Structural origin of the midgap electronic states and the Urbach tail in pnictogen-chalcogenide glasses

Alexey Lukyanov
Department of Chemistry, University of Houston, Houston, TX 77204-5003

Jon C. Golden and Vassiliy Lubchenko*
Department of Chemistry, University of Houston, Houston, TX 77204-5003 and
Department of Physics, University of Houston, Houston, TX 77204-5005
(Dated: June 1, 2018)

We determine the electronic density of states for computationally-generated bulk samples of amorphous chalcogenide alloys As_xSe_{100-x}. The samples were generated using a structure-building algorithm reported recently by us (J. Chem. Phys. 147, 114505). Several key features of the calculated density of states are in good agreement with experiment: The trend of the mobility gap with arsenic content is reproduced. The sample-to-sample variation in the energies of states near the mobility gap is quantitatively consistent with the width of the Urbach tail in the optical edge observed in experiment. Most importantly, our samples consistently exhibit very deep-lying midgap electronic states that are delocalized significantly more than what would be expected for a deep impurity or defect state; the delocalization is highly anisotropic. These properties are consistent with those of the topological midgap electronic states that have been proposed by Zhugayevych and Lubchenko as an explanation for several puzzling opto-electronic anomalies observed in the chalcogenides, including light-induced midgap absorption and ESR signal, and anomalous photoluminescence. In a complement to the traditional view of the Urbach states as a generic consequence of disorder in atomic positions, the present results suggest these states can be also thought of as intimate pairs of topological midgap states that cannot recombine because of disorder. Finally, samples with an odd number of electrons exhibit neutral, spin 1/2 midgap states as well as polaron-like configurations that consist of a charge carrier bound to an intimate pair of midgap states; the polaron’s identity—electron or hole—depends on the preparation protocol of the sample.

I. INTRODUCTION

In contrast with their periodic counterparts, amorphous materials are expressly non-Bloch solids. Despite this complication, many amorphous semiconductors purvey electricity similarly to the corresponding crystals. Indeed, from the viewpoint of the wave-packet representing a charge carrier, the pertinent molecular orbitals are virtually indistinguishable from true, infinitely-extended Bloch states as long as the mean-free path of the carrier is less than the extent of the orbitals. Thus one may still speak of mobility bands even in the absence of strict periodicity.1–4 In addition, many families of crystalline and amorphous compounds alike are expected to exhibit conduction by strongly localized, “polaronic” charge carriers, when the electron-lattice interaction is sufficiently strong.5,6 These ideas are graphically summarized in Fig. 1.

Because of the electron-lattice coupling, on the one hand, and the constant thermal motion of atoms, on the other hand, the optical edge in semiconductors is not sharp whether the solid is periodic or not: This is because optical excitations are much faster than nuclear motions implying that, effectively, electrons are always subject to an aperiodic Born-Oppenheimer potential. The effective band edge turns out to be nearly exponential and is often called the Urbach tail.7–16 Now in amorphous materials, there is no underlying long-range order even with regard to vibrationally-averaged atomic positions. Thus the disorder in the atomic locations is partially frozen-in. Because such frozen glasses can be very far away from equilibrium, the distribution of the energies of the localized states is generally decoupled from the ambient temperature (and pressure) and, furthermore, depends on the preparation protocol of the sample.17,18 Appropriately, exponential tails of localized states generally emerge in models with quenched disorder, within the venerable Anderson-Mott framework of electron localization in disordered media.1–4 Those approaches assume generic forms for the random one-electron potential. Non-withstanding the seemingly general character of the resulting predictions, such generic approaches are not fully constructive in that the random potential they postulate may not be consistent with the actual molecular field in a stable structure. In a constructive treatment, such a potential must arise self-consistently.

An additional, seemingly separate set of electronic excitations have been observed in glassy chalcogenides. These excitations are apparently activated by exposing the sample to macroscopic quantities of photons at supragap frequencies. The excitations reveal themselves as optical absorption at midgap frequencies and concomitant emergence of ESR signal,21–25 and anomalous photoluminescence.26–28 The concentration of these seemingly intrinsic defect-like states is estimated at ~10^{20} cm^{-3},21 i.e. one per several hundred atoms. This is much greater, for instance, than the typical amount of dopants in crystalline semiconductors, and leads to a very effi-
One-electron excitations

Because of put forth specific microscopic and a minor de-
argued that given a certain, large number of
Thus doping the material have produced candidate defected configura-
(ZL) have argued that chalcogenide
proposed candidate malco-
The strained configurations can
the midgap states can
dangling bonds, nearby
pairs
in a frozen glass.
in an equilibrated glassy liquid, as well as for aging
effective negative Hubbard
Drabold
transition-state configurations for activated transport
stance, from lattice polarization and would amount to an
perfect crystalline lattice. In a more recent effort, Li and
actively small.) The effective attraction could stem, for in-
orbital. (The subset of such special orbitals would be rel-
valence alternation pairs (VAP). In addition, Van-
glasses must host special midgap electronic states that
midgap states. The robustness of the ZL midgap states
has a topological aspect, too: They can be thought of as stemming from an extra or missing bond on an
atom while the malcoordination cannot be removed by
elastically deforming the lattice. At most, lattice relax-
also Refs. 5, 19, 20 One-electron excitations occur between two bands, separated by a mobility gap. Lo-
ocalized states exist in the gap. The topological states, whose
abundance is distributed throughout the gap, absorb
light at sub-gap frequencies when singly-occupied. (b) The
hypothetical density of states of itinerant electrons and holes
that have relaxed to form small polarons, after Ref. 5

According to an early microscopic proposal, due to
Anderson\textsuperscript{19,31}, such behavior would be observed, if
electrons exhibited effective mutual attraction when occup-
ying a localized orbital so that per electron, the energy
of a filled orbital is lower than that of a singly-occupied
orbital. (The subset of such special orbitals would be rel-
atively small.) The effective attraction could stem, for in-
stance, from lattice polarization and would amount to an
effective negative Hubbard $U$. Shortly thereafter, Street
and Mott\textsuperscript{32} argued that given a certain, large number of
dangling bonds, nearby pairs of such bonds will be un-
stable toward the formation of intimate pairs of malcoor-
dinated configurations, one negatively and one positively
charged. (The dangling bonds themselves are electrically
neutral.) Kastner et al.,\textsuperscript{33} put forth specific microscopic proposals as to the specific atomic motifs that could host
such valence alternation pairs (VAP). In addition, Van-
derbilt and Joannopoulos\textsuperscript{34} proposed candidate malcoor-
dinated configurations that do not have to come in
pairs but are standalone. In these approaches, the defect
states are viewed as essentially defects in an otherwise
perfect crystalline lattice. In a more recent effort, Li and
Drabold\textsuperscript{35} have produced candidate defecte d configura-
tions by modeling light-induced formation of relatively
malcoordinated motifs in computer generated aperiodic
samples.

In a distinct approach, Zhugayevych and
Lubchenko\textsuperscript{20,36} (ZL) have argued that chalcogenide
glasses must host special midgap electronic states that
would be intrinsic to any glass exhibiting spatially-
inhomogeneous bond saturation. (In the case of the
chalcogenides, the bond strength varies between that of
a canonical single bond and a formally closed-shell,\textsuperscript{37}
secondary bond.) These midgap states are tied to
relatively strained regions that are intimately related to
transition-state configurations for activated transport
in an equilibrated glassy liquid, as well as for aging
in a frozen glass.\textsuperscript{17,18} The strained configurations can
be thought of as domain walls separating distinct
aperiodic minima of the free energy; they must be
present in thermodynamic quantities. The equilibrium
concentration of the domain walls just above the glass
transition has been estimated at $\sim 10^{20} \text{cm}^{-3}$, using
the random first order transition (RFOT) theory,\textsuperscript{38–41}
this figure matches well the apparent quantity of the
light-activated midgap states. Since the structure of
the liquid becomes largely arrested below the glass
transition—apart from some aging\textsuperscript{17} and a minor de-
crease in the vibrational amplitude—the concentration of the strained regions remains steady upon cooling
the glass. When expressed in terms of the size of a rigid
molecular unit, this concentration is nearly universal
for two reasons: On the one hand, it depends only
logarithmically on the time scale of the glass transition.
On the other hand, the dependence of the concentration of
the midgap states on the material constants is only
through the so called Lindemann ratio. This ratio is
defined as the relative vibrational displacement near
the onset of activated transport and happens to be a
nearly universal quantity.\textsuperscript{42–43} Thus doping the material
may shift the glass transition temperature, but will not
significantly affect the concentration of the intrinsic
midgap states. The robustness of the ZL midgap states
has a topological aspect, too: They can be thought of
as stemming from an extra or missing bond on an
atom while the malcoordination cannot be removed by
elastically deforming the lattice. At most, lattice relax-
ation results in the malcoordination being “smeared”
over a substantial region. This smearing is accompanied
by further delocalization of an already surprisingly
extended wave-function of the midgap electronic state.
The overall extent of the wavefunction can be in excess
of a dozen lattice spacings, which is much greater than
what one would expect for a very deep impurity state.

