Learning Silicon Dopant Transitions in Graphene using Scanning Transmission Electron Microscopy

Anonymous Author(s)
Affiliation
Address
email

Abstract

We introduce a machine learning approach to determine the transition dynamics of silicon atoms on a single layer of carbon atoms, when stimulated by the electron beam of a scanning transmission electron microscope (STEM). Our method is data-centric, leveraging data collected on a STEM. The data samples are processed and filtered to produce symbolic representations, which we use to train a neural network to predict transition probabilities. These learned transition dynamics are then leveraged to guide a single silicon atom throughout the lattice to pre-determined target destinations. We present empirical analyses that demonstrate the efficacy and generality of our approach.

1 Introduction

Sub-atomically focused electron beams in scanning transmission electron microscopes (STEM) can induce a broad spectrum of chemical changes, including defect formation, reconfiguration of chemical bonds, and dopant insertion. Several groups have shown the feasibility of direct atomic manipulation via electron beam stimulation, which holds great promise for a number of downstream applications such as material design, solid-state quantum computers, and others [Jesse et al., 2018, Susi et al., 2017b, Dyck et al., 2017, Tripathi et al., 2018, Dyck et al., 2018]. One of the challenges for advances in this space is that these types of atomic manipulation rely on manual control by highly-trained experts, which is expensive and slow.

The ability to accurately automate this type of beam control could thereby result in tremendous impact on the feasibility of atomic manipulation for real use cases. A critical requirement for this automation is accurate estimation of the transition dynamics of atoms when stimulated by focused electron beams. To date, the microscopy community has relied on heuristic estimates for these transition dynamics, with anecdotal evidence of their accurateness. Indeed, the common practice has been to use the physically intuitive, but heuristic, assumption that the optimal beam position is directly on a neighboring atom.

In this paper we present a technique for estimating these atomic transition dynamics using machine learning techniques on collected observations. Our approach consists in a sequence of steps ultimately resulting in a probability distribution over possible beam positions relative to a particular atom, conditioned on the atom's prior position as well as the electron beam's location and dwell time. While evaluated on single silicon atom (typically referred to as a *dopant*) on a lattice of carbon atoms (graphene), our methodology can be generally applied to other materials.

Submitted to 37th Conference on Neural Information Processing Systems (NeurIPS 2023). Do not distribute.

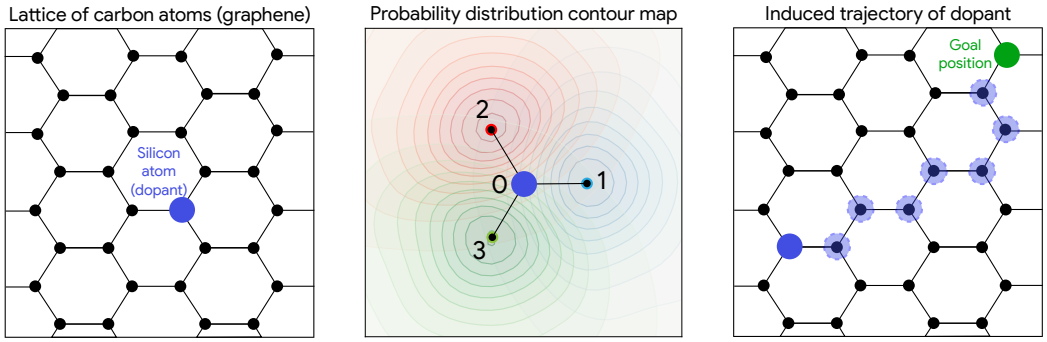

Figure 1: **Left:** Illustration of graphene with a single dopant. Carbon atoms depicted with black circles, while the silicon atom in blue; **Center:** Contour map learned by our method, depicting the probability of transitioning to each of the neighbours for different beam positions. The distributions for each of the neighbours are differentiated using three colours, and the numbers indicate neighbour ordering as discussed in Section 3.3; **Right:** Example trajectory of dopant towards a goal position.

To demonstrate the practical validity of our approach, we use our learned transition probabilities to automate the sequential control of a silicon atom on graphene towards a pre-specified target position. Modern STEMs are capable of this type of automation, and our work paves the way for future advances in automated atomic control.

2 Problem description

Our system consists of *graphene*: a single layer of carbon atoms arranged in 3-fold configuration (i.e. every carbon atom is connected to three other carbon atoms). On this lattice, a single silicon atom (hereafter referred to as *the dopant*) has taken the place of one of the carbon atoms. We focus an electron beam on a position in the area spanned by the dopant and its three carbon neighbours for a specified amount of time (referred to as the *dwell time*). This electron beam stimuli can result in the dopant moving to one of its neighbours (by trading places with the respective carbon atom), or in the configuration remaining unchanged[1]. This configuration has been extensively explored, and hence is an ideal system for exploring this type of automation [Dyck et al., 2017, Markevich et al., 2020, 2021]. We provide an illustration of this configuration in the left panel of Figure 1.

