PCCP

ROYAL SOCIETY OF CHEMISTRY

PAPER

Published on 17 March 2025
Licence and permissions

Check for updates

Cite this: *Phys. Chem. Chem. Phys.*, 2025, **27**, 7053

Determining charge transport regimes in organic molecular crystals: a machine learning framework†

T. S. A. Cassiano, [a] M. L. Pereira Junior, [b] P. H. de Oliveira Neto *[ac] and L. A. Ribeiro Junior [ad]

Received 1st November 2024,
Accepted 20th February 2025

DOI: 10.1039/d4cp04185b

rsc.li/pccp

Charge transport performance in organic molecular crystals (OMCs) is crucial for advancements in nanotechnology. Experiments have shown metallic-like and semiconducting charge transport regimes in OMCs, mediated by free electrons with Bloch-like oscillations (BOs) and polaronic states. In metallic-like regimes, the charge propagates as a wave, while in semiconducting regimes, it travels as a quasi-particle coupling charge with a cloud of lattice phonons. While the conditions for polaronic states in OMCs are well-established, those enabling BOs still need to be understood. In this study, we identify the electronic and structural properties of OMCs that favor the formation of polarons or BOs by analyzing their linear and wave transport properties. We employ semiempirical non-adiabatic dynamical simulations at the picosecond scale and machine learning methods to map the parameter spaces where the charge transport occurs *via* polarons or BOs. The dynamical simulations are based on a general model Hamiltonian developed to address OMCs. Our results reveal that increasing the electronic transfer rate between molecules, the crystal's speed of sound, and its metallicity favors the formation of BOs. BOs in OMCs can exhibit frequencies around 2 THz and current amplitudes up to 3000 $|e|$ ps^{-1}, opening up possibilities for high-frequency applications. Conversely, large polarons are predominantly formed based on the interplay between intra- and intermolecular electron–lattice interactions.

1 Introduction

Organic molecular crystals (OMCs) are one of the flagships in nanoscience, enabling the development of crucial optoelectronic technologies such as organic photovoltaic cells (OPVs),[1,2] organic light-emitting diodes (OLEDs),[3–5] and others.[1,6,7] Much of OMC's appeal resides in its broad electronic and mechanical properties. This diversity arises from the numerous possible compositions that can be used to create new OMCs. Changes in the molecular units and the overall crystalline structure (packing and site orientation) can result in distinct optical/electronic responses. The matter is particularly relevant for charge transport, a defining mechanism for the performance of optoelectronic technologies.

In general, charge transport strongly depends on confinement effects, doping, and impurities.[1,8,9] Substantial effort was devoted to developing theoretical methods to investigate this mechanism. The use of *ab initio* formalisms is especially prevalent.[10–14] For instance, combining density functional theory (DFT) with the Boltzmann transport equation has resulted in accurate carrier predictions on crystals.[15,16] A recent report presented a reliable DFT-based description of InAlN nanorods' electronic properties showing accuracy that is competitive with many-body GW methods.[17] Model Hamiltonians represent an alternative approach. They consist of a combination of simplified interactions to describe crystal physics. In this context, hybrid tight-binding formulations, which contain electronic and lattice components, display notable success in modeling complex phenomena such as quasiparticle formation,[18–24] transport,[8,25–31] and collision mechanisms.[32–35]

Charge migration in organics can unravel over multiple transport regimes. Two of them retain special importance: the polaronic and Bloch oscillation (BO) types. The former consists of self-trapped quasiparticles that propagate, in many aspects, as projectiles.[9,26,27,36] BOs, in turn, exhibit a wave-like

[a] *Institute of Physics, University of Brasília, 70910-900 Brasília, Brazil. E-mail: pedrohenrique@unb.br*
[b] *College of Technology, Department of Electrical Engineering, University of Brasília, 70910-900 Brasília, Brazil*
[c] *University of Brasília, International Center of Physics, 70919-970, Brazil*
[d] *Computational Materials Laboratory, LCCMat, Institute of Physics, University of Brasília, 70910-900 Brasília, Brazil*

† Electronic supplementary information (ESI) available: Proof of the expression used for current calculation (Section 1), details of the machine learning approach, including the autoencoder architecture (Section 2), FFT mean frequency pair plot (Section 3), charge current amplitude pair plot (Section 4), and DC pair plot (Section 5). See DOI: https://doi.org/10.1039/d4cp04185b

propagation due to the crystal translation symmetry.[37,38] Substantial interest in BOs emerged[38–44] after being successfully confirmed on superlattices.[45,46] Harnessing the mechanism could allow the development of fast emitters with tunable frequencies.[47,48] Additionally, the BO concept was exported to photonics, becoming the core of promising optical devices.[49–51]

Despite the advances, the conditions that enable BOs on typical OMCs remain an open topic. Currently, some challenges obstruct this development. The first one is the number of parameters that characterize crystals. Accounting for the different molecular species, their electronic/mechanical properties, and intermolecular interactions substantially increase the parameter space dimension. Consequently, a general approach that deals with BOs on multiple crystalline configurations would result in a large and potentially intractable dataset. Another issue lies in identifying BOs through a reliable and scalable protocol. Although easily recognizable *via* visual confirmation, these states are hard to label through numerical metrics because the wave/polaron profiles can deeply change depending on the crystal characteristics.

Recent progress in machine learning (ML) methods suggests potential routes to overcome these challenges. Employing unsupervised cluster learning is becoming an established strategy for processing high-dimension datasets in quantum chemistry. The analysis consists of grouping data occurrences into sets (clusters) based on proximity criteria, potentially revealing hidden correlations between features and target properties.[52–56] In the same way, substantial improvements have been made in image-processing ML algorithms over the last decades. Now, complex and robust methods such as the convolutional neural network (CNN) can aptly extract complex features from images while condensing them on compact representations.[57,58] In this scenario, integrating quantum chemistry simulations with ML methods presents a promising strategy for investigating charge transport regimes in a general OMC.

In this work, we determine the physical conditions that favor the formation of Bloch oscillations on typical OMCs. The protocol combines the simulation of over 13 000 non-adiabatic charge transport dynamics over different crystal configurations with a machine learning workflow. The ML procedure processes images from the dynamics using CNN. Then, the unsupervised clustering method groups the images into two classes: polaronic or BO. Close analysis of the Bloch cluster reveals which OMC's mechanical and electronic properties enable the oscillating motion. Moreover, we report the presence of rich and well-defined wave patterns with characteristic frequencies. The influence of the electric field strength is also assessed, showing its defining role in shaping the oscillations. In addition, the work proposes general synthesis guide rules to favor BOs, while providing a general (and extendable) protocol to investigate quasiparticle transport mechanisms through machine learning methods.

2 Methodology

Model Hamiltonian

Molecular crystals are described under the Holstein–Peierls model. The Hamiltonian (H) is the combination of intramolecular (H_{intra}) and intermolecular (H_{inter}) contributions.[59–61] Both intramolecular and intermolecular components contain electronic and lattice terms. Fig. 1(a) illustrates a 2D molecular crystal under this formalism. The molecular sites are indexed in the molecular matrix according to their row and column positions. Each site is linked to four neighboring sites by intermolecular bonds. For instance, the molecule at position (i, j) is connected to the sites $(i-1,j)$, $(i+1,j)$, $(i1,j-1)$, and $(i1,j+1)$. The bonds between them are modeled under the harmonic approximation, with a lattice mass M_{inter} and spring constants $K^{x/y}_{\text{inter}}$. It is important to note that each bond can be distorted, relative to the uniform configuration, due to site rearrangement. The resulting intermolecular lattice energy is

$$H^{\text{latt}}_{\text{inter}} = H^{\text{latt},x}_{\text{inter}} + H^{\text{latt},y}_{\text{inter}}, \tag{1}$$

such that, for instance,

$$H^{\text{latt},x}_{\text{inter}} = \frac{K^{x}_{\text{inter}}}{2}\sum_{i,j}\left(\Delta x_{i,j}\right)^2 + \frac{M_{\text{inter}}}{2}\sum_{i}\left(\dot{v}^{x}_{i,j}\right)^2. \tag{2}$$

Here, $\Delta x_{i1,j} \equiv v^{x}_{i+11,j} - v^{x}_{i1,j}$ is the bond deformation connecting the sites $(i1,j)$ and $(i+11,j)$, and the difference between the change in the position of the sites $(i1,j)$ and $(i+11,j)$. In an analogous definition, $\Delta y_{i1,j}$ and $v^{y}_{i1,j}$ are the equivalent variables for the y-direction intermolecular lattice component, $H^{\text{latt},y}_{\text{inter}}$. We present an illustration of $v^{x}_{i1,j}$ in Fig. 1(b).

The electronic transfer mechanism is modeled according to the tight-binding approximation, where each site is connected

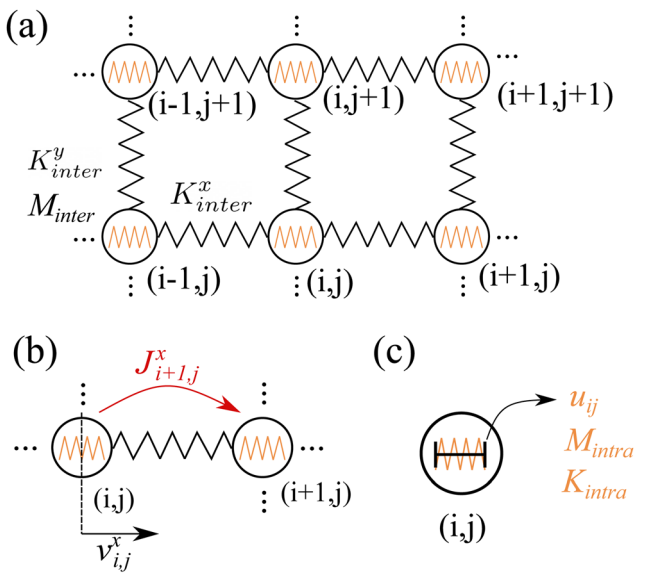

Fig. 1 Illustration of a 2D crystal within the Holstein–Peierls model. (a) A slice of the lattice composed of molecular sites linked through harmonic springs representing intermolecular bonds. The molecules are labeled according to their row and column position in the molecular matrix. (b) Two adjacent molecules are highlighted. The x-direction hopping integral $J_{i+1,j}$ is illustrated, along with the deformation of the intermolecular position of the site (i,j) in the x axis, $v^{x}_{i,j}$. (c) The intramolecular variables, modeled as an internal harmonic oscillator.

 This journal is © the Owner Societies 2025

with its first neighbors through a hopping interaction:

$$H_{\text{inter}}^{\text{elec}} = \sum_{i,j}\left(J_{i+1,j}^{x}C_{i+1,j}^{\dagger}C_{i,j} + \text{H.C.}\right) + \sum_{i,j}\left(J_{i,j+1}^{y}C_{i,j+1}^{\dagger}C_{i,j} + \text{H.C.}\right), \tag{3}$$

in which $C_{i1,j}^{\dagger}$ and $C_{i1,j}$ are, respectively, the creation and annihilation operators of the lowest unoccupied molecular orbital (LUMO) located in the site $(i1,j)$. Moreover, $J_{i+11,j}^{x}$ is the hopping integral linking the molecules $(i1,j)$ and $(i+11,j)$, as depicted in Fig. 1(b). $J_{i1,j+1}^{y}$ is the equivalent hopping term for the y direction.

