



This is an electronic reprint of the original article. This reprint may differ from the original in pagination and typographic detail.

Järvi, Jari; Li, Jingrui; Rinke, Patrick

Multi-scale model for the structure of hybrid perovskites

Published in: New Journal of Physics

DOI: 10.1088/1367-2630/aae295

Published: 10/10/2018

Document Version Publisher's PDF, also known as Version of record

Published under the following license: CC BY

Please cite the original version:

Järvi, J., Li, J., & Rinke, P. (2018). Multi-scale model for the structure of hybrid perovskites: Analysis of charge migration in disordered MAPb1_structures. *New Journal of Physics*, *20*(10), 1-19. Article 103013. https://doi.org/10.1088/1367-2630/aae295

This material is protected by copyright and other intellectual property rights, and duplication or sale of all or part of any of the repository collections is not permitted, except that material may be duplicated by you for your research use or educational purposes in electronic or print form. You must obtain permission for any other use. Electronic or print copies may not be offered, whether for sale or otherwise to anyone who is not an authorised user.

PAPER • OPEN ACCESS

Multi-scale model for the structure of hybrid perovskites: analysis of charge migration in disordered MAPbl₃ structures

To cite this article: Jari Järvi et al 2018 New J. Phys. 20 103013

View the article online for updates and enhancements.



IOP ebooks[™]

Bringing you innovative digital publishing with leading voices to create your essential collection of books in STEM research.

Start exploring the collection - download the first chapter of every title for free.

New Journal of Physics

The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft **DPG IOP** Institute of Physics Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

PAPER

OPEN ACCESS

CrossMark

RECEIVED 14 June 2018

REVISED 6 September 2018

ACCEPTED FOR PUBLICATION

19 September 2018
PUBLISHED

10 October 2018

Original content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence.

Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.



Multi-scale model for the structure of hybrid perovskites: analysis of charge migration in disordered MAPbl₃ structures

Jari Järvi^{1,2}, Jingrui Li¹ and Patrick Rinke¹

¹ Department of Applied Physics, Aalto University, PO Box 11100, FI-00076 AALTO, Finland

 2 $\,$ Department of Physics, University of Helsinki, PO Box 64, FI-00014 University of Helsinki, Finland

E-mail: jari.jarvi@aalto.fi

Keywords: photovoltaics, solar cells, hybrid perovskites, multi-scale modeling, charge mobility, DFT

Abstract

We have developed a multi-scale model for organic–inorganic hybrid perovskites (HPs) that applies quantum mechanical (QM) calculations of small HP supercell models to large coarse-grained structures. With a mixed quantum–classical hopping model, we have studied the effects of cation disorder on charge mobilities in HPs, which is a key feature to optimize their photovoltaic performance. Our multi-scale model parametrizes the interaction between neighboring methylammonium cations (MA⁺) in the prototypical HP material, methylammonium lead triiodide (CH₃NH₃PbI₃, or MAPbI₃). For the charge mobility analysis with our hopping model, we solved the QM site-to-site hopping probabilities analytically and computed the nearest-neighbor electronic coupling energies from the band structure of MAPbI₃ with density-functional theory. We investigated the charge mobility in various MAPbI₃ supercell models of ordered and disordered MA⁺ cations. Our results indicate a structure-dependent mobility, in the range of 50–66 cm² V⁻¹ s⁻¹, with the highest observed in the ordered tetragonal phase.

1. Introduction

Hybrid organic–inorganic perovskites (HPs) are the primary candidates to substitute silicon as the cheap and efficient solar cell material of the future [1]. Methylammonium lead triiodide CH₃NH₃PbI₃ (shortened to MAPbI₃ hereafter) is the spearhead of this novel photovoltaic (PV) materials class with an unprecedented development in its power conversion efficiency (PCE), from ca. 10% in 2012 to the present 23%, approaching the best silicon-based single-junction devices. However, the most efficient HPs are currently unstable due to structural degradation in atmospheric water [2]. Conversely, stable HPs are still too inefficient in their PCE. New computational models, developed in parallel with experimental research, aim to improve the PCE of HPs via detailed understanding of the atomic structure and expedite the introduction of HPs into consumer-grade solar devices. While several mixed-halide HP compositions have recently entered the field, MAPbI₃ has maintained its status as the archetypal HP in the latest research.

 $MAPbI_3$ combines the key properties of a good PV material—a suitable band gap of ca. 1.6 eV [3], good visible-light photon absorption [4] and long diffusion length [5]. At present, the nature of the charge-carrier mobility in MAPbI_3 remains unclear. Recent studies suggest that the mobility is fairly modest in comparison to typical inorganic semiconductors, less than 100 cm² V⁻¹ s⁻¹ for electrons and holes [6, 7]. This indicates that the observed long diffusion length and the resulting high PV performance likely arise from a long charge-carrier lifetime. To optimize the charge mobility, we need detailed insight into the relationship between the mobility and the corresponding MAPbI_3 structure.

Recent research has focused on solving the atomistic origins of these properties, yet the structural complexity of MAPbI₃ and other HPs has hampered progress. Structural investigations have focused on the role of the organic methylammonium (MA⁺) cations, specifically the rotational dynamics of the cations and their interplay with the inorganic PbI₃ lattice [8–16]. Depending on the cation orientation, coupling with the lattice via hydrogen bonding distorts the framework geometry and affects the electronic properties, for example by altering

the nearest-neighbor (NN) atomic orbital overlap and the lattice vibrational modes. The exact relationship between this dynamical structure and the material's PV properties remains debated. Suggested mechanisms include for example charge-carrier pathways formed by ferroelectric domains of ordered MA⁺ cations [16], the effects of the lattice distortion on charge localization [17] and potential fluctuation effects on the charge-carrier mobility due to dynamical MA⁺ disorder [18]. However, detailed computational studies of these processes require electronically accurate models, which significantly restricts the size of the studied systems.

To describe properties that derive from the atomic and electronic structure with sufficient accuracy, we need to incorporate quantum mechanical (QM) information. Methods such as density-functional theory (DFT) [20] deliver this QM insight, but are limited to relatively small structures (<100 unit cells). However, due to the orientational freedom of the MA⁺ cations in each unit cell, HPs will be disordered and require modeling on much larger length scales than accessible by DFT. We approach this accuracy-size conundrum with a new multiscale model, which we have developed specifically for HP structures. Our model is based on a DFT parameterization of small supercell models and then applicable to large MAPbI₃ structures.

In our previous MAPbI₃ unit-cell energy-optimization study [13] we recognized two stable low-energy structures, corresponding to diagonal and face-to-face MA⁺ orientations in the unit cell. We then parametrized the MAPbI₃ atomic structure by studying the MA⁺s in neighboring unit cell pairs [14]. Each pair of the NN MA⁺s share a common face of the cubic unit cell. We focused on the MA⁺ orientations, describing them by a C–N dipole vector, and defined the dipole pair configurations as 'pair modes' (PMs hereafter, see section 2.1). We analyzed the effect of energy optimization on the PM distribution in 33 4 × 4 × 4 MAPbI₃ supercells. In the optimized structures, we observed an absence of the diagonal dipoles and a notable change in the PM distribution, for example an increase in the number of perpendicular PMs. As a good approximation, the observed changes were also independent of the initial structures before optimization. Moreover, it enabled us to generate new large-scale structures using the optimized PM distribution for further analyses, such as charge migration.

