PHYSICAL REVIEW B **109**, 104203 (2024)

Understanding localized states in the band tails of amorphous semiconductors exemplified by *a*-Si:H from the perspective of excess delocalized charges

Yuezhou Luo [*] and Andrew John Flewitt
Electrical Engineering Division, Department of Engineering, University of Cambridge, Cambridge CB3 0FA, United Kingdom

(Received 29 September 2023; revised 28 January 2024; accepted 12 February 2024; published 13 March 2024)

In this paper, we use a perturbation strategy to show that the band tails in the density of states (DOS) distribution of amorphous semiconductors form due to the existence of excess delocalized charges. These charges satisfy a Gaussian distribution with a mean of zero and vary slowly in space due to the short- and medium-range order of amorphous materials. The charges exist due to the bond angle and bond length distortion. They induce an extra potential energy distribution that leads to energy band fluctuation in the energy-space diagram and consequently gives rise to localized states. A 150 × 150 × 7.5 nm large-scale finite-element excess charge model for hydrogenated amorphous silicon (*a*-Si:H) is developed using a moving average smoothing that filters a Gaussian array of random charge values. Thanks to the analytical and computational simplicity of the theory in this paper (compared with conventional approaches considering atomistic details and complex electron-electron interactions), this large-scale model is calculated in only 90 s and reproduces the typical exponential and linear features found in the conduction band tail of *a*-Si:H and other amorphous semiconductors. Through a satisfactory fitting to experimental *a*-Si:H DOS data in the literature, model parameters are semiquantitatively determined. Unlike previous analytical and computational efforts, this large-scale model is physically unambiguous, computationally tractable, and satisfactorily accurate. The large-scale modeling capability allows reliable insights into the geometric features of localized and extended states, which are visualized in a nonschematic manner. This further leads to reinvestigations of several established concepts and conclusions. First, calculations in this paper challenge the description that the wave function envelope of a localized state decays away from its center in an exponential manner and that the spread of this exponential decay varies with energy in a power-law manner. Second, impurity states and/or extended states are critical to enable band tail hopping. Third, low-energy states are spatially included inside high-energy states, so the existing concept of average spatial separation of localized states is meaningless. Fourth, the mobility edge obtained from conductivity activation energy measurements turns out to be higher than the actual critical energy E_{ct} that demarcates extended states from localized states; this judgment is supported by the electron mobility-energy relation that is inferred from the geometry of states, which validates a continuous increase of energy-specific mobility as electron energy increases from E_{ct}.

DOI: 10.1103/PhysRevB.109.104203

I. INTRODUCTION

Hydrogenated amorphous silicon (*a*-Si:H) was one of the earliest candidates for the channel semiconductor in thin film transistors that are critical in large-area electronics (LAE) and Internet of Things (IoT) applications. The most noticeable drawback of *a*-Si:H is the low electron mobility, which mainly results from localized electron states induced by structural disorder. Most of these localized states exist in the density of states (DOS) distribution as a nearly exponential tail at the bottom of the conduction band (CB) [1–3]. Similar tails have been observed in other disordered semiconductors as well [4].

Revealing the physical origin of these band tails was a matter of significant discussion in the last century. Kane [5] understood the structural disorder from the perspective of potential fluctuation that follows a Gaussian probability distribution. A Thomas-Fermi-type approximation was adopted, and the DOS distribution of deep localized states was reproduced for highly impure semiconductors. The DOS distribution of shallow states was simulated in an improved model by Halperin and Lax [6], where a spatial autocorrelation function was introduced to model the short-range order in amorphous solids. More accurate modeling schemes were developed later by John *et al.* [7], where different autocorrelation functions were studied for disordered solids. From a different perspective, randomly positioned charged centers were considered in the model of Overhof and Beyer [8], which they thought could account for the existence of potential fluctuations, but the essence of the charged centers and their relation to the material structure of *a*-Si:H are vague. The model developed by Silver *et al.* [9] was like the above approach, but the effect of charged defects was superimposed onto a backbone Gaussian distribution of potential fluctuations, which in turn

---

[*] yl778@cam.ac.uk

generated the exponential feature in the DOS distribution of *a*-Si:H.

Despite their analytical rigor, none of these models probe into and establish direct connections with the structural disorder of amorphous semiconductors. In addition, though capable of reproducing DOS distributions, these models fail to unambiguously exhibit the nature and properties of localized tail states. In this regard, the earlier Anderson localization theory elucidates the physical mechanisms of electron localization [10]. However, insights are not straightforwardly obtained as quantum-mechanical solutions of the potential wells model are necessary, the computational cost of which prevents meaningful results at a reasonable three-dimensional (3D) scale. Returning to the above DOS models, the exponential feature of band tails is itself merely an approximation. Real experimental data from field-effect techniques [11–13], photoelectron spectroscopy [14], etc., never show a strict exponential characteristic. It is usually reported that a linear transitional region exists in and above the band tail [2,11,14,15]. Some researchers even adopted double-exponential DOS modeling to describe the band tail of *a*-Si:H [16–18]. Hence, attempts to reproduce a purely exponential band tail, as pursued by most earlier models, may lead to incorrect interpretation of the underlying physics.

Aside from analytical and theoretical endeavors, computational approaches based on first-principles studies have also been used to reproduce the band structure of amorphous semiconductors such as *a*-Si:H [19–21]. Despite their high accuracies, direct physical insights are missing, and the origins of localized tail states remain ambiguous. Moreover, taking detailed atomistic structures and complex electron-electron interactions into consideration leads to such a high computational load that material models are limited to only several hundreds of atoms even with the aid of high-performance computing facilities. This prevents a reliable understanding of the properties of localized states. The periodic boundary condition may be used to virtually expand the material size, but the reliability of such a treatment is questionable given the structural randomness of amorphous materials. A physically unambiguous, computationally tractable, and satisfactorily accurate large-scale model is therefore lacking, and this gap needs to be filled.

a-Si:H and similar covalently bonded amorphous semiconductors have been gradually overshadowed by the emerging amorphous oxide semiconductors (AOSs) since the breakthroughs of the latter in the early 21st century (e.g., [22]). The superior electronic performances of AOSs (e.g., higher carrier mobilities) signify a higher technological potential and therefore place significant demand on the experimental and application study of these materials, while old theoretical topics, such as the focus of this paper, have received less attention.

In this paper, we aim to reinvestigate this old topic because it is believed that fundamentally understanding the nature and properties of localized tail states will shed light on the electronic conduction mechanisms of amorphous materials and thereby provides practical guidance on the selection and engineering of this technologically important class of materials. In Sec. II, a large-scale *a*-Si:H model based on delocalized charge density has been developed following a perturbation approach that obviates the expensive need of resolving the atomistic details and electron-electron interactions from the beginning. In Sec. III, with the aid of a recent computational algorithm, the model has easily and accurately reproduced the band tail features reported so far for *a*-Si:H. Not only has this model unambiguously revealed the essence and properties of localized tail states, but also it leads to insights and critical modifications to existing concepts and conclusions that were widely accepted. These insights will be discussed in Sec. IV. They would not have been achieved without the large-scale modeling capability reported in this paper. An important modification to the room-temperature multiple trapping and releasing (MTR) transport theory [12,23–25] will be separately elaborated in a sequel to this paper [26]. Further, the model developed in this paper possesses the merit of being capable of downscaling which, unlike existing theories, can be reliably applied to modeling nanoscale amorphous materials. It will be separately presented in a sequel that the electron mobility of a nanoscale *a*-Si:H device is estimated to be ~ 14 times that of a standard *a*-Si:H thin film [27], thereby providing an engineering strategy that fundamentally overcomes the notorious low mobility of *a*-Si:H and bringing this otherwise excellent material back into the spotlight of LAE and IoT.

II. THEORY AND MODELING

The structure of high-quality *a*-Si:H can be considered in a thought experiment where it evolves from crystalline silicon (*c*-Si) through three phases. First, silicon atomic cores rearrange themselves to form a continuous random network (CRN) where the long-range translational symmetry of *c*-Si is lost [28]. Second, silicon sp^3 orbitals attempt to rearrange to adjust to the CRN; this leads to two main results: the variation of bond angle and length (altered overlap) and the formation of dangling bonds (DBs; lost overlap). Third, hydrogen atoms come into the network and passivate most DBs [29].

Assuming that the disorder in the high-quality *a*-Si:H is weak, it is reasonable to take a perturbation approach and assume that the weak disorder only causes perturbations to the atomistic details and electron-electron interactions from those of *c*-Si. Instead of rigorously resolving them from the beginning as a first-principles study would pursue, it is acceptable to analyze the band structure of *a*-Si:H using *c*-Si DOS distribution as a backbone and subsequently consider effective perturbation factors.

Due to their spatial confinement, inner orbital electrons of silicon atoms are not separately resolved. Instead, they are analyzed alongside the silicon nuclei; the ensembles are atomic cores. The structural disorder of *a*-Si:H means that there are displacements of silicon atomic cores from their previous positions in the *c*-Si. However, the average spacing of neighboring silicon atomic cores in *a*-Si:H is similar to that of *c*-Si given the similar densities of the two materials even after excluding hydrogen [30]. Further, given that the absolute displacements of atomic cores are random with no spatial correlations, they do not accumulate a noticeable net change of electric environment within the material. These facts mean that the displacements of atomic cores alone cannot noticeably modify the backbone *c*-Si DOS. The above approximations are similar to that adopted by the reputable Anderson lattice model which also circumvents analysis of the details and displacements of atomic cores [10].

The existence of hydrogen atoms adds more complexity, as they bring in extra electrons and protons into the network which perturb the electric environment. Experimental measurements show that hydrogen does leads to characteristic peaks in the DOS distribution of *a*-Si:H [31,32]. These peaks correspond to the local DOS of the hydrogen atoms themselves and the local DOS of the silicon atoms that are spatially adjacent to these hydrogen atoms [32]. Nevertheless, these characteristic peaks prevail outside the CB tail range which is solely of interest in this paper [32]. Furthermore, not only is the hydrogen concentration in high-quality *a*-Si:H films low (typically ∼10 at. %) [2], but also hydrogen possesses a low valency compared with silicon. The perturbation by hydrogen to the band tail region is thus further limited. It is also worthwhile to note that hydrogen exhibits a higher electronegativity than silicon. Hence, the extra electrons brought by hydrogen are in closer proximity to hydrogen protons than to the rest of the random network linked via silicon atoms. The hydrogen atoms can therefore be treated as isolated neutral points when analyzing the band tail structure.

Hence, the formation of the CB tail must be mainly attributed to the bond angle and bond length distortion. The change of orbital overlap due to the distortion has a consequence of generating an extra nonuniformity to the distribution of delocalized valence band (VB) and CB electrons of silicon (see the note in [33]), which, under thermal equilibrium, may be expressed as

$$\rho(\boldsymbol{r}) = \rho_c(\boldsymbol{r}) + \delta\rho(\boldsymbol{r}), \tag{1}$$

where $\boldsymbol{r}$ is a position vector, ρ_c and ρ refer to the delocalized electron density distribution in *c*-Si and *a*-Si:H, respectively, and $\delta\rho$ is the excess charge density distribution that quantifies the extra nonuniformity. Such a nonuniformity is not unknown either in theories or in experiments [34–37]. Since the number of delocalized electrons from silicon is conserved despite the structural evolution from *c*-Si to *a*-Si:H, the neutrality rule:

$$\iiint_{\boldsymbol{\Omega}} \delta\rho(\boldsymbol{r})dv = 0, \tag{2}$$

holds true, where dv is an infinitesimal volume in the whole spatial region ($\boldsymbol{\Omega}$) of the studied *a*-Si:H.