It is important to note that a fixed hopping integral alone is not adequate to realistically describe organic materials. For this reason, it is important to explicitly account for the spatial dependence of $J_{i+11,j}^{x}$ and $J_{i1,j+1}^{y}$. This adaptation can be done through a first-order expansion over the bond distortion:[9,18,26,27,60,62]

$$\begin{aligned} J_{i+1,j}^{x} &= J_{0}^{x} - \alpha_{\text{inter}}^{x}\Delta x_{i,j}, \\ J_{i,j+1}^{y} &= J_{0}^{y} - \alpha_{\text{inter}}^{y}\Delta y_{i,j}, \end{aligned} \tag{4}$$

where $J_{0}^{x/y}$ is the hopping integral for the pristine configuration and $\alpha_{\text{inter}}^{x/y}$ is the intermolecular electron–phonon constant. This scaling parameter controls the interplay between lattice and electronic interactions.

The intramolecular variables are represented in Fig. 1(c). Each site $(i1,j)$ has its internal lattice degree of freedom, $u_{i1,j}$, mass, M_{intra}, and harmonic spring constant K_{intra}. The corresponding lattice component for H_{intra} is

$$H_{\text{intra}}^{\text{latt}} = \frac{M_{\text{intra}}}{2}\sum_{i,j}\dot{u}_{i,j}{}^{2} + \frac{K_{\text{intra}}}{2}\sum_{i,j}u_{i,j}{}^{2}, \tag{5}$$

while the electronic term is, under the second quantization formalism, given by

$$H_{\text{intra}}^{\text{elec}} = \sum_{i,j}\alpha_{\text{intra}}u_{i,j}\left(C_{i,j}^{\dagger}C_{i,j}\right), \tag{6}$$

in which α_{intra} is the intramolecular electron–phonon constant.

We emphasize that the electronic and lattice interactions are coupled. Therefore, one cannot directly diagonalize the electronic Hamiltonian $H_{\text{elec}} = H_{\text{intra}}^{\text{elec}} + H_{\text{inter}}^{\text{elec}}$. Its numerical evaluation requires prior knowledge of the lattice degrees of freedom. Then, a self-consistent algorithm has to be employed to obtain the optimal configuration. This method consists of solving both the lattice and electronic parts interactively. Here, we adopted the backpropagation (RPROP)[63] method for this task. Its application returns the stationary set of eigenvectors, $|\psi_0\rangle$, of the electronic Hamiltonian and lattice configuration, $\{u_{i1,j}\}$ and $\{v_{i1,j}^{x/y}\}$.

The dynamics of the charged states are simulated by placing the converged stationary states under the influence of an external electric field. Here, we include this effect through the Peierls substitution with a convenient choice: $A_x(t) = -cE_0t$, where $A_x(t)$ is the projection of the potential vector along the crystal (x) chain and E_0 is the electric field strength. As a result, the hopping integral gains a time-dependent factor:[26,27,62]

$$J_{i+11,j}^{x} \rightarrow \mathrm{e}^{i\gamma A_x(t)}J_{i+11,j}^{x}, \tag{7}$$

in which $\gamma \equiv ea_0/c\hbar$ and a_0 is the lattice parameter.

With this modification, the electronic Hamiltonian becomes time-dependent, such that the time evolution of the electronic states follows the time-dependent Schröedinger equation. The updated $|\psi(t)\rangle$, after a time increment $\mathrm{d}t$, is

$$|\psi(t+\mathrm{d}t)\rangle = \mathrm{e}^{-iH_{\text{elec}}(t)\mathrm{d}t/\hbar}|\psi(t)\rangle. \tag{8}$$

It can be shown[64] that the projection of this state at the site $(i1,j)$, $\langle i1,j|\psi(t+\mathrm{d}t)\rangle \equiv \psi(i1,j,t+\mathrm{d}t)$, evolves as

$$\psi(i,j,t+\mathrm{d}t) = \sum_{l,m,n}\left(\mathrm{e}^{-i\varepsilon_l(t)\mathrm{d}t/\hbar}\phi_l(m,n,t)\psi(m,n,t)\phi_l(i,j,t)\right), \tag{9}$$

where $\{\varepsilon_l(t)\}$ and $\{\phi_l(i1,j,t)\}$ are, respectively, the instantaneous electronic eigenenergies and the $(i1,j)$ projection of the l-th eigenvector. Applying eqn (9) allows time progression of the electronic state by evaluating the change at each component $(i1,j)$.

Due to the coupling, the lattice and electronic parts have to be evolved simultaneously. The lattice evolves according to the Euler–Lagrange equations of the Lagrangian expected value. Explicit application of the expression yields equations of motion. They dictate the change in time for the lattice degrees of freedom. The intramolecular motion equations are

$$M_{\text{intra}}\ddot{u}_{i1,j}(t) = -K_{\text{intra}}u_{i1,j}(t) - \alpha_{\text{intra}}\rho_{ij;ij}, \tag{10}$$

where $\rho_{i1,j;k,l}(t) \equiv \psi(i1,j,t)\psi(k,l,t)^{*}$ is the molecular charge density matrix.

The motion equations of the intermolecular lattice deformation are the sum of the lattice $(F_{i1,j}^{x/y,\text{latt}})$ and electronic forces $(F_{i1,j}^{x/y,\text{elec}})$:

$$M_{\text{inter}}\ddot{v}_{i1,j}^{x/y}(t) = F_{i1,j}^{x/y,\text{latt}} + F_{i1,j}^{x/y,\text{elec}}. \tag{11}$$

The explicit form of the x-direction components are

$$\begin{aligned} F_{i,j}^{x,\text{latt}} &= -K_{\text{inter}}\left(2v_{i,j}^{x}(t) - v_{i+1,j}^{x}(t) - v_{i-1,j}^{x}(t)\right), \\ F_{i,j}^{x,\text{elec}} &= -\alpha_{\text{inter}}\left(\left[\rho_{i,j;i-1,j}(t) - \rho_{i+1,j;i,j}(t)\right]\mathrm{e}^{-i\gamma A_x(t)} + \text{c.c.}\right), \end{aligned} \tag{12}$$

while the y-direction motion follows

$$\begin{aligned} F_{i,j}^{y,\text{latt}} &= -K_{\text{inter}}\left(2v_{i,j}^{y}(t) - v_{i,j+1}^{y}(t) - v_{i,j-1}^{y}(t)\right), \\ &\text{and,} \\ F_{i,j}^{y,\text{elec}} &= -\alpha_{\text{inter}}\left(\left[\rho_{i,j;i,j-1}(t) - \rho_{i,j+1;i,j}(t)\right] + \text{c.c.}\right). \end{aligned} \tag{13}$$

The lattice motion is evolved at each time step $\mathrm{d}t$ through a Verlet integration method.

It should be noted that the electronic Hamiltonian has to be diagonalized after each time step progression. The reason is

because the operator depends on the lattice variables. The resulting set of eigenvectors ($\{|\phi_l(t)\rangle\}$) of the updated H_{elec} are then applied in the electronic state time evolution through eqn (9).

The description of the charge transport mechanism primarily focuses on modeling events that occur along the transport line. In other words, a good representation of the fundamental physical properties requires considering only the sites along the direction of the electric field. Therefore, we will investigate charge dynamics in the x direction.

In the present study, the fixed Hamiltonian parameters are set to represent a typical organic crystal,[59–61] that is N = 100, $M_{\text{intra}} = 1.00 \times 10^{10}$ eV (as Å)$^{-2}$, $M_{\text{inter}} = 2M_{\text{intra}}$, K_{intra} = 5.00 eV Å^{-2}, a_0 = 3.50 Å, and α_{intra} = 0.75 eV Å^{-1}.

Charge current density

Characterizing BO is a paramount step in studying the properties of the Hamiltonian parameter space. To this end, metrics must be applied in the simulation to extract the characteristic features of each configuration for further analysis. Here, we calculated the charge current density from the site $(i1,j)$ to $(i+1,j)$. After rearrangement of the discrete continuity equation, the charge density current can be written as[38]

$$j_{i,j}(t) = -\frac{2|e|}{\hbar}\mathrm{Im}\Big(H_{i,j;i+1,j}\rho_{i+1,j;i,j}(t)\Big), \tag{14}$$

such that $H_{i1,j;i'1,j'} = \langle i1,j|H|i'1,j'\rangle$ is the Hamiltonian element in the localized basis. The total current density, along a particular line j, is the sum of all contributions along the row of N sites:

$$j_j(t) = \sum_i^N j_{i,j}, \tag{15}$$

and the corresponding DC component is the time-average of j. In other words,

$$\mathrm{DC} = \frac{1}{\tau}\int \mathrm{d}t j_j(t), \tag{16}$$

where τ is the simulation time window, which in our case is five ps. The ESI† demonstrates eqn (14).

Machine learning workflow

Within the Holstein–Peierls model, polaron and Bloch solutions can emerge depending on the Hamiltonian parametrization. We implement an unsupervised machine learning train over the dynamical solutions to label the different simulations. The first step is to build a dataset of dynamical simulations over district parametrizations. Configurations are chosen based on the sampling of uniform distributions for K_{inter}, J_0^x, and α_{inter}. Each set is evaluated in five electric field values: 0.3, 0.5, 1.0, 2.0, 3.0, and 5.0 mV Å^{-1}. The $\rho_{ij;ij}(t)$ heatmap is calculated, representing the crystal dynamics. The features of this image will determine whether the solution exhibits polaron or Bloch characteristics. In other words, the classification process is based on the charge density heatmaps.

An autoencoder pipeline processes the heatmaps *via* a convolutional neural network approach. Its task is to store the image information in a more appropriate form for cluster training. This means capturing characteristic patterns/shapes that constitute the image as features.[65] That is a much preferable approach than, for example, pixel-by-pixel treatment because it accounts for the image's collective properties while reducing the data dimensionality. The ESI† contains more details of the autoencoder neural network architecture. Next, we implement an unsupervised clustering train over the data through the agglomerative clustering algorithm.[66] The method consists of a bottom-up clustering hierarchical machine learning algorithm. In other words, each charge transport dynamics is seen as a single cluster. Then, the merging of two sets is made depending on the corresponding variance *via* the Ward linkage.[66,67] Analysis shows two predominant clusters of transport regimes: wave propagation (Bloch-like) and projectile drift (polaron-like).

Fig. 2 summarizes the overall workflow through three steps: (1) build the dataset; (2) process the images *via* the CNN algorithm and implement the unsupervised cluster train; (3) analyze the clusters. The first one is represented by four figures. These are the charge density heatmaps of different crystal intermolecular configurations. In the second step, the processed images are fed into a machine-learning pipeline to categorize the dynamics. The ML categorization process is an unsupervised training based on clustering. The final step consists of analyzing the sets returned from the training. The objective is to identify the common physical properties within the clusters. Note that the methodology combines two separate ML techniques (CNN + unsupervised clustering) to filter out the different charge transport regimes, constituting an integrated machine-learning workflow.