In this paper, we study the charge migration in large-scale MAPbI₃ structures with a semi-classical site-tosite hopping model. We generated the structures based on our previous observations in our supercell DFT studies, applying the PM description and a self-developed structure-generation algorithm. We then analyzed the effect of dipole orientations on the charge migration velocity, by comparing the generated structures to various reference systems, for example ordered dipole configurations (corresponding to tetragonal and orthorhombic phases), parallel and antiparallel dipoles and randomly oriented dipoles. We also analyzed the effect of structural defects (such as planar defects and precipitates) on the charge velocity. Finally, we estimated the charge mobility in our model based on an experimental reference value for the scattering time at room temperature [7].

This paper is organized as follows. In section 2 we review the PM concept and introduce methods for the MAPbI₃ structure generation and the charge migration study. In section 3 we introduce and discuss the results, and then conclude our findings in section 4.

2. Computational models and methods

To analyze the large-scale MA^+ cation order (or disorder), we generated new systems based on the PM distribution of the energy-optimized MAPbI₃ supercell models from our previous study [14]. We developed a semi-classical site-to-site hopping model to study charge migration in MAPbI₃ structures and compared the statistical migration velocities in various dipole configurations. In the following, we will first review the PM concept and then introduce the methods for the large-scale structure generation (see section 2.2) and the charge migration study (see section 2.3).

2.1. Pair mode description

We developed the PM concept based on our previous studies in MAPbI₃ structure optimization, using the allelectron numeric-atom-centered orbital DFT code FHI-aims [20–22]. In these studies we applied the Perdew– Burke–Ernzerhof (PBE) generalized gradient approximation [24] with Tkatchenko–Scheffler van der Waals (vdW) interactions [25] as the exchange-correlation functional. Based on our single-unit-cell structure optimization study [13], the MA⁺ cation has 32 discrete stable orientations in a unit cell—8 diagonal and 24 face-to-face with a small deviation angle to the lattice plane. In these low-energy states the ammonium end of the cation forms hydrogen bonds with the lattice I⁻ atoms, pulling them towards the cation. The amount of displacement is the largest in the face-to-face MA⁺ orientations, which subsequently has ca. 21 meV lower energy than the diagonal orientation. We have parametrized the structural energies of this framework distortion by describing the MA⁺ cation as a dipole vector in the C–N bond direction and studying the dipole configurations in neighboring unit cell pairs, the PMs (see figure 1(a)). The displacement of the I⁻ atom depends

2



Figure 1. (a) Coarse-granied model of a MAPD₃ structure, in which MA⁻ cations are modeled as C=N dipole vectors, and the PDI lattice deformation is parameterized into neighboring dipole pairs, the 'pair modes'. The figure shows a single lattice plane, in which the dipoles are oriented predominantly face-to-face in the unit cell, according to the optimized PM distribution (see figure 2), and the diagonal dipoles are rare. The out-of-plane dipoles are illustrated with a dot-circle (up) and a cross-circle (down) symbols. A solid diagonal arrow illustrates a diagonal dipole that points towards an upper corner in the unit cell. (b) $4 \times 4 \times 4$ supercell model of a DFT-optimized MAPbI₃ structure, showing the I displacement in the lattice. Pb, I, C, N and H atoms are colored in gray, purple, green, orange and white, respectively.



not only on the dipole orientation in a single unit cell, but also on the orientation in the neighboring cell. By considering the rotational and reflectional symmetries of the dipole pairs, we were able to reduce the amount of $32^2 = 1024$ pairs to only 86 unique PMs. Furthermore, by approximating the face-to-face orientations in each unit cell to be identical (i.e. four-fold degenerate without the lattice-plane deviation angle), the amount of unique PMs in this 14-orientation framework (with 8 diagonal and 6 face-to-face orientations) was reduced to 25.

In our recent study [14] we optimized $334 \times 4 \times 4$ MAPbI₃ supercells (see figure 1(b)) with DFT and studied the effect of optimization on the PM distribution (see figure 2). We observed that PMs 1–17, in which one or both dipoles are diagonal, are extremely rare in the optimized structures; much rarer than in an 'intrinsic' random distribution (see red bars in figure 2). This is an obvious consequence of the higher energy of the diagonal structure. Comparing our optimized PM distribution to then only the face-to-face 'intrinsic' PM probabilities of a random system (green data in figure 2), we observe a notable increase in the frequency of PMs 20 and 22. These correspond to perpendicular-aligned dipole pairs. Accordingly, the frequency of the other PMs decrease. To study the dipole order in MAPbI₃ on a large scale and its effect on the charge migration properties, we then generate new structures based on this optimized PM distribution.

2.2. Large-scale structure generation

We generated the structures using our own algorithm, in which the dipoles are reoriented one-by-one in a random order to reach the given PM distribution (the 'target' distribution, hereafter). In this study, the target distribution corresponds to the optimized PM distribution, shown in figure 2. The process is started from an initial state, in which the dipoles are oriented randomly in the 24 face-to-face orientations. These orientations correspond to the minimum-energy structures found in our previous study [13]. New dipole orientations are then determined by evaluating the change that any of the other orientations would induce in the PM distribution. By subtracting the present distribution from the target distribution, the algorithm calculates a difference value for each PM, which tells if the PM is too rare (positive value) or too frequent (negative value) in the present system. Each new dipole orientation is then evaluated by calculating the sum of the difference values for the 6 NNs. The most favorable orientation is the one that produces the largest sum, i.e. provides the largest increase in the deficient PMs. If a more favorable orientation is found, the dipole is rotated to the new orientation, after which the algorithm proceeds to the next random dipole. Via this process, the system evolves towards the target distribution, until a specified error threshold between the distributions has been reached. The error is calculated as the root mean square error (RMSE) of the PM frequency between the present and the target distribution. Due to a noisy variation in the present distribution, the threshold is evaluated with respect to the mean of the 10 previous RMSEs. In practice, the structure generation algorithm consists of an iteration loop, in which the following steps are executed.

- (i) Select a random dipole.
- (ii) Rotate the dipole through all 32 orientations to see, which one (if any) produces the most favorable change towards the target PM distribution.
- (iii) If another orientation is more favorable, reorient the dipole to the most favorable orientation.
- (iv) Calculate the present PM distribution.
- (v) Calculate the RMSE between the present and target distributions. Stop if the specified accuracy is reached.
- (vi) Return to (i).

In this process, we also analyzed the electrostatic dipole-dipole interaction energy of the structure. The energy is

$$U = -\sum_{i} \sum_{j \neq i} p_i \cdot E_{ij} = -\frac{1}{4\pi\epsilon_0 \epsilon_r} \sum_{i} \sum_{j \neq i} p_i \cdot \frac{3(p_j \cdot \hat{r}_{ij})\hat{r}_{ij} - p_j}{|r_{ij}|^3},$$
(1)

in which p_i and p_j are the electric dipole moments of dipoles *i* and *j*, E_{ij} is the electric field at dipole *i* by dipole *j*, r_{ij} is the distance vector from dipole *j* to dipole *i*, ϵ_0 is the vacuum permittivity and ϵ_r is the dielectric constant. The results of the structure generation are presented in section 3.1.