Analogous in many ways to the solitonic midgap
states in trans-polyacetylene,\textsuperscript{44} the midgap states
can be thought of as composed of an equal measure of the
states from the valence and conduction band. As a re-
sult, their energy is efficiently pinned near the center
of the mobility gap, if the states are singly-occupied and,
thus, neutral. Such neutral states absorb light at sub-gap
frequencies. However in a pristine sample, the midgap
states are occupied or vacant—corresponding with being
negatively and positively charged respectively—and thus
stabilized owing to lattice polarization.\textsuperscript{20,36} Because of
the lattice distortion, charged midgap states absorb at
supra-gap frequencies.\textsuperscript{19,31,36} These notions underlie
the light-induced emergence of ESR signal and midgap
absorption.\textsuperscript{36} Supra-gap irradiation excites electrons from
filled midgap states into the conduction band; in addition
filled midgap states can capture the oppositely charged
free carriers produced by the irradiation. As a result, the
midgap states become electrically neutral and begin to
absorb light at sub-gap frequencies. Thus in the ZL sce-

cenario, the specific atomic motifs giving rise to the midgap
states are not light generated defects; instead they are in-
trinsically present in the structure. Light only serves to

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1.png}
\caption{(a) Hypothetical electronic structure of amorphous chalcogenides, see also Refs. 5, 19, 20 One-electron excitations occur between two bands, separated by a mobility gap. Localized states exist in the gap. The topological states, whose one-particle energy is distributed throughout the gap, absorb light at sub-gap frequencies when singly-occupied. (b) The hypothetical density of states of itinerant electrons and holes that have relaxed to form small polarons, after Ref. 5.}
\end{figure}
make the defects ESR- and optically-active—at sub-gap frequencies—by making them half-filled.

Some of these aspects of the topological midgap states are reminiscent of the negative-$U$ model of Anderson\textsuperscript{19,31} and subsequent ad hoc defect-based models mentioned above.\textsuperscript{32,34,45} In fact, some of those defect states can be viewed as an ultra-local limit of the ZL theory.\textsuperscript{36} In those earlier developments, however, the concentration of the defects is tied to the number of specific molecular motifs whose quantity would seem to depend on the stoichiometry. Nor is it clear whether such defected configurations could combine to form a stable lattice. In the ZL treatment, the concentration is predicted to be inherently at roughly one defect per several hundred atoms, irrespective of the precise stoichiometry. The defect concentration is determined by an interplay between the enthalpic cost of forming the domain walls, on the one hand, and their entropic advantage, on the other hand. This entropic advantage can be ultimately be traced to the excess liquid entropy of the supercooled liquid relative to the corresponding crystal.

ZL\textsuperscript{36} have also proposed specific malcoordinated motifs underly the midgap states. The motifs consist of relatively extended $p$-$p$-bonded chains connecting an odd number of orbitals, at half filling. The presence of such motifs is expected on a systematic basis, insofar as one may regard the chalcogenides as distorted versions of relatively symmetric, parent structures defined on the simple cubic lattice.\textsuperscript{46} Parent structures defined on the simple cubic lattice have been indeed obtained for all known lattice types found in stoichiometric crystalline compounds of the type $\text{P}n_2\text{Ch}_3$ ($\text{P}n = \text{P, As, Sb, Bi, Ch} = \text{S, Se, Te}$),\textsuperscript{47} and are themselves periodic of course. ZL anticipated that aperiodic parent structures, upon deformation, would give rise to bona fide glassy structures. Thus the aperiodicity and the sporadic malcoordination of the lattice would arise self-consistently. ZL\textsuperscript{36} have generated standalone, molecular malcoordinated motifs, which were properly passivated to emulate proper coordination appropriate for a 3D solid. It was argued that the interaction of defect-bearing chains with the surrounding solid amounts to a renormalization of the on-site energies and electron hopping elements along the chain but would not change the physics qualitatively—a notion that will be revised in this article. No attempt was made to generate a bulk three-dimensional structure that would host such motifs while obeying local chemistry throughout.

Here we report what we believe is the first realization of the topological midgap states in bulk samples. The samples were generated using a structure-building algorithm reported earlier by us.\textsuperscript{48} In this algorithm, one first generates so called parent aperiodic structures that exhibit octahedral coordination, locally. A subset of the lattice sites in the parent structure form a random-close packed structure,\textsuperscript{49} thus ensuring that within a certain range of wavelengths, the structures are at what is believed to be the highest achievable density for aperiodic solids. These parent structures are subsequently optimized using quantum-chemical force fields while coordination becomes distorted-octahedral.\textsuperscript{46} Thus the sample is not generated by quenching a melt that had been equilibrated at some high temperature. In principle, such quenched samples should approximate real materials subject only to the accuracy of the effective inter-atomic force fields and, possibly, finite-size effects. In practice, however, the dynamical range of molecular dynamics simulations is very limited even for relatively simple model liquids let alone the chalcogenides, where the effective force fields between the atoms must be determined using computationally expensive, quantum-chemical approximations. As a result, one can equilibrate a chalcogenide melt only at temperatures much exceeding the laboratory glass transition. The resulting samples are thus hyper-quenched while their structure exhibits exaggerated effects of mixing entropy in the form of excess homo-nuclear contacts even in the stoichiometric compound $\text{As}_{40}\text{Se}_{60}$, in contrast with observation.\textsuperscript{50}

The parent structures generated in Ref.\textsuperscript{48} do not contain homonuclear contacts by construction, nor do such defects seem to appear in significant quantities following the geometric optimization, at least in the stoichiometric compound $\text{As}_{40}\text{Se}_{60}$. We have argued\textsuperscript{48} this circumstance explains why the resulting amorphous samples consistently exhibit the first sharp diffraction peak (FSDP) in the structure factor, including its trends with pressure and arsenic content. (The FSDP is the hallmark of the poorly understood medium-range-order\textsuperscript{51,52} in inorganic glasses.) In contrast, samples generated using first principles molecular dynamics often fail to exhibit the FSDP.\textsuperscript{53}

Likewise, the electronic density of states (DOS) for the presently generated samples exhibit properties expected of the amorphous chalcogenides. On the one hand, the gross part of the electronic spectrum does not vary significantly from sample to sample; this part of the DOS thus can be attributed to the mobility bands. On the other hand, the states near the band edges are found to fluctuate in energy substantially, as is expected for the localized Urbach-tail states. In fact, the magnitude of the fluctuation matches well the width of the Urbach tail observed in experiment. Perhaps more interestingly, samples generated according to the procedure from Ref.\textsuperscript{48} do indeed exhibit very deep-lying midgap states that are close to the gap center, closer than could possibly take place for an Urbach-tail state.

Both the presence of the midgap states in the computed spectra and the value of the gap itself are found to depend rather sensitively on the detailed quantum-chemical approximation, consistent with earlier studies requiring the relatively higher-end, computationally expensive hybrid DFT to solve for the electronic spectrum. (Plain DFT seems to suffice for structure optimization.) Consistent with the general predictions by ZL,\textsuperscript{20} the wavefunctions of a subset of the midgap states are elongated preferentially in one direction, in contrast with the extended states comprising the mobility bands,
which are isotropic. In addition, here we argue that a somewhat distinct physical possibility can be realized in which the wave function consists of a few linear fragments emanating from the same spot in space; this possibility was overlooked by ZL. Some of the samples we have generated appear to exhibit this urchin-like shape. At least in one case, we were able to clearly identify a motif that can be thought of as a passivated odd-numbered chain that must necessarily host a topological midgap state. Such states correspond to ZL chains with closed ends. Now, the aforementioned deep midgap states were unforced in the sense that the structures contained an even number of electrons and were geometrically optimized so that there are no intentionally broken bonds in the sample. By way of contrast, we have also generated samples containing an odd number of electrons, thus forcing the system to have a dangling bond. Here we find the resulting “defect” is often very deep in the forbidden gap. This is in contrast with what would be expected for a generic impurity state. In some cases, the defect states are so deep inside the gap that they can be thought of as electrically-neutral entities, where the electron charge is largely compensated by lattice polarization. Yet in some cases, the configurations are more reminiscent of the polarons in conjugated polymers.

Finally, we have established that the Urbach states turn out to be intermediate—both in terms of their shape and degree of localization—between the extended band states and the topological states thus suggesting the Urbach states and the topological states are intimately related.