Normally, when the dataset contains a high degree of complexity, forcing a clusterization with a low number of clusters, even if they are motivated by physical insight, they might deliver underperformed separations. Here, we experienced this very limitation when directly forcing a separation into two large clusters. To resolve this issue, we progressively increased the size of the clusters until self-consistency inside each set was attested.

Specifically, we apply clustering to divide the dataset into several subclusters. Then, we look over each generated subset to check if its elements (within a given subset) shared characteristic patterns (*e.g.* wave shape or polaronic motion). Next, we manually labeled each subcluster as Bloch or polaron. Because of this clustering subdivision, the search for consistency became viable since we only needed to spot visual discrepancies over clear patterns within each cluster. Attested that a given subset is internally consistent, all the dynamics inside of it received the same transport-type label. While we did not look at every case, it became possible to cover a significant portion of the parameter space through this approach, allowing a broad assessment of the clustering quality. With this method, we could benefit from the clustering features without restricting to only two subsets (which can have an underperforming

 This journal is © the Owner Societies 2025

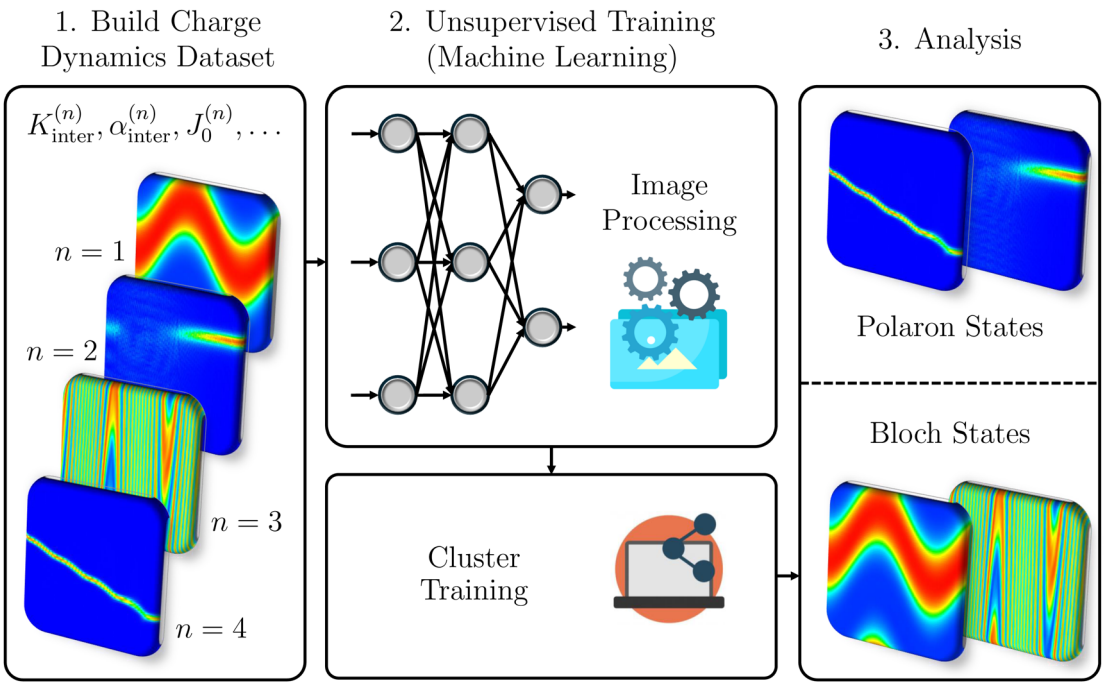

Fig. 2 Framework used to filter the charge transport regimes of a general molecular crystal. Initially (step 1), multiple charged crystal configurations are simulated under the influence of an external electric field. Subsequently, charge density heatmaps are calculated, with each configuration having its corresponding heatmap image. Next (step 2), an autoencoder pipeline processes these images. The data are then subjected to unsupervised clustering. Finally (step 3), the clusters are analyzed to identify the crystal properties that differentiate BOs from polaron-type solutions.

separation), and the subdivision aided the quality check of the model. In the present case, we found that 25 subclusters led to an appropriate division, allowing self-consistency within the subsets.

The choice of parameters' ranges is based on the collective consideration of multiple works devoted to describing specific crystals. In general, parametrization is made using *ab initio* calculations.[68–71] The transfer integrals are estimated from band structure analysis, while electron–phonon constants are obtained from the numerical derivatives of the integral with respect to nuclear coordinates.[68] When this approach is demanding or the object of study is a group of crystals, the parameters are not tightly specified, but a minor range is set in which transport is stable.[59,64] It is important to realize that aiming for a general description of crystals requires wider parameter ranges. Otherwise, there is a risk of particularizing to a specific configuration. In other words, we aimed to simultaneously have a broad range with physically sound values.

We report that the parameter sampling was limited to a subset of the 4-dimensional parameter space. This zone was determined through an initial assessment, which involved simulating crystals with extreme parameter variations. The objective was to roughly locate the regions where the non-oscillating and oscillating solutions may intersect. We opted for this route instead of an unrestricted variation to ease the computational burden of building and processing the dataset. Table 1 presents the intervals determined after the initial assessment. The cluster training was done over $\approx$13 000 simulations within these intervals.

As can be seen, all the values are consistent with the magnitudes for typical crystals.[59–61] The parameters K_{inter}, J_0^x, and α_{inter} were sampled from the uniform distributions delimited by the Table 1 ranges. We emphasize that the sampling is over a continuous set, not requiring a discrete representation of the parameter space. The exception is the electric field strength, which we sampled over a discrete array of values. The E_0 sampling values were fixed to reduce numerical instability during the clustering train process. Despite the discrete nature of the sampling, the wide range of field magnitudes ensures a good representation of molecular crystals, covering relevant regimes associated with realistic physical scenarios.

The database was constructed by keeping the Hamiltonian parameters M_{intra}, M_{inter}, K_{intra}, and α_{intra} fixed, as previously mentioned. We adopted this approach after preliminary calculations indicated that intramolecular variables had minimal effects on the emergence of Bloch oscillations (BO). Additionally, the lattice parameter is kept constant. This decision is

Table 1 Parameters' intervals used to build the dataset. All parameters except E_0 were sampled over a uniform distribution limited by these intervals. Each crystal configuration was evaluated at all electric field strength values shown in this table

Parameter	Interval
K_{inter} (eV Å^{-2})	[0.1–100.00]
J_0^x (eV)	[0.01–0.50]
α_{inter} (eV Å^{-1})	[0.001–5.0]
E_0 (mV Å^{-1})	[0.3, 0.5, 1.0, 2.0, 3.0, 5.0]

justified because a_0 appears explicitly in the model only through eqn (7). The same applies to the electric field strength. Therefore, the parameter controlling the variation in $J^x_{i+1,i}$ is the product E_0a_0. Thus, increasing either E_0 or a_0 should produce the same effect. Here, we chose to fix a_0 and increase E_0, as this change is more readily interpretable in the current context.

Appropriate representation of the images is paramount for meaningful clustering. In a general approach, the CNN architecture has to be carefully tuned to ensure that the key elements of the figures remain intact. It is important to realize that the role of CNN is to represent the data. Therefore, as long as an appropriate representation is assured, further optimizations are, although desirable, not essential. Here, by trial and error, we found that the structure detailed in the ESI† was enough to process the charge dynamics. Validation was made by checking the convergence of the loss function to standard thresholds and the appearance of constructed images of selected cases. The convergence evolution and the constructed images are shown in, respectively, Fig. S2 and S3 (ESI†). Naturally, another architecture may deliver similar or even better performance. Therefore, generally, tuning the CNN hyperparameters through a systematic approach is a recommended strategy to ensure a good representation and extrapolation. Here, we moved to a more practical approach, showing that insight extraction is possible without relying on highly tuned layer structures since appropriate representation was attested. This relaxed constraint eases the computational burden of applying this method while keeping its usefulness.

The following discussions will address different crystal parameter sets using vector notation. Specifically, K_{inter}, J^x_0, α_{inter}, and E_0 will be represented as directions in a 4-dimensional parameter space. As a result, a given crystal configuration will be mapped to a vector $\vec{c}$ with components:

$$\vec{c} = (K_{\text{inter}}, J^x_0, \alpha_{\text{inter}}, E_0), \tag{17}$$

where all components follow the units adopted in Table 1.

3 Results & discussion

Visualizing clusters is crucial for validating the machine learning model. Dimensionality reduction analysis is a valuable tool for this purpose, as it identifies a parameter subspace where the clusters are distinguishable. For this purpose, we use the t-distributed stochastic neighbor embedding (TSNE) method[54,66,72] to present the data in a 2D space spawned by two components: TSNE1 and TSNE2. These are non-linear combinations of the latent space, composed of the features of the CNN image training basis. Fig. 3(a) displays the dataset in this representation, with orange points corresponding to configurations (specific parameter sets of H) associated with oscillating solutions and blue points indicating polaron-like behavior.

Two distinct zones are visible in the TSNE plot, revealing that different transport regimes populate separate regions in the parameter space. Fig. 3(b) and (c) display representative charge density dynamic heatmaps for Bloch and polaron solutions, respectively. Hot colors correspond to high charge accumulation, while cold tones indicate low density. In the BOs example, the charge pulse does not travel indefinitely in the field direction; instead, the carrier velocity progressively shifts, causing it to move in the opposite direction, resulting in spatial oscillation. It is important to note that the simulations include periodic boundary conditions.

Naturally, the frontier separating the zones in Fig. 3(a) is not absolute. There are some orange points entering the dominantly blue zone and *vice versa*. This trend occurs because the transition from charge carrier to oscillating motion is smooth, with many intermediate states exhibiting BOs and polaron behavior characteristics. In other words, these states do not exactly fit into a binary clustering. This limitation is an inherent trade-off when clustering divides the parameter space into discrete regions. Nonetheless, the results confirm a zone with a higher probability of finding BOs. Therefore, the clustering procedure is validated.

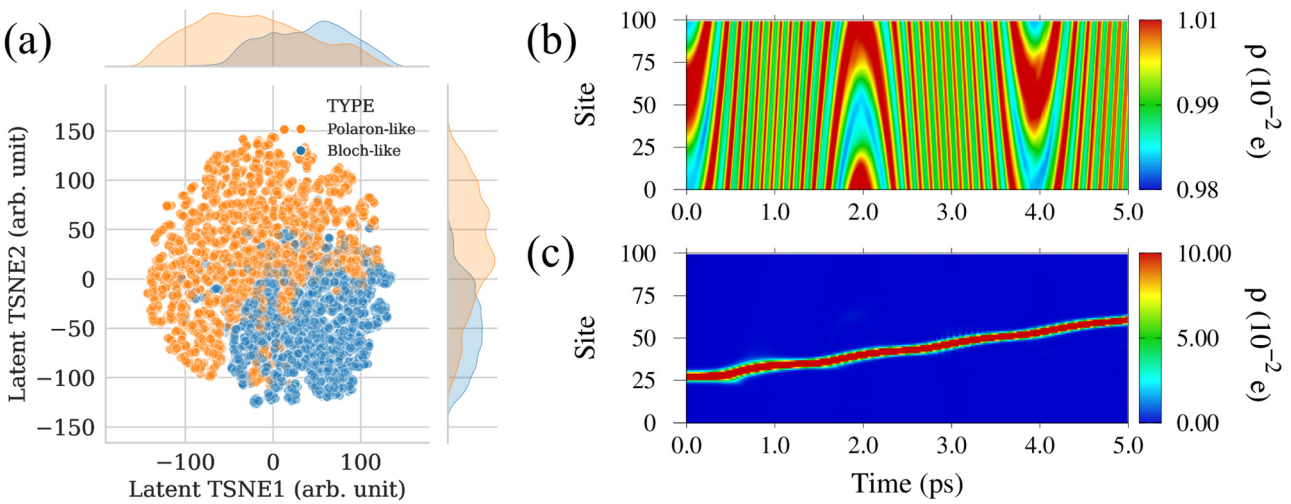

Fig. 3 Results of the unsupervised training. Panel (a) presents the trained data in a subspace generated by the t-distributed stochastic neighbor embedding (TSNE) method. Configurations exhibiting wave-like solutions are shown in orange, while those behaving as projectiles are shown in blue. Panels (b) and (c) display charge density heatmaps, representing the dynamics of BOs and polarons, respectively.