2.3. Hopping model for charge migration

We analyzed the effect of dipole order on the charge migration using a self-developed algorithm (see figures 3 and 4), in which a fully localized charge-carrier executes a classical random walk in a simulation box, based on PM-specific NN hopping probabilities. The QM behavior of the charge-carrier is evaluated at each hop based on an analytical solution of the hopping probabilities and the time step (see section 2.3.1), without the need to compute the time evolution of the QM wave-function explicitly. The analytical solution applies the PM-specific electronic coupling energies between the NNs, which are pre-computed transfer integrals (TIs) from the DFT band structure of MAPbI₃ (see section 2.3.2). Importantly, the analytical solution results in an extremely low computational simulation cost. This facilitates the calculation of a large number of migration paths in large systems. For example, in this study, each sample set consists of 10⁴ paths, which were calculated in systems of 10⁶ unit cells and above.

Each hopping path was started from a random initial location at the bottom plane (z = 0) of the system (see figure 4(a)). To increase the number of the paths that migrate through the system and not exit from the sides, the initial sites were selected from the middle half of the bottom plane (in range [d/4, 3d/4], in which d is the dimension of the system in x or y direction). The hopping was then executed (see figure 4(b)) until the path exited the system, from any of the 6 sides of the system (see figure 4(c)). To simulate the force of an external electric field on the charge-carrier, we applied a constant bias multiplier to the hopping probabilities in the direction of the electric field, which was set to +z. The multiplier increased the number of the paths that







migrated through the system, which was essential in obtaining a sufficiently large set of migration times for the statistical charge velocity. For this model, we first derived the analytical solution for the NN hopping probabilities (see section 2.3.1). The solution applies the electronic coupling energies V_n between the NNs, which we calculated from the DFT band structure of MAPbI₃ (see section 2.3.2).

2.3.1. Analytical solution for hopping probabilities

The NN hopping probabilities and the time step have an analytical QM solution (see appendix A for details). The Hamiltonian consists of the NN coupling energies V_n between the initial site 0 and the neighboring sites (i.e. the 6 NNs that share a common face of the cubic unit cell with the initial site) $n \in [1,6]$, such that

$$\hat{H} = \sum_{n=1}^{5} (|0\rangle V_n \langle n| + |n\rangle V_n \langle 0|).$$
⁽²⁾

J Järvi et al

We defined the squared sum of the coupling values as

$$V = \left(\sum_{n=1}^{6} V_n^2\right)^{1/2}.$$
 (3)

From these, we derived the time evolution of the wave function and the probabilities for the charge to be at the initial site or at a neighboring site. The probability for the charge to be at the initial site at time *t*, starting from a fully localized state of the wave function, is

$$P(0; t) = \cos^2\left(\frac{Vt}{\hbar}\right). \tag{4}$$

Importantly, the probability P(0; t) is a simple function of V and does not require solving the QM time evolution of the wave function explicitly in each hopping step, which is computationally expensive. As can be seen from the cosine function, the probability for the charge to be at the initial site starts to decrease as time progresses. Subsequently, the charge has a non-zero probability to exist at 7 different sites, i.e. at the initial site or at any of the 6 neighboring sites. Thus, we approximated that the charge will hop, on average, once its initial-site probability has decreased below 1/7. The hopping time is therefore

$$t_{\rm hop} = \frac{\hbar}{V} \cos^{-1} \left(\frac{1}{\sqrt{7}} \right). \tag{5}$$

At the time t_{hop} , the probability of the charge to be at site n is

$$P(n; t_{\rm hop}) = \left(\frac{V_n}{V}\right)^2 \sin^2\left(\frac{Vt_{\rm hop}}{\hbar}\right).$$
(6)

We then applied the bias multiplier to the probability of the neighbor n in the direction of the electric field and determined the site to which the charge hopped according to the thus-obtained probability distribution. Next, we will introduce the calculation of the electronic coupling energies V_n from the DFT band structure of MAPbI₃.

2.3.2. Electronic coupling between neighboring unit cells

The DFT band structure of MAPbI₃ (see figures 6(a) and C1) exhibits a cosine-type dispersion relation, which is known from the tight-binding (TB) theory [26]. Based on this, we adopted the TB model in the calculation of the electronic coupling between the neighboring unit cells. The coupling energies correspond to the TIs, from which we calculated the NN hopping probabilities in the charge-migration model via the analytical solution.

In the TB approximation, the wave function Ψ_k of a particle in a periodic lattice, in which k is the momentum quantum number, is written as a linear combination of the atomic wave functions ϕ_n , in which n denotes the site index. In the one-dimensional single-band TB model:

- (i) ϕ_n are considered to form a complete orthonormal set, i.e. $\langle \phi_n | \phi_{n'} \rangle = \delta_{nn'}$.
- (ii) All sites have the same on-site energy, i.e. $\langle \phi_n | \hat{H} | \phi_n \rangle = \text{constant}$. This constant can be set to 0 without loss of generality.
- (iii) Only the couplings between the NNs are non-zero, which we will denote by $-\Lambda$, i.e. $\langle \phi_n | \hat{H} | \phi_{n'} \rangle = -\Lambda \delta_{nn'\pm 1}$.

The one-dimensional TB Hamiltonian can thus be written as

$$\hat{H} = -\Lambda \sum_{n=1}^{\infty} (|\phi_n\rangle \langle \phi_{n+1}| + |\phi_{n+1}\rangle \langle \phi_n|).$$
(7)

In practice, we will work on finite systems, in which the total number of sites is *N*. For this, we will apply periodic boundary conditions by coupling the first and the last (*N*th) site with $-\Lambda$, such that

$$\hat{H} = -\Lambda \left[\sum_{n=1}^{N-1} (|\phi_n\rangle \langle \phi_{n+1}| + |\phi_{n+1}\rangle \langle \phi_n| + |\phi_N\rangle \langle \phi_1| + |\phi_1\rangle \langle \phi_N|) \right].$$
(8)

The solutions to this eigenvalue problem are Bloch waves Ψ_k , and the corresponding eigenvalues ε_k are cosine functions:

$$\Psi_k = \frac{1}{\sqrt{N}} \sum_{n=1}^{N} \exp(ikna)\phi_n \tag{9}$$

$$\varepsilon_k = -2\Lambda\cos(ka),\tag{10}$$

in which *a* is the distance between the NN sites. With (10), we can calculate the TIs as the coupling energies $-\Lambda$ from the amplitude of the conduction band in the DFT band structure of the material.

The effective mass m^* of the charge-carrier can be calculated from the curvature of the conduction band (see for example [27]) as

$$\frac{1}{m^*} = \frac{1}{\hbar^2} \frac{\mathrm{d}^2 \epsilon_k}{\mathrm{d}k^2}.$$
(11)

Using the cosine function of the energy bands (10) we get

$$m^* = \frac{\hbar^2}{2a^2\Lambda},\tag{12}$$

in which we have applied ka = 0 in the conduction-band minimum. With the effective mass, we can then calculate the mobility of the charge-carrier as

$$\mu = \frac{q\tau}{m^*} = \frac{2a^2\Lambda q\tau}{\hbar^2},\tag{13}$$

in which q is the charge of the charge-carrier and τ is the scattering time. The processes that affect the scattering time in MAPbI₃, for example the electron–phonon coupling, are outside the scope of our charge-migration model. Thus, we calculated the charge-carrier mobility based on an experimental reference value for the scattering time in the normal-temperature tetragonal phase. We obtained the mobilities for the other analyzed structures by scaling the computed velocities in our charge-migration model linearly in relation to the tetragonal-phase mobility. The results of this analysis are presented in section 3.3.