Focusing on the CB tail, the excess delocalized charge density distribution serves as a perturbation factor that introduces a minor correction to the well-known near-band-edge parabolic DOS of *c*-Si:

$$g_c(E) = \frac{4\pi(2m_n^*)^{3/2}}{h^3}\sqrt{E - E_{C0}}, \tag{3}$$

where m_n^* represents the effective electron mass in *c*-Si, E_{C0} denotes the CB edge of *c*-Si and h is the Planck constant. The minor correction stems from the extra electric potential distribution induced by $\delta\rho(\mathbf{r})$. As has been mentioned, the displacements of silicon atomic cores produce trivial extra potential because the absolute displacements of individual atomic cores are not spatially correlated as revealed by the network randomness. By contrast, the extra potential induced by $\delta\rho(\mathbf{r})$ is more significant. Here, $\delta\rho(\mathbf{r})$ is determined by the spacing between neighboring silicon atomic cores. Thus, it is the relative displacements of atomic cores that matter, see the note in [38]. The radial distribution function (RDF) obtained from neutron scattering indicates that there is short-range order in *a*-Si:H and that the tetrahedral bonding structure is approximately maintained locally [2,39]. In addition, *a*-Si:H also possesses medium-range order. Instead of a pure CRN model, a blended model of CRN and paracrystallites better accounts for the medium-range order characteristic observed from fluctuation electron microscopy [40]. The coexistence of short-range order and paracrystallite-like medium-range order means that the interatomic spacings of silicon at nearby locations are similar. This implies that $\delta\rho(\mathbf{r})$ exhibits certain spatial correlation and that the value of $\delta\rho$ varies slowly in space and allows the accumulation of noticeable electric potential inside the material.

Like the approximation in the model of Kane [5], it is further assumed that the superimposed potential fluctuates considerably slowly in space relative to the spatial oscillations of wave functions. Hence, it causes the previous *c*-Si energy levels to fluctuate in the energy-space diagram in a similar manner. Specifically, the conduction band minimum (CBM) fluctuates as

$$E_C(\boldsymbol{x}) = E_{C0} + (-e)\iiint_{\boldsymbol{\Omega}} \frac{\delta\rho(\boldsymbol{r})}{4\pi\varepsilon_r\varepsilon_0|\boldsymbol{x}-\boldsymbol{r}|}dv. \tag{4}$$

Here, ε_r is the relative permittivity which is deemed constant throughout the material (12 for *a*-Si:H [9]), ε_0 is the vacuum permittivity, and e is the elementary charge. As will be utilized, it is useful to assume that energy levels fluctuate synchronously, but this is acceptable only within a narrow energy range as is the focus of this paper; it may lead to nontrivial inaccuracies when analyzing energy levels that are well separated (see Sec. IV A for further information).

Despite the band fluctuation, the same Fermi level is shared throughout the amorphous system under thermal equilibrium. This can be understood as follows. If a zero temperature was assumed in the beginning when there were only VB electrons that generated band fluctuation in the amorphous system, the Fermi level should fluctuate in a similar manner to the other energy levels. This is because all local regions could be deemed thermally isolated from each other. At room temperature, however, a portion of electrons are thermally excited from VB to CB at every local region, immediately after which is the spatial rearrangement of CB electrons; regions with higher CBM energies tend to lose CB electrons to regions where CBM energies are lower. Holes in VB act similarly in response to the valence band maximum (VBM) fluctuation. The redistribution of CB electrons and VB holes effectively superimposes an extra charge distribution onto the previous valence charge distribution at 0 K. This superimposed charge distribution in turn induces an extra potential distribution that tends to counteract the previous CBM and VBM fluctuation at 0 K. Nevertheless, given the far lower number of CB electrons and VB holes than that of VB electrons, at room temperature, the effect of free-carrier redistribution is merely to smooth the band fluctuation rather than canceling it. In another word, the free carriers can be treated as an attenuator which reduces the extent of band fluctuation. The change of band fluctuation triggers further charge redistribution. These two dynamically interacting factors are meanwhile coupled with the effect of the concentration gradients of CB electrons and VB holes. An equilibrium state is eventually established, which is when the Fermi level is unified throughout the whole amorphous system whereas the other energy levels fluctuate. Regions with lower CBM energies eventually hold more electrons than holes, indicative of lower local activation energies. Returning to Eq. (1), it is the eventual steady-state distribution of VB and CB electrons that is depicted. Therefore, $\delta\rho(\mathbf{r})$ is said to be a macroscopic implicit reflection of the additional electron-electron interactions induced by structural disorder. This perturbation factor is explicitly superimposed onto the basic electron-electron interactions in *c*-Si which have been reflected by the backbone *c*-Si DOS distribution.

Like crystalline systems, the DOS in an amorphous system is defined based on the maximum possible number of electrons per unit energy interval per unit volume. Taking a finite-element approach, *a*-Si:H is segmented into elements termed *short-range localities* (SRLs), which are cubic regions with a side length of a_0 that is similar to the short-range order length which typically spans from 2 to 5 Å [41]. In this paper, a_0 is chosen to be 3 Å. Consider an energy interval $[E, E + \delta E)$ and a SRL at position $\boldsymbol{x}_{i,j,k}$, within which the band is deemed flat in the energy-space diagram. Here, i, j, and k are the position indices of the SRL. Under the assumption of fluctuation synchronicity, the maximum number of electrons in this SRL within the energy interval is

$$\frac{4\pi(2m_n^*)^{3/2}}{h^3}\sqrt{E - E_C(\boldsymbol{x}_{i,j,k})}\delta E a_0^3, \tag{5}$$

in accordance with Eq. (3). Equation (5) maintains accuracy near the local band edge $E_C(\boldsymbol{x}_{i,j,k})$. Here, another assumption is that any energy level (after resolving spin degeneracy) within any SRL holds a maximum number of a_0^3/V electrons, where V is the volume of the material (see Sec. IV B for a brief proof). Scaling up to the whole material, the maximum number of electrons within the specified energy interval is

$$g(E)V\delta E = \sum_i\sum_j\sum_k\left\{\frac{4\pi(2m_n^*)^{3/2}}{h^3}\times\Re\left[\sqrt{E - E_C(\boldsymbol{x}_{i,j,k})}\right]\right\}a_0^3, \tag{6}$$

where $\Re$ extracts the real part of a function. The DOS distribution $g(E)$ of the amorphous material can therefore be calculated if the energy distribution of $E_C(\boldsymbol{x}_{i,j,k})$ is known. To this end, a Gaussian probability density function (PDF) centered at E_{C0} has been usually assumed [5,9,37,42]. However, this assumption is not adopted in this paper for three reasons:

(1) The Gaussian assumption is based on multiple further approximations which might not necessarily be accurate [5].

(2) The standard deviation σ_0 of the energy distribution of $E_C(\boldsymbol{x}_{i,j,k})$ would be the only effective parameter that determines the DOS profile through Eq. (6). This oversimplification significantly reduces the reliability of the theory and limits its capability of fitting to existing DOS data.

(3) The origin and physical meaning of σ_0 is unknown, so it is not a good quantity to unambiguously reflect the extent of structural disorder. Moreover, it is hard to evaluate how σ_0 changes with material size. This prevents a reliable application of this modeling method to nanoscale amorphous semiconductors which are of significant research interest.

A more prudent method is therefore developed in this paper, where the charge density distribution $\delta\rho(\mathbf{r})$ serves as the basis of modeling. Given the short-range order, $\delta\rho(\mathbf{r})$ is equivalently represented by an array of point charges sitting at the centers of individual SRLs. These equivalent excess delocalized charges are quantified by

$$q_{l,m,n} = \iiint_{\boldsymbol{\Omega}_{l,m,n}} \delta\rho(\boldsymbol{r})dv, \tag{7}$$

where $\boldsymbol{\Omega}_{l,m,n}$ denotes the region of the SRL indexed by l, m, and n. Modeling an amorphous thin film now reduces to modeling a discrete 2.5-dimensional (2.5D) array with two attributes:

(1) Globally, the PDF of $q_{l,m,n}$ [instead of $E_C(\boldsymbol{x}_{i,j,k})$] follows a Gaussian distribution with a zero mean. This is based on four reasons. First, $\delta\rho(\mathbf{r})$ meets the neutrality criterion in Eq. (2), and so does $q_{l,m,n}$. Second, given the randomness of structural evolution from *c*-Si to *a*-Si:H and the similar densities of the two materials, the increase and decrease of atomic spacing should occur with similar probabilities. As a result, there should be no evident mechanism by which negative excess charges are created more easily than positive excess charges, and vice versa. Hence, the charge distribution should be symmetric around zero. Third, according to the weak disorder assumption, there are more SRLs whose $q_{l,m,n}$ is nearer zero, indicative of a PDF peaked at zero. Fourth, the creation of excess charge density should obey the central limit theorem given the various independently occurring noises during the structural evolution processes from *c*-Si to *a*-Si:H.

(2) Locally, $q_{l,m,n}$ varies slowly in space. This property has not been specially considered in previous literature (e.g., Refs. [8,9]), but its importance is emphasized here, as it is an alternative expression of the short- and medium-range order detailed earlier.

The autocorrelation function has been used by earlier researchers to model the short- and medium-range order of disordered systems [6,7], but they targeted the potential fluctuation. Though adoptable in this paper, it is hard to ensure that the charge array modeled in this way exhibits a Gaussian distribution. This difficulty increases drastically for 2.5D modeling. Therefore, an alternative method is investigated, where the slow spatial variation is quantified by

$$\begin{aligned}\max(&|q_{l,m,n} - q_{l+1,m,n}|,\ |q_{l,m,n} - q_{l-1,m,n}|,\\ &|q_{l,m,n} - q_{l,m+1,n}|,\ |q_{l,m,n} - q_{l,m-1,n}|,\\ &|q_{l,m,n} - q_{l,m,n+1}|,\ |q_{l,m,n} - q_{l,m,n-1}|) \leqslant \Delta q.\end{aligned} \tag{8}$$

This needs to hold true for almost all (defined here as $\geqslant$98%) indices in the array, where Δq is a constant charge and is defined in a scaled form as

$$\Delta q = \delta \times 4\sigma, \tag{9}$$

with σ being the standard deviation of the Gaussian PDF of $q_{l,m,n}$ and δ being a dimensionless filtering constant. The smaller the filtering constant, the more gradual the spatial variation.