Our objective is to learn a probability distribution over the position of the dopant, conditioned on its current position, beam location, and beam dwell time. If accurate, we can use this distribution to determine the optimal beam location and dwell time so as to induce the dopant to move to one of its neighbouring positions. In the center panel of Figure 1 we can see a heat map depicting the transition probabilities for each of the dopant's neighbours (differentiated with colors) for varying beam locations.

Equipped with these probability maps, we can repeatedly induce transitions of the dopant to neighbouring atoms, resulting in a full trajectory. In other words, we can use a fully greedy strategy to move the dopant to any pre-specified target position on the lattice via a simple shortest-distance path. In the right panel of Figure 1 we depict one such possible trajectory.

3 Description of Methodology

In this section we detail the methodology we used in order to learn the transition dynamics of electron-beam induced atomic manipulation, illustrated in Figure 2. We used a Nion UltraSTEM 100, which grants us access to nearly every microscope control via a Python

[1]Note that for long dwell-times this "unchanged" outcome can be a result of the dopant moving twice: once to its neighbour and then back to its original position.

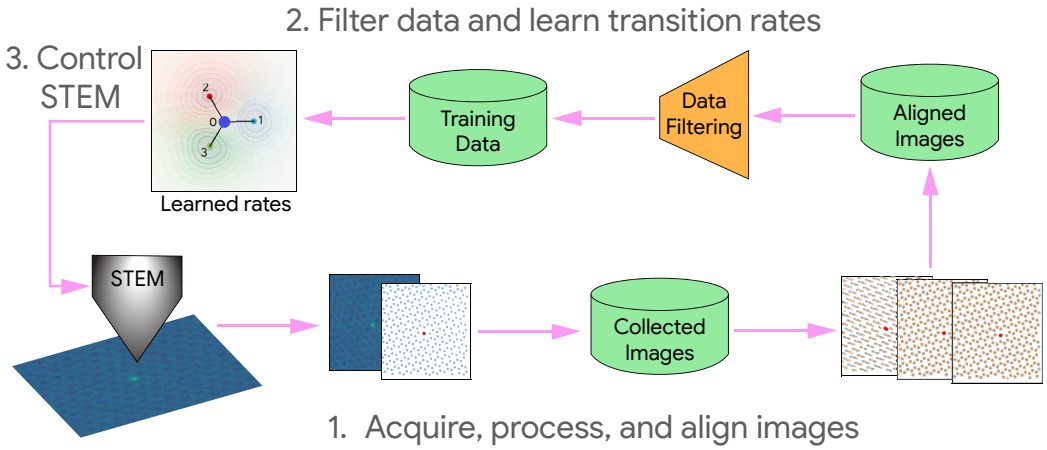

Figure 2: An overview of the full pipeline for learning the transition probabilities.

API. The silicon dopant atoms have been inserted into the lattice in a previous step at a higher (100 kV) accelerating voltage, which is described elsewhere [Roccapriore et al., 2023].

3.1 Data collection

Our methodology relies on real data collected with a STEM device, so it is important that the data gathered is informative for the task at hand. Since we are concerned with the transition dynamics of the dopant, our data collection approach is as follows:

1. Acquire an initial image of the graphene. For this, we used a field of view of 3 nm with the dopant in the center. A sample image is shown on the bottom-left panel of Figure 3.
2. Sample a position uniformly within a 2.84 Å radius[2] of the atom. In the top-center panel of Figure 3 we display the normalized beam positions.
3. Focus the electron beam at that position for a dwell time drawn from a distribution mostly between 1 and 10 seconds. The top-left corner of Figure 3 displays the distribuion of dwell times.
4. Acquire a final image of the graphene.

There are a few considerations that are worth mentioning. First, it is important to gather multiple samples using the same beam position and dwell time, as transitions are probabilistic. Second, image acquisition is done using the same electron beam, which implies that imaging itself can cause a transition; to mitigate this, the imaging electron dose should be minimized. Third, placing the electron beam in a known position relative to the silicon should be conducted in a controlled and automated fashion – for this, atomic coordinates must be known in as close to real time as possible, and flexible control of the beam position is needed.

3.2 Atomic alignment

In the bottom-left panel of Figure 3 we display one of the raw images captured with the STEM device, and to the right of it the processed output. The processed image is an "idealized" configuration as illustrated in Figure 1, which are easier to operate on. While it is relatively simple to identify the atoms in a single image, there are some challenges that arise when using more than one raw image, as discussed in Section 3.1: the graphene sheet may have physically moved between image acquisition steps (specimen drift), and there may be aberrations (such as warping) caused by the electron beam. To be able to use the images acquired for learning transition dynamics, we need not only need to detect the atoms in each

[2]Chosen to be the cumulative length of two carbon-carbon bonds.

