Title
Investigating Diffusion Using Video Scanning Tunneling Microscopy and Deep Learning

Author
Aikawa, Andrew Shunichi

Publication Date
2022

Investigating Diffusion Using Video Scanning Tunneling Microscopy and Deep Learning

by

Andrew Aikawa

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Physics

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Michael F. Crommie, Chair
Professor Jeffrey Neaton
Professor Felix R. Fischer

Summer 2022

Investigating Diffusion Using Video Scanning Tunneling Microscopy and Deep Learning

Abstract

Investigating Diffusion Using Video Scanning Tunneling Microscopy and Deep Learning

by

Andrew Aikawa

Doctor of Philosophy in Physics

University of California, Berkeley

Professor Michael F. Crommie, Chair

Diffusion is a ubiquitous phenomenon encompassing a wide range of physical systems from atomic scale motion to cellular locomotion. Diffusion plays a central role in practical applications such as electrochemistry, batteries, and thin film growth. The external control of diffusion properties presents new possibilities for understanding fundamental physical phenomena as well as a basis for next generation technological advances.

The contributions of this thesis include the fabrication, modeling, and characterization of diffusion on gate-tunable nanodevices. We first derive the fundamental results of diffusion which form the basis of our experimental analysis. Afterwards, we describe the procedures for fabrication and algorithms for characterization of diffusion in video scanning tunneling microscopy (STM). The physical platform upon which most of these results are based on is a single-layer graphene on hexagonal boron nitride (h-BN) field-effect transistor (FET) with fluorinated fluorinated tetracyanoquinodimethane (F4TCNQ) adsorbates, which were chosen due to favorable gate-tunable electronic properties. This device architecture enables control of diffusion through external temperature control, application of lateral and vertical electric fields, and substrate engineering. The analysis presented in this work provides rich sandbox for locally probing diffusion behavior in different system regimes.

In particular, we find that device level control allows us to tune the density of adsorbates on the device due to the proximity of molecular orbitals to the device Fermi level and the device classical capacitance. Charged molecules form uniform arrays in response to the gate voltage to screen the applied field. In addition, we see that gate voltage can directly control diffusion through the modification of transition state energies as a result of molecule charging.

We show that substrate level control in the form of moiré superlattice engineering can suppress diffusion by inducing anomalous subdiffusion in short time scales. Such behavior is explained using thermal equilibrium statistics for diffusion which show that the introduction of the moiré superlattice increases the number of unique energy states available to the diffus-

ing particle. As a result, the size of the state space increases which in turn increases the time required to visit all available states. For short times, the entirety of the state space is not visited with high likelihood, a violation of the thermodynamic hypothesis that all accessible states are equiprobable, a phenomenon known as *ergodicity breaking*. Despite the challenges inherent in uncovering this behavior under experimental constraints of data sparsity using canonical statistical methods, we demonstrate that it is possible to observe this phenomenon by using deep learning.

To family and friends. *We did it.*

Contents

List of Figures

List of Tables

List of Terms

AFM atomic force microscopy. 2, 62

CTRW continuous time random walk. 4, 9, 12, 13, 32, 41–43, 45, 46, 48, 51

DFT density functional theory. 65, 66

EBP ergodicity breaking parameter. 44–46, 48

EMSD ensemble averaged mean squared displacement. 12, 13, 43–46, 48

F4TCNQ fluorinated tetracyanoquinodimethane. 1, 15, 16, 18–20, 22–25, 27, 29, 31–34, 36, 46, 49, 50

FET field-effect transistor. iv, 1, 3, 15, 16, 18–23, 25–27, 46

FIB focused ion beam. 62, 64

h-BN hexagonal boron nitride. 1, 16, 20, 40, 41, 46, 53

i.i.d independent, identically distributed. 5, 7

LDOS local density of states. 17

LUMO lowest unoccupied molecular orbital. 15, 18, 22–24, 26–28

MSD mean squared displacement. 2, 5, 7, 29, 53

SiO_2 silicon dioxide. 16, 53

STM scanning tunneling microscopy. 1–3, 15, 18, 20–23, 25, 27, 29, 30, 34, 40, 46, 49, 51, 53, 62

STS scanning tunneling spectroscopy. 24, 25, 27, 28

TAMSD time averaged mean squared displacement. 12, 13, 45

UHV ultra-high vacuum. 15, 16, 20, 53

UMAP universal manifold approximation and projection. 43, 52

Acknowledgments

As I reflect on the past 5 years of my Ph.D., I want to acknowledge all the people who have supported me and helped make this journey possible.

I've had the chance to work with Mike since I was a freshman undergraduate at Berkeley. Mike instilled in me the importance of communicating science at an intuitive level. I'd like to thank Mike for his willingness to support my crazy ideas and for his 9 years of mentorship.

I want to recognize my qualifying exam/thesis committee members: Zi Qiu, my qual committee chair, and Jeffrey Neaton and Felix Fischer.

I was tremendously lucky to have Hsin-zon Tsai and Franklin Liou as my mentors and colleagues starting from undergrad and all throughout grad school. Because of their guidance, I was able to navigate the most stressful parts of my physics career, applying to grad school, prelims, and the qual. I will miss their company, banter, and conversations about nothing in particular.

I am grateful to have worked with Vida Jamali. It was from my collaborations with Vida that I was empowered to use my background in computer science to apply deep learning to my own physics research. I'm thankful for her endless support in my paper writing as well as professional guidance.

This work would not have been possible without the meticulous theory done by Johannes Lischner, Zachary Goodwin, and Young Woo Choi.

I want to recognize the undergraduates who helped me along the way, namely Eric Tang, Ethan Ha, Kyler Natividad, Brian Angeles, Kevin Qi, Michael Hu and Elvis Hsieh.

I thank professors Susumu Okada, Mina Maruyama, and Yanlin Gao for hosting me during my stay at the University of Tsukuba. My time in Japan gave me a closer look at how computational physics is done and taught me how to best collaborate with theorists. I'm grateful for their hospitality and and patience with my broken Japanese and letting me play with the dog! Some of my most memorable moments of graduate school was during my time in Japan.

This work was made possible thanks to my graduate school teachers. In particular, I want to recognize Ren Ng, for providing professional and academic guidance to me. I had shortly worked with Ren on an optometry research project which was my first heavy foray into research programming. It was from this formative experience that I gained the confidence to program for my own research projects and tackle seemingly impossible problems. During this time, I had the pleasure of working closely with Professor Austin Roorda, Dr. Emma Alexander, James Fong, Jay Shenoy, and Rishi Upadhyay.

I'm grateful for my graduate school teachers Thomas Courtade, Jonathan Shewchuk, Sergey Levine, Alexei Efros, Jitendra Malik, Stella Yu, Josh Hug, Deborah Nolan, Dung Hai Lee, Martin White, Robert Littlejohn, Steve Louie, and James Analytis.

I want to give a special thanks to my Japanese language teachers Kayoko Imagawa, Yuriko Miyamoto-Caltabiano, Chika Shibahara, and Yasuko Konno-Baker. Amidst the busy days of graduate school, Japanese class gave me a chance to take a break everyday. Thank you for filling every class with laughter and fun.

I thank my friends from Berkeley, including Raymond Ng, Brian Sang, Sang Hoon Maeng, Quincy Huynh, Sahil Hasan, Kevin Mo, Derek Phan, Ja Wattanawong, Takeshi Mochida, Douglas Wong, Tony Situ, Eason Chan, Bernard Zhao, Edric Xiang, Ryan Chan, George Lee, Tony Liu, Jason Perrin, William Ho, Mathew Cha, Harrison Kuo, Abizer Lokhandwala, Alex Welty, Britney Chieng, Peggy Shih, Jayne Yokoyama, Claudia Lim, Zoey Kenny, Weylan Wang, David Chen, Jessica Han, Steven Veshkini, John Okahata, Kacey Ching, Kelly Hamamoto, Kenyou Sasaki, Elaine Hsieh, Grace Yeo, James Liao, Alyssa Jolene, Jeremy Chan, Minami Sugiwara, Daniel Nguyen, Patrick Truong, John Ham, Jon Bertsch, Kenny Higa who have made my time at Berkeley so memorable.

I'm lucky to have my friends from back home, William Crosson, Dylan Feinhals, Edward Truong, Nicolas Ott. The time we spent over Discord helped make the pandemic a little less lonely. I want to thank them for listening to my practice qual talk as well as their constant support.

Finally I want to recognize my family especially Megan, mom, dad. Words are insufficient to describe the countless sacrifices they have made to bring me to this milestone. From the calls keeping me positive through stressful times, videos of my baby cousins to cheer me up, care packages to remind me of the food from home, and the spontaneous visits to the Bay, I was able to weather through my Ph.D. This thesis is dedicated to them.

Chapter 1

Introduction

Diffusion plays a central role in fundamental natural phenomena in addition to practical technologies. This ranges from understanding reaction rates in chemistry(*1*) to cellular motion(*2*). An understanding of diffusive properties has enabled an understanding of phenomena such as ionic motion in batteries(*3*) as well as growth selectivity in thin film growth(*4*). Even outside of physical sciences, diffusion has been used in understanding movement of financial asset prices(*5*) and even to generate fake images of faces(*6*).

For nanoscience, diffusion can be thought of as just one component of the wider goal of spatial control of individual atoms and molecules. Control over the motion of constituent matter presents exciting possibilities for new physics such as molecular assemblies and phase transitions. Oftentimes, especially for physical systems, the only lever to manipulate diffusion is temperature. It remains an open problem to determine the full range of physical quantities and properties that control diffusive behavior, but microscopic physical interactions that manifest diffusion present the possibility for material control. The work of this thesis focuses on how diffusion properties can be modified using nanomaterial properties and device control.

1.1 Diffusion Theory

Diffusion can be viewed from two perspectives: a stochastic, single particle view or a deterministic ensemble view. In the single particle picture, particles interact with random forces, which, on average, cause each particle to stray from its original position. In the macroscopic view, diffusion is characterized over a length scale significantly larger than the size of any single constituent. In this picture, matter is handled in terms of concentration gradients with the flow of mass going from high concentration to low concentration. While both pictures may seem incompatible, they predict the same ensemble behavior due to the law of large numbers. Over many particles, the behavior of the stochastic model converges with certainty to that predicted by macroscopic modeling.

Historically, microscopic models have found success in both experimental and computational condensed matter studies(*7*, *8*) for their simple dependence on a few key physical quantities. For crystals, these are just the lattice constant, temperature, diffusion barrier, and attempt frequency. Microscopic models are attractive for their flexibility to be extended and to other types of physical interactions since one only needs to consider the forces on a single particle rather than on how a concentration gradient should behave. These can include pair-wise interactions, external fields, and substrate effects.

One area of diffusion that has not been so well studied is “anomalous diffusion”. Whereas “normal” diffusion is characterized by a linear time scaling in the spread of particles, anomalous diffusion exhibits nonlinear dependence. While mathematical models based on probabilistic modeling of particle trapping and correlated motion exist, studies investigating the physical mechanisms responsible for anomalous diffusion are only recently emerging(*9*, *10*) due to the higher mathematical complexity of anomalous diffusion models and a lack of universal tools to map physical systems onto these models. This motivates the question of which physical properties give rise to anomalous diffusion and how can they be simulated and controlled.

1.2 Diffusion in Experiments

Direct characterization of diffusion from particle motion in videos typically revolves around identification of the mean squared displacement (MSD), a statistical quantity that measures the “spread” of particles over space. In order to extract MSDs, particle positions are first estimated from video frames and then trajectories are constructed by associating successive detection between consecutive frames(*11*).

The selection of the imaging modality for analyzing diffusion can significantly affect the viability of this type analysis. Whereas optical and even electron microscopy can capture up to hundreds of frames per second, raster scan imaging methods such as atomic force microscopy (AFM) and scanning tunneling microscopy (STM) can typically only capture quality frames several minutes at a time, although video STM studies have been done with high temporal resolution to characterize adsorbate diffusion on metals both in solution and in vacuum.(*7*, *12–16*) It has been demonstrated that it is possible to perform lateral atom tracking in STM with high temporal resolution(*17*). However, such studies with high frame rates appear to be limited to room temperature experiments. At low temperature, low temporal resolution combined with a finite field of view constrains analysis of STM videos to experiments containing only a few trajectories that are limited both in length and time span.(*18*) This can hamper the accuracy of canonical statistical methods and values which require a large number of samples to be accurate. These difficulties are further exacerbated by the possibility of anomalous modes of diffusion in which the functional form is permitted to be nonlinear in general, which begs the question of model selection.

Given the desirability for atomic scale resolution for direct observation of molecule and atomistic diffusion, this motivates the creation of a flexible toolkit for analyzing experimental diffusion data that can handle sparse trajectory data and different classes of diffusion.

1.3 Contributions of This Thesis

The goal of this work is to advance the understanding of diffusion and how to control diffusion in the context of nanoscience. This work focuses on the creation of devices through nanomaterial engineering for facilitating both diffusion manipulation and characterization. We demonstrate that it is possible through device and substrate level control to artificially engineer diffusion speeds and to tune between classical and anomalous modes of diffusion. Further, we show that it is possible to observe these phenomena despite the experimental constraints of STM imaging through the use of deep learning for model selection.

The technical contributions of this project are the following:

- The design of a tuneable platform for engineering a wide range of diffusion regimes through the creation of a graphene-hBN silicon-based field-effect transistor (FET). Gate tuneable adsorbates are used to demonstrate the ability for device level control to manipulate molecule charge states, adsorbate density, and diffusion rates,
- A physical model to explain the existence of subdiffusion for surface adsorbates which can be controlled through moire heterostructure engineering.
- A deep learning framework for analyzing experimental diffusion data for diffusion model selection.

1.4 Thesis Structure

This thesis is structured as follows:

- In chapter 2, I review diffusion theory, highlighting some key results of classical/anomalous diffusion.
- In chapter 3, I review the experimental techniques and characterization necessary for graphene field-effect transistor device fabrication and video STM. Gate control is demonstrated through the imaging of variable density molecule arrays.
- In chapter 4, I share results demonstrating the ability to tune diffusivity through the use of device level controls. Our experimental results here are corroborated by diffusion barrier calculations which demonstrate a gate dependent diffusion which switches with molecular charge state.
- In chapter 5, I present the application of deep learning on the problem of diffusion model selection. I explain the observation of ergodicity breaking subdiffusion as a result of substrate level control over moire twist angles. Subdiffusion is explained using a Markov chain model. Here, we use deep learning to handle the problem of sample availability in experiment in order to identify anomalous modes of diffusion.

Chapter 2

Fundamentals of Classical and Anomalous Diffusion

The diffusion of particles from high concentration to low concentration is the result of the stochastic motion of single particles induced by thermal coupling with the environment. Much of what is understood about macroscopic diffusive motion is explained by single-particle stochastic dynamics. Robert Brown initially observed the diffusion of suspended pollen in the early 19th century, commenting:

> These motions were such as to satisfy me, after frequently repeated observation, that they arose neither from currents in the fluid, nor from its gradual evaporation, but belonged to the particle itself. -**Robert Brown**, 1828 (*19*)

It was not until the 20th century that Einstein and Smoluchowski successfully created mathematical models to explain the motion of molecules in probabilistic terms. Their models connected physical quantities to the microscopic pictures of diffusion and the macroscopic results of Fick's second law of diffusion.

Meanwhile, deviations from Brownian motion, collectively referred to as *anomalous diffusion*, have been observed in systems, including crowded cellular motion(*2*), single atom diffusion on metals(*9*), and nanoparticle surface diffusion in liquids(*10*). Such anomalous diffusion models extend the classical Brownian motion to model additional effects such as particle trapping from defects and correlated particle motion. While the set of all anomalous diffusion processes contains many different types, this thesis will only focus on one kind, the subdiffusive continuous time random walk (CTRW), from the perspective of probabilistic, atomistic modeling.

2.1 Random Walks

The simplest atomistic model of particle diffusion is the 1D discrete unbiased random walk. To illustrate the physical picture behind the discrete random walk, we imagine a particle

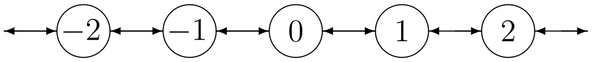

Figure 2.1: Random Walk Markov Chain

moving in a lattice that is thermally coupled to its environment and driven by thermal forces. At equilibrium, the particle has no preference to move left or right. A Markov chain, like the one illustrated in Figure 2.1, can describe this behavior where the state represents the current position of the particle within the lattice and the allowed transitions are nearest neighbors with equal probability of moving left or right. Put mathematically, the displacement at a time, $t \in \mathbb{Z}^+$, of the particle, X_t, is:

$$X_t = \sum_{i=1}^{t} \Delta_i \tag{2.1}$$

$\{\Delta_i\}$ represents the nearest neighbor transition and are characterized by independent, identically distributed (i.i.d) Rademacher random variables. Having an identical and independent distribution over the steps assumes that the statistics of motion are stationary over time and that the particle's future position is uncorrelated with its past given its present state.

In order to verify experimental results of diffusion, we are interested in expectation values since these can be empirically estimated by taking sample averages. For the random walk process, the first and second moments respectively of the displacement vector are (in units of the lattice constant):

$$\mathbb{E}[X_t] = \mathbb{E}\left[\sum_{i=1}^{t} \Delta_i\right] = 0 \tag{2.2}$$

$$\mathbb{E}[X_t^2] = \mathbb{E}\left[\left(\sum_{i=1}^{t} \Delta_i\right)^2\right] = t \tag{2.3}$$

The equations 2.2 and 2.3 are key results from diffusion theory which predicts that the variance of particle position, which measures how far diffusive particles "spread out", is proportional to time. This variance over positions in equation 2.3 is known as the mean squared displacement (MSD). While this derivation was for one dimension, the result remains the same for any finite number of dimensions. This can be understood by recognizing that the squared displacement in multiple dimensions is a sum of one-dimensional processes. For example, in a d-dimensional isotropic random walk, the particle can move along either the positive or negative direction of any dimension independent of the others. Here i represents the dimensional index.

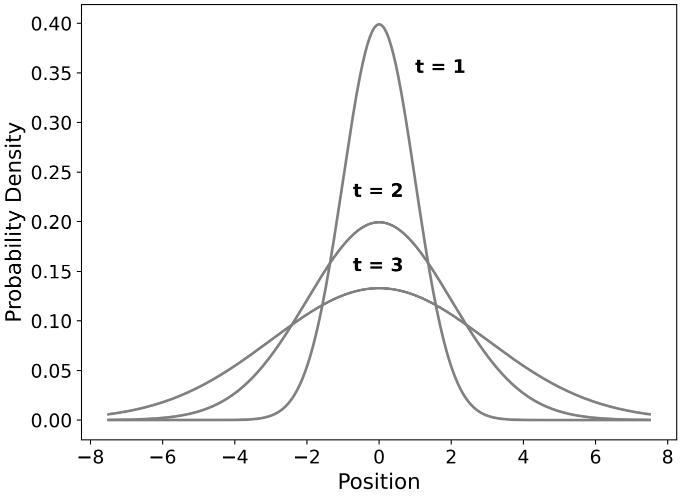

Figure 2.2: Time evolution of a 1D Brownian particle with $D = 1$. The probability density can also be interpreted in the sense of a physical density when talking about large numbers of particles such as concentrations in solution

$$r^2(t) = \sum_{i=1}^{d} X_i^2(t) \tag{2.4}$$

$$\mathbb{E}[r^2(t)] = \sum_{i=1}^{d} \mathbb{E}[X_i^2(t)] = \sum_{i=1}^{d} t = dt \tag{2.5}$$

2.2 Gaussians in Classical Diffusion

When looking at stochastic quantities such as a diffusive particle's position, we are often interested in the probability distribution of values. Distributions over positions provide a test for identifying random processes from observations.

Classical diffusion follows a Gaussian distribution which can be derived from different perspectives. Taking the macroscopic perspective, we start from Fick's second law of diffusion:

$$\frac{\partial \psi}{\partial t} = D \nabla_{\mathbf{x}}^2 \psi(\mathbf{x}, t) \tag{2.6}$$

with $\psi(\mathbf{x}, t)$ as the particle concentration, t as time, $\mathbf{x}$ as position, and D is a quantity known as the diffusion coefficient or diffusion constant and quantifies the proportion between particle flux to a concentration gradient. Given a point source initial boundary condition, using Green's functions, the solution to this differential equation is in d dimensions:

$$\psi(\mathbf{x},t) = \left(\frac{1}{4\pi Dt}\right)^{d/2} \exp\left(\frac{-|\mathbf{x}|^{\mathbf{2}}}{4Dt}\right) = \prod_{i=1}^{d}\left(\frac{1}{4\pi Dt}\right)^{1/2} \exp\left(\frac{-x_i^2}{4Dt}\right) \tag{2.7}$$

From equation 2.7, we see that the multidimensional concentration reduces to the product of independent, identically distributed 1D gaussians, $\mathcal{N}(0, 2Dt)$, and establishes the connection between the variance derived using the atomistic random walk model and macroscopic perspective of diffusion. It follows from this that the MSD is $2Dtd$. While D doesn't make an appearance in equation 2.3, this is due to the choice of units to match the lattice dimension and frequency of hops such that $D = 1/2$.