 This journal is © the Owner Societies 2025

We emphasize that the overlap at the parameter space frontier is not the result of an underperforming CNN structure. Instead, it is a mixed motion in which the binary classification fails due to the absence of a clear predominant pattern. Therefore, even if the autoencoder had perfectly represented the data, this issue would remain because it is an intrinsic transport feature. In other words, the purely Bloch or polaronic cases benefit the most from the filtering scheme. The clustering applied in this work falls into this case, as shown in Fig. 3(a), validating the methods. Moreover, our study focuses on identifying the physical conditions behind typical BO-like and polaronic states. Transient cases, although retaining interest from a theoretical standpoint, are not the main focus of this study.

The TSNE decomposition reveals well-defined zones for Bloch-like and polaron-like states. However, the technique does not provide a direct interpretation of the physical conditions that constitute the TSNE subspace basis. For deeper insights, Fig. 4 displays the pairplots of K_{inter}, J_0^x, and α_{inter}. This figure consists of a matrix of scatter plots, with each element representing a projection in a 2D space formed by a pair of parameters. The row and column position of the non-diagonal plots determine the parameters pair. Therefore, the upper and the lower triangles contain the same information. Additionally, the diagonal plots show the parameter distributions segregated by solution type. Points in orange represent dynamical simulations that were recognized as Bloch-like solutions, while points in blue indicate polaron-like solutions.

Certain projections reveal frontiers that separate the transport regimes. As guides for the eye, free-hand red lines are added to highlight these intersection areas for better visualization. First, consider the α_{inter} plot as a function of J_0^x. Here, an inclined line separates the blue and orange regions. This

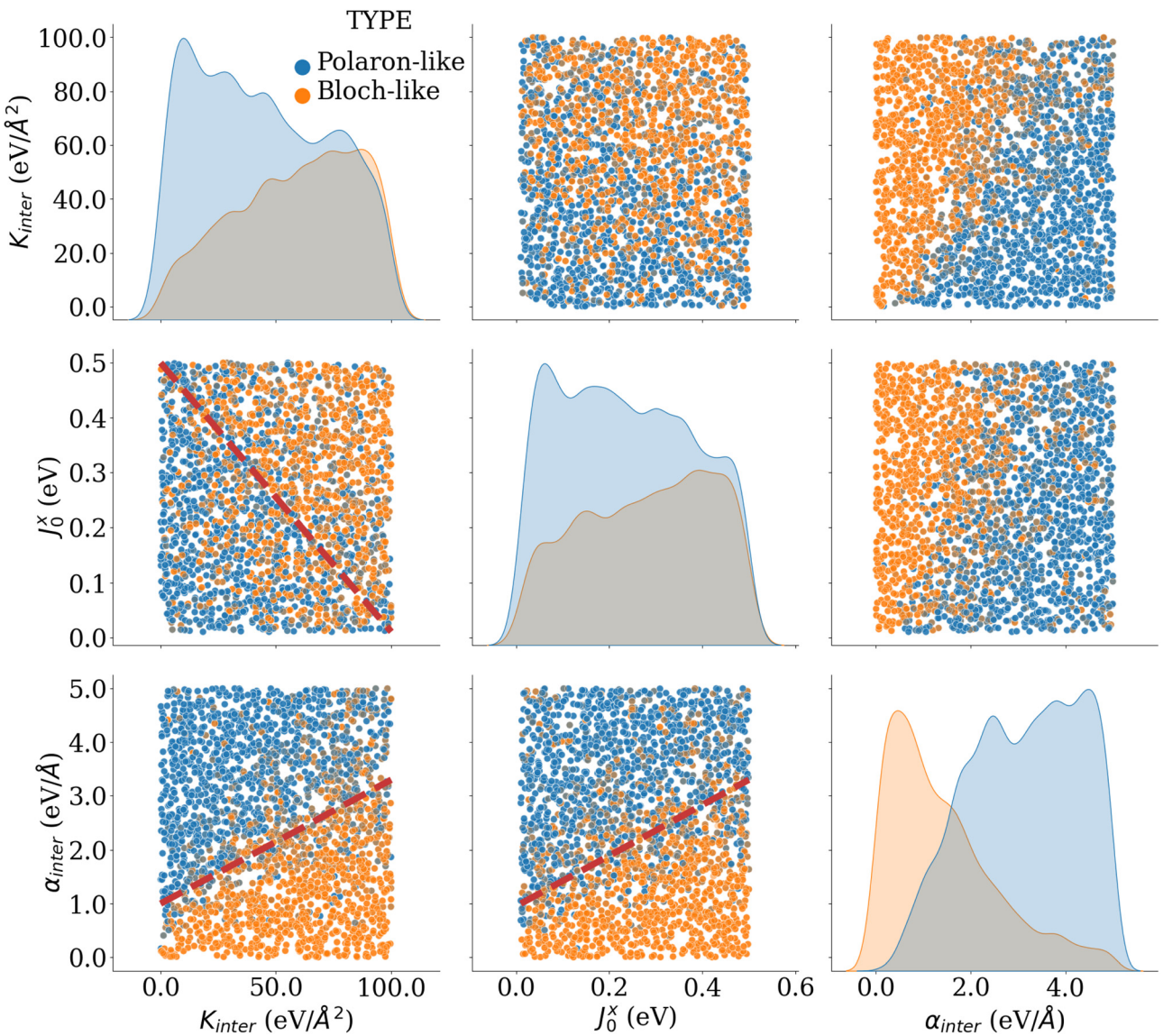

Fig. 4 Scatter pairplot of K_{inter}, J_0^x, and α_{inter}. Crystal configurations in blue refer to polaronic motion, while those in orange exhibit BO. The diagonal plots are the probability distributions of each parameter for the Bloch and polaron dynamics.

intersection line suggests that the region with oscillating dynamics is roughly a triangle in this projected space. Therefore, increasing α_{inter} induces the transition from wave-like to polaron behavior. However, the bigger J_0^x, the higher the α_{inter} threshold connecting the two transport regimes. In other words, J_0^x facilitates BO solutions. This result can be understood by considering how BOs emerge. Wave-like propagation arises when there is significant dissociation between electronic and lattice interactions. This kind of disruption mitigates potential lattice-related scattering effects, enabling the emergence of the wave pattern. Additionally, when there is this semi-separation, electrons can move beyond the mechanical acoustic limit of phonons.

This defining effect of the hopping integral is expected, as its magnitude calibrates the strength of the electronic interaction. Therefore, increasing J_0^x might shift the balance of lattice and electronic energies, effectively facilitating dissociation. Conversely, a substantial α_{inter} does the opposite. The coupling parameter enforces the connection between the two mechanisms. Thus, a higher α_{inter} means a stronger coupling. When the coupling is strong enough, the self-trapping nature of polaronic defects becomes energetically favorable.

The intermolecular spring constant plays a similar role of J_0^x in aiding the formation of BO when increasing α_{inter}. This effect is evident in the α_{inter} pair plot as a function of K_{inter}. Here, an orange zone resembling a triangle is also present. Therefore, raising K_{inter} can also limit the formation of polarons for strong couplings. Note that K_{inter} controls the intermolecular lattice interaction strength. Therefore, raising K_{inter} may make it more challenging to rupture the two mechanisms. That is an unexpected result since K_{inter} is often associated with aiding charge localization in quasiparticle formation. This finding highlights the complex relationship between the Hamiltonian parameters. It evidences the need for careful analysis to determine the parameter zones associated with each solution type.

The diagonal plots offer additional insights. First, consider the distribution of K_{inter}. Bloch-like and polaron-like solutions span the range of 0–100 eV Å^{-2}, with a complete overlap. However, the polaron and BO-like peaks are located in opposite regions. The polaron peak is $\approx$20 eV Å^{-2}, while BO maximum occurrence is around 100 eV Å^{-2}. The hopping integral follows a similar pattern. This result is anticipated considering that their corresponding pair plots with α_{inter} yielded equivalent zones.

In contrast with the previous distributions, there is a strong separation between the two transport regimes for the electron–phonon constant. The overlapping is notably smaller. Centering in $\alpha_{inter} \approx 0.5$ eV Å^{-1}, the BO curve is sharper compared to the K_{inter} and J_0^x distributions. At the other extreme, when $\alpha_{inter} \approx 4.0$ eV Å^{-1}, the occurrence of BOs is negligible, while the polaron distribution is at its peak. This response suggests that α_{inter} has a stronger influence on determining whether the transport will exhibit wave-like or polaronic behavior.

Our findings show that the two solutions have clear zones in the parameter space. Therefore, rationalized synthesis routes could be devised to obtain molecular crystals that favor (or not) Bloch oscillations. The results reveal that minimizing α_{inter} effectively achieves this. We recall that the electron–phonon coupling can be associated with the metal-insulator Peierls transition.[9] Therefore, our results imply that molecular crystals exhibiting metallic behavior are more likely to show BOs. This observation is consistent with previous research on conjugated polymers.[38] In addition, high values of J_0^x and K_{inter} can induce a similar effect. Consequently, crystal synthesis that promotes molecular electron transfer can enhance the observation of Bloch oscillations. Similar outcomes are expected for syntheses that improve the crystal speed of sound (v_s). We recall that $v_s = a_0\sqrt{K_{inter}/M_{inter}}$. Therefore, maximizing K_{inter} means increasing v_s.

We emphasize that all simulations are limited to a time window of five ps. Therefore, our methodology can only identify BO with frequencies equal to or greater than 0.2 THz. Any oscillatory motion below this limit was categorized as non-oscillating. However, this does not mean that lower-frequency oscillatory motion cannot develop. Some systems might exhibit oscillatory characteristics given enough time. Here, we restricted the dynamics to the picosecond scale because arbitrary long wavelengths, while theoretically possible, are irrelevant. This is because scattering effects can disrupt the wave picture. The longer the pulse takes to complete a cycle, the bigger the probability of facing these effects.[73] Consequently, oscillations may not be detected even when the idealized crystal indicates otherwise.

After determining the physical conditions to induce BO, assessing the general wave properties is a natural step. Due to the wide range of parameters, the profile of two Bloch states can vary significantly. Therefore, individual analysis of every possible case is not feasible. In this context, statistical treatment is advantageous. It allows for the estimation of general behavior through distributions and averaged quantities.