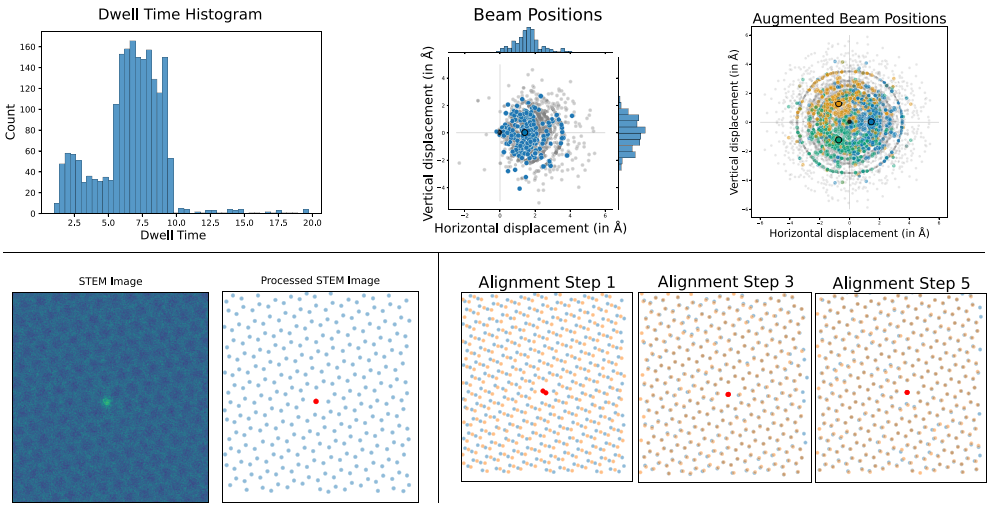

Figure 3: **Top row:** Histogram of dwell times (left) and standardized beam positions (center) used for data acquisition; beam position data augmented for three neighbours (right). In the beam position plots, grey and colored circles represent negative and positive transitions, respectively. The colored circles with a black border represent the neighbouring atoms. **Bottom left:** Raw image acquired from the STEM (left), and processed image after atom detection (right). **Bottom right:** The result of conducting multiple alignment iterations with a trained convolutional aligner. Orange dots reflect estimated atom positions in the previous scan, blue dots reflect atom positions in the current scan, and red dots are dopant positions. As can be observed, the discrepancies diminish with additional iterations.

of the separate images, but also to be able to *map* each of the atoms from one image to the next.

We considered three solutions to this problem. A classical approach would be to take the cross-correlation between the two scans and estimate the drift to be the arg max of this cross-correlation. Unfortunately, as images of this system are generally dominated by the very bright silicon dopants, this has the net effect of always aligning the dopant positions between time steps – leading to a conclusion that the dopants never move, which is known to be false. A second, more sophisticated alternative considered was to use the iterative closest points (ICP) algorithm on the extracted atom positions (a similar technique as used by Roccapriore et al. [2021]). This allows us to equally weight the dopant and non-dopant atoms, simply aligning the lattices together. This method led to acceptable alignments in most cases, but it was quite sensitive to failures in atom detection.

Our final, most robust solution was to use a denoising convolutional neural network to solve the alignment problem directly from scans of the system. Given a stack of image observations $(o_1, ..., o_n)$, the network is trained to predict the drift between o_{n-1} and o_n, d_n, as a two-dimensional vector. Historical observations $o_1, \dots, o_{n-2}$ serve to provide context and noise reduction, but are not used in the loss; as they have already been approximately aligned with o_{n-1}, they provide additional information about the needed shift. After the network has been applied, we take its prediction $\hat{d}_n$ and shift the current observation by it to align it to the prior observations, and add it to the current stack of observations.

To train this network, we generated a dataset of synthetic trajectories, each consisting of sequential image observations of a doped graphene system under random, correlated drift. To add robustness, we applied both trajectory-wide augmentation by randomly dropping atoms, introducing regions of bright contamination, and adding synthetic large holes to the system. We then simulated drift, treating the direction and magnitude of the drift as a temporally correlated random variables. Finally, we took a series of synthetic scans with the simulated cumulative drifts applied; for each scan, we also randomly perturbed the system, occasionally dropping or moving atoms. We parameterized the network as six convolutional

layers followed by downsampling, followed by a single fully-connected layer. We trained drift correction prediction with mean squared error.

Note that the estimation of d_n is a denoising task. Given this, we can iteratively apply our network for more precise drift correction. We find that this has a very large impact on performance. Qualitatively it is clear that the alignment quality increases with additional iterations (see bottom-right panel of Figure 3). Quantitatively, a transition model trained on only single-step-aligned data produces markedly different estimates of the optimal beam position, which are not able to successfully cause transitions when applied to the greedy controller (see "Past Neighbor" in Figure 4).

3.3 Data filtering, augmentation, and structure

As there are many possible sources of noise and error in our data, we conduct aggressive filtering prior to training. Out of a starting data set size of $6,754$ examples, we discard transitions where: **(1)** There is no recorded beam position (793 examples); **(2)** There is not exactly one (1) detected dopant atom before and after the transition (4 examples); **(3)** The dopant does not have the expected number of neighbors (3) before and after the transition ($3,593$ examples); and **(4)** The neighbors are not roughly the expected distance (1.42-1.7 Å) from the dopant (411 examples).

These led us to discard approximately 80% of the data we received from the microscope, resulting in a final dataset of $1,953$ examples. Once filtering has been done, we further post-process our observations to create a uniformized training set. To express beam positions in a consistent format, we translated them to a frame-of-reference relative to the current position of the dopant to be moved, with this atom at the origin. We then label the neighbor closest to the beam position as neighbor 1; we rotate the system, including the beam, such that this atom lies on the x-axis. The neighboring atoms are then numbered accordingly in counter-clockwise order; these are the indices we predict in our classification. We denote "no movement" as index 0. This labeling is illustrated in the top-center panel of Figure 1.