To satisfy both attributes, firstly, a 2.5D random number array with a Gaussian PDF of zero mean and a tentative

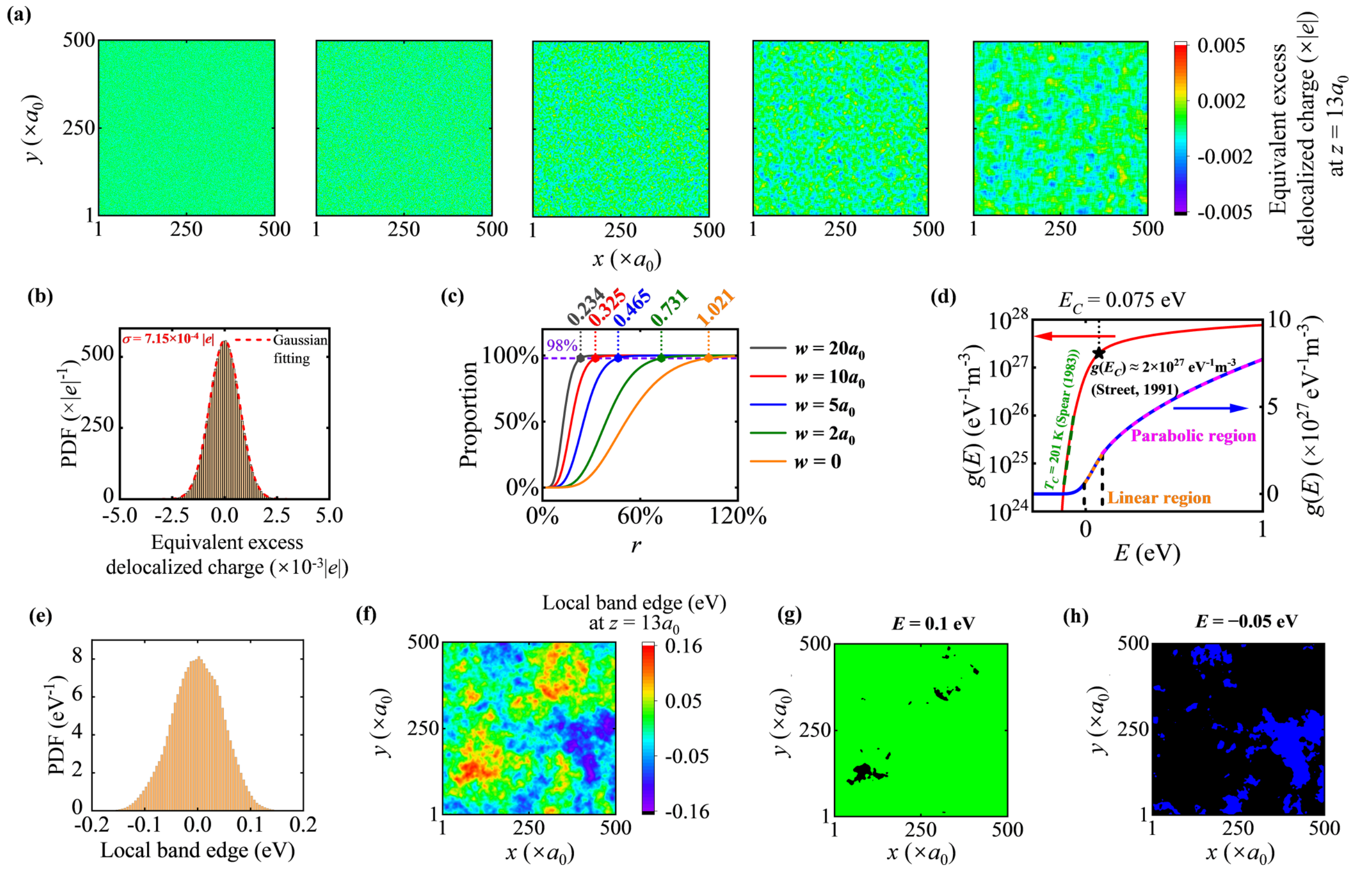

FIG. 1. Simulation of an a-Si:H thin film. (a) The equivalent excess delocalized charge distributions in the middle layer of the 2.5-dimension (2.5D) model under different window dimensions. From left to right, the window dimension w ranges from 0, $2a_0$, $5a_0$, $10a_0$, to $20a_0$. (b) The probability density function (PDF) of charge distribution shared among the five cases in (a). (c) The dependences of the proportion on r (as defined in text) for the five cases in (a). The corresponding filtering constants (δ's) are derived and labeled on the top. (d) The density of state (DOS) distribution of the 2.5D model with $w = 5a_0$. The data are plotted to both the logarithmic scale (red, left) and the linear scale (blue, right). The exponential, linear, and parabolic regions are labeled by the dashed lines on the curves. Quantitative and qualitative features reported in the literature are reproduced as elaborated in text. (e) and (f) The energy distribution and spatial distribution of the local band edges of the 2.5D model fitted in (d). (g) and (h) The geometries of states derived from (f). The green region in (g) and the blue region in (h) denote a typical extended state and a typical localized state, respectively. The energies in (d)–(h) are evaluated relative to E_{C0}.

standard deviation is generated in MATLAB using the built-in NORMRND function. The numbers in this array vary drastically in space. A 3D moving average algorithm is then adopted to smooth the spatial variation. The window dimension (w), which quantifies the side length of the cubic region for the averaging operations in the algorithm, is quantitatively related to the filtering constant δ. It will be shown shortly that the PDF of the smoothed array still maintains the feature of a Gaussian distribution if the array size is sufficiently large relative to w.

III. SIMULATION AND RESULTS

The direct calculation of the band edge distribution from the excess charge distribution via Eq. (4) is essentially a many-body interaction, which scales with N^2 with N being the number of SRLs involved. Like massive celestial and particle simulations, advanced computing algorithms can alleviate the numerical burden while allowing the increase of model size [43–46]. The fast multipole method (FMM) has been chosen to accelerate the computation in this paper, with which the computational load only approximately scales with N [45]. For the 2.5D model here, the reputable FMM3D algorithm developed by the Flatiron Institute has been utilized [47]. The execution time of this Fortran-based parallel computing algorithm has been evaluated, based on which the size of the model in this paper has been chosen as $500a_0$ long, $500a_0$ wide, and $25a_0$ thick, corresponding to a 150 × 150 × 7.5 nm a-Si:H thin film. The band edge calculation takes only ∼ 90 s with an Intel Core i7-12700H CPU (14 cores, 20 threads) and a 32 GB memory (DDR5, 4800 MHz). Though it is still computationally expensive to approach the submillimeter-scale size of a typical amorphous thin film from which most experimental DOS data is produced, the submicrometer scale model here is massive enough to yield semiquantitatively reliable results that are sufficient to validate the theory in this paper.

Shown in Fig. 1(a) is the spatial distribution of the equivalent excess delocalized charges in the middle layer (at $z = 13a_0$) of the modeled thin film. From left to right, w ranges from 0, $2a_0$, $5a_0$, $10a_0$ to $20a_0$. Although the moving average smoothing maintains the Gaussian distribution feature, it changes the standard deviation. The results after rescaling the five systems to match a unified standard deviation are shown in this figure. Its exact value, as labeled by σ in the Gaussian PDF plot in Fig. 1(b), has been retrospectively derived from the subsequent fitting to existing experimental DOS data, which will be discussed in reference to Fig. 1(d).

For every SRL, the differences between its excess charge value and those of its nearest neighboring sites have been evaluated, and the maximum charge difference is recorded. In accordance with the left-hand side of Eq. (8), a factor r_0 is assigned to every SRL, which is defined as the maximum charge difference at the SRL after being divided by 4σ. There are only a proportion of SRLs whose r_0 is less than a certain value r. The dependence of this proportion on r has been calculated in Fig. 1(c). According to the definition in Eqs. (8) and (9), the value of the filtering constant δ for each choice of w has been retrieved and is labeled in the figure. These values are determined in such a way that 98% SRLs in the modeled material possess an r_0 that is less than δ.

Using the FMM3D algorithm, the local band edge distribution has been calculated, and the DOS distribution has been derived via Eq. (6). Both the Gaussian standard deviation σ and the window dimension w were initially unknown. For this semiquantitative study, w is tentatively chosen to be $5a_0$ (see Sec. IV C for further discussions). The value of σ is then determined through fitting the calculated DOS curve to experimental DOS data obtained from a glow discharged a-Si:H sample using field-effect DOS characterization [12] (see Sec. IV C for the reasons why these experimental data are chosen). It is emphasized here that only the tail region DOS data (from 0.2 to 0.14 eV below the CB mobility edge E_C) are selected for fitting. Below this range, the dangling-bond-related midgap defect states begin to interfere. Above this range, the data come from extrapolations [11,12,17,48,49], which are not quantitatively reliable. The fitted DOS curve is shown in Fig. 1(d), where an exponential region of a characteristic temperature $T_C = 201$ K fits well with the experimental data. The position of E_C is determined at the energy that is 0.14 eV higher than the end of the exponential branch. The DOS value at E_C matches the estimation in Ref. [2]. As indicated by photoemission methods [2,31,32], the exponential curve will gradually transition to a parabolic curve at higher energies via a linear region that typically begins from ~ 0.1 eV below E_C [15]. These quantitative and semiquantitative features can all be found in Fig. 1(d). In addition, it will be shown in a sequel to this paper that using the calculated DOS distribution in this paper, an excellent fitting to experimentally obtained temperature-dependent electron drift mobility data is achieved [26]. All these facts validate the accuracy of the band fluctuation theory developed in this paper.

The band edge energy distribution is given in Fig. 1(e). Unlike the assumption in Refs. [5,9], this distribution does not follow a Gaussian profile. The spatial distribution of the band edge in the middle layer is shown in Fig. 1(f). As the thickness of the modeled sample is only $25a_0$, which is much shorter than the characteristic length of band edge fluctuation, it is expected that the spatial distributions within the other layers are almost identical to Fig. 1(f). Thus, only the in-plane state geometry is investigated. Shown in Fig. 1(g) is the geometry of the extended state (see the note in [50]) at the energy of 0.1 eV, where the green regions are allowed and the black regions are prohibited. An electron at this energy can carry net global currents because it can migrate from an electrode on one boundary to the counterelectrode on the opposite boundary. By contrast, in Fig. 1(h), the localized state at the energy of −0.05 eV, which is 0.125 eV below E_C, does not connect the boundaries, so an electron at this energy carries zero net current. These results have partly validated the schematic envisagement of state geometry in Ref. [51].

It is highly significant that the modeling scheme in this paper is scalable; it can be reliably applied to nanoscale amorphous semiconductors. A study on this subject has been carried out, and the results will be reported in a sequel to this paper [27]. The reason for the scalability is that the fundamental bonding structure of a nanoscale amorphous semiconductor is like that of its bulk counterpart if the growth or deposition conditions of the two materials are identical. This is especially so if the nanoscale material is directly etched from the bulk. This means that a fraction of the modeled charges in the bulk material can be directly extracted and reliably passed on to the modeling of the nanoscale material. By contrast, if it was only the potential fluctuation that was modeled at the start (e.g., in Ref. [5]), it would be hard to tell if and how the potential fluctuation in the nanoscale material would be different from that in its bulk counterpart, thus leading to ambiguities in understanding the difference between the properties of the two materials.

IV. DISCUSSIONS

A. Discussion on the different sensitivities of orbitals to the excess delocalized charges

The calculation of DOS distribution in this paper assumes that all energy levels in a previous crystalline counterpart now fluctuate in the energy-space diagram in synchronicity such that the local DOS distribution within a SRL of an amorphous semiconductor is treated as being the same as that of the crystalline counterpart. It has been stated in Sec. II that this assumption only holds true within a narrow energy range such as the band tail region studied in this paper. For energy levels that are well separated, this assumption could lead to noticeable errors. Evidence is that, in experimental DOS data, the CB tail is steeper than the VB tail, which cannot be explained if the fluctuation synchronicity is assumed.