Metrics must be applied to represent BOs. In this context, we characterize the motion by examining the oscillation of the total current.[38] Fig. 5 illustrates the procedure. Fig. 5(a) shows the charge density heatmap of the configuration (0.10, 0.07, 0.01, 0.50). Fig. 5(b) shows the corresponding current time evolution. Note that j oscillates with the same frequency as the charge density heatmap. Moreover, the time frames where the carrier has no velocity also display zero current. This response validates that j accurately describes the charge dynamics in the subsequent analysis.

Characterizing Bloch oscillations involves calculating the mean frequency of this oscillatory motion. Thus, it is convenient to decompose the signal using fast Fourier transform (FFT). The resulting spectrum is shown in Fig. 5(c), where each peak represents a pure oscillation at a specific frequency. This distribution serves as a fingerprint of the BO. The characteristic frequency of the BO will be the mean frequency weighted by the FFT amplitudes. A notable peak appears at 0.4 THz in the present case, representing the primary mode. The red line in Fig. 5(b) is the inverse FFT corresponding to this particular frequency. As can be seen, this single component recovers most of the current density dynamics. Another critical attribute of an oscillating signal is its amplitude. Here, we define the charge

 This journal is © the Owner Societies 2025

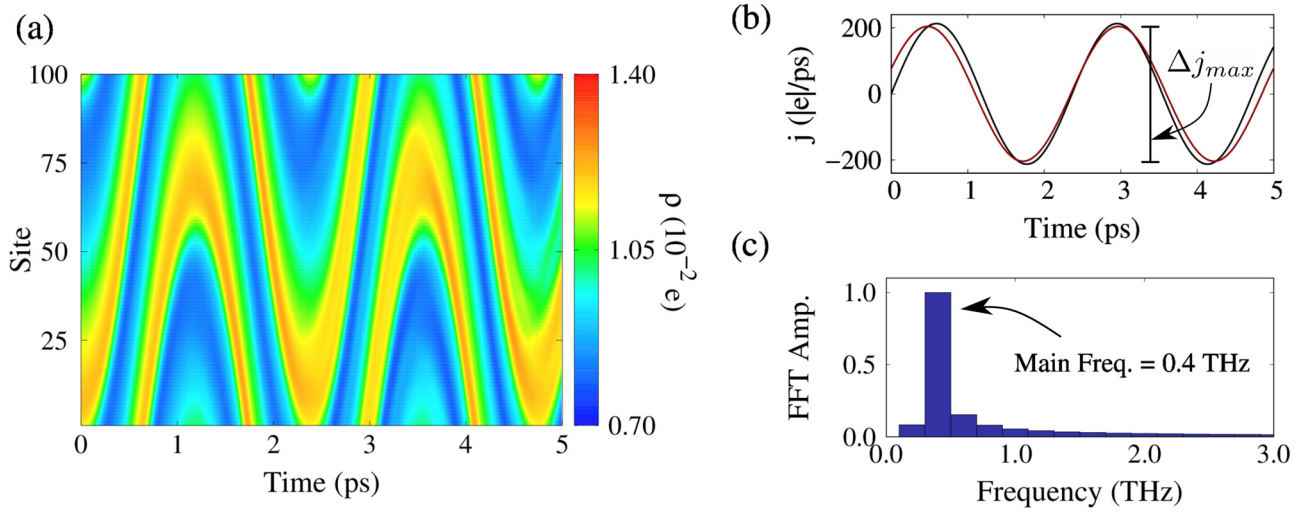

Fig. 5 Illustration of the metrics applied in the BO. (a) The charge density heatmap of a typical BO. The corresponding charge current time evolution is shown in (b). (c) The corresponding FFT spectra, presenting a predominant peak at the characteristic frequency of 0.4 THz.

current amplitude, $\Delta j_{\max}$, as the difference between the maximum and minimum charge current values during the dynamics. Fig. 5(b) represents the calculation of this metric. The mean frequency, $\Delta j_{\max}$, and the DC component will constitute the metrics that characterize the BO.

Fig. 6 presents the general properties of BO. The charge density heatmaps (a)–(f) represent the different possible profiles. As can be seen, there is a significant diversity in the propagation. The dynamics in Fig. 6(a)–(c) are from the same crystal (0.1, 0.3, 0.001, E_0) with, respectively, E_0 = 0.3, 1.0, and 5.0 mV Å^{-1}. In other words, moving from (a) to –(c) means an increase in the electric field strength. Analysis of this transition shows that E_0 increases the oscillation frequency, an expected outcome. We recall that, under a semi-classical approach, the frequency of Bloch waves in a homogeneous lattice is eE_0a_0/h.[37] Numerous recent model Hamiltonian-based studies strongly corroborate this relationship.[40,41,74,75] Our results agree with this critical physical connection, further validating the description.

The other profiles emerge by altering the crystal properties. Fig. 6(d) shows the case of (0.1, 0.01, 0.001, 0.3). Although fairly delocalized, the pulse does not extend over the entire chain. It exhibits approximately one cycle during five ps. Therefore, the motion has one precise frequency of $\approx 1/5$ = 0.2 THz. The carrier might have a high transport inertia depending on the

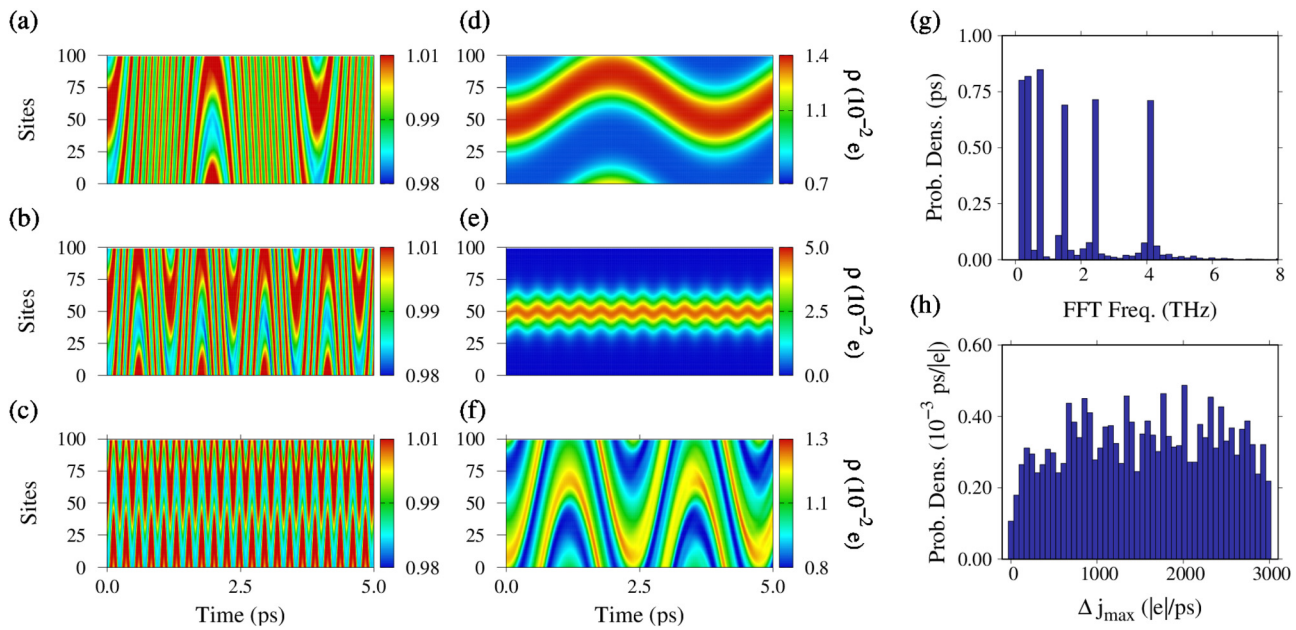

Fig. 6 General properties of BO in molecular crystals. (a)–(f) Representative charge density heatmaps of Bloch states. In (g), the frequency distribution within the cluster wave-like propagation is presented. (h) The distribution of charge current amplitude.

crystal attributes. Then, although oscillating, the pulse is heavily confined. Consequently, the motion is a high-frequency fluctuation restricted to a specific area. That is the case of Fig. 6(e) heatmap. Alternatively, intermediate wave patterns can appear. For instance, the dynamics in Fig. 6(f) is a combination of the profiles shown in Fig. 6(a) and (d).

The heatmaps depicted in Fig. 6(a)–(f) are good representations of the cluster. Still, a statistical treatment may provide broader insights. Fig. 6(g) and (h) are the averaged frequency and amplitude probability density distributions. The frequency distribution extends over $\approx$0.2 to 6 THz with multiple well-defined peaks. The heatmaps in Fig. 6(c) and (e) represent the high-frequency configurations. On the other hand, the lower end consists of cases such as those shown in Fig. 6(d). The results demonstrate molecular crystals' BO in the terahertz scale. This finding is consistent with previous experimental and theoretical reports. For instance, BOs in the THz range were observed in AlGaAs/GaAs[45,46] and theorized in graphene[42] and its variants.[43,44] Nonetheless, the amplitude distribution does not have a distinctive peak. It spans from $\approx$0 to 3000 $|e|$ ps^{-1} almost uniformly. The high amplitude BOs are associated with dynamics that travel over the chain multiple times before switching the current direction. An example of this case is shown in Fig. 6(a) that has Δj_{max} = 1823 $|e|$ ps^{-1}. Under those conditions, the current has time to build up before decelerating. Consequently, the fluctuation happens over bigger j magnitudes.

In turn, the low-amplitude configurations have the opposite dynamics. The corresponding pulse does not cross multiple times the chain ends. The heatmap in Fig. 6(d), with Δj_{max} = 60 $|e|$ ps^{-1}, represent those individuals. Note that a low crossing ratio does not necessarily induce a high frequency. This can be verified when comparing Fig. 6(a) and (d). They have similar frequencies ($\approx$0.2 THz) but strikingly different Δj_{max}. Note that Fig. 6(a) and (d) also have a significant difference in charge localization. This trend suggests that Δj_{max} may correlate with the delocalization length.

The distributions provide a general understanding of the scale of oscillations. However, the analysis must reveal how the Hamiltonian parameters can modulate the wave properties. To better understand the mechanism, we present in Fig. S4 (ESI†) the FFT frequency pair plot of K_{inter}, J_0^x, and α_{inter} for the BO configurations. The scatter points in blue are the low-frequency BO near 0.20 THz. Those in gray and red are close to 1.74 and 3.37 THz, respectively. Analysis of the distributions (and the scatter plots) indicate no evident influence in the BO frequency at first sight.

However, a closer look at the pairplot elements reveals an intriguing result. For instance, consider the plot of α_{inter} *versus* J_0^x, shown in Fig. 7(a). Although high-frequency configurations are distributed uniformly across the subspace, this is not true for those below 1.74 THz. These solutions are more abundant when $\alpha_{inter} < 2$ eV Å^{-1}. In other words, higher values of α_{inter} favor, simultaneously, the formation of high-frequency waves and non-oscillating states. This behavior occurs because increasing α_{inter} leads to dynamics similar to those presented in Fig. 6(e). Stronger electron–phonon coupling forces the charge to be highly localized, effectively trapping it. Due to the high energy barrier, the electric field cannot release the carrier. Instead, the quasiparticle remains confined within the trapping region.