In the structure noted above, there is no systematic difference between the three neighboring atoms; only their distance from the beam separates them. As we expect our system to be invariant to both rotation and reflection, we enforce this by adding data augmentation. To do this, we first reflect across the x-axis with 50% probability. This exchanges the second and third neighbors. We then apply 0°, 120°or 240°rotations with equal probability, rotating the neighbor indices accordingly. This leads to an effective sixfold increase in data coverage. The top-right panel of Figure 3 visualizes the beam positions post-augmentation.

3.4 Learning the transition dynamics

Experimentally, we modeled our problem as a classification task, estimating $P(S'|a)$, where S' is the position of the silicon at the next time step, s_0 is its current position, and a is the beam dwell action chosen by the user, specified as a two-dimensional coordinate x and a duration in seconds Δt. To constrain our predictions to respect the observed physical reality (i.e., that the probability of moving to another state should be monotonically increasing in Δt), we further formulate the problem as predicting transition *rates* [Voter, 2007]. We decompose this as a total rate λ and a categorical distribution over possible next states y; this corresponds to decomposing $P(S'|a)$ into a probability of any transition, $P(S' \neq s_0|s_0, a)$, and a distribution over which next state is chosen if a transition occurs, $P(S'|s_0, a, S' \neq s_0)$. We parameterize $P(S' \neq s_0|s_0, a)$ as $1 - e^{-\lambda \Delta t}$ – i.e., as an exponential CDF. If desired, per-neighbor rates are simply given by λy. We use then formulate maximum likelihood loss functions for the total rate and distribution over neighbors:

$$\mathbf{J}_{rate} = -\mathbb{1}_{(S' \neq s_0)} \cdot (-\lambda \Delta t) - \mathbb{1}_{(S' = s_0)} \log(1 - \exp(-\lambda \Delta t) \tag{1}$$

$$\mathbf{J}_{next} = -\mathbb{1}_{(S' \neq s_0)} (S' \cdot \log y) \tag{2}$$

$$\mathbf{J}_{total} = \mathbf{J}_{rate} + \mathbf{J}_{next} \,. \tag{3}$$

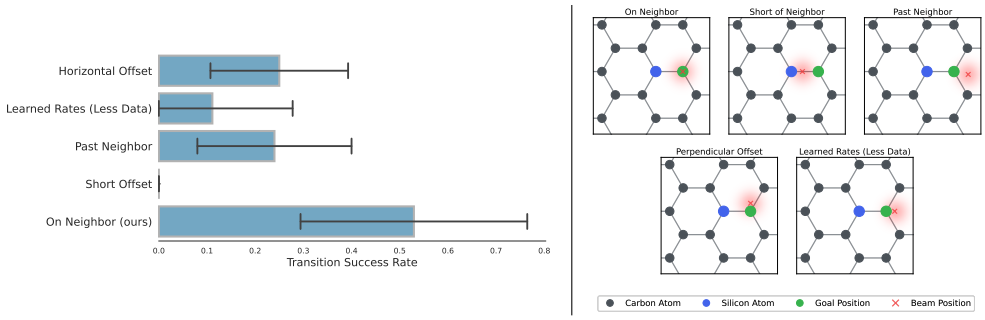

Figure 4: **Left:** The proportion of transitions that induced the intended transition in under a 5 second dwell. **Right:** The beam placement strategies considered in our experiments.

Given the above loss function, we trained a three-layer neural network using Adam with weight decay [Kingma and Ba, 2015] and ReLU hidden layer nonlinearities, trained with the cross-entropy loss for 500 epochs using batch size 256. We predicted λ with a softplus activation and y with a softmax. We whiten all inputs to the network prior to training, for stability.

To improve the robustness of our transition model to initializations, we trained an ensemble of transition predictors of bootstrap-resampled datasets. In addition to giving us the ability to esimate uncertainties, this significantly improved our overall accuracy and robustness To support more rapid inference, we also distilled this ensemble to a single transition predictor. By using widely-sampled random beam positions and a large number of training steps, this distillation could be made to fairly precisely match the ensemble predictions.

We display an example of the probability contours found by our system in the center panel of Figure 1. Note that we are overlaying three different probability distributions (one for each neighbour), where the colours are used to distinguish them. Our main finding confirms what was anecdotally held to be true by the microscopy community:

> To induce the dopant to transition to one of its neighbours, the optimal beam placement is directly on the neighbour, with a 50% probability of causing a transition with a five second dwell time.

Perhaps more relevant than dwell time is the number of electrons emitted, which is a function of both dwell time and the beam current. For our experiments we used a beam current of 90 pA, resulting in approximately 3 billion electrons in a five second period.