In the derivation of the energy level fluctuation in Sec. II, it is assumed that the excess charge density is distributed uniformly within individual SRLs such that equivalent excess charges are defined, which sit at the centers of the SRLs. In addition, when calculating the potential at a SRL via Eq. (4), this uniform charge distribution assumption allows the omission of the contribution to the potential from the charge within the SRL itself, which otherwise leads to a singularity. Most delocalized charges are valence electrons, which comprise $\frac{3}{4}$ p electronic states and $\frac{1}{4}$ s electronic states. Valence orbitals therefore result from the coupling of the two types of orbital components, and the valence electron cloud possesses a specific spatial geometry. Hence, delocalized charges are not microscopically uniform; they prevail at certain regions inside individual SRLs. Similarly, the electrons at the CBM and VBM also follow specific spatial distributions within individual SRLs. While the VBM comprises mainly p-like orbital

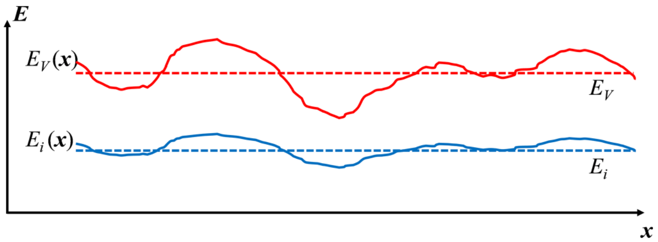

FIG. 2. A one-dimensional (1D) illustration of the fluctuations of valence band (VB) energy levels in *a*-Si. The dashed lines represent valence band maximum (VBM; red) and the *i*th energy level (blue) in the VB of *c*-Si, while the solid lines represent the fluctuation of these levels in *a*-Si.

components, the CBM contains mainly *s*-like orbital components [2]. Thus, it is expected that, compared with CBM electrons, the VBM electrons would have more spatial overlap with the valence electron cloud because the valence electron cloud contains a high proportion of *p* components as well. As a result, the contribution from the excess charges within a SRL to the potential sensed by a VBM electron at this SRL can no longer be omitted. The Coulomb interaction between the VBM electron and the other valence electrons within the SRL is significant due to the spatial proximity of these charges. The potential fluctuation sensed by the VBM electrons are thus underestimated by Eq. (4). According to Eq. (6), the actual VB tail should therefore appear more gradual.

The entire CB spans ∼9 eV [32], whereas the CB tail in this paper only spans an energy range of ∼0.2 eV. Hence, for the purpose of studying the band tail, only a very narrow range of energy levels are involved; they possess similar orbital compositions. The fluctuation synchronicity is thus still a good approximation in this paper.

B. Justification of the assumption on the maximum number of electrons allowed at an energy level within a SRL

It is stated in Sec. II that the maximum number of electrons which can be allowed at any energy level within any SRL is a_0^3/V, where a_0 is the side length of the SRL and V is the whole volume of the material. A *reductio ad absurdum* approach can be used to justify this statement.

First, it should be recalled that the existence of hydrogen atoms is neglected when analyzing the formation of CB tail. DBs in *a*-Si:H are also trivial, as most of them have been passivated by hydrogen. It is thus more straightforward to investigate a fictitious fourfold-coordinated unhydrogenated amorphous silicon (*a*-Si) with neither hydrogen nor DBs, which evolves from *c*-Si through structural disorder. The statement being investigated is regarding the nature of all energy levels, either in CB or VB. For ease of analysis, the VB is considered. As illustrated in Fig. 2, due to the potential fluctuation induced by the excess delocalized charge density, the previously flat VB energy levels now fluctuate in the energy-space diagram; depending on material deposition conditions, this fluctuation can take any possible form and differs for different energy levels as has been mentioned in Sec. IV A.

Suppose that there are M energy levels in the VB of *c*-Si. At 0 K, all the energy levels in the VB of *a*-Si are fully occupied by electrons. If the fact was opposite to the statement under investigation, the maximum number of electrons (N_{ij}) held at the ith energy level within the jth SRL of *a*-Si would be quantitatively correlated with the local energy offset $[E_i(\boldsymbol{x}_j) - E_i]$ via

$$N_{ij} = \frac{a_0^3}{V} + H[E_i(\boldsymbol{x}_j) - E_i], \tag{10}$$

where H is a function whose form is unknown. The total number of valence electrons (M') in *a*-Si at 0 K can be calculated by considering all energy levels from all SRLs:

$$M' = \sum_{i=1}^{M}\sum_{j=1}^{V/a_0^3}\left\{\frac{a_0^3}{V} + H[E_i(\boldsymbol{x}_j) - E_i]\right\}. \tag{11}$$

It is apparent that the evolution from *c*-Si to *a*-Si does not change the total number of valence electrons at 0 K; this requires M=M', the only way to ensure which is

$$\sum_{i=1}^{M}\sum_{j=1}^{V/a_0^3}\{H[E_i(\boldsymbol{x}_j) - E_i]\} = 0. \tag{12}$$

This means that there needs to be a constraint between the energy distributions of all fluctuating levels and the specific form of the function H. While there must be mechanisms leading to a unique form of H (if not always zero), the energy distributions of the fluctuating levels result from the inherent material structure that can take numerous possible forms. Thus, it is impossible to ensure that Eq. (12) always holds true. A contradiction occurs, so the initial assumption is justified.

C. The choice of window dimension *w* for the moving average smoothing and the choice of experimental DOS data

It has been stated in Sec. III that the window dimension w of $5a_0$ is tentatively chosen for the purpose of a semi-quantitative study. It has been subsequently found that the fitting of the DOS curve to experimental data from Ref. [12] is attainable under other choices of window dimension. This fundamentally stems from the fact that the fitted DOS data are limited to a narrow energy range. A stricter fitting requires data spanning a wider energetic window to include the linear and parabolic region as well. The limited energy measurement range is an inherent issue of the field effect DOS characterization used in Ref. [12]. In the field-effect measurement, the Fermi level of the channel semiconductor shifts upward with gate voltage, which induces charges and thereby changes the conductivity [11]. DOS at different energies are derived based on the assumption that the induced charges are all in localized states [11]. To avoid inducing free carriers, the Fermi level must therefore be kept sufficiently below the CB mobility edge of the channel semiconductor. This sets the upper limit of the measurable energies in *a*-Si:H to ∼ 0.14 eV below E_C [12]. Extrapolations beyond the measurable range up to E_C, though widely accepted [11,12,17,48,49], can be doubted, especially in terms of the DOS value at E_C, which is smaller than the values obtained via other techniques (e.g., Ref. [52]). As labeled in Fig. 1(d), Street [2] estimated that this value

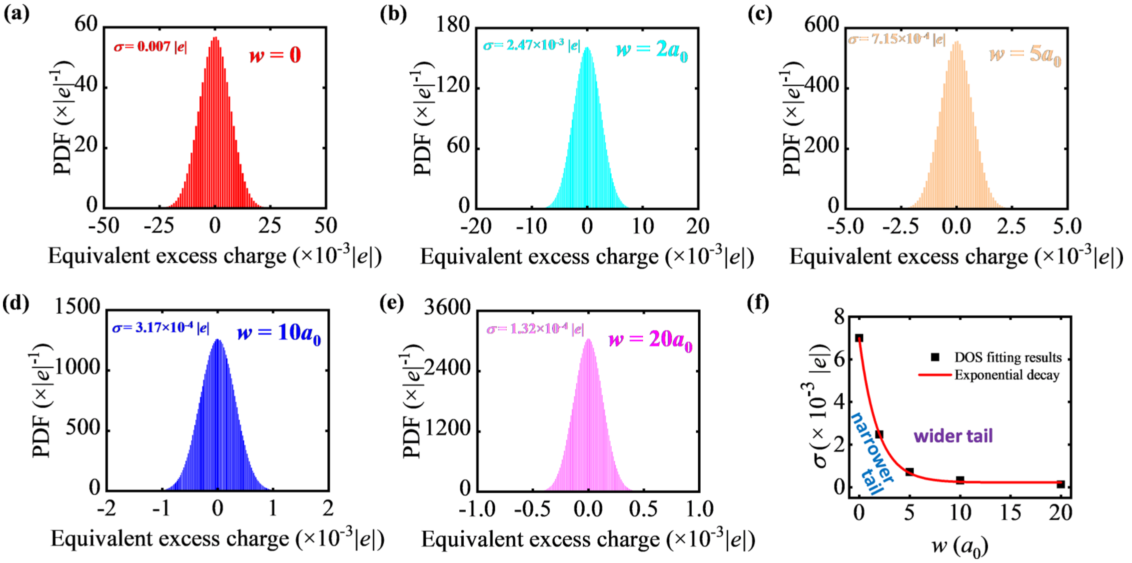

FIG. 3. Different combinations of window dimension w and standard deviation σ to achieve a good fitting to the experimental band tail density of state (DOS) data. (a)–(e) The probability density functions (PDFs) of the equivalent excess delocalized charges, which enable fitting to DOS data under different choices of w. (f) The constraint between w and σ to ensure the DOS fitting. The combinations in (a)–(e) are labeled using black squares; these follow an exponential decay function (red line). A higher σ and/or a larger w lead to a wider calculated band tail than the experimental result, whereas a lower σ and/or a smaller w lead to a narrower band tail.

is $\sim 2 \times 10^{27}$ eV^{-1} m^{-3}. Thus, only the data that are directly obtainable from the field-effect characterization are adopted [i.e., the dashed olive line in Fig. 1(d)].

The extended states in the CB and VB of amorphous materials may be characterized by other techniques. The photoemission measurement is a commonly used technique [32], but it suffers from inaccuracies due to the ambiguities in the estimation of the spatial volume from which photoelectrons emit. Meanwhile, it is limited in energy resolution such that the band tail features are difficult to resolve. Total-yield photoelectron spectroscopy has been a direct technique to reveal the DOS both in the bands and within the mobility gap [14,53]. Nevertheless, the interpretation of DOS from the raw spectroscopic data is not straightforward. Quantitative approximations need to be made. Further, the raw spectroscopic data are essentially count numbers. The inaccuracy further increases due to the necessity to normalize the raw data to a characteristic DOS peak [53], but the peak itself might differ from one sample to another. A nearly perfect exponential VB tail is revealed through this spectroscopy [14]. In addition to the mentioned inaccuracies, in Ref. [14], there is a seemingly intentional quantification of the average dipole transition matrix element $R(\hbar\omega)$ as $|R(\hbar\omega)|^2 \propto (\hbar\omega)^{-5}$ [14], where $\hbar\omega$ denotes the incident photon energy. In fact, a more complex form is elaborated in Ref. [52]. Given that the band tail DOS values are low such that they are prone to be affected by the above characterization inaccuracies, total-yield photoelectron spectroscopy is thus not favored in the energy range of interest in this paper. Nevertheless, if the inaccuracy of this technique decreases for the higher CB where DOS values are much higher, reliable full-spectrum DOS data may be obtained through marrying the extended-state DOS data from photoelectron spectroscopy with the band tail DOS data from the field-effect technique. Apparently, an accurate and reliable calibration is needed to bridge the two types of data due to the use of different methods. More work is needed.

There are also other techniques that characterize the localized-state DOS. Examples include deep level transient spectroscopy [54], the space-charge-limited current measurement [55], the space charge capacitance method [56], and scanning tunnelling spectroscopy [57]. Compared with these, the field-effect technique is more advantageous in terms of its accuracy and the fact that it can reliably characterize both the deep midgap states and shallow tail states. Wide-spectrum optical absorption techniques [58,59] may appear equally favorable, but like the issue of total-yield photoelectron spectroscopy, determining the absolute DOS values from the raw absorption data is not straightforward and may cause extra inaccuracies.