Additionally, the density of points indicates that the BOs have a delicate stability in this region. Polaron-like solutions predominantly populate zones with visible white spots. Therefore, the lack of scatter points for $\alpha_{inter} > 2$ eV Å^{-1} in Fig. 7(a) suggests that slight variations may trap the carrier in a single site, halting the oscillating motion. The results also provide a potential design rule to obtain low-frequency BO by reducing α_{inter}.

Fig. 7(b) presents the Δj_{max} pairplot element $\alpha_{inter} - J_0^x$. The points in yellow, orange, and brown are near, respectively, to Δj_{max} = 3.41, 763.10, and 1524.50 $|e|$ ps^{-1}. Note that three zones, with increasing amplitude, are present: $J_0^x < 0.15$ eV, 0.15 eV $\leq J_0^x < 0.25$ eV, and $J_0^x \geq 0.25$ eV. Therefore, increasing J_0^x leads to wider BOs. This result can be rationalized when considering the meaning of J_0^x. The hopping integral is related to the electronic transfer rate. The higher the value of J_0^x, the easier charges hop between sites. In other words, a strong

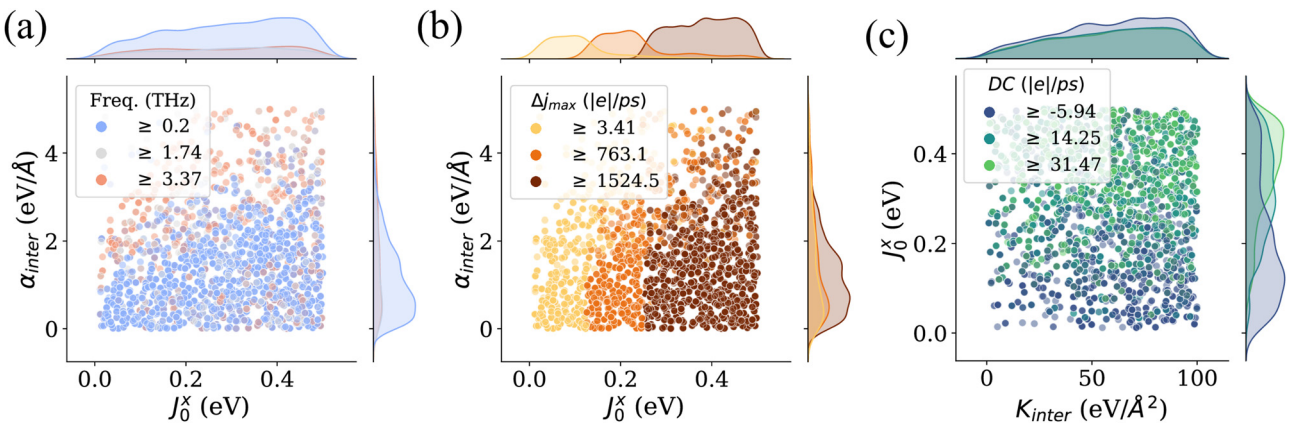

Fig. 7 Pairplots of the wave properties. (a) Frequency scatter plot for the interplay between α_{inter} and J_0^x. (b) Amplitude scatter plot for the interplay between α_{inter} and J_0^x. (c) DC component scatter plot of J_0^x and K_{inter} interplay.

 This journal is © the Owner Societies 2025

hopping induces a higher charge current. Notably, the effect of the other parameters is not as clear as J_0^x. To see that, we present in Fig. S5 (ESI†) the Δj_{max} pairplot of K_{inter}, J_0^x, and α_{inter}.

Note that Fig. 7(b) corroborates the trend shown after comparing Fig. 6(a) and (d). It is known that increasing the hopping integral promotes charge delocalization. We observe that increasing the same parameter leads to higher current amplitudes. Therefore, high Δj_{max} are indeed associated with delocalized oscillations. An important metric for characterizing the BO is the DC component. Fig. 7(c) shows the element $K_{inter} - J_0^x$ of the DC pairplot. The configurations in green have a high DC component, while those in blue have a reduced magnitude. As can be seen, K_{inter} poses no significant influence on this wave property. The electronic hopping, in turn, can dramatically change it. Increasing J_0^x from 0.01 to 0.50 eV shifts the DC from ≈ 0 $|e|$ ps^{-1} to ≥ 62.94 $|e|$ ps^{-1}. The hopping integral affects the DC because the parameter controls the charge transport efficiency. Crystals with a strong J_0^x have charge carriers with a lower effective mass. Consequently, the electric field can induce a stronger drift.

The DC pairplot of K_{inter}, J_0^x, and α_{inter} is shown in Fig. S6 (ESI†). The colors follow the same meaning as in Fig. 7(c). In contrast to the electronic hopping, the change in DC due to K_{inter} and α_{inter} variations follow an unclear pattern. However, some insight is available when considering certain pairplot combinations. For instance, the scatter plot of α_{inter} as a function of K_{inter} shows that high-DC configurations are associated with a low α_{inter}. The setting is similar to Fig. 7(a), where the BO below $\alpha_{inter} = 2$ eV Å^{-1} shows a common behavior. Here, low DC for strong coupling is due to the lattice-electronic trapping. The charge can no longer be transferred when the confinement is too strong. The results extracted from Fig. 7(b) and (c) indicate that increasing J_0^x can simultaneously enhance the oscillation amplitude and DC component in molecular crystals. This outcome has significant practical implications as both properties might be relevant for the performance of optoelectronic devices.

Besides the crystal's internal parameters, external agents such as the electric field can profoundly alter the charge dynamics. Therefore, it is essential to analyze the impact on the wave properties. Fig. 8 presents the E_0 probability densities according to the frequency (a) and DC (b) subsets. The color legends of Fig. 8(a) and (b) retain the meaning of Fig. 7(a) and (c), respectively. As can be seen in Fig. 8(a), the low-frequency distribution (blue) is limited to $E_0 \leq 2$ mV Å^{-1}. The distributions of higher frequency BO are shifted about the blue curve. In other words, raising E_0 increases the oscillation frequency. That is an expected result, already anticipated from the analysis of the charge density heatmaps in Fig. 6(a)–(c). Here, this response is confirmed to happen in the whole dataset of Bloch-like states.

The dependence of DC with E_0 shows a clear trend. In Fig. 8(b), the distribution with the highest DC is confined to $E_0 < 2$ mV Å^{-1}. Waves with progressively lower DC are located at high-field regions. Therefore, raising E_0 reduces the DC component. This surprising result can be understood when reconsidering Fig. 8(a). We recall that the electric field

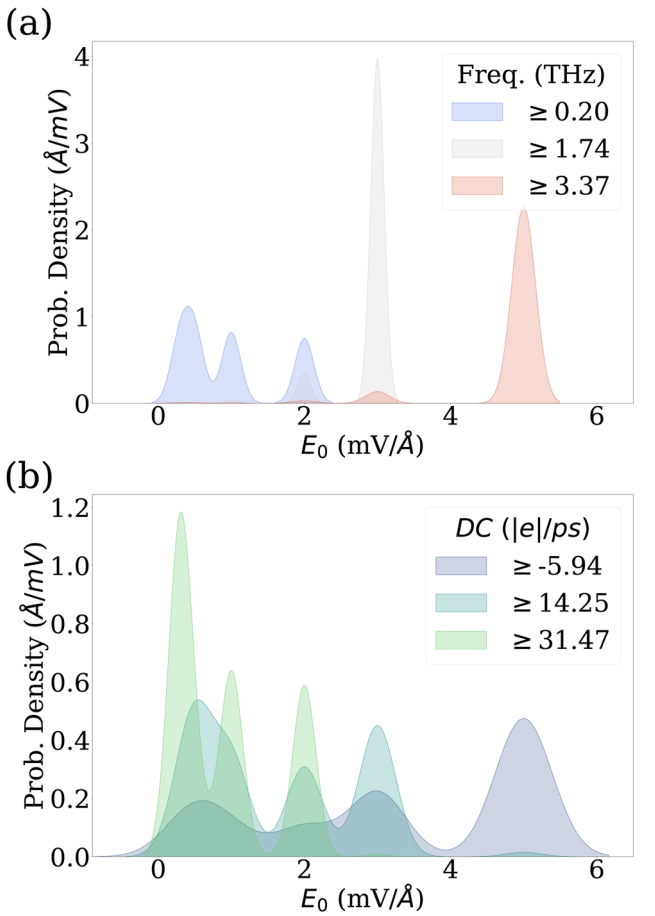

Fig. 8 (a) Probability density distribution of wave dynamics with frequencies near 0.20 (blue), 1.74 (gray), and 3.37 THz (red). (b) Probability density distribution of wave dynamics with DC near −5.94 (blue), 14.25 (dark green), and 31.47 |e| ps^{-1} (light green).

increases the wave frequency. In other words, this external agent modulates the charge transport to have a stronger alternating component. The current energy is limited. Then, increasing the alternating component is bound to reduce the DC counterpart. A similar outcome was reported for conjugated polymers.[38] In essence, the effect of E_0 in the wave properties found in our results corroborates previous observations, validating our methodology.

The same study also indicates that temperature may lead to an inverse response for typical atomic crystals.[38] In general, modeling thermal effects on hybrid Hamiltonian stationary solutions is possible by introducing random noise distortions on the lattice or/and employing the Fermi–Dirac fermion distribution on the electronic states.[76,77] Implementing the fluctuation–dissipation theorem is a common strategy to explicitly add fluctuations during the dynamics. Through this route, the lattice motion equation has an empirical damping term representing the interaction with the thermal bath. Previous investigations based on these modifications often indicate that thermal effects degrade transport efficiency.[25,78] We believe the same conclusions extend to molecular crystals in the physical regimes considered, a hypothesis to be tested.

Fig. 8 evidences that BOs can emerge on molecular crystals when subjected to relatively weak electric fields (≈ 0.3 mV Å^{-1}). This result is surprising. Most theoretical works report BO on conjugated polymers only when the material is exposed to intense E_0 as a form to compensate for the small lattice parameter. For instance, fields up to 20 mV Å^{-1} were applied on polyacetylene.[38] Inorganic compounds seem to follow a similar requirement. For example, the critical electric field required to observe BO in polysilane is $E_0 = 5.2$ mV Å^{-1}.[40]

Electronic-field interactions drive the formation of BOs. However, the crystal is a coupled system. Therefore, looking at the lattice as during the charge carrier dynamics should be instructive. Fig. 9 presents the heatmaps of the internal lattice degree. Fig. 9(a)–(f) refer to the configurations shown in Fig. 6(a)–(f). Notably, much of the oscillating character of the charge density is present. First, consider the highly delocalized profiles shown in Fig. 6(a)–(c). Note that the corresponding heatmaps, present in Fig. 9(a)–(c), exhibit almost the same characteristic frequency. Moreover, the oscillation does not show an indication of a pulse-like propagation. Instead, the site's distortions oscillate in unison. This response suggests that the lattice is oscillating on the acoustic limit, the speed of sound. In this regime, the electric field can no longer enforce higher frequency motion for the lattice. Instead, the deformations become stronger due to the additional energy input.