The article is organized as follows: Section II discusses the salient features of the presently obtained spectra, the forbidden gap and the Urbach tail states. In Section III we first briefly survey pertinent general properties of the topological midgap states for isolated chains and then use model calculations to demonstrate that such midgap states are robust even if the defected chain are coupled to other chains. In Section IV, we quantitatively characterize the very deep midgap states obtained in the present study and argue that they are, in fact, the topological midgap states predicted earlier by ZL. Section V provides a brief summary.

II. SALIENT FEATURES OF THE DENSITY OF STATES: MOBILITY BANDS AND THE URBACH TAIL OF LOCALIZED STATES

We have generated disordered structures for the binary compound arsenic selenide As$_x$Se$_{100-x}$ for three distinct stoichiometries $x = 20$, 40, and 50, with the help of the algorithm described in Ref. 48. In this algorithm, we first generate a chemically-motivated aperiodic parent structure, in which a subset of atoms are placed at the vertices of a random close-packed (RCP) lattice, called the “primary lattice.” A complementary subset of atoms is then placed at the vertices of the so called “secondary lattice,” which is generated according to a detailed algorithm to maximize octahedrality in local bonding while maintaining desired stoichiometry. This detailed algorithm uses a threshold parameter $A$, which prescribes the precise way in which we break up the interstitial space of the primary lattice into cavities. For instance, for very small values of $A$, the cavities will be all tetrahedral, the tetrahedra determined using the Delaunay triangulation. For sufficiently large values of $A$, on the other hand, some of these tetrahedra are merged into higher order polyhedra. These higher order polyhedra host the vertices of the secondary lattice. To minimize the number of homonuclear bonds, the primary and secondary lattice are populated by distinct species. Here we focus exclusively on the so called C-type structures, for which the primary lattice is made of Chalcogens (selenium in this case).

By varying the value of the parameter $A$, one can effectively control the amount of vacancies in the parent structures. This amount is equal to the number of non-tetrahedral cavities in excess of the number of pnictogens (for C-type structures). For instance, $A = 0.1$ is convenient for the stoichiometric compound As$_{40}$Se$_{60}$ because the number of vacancies that must be made is small yet not too small so that one can generate sufficiently distinct structures by randomly choosing the locations for the vacancies. In addition to $A = 0.1$, we have also generated samples with $A = 0.014$ and 0.24 for the As$_{20}$Se$_{80}$ and As$_{50}$Se$_{50}$ compounds, respectively. Five samples were generated for each value of $A$, for each compound, except for As$_{60}$Se$_{50}$, where we have generated eleven. Finally, the parent structures are geometrically optimized using plane-wave DFT as implemented in the package VASP in the Perdew-Wang (PW91) generalized gradient approximation (GGA) for the exchange-correlation energy. Optimization was performed both at constant pressure and volume; the corresponding structures are labelled const-$P$ and const-$V$, respectively. We note that optimization at constant pressure, which thus allows for unit cell optimization, produces much better results for the first sharp diffraction peak. Throughout

![FIG. 2. Electronic density of states for five distinct amorphous samples of the compound As$_{20}$Se$_{80}$ and two distinct values of the parameter $A$, optimized at constant pressure and volume, respectively. The vertical dashed line denotes the Fermi energy.](image-url)
All methodologies seem to reproduce well the density of states is truly continuous where the density of states is sufficiently away from the gap, where the density of states is truly continuous and self-averaging. It is this averaged spectrum that is pertinent to experimentally observed absorption spectra because absorption experiments represent a bulk measurement. The electronic density of states averaged over several samples are shown in Fig. 3, along with experimental data, for all three stoichiometries. Clearly, the present results reproduce the gross features found in actual spectra, particularly in the higher energy part of the valence band. The agreement is especially notable for the 20-80 mixture.

To infer the atomic-orbital makeup of the electronic density of states, we consider the s and p atomic-orbital contributions to the total density of states as determined using VASP’s built-in projection technique, see Fig. 4. According to the latter Figure, the higher energy portion of the valence band is largely due to the p-orbitals. We note that the present spectra for As_{40}Se_{60} and As_{50}Se_{50} seem to overestimate the “dip” near -2 eV, which separates the bonding p-orbitals from the selenium-based lone pairs. According to Fig. 4, this may stem from a (modest) underestimation of the density of states corresponding to the As-based p-orbitals. The overall good agreement between theory and experiment in this part of the spectrum is reassuring since the bonding is indeed expected to be largely due to the p-orbitals. The corresponding figures for the compounds As_{30}Se_{70} and As_{50}Se_{50} can be found in the Supplemental Material. There, we also provide an animation that allows one to efficiently survey how the shape and extent of the wave function is correlated with the corresponding energy eigenvalue. The aforementioned sample-to-sample variation is indeed expected for disordered samples. The variation is expected to be relatively small for electronic states sufficiently away from the gap.

The electronic spectra, as illustrated for several As_{20}Se_{80} samples in Fig. 2, are overall similar yet clearly vary in detail from sample to sample, and especially so around the forbidden gap. The corresponding figures for the compounds As_{30}Se_{70} and As_{50}Se_{50} can be found in the Supplemental Material. There, we also provide an animation that allows one to efficiently survey how the shape and extent of the wave function is correlated with the corresponding energy eigenvalue. The aforementioned sample-to-sample variation is indeed expected for disordered samples. The variation is expected to be relatively small for electronic states sufficiently away from the gap. In the Supplemental Material, we compare the present results with the electronic spectra obtained in two earlier studies, due to Bauchy et al. and Li et al. All methodologies seem to reproduce well the p-
portion of the spectrum, Li et al.’s results for As$_{50}$Se$_{50}$ standing out. There is significantly less agreement with experiment in the s-orbital portion.

In contrast with the states comprising the mobility bands, states near the gap are expected to be relatively localized; their quantity and energy should strongly depend on the specific realization. Only upon averaging over many realizations, these localized states should yield a relatively smooth spectrum, which is expected to be exponential. This is indeed born out by the present data. In view of limited statistics and modest sample sizes, however, it is difficult to determine, based on the spectra alone, which states comprising the individual spectra in Fig. 2 should be assigned to the mobility band and which to the Urbach tails of localized states or, potentially, to the topological midgap states predicted by ZL. This complication makes determination of the width of the forbidden gaps in Fig. 2 ambiguous. In experiment, one conventionally places the edges of the mobility bands at energies corresponding to the onset of the Urbach tail, the latter located by fitting, see Figs. 1 and 2 of Ref. 69. The functional form used to fit the Urbach-tail states is:

\[ n(\epsilon) \propto e^{-\epsilon/E_U}, \]

where \( \epsilon \) stands for the depth of a localized state relative to the edge of the corresponding mobility band, into the gap.

According to Ref. 69, the onset of the exponential tail of the localized states corresponds, by convention, to roughly 1% of the absorption strength characteristic of transitions between bona fide extended states. The full range of absorption strength available in experiment spans a range of at least five orders of magnitude thus permitting one to identify the edge states relatively unambiguously. In contrast, the present DOS barely spans two orders of magnitude. To work around this complication, we employ two distinct methods to assess the gap. In method one, we fit the DOS, near but not exactly at the gap edge, using the functional form \((E_{\nu} - E)_{1/2}^{1/2}\) and \((E - E_{c})_{1/2}^{1/2}\) for the valence and conduction band, respectively. The square root scaling would be exact for a translationally invariant system in 3D, of course. An example of the resulting fit is shown in one of the panels in Fig. 5. (As a practical matter, we obtain the fits by plotting the square of the DOS.) The resulting values of the gap, \(E_g = E_{c} - E_{\nu}\), are given as \(E_g^{\exp}\) in Table I. We note the apparent deviation of the computed DOS from the

| const | \(E_{\nu}^{\exp}\) | \(E_{c}^{\exp}\) | \(\sigma_U\) |
|-------|----------------|----------------|----------|
| \(A_{50}S_{50}\) | 2.00, 1.91 | 1.84$_{70}^{70}$, 1.84$_{71}^{71}$ | 0.1 |
| \(A_{50}S_{50}\) | 1.94, 1.78 | 1.74$_{72}^{73}$, 1.64$_{74}^{74}$ | 0.07 |
| \(A_{50}S_{50}\) | 1.94, 1.60 | 1.74$_{75}^{76}$, 1.72$_{77}^{78}$, 1.64$_{79}^{80}$ | 0.27 |

TABLE I. Band gap values \(E_g\) estimated using two different approaches for samples optimized at constant pressure (const-P) and volume (const-V) at three distinct stoichiometries. The corresponding experimental values are also provided. The quantity \(\sigma_U\) yields the steady state value of the standard deviation for the level position and is used to estimate the width \(E_U\) of the Urbach tail, see text.
simple square-root scaling is consistent with the expectation that the presently generated disordered samples should host localized states.