4 Empirical evaluation

While our main finding is consistent with previously held beliefs in the community, we do not have ground-truth data for the learned transition probabilities to quantitatively assess the accuracy of our predictions. However, as discussed in Section 2, the purpose of learning these transition probabilities is to be able to automate atomic manipulation, so in this section we evaluate the efficacy of our learned transition functions for this purpose. Specifically, our experiments are conducted as follows:

1. Start from a configuration with one dopant and 3-fold connections to neighbours, with the field-of-view (FOV) centered at the dopant.
2. Pick an arbitrary carbon atom as the goal position.
3. Focus the beam on a position dictated by one of the strategies (defined below).
4. Acquire an image to determine if we caused a transition.
5. Repeat above steps until the dopant has arrived at the goal position, or until we have reached the maximum allowable attempts.

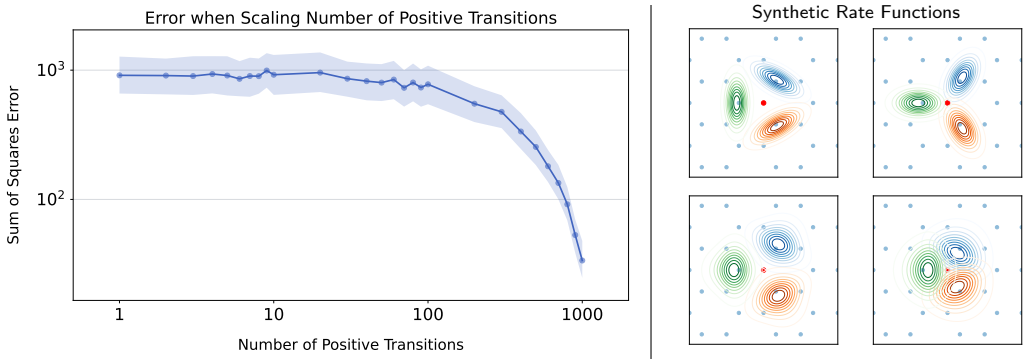

Figure 5: **Left:** Number of positive transitions in dataset versus the sum of squared prediction error. **Right:** A visualization of some of the synthetic rate functions overlayed on a sheet of graphene. The contours represent the rate of transitioning to the associated neighbor.

In order to determine that on-neighbour is in fact the optimal beam placement, we use a few different strategies for beam placement, detailed below and illustrated in Figure 4 (Right).

- **On neighbour:** our proposed optimal strategy
- **Short of neighbour:** place the beam in between the dopant and its neighbour)
- **Past neighbour:** place the beam beyond the neighbour atom
- **Perpendicular offset:** offset the beam perpendicularly from neighbour
- **Learned dynamics (less data):** we ran the learning method detailed in Section 3.4, but with approximately half of the collected data. When doing so, the resulting "optimal" beam placement was just past the neighbour.

For consistency, we used a constant 5s dwell time for all agents. We measured the number of times the electron beam was used to try and induce a dopant transition and report the findings in Figure 4. We observe that the on neighbor strategy induces the intended transition on average over 50% of the time, whereas the other approaches are below 25%.

5 In-depth empirical analyses

5.1 Synthetic data

To test the learning behavior of our transition model without facing the risk of overfitting to our relatively small real-world dataset, we generated many datasets of simulated microscope interactions. We used these synthetic datasets both to perform hyperparameter selection on our transition model and to demonstrate the scaling and learning performance of the model.

To generate these datasets, we sampled synthetic transition probability distributions as mixtures-of-Gaussians, each giving the non-normalized transition rates to neighboring states given a certain beam position (equivalent to the predicted per-state rates λy in our neural network model). Figure 5 (Right) is representative of the synthetic rate functions used throughout this section. Doing so grants us ground-truth data for transition probabilities, allowing us to quantitatively assess the accuracy of our learned transition model.

When creating a dataset, we first generate a simulated graphene sheet with a single dopant. We then uniformly sample random actions within 2Å of the silicon and simulate the transitions of the silicon according to the rate function, continuing until a certain number of positive transitions have been observed (As positive transitions are generally far rarer than negative transitions, they are the critical determiner of effective dataset size). To simplify comparisons between datasets, we enforce that the synthetic rate functions for each dataset have the same maximum value, preventing us from sampling datasets that are entirely positive transitions or almost entirely negative transitions.

We can then use a synthetic dataset to evaluate a learning algorithm by training on it and directly comparing its predictions (here, the predicted rates λy) to the known true values across a large grid of beam positions (a uniformly-spaced 2d grid surrounding the dopant). For simplicity, we report the sum of squared differences across this grid.[3]

5.2 Data scaling

To test the data scaling of our model – and show that it converges to near-perfect predictions in the limit of large datasets – we use this evaluation procedure at a range of data scales. We start by collecting 30 datasets with at least 1,000 positive transitions each. We then train our models on a range of subsampled scales from these datasets, and report a bootstrapped confidence intervals over our our 30 synthetic datasets at each scale in Figure 5 (Left). We find that while our model is largely incapable of learning from minimal amounts of data, it rapidly converges to near-perfect performance once hundreds of positive transitions are available, matching our own experience of the model's improvement as additional data was collected.