2.4. Computational details

We performed the band-structure DFT calculations with the all-electron numeric-atom-centered orbital code FHI-aims [20–23]. For all calculations we used tier 2 basis sets and the PBE exchange-correlation functional [24] augmented with long-range vdW corrections based on the Tkatchenko–Scheffler method of Hirshfeld electron-density partitioning [25]. In addition, scalar relativistic effects were included by means of the zero-order regular approximation [28]. A Γ -centered 4 × 4 × 4 k-point mesh was used for the supercell models a and b (see figure C1), and a 8 × 8 × 8 k-point mesh for the primitive-cell model c (see figure 6(a)). The results of the band-structure calculations are available from the Novel Materials Discovery (NoMaD) repository [29].

3. Results and discussion

3.1. Large-scale dipole structures

First, we studied the performance of the structure-generation algorithm by generating 50 structures of $20 \times 20 \times 20$ unit cells (see figure 5(b)). With the PM distribution of the 33 energy-optimized MAPbI₃ $4 \times 4 \times 4$ supercells (see figure 2) as the target distribution, we consider the generated structures to describe a low-energy MAPbI₃ structure on a large scale. In the process, we followed the electrostatic dipole–dipole interaction energy (1) within a cutoff radius of 3 unit cells (see figure 5(a)), using the experimental dielectric constant of 25.7 [30] and the dipole moment of 2.3 D [9]. Since the electrostatic dipole–dipole energy decreases with respect to the distance between the dipoles as r^{-3} , we consider the dipole–dipole interactions to be negligible in this study beyond the radius of 3 unit cells (ca.19 Å). We evaluated the convergence of the solution by following the RMSE between the present and the optimized PM distributions. The moderately small system size allowed us to update the PM distribution after each dipole reorientation. The RMSE was updated once in every 100 reorientations, and the system generation was terminated once the mean value of the 10 previous RMSEs decreased below 10^{-7} .

The results show that the optimized PM distribution is reached well within 10^4 generation steps (i.e. dipole reorientations). The electrostatic energy, which is initially ca. 0 in a system of randomly oriented dipoles, decreases in the generation to ca. -0.08 meV per unit cell.

The generated structures suggest that, instead of the polarized domains (see e.g. Frost *et al* [16]), the dipole order in MAPbI₃ on a large scale may predominantly consist of non-aligned dipoles, which can be in perpendicular formations, corresponding to PMs 20 and 22, and the out-of-plane PM 25. This resembles the tetragonal-phase MAPbI₃ structure (see for example [12]), which however is not static at normal temperatures due to dipole reorientations [8, 10]. The dynamic structure may exhibit the tetragonal-phase dipole order in local domains on a small scale, while remaining macroscopically disordered.



Figure 5. (a) Electrostatic dipole–dipole energy per unit cell and the root mean square error (RMSE) from the optimal pair mode distribution in the structure generation. The energy and RMSE are mean values of 50 generated $20 \times 20 \times 20$ structures, starting from initially random face-to-face dipole orientations, and (b) a representative *xy*-plane of the generated $20 \times 20 \times 20$ structures. Solid and outlined diagonal arrows illustrate diagonal dipoles that point towards an upper or lower corner in the unit cell, respectively. Parallel-aligned dipole configurations of length 2, 3 and 4 are highlighted in blue, green and yellow, respectively.



3.2. Transfer integrals

To obtain the NN hopping probabilities in the charge-migration model, we calculated the TIs from the DFT band structure of MAPbI₃. We analyzed the charge migration in structures, which consist of the face-to-face PMs 18–25, thus excluding the rare diagonal dipoles. We obtained the TIs for these PMs from the band structures of one primitive-cell model and two $2 \times 2 \times 2$ supercell models. We calculated the band structure of MAPbI₃ in the primitive-cell model (see figure 6(a)) using lattice parameters a = b = c = 6.313 Å [31]. In the relaxed structure, the dipole is oriented approximately in the +*y* direction. Here we use the usual notation $R \equiv \left(\frac{1}{2}, \frac{1}{2}, \frac{1}{2}\right)$ for the reciprocal lattice of a simple cubic crystal system, in which the conduction-band dispersion along $R-M_{x/y/z}$ reflects the effects of the inter-site coupling along x/y/z. As shown in figure 6(a), they all exhibit the typical cosine character and can be described with the one-dimensional TB theory. We can thus calculate the TI from the amplitude of the conduction band using(10). For example, the TI of the extending PM 18 can be calculated from the primitive-cell model along $R-M_y$ as



Figure 7. 5×5 unit-cell models of the basic structure types in the charge migration study. The models show a single *xy* plane in the structure, with charge-migration in the +z direction. A complete list of the structures is presented in figure B1.

$$\Lambda_y = \frac{|E_R - E_{M_y}|}{4},\tag{14}$$

in which E_R and E_{M_y} are the conduction band energies at points R and M_y , respectively. Also, in these points, ka = 0 and $ka = \pi \text{ in } (10)$, respectively. To obtain the TI for the parallel-aligned PM 23, we calculated the mean value of the TIs in the primitive-cell model along $R-M_x$ and $R-M_z$.

For the remaining PMs, we calculated the TIs from the DFT band structures of two cubic $2 \times 2 \times 2$ supercell models, which mimic the dipole orientations in the low-temperature orthorhombic phase and in the mid-temperature tetragonal phase (see figure C1). In these supercell models, we used the lattice constant of 2a = 12.626 Å. In both models, the conduction band minimum is located at $\Gamma \equiv (0, 0, 0)$, which is different from the primitive-cell model due to band folding. Instead of the common notation for the high-symmetry points in the Brillouin zone, we denote here $X \equiv (\frac{1}{2}, 0, 0)$, $Y \equiv (0, \frac{1}{2}, 0)$ and $Z \equiv (0, 0, \frac{1}{2})$.

From the conduction band dispersion along Γ –*X* and Γ –*Y*, we calculated the TIs associated with the perpendicular PMs 20 and 22. Since both the head-perpendicular and the tail-perpendicular PMs exist in both models and cannot be singled out in the band structure, we took the mean value of the calculated TIs and assigned the same value to both PMs. Finally, the antiparallel PM 24 and the out-of-plane PM 25 correspond to the conduction band dispersion along Γ –*Z* in models a and b, respectively. The opposite-aligned PMs 19 and 21 are extremely rare in the energy-optimized structures (see figure 2) and difficult to obtain via structural relaxation in DFT. Therefore, they were excluded from this TI analysis and assigned a coupling value of 0 eV in the charge-migration model. The TIs for all the calculated PMs are presented in figure 6(b).

3.3. Charge migration analysis

With our hopping model, we simulated the charge migration in various MAPbI₃ structures and analyzed the effect of dipole configurations on the charge mobility. We estimated the charge mobility based on the computed charge velocities, the effective mass of the charge-carrier calculated from the DFT band structure (see (12) and figure 6(a)), and an experimental reference value for the scattering time. Since we computed the velocities from the average charge-migration time through a finite-sized systems, they were dependent on the system size. Therefore, we first analyzed the velocities in systems of different sizes to determine the required size for a good approximation of a bulk material.

We generated dipole configurations of 10-130 unit cells with cubic system dimensions, which are analogous to (dis)ordered MAPbI₃ structures and known crystal phases in the coarse-grained model. The structures (see figures 7 and B1) included the following.