In an alternative strategy, we do not assume any specific functional form for the spectrum either within mobility bands or for the localized states. Instead, we first order the occupied states according to their distance, energy-wise, from the HOMO for each individual sample. Likewise we order, for each sample, the vacant levels according to their separation from the LUMO. Next, we determine the average energy \( E_{v}^{(n)} \) of the \( n \)-th occupied state and the average energy \( E_{c}^{(n)} \) of the \( n \)-th empty state. Call the corresponding standard deviations \( \sigma_{v}^{(n)} \) and \( \sigma_{c}^{(n)} \), respectively. We graph these average energies and the corresponding standard deviations in Fig. 6 for levels 1 through 20. Data for the other stoichiometries, both at constant pressure and volume, can be found in the Supplemental Material. Although the detailed trends of \( \sigma_{v} \) vary somewhat depending on the stoichiometry and the optimization conditions—i.e. const-P vs. const-V—the following two features are securely reproduced in all cases: (a) The energies \( E_{v}^{(n)} (E_{c}^{(n)}) \) decrease (increase) with \( n \). (b) The standard deviations \( \sigma_{v}^{(n)} \) and \( \sigma_{c}^{(n)} \) saturate at large \( n \) at some value \( \sigma_{U} \). We also note that \( \sigma_{v}^{(1)} \) is always relatively small, the likely reason being is that the Fermi level is tied to the HOMO.

Already the energies of the next-to-HOMO and LUMO levels, \( n = 2 \), exhibit the amount of variation comparable to the steady value of \( \sigma_{U} \). At the same time, even if it so happens that these \( n = 2 \) levels belong to the mobility band, they would not be too deep into the band. In addition, there are good reasons to believe the HOMO and LUMO themselves can be sufficiently often associated with the topological midgap states, as we shall see in Section IV. Based on these notions and for the sake of concreteness, we settle on quantifying the band gap using the energy difference between the \( n = 2 \) states:

\[
E_{g}^{11} \equiv E_{c}^{(2)} - E_{v}^{(2)},
\]

see Table I. Given the scatter in the reported experimental values of the gap, the present estimates seem rather satisfactory. Significantly, we observe that the samples optimized at constant pressure reproduce the experimental trend that the gap decreases modestly with arsenic content. (We note that some older data suggest that the gap width, as a function of arsenic content, experiences a shallow “dip” around As\(_{45}Se_{55}\).) In contrast, the const-V samples do not follow this trend. Interestingly, samples optimized at constant pressure also did much better\(^{48}\) with regard to the first sharp diffraction peak (FSDP), which is a structural feature. In contrast, const-V samples often exhibited a small shoulder instead of a well-defined peak. That the detailed characteristics of the band edges would have a structural signature is consistent with the apparent correlation between the strength of the FSDP, pressure, and the phenomenon of photodarkening. (The term photodarkening\(^{79}\) refers to a light-induced narrowing of the optical gap.) This correlation was discussed in detail in Ref. 48.

We next turn to the standard variation \( \sigma_{U} \). First we note that for the exponential distribution \( p(x) = \sigma^{-1} e^{-|x|/\sigma} \), the standard deviation is equal to \( \sigma \). The number of samples we have generated is too modest to conclusively make out the shape of the distribution of the tail states, as already mentioned. Yet insofar the sampling can be regarded as sufficient to infer the standard deviation, we may associate the quantity \( \sigma_{U} \) with the width parameter \( E_{U} \) of the Urbach tail from Eq. (1). The thus estimated value for the width of the Urbach tail for the stoichiometric compound As\(_{40}Se_{60}\), viz. 0.07-0.08 eV, is consistent with the experimentally reported value of 0.066 eV.\(^{80}\) We were unable to find the Urbach energy \( E_{U} \) in the literature for the other two stoichiometries.

We now turn our attention to the features deep in the mobility gap that are sporadically found in the density of states for specific realizations; these are exemplified in Fig. 5. Note that such sporadic features readily become part of the background upon averaging, as in Fig. 3. While our statistics are clearly limited, examination of the available spectra for individual realizations, Fig. 2 and Figs. S3 and S4 in the Supplemental Material, suggest that roughly only one in ten samples display

![Figure 6](image-url)
such deep midgap states already in the stoichiometric compound As$_{40}$Se$_{60}$, the number of incidents seemingly increasing away from the exact 2:3 stoichiometry. We immediately point out that on purely statistical grounds, observing a midgap state like that in Fig. 14 is not at all likely. Indeed, given that the width of the Urbach tail is numerically close to 0.08 eV, our chances of observing that midgap state are roughly one in $e^{0.08eV/0.08eV}$ in view of Eq. (1). That is, roughly one per thousand samples.

Note that the defect-like midgap states in Fig. 2 and 5 are unforced in that the samples contain an even number of electrons and are fully geometrically optimized; no broken bonds are deliberately put in the system. Conversely we note that when occupied, such deep states are very costly—at up to $(E_g/2) \times 2$ per state, viz. ca. 2 eV, and thus would be presumably stabilized during geometric optimization unless prevented from doing so for some special reason. Zhugayevych and Lubchenko (ZL) have argued exactly such a reason exists in glassy chalcogenides. The reader is referred to those publications for detailed discussion. In the following Section, we briefly reiterate some of those notions and present new results that will be helpful in interpreting some of the midgap-state features revealed in this study.

III. THE TOPOLOGICAL ELECTRONIC MIDGAP STATES: GENERAL RESULTS

ZL have argued that at some places along the domain walls separating regions occupied by distinct aperiodic free energy minima in a glassy liquid, local coordination will differ from its optimal value. The missing or extra bond should be perpendicular to the domain wall. Specifically in the chalcogenides, quasi-one-dimensional chain-like motifs can be identified based on a structural model, in which these materials are regarded as distorted, symmetry-broken version of more symmetric parent structures locally defined on a simple cubic lattice. In the simplest case, a defect-free distorted chain exhibits an exact alternation pattern of a covalent bond and a secondary interaction and can be thought of as a chain of weakly interacting dimers. At a malcoordination defect, one atom or more will have one too many or one too few bonds. As was understood in the context of conjugated polymers, such malcoordination defects must host very special midgap electronic states, see illustration in Fig. 7. If singly occupied and close to the middle of the gap, they formally correspond with a neutral particle that has spin 1/2, see Fig. 7(a). (The electron charge is exactly compensated by the polarization of the lattice.) The deviation of the bond length from its value in a perfectly dimerized chain shows a solitonic dependence on the coordinate. The strain has a solitonic profile, hence the midgap states are often called “solitonic.”

In the chalcogenides, one expects that most of the midgap states would be either fully occupied or empty: On the one hand, the chalcogenides exhibit spatial variation in the electronegativity. Suppose the variation is on average equal to $\varepsilon$. Then a midgap state based on the more electronegative site is typically lowered, energy-wise, by the amount $\varepsilon/2$ relative to the middle of the gap. Likewise, defects centered on the less electronegative element would be destabilized by the same amount. In addition, one expects that a singly occupied midgap state will be modestly stabilized by binding an electron or hole—consistent with simple electron counting arguments. These notions are summarized in Fig. 7(b). The apparent charge of the defect will depend not only on the occupation of the midgap state but also on its position in the band, see the informal chart in Fig. 8.

A midgap state cannot be removed by elastic deformation but, instead, only via annihilation with a defect of opposite malcoordination, which resides on the same

---

### FIG. 7. Schematic of quasi-linear motifs—the surrounding solid matrix not shown—which host solitonic midgap states. Zigzag-like distortions of the motif, which would be present in actual materials, are not shown for clarity. The dots denote atoms, either chalcogens or pnictogens. The thickest line stands for a single covalent bond, the thinnest line for the secondary interaction, and the line intermediate in thickness for a bond intermediate in strength between those two interactions. (a) Malcoordination leads to the appearance of a midgap state. (b) In the presence of lattice polarization, such states will be stabilized by becoming filled with electrons (holes). (c) When close-by, but not recombined—two pertinent atomic configurations shown schematically—the midgap states form a resonance, the electrons occupying the lower term. Upon complete recombination, the lower and upper term merge with the valence and conduction band, respectively. (d) The configuration from (c) with an added electron or hole corresponds with a polaron.
chain. For this reason, such oppositely malcoordinated configurations are often called soliton and anti-soliton. When the spatial separation between such oppositely malcoordinated defects is small, the respective energy levels form a resonance. The lower level of the resonance becomes occupied, the upper level vacant, as in Fig. 7(c), where we also sketch the corresponding structural motifs. Once the defects fully recombine, the bottom level merges with the valence band and the top level with the conduction band. We note that odd-numbered cyclic ring molecules at half-filling must host at least one defect that in principle cannot be removed. The latter will thus host a midgap state which is singly-occupied. Finally, Fig. 7(d) schematically shows what happens when an electron or hole is added to a chain and the system is allowed to relax to form a polaron. In contrast with a continuum view, in which a polaron can be thought of as a generic impurity-like state, the lattice deformation results in the appearance of at least two midgap states. The resulting configuration can be thought of as a charged soliton in a close proximity with a neutral soliton, or a charge carrier bound to an intimate soliton-antisoliton pair.