6 Related work

Atomic manipulation was first demonstrated by Stroscio and Eigler [1991] by using the tip of a scanning tunneling microscope (STM) to position individual Xenon atoms on the surface of a single crystal surface to form the IBM company logo. Further demonstrations, such as quantum corrals and molecular cascades, have demonstrated the potential of the method. Perhaps the application that has attracted the most interest is in using tip-induced atomic motion as an enabling tool for the fabrication of P and other atoms in Si qubits, the building blocks for quantum computers.

Despite the feasibility of manual control with STM tips, this type of atomic manipulation is limited to metallic/conducting surfaces. On the other hand, while STEMs can manipulate atoms embedded within a several layer thick specimen, it is still a rather haphazard and unpredictable process relative to using STM tips. To date, electron beam induced effects (with a STEM) have been studied purely by human operation, most typically by scanning a raster pattern (where the electron dose tends to be concentrated non-uniformly on one side of the image) in a selected field of view. The more sophisticated experiments involve manual positioning of the electron beam by a human, but this kind of motion is unpredictable and unreliable, and useful statistics are challenging (if not impossible) to glean from experiments conducted in this manner. We note that other experiments have been performed which control the electron beam in non-standard trajectories, effectively performing direct-write beam patterning processes – but with the critical point that the atomic landscape (i.e., position of atoms) is not considered [Dyck et al., 2023a,b,c].

The potential of the electron beams of scanning transmission electron microscopes to affect matter on the atomic level has been recognized since the early days of the technique. Most of these effects have been generally classified as a beam damage, denoting unwanted changes in materials structure induced by the beam. Indeed, minimization of beam damage, along with the need to increase spatial and energy resolutions, remains one of the three primary drivers behind STEM development, having spurred the high-voltage machines of the 1980s and 1990s and the aberration corrected low-voltage machines of the last two decades. It was also discovered that electron beam effects can be far more subtle, including crystallization and amorphization of oxides and semiconductors [Lulli and Merli, 1993, Yang et al., 1997, I. Jencic, 1995, Robertson and Jenčič, 1996, Frantz et al., 2001].

The emergence of aberration-corrected STEMs have made the atomic-resolution imaging relatively routine, and spurred a new wave of electron beam matter manipulation on the atomic level. The electron beam was shown to be able to deposit single atoms from chemisorbed species [van Dorp et al., 2012] and form ordered vacancy arrays [Jang et al., 2017]. Similarly, electron beams have been shown to induce direct atomic motion and creation of functional defects [Cretu et al., 2012, Yang et al., 2014, Susi et al., 2014].

[3] If the rate-based formulation is not desired, the same procedure could be changed to comparing transition probabilities.

The combination of simple beam control and feedback systems has enabled the direct assembly of crystalline materials with a single unit plane precision via directed crystallization and amorphization [Jesse et al., 2015]. These systems have also demonstrated potential for direct single atom dopant movement [Jesse et al., 2018], and finally, the controlled manipulation of Bismuth dopants in bulk silicon [Hudak et al., 2018]. In 2016, it was proposed that the combination of machine learning with electron beam manipulation can become a third paradigm for direct atomic construction [Kalinin et al., 2016]. In 2017, Dyck et al. [2017], Susi et al. [2017a] and Susi et al. [2017b] demonstrated single atom manipulation and insertion experiments for silicon in graphene [Susi et al., 2017b, Dyck et al., 2017, Tripathi et al., 2018, Dyck et al., 2018], an approach soon extended to direct atomic assembly of homo- [Dyck et al., 2018] and hetero-atomic artificial molecules [Dyck et al., 2019].

A number of theories for beam manipulation have been proposed, including those based on phonon-assisted knock-on and electronic excitations. However the causal relationship of the electron beam position relative to the silicon atom has only been suggested and demonstrated anecdotally. Physical intuition dictates that the damage mechanism is primarily through momentum transfer or so-called "knock-on" processes; therefore, the ideal placement of the electron beam would seem to be positioned exactly centered on a carbon (first) neighbor. While this is intuitive and appears to have been a successful route by multiple groups, damage mechanisms tend to be complex and are dictated by more than one process. For example, ionization or sputtering processes may be occurring as well, meaning it is unclear if the suggested beam position is actually the ideal one for inducing the most efficient transition of a silicon hop. Moreover, the anecdotal but physically intuitive rule of placing the electron beam on the center of a carbon neighbor is mostly valid only for a direct Si substitution (i.e., 3-fold coordinated silicon). For any other configuration, the rules are already not the same, and the optimal beam position for causing a transition event is not clear.

7 Conclusion

The last note left by Richard Feynman stated "What I cannot create, I do not understand." Building solid state quantum computers, creating nano-robots, and designing new classes of biological molecules and catalysts alike requires the capability to manipulate and assemble matter atom by atom, probe the resulting structures, and connect them to the macroscopic world; all this necessitates accurate estimates of the transition dynamics induced by sub-atomically focused electron beams. Until now, the elements of relevant knowledge have been limited to a few research groups, and atomic manipulation has been performed via direct control by human operator one beam positioning at a time. The characteristic timescale of human-operated experiments vastly exceeds the intrinsic latency of the electron microscope, for which hundreds of fabrications steps per second should be possible. Similarly, human control necessarily lacks precision, reproducibility, and systematic error correction capabilities. While sufficient for a proof of concept, atomic scale fabrication with the precision and throughput necessary for applications such as nanopore fabrication for protein sequencing, molecule screening platforms for physics and biology, and particularly quantum communication, sensing, and computing devices requires moving beyond the current human control paradigm.