On the other hand, the pulse-like oscillations display a slightly different setting. The lattice oscillations shown in Fig. 9(d)–(f) are more in tune with the charge density fluctuations displayed in Fig. 6(d)–(f). The reason lies in the electronic-lattice interplay during the formation of charge carriers. The more concentrated the charge is, the stronger the self-trapping mechanism. In other words, the lattice deformation is more profound, raising the transport inertia. Consequently, the charge cannot oscillate through an unrestricted high speed. This allows the electronic and lattice dynamics to oscillate, manifesting the carrier's coupling nature. The results demonstrate that the oscillations are not a purely electronic effect, having a direct impact on the site spatial arrangement. That attribute could serve as a potential source for indirectly detecting Bloch oscillations.

The results suggest that improving the intramolecular hopping integral might favor the formation of Bloch oscillations. In practical terms, this parameter can be improved in a number of ways. Since this attribute is the result of electronic orbital overlap, any chemical engineering approach that enhances the overlap will do. However, given the complexity of the underlying mechanisms involving the molecular electronic structure, these modifications should be adopted with care, as they could (potentially) trigger other competing effects. Nevertheless, one route to increase the hopping integral worth investigating is reducing the lattice parameter, as it would improve the overlap.

A direct way would be subjecting the sample to strong pressure.[79] Since many crystals are held by non-covalent bonds,

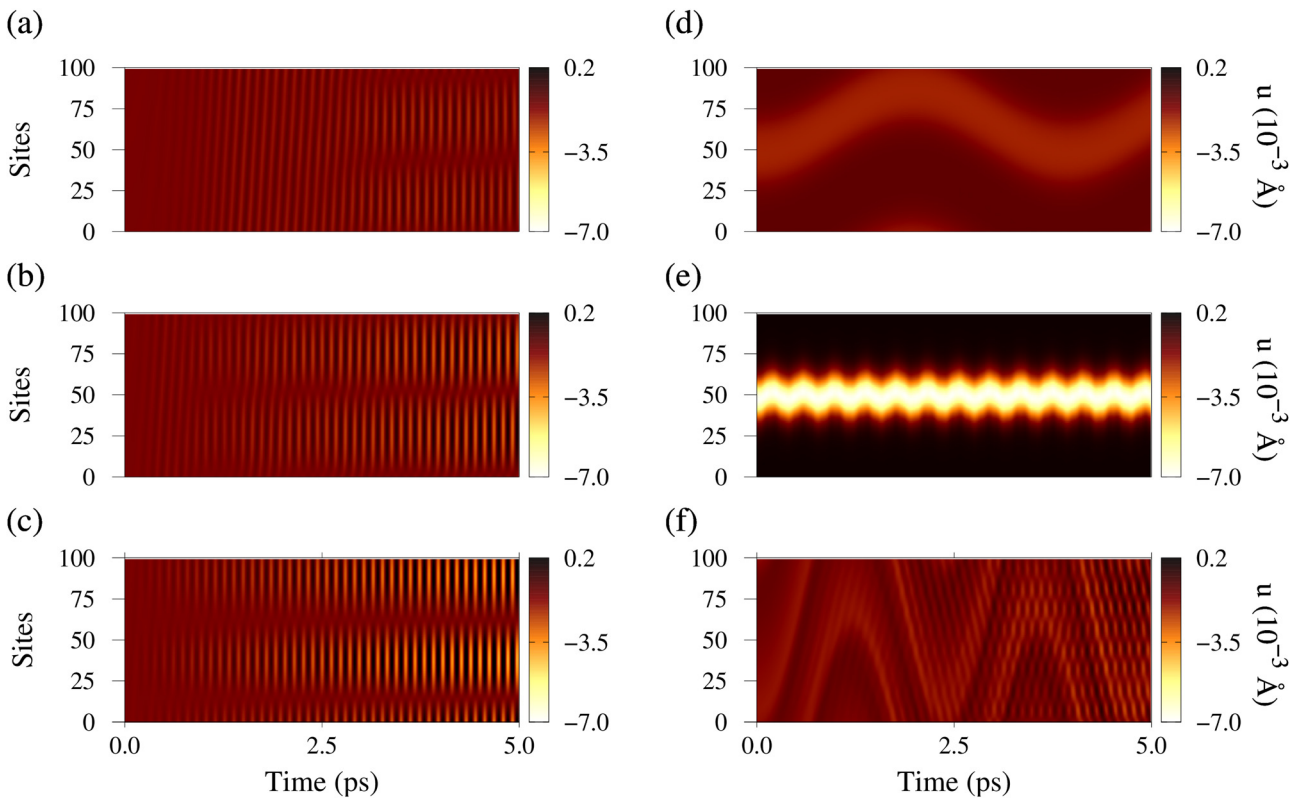

Fig. 9 Heatmaps of the intramolecular lattice degree of freedom of sites in the transport line. The crystal configurations of (a)–(f) are from the same profiles shown in Fig. 6(a)–(f).

 This journal is © the Owner Societies 2025

this external influence could narrow the gap between sites. Conversely, if the objective is to mitigate Bloch-type propagation, a potential way of doing this is increasing the distance between the sites. This can be done by, for instance, improving the steric repulsion between sites through the addition of polymeric radicals at the molecular sites.

When it comes to the electron–phonon coupling, the results suggest that narrowing the gap could favor the formation of Bloch oscillations. This can be achieved by tailoring the electronic band structure by adding mid-gap states. One way of doing this is by doping.[80,81]

As for the harmonic constants, their magnitude is related to the elastic strength holding the sites together. They can be improved by increasing the bond energy. One possible practical approach towards this is by enhancing the Coulomb interactions between molecular sites, effectively increasing the binding energy. This scenario resembles a task of finding ways to increase the slope of the potential energy curves for a generic diatomic molecule.[82]

We emphasize that metrics such as the FFT frequency and $\Delta j_{\max}$ are best suited to characterize BO dynamics. Their usefulness in representing the motion does not extend to projectile-type transport. A clear manifestation of this limitation appears when considering the Fourier expansion of a non-periodic signal, as would be the case of a polaronic propagation. Given enough terms, the series will converge to a large linear combination of sin and cos components. This convoluted composition affects interpretation, yielding little insight. More importantly, the most prominent component does not represent the motion because there is no oscillating motion in this example. For this reason, separating the transport regimes based on BO metrics might be intricate and not scalable.

This scenario resembles the classical ''cats *vs.* dogs'' problem in data science.[83] The objective is to cluster the images from a dataset containing dogs and cats. One could try to manually define characteristics such as skull shape, fur pattern, eye color, and other criteria to differentiate dogs from cats. However, this approach would require deep manual tuning of threshold values and continuous validation of each picture. Visual confirmation *via* machine learning is a more manageable and less demanding strategy. Here, we moved in the same direction, relying on the ML ability of image recognition rather than transport metrics to obtain the clusters. A clear advantage of proceeding with the proposed workflow is the lack of artificial thresholds to filter out the clusters. Using a metric (or a combination of them) to delimit the regions would require the determination of threshold values. Eventually, validation of the criteria would require analysis case-by-case of the whole dataset. This kind of imposition defeats the entire purpose of analyzing the parameter space as a whole.

We report that other clustering approaches were contemplated during the development process, namely the *k*-means and the affinity propagation methods. The choice for the propagation approach was based on the quality of the heatmaps subcluster separation and their number. The affinity approach, while performing well on recognizing the transport profiles, generated an untractable number of subclusters nearing 400. On the other hand, the *k*-means provided a good separation using a reasonable number of subclusters. However, a qualitative comparison between the separations indicated that the agglomeration method was performing slightly better, which motivated our choice.

The clustering algorithm selection process was by no means exhaustive, not involving a systematic search using the many available clustering methods. Here, we resorted to a practical approach in which we chose the algorithm able to adequately perform the task. A systematic benchmark of different methods is undoubtedly an open venue to explore, which could represent significant improvements in the overall workflow.

4 Conclusion

In summary, this study investigated charge transport regimes in OMCs, focusing on Bloch-like and polaron-like solutions. We used machine learning to cluster dynamic solutions based on hybrid Hamiltonians representing different crystal configurations. By analyzing charge current densities, we applied an unsupervised learning approach to distinguish between the two types of solutions.

Our analysis revealed distinct parameter regions associated with BOs and polaron-like solutions. Parameters such as the intermolecular spring constant, electronic transfer integral, and electron–phonon coupling constant were found to influence the type of transport regime significantly. Specifically, higher values of α_{inter} tend to favor polaron-like behavior, while increased J_0^x and K_{inter} generally enhance the likelihood of BOs.

Dimensionality reduction techniques visually confirmed the separation between Bloch and polaron states within the parameter space. Our clustering results, supported by heatmap analysis, demonstrated clear zones where one or the other transport regime predominates. However, some intermediate states exhibited mixed characteristics.

We observed BOs with variable amplitudes across a broad frequency range, from 0.2 THz to 6 THz. At the same time, the DC component of the current density highlighted how J_0^x affects charge transport efficiency. These findings suggest that adjusting parameters such as J_0^x, α_{inter} and K_{inter} can optimize BOs and polaron-like solutions in molecular crystals, pertinent for designing advanced optoelectronic devices. Our study provides a comprehensive framework for understanding and controlling BOs in molecular crystals, laying the groundwork for future research and practical applications in electronic materials.

Author contributions

T. S. A. C.: data curation, formal analysis, methodology, software, and writing – original draft preparation; M. L. P. J.: supervision, formal analysis, visualization, writing – original draft preparation. P. H. O. N. and L. A. R. J.: supervision, conceptualization, funding acquisition, methodology,

visualization, formal analysis and writing – reviewing. All authors reviewed the manuscript.

Data availability

Due to the large size of the dataset, the data supporting the findings of this study are available from the corresponding author upon reasonable request.

Conflicts of interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work received partial support from Brazilian agencies CAPES, CNPq, and FAPDF. M. L. P. J. acknowledges the financial support of the FAP-DF grant 00193-00001807/2023-16. Thanks also to CENAPAD-SP (National High-Performance Center in São Paulo, State University of Campinas – UNICAMP, project: proj960) and NACAD (High-Performance Computing Center, Lobo Carneiro Supercomputer, Federal University of Rio de Janeiro – UFRJ, project: a22002) for the computational support provided. P. H. O. N. acknowledges the financial support from CNPq grant 306080/2022-2, FAP-DF grants 00193-00001853/2023-15 and 00193-00001831/2023-47. L. A. R. J. acknowledges the financial support from FAP-DF 00193.00001808/2022-71 and 00193-00001857/2023-95 grant and FAPDF-PRONEM grant 00193.00001247/2021-20, and CNPq grant 350176/2022-1.