We can similarly recover a Gaussian if we consider the atomistic random walk model. Notice that the equation for displacements (eqn. 2.1) is the sum of i.i.d random variables. We use the central limit theorem as shown below:

Theorem 1 *If* $(X_n)_{n\in\mathbb{N}}$ *is a sequence of i.i.d random variables with common mean* μ *and finite variance* σ^2*, then in convergence in distribution*

$$\frac{\sum_{i=1}^{n} X_i - n\mu}{\sigma\sqrt{n}} \xrightarrow[n\to\infty]{d} \mathcal{N}(\mu, 1) \tag{2.8}$$

Using the theorem 1 and $\mu = 0, \sigma^2 = 2D, n = t$, we have for the distribution over displacements.

$$X_t = \sum_{i=1}^{t} \Delta_i \xrightarrow[\text{t}\to\infty]{\text{d}} \mathcal{N}(0, 2Dt) \tag{2.9}$$

2.3 The Physical Picture of Surface Diffusion

With the key quantity defining diffusion being the diffusion constant, a natural question that arises when studying diffusive systems is how is the diffusion constant physically determined, which generally depends on the physical system being considered. For surface diffusion on metals in particular (*20*), a typical model for diffusion is a particle coupled with a thermal bath trapped in a periodic potential landscape. In this picture, adsorbates oscillate within a lattice minimum until receiving a thermal excitation sufficient to overcome the energy barrier between lattice sites, E_{diff}, a quantity known as the diffusion barrier. To model the rate an adsorbate can escape the lattice site, we use Boltzmann statistics:

$$\Gamma = \nu \exp\left\{\frac{-E_{diff}}{k_B T}\right\} \tag{2.10}$$

where Γ represents the escape rate, ν is a temperature independent attempt frequency which roughly corresponds to the oscillation of the particle within the trap and how often it tries to

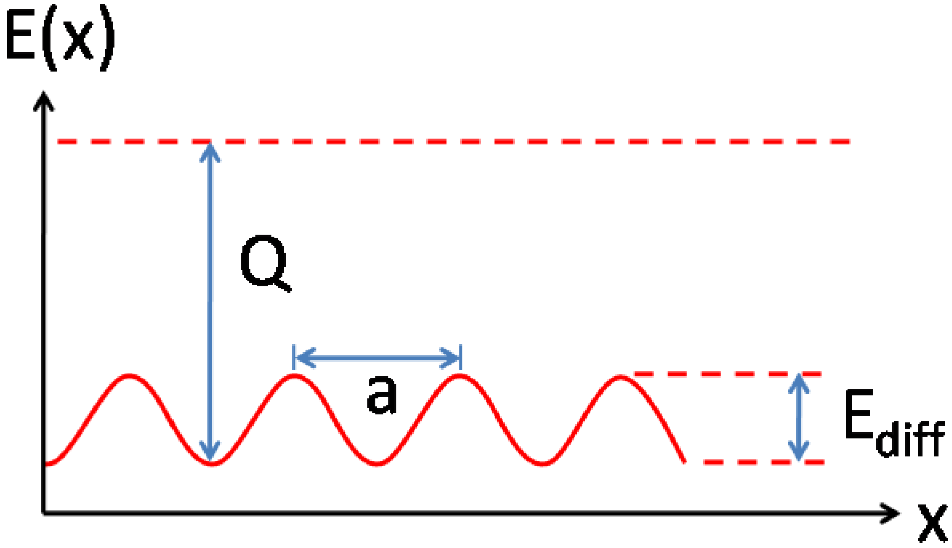

Figure 2.3: Adsorption energy landscape for diffusion in one dimension. The vacuum level, Q, lies far above the diffusion barrier energy, E_{diff}, shown for a lattice with a spatial period of a.

overcome the diffusion barrier, and k_B and T are the Boltzmann constant and temperature respectively. Equation 2.10 models the effective hopping rate by just taking the successful fraction of all attempts made by the particle per unit time. The quantity Γt thus encodes the total number of hops made by the particle up to time t. It follows from the discussion of equations 2.3 and 2.7 that by equating the variances of the random walk model with the solution from Fick's second law, we can express the MSD and diffusion constant solely in terms of the physical system parameters.

$$\mathbb{E}[r_t^2] = 2Dtd = a^2\Gamma t \tag{2.11}$$

$$D = \frac{\Gamma a^2}{2d} \tag{2.12}$$

2.4 Anomalous Subdiffusion: Continuous Time Random Walks

While the aforementioned diffusion models, which I will collectively refer to as "classical diffusion", have been successfully applied to a wide range of different physical systems, these models are not general enough to describe all types of physical diffusion. Diffusion processes where the MSD grows sublinearly with time, so called subdiffusive processes, have been

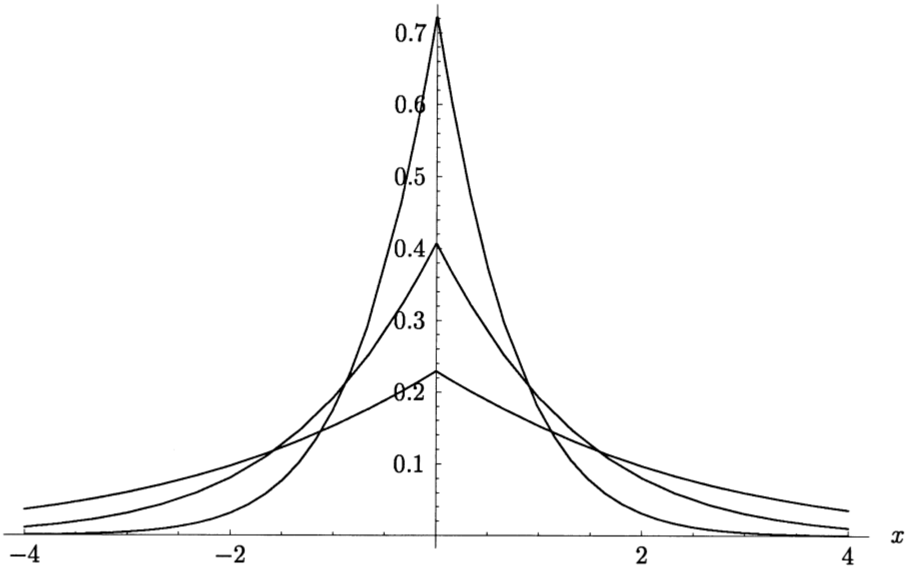

Figure 2.4: Probability density function (PDF) for a subdiffusive continuous time random walk. CTRW with anomalous exponent $\alpha = 1/2$, drawn for the consecutive times $t = 0.1, 1, 10$. The cusp shape of the PDF is distinct(*25*)

modeled mathematically (*21–23*) and observed in various physical systems(*9*, *24*). A subset of these processes are referred to as subdiffusive continuous time random walk (CTRW), which at a high level aim to model the phenomenon of inhomogeneous diffusion or diffusion in the presence of particle traps.

Counting Processes and Wait Times

In the previous treatment of diffusion, we looked at the evolution of the particle position in terms of discrete time. However, since a particle's position is defined in continuous time, it's desirable to have a description of the diffusion process in terms of continuous time. In order to factor this change into our model, we can incorporate a counting process into our random walk. A counting process, $N(t)$, is a stochastic process that counts the number of events that have occurred up to some time t, which in the context of diffusion counts the number of hops up to time t. With this change, equation 2.1 becomes:

$$X(t) = \sum_{i=1}^{N(t)} \Delta_i \tag{2.13}$$

If the spatial portion of the diffusion process, $\{\Delta_i\}$, is independent from the temporal component, $N(t)$, we call this diffusion process separable. For separable diffusion, we use Wald's equation to obtain the following result for the MSD similar to equation 2.3.

$$\mathbb{E}[X_t^2] = \mathbb{E}\left[\left(\sum_{i=1}^{N(t)} \Delta_i\right)^2\right] = \mathbb{E}[N(t)]\mathbb{E}[\Delta_i^2] \tag{2.14}$$

Equation 2.14 offers a similar insight as equation 2.11 demonstrating that the lattice length scale and the hopping rates are simply multiplied together to obtain the mean squared displacement. However, it generalizes the concept of having equally spaced hops in time to any counting process in general.

For the case where events occur with some well defined positive rate, say a Poisson process of rate λ and holding/waiting times $\{T_1, T_2, ...\}$, which we could choose to model the jumping process of a Brownian particle, the number of hopping events is proportional to time $N(t) = \lambda t$. This result is derived from Renewal Theory as,(*26*)

$$\lim_{t\to\infty} \frac{N(t)}{t} \overset{a.s.}{\to} \frac{1}{\mathbb{E}[T_i]} = \lambda \tag{2.15}$$

where the limit converges almost surely. While the above equation 2.15 is clearly true, the above result proves an important property about Brownian motion known as *ergodicity*. The left-hand side of the limit is an expression about a single particular realization whereas the right-hand side is a measure over an ensemble, indicated by the expectation value over waiting times. This says that we can completely characterize the statistics of a Brownian process by tracking a single particle for a long time rather than requiring many distinct samples.

The previous derivation critically relied on the average waiting times having a finite mean. If we instead consider a wait time distribution with infinite mean(*28*), for example a distribution with a heavy tail, $\psi(t) := \psi_{T_i}(t) \sim \frac{1}{t^{\alpha+1}}$, defining(*29*)

$$\Psi(t) = \int_t^\infty \psi(s)ds \tag{2.16}$$

which is the probability no hops occur before time t. Letting $P(N,t) = P(N(t) = N)$, It follows that

$$P(N(t) = 1) = \int_0^\infty \psi(s)\Psi(t-s)ds = (\psi \circledast \Psi)(t) \tag{2.17}$$

which is a convolution. Doing an N-fold convolution yields $P(N,t) = (\psi^{\circledast N}\Psi)(t)$. Using a Laplace transform and integrating by parts gives

$$\tilde{P}(N,s) = \tilde{\psi}(s)^N \tilde{\Psi}(s) = \tilde{\psi}(s)^N \frac{1-\tilde{\psi}(s)}{s} \tag{2.18}$$

We can use this expression to obtain the expected value counts, $\mathbb{E}[N(t)]$, but in Laplace space $\mathbb{E}[\tilde{N}(s)]$:

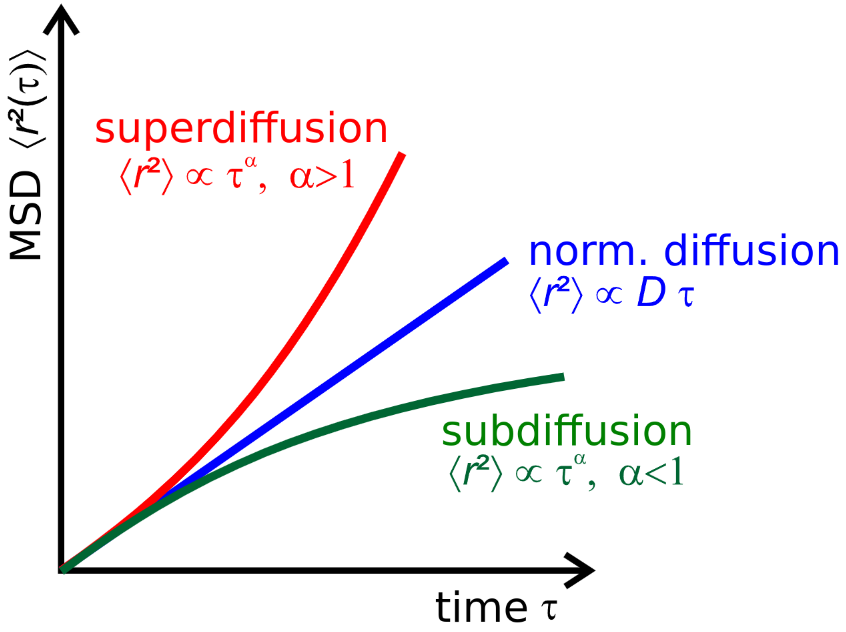

Figure 2.5: Mean squared displacements for different types of diffusion (*27*).

$$\begin{aligned}
\mathbb{E}[\tilde{N}(s)] &:= \sum_{N=0}^{\infty} \tilde{P}(N,s)N && (2.19)\\
&= \frac{1-\tilde{\psi}(s)}{s} \sum_{N=0}^{\infty} N\tilde{\psi}(s)^N && (2.20)\\
&= \frac{\tilde{\psi}(s)}{s(1-\tilde{\psi}(s))} && (2.21)
\end{aligned}$$

For diffusion we're often interested in the long term behavior where $t \to \infty$, which corresponds to the $s \to 0$ limit in Laplace space. The Tauberian theorem relates the asymptotics of a function to its representation in Laplace domain, in particular

$$s \to 0 \Longleftrightarrow t \to \infty \tag{2.22}$$

$$\tilde{\psi}(s) \sim 1 - As^{\alpha} \Longleftrightarrow \psi(t) \sim \frac{1}{t^{1+\alpha}}\ (0 < \alpha < 0) \tag{2.23}$$

$$\tilde{\psi}(s) \sim s^{\beta} \Longleftrightarrow \psi(t) \sim \frac{t^{\beta-1}}{\Gamma(\beta)}\ (\beta > 0) \tag{2.24}$$

If we have a heavy tailed wait time distribution, upon substituting into the counts equation 2.21, the resulting scaling law is,

$$\mathbb{E}[\tilde{N}(s)] \sim s^{-(\alpha+1)}/A \Longleftrightarrow \mathbb{E}[N(t)] \sim t^{\alpha} \ (0 < \alpha < 1) \tag{2.25}$$

yielding a sublinear mean squared displacement characteristic of subdiffusion.

2.5 Ergodicity

A key assumption underlying much of statistical mechanics is the idea of *ergodicity*, which means that at thermal equilibrium, the likelihood of a system being in any given state depends only on the energy of the state, which is said to follow Boltzmann statistics. A random process is considered ergodic if the time averaged statistics of a single sample over a long period of time is identical to that over an ensemble of samples. An important question to ask for a number of physical systems is whether the system will converge to its Boltzmann distribution and the speed of convergence.

For diffusion, it is common to use the time averaged mean squared displacement (TAMSD) to characterize a physical system, defined mathematically as:

$$\overline{\delta^2}(\Delta, t) = \frac{\int_0^{t-\Delta} [x(t' + \Delta) - x(t')]^2 dt}{t - \Delta} \tag{2.26}$$

which just averages over all displacements occurring over a window in time of length Δ for a system that has been measured up to time t. From our previous treatment of classical and anomalous subdiffusive random walks, we saw that the time scaling was only dependent on the hopping process, scaling as $\mathbb{E}[N(t)] \sim t^{\alpha}$. Taking a time average of the hopping process yields a linear scaling law with the window size for $\Delta \ll t$, for which the TAMSD and ensemble averaged mean squared displacement (EMSD) are unequal in the case of a subdiffusive random walk.

$$\mathbb{E}[\overline{\delta^2}(\Delta, t)] \sim \mathbb{E}[N(t + \Delta) - N(t)] = \mathbb{E}[N(t + \Delta)] - \mathbb{E}[N(t)] \tag{2.27}$$

$$\sim \frac{t^{1+\alpha} - \Delta^{1+\alpha} - (t - \Delta)^{1+\alpha}}{(1 + \alpha)(t - \Delta)} \tag{2.28}$$

Taking the limit of $\Delta \ll t$

$$\mathbb{E}[\overline{\delta^2}(\Delta, t)] \sim \frac{\Delta}{t^{1-\alpha}} \tag{2.29}$$

From 2.29 it's evident that at smaller lag times the linear behavior of the TAMSD can never match the sublinearity of the EMSD for a subdiffusive CTRW.

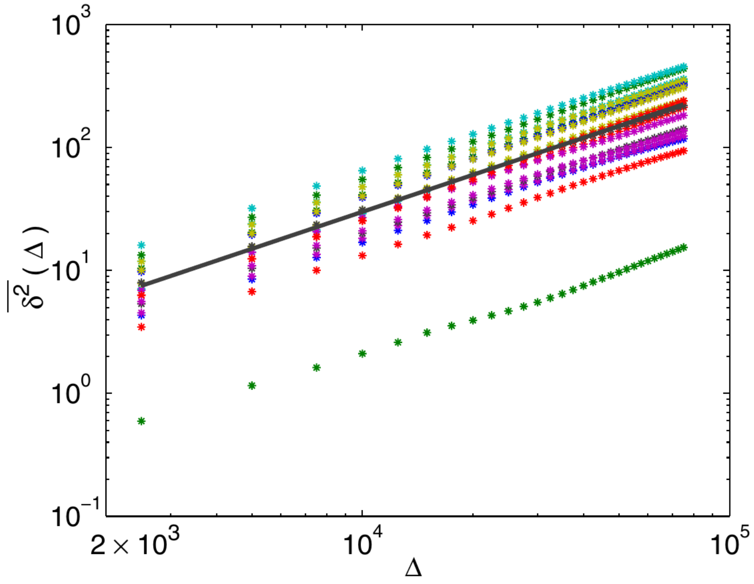

Figure 2.6: TAMSD Simulations of the subdiffusive CTRW process. TAMSDs for a CTRW with $\alpha = 3/4$ and free boundary conditions. Different realizations of the CTRW result in different estimated diffusion constants. (*24*)

2.6 Discussion

This chapter presents atomistic models of both classical and anomalous diffusion. To create these models, we argued that steps occur at random times and are uncorrelated. By permitting different types of hopping processes we found different resulting scaling relations for the mean squared displacement. This theoretical toolkit provides a generalized framework for understanding the physics of stochastic motion. These models, which rely only on probabilistic constructions, can be used to describe a wide array of systems.

Diffusion provides a rich theoretical framework for answering fundamental questions about non-equilibrium behavior through the lens of ergodicity. An important question when studying non-equilibrium systems is the time scale, if it exists, required to relax to equilibrium conditions. (*30*) For a subdiffusive CTRW, we showed that there existed a non-ergodic phase throughout all time, a result of the aging of a CTRW process (*24*, *31*). However, for real physical systems, there are many deviations from the CTRW model(*32*), such as smooth cutoffs in the waiting time distribution due to finite bounds on energy depth of particle trapping sites, which allow the system to relax to an ergodic state given infinite time, albeit slowly. By comparing the TAMSD and EMSD of a system, the relaxation times of a stochastic system can be estimated.

A challenge that remains when understanding real physical systems is how microscopic physical properties and interactions manifest into realizations of the models discussed here.

In chapter 4, we use this framework to identify and characterize gate-tuneable diffusion quantities. By first identifying the correct diffusion model, we can infer or propose the types of microscopic physical interactions that give rise to the ensemble macroscopic diffusion observations. In chapter 5, we explore this problem of model selection, first on a proposed physical diffusion model and then on real experimental data through deep learning and statistical tests, which are based on the properties discussed here.

Chapter 3

Fabrication and Characterization of Gate Tuneable Molecule Arrays Through Scanning Tunneling Microscopy

This chapter focuses on methods for creating and characterizing molecules on a gate tuneable device. Here, we have used scanning tunneling microscopy (STM) to image reversible, gate-tuneable control of the surface concentration of charged molecules on a graphene field-effect transistor (FET). We demonstrate precise, all-electrical control over the inter-particle density of charged F4TCNQ molecular arrays on graphene FETs held in ultra-high vacuum (UHV) at T = 4K. The resulting tuneable surface arrays enable controlled charge doping and Fermi-level pinning of the graphene FET, and provide a new method for determining molecular energy level alignment based on measuring molecular surface concentration. The mechanism underlying gate-based tuning of molecular arrays is well-described by a simple capacitor-based model in which molecules each carry one electron of charge and dynamically rearrange themselves to screen back-gate-induced electric fields and reduce overall electronic energy. Ionic surface array concentrations are fully determined by the device back-gate voltage, the device geometric capacitance, and the energy difference between the Dirac point of the graphene device and the lowest unoccupied molecular orbital (LUMO). Much of the work here is presented in our published manuscript: F. Liou et. al. "Imaging Reconfigurable Molecular Concentration on a Graphene Field-Effect Transistor". (*33*)

3.1 Related Work

Ions are an effective tool for modifying the behavior of surfaces and 2D materials.(*34–37*) For surfaces a common route to control ion behavior is through techniques involving electrochemical cells and electrolyte solutions.(*38–40*) This has enabled advancements in the understanding of ionic surface self-assembly,(*41*) surface diffusion,(*7*, *42*) and catalysis.(*43*) For 2D materials ions are employed largely as intercalants and there is great excitement at the prospect of using them for superdense energy storage(*44*, *45*) to drive useful phase-change processes,(*46*) and to create neuromorphic memristive devices.(*47*) At the intersection of surface science and 2D materials, however, there is less activity focused on controlling ion motion in single-layer devices at the individual ion level.(*48*) This topic is important since better control of ionic behavior at device surfaces could enable ions to be positioned into useful new device-relevant configurations.(*49*, *50*) In some sense the Coulomb field of a single ion can be thought of as the smallest possible "local gate" (in the parlance of device physicists),(*51*) and deterministic placement of ions at a device surface thus represents the ultimate microscopic limit of electrostatic control. For individual ions to be effective elements in 2D nanodevices, however, there are some practical matters to consider. For example, the control of ionic mechanical arrangements should occur in the absence of bulky electrolyte solutions, and precise ionic surface concentration and placement should ideally be responsive to applied electric fields and currents within the device itself.

