar across all four models, with a fully bonding contribution below E_F (indicated by positive values in the plot), and a fully antibonding contribution (negative values) above E_F .

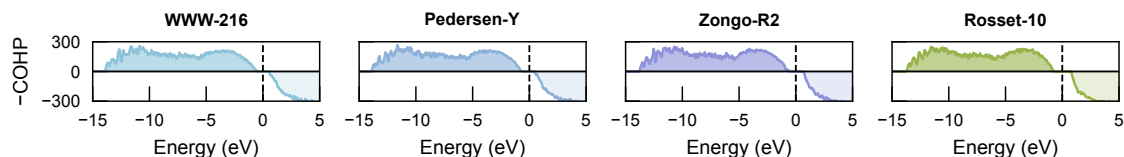


Figure S9: Negative crystal-orbital Hamilton population ($-COHP$) for all four structural models, where bonding and antibonding regions are indicated with a positive ‘+’ and negative ‘-’ sign, respectively, on the y-axis. A dashed vertical line indicates the Fermi level (E_F), defined as being at the middle of the bandgap.

We plot ICOBI against ICOHP values for each bond in Fig. S10, showing the relationship between the bond order (ICOBI) and the bond covalency ($-ICOHP$, in eV). The bond order

increases with higher bond covalency in a linear fashion, excluding the bonds arising from coordination defects in the ‘Pedersen-Y’ model.

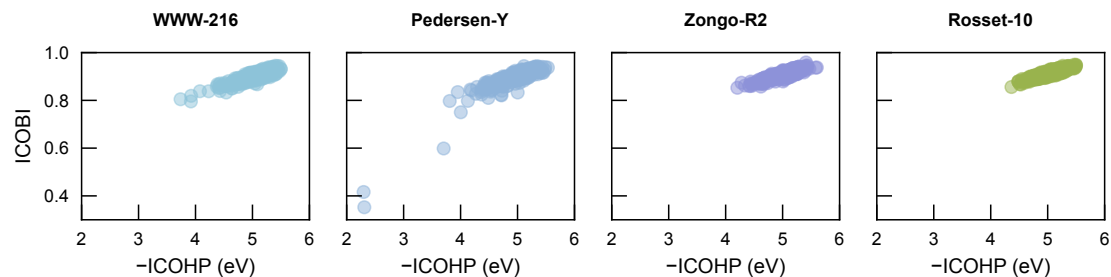


Figure S10: Scatter plots of the integrated crystal orbital bond index (ICOBI) against negative integrated crystal orbital Hamilton population ($-ICOHP$) for all four structural models.

Supplemental references

- S1. L. A. M. Rosset, D. A. Drabold, and V. L. Deringer, [Nat. Commun. **16**, 2360 \(2025\)](#).
- S2. G. Kresse and J. Hafner, [Phys. Rev. B **49**, 14251 \(1994\)](#).
- S3. G. Kresse and J. Furthmüller, [Comput. Mater. Sci. **6**, 15 \(1996\)](#).
- S4. P. E. Blöchl, [Phys. Rev. B **50**, 17953 \(1994\)](#).
- S5. G. Kresse and J. Furthmüller, [Phys. Rev. B **54**, 11169 \(1996\)](#).
- S6. G. K. M. Thutupalli and S. G. Tomlin, [J. Phys. C: Solid State Phys. **10**, 467 \(1977\)](#).
- S7. J. R. A. Carlsson, X. Li, S. F. Gong, and H. T. G. Hentzell, [J. Appl. Phys. **74**, 891 \(1993\)](#).
- S8. C.-C. Huang, Y. Tang, M. van der Laan, J. van de Groep, A. F. Koenderink, and K. Dohnalová, [ACS Appl. Nano Mater. **4**, 288 \(2021\)](#).
- S9. R. H. Klazes, M. H. L. M. van den Broek, J. Bezemer, and S. Radelaar, [Philos. Mag. B **45**, 377 \(1982\)](#).
- S10. A. J. M. Berntsen, W. F. Van Der Weg, P. A. Stolk, and F. W. Saris, [Phys. Rev. B **48**, 14656 \(1993\)](#).
- S11. N.-M. Park, C.-J. Choi, T.-Y. Seong, and S.-J. Park, [Phys. Rev. Lett. **86**, 1355 \(2001\)](#).
- S12. A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, [Phys. Rev. X **8**, 041048 \(2018\)](#).
- S13. V. L. Deringer, N. Bernstein, A. P. Bartók, M. J. Cliffe, R. N. Kerber, L. E. Marbella, C. P. Grey, S. R. Elliott, and G. Csányi, [J. Phys. Chem. Lett. **9**, 2879 \(2018\)](#).
- S14. V. L. Deringer, N. Bernstein, G. Csányi, C. Ben Mahmoud, M. Ceriotti, M. Wilson, D. A. Drabold, and S. R. Elliott, [Nature **589**, 59 \(2021\)](#).
- S15. K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J. L. Robertson, J. Wang, and S. C. Moss, [Phys. Rev. B **60**, 13520 \(1999\)](#).
- S16. J. Fortner and J. S. Lannin, [Phys. Rev. B **39**, 5527 \(1989\)](#).
- S17. K. Laaziri, S. Kycia, S. Roorda, M. Chicoine, J. L. Robertson, J. Wang, and S. C. Moss, [Phys. Rev. Lett. **82**, 3460 \(1999\)](#).
- S18. S. M. Nakhmanson, P. M. Voyles, N. Mousseau, G. T. Barkema, and D. A. Drabold, [Phys. Rev. B **63**, 235207 \(2001\)](#).
- S19. P. Grigorev, L. Frérot, F. Birks, A. Gola, J. Golebiowski, J. Griebner, J. L. Hörmann, A. Klemenž, G. Moras, W. G. Nöhring, J. A. Oldenstaedt, P. Patel, T. Reichenbach, T. Rocke, L. Shenoy, M. Walter, S. Wengert, L. Zhang, J. R. Kermode, and L. Pastewka, [J. Open Source Softw. **9**, 5668 \(2024\)](#).

- S20. A. Stukowski, [Model. Simul. Mater. Sci. Eng. **18**, 015012 \(2010\)](#).
- S21. N. Bernstein, B. Bhattarai, G. Csányi, D. A. Drabold, S. R. Elliott, and V. L. Deringer, [Angew. Chem. Int. Ed. **58**, 7057 \(2019\)](#).
- S22. P. A. Fedders, D. A. Drabold, and S. Nakhmanson, [Phys. Rev. B **58**, 15624 \(1998\)](#).
- S23. Y. Pan, F. Inam, M. Zhang, and D. A. Drabold, [Phys. Rev. Lett. **100**, 206403 \(2008\)](#).
- S24. D. A. Drabold, Y. Li, B. Cai, and M. Zhang, [Phys. Rev. B **83**, 045201 \(2011\)](#).
- S25. R. Dronskowski and P. E. Blöchl, [J. Phys. Chem. **97**, 8617 \(1993\)](#).
- S26. V. L. Deringer, A. L. Tchougréeff, and R. Dronskowski, [J. Phys. Chem. A **115**, 5461 \(2011\)](#).