(i) The low-temperature orthorhombic phase in two different orientations with respect to the charge migration direction (i.e. the direction of the bias multiplier in the NN hopping probabilities).



- (ii) The mid-temperature tetragonal phase, in two different orientations, as above.
- (iii) Parallel and antiparallel dipoles in *z* and *x* directions, which correspond to the parallel and perpendicular orientations with respect to the charge migration direction, respectively.
- (iv) Randomly oriented dipoles.
- (v) Optimized dipole configurations, generated with the structure-generation method based on the optimized PM distribution (see section 2.2).

In all the charge-migration calculations, we set the bias multiplier to 5, which resulted in ca. 80% of the migration paths going through the system. Since the paths were started within the center half of the bottom plane, they rarely exited at the side-faces of the cubic system, particularly in large systems (see figure 8). Typically, a path was terminated due to exiting through the bottom plane during the first few hops in the migration. We computed 10⁴ paths in each of the structures, which based on test calculations on larger sample sets, provided a sufficient accuracy for the statistical velocity calculation.

The results (see figure 9(a)) show that in all the analyzed structures, a good approximation of a bulk system was obtained with a system dimension of ca. 100 unit cells or larger. We observed the highest velocities in the mid-temperature tetragonal phase in the *xy* direction. With respect to the charge migration in the +z direction, the tetragonal *xy* structure consists of the out-of-plane PMs 25, which have the strongest electronic coupling, i.e. the highest TI value (see figure 6(b)). Conversely, the tetragonal *xz* structures had the lowest computed velocities. This can be attributed to the low electronic coupling of the perpendicular PMs 20 and 22 in the direction of the charge migration. In this case, the highly coupled out-of-plane PMs in the *x* and *y* directions effectively deviate the charge sideways, which results in a prolonged migration time through the system and thus a low velocity. The same deviation effect can be seen in the parallel and antiparallel *z* structures, in which the higher coupling of the parallel PM 23 results in more deviation in the sideways direction and thus a lower velocity than in the antiparallel structure. With respect to the disordered structures, the velocities were slightly higher in the generated optimized structures than in the fully random structures. This likely results from the reduced number of the non-coupling PMs 19 and 21 in the optimized structures.

We analyzed the effect of structural defects on the charge-migration velocity by introducing volumes of randomly oriented dipoles in the tetragonal *xy* structures of $100 \times 100 \times 100$ unit cells. We studied the effect of two different types of defects: planar defects and precipitates (see figure 10(a)). In the planar defects, we introduced a system-wide plane in the middle of the system, such that the orientation of the plane (*xy*) is perpendicular to the charge migration direction (+*z*). We then varied the thickness of the plane and analyzed the charge velocity with respect to the defect volume in the system, i.e. $\widetilde{V}_{defect} / \widetilde{V}_{system}$. In the precipitates, we



Figure 9. (a) Charge migration velocities in various dipole configurations, showing the dependence of the velocity on the cubic system dimension, and (b) charge mobilities in the analyzed structures, calculated in relation to the reference mobility of the tetragonal *xy* structure.





introduced 25 randomly located non-overlapping cubic volumes of randomly oriented dipoles in the system. To reduce the noise that originates from the randomly located precipitates, we generated 20 systems of each defect concentration in 0%–20% volume range, and computed the velocities as their mean value.

The results (see figure 10(b)) show that the velocity decreases fairly linearly with defect concentration. By extrapolating the linear velocity decrease, which starts from the tetragonal *xy* phase velocity of ca. 7.4 Å fs⁻¹ at 0% defect concentration, the velocity approaches the fully random system with a velocity of ca. 5.8 Å fs⁻¹. The linear velocity decrease is qualitatively similar in both defect types, with the precipitates having slightly lower velocities.

Finally, we calculated the charge mobility in the analyzed structures, based on (i) the computed velocities in the largest computed structures of size $130 \times 130 \times 130$ unit cells, (ii) the effective mass of the charge-carrier, calculated from the tetragonal-phase DFT band-structure (12), and iii) an experimental reference value of the scattering time in the tetragonal-phase MAPbI₃. Karakus *et al* [7] have studied the electron–phonon scattering in MAPbI₃ with time-resolved terahertz spectroscopy, measuring a scattering time of ca. 4 fs in the tetragonal phase at 300 K. We associated this value with the best-performing tetragonal *xy* structure in our charge-

migration model, which consists of the out-of-plane PMs 25 with a TI of 0.895 eV. With these values and the lattice constant of a = 6.313Å, we estimated the best charge mobility in our model from (13) as

$$\mu_{\text{tet}} = \frac{2a^2 \Lambda |q_e|\tau}{\hbar^2} = 65.865 \text{ cm}^2 \text{ V}^{-1} \text{ s}^{-1},$$
(15)

in which q_e is the charge of an electron. The resulting mobility is more than twice the experimental reference mobility of ca. 27 cm² V⁻¹ s⁻¹, measured by Karakus *et al* [7]. The difference may arise e.g. from the calculation of the effective mass, which we have derived from the DFT band structure of the tetragonal-phase 2 × 2 × 2 supercell model with the TB approximation. In reality, the atomic structure of MAPbI₃ is likely disordered on a macroscopic scale at room temperature. Despite the simplified model in this study, the result is of the correct order of magnitude.

We obtained the mobilities for the other structures by relating the computed velocities to the tetragonalphase mobility in (15). First, we associated the tetragonal-phase mobility μ_{tet} with the computed velocity of the corresponding structure ν_{tet} . Then, we assigned mobilities for the other structures in relation

$$\mu_i = \mu_{\text{tet}} \frac{\nu_i}{\nu_{\text{tet}}},\tag{16}$$

in which μ_i is the charge mobility in structure *i*, and v_i is its computed velocity in the charge-migration model. The results (see figure 9(b)) show that the mobilities vary in a range of ca. 50–66 cm² V⁻¹ s⁻¹. The relative difference between the best- and worst-performing structures is ca. 30%. Also, the variation of 20 cm² V⁻¹ s⁻¹ is of the order of the experimentally measured charge mobility in the reference study by Karakus *et al.* The result suggests that the dipole order has a significant effect on the charge mobility in MAPbI₃. Interestingly, the highest and the lowest charge mobilities arise from the same tetragonal-phase structure, the difference being the direction of the charge migration. In MAPbI₃ preparation, correct orientation of the structure is therefore important to achieve optimal PV performance.

3.4. Computation time

Overall, the computation times in this study are extremely low, measured in seconds or minutes. In comparison, calculations with the conventional DFT methods can typically take hours or days. The structure-generation algorithm (see section 2.2) scales approximately as $O(n^2)$ with respect to the number of unit cells *n*. The scaling factor arises non-trivially from the way that the system size relates to reaching the target PM distribution within the set RMSE threshold. Since the PM distribution of the whole system is updated at regular intervals during the generation process, the computation time depends heavily on the system size. As an example, a system of a cubic dimension of 40 unit cells takes approximately 1 min to generate, whereas a system of a dimension of 50 unit cells requires ca. 3.5 min.

The charge-migration algorithm (see section 2.3) relies on the analytical solution for the NN hopping probabilities (see section 2.3.1). Due to this, the process is extremely rapid, and the system size affects directly only on the path length in the migration direction. In large systems, in which the overhead of initiating the process is negligible, the algorithm scales approximately linearly as O(n) with respect to the number of unit cells. For a given system, the charge-migration analysis can be executed in a fraction of the time that it takes to generate the system. For example, computing 10^5 migration paths in a system of a cubic dimension of 50 unit cells (with the optimized PM distribution) takes ca. 2 s.