A distinct feature of the midgap states is their shape: Since the wavefunctions of the topological states are hosted by chain-like dimensional motifs, they are relatively localized in the direction perpendicular to the motif, but could be substantially delocalized along the motif itself. The delocalization along the chain is surprisingly large given how deep the states are. This is because the localization length is no longer determined by the depth of the state but, instead, by the ratio of the transfer integrals of the strong and weak bond. The amount of delocalization is additionally increased because the malcoordination is smoothly distributed over the chain, typically over ten bond lengths or so, see the Supplementary Material for more detail and Ref. 36 for specific molecular realizations. It was argued in that work that the effect of the solid matrix housing the chain is, largely, to renormalize the on-site energies and electronic transfer integrals along the chain. Here we revisit this proposition by considering two or more chains crossing at one site. One can think of one chain as made, for instance, primarily of $p\sigma$-bonded $p_x$ orbitals and the other chain of $pp\sigma$-bonded $p_y$ orbitals. Exactly one of the chains hosts a malcoordination defect by construction. A graphical illustration of this set-up is shown in Fig. 9, panels (a) and (b) corresponding to over-coordination and undercoordination, respectively. The chains are depicted as parallel to avoid crowding the picture, however they are not parallel in the physical space. The transfer integral between the two orbitals at the intersection between the chains is denoted with $t$. These two orbitals belong to the same atom; the non-zero overlap between the orbitals can come about because the chains are not exactly perpendicular and/or because $sp$-mixing is present. For simplicity, we allow the bond lengths to have only two values. These values would correspond to the covalent and secondary bond in a perfectly dimerized chain; the electronic hopping integrals are $t_1$ and $t_2$ ($t_1 > t_2$) for the covalent and secondary bond, respectively. Also for simplicity, we set all of the on-site energies at the same value $\epsilon = 0$. The results can be straightforwardly generalized for a non-vanishing, sign-alternating on-site energy. The latter situation would be directly relevant to linear chain-like motifs in which chalcogen and pnictogen alternate in sequence. Finally, we set transfer integrals for next-nearest and farther neighbors at zero.

The resulting Hamiltonian is easily numerically diagonalized, the corresponding spectra and the midgap wave-functions shown in Figs. 10 and 11. We observe that in the case of overcoordination, the wave-function is confined to the chain housing the defect, consistent with the analysis of Zhugayevych and Lubchenko. 36,46

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure8}
\caption{The charge of a state depending on its position in the forbidden gap and the filling.}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure9}
\caption{(a) Graphical description of the tight-binding Hamiltonian used here to infer the electronic structure for two coupled chains, one hosting an overcoordinated atom and one perfectly dimerized. The sites connected by the transfer element $t$ are physically located on the same atom, as is signified by the dashed-line frame. On the r.h.s., we show the resulting shape of the wave function, according to the data in Fig. 10. (b) Same as (a), but the defective chain now hosts an under-coordinated atom. In this case, the wavefunction of the midgap state is shared between the chains, see Fig. 11. In the case of two intersecting chains, one expects a T-shaped wavefunction, in which the angle between the two linear fragments is not necessarily 90°. For a greater number of intersecting chains, the shape could be described as urchin-shaped. Only two of the “rays” are expected to be approximately co-linear.}
\end{figure}
FIG. 10. The energy spectrum for the two intersecting chains from Fig. 9(a), the eigenvalues ordered in the ascending order. The lengths of the defected and defect-free chains are 103 and 100, respectively. The midgap state at $E = 0$ is clearly seen, the corresponding wavefunction shown in the Inset, where sites 1 through 103 comprise the defected chain and sites 104 through 203 the defect-free chain. We observe that the midgap state based on overcoordinated atoms stays on the defected chain. The delocalization length is seen to be around twenty sites. Note the defect also gives rise to two bound states outside the bands, wavefunctions not shown, which is shared between the chains.

In contrast, when the midgap state is caused by undercoordination, a substantial part of the the wave function now “spills” into the crossing, defect-free chains. We notice that the wavefunction of the midgap state, on such a defect-free chain, is non-vanishing only on one side of the intersection. We have directly checked that the same conclusions apply when three chains intersect. In cases when the wavefunction of the midgap state is shared between a number of intersecting chains, its shape can be thought of as a set of linear portions emanating from the same lattice site; only two of these portions are expected to be approximately co-linear.

These contrasting behaviors with respect to the wavefunction’s shape are discussed formally in the Supplemental Material, but can be rationalized already using the following, informal line of reasoning, which is graphically illustrated in Fig. 12. In the absence of geometric optimization, which would distribute the malcoordination over an extended portion of the chain, an over-coordinated center can be thought of as a trimer weakly coupled to a perfectly dimerized chain via two weak bonds, as in Fig. 12(a). The bottom and top level of the trimer will be, respectively, stabilized and destabilized as a result of the coupling. The non-bonding state of the trimer, on the other hand, will remain in the middle of the gap by symmetry, since it is coupled equally strongly to both the bottom and top band of the chain. The corresponding wavefunction vanishes at the central site of the trimer, either without or with coupling.

In contrast, an under-coordinated center in a non-relaxed geometry can be thought of as a single orbital at $E = 0$ coupled to a perfectly dimerized chain, as in Fig. 12(b). The wave function of the resulting midgap state will be non-zero and, in fact, will have its largest value at that orbital. When a defected chain is coupled to a perfectly dimerized chain, the electron will be also able to tunnel to some extent into the dimerized chain, but will do so only on one side. This is because a state

FIG. 11. Same as Fig. 10, but for the under-coordinated situation from Fig. 9(b). The defected and defect-free chains contain 101 and 100 sites, respectively. In the inset, sites 1 through 101 comprise the defected chain and sites 102 through 201 the defect-free chain. Here we observe that the midgap state wavefunction is shared between the chains. The wavefunction penetrates the defect-free chain only on one side of the crossing point.

FIG. 12. (a) An informal molecular orbital (MO) diagram illustrating why the midgap state centered on an over-coordinated atom in an odd function with respect to reflection about the central site of the midgap state. The size and color of the circles on the l.h.s. reflect, respectively, the magnitude and sign of the coefficient at the pertinent atomic orbital in the MOs. Clearly, the MO corresponding to the $E = 0$ state is odd and vanishes on the central site. (b) Same as (a), but for an under-coordinated site. Here the midgap state’s wavefunction is even.
Motivated by the notions made in Section III, we next quantify both the extent and compactness of states within the gap and its immediate vicinity. A concrete way to quantify the localization of an orbital $\psi$ is to compute the so called inverse participation ratio (IPR):

$$\text{IPR} = \frac{\int d^3r |\psi(r)|^4}{\left(\int d^3r |\psi(r)|^2\right)^2} \tag{3}$$

where the denominator is included in case the orbital is not normalized to unity. A moment thought shows that the above expression generically scales as the inverse volume of the region occupied by the orbital. Thus the inverse participation ratio gives a volumetric measure of localization of the respective wavefunction. The so evaluated measure of localization is shown with the red dots in Fig. 5. Consistent with expectation, the states comprising the mobility band are delocalized over the whole sample. There may be an ever so slight increase in localization going toward the gap. This increase picks up closer to the edge of mobility band so that by the end of the Urbach tail, the states are seen to occupy a volume that is about one order of magnitude less than the extended states. (This difference depends on the sample size, of course.) We observe that the overall degree of localization of the deep midgap states is greater still than that of the Urbach states, but not dramatically so.

To assess the shape of the wavefunctions, we compute its “tensor of inertia” according to:

$$I_{ij} = \int r_ir_jd\rho, \tag{4}$$

where $\rho$ is the charge density at a point $r$. There is a technical complication in computing the above quantity since the sample is periodically continued, by construction, and so one must make a decision as to the precise location of the repeat unit so that the wavefunction is most connected and, hence, compact. After choosing the optimal location for the centroid of the wavefunction, we compute the corresponding inertia tensor and bring it to a diagonal form. We next sort the resulting principal moments of inertia in ascending order and label them as follows: $I_{xx} \leq I_{yy} \leq I_{zz}$. Since here we are interested only in the shape, not the absolute extent of the wavefunction, we consider only the aspect ratios of $I_{yy}/I_{xx}$ and $I_{zz}/I_{xx}$. According to the ZL predictions, these ratios should significantly exceed unity at least for the midgap states that are confined to one chain. One may also generically expect the aspect ratio to be numerically close to one for the extended states within the mobility bands.