3.2 Fabrication of Graphene - Hexagonal Boron Nitride Field-Effect Transistors with F4TNCQ

Gate tuneable graphene/h-BN FET devices are composed of a conductive channel material (graphene), dielectric material (h-BN/SiO_2), contact material (Au) and the device package (sample mounting and wire bonding). The content here is based on our published paper: H.S. Jung et al., "Fabrication of Gate-tuneable Graphene Devices for Scanning Tunneling Microscopy Studies with Coulomb Impurities".(*52*)

The device consists of (from top to bottom): F4TCNQ adsorbates, graphene, h-BN, SiO_2 and bulk silicon. h-BN acts as an insulating thin film with an atomically flat and electrically uniform substrate for graphene.(*53*, *54*) SiO_2 provides a dielectric layer, while bulk silicon acts as a back-gate.

Single layer graphene is placed on top of a h-BN/SiO_2/Si chip via mechanical exfoliation.(*55*) Electrical source and drain contacts were fabricated by depositing 3nm Cr and 10 nm Au through a stencil mask. The silicon substrate was used as the back-gate. The completed device geometry is illustrated in Figure 3.1. After placing in UHV, the graphene surface was cleaned by high-temperature annealing in vacuum at 400C for 12 hours. F4TCNQ molecules were deposited onto the device surface while holding it at room temperature by evaporation from a Knudsen cell.

3.3 Principles of Scanning Tunneling Microscopy and Spectroscopy

Scanning Tunneling Microscopy enables atomically resolved characterization of topography and electronic structure of materials through the phenomenon of quantum tunneling. Based

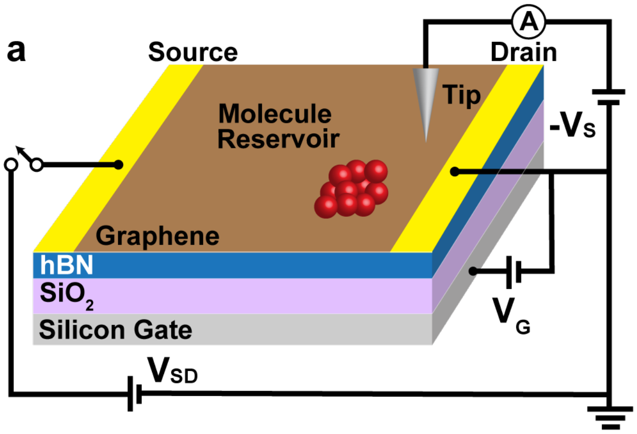

Figure 3.1: Graphene FET device geometry.

on the design developed by Binning, Rohrer, Gerber and Weibel(*56*). The setup uses a metal probe connected to a set of piezo motors that operate laterally and vertically. The probe is brought near (∼1nm) a conductive surface with a finite voltage applied across the probe and the sample and electrons can flow can tunnel across the vacuum barrier between the tip and the sample with exponential sensitivity to the tip-sample distance. In constant current mode, the feedback loop is used to maintain this tunneling current whilst adjusting the vertical piezo. To create an image, the sample is raster scanned with the probe while recording the response of the vertical piezo motor.

The theoretical description of quantum tunneling was created by Bardeen, Tersoff, and Harmann.(*57*) The transition rate of electrons from the tip to the sample can be described by Fermi's golden rule:(*58*)

$$W_{t\to s} = \frac{2\pi}{\hbar}|M_{s,t}|^2\delta(E_F - E) \tag{3.1}$$

where the tunneling matrix element is $M_{s,t} := \langle\psi_s|\, H'\, |\psi_t\rangle$, and the sample and tip states are, $|\psi_s\rangle$ and $|\psi_t\rangle$ respectively, and H' is the first order perturbation Hamiltonian describing the tip-sample junction. The total current can be obtained by summing over the contribution of sample surface states. Tersoff showed that in the limit of small voltage and temperature, for a point probe, the current across the junction is:

$$I \propto_s \sum_s |\psi_s(\mathbf{r}_0)|^2\delta(E_s - E_F) \tag{3.2}$$

where $\mathbf{r}_0$ is the position of the point probe, E_s is the energy of state ψ_s, and E_F is the Fermi level. This equation expresses that the current is proportional to the surface local density of states (LDOS) at the Fermi level. Tersoff shows that

$$|\psi_s(\mathbf{r}_0)|^2 \propto \exp\{-2\kappa d\} \tag{3.3}$$

where $\kappa = \hbar^{-1}(2m\phi)^{1/2}$ is the minimum inverse decay length for wave fucntions in vacuum, ϕ is the work function, and d is distance from the tip to the surface. This demonstrates the exponential sensitivity of the tunneling current to the tip-sample distance. Often, we want to probe the local electronic structure in STM. When a bias voltage, V, is applied on the sample, the band structure of the sample is shifted by eV, which shifts energy values of the Fermi-Dirac distribution and density of states with respect to the Fermi level. This modifies the current by including the participation of states up to energies of eV. This can be summed over by integrating over the continuum of tip energies (assuming constant density of states). At low temperature, the Fermi-Dirac distribution is a unit function from 0 to eV and zero otherwise.

$$I(V) \propto \int_0^{eV} |\psi_s(\mathbf{r}_0)|^2 \delta(E_s - eV - E_F) \tag{3.4}$$

We can measure the differential conductance to perform Scanning Tunneling Spectrocopy by taking the derivative, which yields:

$$\frac{dI}{dV} \propto_s |\psi_s(\mathbf{r}_0)|^2 \delta(E_s - eV - E_F) = \mathrm{LDOS}(E_F + eV) \tag{3.5}$$

3.4 Gate Tuneable Density of Molecules

We demonstrate reversible nanoscale control over the concentration of ions at the surface of a graphene field-effect transistor (FET) by applying current and voltage signals to the FET in the absence of any electrochemical solvents. Reversible control over the surface concentration of charged F4TCNQ molecules was accomplished by simultaneously applying a source-drain current and an electrostatic back gate voltage to conventional graphene FET devices at cryogenic temperature. Under these conditions the primary effect of the source-drain current is to heat the graphene surface and promote molecular diffusion. When an electrostatic back-gate is simultaneously applied with the source-drain current, the diffusing molecules adjust their surface density in response to the gate voltage. Setting the source-drain current to zero causes the molecules to freeze in place so that their gate-adjusted density can be measured directly via STM imaging. This “freeze-frame” imaging technique allows us to directly visualize the effects of molecular motion in response to electrical signals applied to our devices. We observe that the surface density of charged molecules varies linearly with applied back-gate voltage, suggesting that the molecules contribute to the screening of the gate electric field. The graphene Fermi level also appears to be pinned by the F4TCNQ lowest unoccupied molecular orbital (LUMO) energy level as the gate voltage is changed during the density manipulation process. This implies that charged molecules on the FET surface dynamically change their concentration in order to suppress energetically unfavorable filling of graphene band states. The connection established here between molecular electronic structure and molecular surface concentration provides a new technique for quantitatively determining molecular energy levels by simply counting the number of molecules on a surface.

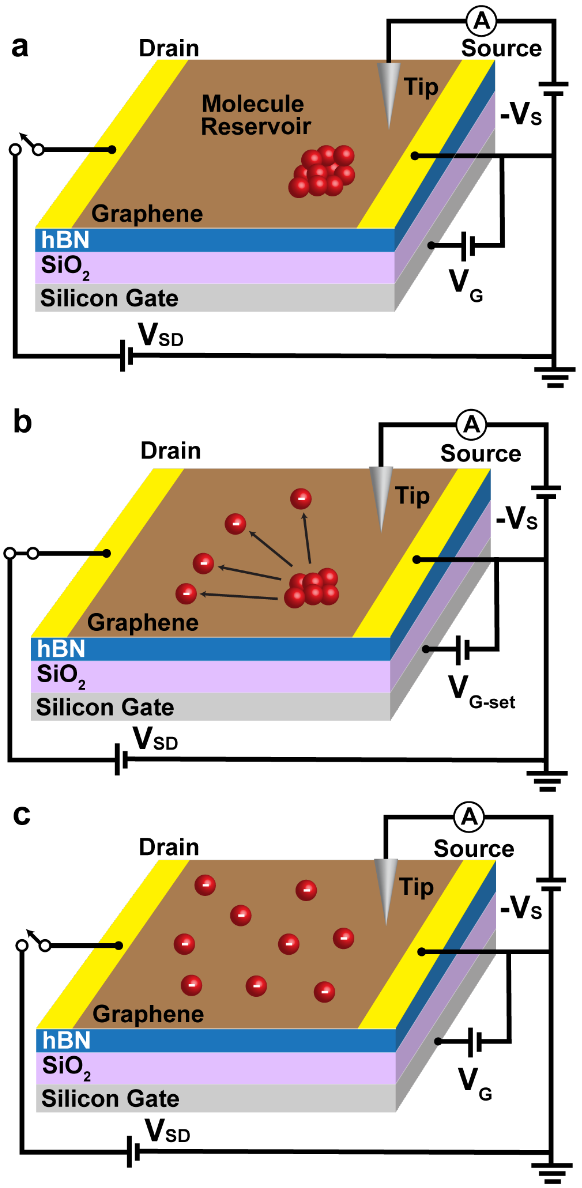

Figure 3.2: Controlling molecular concentration on a graphene FET. (a) Sketch representing the F4TCNQ molecular configuration (red balls) in the "as grown" state after thermal evaporation onto a graphene FET. (b) Application of a back-gate voltage to the FET while simultaneously flowing a source-drain current through the graphene causes the molecules to diffuse out onto the graphene surface. (c) When the source-drain current is turned off the device cools and the molecules freeze in place.

Our technique for reversibly changing molecular concentration on graphene starts with the deposition of F4TCNQ molecules onto a clean graphene/h-BN field-effect transistor held at room temperature in UHV. We then cool our devices down to 4.5K and insert them into our STM stage without breaking vacuum. At 4.5K the molecules remain stationary and can be stably imaged by our STM. In order to controllably set the molecular density to a desired concentration, a set-voltage ($V_{\text{G-set}}$) is applied to the device back-gate while a source-drain current of $I_{sd} \sim 1.5$mA is simultaneously passed through the graphene (i.e., a current density of $\sim$ 40 A/m), as sketched in Figure 3.2 Joule-heating induced by the source-drain current provides the molecules with enough thermal energy to diffuse on the surface and establish a new molecular concentration. As long as $V_{\text{G-set}}$ is above the threshold $-10\text{V} < V_{\text{G-set}}$ (which ensures that the molecules are negatively charged(*59–61*)) then a well-defined molecular surface concentration will be established after allowing enough time for equilibration (usually 100 seconds). The molecules freeze in place as soon as the source-drain current is turned off and the new molecular concentration can then be imaged by the STM.

The results of this protocol can be seen in the different prepared molecular concentrations shown in Figure 3.3a-f. At a high set-voltage of $V_{\text{G-set}} = 60$ V the resulting molecular density is correspondingly high ($n_M = 6 \cdot 10^{12}\text{cm}^{-2}$) and repulsive interactions between the charged molecules produce a roughly evenly-spaced molecular distribution (Figure 3.3a). As $V_{\text{G-set}}$ is reduced, the resulting molecular concentration correspondingly reduces (Figure 3.3b-f). A plot of molecule density vs. $V_{\text{G-set}}$ (Figure 3.3g) shows that the molecular concentration is almost perfectly linear over the gate-voltage range $-10\text{V} < V_{\text{G-set}} < 60\text{V}$, and remains nearly zero for $V_{\text{G-set}} < -10\text{V}$. The gate-tuneable molecular concentration is observed to be non-hysteretic, as seen in Figure 3.3g where the forward and backward sweep data lie almost perfectly on top of each other. The molecular concentration within our STM field of view is thus precisely and reversibly controlled by tuning the graphene FET electrical device parameters.

The molecular concentration across the entire FET surface, however, cannot be completely uniform since the molecules have to “go somewhere” when their concentration is reduced and “come back from somewhere” when their concentration is increased, and this “somewhere” cannot have the same density as the STM field of view seen in Figure 3.3. However, the simple process described above for setting the molecular surface concentration is surprisingly robust and can be observed over a majority of the surface area that we have scanned ($\sim$ 80%) on multiple devices. A fraction of the device surface is defective/contaminated ($\sim$ 20%) (and therefore unusable for this purpose) since the devices are fabricated under ambient conditions before being placed in our UHV chamber for measurement. We hypothesize that defective regions of the surface act as reservoirs that store F4TCNQ molecules in a compact, condensed phase. F4TCNQ molecules must flow in and out of these condensation regions (due to electrostatic potential inhomogeneities) to enable the tight concentration control that we observe on the majority of the surface. A definitive test of this “reservoir hypothesis” is difficult since mapping the entire surface of a graphene FET at the atomic scale is currently not possible (and defective regions are difficult to im-

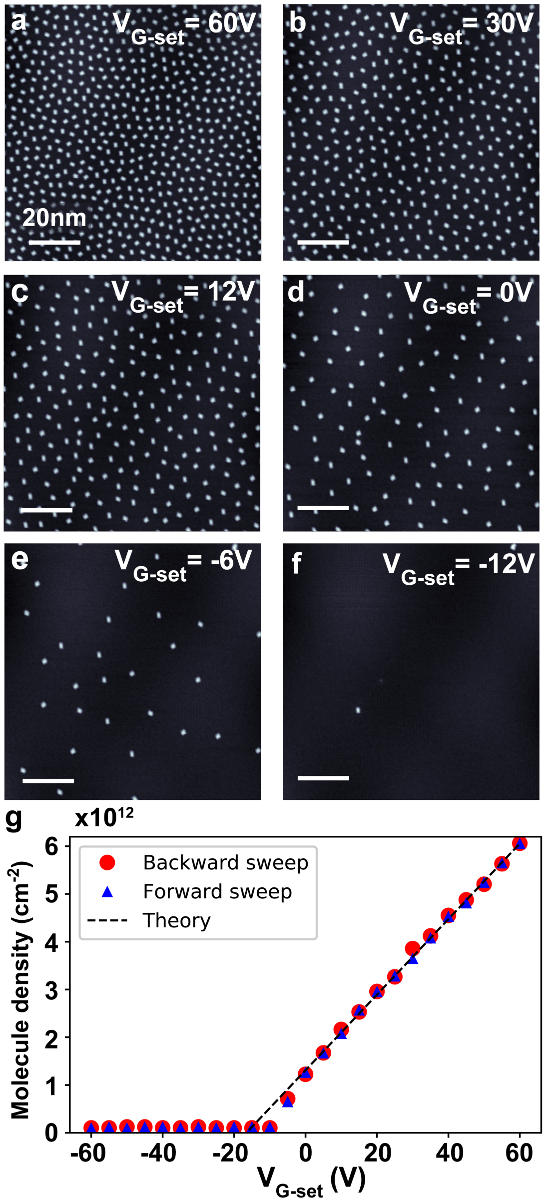

Figure 3.3: Tuning molecular concentration by using FET gate voltage. (a)-(f) STM topographs of the same area on the surface of a graphene field-effect transistor after tuning the molecular density with different values of gate voltage over the range $-12\text{V} < V_{\text{G-set}} < 60\text{V}$. (g) Measured molecular concentration as a function of gate set-voltage, $V_{\text{G-set}}$. No hysteresis is observed for forward and backward sweeps. The predicted molecular density from a simple capacitor-model is plotted as a dashed line.

age), but we do note that a condensed phase of F4TCNQ molecules has been observed on F4TCNQ-decorated graphene FETs in the past.(*62*)

In order to understand the microscopic mechanism that underlies the gate-tunability of molecular surface concentration, we must understand how charge readjusts itself in a graphene device decorated by mobile, charge-tuneable molecules as we tune the gate voltage. Here the molecule-decorated graphene forms one side of a capacitor while the back-gate electrode forms the other, and so the total charge that flows to the molecule/graphene system can be determined straightforwardly from the device capacitance and applied gate-voltage ($Q_T = -V_{\text{G-set}}C_T$). It is important to understand where this charge ends up in the device since it has two places to go. On the one hand it can occupy Dirac fermion band states while on the other it can occupy the LUMO states of the adsorbed molecules. In order to untangle where the charge goes, we performed STM spectroscopy directly on adsorbed F4TCNQ molecules as well as on bare graphene patches for different applied gate-voltages and molecular concentrations. This enabled us to track the energetic alignment of the F4TCNQ LUMO energy (E_L) and the graphene Dirac point energy (E_D) with respect to the graphene chemical potential (i.e., the Fermi level, E_F) for different gate voltages, thus providing snapshots of how charge in the device is distributed under different conditions.

We start with the negative gate-voltage regime, $V_G < -15$V, since this is the most straightforward regime to understand from an electronic viewpoint. The electronic configuration of the molecule/graphene system in this regime is qualitatively represented by the sketch in the inset to Figure 3.4c. Here the graphene is p-doped since E_F lies below E_D. The molecular LUMO level, E_L, lies below E_D but remains unfilled by electrons since it is still higher than E_F. Experimental evidence for this type of electronic behavior is seen in the top dI/dV spectra of Figure 3.4a and 3.4b (all dI/dV spectra shown here were acquired with molecules in the "frozen" state with the source-drain current set to zero). The top curve in Figure 3.4a shows a dI/dV spectrum measured with the STM tip held above the center of an F4TCNQ molecule for a gate-voltage of $V_G = -60$V and a fixed molecular density of $n_M = 0.8 \cdot 10^{12}\text{cm}^{-2}$. The F4TCNQ LUMO level is resolved as a peak at $V_S \approx 0.2$V above the Fermi energy (E_F is always at $V_S = 0$). The precise energy position of the LUMO is determined following the protocol of ref. (*59*) which takes into account known vibrational effects.(*59*, *63*) The top curve in Figure 3.4b shows a corresponding "off-molecule" dI/dV spectrum taken with the STM tip held over a bare patch of graphene 100 Angstroms away from nearby F4TCNQ molecules for the same gate-voltage and molecular concentration. A depression in dI/dV over the range $-60\text{mV} < V_s < 60\text{mV}$ can be seen here that is known to occur due to phonon inelastic tunneling effects,(*64*) and another depression is seen at $V_s \approx 0.35$V that is known to mark the location of the Dirac point (E_D).(*54*, *64*) E_D (from Figure 3.4b) is observed to lie above E_L (from Figure 3.4a) by $\propto 150$mV at this gate voltage and both quantities shift downward in energy together as the gate voltage is increased from $V_G = -60$V to -10V (Figure 3.4c). Such behavior is expected as E_F moves upward in the band structure shown in the inset to Figure 3.4c.