Notes and references

1 O. Ostroverkhova, *Chem. Rev.*, 2016, **116**, 13279–13412.
2 H. Lai and F. He, *Adv. Energy Mater.*, 2020, **10**, 2002678.
3 X. Yang, X. Feng, J. Xin, P. Zhang, H. Wang and D. Yan, *J. Mater. Chem. C*, 2018, **6**, 8879–8884.
4 D. Liu, F. Zhu and D. Yan, *J. Mater. Chem. C*, 2022, **10**, 2663–2670.
5 P. Sun, D. Liu, F. Zhu and D. Yan, *Nat. Photonics*, 2023, **17**, 264–272.
6 T. He, M. Stolte, C. Burschka, N. H. Hansen, T. Musiol, D. Kälblein, J. Pflaum, X. Tao, J. Brill and F. Würthner, *Nat. Commun.*, 2015, **6**, 5954.
7 Q. Di, L. Li, X. Miao, L. Lan, X. Yu, B. Liu, Y. Yi, P. Naumov and H. Zhang, *Nat. Commun.*, 2022, **13**, 5280.
8 S. Stafström, *Chem. Soc. Rev.*, 2010, **39**, 2484–2499.
9 A. J. Heeger, S. Kivelson, J. Schrieffer and W.-P. Su, *Rev. Mod. Phys.*, 1988, **60**, 781.
10 R. Freitas, R. Rivelino, F. D. B. Mota, G. K. Gueorguiev and C. de Castilho, *J. Phys. Chem. C*, 2015, **119**, 15735–15741.
11 R. Zhuravel, H. Huang, G. Polycarpou, S. Polydorides, P. Motamarri, L. Katrivas, D. Rotem, J. Sperling, L. A. Zotti and A. B. Kotlyar, *et al.*, *Nat. Nanotechnol.*, 2020, **15**, 836–840.
12 D. G. Sangiovanni, G. Gueorguiev and A. Kakanakova-Georgieva, *Phys. Chem. Chem. Phys.*, 2018, **20**, 17751–17761.
13 P. V. Medeiros, G. Gueorguiev and S. Stafström, *Carbon*, 2015, **81**, 620–628.
14 D. G. Sangiovanni, R. Faccio, G. K. Gueorguiev and A. Kakanakova-Georgieva, *Phys. Chem. Chem. Phys.*, 2023, **25**, 829–837.
15 N.-E. Lee, J.-J. Zhou, L. A. Agapito and M. Bernardi, *Phys. Rev. B*, 2018, **97**, 115203.
16 J. Ma, A. S. Nissimagoudar and W. Li, *Phys. Rev. B*, 2018, **97**, 045201.
17 R. R. Pela, C.-L. Hsiao, L. Hultman, J. Birch and G. K. Gueorguiev, *Phys. Chem. Chem. Phys.*, 2024, **26**, 7504–7514.
18 W. P. Su, J. R. Schrieffer and A. J. Heeger, *Phys. Rev. Lett.*, 1979, **42**, 1698.
19 W. P. Su, J. R. Schrieffer and A. J. Heeger, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1980, **22**, 2099.
20 L. Yang, D. Yi, S. Han and S. Xie, *Org. Electron.*, 2015, **23**, 39–43.
21 X.-X. Li, H. Peng, D. Wang and D. Hou, *Chin. Phys. B*, 2024, **33**, 037201.
22 L. Zhang, Y. Hao, W. Qin, S. Xie and F. Qu, *New J. Phys.*, 2020, **22**, 103034.
23 Q. Lu, X. Liu, M. Zhang and Z. An, *J. Phys. Chem. B*, 2024, 5500–5505.
24 Z. Sun, S. Li, S. Xie and Z. An, *J. Phys. Chem. C*, 2019, **123**, 21336–21344.
25 A. Troisi and G. Orlandi, *Phys. Rev. Lett.*, 2006, **96**, 086601.
26 Y. Ono and A. Terai, *J. Phys. Soc. Jpn.*, 1990, **59**, 2893–2904.
27 G. M. e Silva, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2000, **61**, 10777.
28 T. Astakhova and G. Vinogradov, *Eur. Phys. J. B*, 2019, **92**, 1–9.
29 A. A. Johansson and S. Stafström, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2004, **69**, 235205.
30 S. Sun, Y. Zhang, X. Liu and Z. An, *J. Phys. Chem. C*, 2019, **123**, 28569–28575.
31 Z. Sun, S. Li, S. Xie and Z. An, *Synth. Met.*, 2021, **279**, 116841.
32 M. B. Falleiros and G. M. e Silva, *J. Phys. Chem. A*, 2019, **123**, 1319–1327.
33 R. Meng, Y. Li, C. Li, K. Gao, S. Yin and L. Wang, *Phys. Chem. Chem. Phys.*, 2017, **19**, 24971–24978.
34 X. Li, D. Hou and G. Chen, *Phys. E*, 2019, **108**, 1–6.
35 Z. Sun, Y. Li, K. Gao, D. Liu, Z. An and S. Xie, *Org. Electron.*, 2010, **11**, 279–284.
36 J. Bredas, J. Scott, K. Yakushi and G. Street, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1984, **30**, 1023.
37 C. Zener, *Proc. R. Soc. London, Ser. A*, 1934, **145**, 523–529.
38 S. Sun, *ACS Omega*, 2023, **8**, 40456–40462.
39 R. Khomeriki, *Phys. Rev. A:At., Mol., Opt. Phys.*, 2010, **82**, 033816.
40 L. A. Ribeiro, W. Ferreira da Cunha and A. L. de Almeida Fonseca, *et al.*, *J. Chem. Phys.*, 2017, **146**, 144903.
41 D. Morais, P. de Brito, H. Nazareno and W. Dias, *J. Phys.: Condens. Matter*, 2022, **34**, 455302.
42 D. Dragoman and M. Dragoman, *Appl. Phys. Lett.*, 2008, **93**, 103105.

 This journal is © the Owner Societies 2025

43 H. Cheng, C. Li, T. Ma, L.-G. Wang, Y. Song and H.-Q. Lin, *Appl. Phys. Lett.*, 2014, **105**, 072103.
44 R. Musah, S. Y. Mensah and S. S. Abukari, *Phys. E*, 2014, **61**, 90–94.
45 T. Dekorsy, P. Leisching, C. Waschke, K. Kohler, K. Leo, H. Roskos and H. Kurz, *Semicond. Sci. Technol.*, 1994, **9**, 1959.
46 H. G. R. Kö and H. K. H. Kurz, *Jpn. J. Appl. Phys.*, 1995, **34**, 1370.
47 D. Cai, A. Bishop, N. Grønbech-Jensen and M. Salerno, *Phys. Rev. Lett.*, 1995, **74**, 1186.
48 A. Andronov, I. Nefedov and A. Sosnin, *Semiconductors*, 2003, **37**, 360–366.
49 Y. Zhang, D. Zhang, Z. Zhang, C. Li, Y. Zhang, F. Li, M. R. Belić and M. Xiao, *Optica*, 2017, **4**, 571–575.
50 T.-B. Wang, N.-H. Liu, X.-H. Deng and Q.-H. Liao, *J. Opt.*, 2011, **13**, 095705.
51 R. Morandotti, U. Peschel, J. Aitchison, H. Eisenberg and Y. Silberberg, *Phys. Rev. Lett.*, 1999, **83**, 4756.
52 Y. Zhu, J. Peng, C. Xu and Z. Lan, *J. Phys. Chem. Lett.*, 2024, **15**, 9601–9619.
53 K. Acheson and A. Kirrander, *J. Chem. Theory Comput.*, 2023, **19**, 6126–6138.
54 H. Hadipour, C. Liu, R. Davis, S. T. Cardona and P. Hu, *BMC Bioinf.*, 2022, **23**, 132.
55 A. Glielmo, B. E. Husic, A. Rodriguez, C. Clementi, F. Noé and A. Laio, *Chem. Rev.*, 2021, **121**, 9722–9758.
56 Y. Guan, S. Yang and D. H. Zhang, *J. Phys. Chem. A*, 2018, **122**, 3140–3147.
57 R. Yamashita, M. Nishio, R. K. G. Do and K. Togashi, *Insights into imaging*, 2018, vol. 9, pp. 611–629.
58 Z. Li, F. Liu, W. Yang, S. Peng and J. Zhou, *IEEE Trans. Neural Networks Learn. Syst.*, 2021, **33**, 6999–7019.
59 E. Mozafari and S. Stafström, *J. Chem. Phys.*, 2013, **138**, 184104.
60 E. Mozafari and S. Stafström, *Phys. Lett. A*, 2012, **376**, 1807–1811.
61 M. L. Pereira Junior and L. A. Ribeiro Junior, *J. Mol. Model.*, 2017, **23**, 1–7.
62 Y. Arikabe, M. Kuwabara and Y. Ono, *J. Phys. Soc. Jpn.*, 1996, **65**, 1317–1324.
63 M. Riedmiller and H. Braun, *IEEE international conference on neural networks*, 1993, pp. 586–591.
64 T. S. Cassiano, M. Pereira, G. e Silva, P. H. de Oliveira Neto and L. Ribeiro, *Nanoscale*, 2024, **16**, 2337–2346.
65 F. Chollet *et al.*, *Keras*, 2015, **https://keras.io**.
66 F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot and E. Duchesnay, *J. Mach. Learn. Res.*, 2011, **12**, 2825–2830.
67 J. H. Ward Jr, *J. Am. Stat. Assoc.*, 1963, **58**, 236–244.
68 H. Ishii, K. Honma, N. Kobayashi and K. Hirose, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2012, **85**, 245206.
69 A. Troisi, G. Orlandi and J. E. Anthony, *Chem. Mater.*, 2005, **17**, 5024–5031.
70 A. Troisi and G. Orlandi, *J. Phys. Chem. A*, 2006, **110**, 4065–4070.
71 J. Lei and Y. Shimoi, *J. Phys. Soc. Jpn.*, 2011, **80**, 034702.
72 D. Kobak and P. Berens, *Nat. Commun.*, 2019, **10**, 5416.
73 O. Schubert, M. Hohenleutner, F. Langer, B. Urbanek, C. Lange, U. Huttner, D. Golde, T. Meier, M. Kira and S. W. Koch, *et al.*, *Nat. Photonics*, 2014, **8**, 119–123.
74 Y. Li, X.-J. Liu, J.-Y. Fu, D.-S. Liu, S.-J. Xie and L.-M. Mei, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 2006, **74**, 184303.
75 V. Lakhno and A. Korshunova, *Eur. Phys. J. B*, 2011, **79**, 147–151.
76 T. Gao, Q. Lu, C. Li, K. Gao and S. Xie, *J. Phys. Chem. C*, 2019, **123**, 28527–28532.
77 L. Wang, O. V. Prezhdo and D. Beljonne, *Phys. Chem. Chem. Phys.*, 2015, **17**, 12395–12406.
78 L. Wang, A. V. Akimov, L. Chen and O. V. Prezhdo, *J. Chem. Phys.*, 2013, **139**, 174109.
79 F. Munoz, H. O. Collado, G. Usaj, J. O. Sofo and C. A. Balseiro, *Phys. Rev. B*, 2016, **93**, 235443.
80 E. J. Mele and M. J. Rice, *Phys. Rev. B:Condens. Matter Mater. Phys.*, 1981, **23**, 5397.
81 S. H. Simon, *The Oxford solid state basics*, OUP, Oxford, 2013.
82 H. Ladjimi and M. Tomza, *Phys. Rev. A*, 2024, **109**, 052814.
83 O. M. Parkhi, A. Vedaldi, A. Zisserman and C. Jawahar, 2012 IEEE conference on computer vision and pattern recognition, 2012, pp. 3498–3505.