These expectations are well borne out by our data, which are displayed in Fig. 14. We observe that the aspect ratio largely echoes the participation ratio: The deeper inside the gap, the more non-compact the wavefunctions tend to be. Again, the difference between the
in the present study, the wavefunctions of non-extended states are rather anisotropic, and more so the deeper the corresponding state is into the mobility gap. While this is consistent with the general analysis of ZL,\textsuperscript{36} the atomic motifs housing the wavefunctions in the present study are significantly more complex than the standalone molecular liner motifs generated by ZL. In any event, all of the motifs in Fig. 15 contain at least one malcoordinated atom. Specifically, motif (a) contains an under-coordinated arsenic, (b) under-coordinated selenium and over-coordinated arsenic, (c) an under-coordinated arsenic, and (d) two over-coordinated arsenics. Consistent with the discussion in Section III, the midgap state wavefunction stemming from under-coordination in panel (a) appears to branch out, while the defect in (d) contains over-coordinated atoms and is elongated preferentially in one spatial direction. Defect (c) is notable in that the bulk of electron density resides on a 5-member closed ring. As mentioned earlier, such odd-numbered rings must host midgap states.

We note that when quasi-linear, the midgap states recovered in the present study are more complicated objects than what was foreseen by ZL predictions.\textsuperscript{36} To illustrate this notion, we provide in Fig. 16 a different view of the quasi-linear midgap wavefunction from panel (d). In Fig. 16, we have highlighted the bonds connecting the atoms that contribute significantly to the wavefunction. Importantly, the bond lengths along the bulk portion of the wavefunction, highlighted in purple, significantly exceed the length of the covalent As-Se bond, i.e. 2.4 Å.\textsuperscript{48} That the bond length near a topological defect should have an intermediate value than that for the covalent and secondary bond is a hallmark feature of the midgap states.\textsuperscript{20,36} These bonds clearly form a continuous, chain-like pattern. In addition, we observe a “satellite” portion of the wavefunction running alongside its primary portion. This satellite portion evidently has a substantial contribution from lone pairs and likely stems from mixing between the $p$-orbitals that are, respectively, parallel and perpendicular to the line of the defect.

If standalone, a filled state that is close to the center of the gap would be similar to the negatively charged topological state from Fig. 7(b). On the other hand, we observe that vacant midgap states are also generally present in the same sample implying that the just mentioned filled state may, instead, be the lower level of a resonance formed by two or more neutral midgap states, as in Fig. 7(c). There appears to be no conclusive way to distinguish between those situations in the present computational setup since in such modestly-size samples, it might be difficult to produce even a relatively isolated, let alone truly standalone defect. The situation is however clearer with the filled states close to the valence band, i.e., the Urbach states, since they certainly have vacant counterparts close to the conduction band. Thus in view of Fig. 7(c), one may think of the Urbach states as resulting from intimate soliton-antisoliton pairs. This notion adds quite a bit of microscopic detail to the conventional idea

![FIG. 14. Aspect ratios $I_{yy}/I_{xx}$ and $I_{zz}/I_{xx}$ and the corresponding inverse participation ratio (IPR) for select states in the band gap and its vicinity for a As$_{40}$Se$_{60}$ sample that hosts at least one midgap state. The thin solid line shows the electronic density of states, shading indicates filled states.](image)

![FIG. 15. Charge density corresponding to the wavefunctions of the select deep midgap states for four samples marked by the arrows in Fig. 5. In each panel, we have added arrows pointing toward malcoordinated atoms.](image)

deep-midgap and Urbach states is substantial but not dramatic. It is instructive to visualize some of those deep midgap states. We have selected four representative examples shown in Fig. 15; the location of these states in the respective electronic spectra can be found in Fig. 5.

We observe that at least in the cases encountered
of Urbach-tail states as a generic consequence of disorder. Indeed, we observe in Fig. 14 that the characteristics of a filled state change only gradually with the energy of the state. Thus the Urbach states share, to an extent, some of the properties with the deep midgap states, such as the anisotropy in shape.

We have observed above that very deep midgap states spontaneously arise in some samples and match the characteristics of the topological midgap states predicted by Zhugayevych and Lubchenko. Next we study what happens when the system is forced to have at least one dangling bond, which can be accomplished by using an odd number of electrons. Because selenium and arsenic have an even and odd number of electrons, respectively, we can ensure that the system has at least one unpaired spin—and hence a dangling bond—by using an odd number of arsenic atoms. This we accomplish by removing one of the arsenic atoms from a sample that contains an even number of such atoms using three distinct protocols: In protocol (I) we remove an arsenic atom randomly from a sample that has been already geometrically optimized. No further optimization is performed. In protocol (II), we further optimize the sample obtained in protocol (I). In protocol (III), one arsenic atom is removed already from the parent structure. Only after this is the structure geometrically optimized. The logic behind these protocols is as follows: In protocol (I), we expect the defect to be as close as possible in character to a vacancy. In protocol (II), we allow this “vacancy” to relax but subject to an environment that is already mechanically stable. In protocol (III) the system is given the greatest amount of freedom to relax afforded by the structure-building algorithm in Ref. 48. The amount of residual strain in the structure is expected to be the greatest in structures of type I and the least in structures of type III. Everywhere below, we limit ourselves to the stoichiometric compound $\text{As}_{0.4}\text{Se}_{0.6}$. (The stoichiometry is obeyed only approximately.)

The electronic density of states for individual realizations for all three protocols is illustrated in Fig. 17, alongside that for the original structure that contained an even number of electrons. We observe that the gross features of the electronic spectrum are not sensitive to the detailed preparation protocol.

We next focus on the electronic density of states within the mobility gap and its immediate vicinity. Nine realizations for protocols I, II, and III are shown, respectively in Figs. 18, 19, and 20. In all figures, shading indicates filled levels. Consistent with the expectation that samples of type I should exhibit most strain, such samples show on average the greatest number of midgap states. Yet even though samples of type II and III are allowed to relax, some of them still host very deep-lying midgap states. We observe that the midgap states in some samples of type I are similar, at least superficially, to the hole-polaron configuration in Fig. 7(d), see for instance panel 8. This result is perhaps not too surprising since the region centered on the removed arsenic atom has a lower electron density or, equivalently, increased hole density. In contrast, samples of type II seem to house primarily neutral states from Fig. 7(a). Samples of type
FIG. 18. Band structures of nine different samples with an odd number of electrons, as prepared using protocol I.

FIG. 19. Band structures of nine different samples prepared using protocol II. These correspond, panel-wise, to Fig. 18.

FIG. 20. Band structures of nine different samples prepared using protocol III. These correspond, panel-wise, to Figs. 18 and 19.

(III) exhibit such states as well and, in addition, states similar to the electron-polaron states from Fig. 7(d), as in panel 9. While it seems reasonable that structures of type III would exhibit the highest electron density of the three structure types, it is not at all obvious why electron-polaron-like configurations should be so stable as to be readily found already in the small ensemble structures we have generated in the present study.

Finally we refer the reader to the Supplemental Material for the information on the localization and anisotropy of the electronic states for systems with an odd number of electrons, which is presented there in the format of Fig. 14. The main conclusion is that the midgap states with systems with an odd-number of electrons are entirely analogous to systems with filled states. In a distinction, the midgap states in systems of type I are significantly more localized than when the sample is allowed to relax. This is expected since upon relaxation, the lone pair residing in the cavity formed by removing an atom is able to mix more readily with the rest of the orbitals.

V. SUMMARY

We have computationally-generated samples of glassy arsenic selenide, for several stoichiometries, that appear to exhibit all putative types of electronic states thought to exist in these materials. The gross features of the electronic density of states do not vary from sample to sample. The corresponding portions of the density of states are attributed to the mobility bands; these states are extended and their density of states is self-averaging. In contrast, states near the edge of the mobility bands strongly fluctuate in energy. The statistics of this variation match well the distribution of the venerable Urbach-tail states.

Most significantly, we recover a special set of electronic states whose characteristics match those of the topological midgap states predicted earlier by Zhugayevych and Lubchenko (ZL). These special midgap states are very deep into the mobility gap and are not simply a generic consequence of static disorder in the atomic position. Instead, they stem from the vast degeneracy of the free energy landscape of a glassy liquid. The midgap states have been predicted to reside on domain...
walls separating distinct minima on the free energy landscape. We have shown that in addition to the chain-like shape for the midgap states predicted in Ref. 36, more complicated, urchin-like and cyclic shapes are possible, too.