When V_G is raised above $V_G = -10$V, however, the device enters a different electronic regime. This can be seen even without STM spectroscopy since the gate-dependent molecular

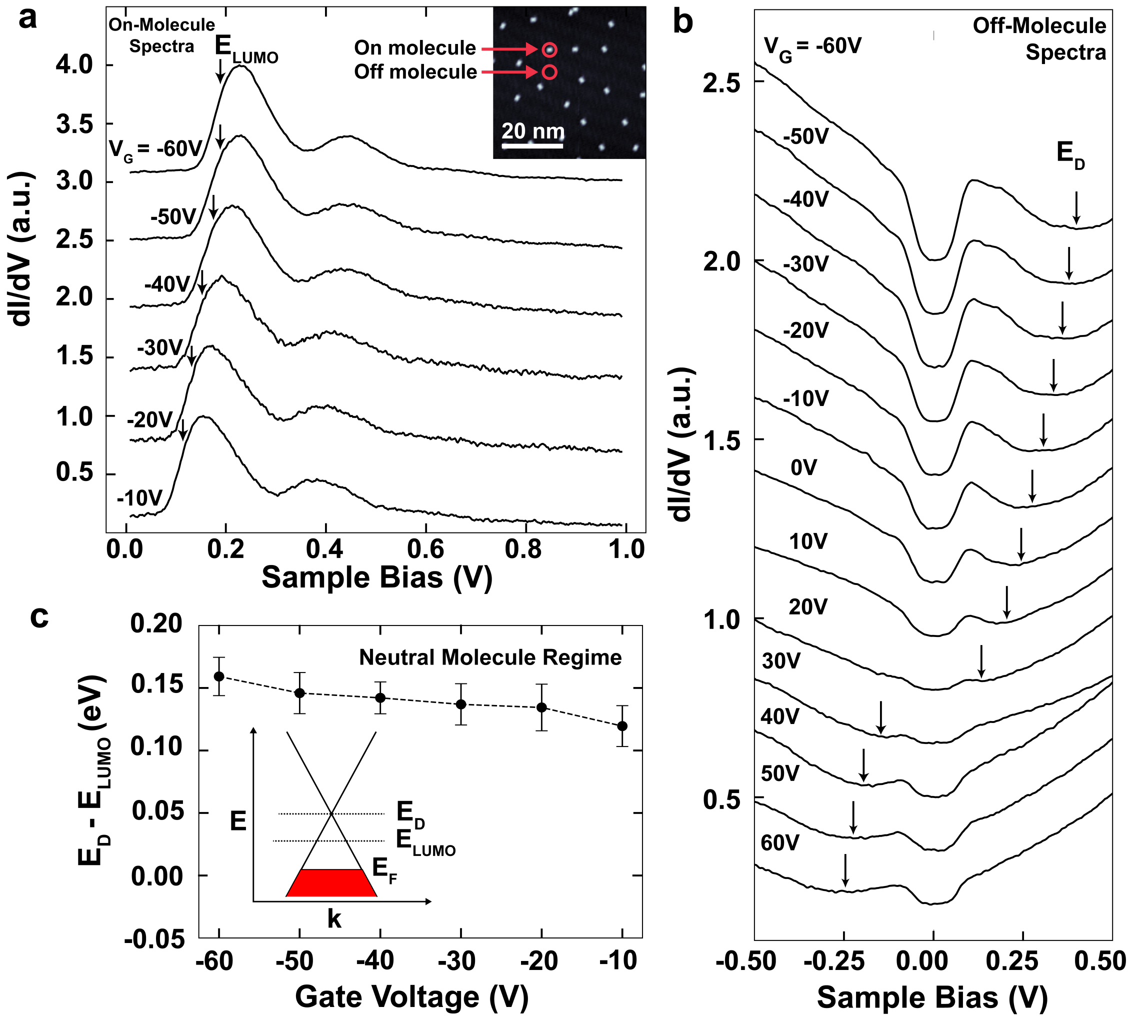

Figure 3.4: STM spectroscopy of F4TCNQ-coated graphene field-effect transistor (FET). (a) dI/dV spectra measured while holding the STM tip directly above an F4TCNQ molecule for a molecular concentration of $0.8 \cdot 10^{12}$ cm^{-2} at the surface of a graphene FET ($-60\text{V} < V_G < -10\text{V}$). Energy location of lowest unoccupied molecular orbital (E_L) is marked. Inset shows a representative image of the surface at this molecular concentration. (b) dI/dV spectra measured while holding the STM tip over bare patches of graphene ~ 100 Angstroms away from nearby F4TCNQ molecules for the same surface conditions measured in (a) ($-60\text{V} < V_G < 60\text{V}, n_M = 0.8 \cdot 10^{12}$ cm^{-2}). Graphene Dirac point energy (E_D) is marked. Inset shows a simplified representation of the electronic structure of F4TCNQ molecules on graphene for large negative gate voltages (E_F = Fermi energy). Spectroscopy parameters: $I_{\text{setpoint}} = 50$ pA, wiggle-voltage amplitude = 15 mV, frequency = 401Hz. (c) Gate-voltage dependent values of the molecular LUMO energy (E_L) shown for molecular density $n_M = 0.8 \cdot 10^{12}\text{cm}^{-2}$.

concentration data of Figure 3.3g shows the concentration rising dramatically (from near zero) for $V_G > -10\text{V}$. The rise in molecular concentration shown in Figure 3.3g occurs at the same gate voltage where the F4TCNQ LUMO begins to fall beneath E_F and change from being empty to being filled (Figure 3.4a). Unfortunately we cannot show any "on-molecule" spectroscopy data for $V_G > -10\text{V}$ because negatively charged molecules (i.e., those with a filled LUMO) tend to become unstable during scanning tunneling spectroscopy (STS) measurement. The off-molecule spectroscopy, however, is robust at all gate voltages, as shown in Figure 3.4b. Here the Dirac point is seen to continuously fall in energy as the gate voltage increases above $V_G = -10\text{V}$, causing the graphene to transition from being hole-doped ($E_D > E_F$) to being electron-doped ($E_D < E_F$) at $V_G \approx 20\text{V}$ for this molecular density ($n_M = 0.8 \cdot 10^{12}\text{cm}^{-2}$).

3.5 Discussion

Close inspection of the gate-dependence of E_D for different fixed molecular densities reveals evidence of Fermi-level pinning in our device induced by the distributed F4TCNQ LUMO levels (Figure 3.5). This phenomenon is useful to understand because it provides information on how device charge distributes itself on the surface at E_F. We measured Fermi-level pinning as a function of molecular concentration by first setting the concentration to the desired value (as shown in Figure 3.5a) and then measuring the gate-voltage dependence of E_D (Figure 3.5b) from off-molecule dI/dV spectra using the procedure described for Figure 3.4b. In the absence of any molecules on the graphene surface E_D moves smoothly down in energy with increasing V_G according to the expression(*64*):

$$E_D(V_G) = -\text{sgn}(V_G)\hbar v_F \sqrt{\pi C|V_G - V_0|} \tag{3.6}$$

where C is the unit area capacitance of the device and V_0 reflects background doping. By fitting Eq. 3.6 to the gate-dependent Dirac point energy of our device before depositing molecules we are able to extract the capacitance between the graphene and the back-gate electrode: $C = 7.8 \cdot 10^{10}\text{cm}^{-2}V^{-1}$.

For nonzero molecular concentrations, however, the E_D vs. V_G curve deviates from Eq. 3.6 and forms a "pinning" plateau at $\approx 140\text{meV}$ above E_F, with the width of the plateau increasing with increased molecular surface density (Figure 3.5b). It is useful to note that the start of the plateau (for increasing V_G) coincides with the gate voltage value where the molecular LUMO begins descending below E_F (Figure 3.4a), allowing the plateau to be associated with the charging of F4TCNQ molecules. The value of E_D at the plateau ($\approx 140\text{meV}$ above E_F) is nearly equal to the energy difference observed between E_D and E_L in Figure 3.4c, strongly suggesting that the plateau in E_D arises due to pinning of E_F by the molecular LUMO level. This interpretation is quantitatively supported by the increased width of the pinned region as molecular concentration is increased (i.e., since higher molecular concentrations can store higher concentrations of charge). For example, the E_D plateau at a fixed molecular concentration of $4.3 \cdot 10^{12}\text{cm}^{-2}$ has a width of $\Delta V_G = 50 \pm 5\text{V}$ which

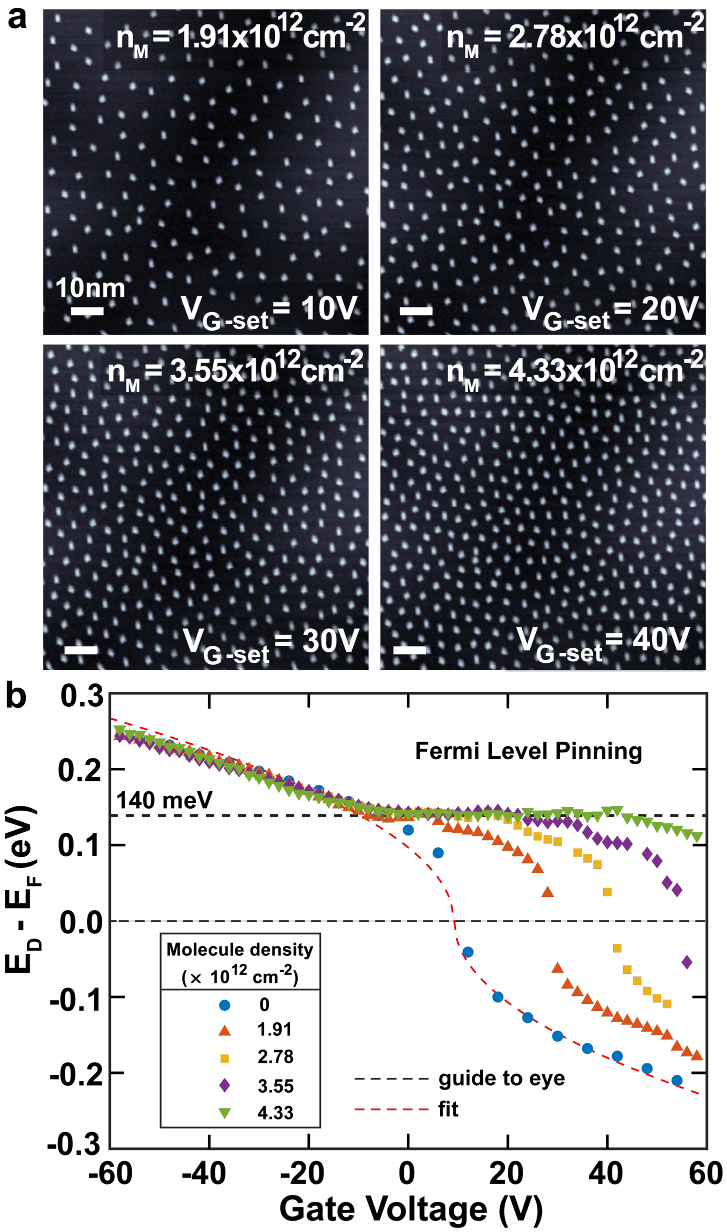

Figure 3.5: Fermi level pinning of F4TCNQ-doped graphene field-effect transistor (FET). (a) STM images of graphene FET surface decorated with different molecular densities for measurement of molecule-induced Fermi-level pinning shown in (b). (b) Gate-voltage dependence of Dirac point energy (E_D) measured via STS on graphene FET surface for the different molecular densities shown in (a). A density-dependent plateau in E_D indicates Fermi level pinning. Red dashed line shows fit of Eq. 3.6 to data for zero molecular density ($n_M = 0$).

corresponds to an increase in surface charge density by an amount $\Delta\sigma = (3.9 \pm 0.4) \cdot 10^{12} e^{-}/\text{cm}^2$ which nicely matches the molecular concentration (charge is calculated using the capacitance value acquired via Eq. 3.6).

The Fermi level pinning described here for the static molecular configurations of Figure 3.5a is intimately related to the dynamic molecular reconfigurations that are at the heart of gate-tuneable molecular concentration, the main phenomenon that we are trying to understand. When the graphene Fermi level is securely pinned by molecular LUMO states, new electrons added to the device (e.g., by an increase in V_G) do not cause the Fermi level to rise in energy since the LUMO levels absorb charge added to the graphene at E_F. On the other hand, if there are not enough molecules on the surface to pin the Fermi level, then increasing the gate voltage will cause electrons to occupy graphene band states at higher energies above E_L. This is the origin (from an energetic perspective) of the force that drives the molecules to move on the surface and to dynamically change the molecular concentration when V_G is modified under "diffusive conditions" (i.e., when lateral source-drain current flows through the device). Under these conditions as gate-voltage is increased the molecular concentration must also increase to maintain Fermi level pinning (the overall lowest energy state) by enabling charge to flow into lower-energy LUMO levels rather than higher-energy Dirac band states.

These concepts allow us to quantitatively predict the concentration of molecules expected on the graphene surface for a given gate voltage. We start with the condition that the lowest-energy electronic configuration under diffusive conditions (and when $-10\text{V} < V_G$) occurs when the Fermi energy is pinned at E_L. For a given gate voltage ($V_{\text{G-set}}$) the total charge density on the molecule-decorated FET surface ($\sigma_T = -CV_{\text{G-set}}$) will then divide itself between charge carried by the molecules (σ_M) and charge carried by the graphene Dirac band (σ_G):

$$-CV_{\text{G-set}} = \sigma_M + \sigma_G \tag{3.7}$$

If each charged molecule contains one electron in its LUMO state then the total molecular charge is $\sigma_M = -n_M$ where n_M is the surface density of molecules and σ_M has units of $|e|$. Eq. 3.7 then leads to the following expression for n_M:

$$n_M = CV_{\text{G-set}} + \sigma_G \tag{3.8}$$

Because E_F is pinned at E_L by the molecular coverage, σ_G, can be found by integrating the charge density in the graphene band from the Dirac point to E_L, resulting in the following well-known expression:(*65*)

$$\sigma_G = \frac{|E_D - E_L|^2}{\pi \hbar^2 \nu_F^2} \tag{3.9}$$

Combining Eqs. 3.7 and 3.9 leads to the final expression for molecular density as a function of $V_{\text{G-set}}$:

$$n_M = CV_{\text{G-set}} + \frac{|E_D - E_L|^2}{\pi \hbar^2 \nu_F^2} \tag{3.10}$$

Using our known device parameters of $C = 7.8 \cdot 10^{10}\text{cm}^{-2}\text{V}^{-1}$ and $V_F = 1.1 \cdot 10^6\text{m/s}$, we are able to fit Eq. 3.10 to the n_M vs. $V_{\text{G-set}}$ data shown in Figure 3.3g by using $|E_D - E_L|$ as a fitting parameter. As shown by the dashed line in Figure 3.3g, Eq. 3.10 fits the data well for a value of $|E_D - E_L| = 142 \pm 23\text{meV}$.

We can check the validity of this theoretical framework by comparing the value of $|E_D - E_L|$ obtained from our molecular concentration measurements described above with the value obtained using the STS measurements shown in Figures 3.4 and 3.5. STS enables us to obtain $E_D - E_L$ in two ways: first by extracting E_L and E_D directly from our dI/dV spectra (Figure 3.4) and subtracting them, and second by determining the E_D energy plateau caused by Fermi level pinning (Figure 3.5). From Figure 3.4 we see that for $V_G = -30\text{V}$ our on-molecule dI/dV spectrum yields $E_L = 165\text{meV}$ while our off-molecule spectrum yields $E_D = 305\text{meV}$. This results in a value of $E_D - E_L = 140 \pm 20\text{meV}$ which agrees well with the value of $|E_D - E_L| = 142 \pm 23\text{meV}$ obtained from the concentration-based analysis above. The average $|E_D - E_L|$ value determined from STS over the entire gate-voltage range of $-60\text{V} < V_G < -10\text{V}$ for this concentration is $143 \pm 9\text{meV}$, in agreement with the value obtained from the concentration analysis. The Fermi-level pinning data of Figure 3.5 also reveals a Dirac point plateau at $E_D - E_F = 140 \pm 5\text{meV}$. Since E_F is pinned at E_L under these conditions, this reflects a value of $E_D - E_L = 140 \pm 5\text{meV}$, again in agreement with the value obtained above from the concentration analysis of Eq. 3.10.

One consequence of this analysis is confirmation that each F4TCNQ molecule involved in graphene Fermi-level pinning carries a single electron of charge. Even at $V_G = 0$ electrons transfer from the Dirac bands above E_L into molecular frontier orbitals at E_L in order to minimize the total electronic energy of the device (i.e., leading to $n_M = |E_D - E_L|^2/(\pi\hbar^2\nu_F^2)$ at $V_{\text{G-set}} = 0$ from Eq. 3.10). Another consequence is that measuring molecular concentration on a graphene FET by this technique provides a new method for quantitatively determining the energy alignment of molecular frontier orbitals with respect to the graphene Dirac point (i.e., $E_D - E_L$). This method is potentially valuable for determining the energy alignment of highly mobile adsorbates since it is difficult to prevent such adsorbates from moving inopportunely while under an STM tip during the bias sweeps required for STS. In our previous work, for example, we found it necessary to anchor F4TCNQ molecules to secondary immobile molecular structures in order to probe their LUMO levels in the charged state via STS.(*59*) Such co-adsorption, however, can alter local dielectric environments and influence molecular orbital energies.(*66*, *67*) Our new technique of measuring gate-dependent molecular concentration allows one to bypass this problem and to access molecular energy-level information through a completely different channel.

In conclusion, we have demonstrated that molecular concentration at the surface of a graphene FET can be reversibly manipulated via a back-gate voltage applied simultaneously with the "thermal lubricant" of source-drain current. The equilibrium molecular density is precisely determined by the capacitance between the back-gate electrode and the graphene, in combination with the energy difference between the Dirac point and the molecular LUMO level. The driving force behind this dynamic mechanical reconfiguration of molecular concentration is the energetic favorability of molecular Fermi-level pinning over the charge-filling of graphene Dirac bands. The energy alignment of adsorbed molecular LUMO levels obtained from a concentration-based analysis using these concepts compares well with values determined from direct spectroscopic probes such as STS. Using this procedure, we can thus prepare systems of tuneable density, which should give rise to qualitatively different diffusive behaviors.

Chapter 4

Gate Tunable Diffusivity

In this chapter we demonstrate gate tunability of a molecular diffusion barrier. We first discuss the procedures and analysis for characterizing diffusion in a low temperature STM experiment which involves creating videos from a series of STM images. Videos are acquired by repeatedly heating the sample to drive diffusion and then quenching the sample to liquid helium temperature in order to take an STM "snapshot" of the surface. Videos are processed using classical computer vision techniques in order to detect and link particles across video frames. This enables analysis of trajectories by measuring MSDs to extract diffusion constants/barriers. Our work highlights a new avenue for enhancing diffusion on a device outside of more common temperature control by using a backgate electric field to directly change the diffusion barrier.

4.1 Related Work

In-situ microscopic characterization of adsorbates on metal surfaces has been performed in ultra-high vacuum settings(*68*, *69*) through STM(*12–16*, *18*) and in solution(*7*) via individual adatom tracking. Previously, diffusion barrier tunability through an externally applied electric field has been demonstrated via tunability of adsorbate-substrate interaction due to changes in the adsorbate dipole moment (*7*), as well as through tunability of adsorbate-adsorbate interaction.(*18*) However, as of yet, control of diffusion in the context of a device has yet to be achieved.

4.2 Methods

Low Temperature Video STM

Low temperature STM is performed at liquid helium temperature ($\sim$ 4K), which is much lower than the required temperature to observe diffusive motion for most adsorbates. Diffusion for our molecule of interest, F4TCNQ, is observable around 20K. However, at this

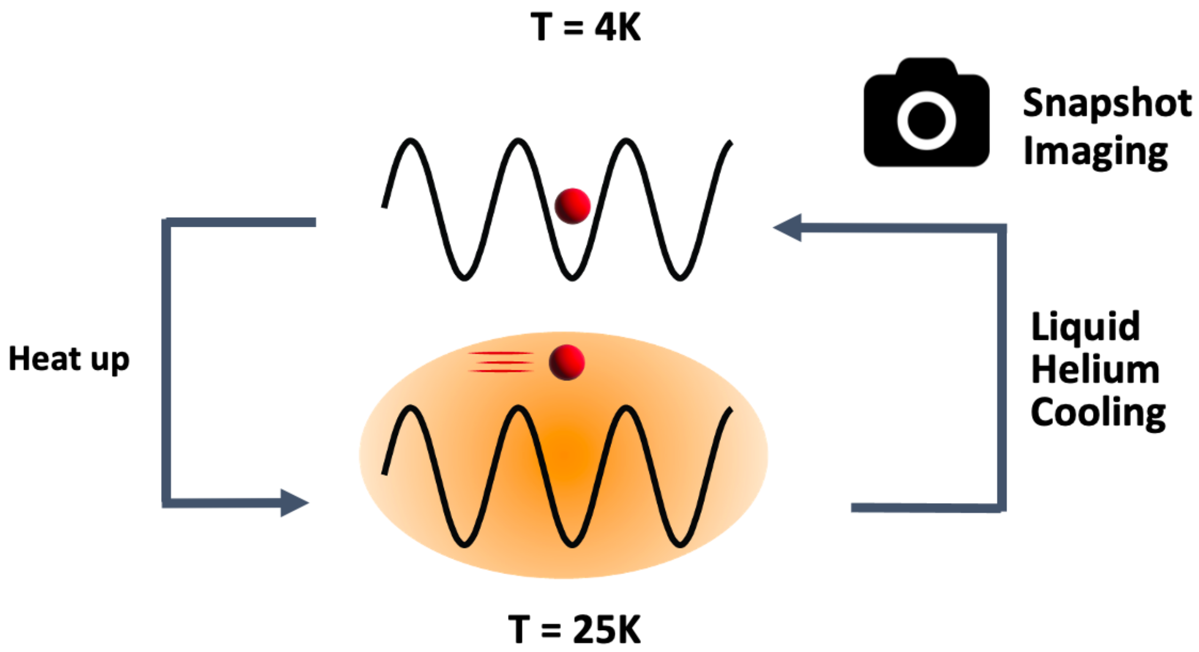

Figure 4.1: Control loop for video STM of diffusion

temperature, the low frame rate of conventional low temperature STM makes it impossible to directly video capture diffusion in STM while the system is held at 20K. In order to study the time evolution of diffusion, we adopted a "clay-mation" method of video imaging in which the system is driven at high temperatures (≈ 20K) for a short amount of time (≈ 10s) and then cooled to liquid helium temperature for imaging. This process, portrayed in Figure 4.1, is repeated many times in order acquire successive video frames.