For the above reasons, the DOS data from the field-effect technique are selected in this paper. Due to the limited energy range of the data, this work has to degrade to a semiquantitative study. Nevertheless, w may still be uniquely determined through additional characterization techniques. Shown in Figs. 3(a)–3(e) are a set of excess delocalized charge distributions modeled under different choices of w, which are all capable of yielding a satisfactory fitting to the limited experimental DOS data. It is apparent that the required standard deviation σ, which is the root mean square (RMS) excess charge per SRL, needs to decrease as w increases, to ensure a decent fitting. This tendency is quantified in Fig. 3(f), which can be fitted to an exponential decay function. Deviation from the σ-w constraint leads to a mismatch of the calculated band tail from the experimental result, as detailed in Fig. 3(f). As will be mentioned in Sec. IV F, though not quantitatively reliable, there have been theoretical and experimental attempts to estimate the RMS excess delocalized charge in a-Si. If a more reliable technique is developed in the future to reveal the correct RMS charge value, w will be retrospectively determined. For now, further quantitative judgments on the exact choice of w are meaningless. Instead, it is the plausibility and the significance of semiquantitatively explaining the formation of band tails using the framework developed in this paper that should be highlighted.

D. Reinvestigation of the localized state wave function—Definition and profile

In a crystalline solid, the wave function specifies an electronic eigenstate which exhibits a specific wave number and energy. Since energy levels fluctuate in a disordered solid, a constant energy can now correspond to different local eigenstates at different SRLs, the straightforward consequence of which is that an electron drifting at a constant energy is scattered from one Bloch state to another. The wave number (thus the momentum) ceases to be a good quantum number [2]. Nevertheless, though no longer related to an eigenstate specified by momentum, a wave function can still be defined in the setting of an amorphous material, which reflects the electron probability density distribution in the vicinity of a specific energy.

In the literature from the last century, there was a stereotypical description of the wave function of a localized state in amorphous materials: the envelope of wave function decays in an exponential manner away from its spatial center. A concept called *localization length* was introduced to characterize the exponential decay rate, which was believed to vary with electron energy in a power-law form [2,60–64]. Nevertheless, based on the understanding of band fluctuation and the definition as given in the note in [50], a localized state in an amorphous material is energy specific and is an ensemble made up of basic electronic states with the same energy at different SRLs. A localized state can therefore possess satellite sites with definite boundaries [as shown in Fig. 1(h)]. This casts three doubts on the stereotypical description of the wave function profile. First, a pure decay function cannot be used to depict a complete localized state that contains multiple satellite sites. The decay function may be used to describe a single satellite site, but the spatial spreads of different satellite sites are different even though they are at the same energy. The localization length is thus a poorly defined quantity unless it quantifies the average spread of all the satellite sites of a localized state. Second, a localized state and its satellite sites possess clear spatial boundaries, see the note in [65], outside which the electron probability is zero, so the asymptotic decay is not a favored description of wave functions. Third, even if a single satellite site is focused, endeavors are needed to validate if the wave function envelope within the satellite site approximately follows an exponential profile. It is thus necessary to reinvestigate this old field considering the understanding of band structure from the perspective of band fluctuation in this paper.

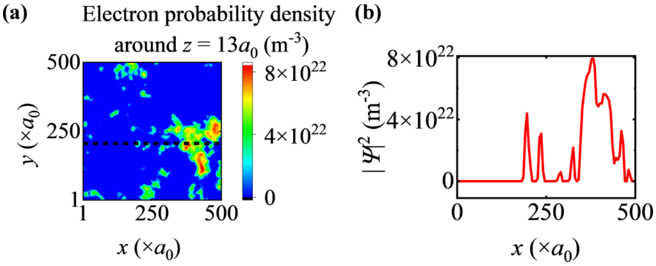

FIG. 4. Calculation of the wave function at $E = -0.05$ eV for the modeled 2.5-dimensional (2.5D) thin film in Fig. 1. (a) The electron probability density distribution near the middle layer. (b) The one-dimensional (1D) probability density distribution near $y = 201a_0$ at the middle layer [labeled by the dashed line in (a)]. These low-resolution calculation results do not resolve high-frequency Bloch components, so they approximate the envelope of wave function.

Consider an infinitesimally small energy interval $[E, (E + \delta E)]$ and a spatial set $\boldsymbol{V}_s$ that includes the centers of multiple neighboring SRLs within an amorphous solid. According to the expression in Eq. (5), the number of electrons within the spatial region denoted by $\boldsymbol{V}_s$ and within the energy interval is calculated as

$$N_e(E, s) = \frac{\sum_{(i,j,k)\in \boldsymbol{V}_s} \frac{4\pi (2m_n^*)^{3/2}}{h^3} \Re[\sqrt{E - E_C(\boldsymbol{x}_{i,j,k})}]\delta E a_0^3}{1 + \exp\left(\frac{E - E_{Fa}}{kT}\right)}, \tag{13}$$

where E_{Fa} denotes the Fermi level of the amorphous solid, k is the Boltzmann constant, and T denotes temperature. Equation (13) is meaningful only when $\boldsymbol{V}_s$ is large enough to allow the Fermi-Dirac statistics which does not resolve the high-frequency Bloch components of electron distribution. The corresponding low-resolution electron probability density distribution is therefore

$$|\Psi(E, s)|^2 = \frac{N_e(E, s)}{V_s \sum_s N_e(E, s)}, \tag{14}$$

where s indicates the position of the set $\boldsymbol{V}_s$, $\Psi(E, s)$ denotes the low-resolution wave function, and V_s is the volume of the spatial region denoted by $\boldsymbol{V}_s$. Using Fig. 1(h) as an example (i.e., at $E = -0.05$ eV) and grouping every 5^3 neighboring SRLs into a set $\boldsymbol{V}_s$, the electron probability density distribution is calculated in Fig. 4, where (b) is a one-dimensional (1D) presentation near $y = 201a_0$ and $z = 13a_0$ as denoted by the dashed black line in (a). In the calculation, the Fermi level is assumed to be at 0.62 eV below E_C based on the measurement of a real a-Si:H sample [25]. Subsequent calculations do not observe any variation of the wave function profile with Fermi level position. This is at least true within a reasonable range of Fermi level positions from 0.4 to 0.8 eV below E_C. The robustness of the calculated wave function to the choice of Fermi level essentially results from the normalization in Eq. (14), where the change brought by different choices of Fermi level tends to cancel in the ratio form.

It should be noted again that the actual wave function contains Bloch components that are of high spatial frequencies. The existence of these Bloch components validates an assumption made in Sec. II that the wave function oscillation is significantly faster than the potential fluctuation. Exclusive of these high-frequency components, the calculated results here are sufficient for the purpose of studying the envelope of wave functions, which casts doubts on the stereotypical exponential decay assumption.

As a localized state does not necessarily exist in the form of a single cluster in space, its localization length is better defined by the average spread of its satellite sites. Not shown here, it is easy to find that, unlike what was widely accepted (e.g., Ref. [64]), the localization length defined this way does not vary with energy in a power-law manner.

E. Reinvestigation of the mobility-energy relation and the position of mobility edge

The conductivity of an *a*-Si:H can be evaluated via [2]

$$\sigma = \int_{E_L}^{+\infty} g(E) e \mu(E) f(E) dE, \quad (15)$$

where the contribution from holes is neglected and the integral starts from the lowest state of electron carriers at E_L below E_C. Here, $f(E)$ is the Fermi-Dirac distribution, and $\mu(E)$ is defined as an energy-specific electron mobility, the form of which was of dispute in the last century. The main argument was on whether, assuming a 0 K temperature, $\mu(E)$ would discontinuously drop to zero as the electron energy crosses E_C [2,4]. This mobility-energy relation can now be reinvestigated considering the modeling developed in this paper.

The mobility of an extended-state electron at a certain energy depends on two factors: the connectivity and the scattering strength. The connectivity evaluates the ease for an electron at a constant energy to find a complete path to migrate from one boundary of the material to the counterboundary. The connectivity at the critical energy (E_{ct}) that demarcates extended states from localized states is just above zero, and the motion of electrons at this energy is highly constrained as the conduction path is just narrowly formed. The higher the energy of a state, the better the connectivity, thus the higher the electron mobility at this energy. Due to the band fluctuation, extended-state electrons are frequently scattered from one momentum to another. Meanwhile, forbidden regions [such as the black regions in Figs. 1(g) and 1(h)] serve as barriers that also scatter the electrons. These two scattering mechanisms significantly reduce the mean free path of an extended-state electron, thus reducing its mobility by around two orders of magnitude compared with the free electron mobility in *c*-Si [47,66]. The change of mobility with energy should be continuous given the continuous evolution of state geometry, as can be inferred from Fig. 1(f). Hence, it is believed that, at 0 K, the electron mobility starts from zero at E_{ct} and continuously increases with an increasing energy. This favors Cohen's point of view [4].

The classical percolation theory may be used to give a rough quantification of the mobility-energy relation [64,67], but this is only reliable near E_{ct} where the connectivity is highly limited such that the effect of scattering is secondary. The connectivity eventually saturates at higher energies in the CB, and these are where the scattering strength becomes the dominant factor that affects the electron mobility. However, the scattering strength at these high energies depends on the extent of the fluctuation of the relevant energy levels, which cannot be predicted now. Moreover, the electron mobility also depends on the effective electron masses, but their values at these high energies are unknown, which can be significantly different from those near E_{ct}. Nonetheless, these ambiguities can be neglected, as the mobility-energy relation is much more important near the CBM where most electronic transport takes place.

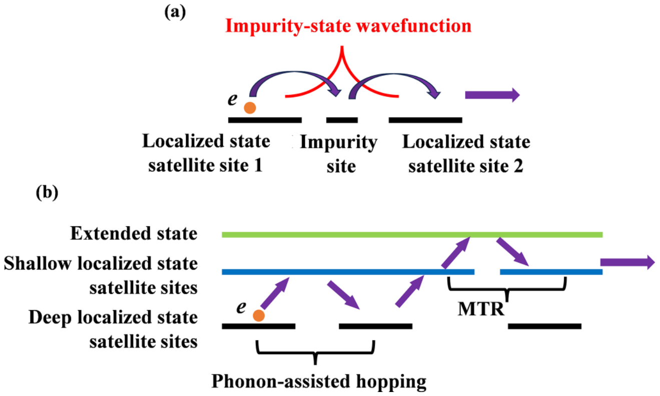

FIG. 5. Illustrations of two possible band tail hopping mechanisms. (a) Impurity-assisted hopping. (b) Combined phonon-assisted hopping and multiple trapping and releasing (MTR). The orange sphere denotes an electron. The purple arrows indicate the migration of the electron.

Under the 0 K assumption, electron hopping is prohibited, so the mobility below E_{ct} would be zero, as these localized states do not connect the boundaries of the material. If a finite temperature is considered instead, the hopping mechanism exists and dominates the transport at low temperatures [25], which therefore maintains nonzero mobilities for electrons below E_{ct}. In a stereotypical description of electron hopping, the wave functions of localized states decay away from their spatial centers in an exponential manner such that overlaps exist between spatially and energetically adjacent localized states to allow tunnelling [2,60–62]. This description of wave function spreads, however, has been questioned in Sec. IV D based on the band fluctuation theory. The following mechanisms are thus proposed to account for the band tail hopping.