Different charge states of the midgap states have been observed that could be potentially identified with standalone midgap states or their intimate pairs, and also with added charge carriers. Clearly, such polaron-like states are significant in the context of the mechanism of electrical conductance in amorphous chalcogenides. The current view is that in the chalcogenides, electrical current is carried by small polarons,5,6 each of which can be thought of as compound particle consisting of a charge carrier proper and the polarization of the lattice, largely in the spirit of the Born model of solvation of charge in a polar solvent. In contrast with the continuum Born picture, the present treatment directly identifies sets of orbitals that house the charge carrier. Such sets greatly exceed in size and complexity what would be expected of a small polaron. The present results also indicate that the polarization of the lattice involves very particular changes in local bonding such as emergence of malco-ordination.

The above findings were enabled by the availability of a quantum-chemical approximation that recovers the width of the mobility gap with reasonable accuracy, viz., a specific flavor of a hybrid DFT approximation. Standard DFT approaches tend to underestimate the width of the gap thus effectively concealing many of the features of the electronic spectrum this work focused on.

We hope that a more conclusive identification of the presently reported deep midgap states with the topological midgap states predicted by ZL will be possible when we learn how to simulate transitions between distinct free energy minima in a glassy chalcogenide or similar compounds. When approached directly, such a simulation is computationally very costly because of excessively high relaxation barriers.41 Circumventing this complication is work in progress.

VI. ACKNOWLEDGMENTS

This work has been supported by the National Science Foundation Grants CHE-0956127, CHE-1465125, and the Welch Foundation Grant E-1765. We gratefully acknowledge the use of the Maxwell/Opuntia Cluster and the uniring support from the Center for Advanced Computing and Data Systems at the University of Houston. Partial support for this work was provided by resources of the uHPC cluster managed by the University of Houston and acquired through NSF Award Number ACI-1531814.
I. ON THE CHOICE OF THE QUANTUM-CHEMICAL APPROXIMATION

The band gaps produced by a number of distinct approximations are listed in Table S1 for several optimized structures for the stoichiometric compound As₄₀Se₆₀. The structures differ by the amount of vacancies that must be introduced in the parent structure to achieve the desired stoichiometry. The amount of vacancies also affects the mobility of the atoms during the optimization. Values $A < 0.1$ imply vacancies must be introduced in the chalcogen sites, $A > 1$ at the pnictogen sites. We found in Ref. 48 that the optimized structures are not overly sensitive to the value of $A$. Likewise, we find here that the band gap values are also quite robust.

According to Table S1, the band gaps obtained with B3LYP and B3PW91 hybrid functionals are significantly more consistent with observation. Because of this circumstance and the availability of the B3LYP functional in the default distribution of VASP, we have chosen B3LYP for the rest of the calculations. We have tested for the effects of varying the plane wave energy cut-off, within the range 200-400 eV, on the quality of the spectra. Larger values provide for better accuracy but incur greater computational cost. We have found that the value 300 eV provides adequate accuracy in the full spectral range of the DOS, as assessed using the Bivariate (Pearson) correlation for each pair of spectra. All further simulations were performed at the Γ point in the first Brillouin zone with the energy cut-off set to 300 eV and the threshold area parameter $A = 0.100$ unless specifically noted otherwise.

| Functional | Method 1 (DOS) | Method 2 (STDev) |
|------------|----------------|------------------|
| B3LYP      | 1.94           | 1.82             |
| B3PW91     | 1.94           | 1.80             |
| HSE03      | 1.34           | 1.24             |
| HSE06      | 1.49           | 1.40             |
| PW91       | 0.85           | 0.84             |
| PBE        | 0.87           | 0.85             |
| Experiment | 1.78[11], 1.76[12], 1.74[13], 1.64[14] |

TABLE S1. Table of the band gaps calculated using two different techniques: in Method 1 we assume that conduction and valence bands have the functional forms $\propto \sqrt{E - E_c}$ and $\propto \sqrt{E_v - E}$ respectively; in Method 2 we calculate a gap between the second energy levels within the valence and conduction bands counting away from the gap. These methods were applied to the As₄₀Se₆₀ compound, parent structures were generated at $A = 0.1$, the mobility gap was estimated based for a density of states averaged over a number of samples. (The averaging is over 10 samples for B3LYP and PW91, over 5 samples for the rest of the approximations.)

Within chosen cut-off energy limits the band gap changes within 0.02%, a negligibly small value (Fig. S1). The densities of states were compared pairwise. For each pair the Pearson’s correlation coefficient was calculated. In Fig. S2 the correlation matrix demonstrates a correlation over 99% between different simulations, based
on which we can make a conclusion that eDOS is not overly sensitive to varying the energy cut-off.

II. INDIVIDUAL ELECTRONIC SPECTRA FOR OTHER STOICHIOMETRIES

FIG. S3. Electronic density of states for five distinct amorphous samples of the compound As$_{40}$Se$_{60}$, optimized at constant pressure and volume, respectively. C.f. Fig. 2 of the main text.

FIG. S4. Electronic density of states for five distinct amorphous samples of the compound for As$_{50}$Se$_{50}$ and two distinct values of the parameter $A$, optimized at constant pressure and volume, respectively. C.f. Fig. 2 of the main text and Fig. S3.

III. COMPARISON WITH ELECTRONIC SPECTRA OBTAINED IN EARLIER STUDIES

FIG. S5. Predicted total electronic density of states (eDOS) for the As$_{40}$Se$_{60}$ glass, const-$P$, compared with predictions due to Bauchy et al. (BKM). Experimental eDOS inferred from X-ray photo-emission spectroscopy studies due to Kozyukhin et al. (KGKSJ).

FIG. S6. Same as Fig. S5 but at constant volume.
FIG. S7. Predicted total electronic density of states (eDOS) for the As$_{20}$Se$_{80}$ glass, const-$P$, compared with predictions due to Bauchy et al. (BKM).$^{53}$ Experimental eDOS due to Golovchak et al. (GKMJS).$^{64}$

FIG. S8. Same as Fig. S7 but at constant volume.

FIG. S9. Predicted total electronic density of states (eDOS) for the As$_{50}$Se$_{50}$ glass, const-$P$, compared with predictions due to Li et al. (LDKCJ).$^{53}$ Experimental eDOS inferred from X-ray photo-emission spectroscopy studies by Kozyukhin et al. (KGKSJ).$^{63}$

FIG. S10. Same as Fig. S9 but at constant volume.
IV. STATISTICS OF ENERGY LEVELS

FIG. S11. Same as Fig. 6 of the main text, but at constant volume.

FIG. S12. (a) The average energies $E_v^{(n)}$ and $E_c^{(n)}$ of the $n$-term as counted starting, respectively, from the HOMO and LUMO respectively. (b) The corresponding standard deviation. As$_{20}$Se$_{80}$, const-$P$.

FIG. S13. Same as Fig. S12, but at constant volume.

FIG. S14. (a) The average energies $E_v^{(n)}$ and $E_c^{(n)}$ of the $n$-term as counted starting, respectively, from the HOMO and LUMO respectively. (b) The corresponding standard deviation. As$_{50}$Se$_{50}$, const-$P$. 
V. COMPARISON OF ELECTRONIC SPECTRA FOR SAMPLES OBTAINED USING DIFFERENT VALUES OF THE THRESHOLD PARAMETER $A$

FIG. S15. Same as Fig. S14, but at constant volume.

FIG. S16. Averaged electronic density of states for $V=\text{const}$ simulations at $A = 0.100$.

FIG. S17. Averaged eDOS for const-$P$ simulations for different values of $A$, all which correspond to a low concentration of vacancies in the parent structure. Experimental data have been shifted to match the position of the valence band.

FIG. S18. Same as Fig. S17 but at constant volume.
VI. BASIC PROPERTIES OF THE WAVEFUNCTIONS OF THE TOPOLOGICAL MIDGAP STATES, ALSO IN THE PRESENCE OF CROSS-LINKING WITH PERFECTLY DIMERIZED CHAINS

To set the stage, we consider an extended, perfectly dimerized chain terminating with the weaker bond on one end, as in Fig. S19(a). In the latter figure, $t_1$ and $t_2$ denote the electronic transfer integrals for the stronger and weaker bond, respectively: $t_1 > t_2$. Everywhere below, we assume the on-site energies are equal to zero, for simplicity. The argument can be extended straightforwardly for a non-vanishing, sign-alternating on-site energy.\(^{36,81}\) Such a situation would be directly relevant to chain-like motifs in which chalcogen and pnictogen alternate in sequence.