Our work is a robust first step for determining transition probabilities via machine learning, and paves the way for further advances in this space. The scenario we considered in this work is somewhat idealized: we limit ourselves to single dopant and 3-way lattices for our learned dynamics. Nevertheless, these settings allowed us to confirm, via a *data-driven approach*, the commonly held belief that placing the electron beam directly on the neighbour has the highest probability of inducing a transition of the dopant. Going forward, we will be exploring broader settings: multiple dopants, graphene with 4-way connections and aberrations (such as holes).

References

O. Cretu, J.A. Rodríguez-Manzo, A. Demortière, and F. Banhart. Electron beam-induced formation and displacement of metal clusters on graphene, carbon nanotubes and amorphous

carbon. *Carbon*, 50(1):259–264, 2012.

Ondrej Dyck, Songkil Kim, Sergei V. Kalinin, and Stephen Jesse. Placing single atoms in graphene with a scanning transmission electron microscope. *Applied Physics Letters*, 111 (11):113104, 09 2017. ISSN 0003-6951. doi: 10.1063/1.4998599. URL https://doi.org/10.1063/1.4998599.

Ondrej Dyck, Songkil Kim, Elisa Jimenez-Izal, Anastassia N. Alexandrova, Sergei V. Kalinin, and Stephen Jesse. Building structures atom by atom via electron beam manipulation. *Small*, 14(38):1801771, 2018.

Ondrej Dyck, Maxim Ziatdinov, David B. Lingerfelt, Raymond R. Unocic, Bethany M. Hudak, Andrew R. Lupini, Stephen Jesse, and Sergei V. Kalinin. Atom-by-atom fabrication with electron beams. *Nature Reviews Materials*, 4(7):497–507, 2019.

Ondrej Dyck, Andrew R. Lupini, and Stephen Jesse. Atom-by-Atom Direct Writing. *Nano Letters*, (23):2339–2346, 2023a.

Ondrej Dyck, Andrew R. Lupini, and Stephen Jesse. A platform for atomic fabrication and in situ synthesis in a scanning transmission electron microscope. *Small Methods*, n/a(n/a): 2300401, 2023b.

Ondrej Dyck, Sinchul Yeom, Andrew R. Lupini, Jacob L. Swett, Dale Hensley, Mina Yoon, and Stephen Jesse. Top-down fabrication of atomic patterns in twisted bilayer graphene (adv. mater. 32/2023). *Advanced Materials*, 35(32):2370228, 2023c.

J. Frantz, J. Tarus, K. Nordlund, and J. Keinonen. Mechanism of electron-irradiation-induced recrystallization in si. *Phys. Rev. B*, 64:125313, Sep 2001. doi: 10.1103/PhysRevB.64.125313. URL https://link.aps.org/doi/10.1103/PhysRevB.64.125313.

Bethany M. Hudak, Jiaming Song, Hunter Sims, M. Claudia Troparevsky, Travis S. Humble, Sokrates T. Pantelides, Paul C. Snijders, and Andrew R. Lupini. Directed atom-by-atom assembly of dopants in silicon. *ACS Nano*, 12(6):5873–5879, 2018.

I. M. Robertson M. A. Kirk I. Jencic, M. W. Bench. Electron-beam-induced crystallization of isolated amorphous regions in si, ge, gap, and gaas. *Journal of Applied Physics*, 78(2): 974–982, 1995.

Jae Hyuck Jang, Young-Min Kim, Qian He, Rohan Mishra, Liang Qiao, Michael D. Biegalski, Andrew R. Lupini, Sokrates T. Pantelides, Stephen J. Pennycook, Sergei V. Kalinin, and Albina Y. Borisevich. In situ observation of oxygen vacancy dynamics and ordering in the epitaxial lacoo3 system. *ACS Nano*, 11(7):6942–6949, 2017.

Stephen Jesse, Qian He, Andrew R. Lupini, Donovan N. Leonard, Mark P. Oxley, Oleg Ovchinnikov, Raymond R. Unocic, Alexander Tselev, Miguel Fuentes-Cabrera, Bobby G. Sumpter, Stephen J. Pennycook, Sergei V. Kalinin, and Albina Y. Borisevich. Atomic-level sculpting of crystalline oxides: Toward bulk nanofabrication with single atomic plane precision. *Small*, 11(44):5895–5900, 2015.

Stephen Jesse, Bethany M Hudak, Eva Zarkadoula, Jiaming Song, Artem Maksov, Miguel Fuentes-Cabrera, Panchapakesan Ganesh, Ivan Kravchenko, Paul C Snijders, Andrew R Lupini, Albina Y Borisevich, and Sergei V Kalinin. Direct atomic fabrication and dopant positioning in si using electron beams with active real-time image-based feedback. *Nanotechnology*, 29(25):255303, apr 2018.