As illustrated in Fig. 5(a), the first is impurity-assisted hopping. Unlike the ordinary localized tail states that form due to band fluctuation, impurity states in the band tails are typically related to extrinsic atoms such as dopants. Their wave functions can sit in the band tail, spread out in space, and might bridge over the forbidden regions between the satellite sites of a localized state, which can enable band tail hopping near a constant energy. Without the impurity states, electron tunnelling between neighboring satellite sites will not be possible because the wave functions of the satellite sites have zero overlap due to their definite boundaries. This casts doubts on the description of tunnelling mentioned in Ref. [51]. Shown in Fig. 5(b) is another mechanism based on a combination of phonon-assisted hopping and multiple trapping and releasing (MTR) processes. Some forbidden regions can be bridged by higher localized states whose spatial spreads are wider. Thus, if an electron is thermally excited from a deep localized state to a shallow localized state, it can migrate through some of the forbidden regions of the previous deep state. However, shallow localized states still cannot eventually connect the boundaries of the material. Hopping conduction cannot occur unless a complete conduction path through the sample is formed. Therefore, for the electron to eventually migrate through the whole sample, it has to, at some points, be thermally released to an extended state; this is when the MTR mechanism comes into play, which describes the interaction between localized states and extended states [12,23–25]. For *a*-Si, electronic transport can be almost entirely described by MTR near room temperature, but MTR mechanism rapidly becomes insignificant as temperature decreases since the time for a localized-state electron to be thermally released to an extended state increases rapidly as temperature decreases [23,25]. In comparison, the pure phonon-assisted hopping is less affected by temperature. Thus, the overall band tail hopping conductivity, which is averaged over both the pure phonon-assisted hopping and the MTR processes, decreases less drastically with temperature than the conductivity of a pure MTR transport [25].

In the old theories, the band tail hopping does not involve any extended states or impurity states because it was believed that localized states were randomly and independently distributed in the amorphous material as scattered islands, whose wave functions extend out to overlap with each other, thus forming conduction paths [2,64]. Nevertheless, based on the framework in this paper, extended states and/or impurity states are essential; in reality, the phonon-assisted hopping mechanism and MTR mechanism should be almost always coupled. In addition, some old hopping theories are based on a stereotypical concept called *average spatial separation of localized states* [68,69]. They believe that this average separation can be easily calculated by taking the $-\frac{1}{3}$ power of the total density of localized states that is evaluated through integrating the DOS over the entire band tail. However, according to the model in this paper, a more convincing depiction should be that low-energy states are always spatially included inside high-energy states. Wherever a low-energy state is allowed, a high-energy state is also allowed. Hence, the concept of average separation is meaningless. This concept only makes sense for localized states that are related to defects (e.g., DBs in unhydrogenated *a*-Si) and impurities (e.g., donors and acceptors) which are physical entities that enable Coulomb interactions with electrons, but it is the disorder-related localized states studied in this paper that are predominant in quantity and form band tails.

Based on the above analyses for the mobility above and below E_{ct}, a convincing mobility-energy relation at a finite temperature is envisaged and schematically illustrated in Fig. 6. As has been analyzed, above E_{ct}, the electron mobility continuously increases as the electron energy increases. Below E_{ct}, the mobility drops asymptotically to zero as the energy decreases since the deeper the state, the more distant and smaller the satellite sites, and the harder the hopping.

According to the original definition of mobility edge, E_C should be the same as E_{ct}. Nevertheless, it is not feasible to experimentally determine the critical demarcation energy, as there is currently no technique to characterize the geometry of states. Instead, the mobility edge is indirectly determined via the activation energy of conductivity in temperature-dependent measurements. Near room temperature, hopping conduction is neglected and transport in *a*-Si:H is dominated by the MTR mechanism [25], which leads to a conductivity:

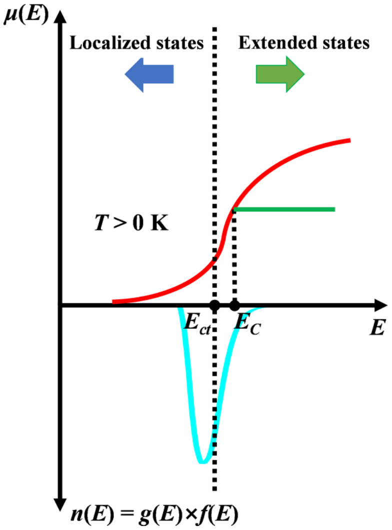

FIG. 6. Illustration of the mobility-energy relation and the energy distribution of electron density. E_{ct} demarcates localized states and extended states. E_C is the experimentally derived conduction band mobility edge through activation energy measurements. The red curve illustrates the actual $\mu(E)$ curve under a finite temperature, while the green line depicts a constant mobility model that is adopted in the interpretation of conductivity and drift mobility results. The cyan curve illustrates the energy distribution of electron density. All the scales are for illustration purposes only.

$$\sigma = \int_{E_c}^{+\infty} g(E) e\mu(E) f(E) dE. \tag{16}$$

In a typical approximation, the mobility is treated as being zero below the mobility edge and being a constant above the mobility edge [2]; this simplifies Eq. (16) to

$$\sigma = e\mu_C n_{\text{ext}}, \tag{17}$$

where μ_C is a constant mobility termed extended-state mobility or free electron mobility [49], and n_{ext} denotes the summed density of all extended-state electrons. Given the energy distribution profile of the extended-state electrons illustrated by the cyan curve in Fig. 6, n_{ext} is further approximated as [49]

$$n_{\text{ext}} = g(E_c) f(E_C) kT, \tag{18}$$

which therefore leads to a conductivity of [2]

$$\sigma = e\mu_C g(E_c) kT \exp\left[-\frac{(E_C - E_F)}{kT}\right]. \tag{19}$$

Like the effect of phonon scattering in crystalline materials, μ_C is believed to decrease with an increasing temperature T. This relation was theoretically derived by Mott as $\mu_C \propto T^{-1}$ [70], which simplifies Eq. (19) to a pure thermally activated form. More generally, assuming $\mu_C \propto T^{-p}$, where p is around (but not necessarily) unity, Eq. (19) still exhibits a thermally activated form which only trivially varies with p. The position of E_C is determined via the activation energy $(E_C - E_F)$ through measuring the logarithmic slope of Eq. (19).

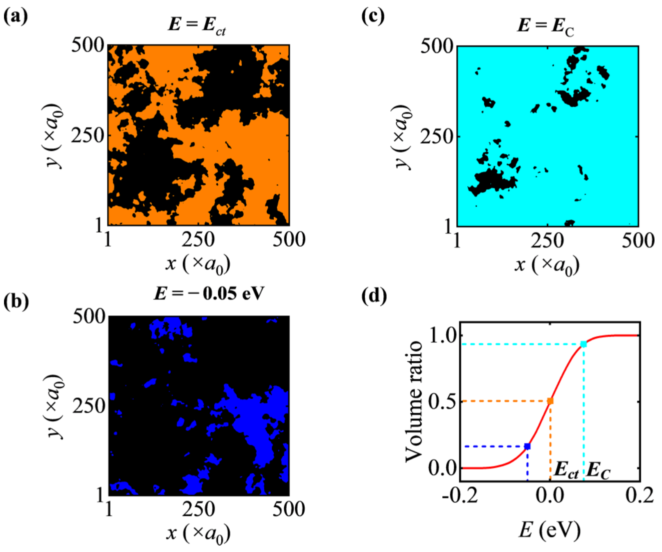

FIG. 7. Investigations of the geometries of states at different energies for the 2.5-dimensional (2.5D) model in Fig. 1(f). (a) The geometry of the state at the critical demarcation energy E_{ct}. (b) A localized state geometry replicated from Fig. 1(h). (c) The state geometry at the fitted mobility edge E_C in Fig. 1(d). (d) The energy dependence of state volume. The energies and volumes of the states in (a)–(c) are indicated using the corresponding colors.

However, considering the actual profile of a $\mu(E)$ curve, it is too crude to treat the mobility above E_C as being constant. Compared with the mobility at higher energies, the mobility in the vicinity of E_{ct} appears too trivial to be incorporated. If one must use a constant mobility approximation, Eq. (17) should instead incorporate only the portion of extended-state electrons beyond an energy that is higher than E_{ct} and neglect the contribution to the conductivity from those in the vicinity of E_{ct}. This is illustrated by the green line in Fig. 6. In Eqs. (18) and (19), E_C should therefore denote an effective mobility edge; this experimentally measurable mobility edge is thus higher than the actual critical energy E_{ct} that demarcates localized and extended states. This result is also favored in Ref. [71]. From a microscopic perspective, a reinvestigation of the MTR transport theory based on the geometry of states in this paper will be elaborated in a sequel to this paper, where a more rigorous study of the positions of the effective mobility edge and E_{ct} is presented [26].

In addition, using the simplified constant $\mu(E)$ approximation, the extended-state mobility μ_C defined in Eq. (17), which can be deduced from time-of-flight transient drift mobility measurements, turns out to be higher than the actual mobility at the demarcation energy. The former is derived to be $\sim 15\,\mathrm{cm}^2/(\mathrm{V\,s})$ [49], while the latter is believed to be $\sim 0.01\,\mathrm{cm}^2/(\mathrm{V\,s})$ at room temperature, which is the lower limit of Brownian motion [4].

Returning to Fig. 1(d), the fitted mobility edge apparently corresponds to an extended state whose energy is higher than the demarcation energy that instead sits near the bottom of the linear DOS region. The state geometry of the latter is shown in Fig. 7(a). For comparison, the geometry of a localized state is replicated in Fig. 7(b) from Fig. 1(h). Figure 7(c) exhibits the geometry of the extended state at the fitted effective mobility edge. The size of state increases with an increasing electron energy; this tendency has been quantified and shown in Fig. 7(d). Given the random nature of the model, it is not surprising to find that the demarcation energy sits at the position where the volume of the state takes up $\sim 50\%$ of the material; this energy position is where a complete path through the material is just formed. The volume ratio at the effective mobility edge is also obtained, but due to the different size of the model from that of the real material where the DOS data are obtained, it is hard to tell if this volume ratio is universally true. Nevertheless, it is the semiquantitative result that matters, and this has been produced.

TABLE I. Reported RMS charge in *a*-Si (summarized from Ref. [36]).

Literature	RMS excess charge ($\times\lvert e\rvert$ per atom)
Guttman *et al.* 1980 (theoretical)[a]	0.2
Klug and Whalley 1982 (experimental)[b]	0.18–0.32
Ley *et al.* 1982 (experimental)[c]	0.11
King *et al.* 1983 (theoretical)[d]	0.04
Kugler *et al.* 1988 (theoretical)[e]	0.021

[a] Reference [34].
[b] Reference [72].
[c] Reference [37].
[d] Reference [35].
[e] Reference [36].

F. Discussions of the RMS excess charge value

Earlier works estimated, either theoretically or experimentally, the RMS excess charge value in *a*-Si [34–37,72]; these values are listed in Table I with the lowest reported value being 0.021 electron units per silicon atom [35]. A closer investigation of Fig. 1(b), however, leads to an RMS excess charge of 7.15×10^{-4} electron units per SRL.

Klug and Whalley [72] held the viewpoint that effective atomic charges are generated in an amorphous material due to the change of atomic dipole moments induced by the displacements of atoms. However, it is hard to tell how such effective charges are correlated with the charges defined in this paper, as the latter act as real electrostatic charges that can induce electric potential, while the former might not. Even if they are related, the RMS charge value derived by Klug and Whalley [72] is not reliable because their estimation based on infrared spectroscopic results contains considerable simplifications [36].