Computer Vision Techniques for Particle Tracking

Particle tracking can be roughly divided into two tasks: detection and linking. Detection is the task of finding particle centers from raw images and linking is making associations between particle detections of successive frames. The techniques described henceforth are described in deeper detail in work by Crocker and Grier and are implemented in TrackPy.(*11*, *70*)

Detection can be done using a technique known as template matching. In template matching, for some image, I_{ij}, we want to find the locations of a smaller image, which we call the template, T_{ij}, within I_{ij}. Given both the image and template, the location of a match can be found by using a sliding window between the image and the template and optimizing an appropriate similarity measure, either minimization of the sum of absolute differences or maximization of normalized cross correlation. Cross correlation methods are

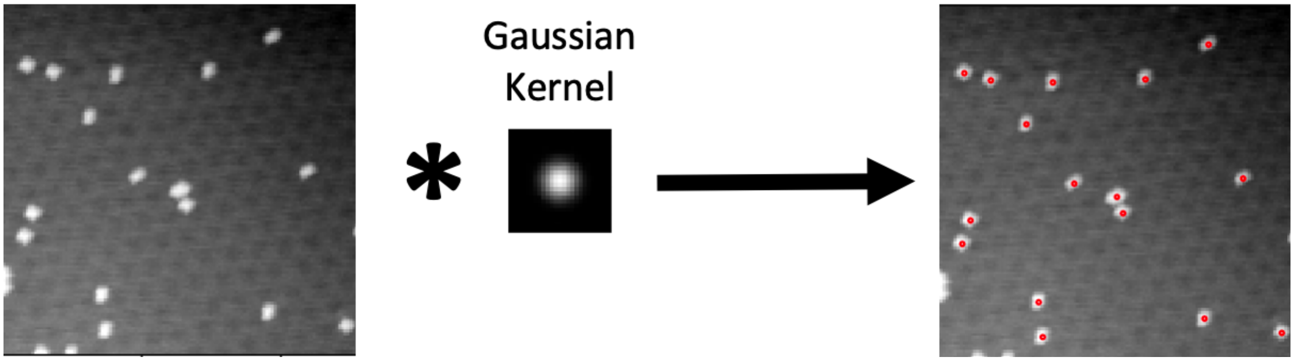

Figure 4.2: Particle detection using Gaussian templates and cross correlation. A Gaussian template is convolved with the original source image in order to detect particle centroids. Gaussian convolution is the same as applying a low pass filter. After filtering, local maxima are thresholded before being accepted as a particle detection.

often favored for their speed due to the possibility of performing the calculation using fast Fourier transform/convolution methods over spatial domain methods, even though spatial methods have techniques such as image pyramids to accelerate the worst case run time of image registration.(*71*, *72*) For the template, a discrete Gaussian kernel of similar width as our particle is chosen as the template. Despite F4TCNQ having a non-zero eccentricity, we find that the choice of an isotropic Gaussian template doesn't adversely affect the quality of detection as shown in Figure 4.2

For linking, the associations between particle detections of successive frames need to be made. If we assume a Gaussian prior on the motion of particles, which is the same as treating the particles as obeying Brownian dynamics, the maximum likelihood particle association roughly corresponds to linking detections that are spatially close to each other. More rigorously, we say the probability density of some displacement $\mathbf{x}_i - \mathbf{x}_j$ occurring is:

$$f(\mathbf{x}_i, \mathbf{x}_j) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(\frac{1}{2}\frac{|\mathbf{x}_i - \mathbf{x}_j|^2}{\sigma^2}\right) \tag{4.1}$$

The maximum likelihood association is the optimal solution to following optimization problem.

$$\max_{\{\mathbf{x}_i\},\{\mathbf{x}'_j\}} = \prod_{ij} f(\mathbf{x}'_i, \mathbf{x}_j) \tag{4.2}$$

where $\{\mathbf{x}_j\}$ is the list of detected particle positions in one frame, and $\{\mathbf{x}_i\}$ is the list of proposed particle associations from detected particles in the following frame. The problem reduces to that of finding the optimal permutation of candidate labels, which naively has combinatorial runtime complexity. In practice, not all particles are likely candidates, so it is common practice to filter candidates by some maximum allowed distance from some starting

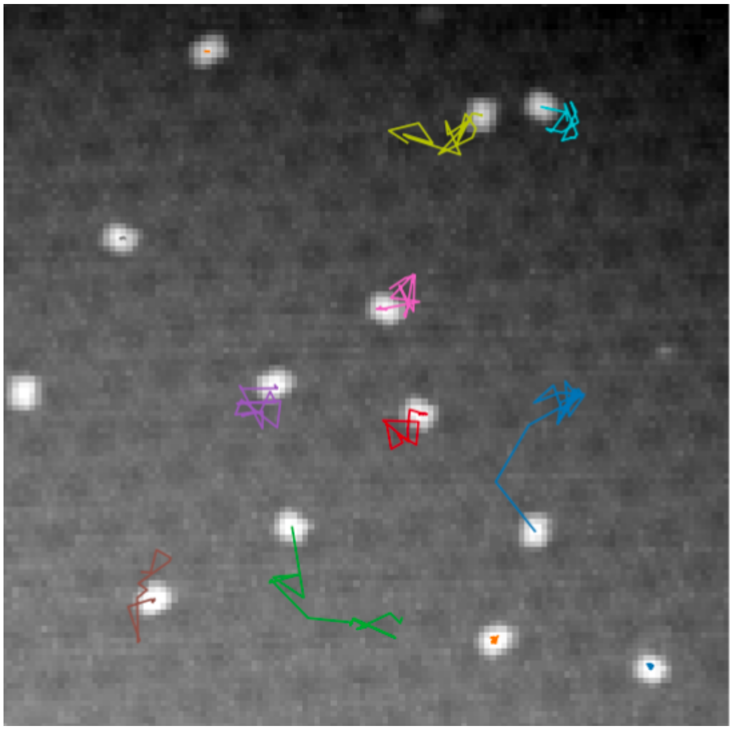

Figure 4.3: Particle Tracking. Example sample tracks generated using Gaussian template matching and Brownian motion priors.

position before considering a successive candidate. While this optimization framework is theoretically based on a Gaussian prior, we find that this technique works well for other types of unbiased random walks that we have observed experimentally, including the CTRW. An example result from tracking is shown in Figure 4.3.

We found that the orientation of the molecule was highly correlated with the diffusive behavior of the system. F4TCNQ prefers to adsorb onto graphene such that the molecule will sit on top of and align to the graphene carbon-carbon bridge site as depicted in Figure 4.4. As a result, there are three allowed F4TCNQ adsorption orientations corresponding to the 3-fold triangular rotational symmetry, each orientation being 120 degrees apart. To determine the orientation of the particle in a given frame, we use a technique known as image moments.(*73*) Image moments fit a multivariate Gaussian kernel to a detected particle instance. It does this by calculating the covariance matrix, Σ, of the particle image intensity, $f(x, y)$, that is:

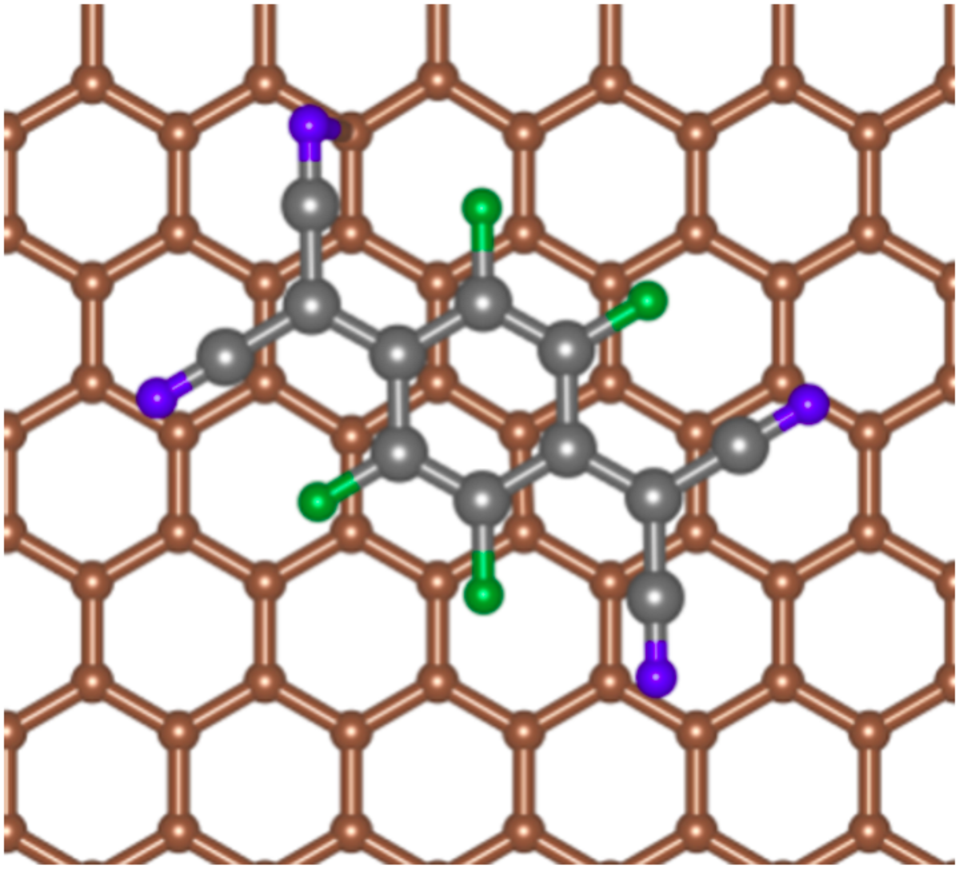

Figure 4.4: F4TCNQ adsorbed on graphene. The preferred F4TCNQ adsorption geometry is such that the center of the molecule is aligned on top of the graphene C-C bond.

$$\mu_{pq} = \sum_x \sum_y (x - \overline{x})^p (y - \overline{y})^q f(x, y) \tag{4.3}$$

$$\mu'_{20} = \mu_{20}/\mu_{00} \tag{4.4}$$

$$\mu'_{02} = \mu_{02}/\mu_{00} \tag{4.5}$$

$$\mu'_{11} = \mu_{11}/\mu_{00} \tag{4.6}$$

$$\Sigma = \begin{bmatrix} \mu'_{20} & \mu'_{11} \\ \mu'_{11} & \mu'_{02} \end{bmatrix} \tag{4.7}$$

The eigenvectors of the covariance matrix correspond to the major and minor axis of the particle. The angle of the major axis, Θ, is then given by:

$$\Theta = \frac{1}{2} \arctan\left(\frac{2\mu'_{11}}{\mu'_{20} - \mu'_{02}}\right) \tag{4.8}$$

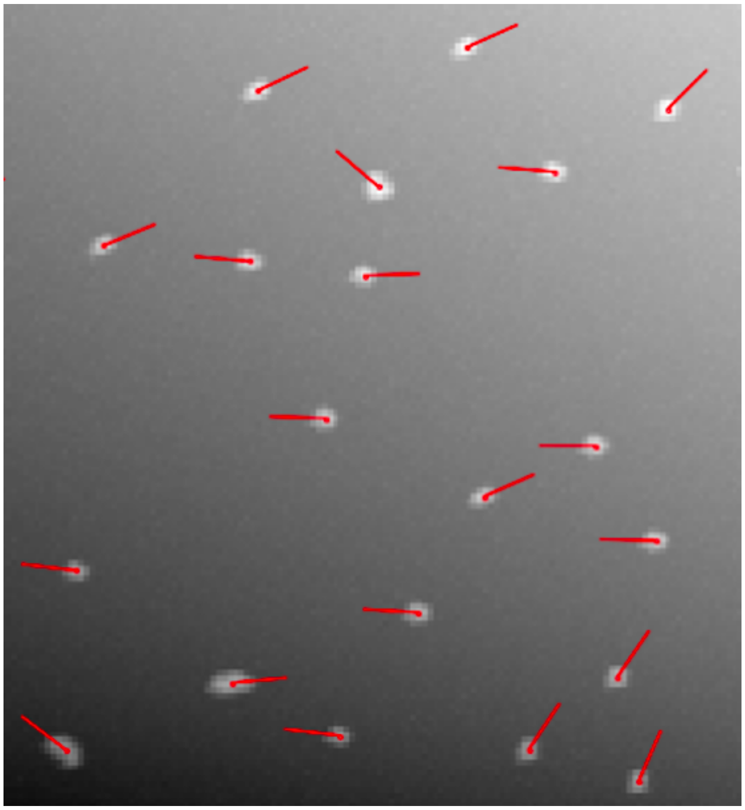

Figure 4.5: Orientation Estimates of F4TCNQ. Estimation of the F4TCNQ major axis via image moments.(*73*)

We demonstrate this capability in Figure 4.5. Although there is some estimation error in the exact orientation of the molecule due to the finite resolution of the STM, we are only interested in which of the 3 orientations any particular molecule belongs to, so we can bin molecule orientations using any 3 triangularly symmetric bins.

4.3 Results

To study the effect of the gate voltage on the diffusion barrier, we collected a series of STM videos of diffusion at varying gate voltages and temperatures. For a given video taken at a fixed temperature and gate voltage, the diffusion constant is extracted by calculating the mean squared displacement through ensemble and time averaging from the measured trajectories obtained via the detection and linking algorithms described in Methods. For a single gate voltage, we repeat this process for different temperatures in order to obtain a functional dependence between the diffusion constant and the temperature. A representative example of one such experiment is shown in Figure 4.6.

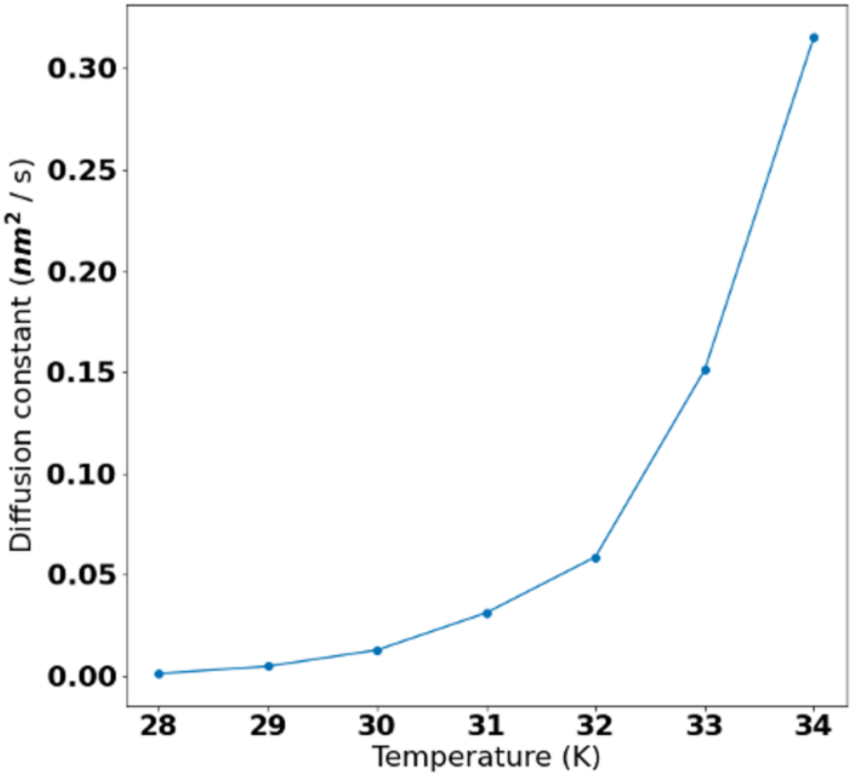

Figure 4.6: Experimental diffusion constants vs. temperature.

Using the physical model of surface diffusion proposed in equations 2.11 and 2.10, we can recover the diffusion barrier, E_{diff}, of the system through an Arrhenius plot. This yields the diffusion barrier for a single gate voltage. We repeated this for differing gate voltages, the result of which is shown in Figure 4.7.

Analysis of the gate dependence of the diffusion barrier reveals a regime change beginning at $V_G = -20$ V, which is near the charge transition point of the molecule as shown in (*59*) and Chapter 3. In the uncharged regime, $V_G < -20$, the diffusion has a sharp gate dependence whereas in the charged regime, there is little charge dependence and the molecule is more mobile in comparison.

Experimentally it was found that the charged and uncharged states of the molecule qualitatively diffuse differently. This is shown in Figure 4.8. Scatter plots of displacements for both charged and uncharged molecules are shown and differentiated by whether the molecule was observed to have rotated or not during diffusion. The scatter plots demonstrate that for the charged molecule the preferred diffusion pathway is purely translational whereas for the uncharged molecule the diffusion pathway also includes rotational modes of motion. This experimental finding is corroborated by constrained density functional theory (DFT) calculations performed by Zachary Goodwin and Johannes Lischner shown in Figures 4.9,

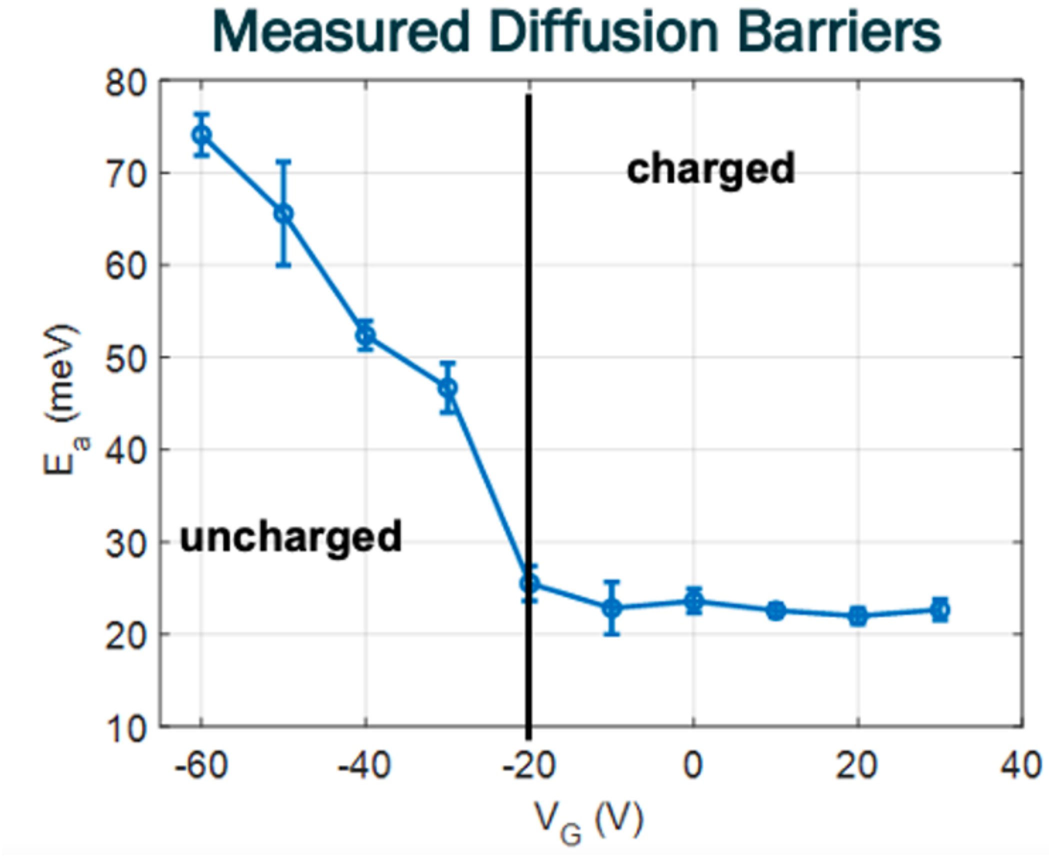

Figure 4.7: Experimental diffusion barrier vs. gate voltage. Gate dependence of F4TCNQ diffusion barrier. Diffusion barrier experiments reveal a regime change corresponding to charged and uncharged molecular states which lead to very different gate responses. In the uncharged case, molecule diffusion has a sharp gate dependence whereas the charged case shows very little gate dependence.

4.10, 4.11. Diffusion pathways are shown in figures 4.9 and 4.10, with the highest energy of a given pathway being the diffusion barrier. Qualitatively, the energies for the high energy reaction coordinates correspond to configurations where some of the atoms of the F4TCNQ lie directly over atoms in the graphene lattice. The energies of these pathways were calculated for different levels of graphene doping and are shown in 4.11. Estimated energy barriers reveal that between the charged and uncharged configurations different diffusion modes are preferred. For the charged molecule, the purely translational diffusion mode is the lowest energy pathway with very little gate dependence, whereas for the neutral molecule the lowest energy diffusion pathway is a 60 degree rotation plus translation which yields a stronger gate dependence. These results demonstrate that gate tunable diffusivity is possible by accessing the different charge states of a molecule.

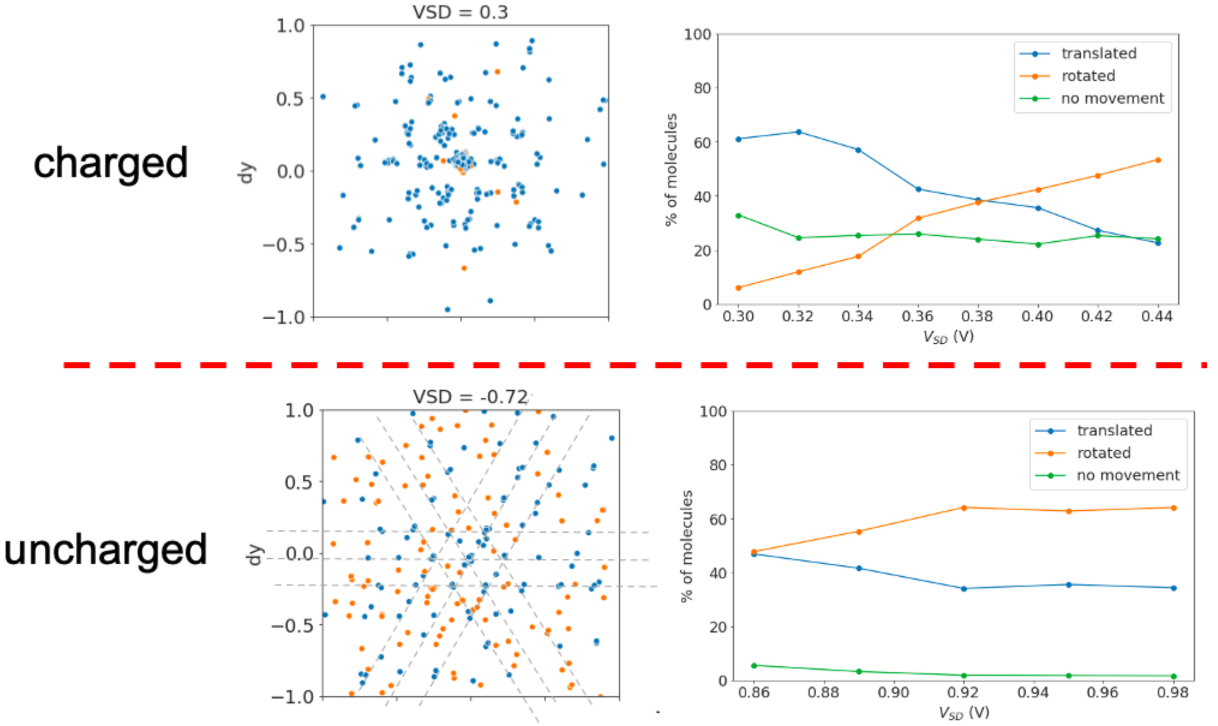

Figure 4.8: Displacement scatterplots for charged and uncharged molecules. Scatter plots show that charged and uncharged molecules use qualitatively different diffusion pathways. Charged molecules are observed to prefer purely translational modes of diffusion (shown in blue) whereas uncharged molecules have access to a rotational pathway (shown in orange).