Irrespective of whether the terminal bond is weak or strong, the chain has a bulk spectrum consisting of two bands, the outer edges of the bands at \( \pm |t_1 + t_2| \), the band-gap edges at \( \pm |t_1 - t_2| \). There is also a midgap, “edge” state for the configuration shown in Fig. S19(a) exactly in the middle of the forbidden gap. To see this, we write down the stationary Schrödinger equation for the components \( \psi_n \) of the wave function, where the subscript \( n \) labels the sites of the chain:

\[
-\frac{t_2}{t_1} \psi_n = E \psi_{n-1} - \frac{t_2}{t_1} \psi_{n+1} - t_1 \psi_{n-1} + t_3 \psi_{n+1} = E \psi_n
\]

\[
-\frac{t_2}{t_1} \psi_{n-1} + t_1 \psi_{n+1} = E \psi_n
\]

\[
-\frac{t_2}{t_1} \psi_{n+1} = E \psi_n
\]

\[
\ldots
\]

Clearly, Eqs. (S1) allow for a midgap solution exactly in the middle of the gap:

\[
E = 0 \quad \psi_{2n} = 0, \quad \psi_{2n+1} = -\frac{t_2}{t_1} \psi_{2n-1}.
\]

The wavefunction vanishes on all even-numbered sites, while on the odd-numbered sites, it alternates in sign while decreasing in magnitude by a factor of \( |t_2/t_1| \) per each pair of sites as one moves away from the terminal site. At the latter site, the wavefunction has its largest value. Thus the wave function will decrease exponentially as a function of the distance \( |x| \) away from the terminus:

\[
\psi(x) \propto e^{-|x|\sqrt{\ln(t_1/t_2)/2a}},
\]

where \( a \) is the average spacing between the sites. (We note a peculiar feature of Eq. (S2): If the transfer integral is inversely proportional to the distance between the sites, the wavefunction as a function of the physical coordinate consists of a set of straight lines crossing the origin at the even-numbered sites and thus can be easily drawn by hand, see the dashed line in Fig. S19(a).)

Using the same logic, one can show that such a midgap edge state would not exist if the terminal bond were the strong one because this would entail an exponential increase of the wavefunction toward the bulk of the chain and, consequently, lack of normalizability for the wavefunction. Again, this is straightforwardly evidenced graphically, see the dashed-dotted line in Fig. S19(a).

Non-withstanding their apparent simplicity, Eqs. (S1) demonstrate that the \( E = 0 \) midgap state will persist even if the bond strength varies somewhat in space, as long as the local value of the \( |t_2/t_1| \) ratio tends to a steady value less than one in the bulk in the chain. (The latter condition is necessary to have well defined bulk bands in the first place.) The wavefunction will remain zero on the even-numbered sites. To avoid confusion we note that depending on the detailed dependence of the local value of \( |t_2/t_1| \) ratio on the coordinate, other midgap states may be present.\(^{83}\)

Now, the above setup can be used to appreciate that the vicinity of an undercoordinated site, in an otherwise perfectly dimerized extended chain, will host a \( E = 0 \) midgap state, as in Fig. S19(b). Hereby, the central site of the defect will correspond to site 1 from Fig. S19(a) while the chain itself and midgap wavefunction to the left of site 1 will be the exact mirror images of the respective entities from the r.h.s. of site 1. Thus the wave function is even with respect to the reflection about the central site.
and vanishes on sites that a separated by an odd number of bonds from the central site. Likewise, the vicinity of an over-coordinated site will also host a midgap state, as in Fig. S19(c). The center of the defect now corresponds to site 2 from Fig. S19(a). The corresponding wave function is odd with respect to the reflection about the central site; it vanishes on the central site and on sites that are separated by an even number of bonds from the central site. These properties of the midgap states are of course well known, see Ref. 44 and references therein. According to those earlier works, if let geometrically optimize, the chain will relax so as to get rid of all midgap states other than the one in the middle, 83 which we have seen is robust with respect to vibrational deformations of the chain. This robustness comes about for very general, topological reasons.44

Next, let us couple a chain containing an under-coordinated center with a perfectly dimerized chain, the corresponding transfer integral set equal to $t$, see Fig. 9 of the main text and Fig. S20. Note that now we number the sites on the chains so that the sites participating in the inter-chain coupling—one of them hosting the defect—are labeled “0”. The wave functions on the defected and perfectly dimerized chain are labeled by the letters $\psi$ and $\varphi$, respectively. It will suffice to write out only two entries of the Schrödinger equation

$$-t_2\psi_{-1} - t_2\psi_1 - t\varphi_0 = E\psi_0 \quad (S4)$$
$$-t_2\varphi_{-1} - t_1\varphi_1 - t\psi_0 = E\varphi_0. \quad (S5)$$

Eq. (S4) and the rest of the entries pertaining to the defected chain still allow for a midgap solution at $E = 0$, where the wavefunction is an even function vanishing on odd-numbered sites, so long as $\varphi_0 = 0$. Eq. (S5), on the other hand, allows for a solution at $E = 0$ such that $\varphi_n = 0, n < 0$, while

$$\varphi_1 = -\frac{t}{t_1}\psi_0. \quad (S6)$$

The rest of the positively-numbered $\varphi$’s obey the same equations as the quantities $\psi$ in Eq. (S2). Clearly, the positively-numbered segment of the dimerized chain is analogous to the setup in Fig. S19(a), since $\varphi_0 = 0$ while $\varphi_1 \neq 0$. The electron occupying the midgap state in the defected chain tunnels within the positively-numbered side of the dimerized chain in the same fashion as the electron occupying the terminal site in Fig. S19(a) tunnels toward the bulk of the chain in that figure. The extent of “spillage” of the midgap state wavefunction from the defected to the defect-free chain will be determined by the strength $t$ of the inter-chain coupling, according to Eq. (S6), and could be significant. Physically, the coupling could be realized, for instance, through $sp$-mixing, which would involve neighboring atoms and/or if the chains are not strictly perpendicular. In the latter case, the coupling $t$ goes roughly $t_1 \cos \alpha$, where $\alpha$ is the angle between the chains. Finally, similar logic can be used to show that a wavefunction centered on an over-coordinated site will not “spill” into the crossing chain. Indeed, coupling two chains does not change the symmetry of the Hamiltonian with respect to reflection about the respective sites through which the chains are coupled. Thus Eq. (S4) still yields a midgap state at $E = 0$ whereby $\psi_{-1} = -\psi_1, \psi_0 = 0$, and thus $\varphi_0 = 0$. Consequently, if the dimerized chain were to co-host a $E = 0$ midgap state, the corresponding wavefunction would have to be an exponential function of the coordinate throughout: $\varphi_{2n+2} = -\frac{t_1}{t_2}\varphi_{2n}$, c.f. the bottom equation in (S2). Since such a wavefunction is not normalizable, we conclude that the $E = 0$ midgap state based on an over-coordinated site would be confined to the chain along which the malcoordination takes place.
VII. MIDGAP STATES INDUCED IN CRYSTALLINE SLABS

FIG. S21. The structure and electronic spectrum of a slab of crystalline $\text{As}_2\text{Se}_3$ passivated by hydrogens. The passivation is to achieve the proper bond number (3 for As and 2 for Se) everywhere except in one place on the surface so as to have exactly one over-coordinated atom. The structure is not optimized. The midgap level is seen deep inside the gap but substantially away from the gap center. The corresponding wave-function is shown as the blobs, the two colors denoting the signs of the wavefunction. The quantum chemistry software: MOPAC2016, Version: 18.117L, James J. P. Stewart. We note that by construction, MOPAC greatly overestimates the magnitude of the forbidden gap, see a related discussion in Refs. 36 and 46.

FIG. S22. Same as Fig. S21, but following geometric optimization. The midgap state is seen to move closer to the gap center while the midgap state wave-function becomes more extended and less compact.

VIII. LOCALIZATION AND SHAPE ANISOTROPY OF ELECTRONIC STATES FOR SAMPLES CONTAINING AN ODD NUMBER OF ELECTRONS

FIG. S23. Odd number of electrons, protocol I: Aspect ratios $I_{yy}/I_{xx}$ and $I_{zz}/I_{xx}$ and corresponding inverse participation ratio (IPR) for select states in the band gap and its vicinity for a $\text{As}_{40}\text{Se}_{60}$ sample that hosts at least one midgap state. The thin solid line shown the electronic density of states, shading indicates filled states.

FIG. S24. Same as Fig. S23 but for protocol II.
FIG. S25. Same as Fig. S23 but for protocol III.
This figure "as40se60-const-P-0100-band-925-bond-len4.jpg" is available in "jpg" format from:

http://arxiv.org/ps/1710.04362v3
This figure "as40se60-const-P-0100-band-925-bond-len4.png" is available in "png" format from:

http://arxiv.org/ps/1710.04362v3