The computation times presented here were calculated on a single core of a modern workstation processor at the time of this research.

4. Conclusion

In this work, we have investigated large-scale MAPbI₃ structures with a multi-scale model that applies the PM concept in the parametrization of the complex HP atomic structure. With the optimized PM distribution, we generated new coarse-grained MAPbI₃ structures to study the dipole configurations on a large scale. Different to the previously suggested polarized dipole order in MAPbI₃ [16], our results indicate that the structure may exhibit local tetragonal-phase domains at room temperature, while being disordered on a macroscopic scale due to reorienting dipoles.

We applied our coarse-grained model in the charge-migration study of MAPbI₃ with a semi-classical hopping model. We estimated the average charge-carrier velocities in various (dis)ordered dipole configurations, applying our analytical solution for the QM hopping probabilities and the electronic NN coupling energies from the DFT band-structure. We related the computed velocities to charge mobilities via an experimental reference value for the scattering time and the effective mass calculated from the DFT band-structure.

We obtained charge mobilities in range of 50–66 cm² V⁻¹ s⁻¹, which are comparable to the experimental mobility of 27–150 cm² V⁻¹ s⁻¹ [7]. Importantly, both the highest and the lowest mobilities were obtained in the same tetragonal phase, the difference being the charge-migration direction (i.e. in the tetragonal-phase structures *xy* and *xz*, respectively). The relative variation of ca. 30% in the charge mobility suggests that the orientation of the dipoles has a substantial effect on the charge mobility in MAPbI₃. In the defected tetragonal-phase MAPbI₃ structures we observed a linear velocity relation with respect to the defect concentration, with little difference between the planar defects and the precipitates.

With our model, we have aimed to advance the present understanding on the order of the organic cations in HPs and their effect on the charge mobility. In addition to MAPbI₃, our model is applicable to other HP compositions with an appropriate analysis of the PMs of the organic cations. Thus, the model provides a cost-efficient method to analyze the cation order and the charge mobility in numerous HPs. With the acquired knowledge, new HP compositions can be offered for experimental research, and modifications to the present materials can be suggested to boost their performance.

Acknowledgments

We gratefully thank H. Levard for insightful discussions in our collaboration. We acknowledge the computing resources by the CSC-IT Center for Science (via the Project No. ay6311) and the Aalto Science-IT project. An award of computer time was provided by the Innovative and Novel Computational Impact on Theory and Experiment (INCITE) program. Computing resources were also provided by the Argonne Leadership Computing Facility, which is a DOE Office of Science User Facility supported under Contract DE-AC02-06CH11357. This research was supported by the Academy of Finland through its Centres of Excellence Programme (2012–2014 and 2015–2017) under project numbers 251748 and 284621, as well as its Key Project Funding scheme under project number 305632.

Appendix A. Analytical solution for hopping probabilities

In the charge-migration model, the probability of a charge-carrier to hop to a neighboring site is based on the QM behavior of the particle. In the QM formalism, the state of a system is described by a complex wave function $\Psi(t)$ in Hilbert space. The time evolution of the wave-function can then be determined via the time-dependent Schrödinger equation (TDSE)

$$\hat{H}\Psi(\mathbf{r},t) = i\hbar\partial_t\Psi(\mathbf{r},t). \tag{A.1}$$

In the charge-migration model, the immediate location of the particle after each hop is known with full certainty, which in the QM formalism corresponds to a fully localized wave function. As time progresses, the wave function expands to the neighboring sites in real space, according to its time-evolution via the TDSE (A.1). In the following, we will show that when considering the hopping probabilities to the NN sites, the wave function resides in a two-state Hilbert space. The corresponding two-state system has a known analytical solution, from which the hopping probabilities can be determined without explicitly calculating the time-evolution of the wave function in each hop. This is extremely valuable for the computational efficiency of the charge-migration model.

First, we construct a localized basis set

$$S = \{|0\rangle; |n\rangle\}, \{n \in \mathbb{Z} \colon 1 \ge n \ge 6\},\tag{A.2}$$

in which states $|0\rangle$ and $|n\rangle$ correspond to the charge being located at the initial site or at any of the 6 neighboring sites, respectively (see figure A1). Here we consider that these 7 states are orthonormal and form a complete set, so that the Hamiltonian and the time-dependent wave function $\Psi(t)$ can be represented in terms of them:

$$\langle 0|n\rangle = 0 \ \forall n \in [1, 6], \tag{A.3}$$

$$\langle n|n'\rangle = \delta_{nn'} \ \forall n, n' \in [1, 6]. \tag{A.4}$$

The Hamiltonian consists of the NN electronic coupling energies V_n between the initial site 0 and the neighboring sites $n \in [1, 6]$, such that

$$\hat{H} = \sum_{n=1}^{6} (|0\rangle V_n \langle n| + |n\rangle V_n \langle 0|).$$
(A.5)

Since we have a freedom of choosing an appropriate basis to represent the system, we can construct a new delocalized basis as a superposition of the localized states $\{|0\rangle; |n\rangle\}$. The new basis set (also orthonormal) is

$$S' = \{ |0\rangle; |\phi\rangle, |\varphi_k\rangle \}, \{ k \in \mathbb{Z} : 1 \ge k \ge 5 \},$$
(A.6)



in which the basis states are

$$|\phi\rangle = \sum_{n=1}^{6} |n\rangle \langle n|\phi\rangle \equiv \sum_{n=1}^{6} \alpha_n |n\rangle, \, \alpha_n \in \mathbb{R},$$
(A.7)

$$|\varphi_k\rangle = \sum_{n=1}^6 u_{kn} |n\rangle, \, u_{kn} \in \mathbb{R}.$$
 (A.8)

We define the coefficients α_n to be

$$\alpha_n = \frac{V_n}{V}, \ V \equiv \left(\sum_{n'=1}^6 V_{n'}^2\right)^{1/2}, \tag{A.9}$$

such that

$$\sum_{n=1}^{6} \alpha_n^2 = \sum_{n=1}^{6} \frac{V_n^2}{\sum_{n'=1}^{6} V_{n'}^2} = 1.$$
(A.10)

State $|n\rangle$ in the delocalized basis is

$$|n\rangle = |\phi\rangle\langle\phi|n\rangle + \sum_{k=1}^{5} |\varphi_k\rangle\langle\varphi_k|n\rangle = \alpha_n |\phi\rangle + \sum_{k=1}^{5} u_{nk} |\varphi_k\rangle,$$
(A.11)

and then

$$\hat{H} = \sum_{n=1}^{6} \left(\left| 0 \right\rangle V_n \left\langle \phi | \alpha_n + \sum_{k=1}^{5} | 0 \right\rangle V_n \left\langle \varphi_k | u_{nk} + \alpha_n | \phi \right\rangle V_n \left\langle 0 | + \sum_{k=1}^{5} u_{nk} | \varphi_k \right\rangle V_n \left\langle 0 | \right) \right)$$

$$= \sum_{n=1}^{6} V_n \left[\alpha_n (| 0 \rangle \left\langle \phi | + | \phi \rangle \left\langle 0 | \right) + \sum_{k=1}^{5} u_{nk} (| 0 \rangle \left\langle \varphi_k | + | \varphi_k \right\rangle \left\langle 0 | \right) \right]$$

$$= V \sum_{n=1}^{6} \left[\alpha_n^2 (| 0 \rangle \left\langle \phi | + | \phi \rangle \left\langle 0 | \right) + \sum_{k=1}^{5} \alpha_n u_{nk} (| 0 \rangle \left\langle \varphi_k | + | \varphi_k \right\rangle \left\langle 0 | \right) \right].$$
(A.12)