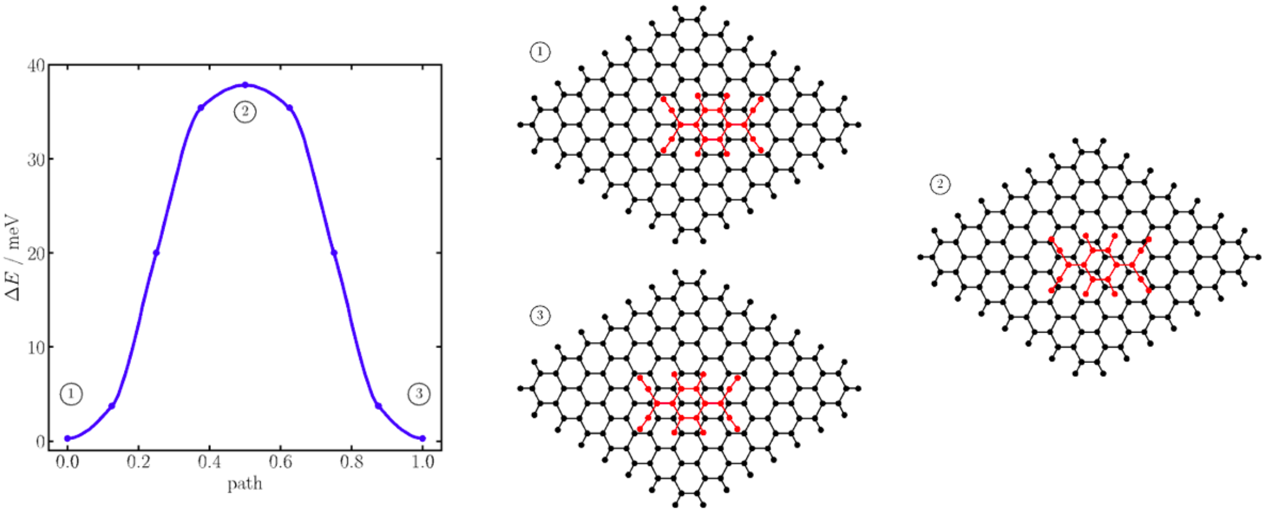

Figure 4.9: Translational diffusion pathway. Calculations done by Z. Goodwin and J. Lischner

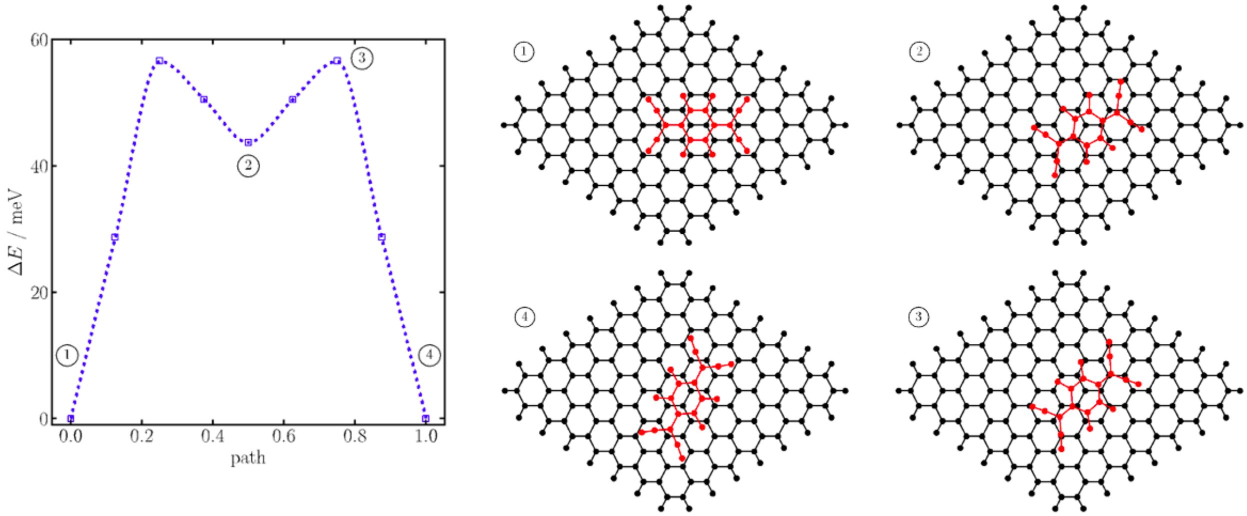

Figure 4.10: Rotational diffusion pathway. Calculations done by Z. Goodwin and J. Lischner.

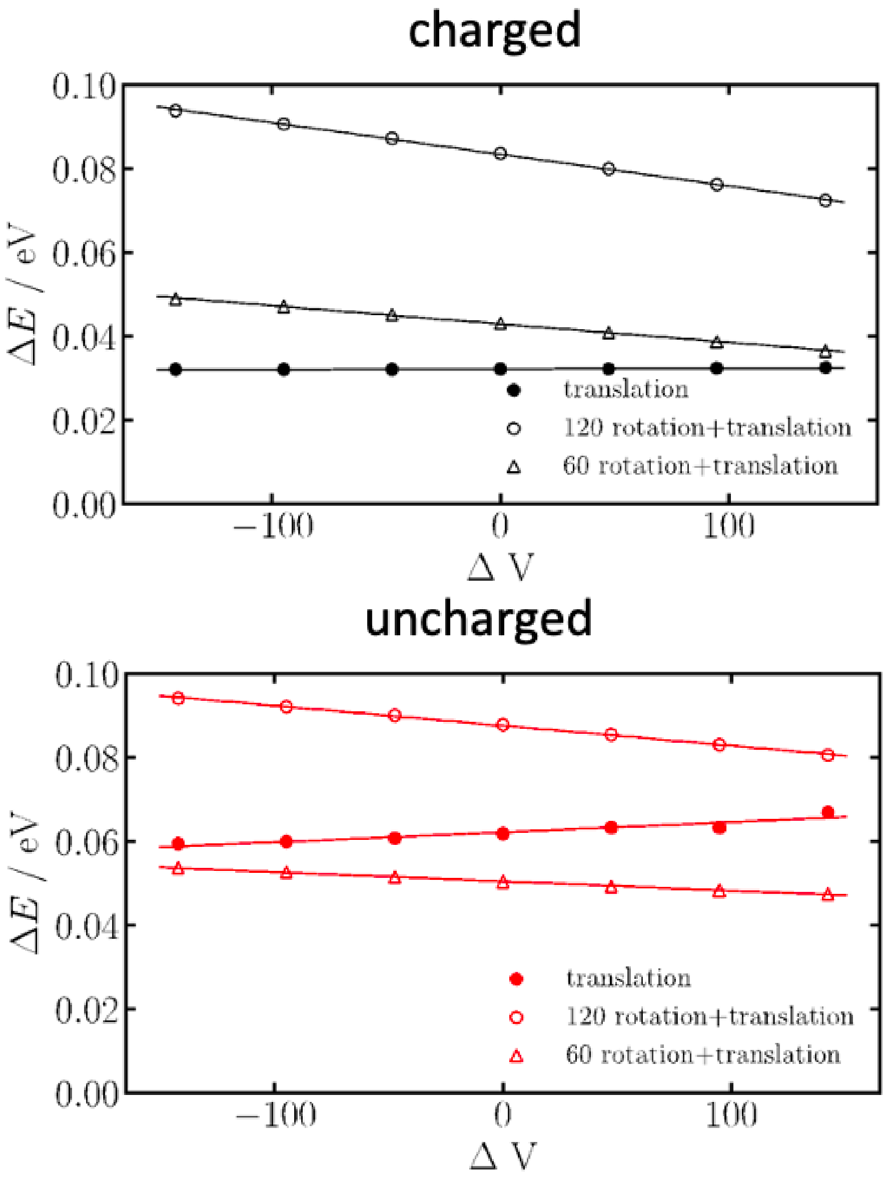

Figure 4.11: *Ab initio* diffusion barrier estimates. Scatter plots show that charged and uncharged molecules use qualitatively different diffusion pathways. Charged molecules are observed to prefer purely translational modes of diffusion whereas the uncharged molecule has access to a rotational motion pathway. Moreover, the translational mode of motion in the charged case possesses little gate dependence in contrast to the uncharged case, where the lowest energy diffusion mode (60 degree rotation and translation) has a significant gate dependence. Calculations done by Z. Goodwin and J. Lischner.

Chapter 5

Tunable Ergodicity of Molecular Adsorbates on Moiré Surfaces with Substrate Energy Landscape Engineering

Ergodicity, the property that any single realization of a physical system given enough time will relax to its ensemble average configuration, is a key assumption in equilibrium statistical mechanics. The amount of time required to reach the ergodic configuration is different for various systems. Here we consider molecular surface diffusion on a moiré superlattice where we find using Monte Carlo simulation that the relaxation time of diffusion is comparable to the observation time scale and that ergodicity is not necessarily held in this time scale. We tested ergodicity experimentally, choosing a model system of organic single molecule adsorbates diffusing over a graphene/h-BN moiré superlattice. We find that the structure of the moiré substrate induces a weak ergodicity breaking in the surface diffusion of adsorbates, particularly at low temperatures and large superlattice wavelengths, allowing us to tune the ergodicity by engineering the moiré surface. Using video scanning tunneling microscopy (STM) experiments on a gated graphene device, we demonstrate that increasing the temperature or decreasing the superlattice wavelength changes the stochastic motion of adsorbates from a moiré superlattice-mediated anomalous subdiffusive continuous time random walk towards the Brownian limit.

5.1 Related Work

Traversing an energy landscape composed of many microstates over a long period of time to relax to an equilibrium condition is key to the stability of disordered systems, which include phenomena such as spin/thermal glass relaxation(*74–76*), motion of laser cooled atoms(*77*), soft particle jamming transitions(*78*, *79*), and simulated annealing(*80*). The ergodic hypothesis posits that at equilibrium, accessing all microstates on an energy landscape are equiprobable. However, for a heterogeneous system, where there are a large number of unique microstates on the landscape, in a short period of time only a fraction of available microstates are sampled.(*24*, *81*) This behavior can be modeled by the subdiffusive continuous time random walk (CTRW), a random walk process which permits the time interval between successive jumps to follow a heavy tailed distribution meant to capture the behavior of state variables trapped in metastable local minima state of varying depth.(*82–84*) Long trapping times cause state sampling to slow over time, a phenomenon known as aging, which results in time averaged quantities being random rather than converging to a single finite value.(*23*, *24*) For such a system, even at equilibrium, the timescale required to access the range of microstates can be extremely large and weak ergodicity breaking may emerge on the observation timescales.(*30*) This stands in stark contrast with the classical model of diffusion, Brownian motion, where the random walk process takes uncorrelated steps in space and time. Whereas the CTRW is a non-ergodic process due to the heavy-tailed time intervals between the jumps, Brownian motion is a random walk with a finite expected time between jumps, resulting in ergodic behavior. While previous work has demonstrated access to both ergodic and non-ergodic regimes of physical systems through access of different timescales(*85*), as of yet, system control in order to directly control the timescales of these regimes has not been reported. Here we demonstrate an example of artificially-induced ergodicity breaking for finite observation timescales with particles sampling an engineered energy landscape through a thermal diffusion process at the molecular length scale.

5.2 Results

A heterogeneous energy landscape for particle diffusion is created by a moiré superlattice of two-dimensional heterostructures. Moiré surfaces form by vertically stacking two slightly different periodic sheets. In two-dimensional heterostructures, moiré superlattices can form by overlaying two sheets of two-dimensional materials with different periodicity or orientation. Thermalized molecular surface adsorbates can diffuse over this engineered superlattice energy surface (Figure 5.1a). The diffusion behavior of adsorbates over the two-dimensional moiré superlattice can be tuned by controlling the temperature, T, and the moiré periodicity, p, of the superlattice. Here, the superlattice potential is constructed by vertically stacking two layers of two-dimensional materials such as graphene and hexagonal boron nitride (h-BN). The superlattice energy potential surface of a graphene/h-BN stack is calculated by *ab initio* methods on AA and AB sites, see Methods for the details. We then extended the potential between AA and AB sites by a quadratic function following the previous literature. Hence, the peak potential value was kept constant while only tuning the moiré periodicity in the simulation. Meanwhile, the graphene lattice creates metastable adsorption sites on the surface at a much smaller spatial wavelength (e.g. 2.46 Angstroms). The moiré superlattice energy induces energy heterogeneity amongst the graphene lattice sites because of the spatial dependence of the superlattice energy.

We simulated the diffusion of molecular surface adsorbates on the two-dimensional moiré superlattice using a Markov chain model in a Monte Carlo simulation – see Methods for

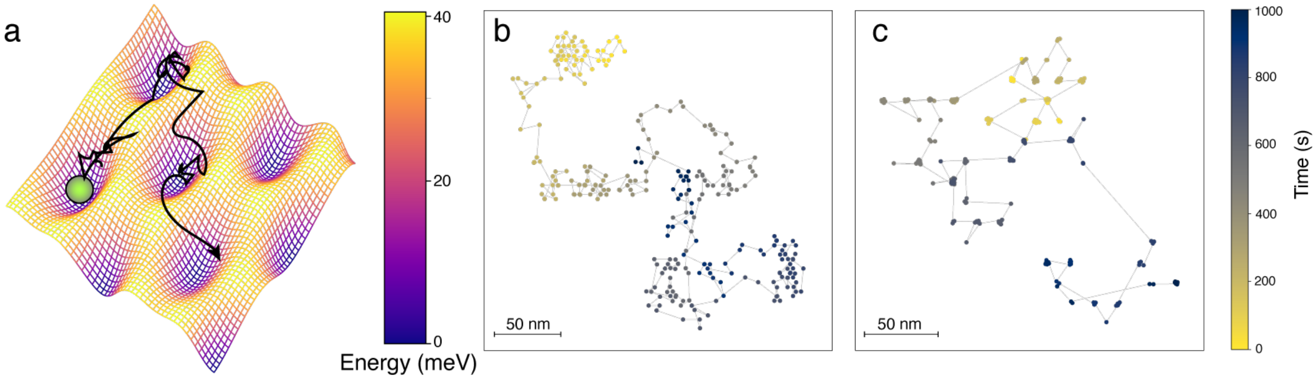

Figure 5.1: Moiré energy surface and sample trajectories. (a) Schematic showing a moiré superlattice energy landscape with an example of molecular adsorbate diffusive motion. (b-c) Example spatiotemporal trajectories simulated at $T = 22$ K for (b) $p = 4$ nm, and (c) $p = 16$nm.

details of simulation. The adsorbates were allowed to thermally diffuse by hopping between nearest-neighbor graphene lattice sites, while the moiré superlattice energy landscape influences the trapping of the adsorbates at each site. Figure 5.1b-c show example spatiotemporal trajectories simulated at $T = 22$ K for moiré periodicities of $p = 4$ nm and $p = 16$ nm, where distinct diffusive behaviors are observed over a long period of observation time of $t = 1000$s. For a small moiré periodicity of $p = 4$ nm, the motion of the adsorbates resembles uncorrelated normal diffusion (Figure 5.1b), whereas for large moiré periodicity of $p = 16$ nm the motion is punctuated by trapping events in the potential wells over the moiré substrate (Figure 5.1c). At such long observation timescales, it is easy to distinguish different diffusion patterns for small and large values of moiré periodicity in Figure 5.1b-c.

To identify the mechanism of diffusion from simulated trajectories of different temperatures and periodicities, we developed a workflow that uses canonical statistical tests in combination with a temporal convolutional deep neural network algorithm – see Methods section.(*86*) Convolution deep neural networks allow us to analyze the diffusion behavior for single particle trajectories with a limited number of diffusion steps, which is often the case in experiments(*87*). The network analyzes the trajectory at different time resolutions through convolutional layers of varying sizes and dilations, taking cue from the p-variation statistical method.(*88*) The network is trained on tens of thousands of simulated trajectories with ground truth diffusive behavior based on methods published in the previous literature. Given any input trajectory, the network classifies the underlying behavior by outputting a vector that contains the probability values for relevant classes of diffusion (e.g. Brownian, continuous time random walk (CTRW), etc). Through this neural network, we classify the diffusion as either Brownian or CTRW. Figure 5.2a shows the probability of Brownian and CTRW classes as outputted by the neural network for 1000 trajectories simulated for 300 frames at 1.5 frames per second for various values of moiré periodicity and temperatures.

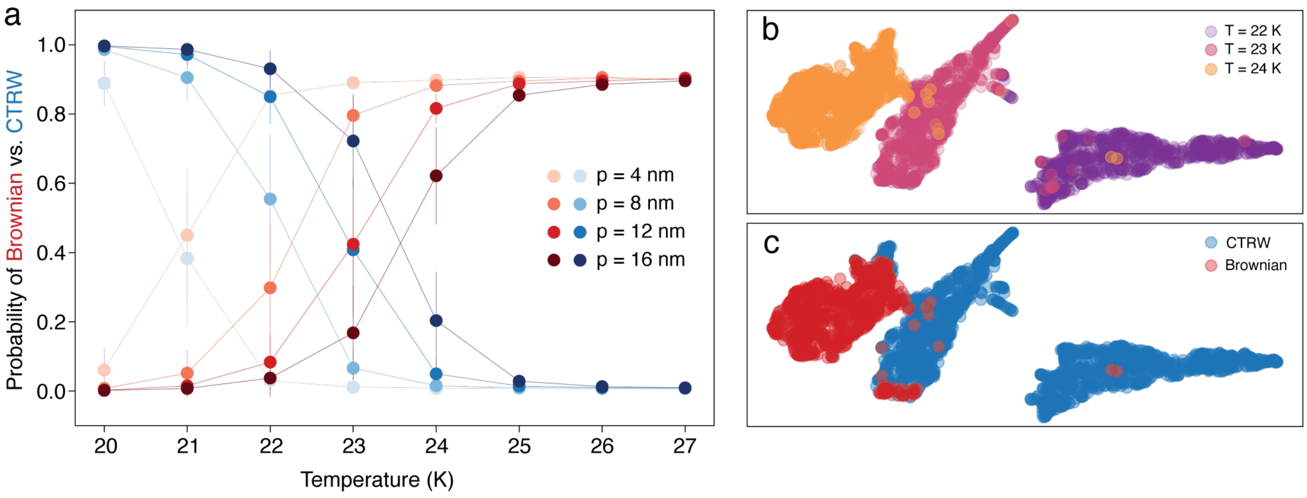

Figure 5.2: Machine learning analysis of simulated trajectories. (a) Diffusion mode probabilities for variable temperature and moiré size for simulated trajectories as outputted by the neural network. Error bars represent the standard deviation of the estimated probabilities which are larger within the transition region. Figures (b)-(c) show the universal manifold approximation and projection (UMAP) representation for each trajectory with (b) showing temperature and (c) showing the diffusion class.

The probability values show that on finite time scales the system smoothly transitions from CTRW to the Brownian limit with increasing temperature. Larger moiré periodicity has the effect of increasing the required temperature to reach the Brownian limit.

Using a universal manifold approximation and projection (UMAP) in Figure 5.2b-c, we visualized the latent space embedding of trajectories analyzed in Figure 5.2a for a moiré superlattice with $p = 16$nm at the transition temperature region of 22K to 24K. Figure 5.2c shows the predicted classes of diffusion for different trajectories in latent space where the Brownian and CTRW predictions formed visible clusters, but with partial overlapping. The Brownian cluster corresponds well with the trajectories obtained at $T = 24$K, while the CTRW cluster corresponds to trajectories at $T = 22$K. The trajectories at the intermediate temperature of $T = 23$K lie in the intermediary region between the 22K and 24K regions in Figure 5.2b, where the CTRW and Brownian regions clusters overlap in Figure 5.2c. This confirms that the crossover in diffusive behavior at this temperature for $p = 16$nm is in agreement with the neural network classification result in Figure 5.2a.