The charges are more clearly defined in the experimental work by Ley *et al.* [37], where they attribute the observed broadening of silicon 2*p* core level spectra in *a*-Si to the charge transfer into or out of the valence shell of each silicon atom. Their quantification of the RMS charge transfer (thus the RMS excess charge per silicon atom) is based on another tight-binding calculation result in which the core level chemical shift scales linearly with the charge transfer and the scaling factor is 2.2 eV per charge transfer of a unit electron [73]. However, the following three aspects should be noted. First, there is no evidence that the scaling is necessarily linear. The scaling factor is crudely calculated based on only a single pair of data coming from the tight-binding calculation of the electronegativity of the Si—O bond in α-quartz and the measured chemical shift of the silicon $2p$ core level. Second, the tight-binding calculation itself can be questioned, as different parameter choices can lead to different results; this is shown in their Fig. 8 [73]. Third, the tight-binding results are obtained in the specific context of α-quartz. The effect of the existence of oxygen atoms on the chemical shift is considered entirely from the perspective of charge transfer between silicon and oxygen atoms, but the actual mechanism is apparently much more complex. The authors do not consider this, so the quantitative result, 2.2 eV chemical shift per unit charge transfer, should be highly specific to α-quartz and should not be generally applicable to other solids such as the *a*-Si studied by Ley *et al.* [37]. Aside from the issues with the tight-binding calculation mentioned above, the experiment targets $2p$ orbitals. Like the analysis in Sec. IV A, considering spatial proximity, the $2p$ electrons are more sensitive to the spatial distribution of valence electrons, while CBM electrons, which are of *s*-like electronic states, are less sensitive. The results in Sec. III of this paper are based on the latter, so the lower RMS charge value is understandable.

The theoretical investigation conducted by Guttman *et al.* [34] assigns net charges to individual silicon atoms through a population analysis of one-electron eigenfunctions. This is built on a CRN where each silicon atom is fourfold coordinated. Their larger RMS net atomic charge value than Fig. 1(b) can be attributed to the reliability issue of the CRN model itself. The structure of a real *a*-Si contains a significant number of coordination defects such as DBs [2], which facilitate the reduction of the overall network coordination number to an ideal value 2.45 [74]. The rigid fourfold coordinated CRN models adopted by these earlier theoretical investigations [34,36] are apparently overcoordinated, where bond strains accumulate. Based on the microscopic origin of excess delocalized charges described in Sec. II, the fourfold CRN models therefore result in higher RMS charge values.

The model used by King *et al.* [35] targets *a*-Si:H. Though each silicon atom is still fourfold coordinated, the incorporation of hydrogen reduces the network coordination such that strains are relaxed. Based on this more reliable model, they have calculated the net atomic charges for both *a*-Si:H and unhydrogenated *a*-Si, the latter having considered DBs by setting the Hamiltonian parameters related to hydrogen to zero (i.e., by eliminating hydrogen atoms from the network). The resultant RMS atomic charges are still high, which can be doubted because the Hamiltonian parameters are taken from the calculations for *c*-Si. Meanwhile, the model is rather crude, as admitted by the authors themselves. Moreover, they calculated the local charges through evaluating the number of occupied electronic states based on their own DOS calculation. The reliability of the calculated DOS distribution and the Fermi level they derived are doubtful. Further, it should be noted that the CRN models adopted by the above studies contain <315 silicon atoms due to their computational complexity [34–36]; the statistical feature may not be sufficiently manifested in these models due to their limited sizes. This issue, coupled with the inaccuracy of model parameters, may lead to qualitative errors.

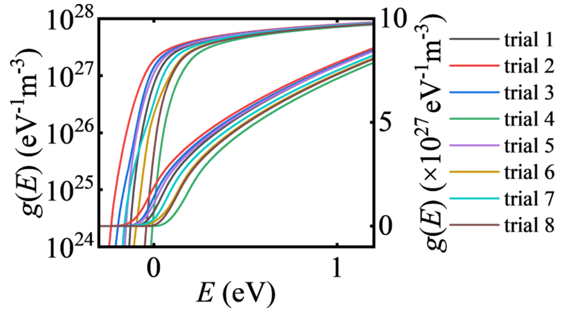

FIG. 8. Density of state (DOS) curves plotted in both logarithmic and linear scales, calculated in different stochastic trials with the same charge distribution standard deviation σ and the same window dimension w.

Another interesting point to be noted is that, though not clearly mentioned in the literature, these theoretical calculations were most likely to have been conducted under a zero-temperature setting. As has been mentioned in Sec. II, the excitation and redistribution of free carriers at an elevated temperature attenuates the 0 K band fluctuation and thus leads to a steeper band tail according to Eq. (6). According to Fig. 3(f), this steeper tail indicates either a lower RMS excess charge value or a smaller w. The redistribution of free carriers is driven by CBM and VBM fluctuations and should thus follow their spatial characteristics. According to Fig. 1, band fluctuation is much slower than typical delocalized excess charge fluctuation, so it is anticipated that the redistribution of free carriers should at least not lead to a faster charge fluctuation. In fact, given that the number of the free carriers are trivial among all the delocalized charges, w at the elevated temperature should be almost identical to that at 0 K. Thus, it is certain that the steeper tail must be correlated with a reduced RMS delocalized charge value. As a result, the lower RMS charge value in this paper, which is under a room-temperature setting, is reasonable.

Looking back on this paper, experimental DOS data are utilized in the derivation of charge distribution. These data come from a glow discharged *a*-Si:H sample [12]. *a*-Si:H is expected to exhibit a lower RMS charge value than the pure *a*-Si studied in the above literature because the passivation of DBs by hydrogen atoms reduces bond length fluctuation in the material [37] and thus also leads to a steeper CB tail.

In the end, the semiquantitative study in this paper is deemed convincing and is sufficient to reveal the origin of the localized states and account for the profile of the CB tail.

G. Investigation of the stochasticity and the robustness of the proposed modeling scheme

Due to the stochastic nature when generating the 3D array using the NORMRND function at the very initial stage, the calculated DOS curves are expected to differ for each trial. To evaluate the extent of such differences and validate the robustness of the proposed modeling scheme, another seven trials of modeling have been conducted. The set of DOS curves shown in Fig. 8 are calculated based on the charge distributions modeled with w and σ that are the same as those of Fig. 1(d). It is evident that the effect of stochasticity is predominantly on the absolute horizontal positions of the DOS curves, whereas the relative energy ranges of the exponential and linear regions, the characteristic temperature of the exponential region, and the slope of the linear region remain almost unchanged regardless of the stochasticity. Without knowing addition information such as electron affinity, which specifies the absolute value of E_C, any trial is acceptable and can lead to generally the same σ and w upon fitting to experimental DOS data. Therefore, σ and w are believed to be rough but reliable quantifications of the extent of material disorder. The robustness of the modeling scheme is thus validated.

V. CONCLUSIONS

In conclusion, in this paper, we have developed a physically unambiguous, computationally tractable, and satisfactorily accurate large-scale model to account for the formation of band tails in the DOS distributions of amorphous semiconductors. Here, *a*-Si:H is used as an example. Rather than resolving detailed atomistic structures and complex electron-electron interactions from the beginning, this model is based on a perturbation approach which attributes the existence of localized band tail states in *a*-Si:H to the band fluctuation induced by an excess delocalized charge density distribution. This excess charge density distribution forms as a result of bond angle and bond length distortion in *a*-Si:H. Through finite-element analyses, equivalent discrete excess charges are modeled. It is then reasonably assumed that these excess charges follow a Gaussian distribution and vary slowly in space due to the short- and medium-range structural order in amorphous semiconductors. A 150 × 150 × 7.5 nm *a*-Si:H thin film model is constructed through a moving average smoothing technique that filters a Gaussian array of 6.25×10^6 random charge values. The band fluctuation induced by the excess charges is easily evaluated using the 3D FMM algorithm developed by the Flatiron Institute. Accordingly, the DOS distribution is obtained, and the key model parameters are determined in a semiquantitative manner through fitting the calculated DOS distribution to existing experimental data. The satisfactory fitting retrospectively validates the reliability of the proposed perturbation approach.

Unachieved by earlier analytical and computational endeavors the authors are aware of, the large-scale model in this paper has nonschematically revealed the geometric features of localized and extended states; this has profound significance on the understanding of charge transport in *a*-Si:H. A sequel to this paper will continue this topic, where important modifications are made to the established MTR theory considering the understanding in this paper [26]. Based on the insight of state geometries, several existing concepts are questioned or reinvestigated in this paper. These include the envelope profile of localized wave functions, the mechanisms of band tail hopping, the mobility-energy relation, and the position of mobility edge.

The model is robust against the stochasticity of random number generation at the initial stage. Most importantly, it is scalable, meaning that it can be reliably utilized to study nanoscale amorphous semiconductors deposited under a similar condition to that of their bulk counterparts. This will become increasingly important as device dimensions continue to be reduced, and this topic will be elaborated in the next paper by the authors, which estimates an ∼14 times increase of electron mobility in a nanoscale *a*-Si:H device compared with that of a standard *a*-Si:H thin film [26].

The supporting data for this article are openly available from the University of Cambridge repository (Apollo) at Ref. [75].

ACKNOWLEDGMENTS

Y.L. appreciates the technical support from Fanyu Ding on the setup of FMM3D. This paper is supported by the UKRI Engineering and Physical Sciences Research Council under Grant No. EP/W009757/1. The authors acknowledge the Rank Prize for their Return to Research Grant. Y.L. is grateful to the Cambridge Commonwealth, European and International Trust for their Ph.D. scholarship.

---

[1] F. Urbach, The long-wavelength edge of photographic sensitivity and of the electronic absorption of solids, Phys. Rev. **92**, 1324 (1953).

[2] R. A. Street, *Hydrogenated Amorphous Silicon* (Cambridge University Press, New York, 1991).

[3] T. Tiedje, B. Abeles, and J. Cebulka, Urbach edge and the density of states in hydrogenated amorphous silicon, Solid State Commun. **47**, 493 (1983).

[4] M. H. Cohen, Review of the theory of amorphous semiconductors, J. Non-Cryst. Solids **4**, 391 (1970).

[5] E. O. Kane, Thomas-Fermi approach to impure semiconductor band structure, Phys. Rev. **131**, 79 (1963).

[6] B. Halperin and M. Lax, Impurity-band tails in the high-density limit. I. Minimum counting methods, Phys. Rev. **148**, 722 (1966).

[7] S. John, M. Chou, M. Cohen, and C. Soukoulis, Density of states for an electron in a correlated Gaussian random potential: Theory of the Urbach tail, Phys. Rev. B **37**, 6963 (1988).

[8] H. Overhof and W. Beyer, A model for the electronic transport in hydrogenated amorphous silicon, Philos. Mag. B **43**, 433 (1981).

[9] M. Silver, L. Pautmeier, and H. Bässler, On the origin of exponential band tails in amorphous semiconductors, Solid State Commun. **72**, 177 (1989).

[10] P. W. Anderson, Absence of diffusion in certain random lattices, Phys. Rev. **109**, 1492 (1958).

[11] W. Spear and P. Le Comber, Investigation of the localised state distribution in amorphous Si films, J. Non-Cryst. Solids **8**, 727 (1972).

[12] W. Spear, The study of transport and related properties of amorphous silicon by transient experiments, J. Non-Cryst. Solids **59**, 1 (1983).

[13] C. Chen, K. Abe, H. Kumomi, and J. Kanicki, Density of states of a-InGaZnO from temperature-dependent field-effect studies, IEEE Trans. Electron Devices **56**, 1177 (2009).

[14] K. Winer, I. Hirabayashi, and L. Ley, Exponential conduction-band tail in P-doped *a*-Si:H, Phys. Rev. Lett. **60**, 2697 (1988).