In the first term, the sum over α_n^2 equals 1 (A.10). The second term vanishes due to the orthogonality of the states

$$\sum_{n=1}^{6} \sum_{k=1}^{5} \alpha_n u_{nk} = \sum_{n=1}^{6} \sum_{k=1}^{5} \langle \varphi_k | n \rangle \langle n | \phi \rangle = \sum_{k=1}^{5} \langle \varphi_k | \phi \rangle = 0.$$
(A.13)

Therefore, \hat{H} in the delocalized basis consists of only two states

$$\hat{H} = V(|0\rangle \langle \phi| + |\phi\rangle \langle 0|). \tag{A.14}$$

The eigenstates of \hat{H} are

$$|\chi_1\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |\phi\rangle), \tag{A.15}$$

J Järvi et al

$$|\chi_2\rangle = \frac{1}{\sqrt{2}}(|0\rangle - |\phi\rangle) \tag{A.16}$$

and their corresponding eigenvalues E_1, E_2 are

$$\hat{H}|\chi_1\rangle = \frac{1}{\sqrt{2}}V(|0\rangle\langle\phi| + |\phi\rangle\langle0|)(|0\rangle + |\phi\rangle) = V|\chi_1\rangle \to E_1 = V,$$
(A.17)

$$\hat{H}|\chi_2\rangle = \frac{1}{\sqrt{2}}V(|0\rangle\langle\phi|+|\phi\rangle\langle0|)(|0\rangle-|\phi\rangle) = -V|\chi_2\rangle \to E_2 = -V.$$
(A.18)

Also, the states $|0\rangle$ and $|\phi\rangle$ can be written in the energy eigenbasis as

$$|0\rangle = |\chi_1\rangle\langle\chi_1|0\rangle + |\chi_2\rangle\langle\chi_2|0\rangle = \frac{1}{\sqrt{2}}(|\chi_1\rangle + |\chi_2\rangle), \tag{A.19}$$

$$|\phi\rangle = |\chi_1\rangle\langle\chi_1|\phi\rangle + |\chi_2\rangle\langle\chi_2|\phi\rangle = \frac{1}{\sqrt{2}}(|\chi_1\rangle - |\chi_2\rangle).$$
(A.20)

Next, we calculate the time evolution of the quantum state of the system, described by $|\Psi(t)\rangle$. Initially, at time t = 0, the charge is fully localized at the initial site, which corresponds to the state $|\Psi(0)\rangle = |0\rangle$. The time evolution of the system is determined by the time evolution operator $\hat{U}(t) = \exp(-i\hat{H}t/\hbar)$, so that

$$\begin{split} |\Psi(t)\rangle &= \hat{U}(t)|\Psi(0)\rangle = \frac{1}{\sqrt{2}} \exp(-i\hat{H}t/\hbar)(|\chi_1\rangle + |\chi_2\rangle) \\ &= \frac{1}{\sqrt{2}} (\exp(-iE_1t/\hbar)|\chi_1\rangle + \exp(-iE_2t/\hbar)|\chi_2\rangle) \\ &= \frac{1}{\sqrt{2}} (\exp(-iVt/\hbar)|\chi_1\rangle + \exp(iVt/\hbar)|\chi_2\rangle). \end{split}$$
(A.21)

In general, the probability *P* of a QM system to be in a state $|\xi\rangle$ is calculated as the squared norm of the wave function amplitude

$$P(\xi; t) = |\langle \xi | \Psi(t) \rangle|^2. \tag{A.22}$$

In our case, the amplitude of the wave function at the initial site is

$$\langle 0|\Psi(t)\rangle = \frac{1}{2} \langle 0|[\exp(-iVt/\hbar)(|0\rangle + |\phi\rangle) + \exp(iVt/\hbar)(|0\rangle - |\phi\rangle)]$$

= $\frac{1}{2} (\exp(iVt/\hbar) + \exp(-iVt/\hbar)) = \cos\left(\frac{Vt}{\hbar}\right).$ (A.23)

The probability of the charge to be located at the initial site is then

$$P(0; t) = |\langle 0|\Psi(t)\rangle|^2 = \cos^2\left(\frac{Vt}{\hbar}\right).$$
(A.24)

Similarly, we get the probability of the charge to be located at any of the neighboring sites n as

$$P(n; t) = |\langle n|\Psi(t)\rangle|^{2} = \left|\frac{1}{2}\langle n|\left[\exp(-iVt/\hbar)\left(\left|0\rangle + \sum_{n'=1}^{6}|n'\rangle\langle n'|\phi\rangle\right\rangle\right)\right] + \exp(iVt/\hbar)\left(\left|0\rangle - \sum_{n'=1}^{6}|n'\rangle\langle n'|\phi\rangle\right\rangle\right)\right]\right|^{2}$$
$$= \left|\frac{1}{2}[\exp(-iVt/\hbar) - \exp(iVt/\hbar)]\sum_{n'=1}^{6}\langle n'|\phi\rangle\frac{\langle n|n'\rangle}{\delta_{nn'}}\right|^{2}$$
$$= \alpha_{n}^{2}\sin^{2}\left(\frac{Vt}{\hbar}\right) = \left(\frac{V_{n}}{V}\right)^{2}\sin^{2}\left(\frac{Vt}{\hbar}\right), \qquad (A.25)$$

in which $|\phi\rangle$ has been written in the localized basis according to (A.7). From this result we see that the probability of the charge to be located at the initial site, or any of the neighboring sites, oscillates at frequency V/\hbar . The

initial-site probability P(0; t) (A.24) is 1 at time t = 0, which corresponds to a fully localized charge at the initial site. The probability then starts to decay according to the squared cosine function, as the wave function propagates to the neighboring sites. Concurrently, site *n* probabilities P(n; t) (A.25) increase according to the squared sine function, so that the higher the coupling energy V_n is, the more rapid is the increase.

Due to the probabilistic nature of the hopping, we cannot precisely determine the time period that the charge spends at each site. However, we can approximate the average time t_{hop} that it takes for the charge to hop to another site, based on the fact that the charge has 7 possible sites to be located at. Therefore, on average, we approximate that the charge will hop once its probability to be located at the initial site has decreased below 1/7. The hopping time can be solved from (A.24) as

$$t_{\rm hop} = \frac{\hbar}{V} \cos^{-1} \left(\frac{1}{\sqrt{7}} \right). \tag{A.26}$$

The probability of the charge to be located at site *n* at the time of a hop can then be calculated from (A.25) as $P(n; t_{hop}) \equiv P_n$. Based on these probabilities, we can then determine the site that the charge hops into. For this, we first form a cumulative array with values

$$\left\{0, P_1, P_1 + P_2, \dots, \sum_{i=1}^{6} P_i\right\},\tag{A.27}$$

and normalize the array into range]0, 1[. Then, we pick a random number $\{r \in \mathbb{R}: 0 > r > 1\}$ and see, between which values it lands in the array. This then determines the site that the charge hops into, by correctly accounting for different probabilities at different sites. Finally, after each hop, the simulation time is advanced by the time step t_{hop} .