To better understand the transition from CTRW to Brownian behavior, we also performed canonical statistical tests on the simulated trajectories. The ensemble averaged mean squared displacement (EMSD) is defined as $< (x(t) - x(0))^2 >$ ($< . >$ denotes ensemble average over many trajectories) and is shown in Figure 5.3a for a range of moiré periodicities from p =4 to 16 nm at $T = 20$ K. Linear fits of the EMSDs in log-log scale yield diffusion exponents starting at approximately 1 for $p = 4$ nm and decreasing down to 0.6 as the moiré

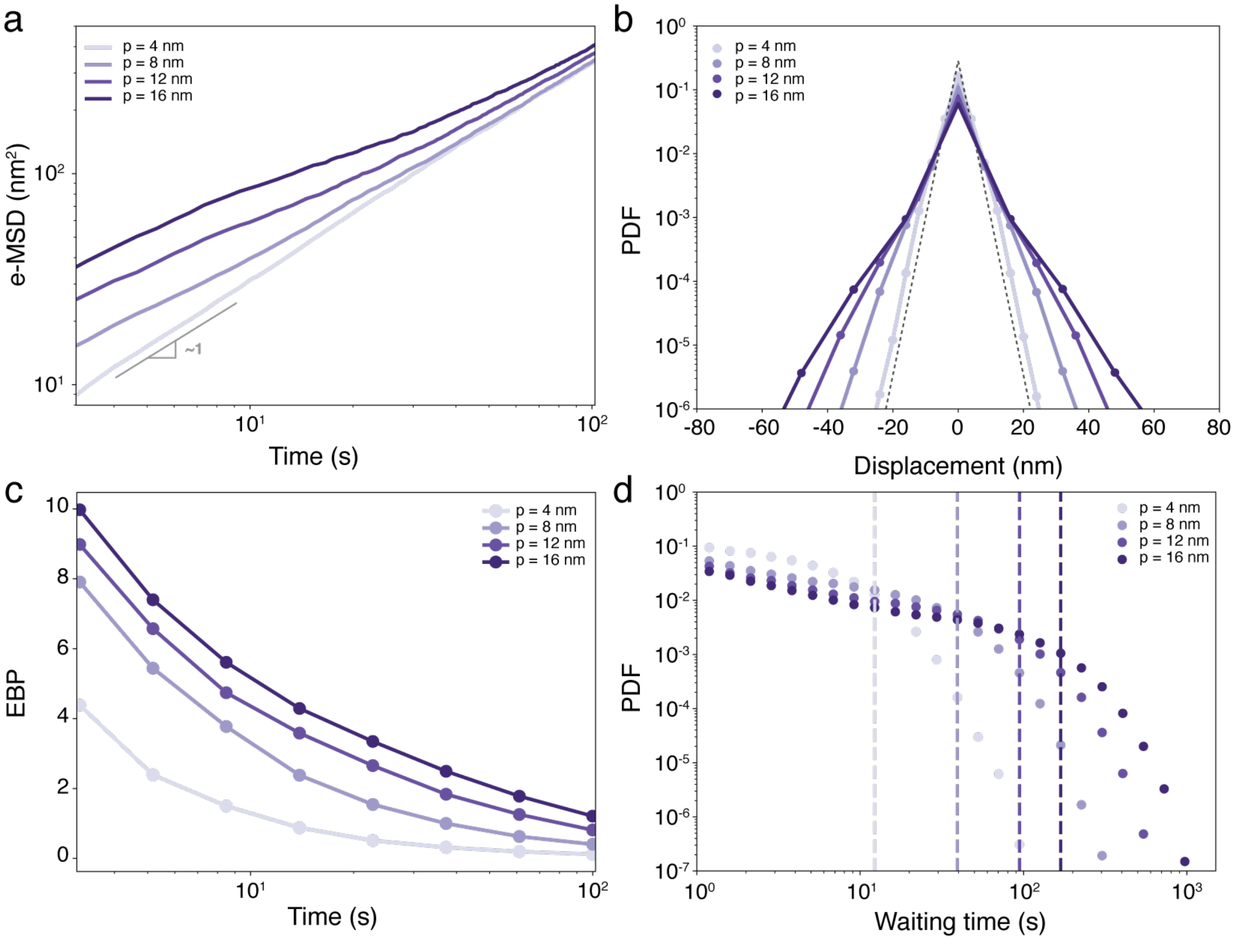

Figure 5.3: Statistics of simulated trajectories. (a) Ensemble averaged mean squared displacements (EMSD) vs. time for trajectories collected at T=20 K and p =4 to 16 nm. (b) Probability distribution of displacements over the time window of Δt = 5s. The dashed line indicates an exponential probability distribution. (c) Ergodicity breaking parameter (EBP) as a function of observation time. (d) Probability distribution of waiting times, τ. The dashed lines indicate the critical timescale where the decay of the distribution becomes faster than τ^{-2}.

periodicity increases to 16 nm, suggesting that anomalous subdiffusion occurs for larger moiré periodicities. We further see evidence of this anomaly by analyzing the probability distribution of displacements over a time window of 5s, shown in Figure 5.3b for the same temperature of $T = 20$K. For a normal Brownian motion the displacement distribution follows a zero mean Gaussian distribution as a result of its finite average time between hops and the central limit theorem. The displacement probability distributions, shown in Figure 5.3b, however, shows increasing non-gaussianity with increasing moiré size as seen from the non-negative curvature of the distribution in semi-log scale. The cusp at the origin of the distribution is consistent with the punctuated trapping events observed in the sample trajectory in Figure 5.1c. Whereas a Gaussian would be concave on a semi-log scale, the displacements for $p = 4$ nm appear to have almost no curvature, with the curvature strictly positively increasing with growing moiré periodicity as evidenced by the enhanced decay of the slope along the tails.

Since ergodicity is broken in the situation of the subdiffusive CTRW(*24*), we investigate the degree of ergodicity breaking by calculating the ergodicity breaking parameter (EBP), defined as

$$\mathrm{EBP}(t,\Delta) = \frac{< \overline{x}^2(t,\Delta)^2 > - < \overline{x}^2(t,\delta) >^2}{< \overline{x}^2(t,\Delta) >^2} \tag{5.1}$$

a dimensionless quantity which estimates the variance of single particle time averaged-MSDs (TAMSD) relative to the EMSD as $\Delta/t \to 0$, where the TAMSD is defined as:

$$\overline{x}^2(t,\Delta) = \frac{1}{t-\Delta}\int_0^{t-\Delta} [x(t'+\Delta) - x(t')]dt' \tag{5.2}$$

Here Δ is the time delay window and t is the total observation time.(*89*) Figure 5.3c plots the EBP for different moiré periodicities at a fixed temperature of $T = 20$K as a function of the observation time t, which we increased up to 100 seconds while fixing the frame rate at 1.5 frames per second. The plot shows that the system slowly and asymptotically approaches 0, suggesting that given infinite time the system would become ergodic. Larger moiré periodicity induces larger ergodicity breaking. For reference, for Brownian motion in the limit where the time between steps is much smaller than the total observation time, the ergodicity breaking parameter should be 4/3 of the ratio between the spacing between steps and the total observation time. For 16 nm periodicity, the EBP is roughly 2 after 100 seconds, over two orders of magnitude larger than the expected value, $9 \cdot 10^{-3}$, for ideal Brownian motion.

The origin of the ergodicity breaking is related to the probability distribution of the waiting time sampling over the moiré superlattice energy landscape. In Figure 3d, we see the probability distribution of waiting times, τ, that the adsorbate spends between hops at each moiré superlattice site. The waiting time probability distribution is characterized by two regimes separated by a relaxation time, defined as the critical waiting time where the decay of the distribution becomes faster than $1/\tau^2$, which in log-log corresponds to where the slope is consistent with a heavy tailed distribution (relaxation time is shown with

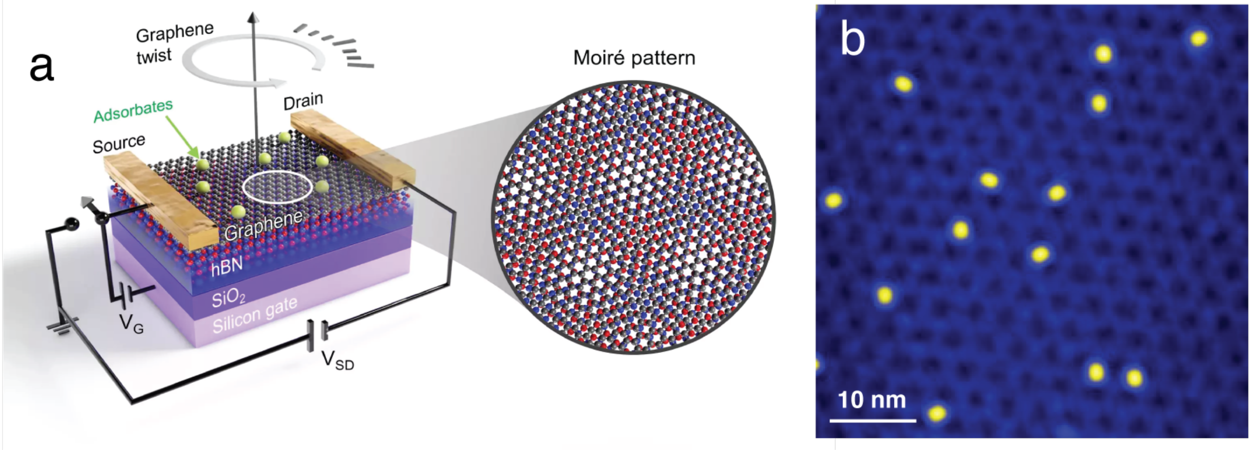

Figure 5.4: Moiré experimental system. (a) Moiré G/h-BN FET with adsorbates. (b) STM topography of F4TCNQ on a $p = 3.8$ nm G/h-BN moiré device.

a dashed line in Figure 5.3d). In the CTRW model, the subdiffusivity occurs when the waiting time distribution decays slower than $1/\tau^2$ (i.e., $\tau^{-\alpha}$, where $|\alpha| < 2$), resulting in an undefined mean waiting time.(*89*) When observation (sampling) time is below the relaxation time, the probability distribution of waiting times shows a slope consistent with the heavy tail distribution of the subdiffusive CTRW behavior. From Figure 5.3d, we see that the relaxation time increases as the moiré periodicity is increased, suggesting that subdiffusion and ergodicity breaking persist longer for larger moiré periodicity. This is consistent with the trend of the EBP being greater for larger moiré periodicities as well as the anomalies present in the displacement distribution and EMSDs in Figure 5.3a-c.

To verify the predictions from our Monte Carlo simulations, we realized this system in experiment by imaging diffusion of organic F4TCNQ adsorbates over a graphene/h-BN moiré superlattice with a moiré periodicity of $p = 3.8$ nm using gate-tuned STM under ultra-high vacuum (Figure 5.4a). In this setup adsorbates are thermalized through resistive heating of the substrate via a source-drain current at three temperatures of $T = 21$K, 23K, and 25 K. We collected over 10 trajectories that are each 30 frames long. Figure 5.4b shows a representative frame collected for a moiré surface with $p = 3.8$ nm at $T = 23$ K. Figure 5.5 shows a scatter plot of the spatiotemporal trajectories collected at different temperatures, where the circles indicate the mean diffusion length after 30 frames. We estimated the average likelihood of each diffusion class as predicted by the convolutional neural network at each of the three temperatures for a moiré periodicity of $p = 3.8$ nm. The results show that we recover the trend of approaching the Brownian limit with increasing temperature.

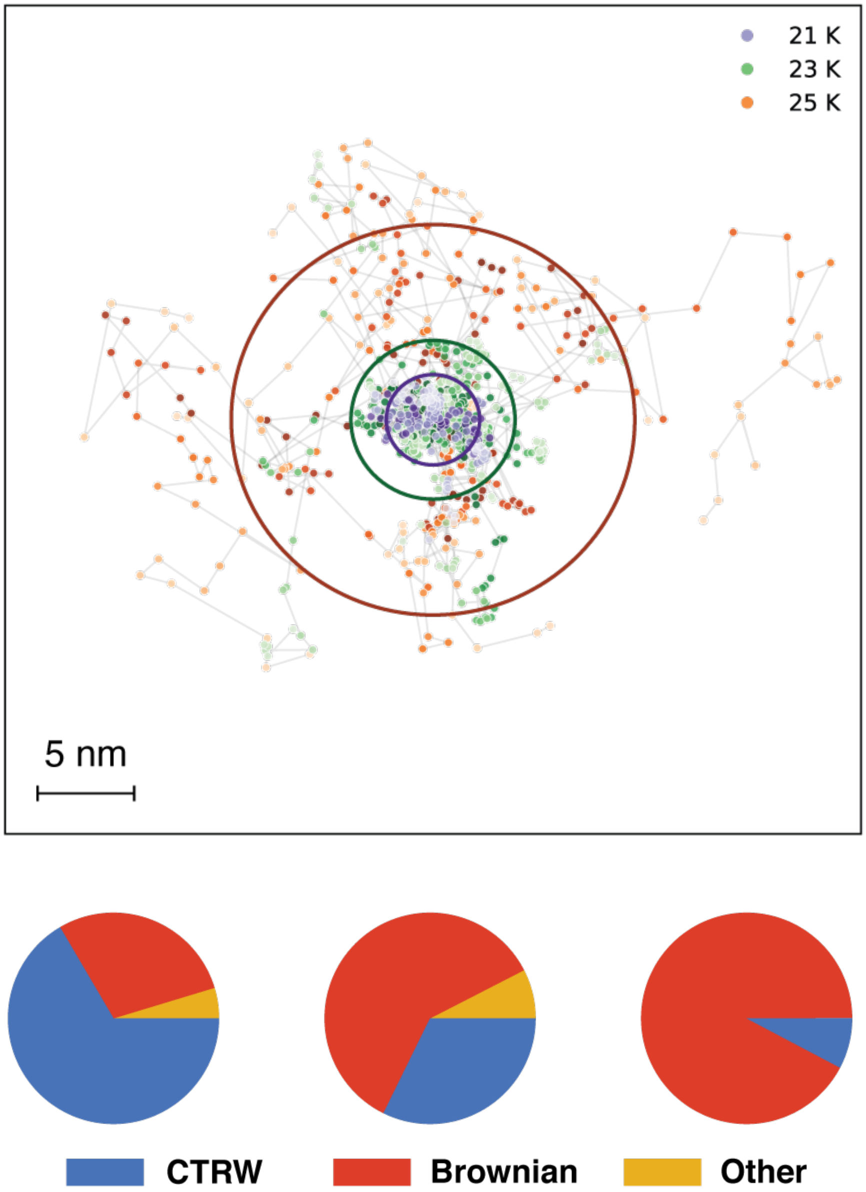

Figure 5.5: Neural analysis of experimental trajectories. Diffusion trajectories for variable temperatures 21K, 23K, and 25K. Circles indicate diffusion length after 30 frames (450 seconds). Diffusion mode average likelihoods for 3.8 nm moiré size and temperatures 21K, 23K, and 25K estimated using deep learning.

5.3 Discussion

To investigate why a CTRW process with a weak ergodicity breaking may arise in this system, we first need to understand the effect of the moiré potential on the waiting times. In the absence of a moiré potential energy surface, every adsorption site is identical from the perspective of the diffusing adsorbate. We can obtain this same behavior with the moiré potential in the high temperature limit. In this setting, we can describe the trapping times everywhere by a single trapping rate, as is the case for Brownian motion. As we cool the system, the corrugation of the potential energy surface induces greater variation in the trapping times of each adsorption site, breaking the adsorption site energy degeneracy. As a result, the waiting time probability distribution becomes a mixture of sites of different rates. To achieve ergodicity in this case, the system should be given enough time so that any single adsorbate visits all microstates. The excess value of the ergodicity breaking parameter, particularly in the case of larger moiré periodicity, suggests that the observation timescales considered here are insufficient for each adsorbate to explore the entire microstate space. Because of the adsorption site energy variation induced by the moiré, increasing the moiré periodicity increases the number of unique energy microstates and in turn raises the time required to explore the entire state space. The increased relaxation time from increasing the moiré periodicity is a direct observation of this consequence, which suggests that the CTRW behavior and consequently ergodicity breaking persist for greater timescales. Therefore, the relaxation time of diffusion becomes significant relative to the observation time and ergodicity is not held.

We note that even for small moiré periodicities, where the diffusion behavior is characterized as Brownian, the diffusion is still not Gaussian. At all moiré periodicity values tested in simulation, the probability distribution of displacements in Figure 5.3b remains non-Gaussian. At large moiré periodicities, the subiffusive slope of EMSD and non-gaussianity of the distribution of displacements indicate the diffusion is categorically a subdiffusive CTRW, in agreement with the neural network classification. Even at the lowest periodicity value of $p = 4$ nm, the probability distribution of displacement is best described by an exponential fit (dashed line in Figure 5.3b). This type of diffusion has been previously reported in the literature as Brownian yet non-Gaussian and has been observed in heterogeneous systems with slowly varying fluctuations including glassy supercooled liquids.(*85*, *90–92*) A proposed model for why displacements should follow an exponential distribution rather than a Gaussian is because the system is described a mixture of Gaussians of different widths.(*91*) Within our system, the width of a single Gaussian component corresponds to the escape rate at a given graphene lattice point. High energy regions on the moiré superlattice correspond to faster moving areas and consequently wider Gaussian components, and likewise for lower energy regions for narrower components. The initial algebraic decay of the waiting time distribution is consistent with a broad distribution over energy values and by extension the widths of these Gaussian components. These components become mixed as the particle visits different areas of the moiré superlattice and different trajectories are aggregated, with the degree of mixing reflected by the EBP decay over time. Previous work has shown that even across subdiffusive, crossover, and Brownian regimes in time, the exponential displacement distribution persists. Similarly, we see that across all moiré periodicities, even when the duration of the subdiffusive regime changes, the exponential behavior in the tails of the displacement distribution is apparent.(*91*) Therefore, we can conclude that when we decrease the moiré periodicity in our system, the non-ergodicity disappears but the non-Gaussianity of the displacement probability distribution remains as a result of the microstate energy heterogeneity.

In summary, we studied the effect of engineering the energy surface of moiré superlattice substrates on surface diffusion of adsorbates at the molecular scale by tuning the twist angle of a moiré using gate-tuned STM and Monte Carlo simulations. Using canonical statistical tests as well as a temporal convolutional deep neural network algorithm, we characterized the underlying diffusion mechanism of spatiotemporal trajectories obtained under various conditions of moiré periodicity and temperature. Our results show that the spatial heterogeneity of the moiré superlattice energy surface induces weak ergodicity breaking in diffusion, which manifests as a continuous time random walk of the molecular surface adsorbates at low temperatures that can persist at time scales comparable to the total observation time, by breaking the microstate energy degeneracy of graphene lattice adsorption sites. Upon increasing the temperature, the diffusion mechanism crosses over toward the Brownian limit where the system approaches ergodic behavior with the ensemble average and the time average converging to the same value in the long time limit due to reduced variation between microstates at high temperature. The analysis presented in this work shows that larger moiré periodicities possessed greater ergodicity breaking due to broader energy variation across a higher number of unique microstates. While our simulations and experiments were done with one physical system in mind, we believe our central finding to be general, that broad spatial energy variation of microstates is a sufficient condition for significant ergodicity breaking. This would suggest that in other physical systems, ergodicity can be tuned through spatial control of energy heterogeneity. Our findings show that engineering the energy surface of the moiré substrate has profound implications in tuning the stochastic motion of molecular surface adsorbates with relevance to bottom-up assembly of nanoscale devices.