Sergei V. Kalinin, Albina Borisevich, and Stephen Jesse. Fire up the atom forge. *Nature*, 539(7630):485–487, 2016.

Diederik Kingma and Jimmy Ba. Adam: A method for stochastic optimization. In *International Conference on Learning Representations (ICLR)*, San Diega, CA, USA, 2015.

G. Lulli and P. G. Merli. Comparison of results and models of solid-phase epitaxial growth of implanted si layers induced by electron- and ion-beam irradiation. *Phys. Rev. B*, 47: 14023–14031, Jun 1993. doi: 10.1103/PhysRevB.47.14023. URL https://link.aps.org/doi/10.1103/PhysRevB.47.14023.

Alexander Markevich, Bethany Hudak, Andrew Lupini, and Toma Susi. Uncovering the mechanism for electron-beam manipulation of dopants in silicon. *Microscopy and Microanalysis*, 26(S2):2560–2561, 2020. doi: 10.1017/S1431927620022035.

Alexander Markevich, Bethany M. Hudak, Jacob Madsen, Jiaming Song, Paul C. Snijders, Andrew R. Lupini, and Toma Susi. Mechanism of electron-beam manipulation of single-dopant atoms in silicon. *The Journal of Physical Chemistry C*, 125(29):16041–16048, 07 2021.

I.M. Robertson and I. Jenčič. Regrowth of amorphous regions in semiconductors by sub-threshold electron beams. *Journal of Nuclear Materials*, 239:273–278, 1996. ISSN 0022-3115. doi: https://doi.org/10.1016/S0022-3115(96)00485-0. URL https://www.sciencedirect.com/science/article/pii/S0022311596004850. Proceedings of the JIM '95 Fall Annual Meeting (117th) on Lattice Defects and Radiation Induced Phenomena.

Kevin M. Roccapriore, Nicole Creange, Maxim Ziatdinov, and Sergei V. Kalinin. Identification and correction of temporal and spatial distortions in scanning transmission electron microscopy. *Ultramicroscopy*, 2021.

Kevin M Roccapriore, Max Schwarzer, Joshua Greaves, Jesse Farebrother, Rishabh Agarwal, Colton Bishop, Maxim Ziatdinov, Igor Mordatch, Ekin D Cubuk, Aaron Courville, et al. Discovering the electron beam induced transition rates for silicon dopants in graphene with deep neural networks in the stem, 2023.

Joseph A. Stroscio and D. M. Eigler. Atomic and molecular manipulation with the scanning tunneling microscope. *Science*, 254(5036):1319–1326, 1991. doi: 10.1126/science.254.5036.1319. URL https://www.science.org/doi/abs/10.1126/science.254.5036.1319.

Toma Susi, Jani Kotakoski, Demie Kepaptsoglou, Clemens Mangler, Tracy C. Lovejoy, Ondrej L. Krivanek, Recep Zan, Ursel Bangert, Paola Ayala, Jannik C. Meyer, and Quentin Ramasse. Silicon–carbon bond inversions driven by 60-kev electrons in graphene. *Phys. Rev. Lett.*, 113:115501, Sep 2014. doi: 10.1103/PhysRevLett.113.115501. URL https://link.aps.org/doi/10.1103/PhysRevLett.113.115501.

Toma Susi, Demie Kepaptsoglou, Yung-Chang Lin, Quentin M Ramasse, Jannik C Meyer, Kazu Suenaga, and Jani Kotakoski. Towards atomically precise manipulation of 2d nanostructures in the electron microscope. *2D Materials*, 4(4):042004, sep 2017a.

Toma Susi, Jannik C. Meyer, and Jani Kotakoski. Manipulating low-dimensional materials down to the level of single atoms with electron irradiation. *Ultramicroscopy*, 180:163–172, 2017b.

Mukesh Tripathi, Andreas Mittelberger, Nicholas A. Pike, Clemens Mangler, Jannik C. Meyer, Matthieu J. Verstraete, Jani Kotakoski, and Toma Susi. Electron-beam manipulation of silicon dopants in graphene. *Nano Letters*, 18(8):5319–5323, 2018.

Willem F. van Dorp, Xiaoyan Zhang, Ben L. Feringa, Thomas W. Hansen, Jakob B. Wagner, and Jeff Th. M. De Hosson. Molecule-by-molecule writing using a focused electron beam. *ACS Nano*, 6(11):10076–10081, 2012.

Arthur F Voter. Introduction to the kinetic monte carlo method. In *Radiation effects in solids*, pages 1–23. Springer, 2007.

Xiangxiu Yang, Renhui Wang, Heping Yan, and Ze Zhang. Low energy electron-beam-induced recrystallization of continuous gaas amorphous foils. *Materials Science and Engineering: B*, 49(1):5–13, 1997.

Zhiqing Yang, Lichang Yin, Jaekwang Lee, Wencai Ren, Hui-Ming Cheng, Hengqiang Ye, Sokrates T. Pantelides, Stephen J. Pennycook, and Matthew F. Chisholm. Direct observation of atomic dynamics and silicon doping at a topological defect in graphene. *Angewandte Chemie International Edition*, 53(34):8908–8912, 2014.