[15] R. Street, J. Kakalios, and M. Hack, Electron drift mobility in doped amorphous silicon, Phys. Rev. B **38**, 5603 (1988).

[16] R. Vandergaghen and C. Longeaud, Experimental determination of hydrogenated amorphous silicon and amorphous silicon-germanium alloy electron transport parameters from time of flight experiment, J. Non-Cryst. Solids **97**, 1059 (1987).

[17] J. Marshall, P. LeComber, and W. Spear, Comments on the calculation of the extended state electron mobility in amorphous silicon, Solid State Commun. **54**, 11 (1985).

[18] C. Longeaud, G. Fournet, and R. Vanderhaghen, Determination of the density of states of the conduction-band tail in hydrogenated amorphous silicon, Phys. Rev. B **38**, 7493 (1988).

[19] K. Jarolimek, R. De Groot, G. De Wijs, and M. Zeman, First-principles study of hydrogenated amorphous silicon, Phys. Rev. B **79**, 155206 (2009).

[20] N. Cooper, C. Goringe, and D. McKenzie, Density functional theory modelling of amorphous silicon, Comput. Mater. Sci. **17**, 1 (2000).

[21] D. A. Drabold, P. Fedders, S. Klemm, and O. F. Sankey, Finite-temperature properties of amorphous silicon, Phys. Rev. Lett. **67**, 2179 (1991).

[22] K. Nomura, H. Ohta, A. Takagi, T. Kamiya, M. Hirano, and H. Hosono, Room-temperature fabrication of transparent flexible thin-film transistors using amorphous oxide semiconductors, Nature (London) **432**, 488 (2004).

[23] T. Tiedje and A. Rose, A physical interpretation of dispersive transport in disordered semiconductors, Solid State Commun. **37**, 49 (1981).

[24] P. Le Comber, A. Madan, and W. Spear, Electronic transport and state distribution in amorphous Si films, J. Non-Cryst. Solids **11**, 219 (1972).

[25] P. Le Comber and W. Spear, Electronic transport in amorphous silicon films, Phys. Rev. Lett. **25**, 509 (1970).

[26] Y. Luo and A. Flewitt, Revisiting the multiple trapping and releasing electronic transport in amorphous semiconductors exemplified by a-Si:H (unpublished).

[27] Y. Luo and A. Flewitt, Increased electron mobility of hydrogenated amorphous silicon in the context of a nanogap device (unpublished).

[28] W. H. Zachariasen, The atomic arrangement in glass, J. Am. Chem. Soc. **54**, 3841 (1932).

[29] J. Poortmans and V. Arkhipov, *Thin Film Solar Cells: Fabrication, Characterization and Applications* (John Wiley & Sons, Chichester, 2006).

[30] T. Haage, U. Schmidt, H. Fath, P. Hess, B. Schröder, and H. Oechsner, Density of glow discharge amorphous silicon films determined by spectroscopic ellipsometry, J. Appl. Phys. **76**, 4894 (1994).

[31] B. Von Roedern, L. Ley, and M. Cardona, Photoelectron spectra of hydrogenated amorphous silicon, Phys. Rev. Lett. **39**, 1576 (1977).

[32] W. Ching, D. J. Lam, and C. C. Lin, Electronic states and bonding configurations in hydrogenated amorphous silicon, Phys. Rev. B **21**, 2378 (1980).

[33] VB electrons and CB electrons are delocalized electrons in the sense that they, unlike core electrons, are not tightly confined around a single atomic core. However, this concept of localization should be differentiated from the localization of states that are used in contexts such as localized-state electrons; the former describes spatial confinement at the atomic scale, whereas the latter describes spatial confinement relative to the scale of the studied solid. CB electrons can be in localized states; this is when they cannot move throughout an amorphous material to carry net currents, but they are essentially delocalized from the atomic perspective.

[34] L. Guttman, W. Ching, and J. Rath, Charge-density variation in a model of amorphous silicon, Phys. Rev. Lett. **44**, 1513 (1980).

[35] H. King, B. Kramer, and A. MacKinnon, Density of states and charge distribution in hydrogenated amorphous silicon, Solid State Commun. **47**, 683 (1983).

[36] S. Kugler, P. R. Surján, and G. Náray-Szabó, Theoretical estimation of static charge fluctuation in amorphous silicon, Phys. Rev. B **37**, 9069 (1988).

[37] L. Ley, J. Reichardt, and R. Johnson, Static charge fluctuations in amorphous silicon, Phys. Rev. Lett. **49**, 1664 (1982).

[38] In a 3D setting, the spatial correlations of the absolute displacements and the relative displacements can differ. This is easily understood by considering the extra degree of freedom brought by the bond angle variation.

[39] J. Fortner and J. Lannin, Radial distribution functions of amorphous silicon, Phys. Rev. B **39**, 5527 (1989).

[40] P. M. Voyles and J. R. Abelson, Medium-range order in amorphous silicon measured by fluctuation electron microscopy, Sol. Energy Mater. Sol. Cells **78**, 85 (2003).

[41] S. Lan, L. Zhu, Z. Wu, L. Gu, Q. Zhang, H. Kong, J. Liu, R. Song, S. Liu, G. Sha, *et al.*, A medium-range structure motif linking amorphous and crystalline states, Nature Mater. **20**, 1347 (2021).

[42] Z.-Q. Zhang and P. Sheng, Density of localized states near the band edge of disordered systems, Phys. Rev. Lett. **57**, 909 (1986).

[43] R. Hockney and J. Eastwood, *Computer Simulations Using Particles* (McGraw-Hill, New York, 1981).

[44] J. Barnes and P. Hut, A hierarchical $O(N \log N)$ force-calculation algorithm, Nature (London) **324**, 446 (1986).

[45] L. Greengard and V. Rokhlin, A fast algorithm for particle simulations, J. Comput. Phys. **73**, 325 (1987).

[46] U. Essmann, L. Perera, M. L. Berkowitz, T. Darden, H. Lee, and L. G. Pedersen, A smooth particle mesh Ewald method, J. Chem. Phys. **103**, 8577 (1995).

[47] T. Askham, Z. Gimbutas, L. Greengard, L. Lu, J. Magland, D. Malhotra, M. O'Neil, M. Rachh, V. Rokhlin, and F. Vico, Flatiron Institute fast multipole libraries, (2021).

[48] A. Madan, P. G. Le Comber, and W. Spear, Investigation of the density of localized states in *a*-Si using the field effect technique, J. Non-Cryst. Solids **20**, 239 (1976).

[49] A. Hourd and W. Spear, Determination of the extended-state electron mobility in *a*-Si, Philos. Mag. B **51**, L13 (1985).

[50] According to the band fluctuation model, in amorphous solids, the same electron energy may correspond to different momenta at different spatial locations. The terminologies *extended state* and *localized state* in the literature typically exist in the context of a specific energy. This is especially so when analyzing electronic transport, where it is the energies of electrons, rather than their momenta, that are vital. These terminologies need to be distinguished from the *electronic states* within individual SRLs. The latter are momentum-specified local eigenstates which are relevant to DOS calculations. Resolving spin degeneracy, an

electronic state allows a maximum of one electron. To avoid ambiguities, in this paper and its sequels, unless in the contexts of electronic state, DOS, or Bloch state, a state in all other contexts refers to an ensemble that is inclusive of all electronic states with the same energy. These electronic states are at different spatial locations.

[51] H. Fritzsche, Optical and electrical energy gaps in amorphous semiconductors, J. Non-Cryst. Solids **6**, 49 (1971).

[52] W. Jackson, S. Kelso, C. Tsai, J. Allen, and S.-J. Oh, Energy dependence of the optical matrix element in hydrogenated amorphous and crystalline silicon, Phys. Rev. B **31**, 5187 (1985).

[53] K. Winer and L. Ley, Surface states and the exponential valence-band tail in *a*-Si:H, Phys. Rev. B **36**, 6072 (1987).

[54] D. Lang, Deep-level transient spectroscopy: A new method to characterize traps in semiconductors, J. Appl. Phys. **45**, 3023 (1974).

[55] W. Den Boer, Determination of midgap density of states in *a*-Si:H using space-charge-limited current measurements, J. Phys. Colloques **42**, C4-451 (1981).

[56] A. Glade, W. Fuhs, and H. Mell, Frequency and temperature dependence of the space charge capacitance in *a*-Si:H films, J. Non-Cryst. Solids **59**, 269 (1983).

[57] D. Bonnell, *Scanning Probe Microscopy and Spectroscopy: Theory, Techniques, and Applications* (Wiley-VCH, New York, 2000), 2nd ed.

[58] M. Vaněček, J. Kočka, J. Stuchlik, and A. Tříska, Direct measurement of the gap states and band tail absorption by constant photocurrent method in amorphous silicon, Solid State Commun. **39**, 1199 (1981).

[59] W. B. Jackson, N. M. Amer, A. Boccara, and D. Fournier, Photothermal deflection spectroscopy and detection, Appl. Opt. **20**, 1333 (1981).

[60] P. Anderson, The size of localized states near the mobility edge, Proc. Natl. Acad. Sci. USA **69**, 1097 (1972).

[61] M. Pepper, S. Pollitt, and C. Adkins, The spatial extent of localized state wavefunctions in silicon inversion layers, J. Phys. C: Solid State Phys. **7**, L273 (1974).

[62] S. Yoshino and M. Okazaki, Numerical study of electron localization in Anderson model for disordered systems: Spatial extension of wavefunction, J. Phys. Soc. Jpn. **43**, 415 (1977).

[63] S. John, C. Soukoulis, M. H. Cohen, and E. Economou, Theory of electron band tails and the Urbach optical-absorption edge, Phys. Rev. Lett. **57**, 1777 (1986).

[64] N. F. Mott and E. A. Davis, *Electronic Processes in Non-Crystalline Materials* (Oxford University Press, New York, 2012), 2nd ed.

[65] The clear boundaries of localized states in an amorphous semiconductor directly result from the clear band edge of its crystalline counterpart where there is no electronic state within the band gap.

[66] D. Neamen, *Semiconductor Physics and Devices: Basic Principles* (McGraw-Hill, New York, 2012), 4th ed.

[67] B. Last and D. Thouless, Percolation theory and electrical conductivity, Phys. Rev. Lett. **27**, 1719 (1971).

[68] D. Adler, Electronic correlations and transient effects in disordered systems, Sol. Energy Mater. **8**, 53 (1982).

[69] D. Monroe, Hopping in exponential band tails, Phys. Rev. Lett. **54**, 146 (1985).

[70] N. F. Mott, Conduction in non-crystalline materials: III. Localized states in a pseudogap and near extremities of conduction and valence bands, Philos. Mag. **19**, 835 (1969).

[71] X. Wang, L. F. Register, and A. Dodabalapur, Redefining the mobility edge in thin-film transistors, Phys. Rev. Appl. **11**, 064039 (2019).

[72] D. Klug and E. Whalley, Effective charges of amorphous silicon, germanium, arsenic, and ice, Phys. Rev. B **25**, 5543 (1982).

[73] R. N. Nucho and A. Madhukar, Electronic structure of SiO_2: α-quartz and the influence of local disorder, Phys. Rev. B **21**, 1576 (1980).

[74] J. C. Phillips, Topology of covalent non-crystalline solids I: Short-range order in chalcogenide alloys, J. Non-Cryst. Solids **34**, 153 (1979).

[75] Y. Luo, University of Cambridge Repository (Apollo) (2024), https://doi.org/10.17863/CAM.106553.