Appendix B. Coarse-grained structures





Appendix C. DFT band structures of MAPbI₃



Figure C1. Band structures of MAPbI₃ calculated with the cubic $2 \times 2 \times 2$ supercell models: (a) model a, and (b) model b, which are analogous to the low-temperature orthogonal and mid-temperature tetragonal phase, respectively. For each model, the dipole orientations at different *xy* planes are given as insets. The cosine-like conduction-band dispersions along Γ -*X*, Γ -*Y* and Γ -*Z* are colored in red, blue and green, respectively.

References

- [1] Green M A, Ho-Baillie A and Snaith H J 2014 The emergence of perovskite solar cells Nat. Photon. 8 506
- [2] Berhe T A, Su W-N, Chen C-H, Pan C-J, Cheng J-H, Chen H-M, Tsai M-C, Chen L-Y, Dubale A A and Hwang B-J 2016 Organometal halide perovskite solar cells: degradation and stability *Energy Environ. Sci.* 9 323–56
- [3] Yamada Y, Nakamura T, Endo M, Wakamiya A and Kanemitsu Y 2014 Near-band-edge optical responses of solution-processed organic–inorganic hybrid perovskite CH₃NH₃PbI₃ on mesoporous TiO₂ electrodes Appl. Phys. Express 7 032302
- [4] De Wolf S, Holovsky J, Moon S-J, Löper P, Niesen B, Ledinsky M, Haug F-J, Yum J-H and Ballif C 2014 Organometallic halide perovskites: sharp optical absorption edge and its relation to photovoltaic performance J. Phys. Chem. Lett. 5 1035–9
- [5] Dong Q, Fang Y, Shao Y, Mulligan P, Qiu J, Cao L and Huang J 2015 Electron–hole diffusion lengths >175 μm in solution-grown CH₃NH₃PbI₃ single crystals *Science* 347 967–70
- [6] Brenner T M, Egger D A, Rappe A M, Kronik L, Hodes G and Cahen D 2015 Are mobilities in hybrid organicinorganic halide perovskites actually high? J. Phys. Chem. Lett. 6 4754–7
- [7] Karakus M, Jensen S A, D'Angelo F, Turchinovich D, Bonn M and Cánovas E 2015 Phonon-electron scattering limits free charge mobility in methylammonium lead iodide perovskites J. Phys. Chem. Lett. 6 4991–6
- [8] Mosconi E, Quarti C, Ivanovska T, Ruani G and de Angelis F 2014 Structural and electronic properties of organo-halide lead perovskites: a combined ir-spectroscopy and *ab initio* molecular dynamics investigation *Phys. Chem. Chem. Phys.* 16 16137
- [9] Frost J M, Butler K T, Brivio F, Hendon C H, van Schilfgaarde M and Walsh A 2014 Atomistic origins of high-performance in hybrid halide perovskite solar cells *Nano Lett.* 14 2584

- [10] Chen T, Foley B J, Ipek B, Tyagi M, Copley J R D, Brown C M, Choi J J and Lee S-H 2015 Rotational dynamics of organic cations in the CH₃NH₃PbI₃ perovskite Phys. Chem. Chem. Phys. 17 31278–86
- [11] Leguy A M A et al 2015 The dynamics of methylammonium ions in hybrid organic-inorganic perovskite solar cells Nat. Commun. 67124
- [12] Lahnsteiner J, Kresse G, Kumar A, Sarma D D, Franchini C and Bokdam M 2016 Room-temperature dynamic correlation between methylammonium molecules in lead-iodine based perovskites: an *ab initio* molecular dynamics perspective *Phys. Rev.* B 94 214114
- [13] Li J and Rinke P 2016 Atomic structure of metal-halide perovskites from first principles: the chicken-and-egg paradox of the organicinorganic interaction Phys. Rev. B 94 045201
- [14] Li J, Järvi J and Rinke P 2018 Multiscale model for disordered hybrid perovskites: the concept of organic cation pair modes Phys. Rev. B 98 045201
- [15] Li J et al 2018 Activation energy of organic cation rotation in CH₃NH₃PbI₃ and CD₃NH₃PbI₃: quasi-elastic neutron scattering measurements and first-principles analysis including nuclear quantum effects J. Phys. Chem. Lett. 9 3969–77
- [16] Frost J M, Butler K T and Walsh A 2014 Molecular ferroelectric contributions to anomalous hysteresis in hybrid perovskite solar cells APL Mater. 2 081506
- [17] Kang B and Biswas K 2017 Preferential CH₃NH₃⁺ alignment and octahedral tilting affect charge localization in cubic phase CH₃NH₃PbI₃ *J. Phys. Chem. C* 121 8319–26
- [18] Ma J and Wang L-W 2017 The nature of electron mobility in hybrid perovskite CH₃NH₃PbI₃ Nano Lett. 17 3646–54
- [19] Hohenberg P and Kohn W 1964 Inhomogeneous electron gas Phys. Rev. 136 B864-71
- [20] Blum V, Gehrke R, Hanke F, Havu P, Havu V, Ren X, Reuter K and Scheffler M 2009 Ab initio molecular simulations with numeric atom-centered orbitals *Comput. Phys. Commun.* 180 2175–96
- [21] Havu V, Blum V, Havu P and Scheffler M 2009 Efficient integration for all-electron electronic structure calculation using numeric basis functions J. Comput. Phys. 228 8367–79
- [22] Levchenko S V, Ren X, Wieferink J, Johanni R, Rinke P, Blum V and Scheffler M 2015 Hybrid functionals for large periodic systems in an all-electron, numeric atom-centered basis framework *Comput. Phys. Commun.* **192** 60–9
- [23] Ren X, Rinke P, Blum V, Wieferink J, Tkatchenko A, Andrea S, Reuter K, Blum V and Scheffler M 2012 Resolution-of-identity approach to Hartree–Fock, hybrid density functionals, rpa, mp2, and gw with numeric atom-centered orbital basis functions New J. Phys. 14 053020
- [24] Perdew J P, Burke K and Ernzerhof M 1996 Generalized gradient approximation made simple Phys. Rev. Lett. 77 3865-8
- [25] Tkatchenko A and Scheffler M 2009 Accurate molecular van der waals interactions from ground-state electron density and free-atom reference data Phys. Rev. Lett. 102 073005
- [26] Slater J C and Koster G F 1954 Simplified LCAO method for the periodic potential problem Phys. Rev. 94 1498–524
- [27] Oberhofer H, Reuter K and Blumberger J 2017 Charge transport in molecular materials: an assessment of computational methods Chem. Rev. 117 10319–57
- [28] van Lenthe E, Baerends E J and Snijders J G 1993 Relativistic regular two-component Hamiltonians J. Chem. Phys. 99 4597-610
- [29] Novel Materials Discovery (NoMaD) repository see http://doi.org/10.17172/NOMAD/2018.06.08-1
- [30] Brivio F, Walker A B and Walsh A 2013 Structural and electronic properties of hybrid perovskites for high-efficiency thin-film photovoltaics from first-principles APL Mater. 1 042111
- [31] Stoumpos C C, Malliakas C D and Kanatzidis M G 2013 Semiconducting tin and lead iodide perovskites with organic cations: Phase transitions, high mobilities, and near-infrared photoluminescent properties *Inorg. Chem.* **52** 9019