5.4 Methods

Simulating Graphene Lattice Random Walks in a Moiré Potential

While there are many different graphene surface binding sites, metastable adsorption sites always correspond to high symmetry positions. For idealized point particles, this can correspond to vacancies within the honeycomb lattice whereas for molecules, such F4TCNQ which has previously been studied on graphene, occupation of the bridge site between atoms is energetically preferred. These differences can only change the lattice geometry up to a constant scale factor, and all within the same order as the graphene lattice. For simplicity we chose the lattice geometry of allowed adsorption sites to be the same as the graphene lattice. For the moiré potential, we model it using a scalar field overlay with triangular lattice symmetry matching that of the graphene/hBN moiré.(*93*, *94*) This is a periodic function as follows:

$$u(x,y) = \delta[3 - \sum_{\mathbf{g}} \cos(\mathbf{g}\cdot\mathbf{r})] \tag{5.3}$$

$$\sum_{\mathbf{g}} \cos(\mathbf{g}\cdot\mathbf{r}) = 2\cos\theta_x \cos\theta_y + \cos 2\theta_y \tag{5.4}$$

$$\theta_x = \frac{2\pi x}{a} \tag{5.5}$$

$$\theta_y = \frac{2\pi y}{a\sqrt{3}} \tag{5.6}$$

$$\tag{5.7}$$

Here δ is the moiré potential energy amplitude, $\mathbf{g}$ are hexagonal reciprocal lattice vectors, $\mathbf{r}$ is a position (x, y) in real space, and a is the moiré superlattice constant. Density functional calculations predict that for F4TCNQ molecules on graphene the difference in adsorption energy from AA to AB moiré stacking is 39.8 meV, which we use for setting the moiré potential energy amplitude. We model the motion of the particle using a continuous time Markov chain where the transition rate out of a given state follows the Boltzmann statistics based on the location of the state.(*95*) In thermal equilibrium, at any given time, the likelihood, P_i, of finding an adsorbate in a given graphene lattice site, i, depends on the energy of the site, $u(x_i, y_i)$, through the Boltzmann distribution, where k_B is the Boltzman constant and T is the temperature.

$$P_i \sim \exp\{-u(x_i, y_i)/k_BT\} \tag{5.8}$$

The rate of escape from this lattice site is inversely proportional to probability at that site, P_i, defined as the rate parameter, w_i,

$$1/w_i \sim \exp\{-u(x_i, y_i)/k_BT\} \tag{5.9}$$

The time between nearest neighbor jumps, τ, are sampled using an exponential distribution, which maintains the memorylessness property and independence of jumps akin to a Poisson process. It's probability density, $f(\tau|w_i)$, is thus:

$$f(\tau|w_i) = w_i \exp\{-w_i\tau\} \tag{5.10}$$

Temporal Convolutional Neural Network

To analyze the mechanism of diffusion from our simulated and experimental trajectories, we trained temporal convolutional neural networks to classify trajectories as either Brownian, CTRW, or other. For the results shown in Figure 5.2, we used supervised learning to train our neural network on 10,000 labeled trajectories of 300 frames, simulated using ideal stochastic processes. However, due to the lower bound (frame rate) and upper bound (total observation time) in experimental trajectories from STM, typical contiguous trajectories contain much fewer frames compared to the simulated ones. As a result, we trained a second neural network model configured for trajectories of 30 frames instead of 300 to handle experimental data, the output of which is shown in Figure 5.5.

Recent progress has shown the ability for neural networks to correctly identify different modes of diffusion.(*96*, *97*) For our work, we took a cue from a convolutional neural network architecture inspired by canonical p-variation tests.(*10*, *88*) This architecture uses convolutional blocks of varying filter sizes consisting of three repetitions of convolutional layers followed by rectified linear units, batch normalization and, at the end, a global pooling layer in order to capture the structure of the trajectory at different time scales. Each convolutional layer contains 32 filters. The convolutional blocks are then concatenated into a single vector prior to a series of fully connected layers. The output layer of the network is a softmax in order to produce class probabilities for each diffusion mode. Our original network used to analyze our simulated data from Monte Carlo simulation is unmodified from the previously published work. However, for video STM, capturing a single frame is very expensive with one frame requiring over 10 minutes. This limited the total size of our videos to 200 frames. In addition, low temporal resolution means that particles often stay in the field of view for much fewer frames. Therefore, we modified the original neural network architecture to perform well with only 30 frame trajectories instead of 300 frames. The key changes to the architecture from the original version were smaller filter sizes and dilations so as not to over downsample a trajectory as it passes through the network, global average pooling instead of max pooling, one additional convolutional block, and an increased width in the first fully connected layer. The details of this architecture are shown in Table 5.1.

For training we used gradient-based supervised learning. Here labeled trajectories are generated corresponding to either Brownian, subdiffusive continuous time random walk, or subdiffusive fractional Brownian motion.(*98*) Following previous literature, we labeled instances of subdiffusive continuous time random walk and fractional brownian motion with diffusivity exponents greater than 0.9 as Brownian, since above this threshold, for only a few frames, these processes are virtually indistinguishable from Brownian motion. Training instances are batched together with 32 instances per batch. We trained the network once over 160,000 unique simulated trajectories with ground truth mode of diffusion by back-propagating the derivative of a categorical cross-entropy loss function, which measures the Kullback-Leibler divergence between the distributions of the probability values in true class probabilities vs. the classified probabilities by the network, with respect to the weights of the network using a stochastic gradient descent algorithm.(*99*, *100*) For our optimizer, we used an Adam optimizer with an initial learning rate of 1e-5. (*101*)

After the training ended, we validate the accuracy of our models on a new set of 3,200 labeled, simulated trajectories that are not seen during the training stage. The 300-frame model achieved a validation accuracy of 90% in agreement with reported accuracies for this

Name, Layer	# of Neurons	Filter sizes	Dilations
x_1, Convolutional Block	N/A	3, 3, 3	1, 2, 2
x_2, Convolutional Block	N/A	4, 3, 3	1, 2, 2
x_3, Convolutional Block	N/A	5, 3, 3	1, 2, 2
x_4, Convolutional Block	N/A	8, 3, 3	1, 2, 2
x_5, Convolutional Block	N/A	6, 3, 3	1, 2, 1
x_6, Convolutional Block	N/A	8, 3, 3	1, 2, 1
x_7, Convolutional Block	N/A	7, 3, 3	1, 1, 1
x_{stack}, Concat($\{x_i\}$)	N/A	N/A	N/A
fc1, Fully Connected	2048	N/A	N/A
fc2, Fully Connected	512	N/A	N/A
out, Fully Connected	3	N/A	N/A
Softmax	N/A	N/A	N/A

Table 5.1: Motion classification neural network architecture.

architecture.(*10*) The 30-frame model achieves a validation accuracy of 84%.

During inference, for a two-dimensional trajectory $(\mathbf{x}, \mathbf{y})$ each coordinate was treated independently and was preprocessed to calculate the discrete difference time series from the original positions of the trajectory, $\Delta\mathbf{x}$. This time series was then shifted to have zero mean and unit variance before finally being passed in as input to the neural network.

Universal Manifold Approximation and Projection (UMAP) Representation

Universal Manifold Approximation and Projection (UMAP) is a dimensionality reduction algorithm that uses manifold learning based on manifold theory and topological data analysis in order to preserve local structure and distances. This results in clusters of similar data points as well as global structure between dissimilar clusters within the learned lower dimensional space.(*102*) In order to generate a two-dimensional UMAP representation, trajectories are fed into our neural network model. Outputs from the last layer preceding the softmax classification layer represent latent features extracted by the network from the trajectory. These features, which form a high dimensional vector, are input into the UMAP algorithm. After the UMAP algorithm fits a latent space manifold, a two-dimensional representation of each trajectory is given, the result of which is shown in Figure 5.2b-c and color coded based on the temperature and the mechanism of diffusion.

Experimental System

We created graphene/h-BN field effect transistor devices on Si/SiO$_2$ through mechanical exfoliation. Source/drain contacts were made using a stencil mask and e-beam evaporation. After placement into ultra-high vacuum (UHV), the device is annealed for 12 hours at 400C.

STM images are acquired under UHV and low temperature conditions (T = 4.5K) using a commercial Omicron LT STM with Pt/Ir probes. Images are acquired in constant current mode. STM tips are conditioned and calibrated on Au(111) before imaging on the graphene/h-BN device.(*103*) To analyze topography data for diffusion, STM images are min-max scaled in order to be analyzed for particle detection and trajectory linking using the TrackPy package in Python.(*11*, *70*)

To create STM videos for analyzing diffusion, the device is driven with a source-drain current for 15 seconds to heat up the graphene through resistive heating. Afterwards, the system is allowed to cool back down to T = 4.5 K for STM imaging. This process is repeated 200 times to collect 200 frames. A negative gate voltage of -20V is applied to ensure that the molecule remains neutral during the diffusion process. While it's not possible to measure the temperature on the graphene directly during resistive heating, we calibrated the temperature using a series of experiments using indirect heating by warming the entire device for 1 minute, during which we could measure the temperature. We do this 5 times in succession to drive diffusion for a total of 5 minutes. To match a particular source-drain current to a temperature, MSDs collected after 5 minutes of diffusion were compared between the resistive and indirect heating methods and matched based on similarity.

Bibliography

1. J. F. Kincaid, H. Eyring, A. E. Stearn, *Chemical Reviews* **28**, 301–365, eprint: `https://doi.org/10.1021/cr60090a005`, (`https://doi.org/10.1021/cr60090a005`) (1941).
2. S. Klumpp, M. Scott, S. Pedersen, T. Hwa, *Proceedings of the National Academy of Sciences* **110**, 16754–16759, ISSN: 0027-8424, eprint: `https://www.pnas.org/content/110/42/16754.full.pdf`, (`https://www.pnas.org/content/110/42/16754`) (2013).
3. A. Ramasubramanian *et al.*, *The Journal of Physical Chemistry C* **123**, 10237–10245, eprint: `https://doi.org/10.1021/acs.jpcc.9b00436`, (`https://doi.org/10.1021/acs.jpcc.9b00436`) (2019).
4. K. Yamaguchi, M. Ogasawara, K. Okamoto, *Journal of Applied Physics* **72**, 5919–5925, eprint: `https://doi.org/10.1063/1.351900`, (`https://doi.org/10.1063/1.351900`) (1992).
5. E. F. Fama, *The Journal of Finance* **25**, 383–417, ISSN: 00221082, 15406261, (`http://www.jstor.org/stable/2325486`) (1970).
6. J. Ho, A. Jain, P. Abbeel, *CoRR* **abs/2006.11239**, arXiv: `2006.11239`, (`https://arxiv.org/abs/2006.11239`) (2020).
7. T. Tansel, O. M. Magnussen, *Phys. Rev. Lett.* **96**, 026101, (`https://link.aps.org/doi/10.1103/PhysRevLett.96.026101`) (2 Jan. 2006).
8. M. Hooshmand, W. Zhong, J. Zhao, W. Windl, M. Ghazisaeidi, *Data in Brief* **30**, 105381, ISSN: 2352-3409, (`https://www.sciencedirect.com/science/article/pii/S2352340920302754`) (2020).
9. T. Furnival *et al.*, *Chemical Physics Letters* **683**, Ahmed Zewail (1946-2016) Commemoration Issue of Chemical Physics Letters, 370–374, ISSN: 0009-2614, (`https://www.sciencedirect.com/science/article/pii/S0009261417304001`) (2017).
10. V. Jamali *et al.*, *Proceedings of the National Academy of Sciences* **118**, ISSN: 0027-8424, eprint: `https://www.pnas.org/content/118/10/e2017616118.full.pdf`, (`https://www.pnas.org/content/118/10/e2017616118`) (2021).

11. J. C. Crocker, D. G. Grier, *Journal of Colloid and Interface Science* **179**, 298–310, ISSN: 0021-9797, (`https://www.sciencedirect.com/science/article/pii/S0021979796902179`) (1996).
12. S. Renisch, R. Schuster, J. Wintterlin, G. Ertl, *Phys. Rev. Lett.* **82**, 3839–3842, (`https://link.aps.org/doi/10.1103/PhysRevLett.82.3839`) (19 May 1999).
13. T. Zambelli, J. Trost, J. Wintterlin, G. Ertl, *Phys. Rev. Lett.* **76**, 795–798, (`https://link.aps.org/doi/10.1103/PhysRevLett.76.795`) (5 Jan. 1996).
14. R. van Gastel, E. Somfai, W. van Saarloos, J. W. M. Frenken, *Nature* **408**, 665–665, (`https://doi.org/10.1038/35047156`) (2000).
15. M. Ø. Pedersen *et al.*, *Phys. Rev. Lett.* **84**, 4898–4901, (`https://link.aps.org/doi/10.1103/PhysRevLett.84.4898`) (21 May 2000).
16. R. van Gastel, E. Somfai, S. B. van Albada, W. van Saarloos, J. W. M. Frenken, *Phys. Rev. Lett.* **86**, 1562–1565, (`https://link.aps.org/doi/10.1103/PhysRevLett.86.1562`) (8 Feb. 2001).
17. P. Rerkkumsup *et al.*, *Review of Scientific Instruments* **75**, 1061–1067, eprint: `https://doi.org/10.1063/1.1651637`, (`https://doi.org/10.1063/1.1651637`) (2004).
18. B. G. Briner, M. Doering, H.-P. Rust, A. M. Bradshaw, *Science* **278**, 257–260, eprint: `https://www.science.org/doi/pdf/10.1126/science.278.5336.257`, (`https://www.science.org/doi/abs/10.1126/science.278.5336.257`) (1997).
19. R. Brown, *The Philosophical Magazine* **4**, 161–173, eprint: `https://doi.org/10.1080/14786442808674769`, (`https://doi.org/10.1080/14786442808674769`) (1828).
20. K. Oura, V. Lifshits, A. Saranin, A. Zotov, M. Katayama, *Surface Science: An Introduction* (Springer Berlin Heidelberg, 2013), ISBN: 9783662051795, (`https://books.google.com/books?id=aeznCAAAQBAJ`).
21. E. Barkai, Y. Garini, R. Metzler, *Physics Today* **65**, 29–35, eprint: `https://doi.org/10.1063/PT.3.1677`, (`https://doi.org/10.1063/PT.3.1677`) (2012).
22. J.-P. Bouchaud, A. Georges, *Physics Reports* **195**, 127–293, ISSN: 0370-1573, (`https://www.sciencedirect.com/science/article/pii/037015739090099N`) (1990).
23. R. Metzler, J.-H. Jeon, A. G. Cherstvy, E. Barkai, *Phys. Chem. Chem. Phys.* **16**, 24128–24164, (`http://dx.doi.org/10.1039/C4CP03465A`) (44 2014).
24. Y. He, S. Burov, R. Metzler, E. Barkai, *Phys. Rev. Lett.* **101**, 058101, (`https://link.aps.org/doi/10.1103/PhysRevLett.101.058101`) (5 July 2008).
25. R. Metzler, J. Klafter, *Physics Reports* **339**, 1–77, ISSN: 0370-1573, (`https://www.sciencedirect.com/science/article/pii/S0370157300000703`) (2000).
26. G. R. Grimmett, D. R. Stirzaker, *Probability and Random Processes* (Oxford university Press, ed. 3, 2001), ISBN: 9788490225370.

27. J. Krieger, *Mean Squared Displacement of Anomalous and Normal Diffusion*, Nov. 2012, (`https://upload.wikimedia.org/wikipedia/commons/e/e6/Msd_anomalous_diffusion.svg`).
28. W. Wang, J. H. P. Schulz, W. Deng, E. Barkai, *Phys. Rev. E* **98**, 042139, (`https://link.aps.org/doi/10.1103/PhysRevE.98.042139`) (4 Oct. 2018).
29. M. Z. Bazant, C. H. Rycroft, (`https://ocw.mit.edu/courses/18-366-random-walks-and-diffusion-fall-2006/03f1932a896d9925ee0a492789756c11_lec23_chris.pdf`) (Feb. 2005).
30. G. Bel, E. Barkai, *Phys. Rev. Lett.* **94**, 240602, (`https://link.aps.org/doi/10.1103/PhysRevLett.94.240602`) (24 June 2005).
31. T. Miyaguchi, T. Akimoto, *Phys. Rev. E* **83**, 031926, (`https://link.aps.org/doi/10.1103/PhysRevE.83.031926`) (3 Mar. 2011).
32. T. Miyaguchi, T. Akimoto, *Phys. Rev. E* **87**, 032130, (`https://link.aps.org/doi/10.1103/PhysRevE.87.032130`) (3 Mar. 2013).
33. F. Liou *et al.*, *Nano Letters* **21**, PMID: 34653333, 8770–8776, eprint: `https://doi.org/10.1021/acs.nanolett.1c03039`, (`https://doi.org/10.1021/acs.nanolett.1c03039`) (2021).
34. X. Zhang, Z. Shao, X. Zhang, Y. He, J. Jie, *Advanced Materials* **28**, 10409–10442, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201601966`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201601966`) (2016).
35. T. Yang *et al.*, *J. Mater. Chem. C* **8**, 7272–7299, (`http://dx.doi.org/10.1039/D0TC00387E`) (22 2020).
36. M. S. Stark, K. L. Kuntz, S. J. Martens, S. C. Warren, *Advanced Materials* **31**, 1808213, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201808213`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201808213`) (2019).
37. J. Wan *et al.*, *Chem. Soc. Rev.* **45**, 6742–6765, (`http://dx.doi.org/10.1039/C5CS00758E`) (24 2016).
38. J. A. A. W. Elemans, *Advanced Functional Materials* **26**, 8932–8951, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/adfm.201603145`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/adfm.201603145`) (2016).
39. X. Peng, F. Zhao, Y. Peng, J. Li, Q. Zeng, *Soft Matter* **16**, 54–63, (`http://dx.doi.org/10.1039/C9SM01847F`) (1 2020).
40. Z.-F. Cai, H.-J. Yan, D. Wang, L.-J. Wan, *Nanoscale* **10**, 3438–3443, (`http://dx.doi.org/10.1039/C7NR08475G`) (7 2018).
41. K. Cui *et al.*, *Chem. Commun.* **50**, 10376–10378, (`http://dx.doi.org/10.1039/C4CC04189E`) (72 2014).

42. J. Repp *et al.*, *Phys. Rev. Lett.* **117**, 146102, (`https://link.aps.org/doi/10.1103/PhysRevLett.117.146102`) (14 Sept. 2016).

43. J.-H. Liu, L.-M. Yang, E. Ganz, *J. Mater. Chem. A* **7**, 11944–11952, (`http://dx.doi.org/10.1039/C9TA01188A`) (19 2019).

44. M. Rajapakse *et al.*, *npj 2D Materials and Applications* **5**, 30, (`https://doi.org/10.1038/s41699-021-00211-6`) (2021).

45. E. Pomerantseva, Y. Gogotsi, *Nature Energy* **2**, 17089, (`https://doi.org/10.1038/nenergy.2017.89`) (2017).

46. Y. Li, H. Yan, B. Xu, L. Zhen, C.-Y. Xu, *Advanced Materials* **33**, 2000581, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.202000581`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.202000581`) (2021).

47. M. T. Sharbati *et al.*, *Advanced Materials* **30**, 1802353, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/adma.201802353`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/adma.201802353`) (2018).

48. M. Kühne *et al.*, *Nature* **564**, 234–239, (`https://doi.org/10.1038/s41586-018-0754-2`) (2018).

49. D. K. Bediako *et al.*, *Nature* **558**, 425–429, (`https://doi.org/10.1038/s41586-018-0205-0`) (2018).

50. J. Zhang *et al.*, *Nature Communications* **9**, 5289, (`https://doi.org/10.1038/s41467-018-07710-z`) (2018).

51. C. Uhlmann, I. Swart, J. Repp, *Nano Letters* **13**, PMID: 23356959, 777–780, eprint: `https://doi.org/10.1021/nl304483h`, (`https://doi.org/10.1021/nl304483h`) (2013).

52. H. S. Jung *et al.*, *JoVE*, e52711, (`https://www.jove.com/t/52711`) (2015).

53. C. R. Dean *et al.*, *Nature Nanotechnology* **5**, 722–726, (`https://doi.org/10.1038/nnano.2010.172`) (2010).

54. R. Decker *et al.*, *Nano Letters* **11**, PMID: 21553853, 2291–2295, eprint: `https://doi.org/10.1021/nl2005115`, (`https://doi.org/10.1021/nl2005115`) (2011).

55. K. S. Novoselov *et al.*, *Science* **306**, 666–669, eprint: `https://www.science.org/doi/pdf/10.1126/science.1102896`, (`https://www.science.org/doi/abs/10.1126/science.1102896`) (2004).

56. G. Binnig, H. Rohrer, *Surface Science* **126**, 236–244, ISSN: 0039-6028, (`https://www.sciencedirect.com/science/article/pii/0039602883907161`) (1983).

57. J. Tersoff, D. R. Hamann, *Phys. Rev. Lett.* **50**, 1998–2001, (`https://link.aps.org/doi/10.1103/PhysRevLett.50.1998`) (25 June 1983).

58. P. A. M. Dirac, N. H. D. Bohr, *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* **114**, 243–265, eprint: `https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1927.0039`, (`https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1927.0039`) (1927).
59. S. Wickenburg *et al.*, *Nature Communications* **7**, 13553 (2016).
60. J. Lu *et al.*, *Nature Communications* **10**, 477, (`https://doi.org/10.1038/s41467-019-08371-2`) (2019).
61. H.-Z. Tsai *et al.*, *Nature Electronics* **3**, 598–603, (`https://doi.org/10.1038/s41928-020-00479-4`) (2020).
62. H.-Z. Tsai *et al.*, *ACS Nano* **9**, PMID: 26482218, 12168–12173, eprint: `https://doi.org/10.1021/acsnano.5b05322`, (`https://doi.org/10.1021/acsnano.5b05322`) (2015).
63. T. L. Cocker, D. Peller, P. Yu, J. Repp, R. Huber, *Nature* **539**, 263–267, (`https://doi.org/10.1038/nature19816`) (2016).
64. Y. Zhang *et al.*, *Nature Physics* **4**, 627–630, (`https://doi.org/10.1038/nphys1022`) (2008).
65. A. H. Castro Neto, F. Guinea, N. M. R. Peres, K. S. Novoselov, A. K. Geim, *Rev. Mod. Phys.* **81**, 109–162, (`https://link.aps.org/doi/10.1103/RevModPhys.81.109`) (1 Jan. 2009).
66. P. Jarvinen *et al.*, *Nano Letters* **13**, PMID: 23786613, 3199–3204, eprint: `https://doi.org/10.1021/nl401265f`, (`https://doi.org/10.1021/nl401265f`) (2013).
67. K. A. Cochrane, A. Schiffrin, T. S. Roussy, M. Capsoni, S. A. Burke, *Nature Communications* **6**, 8312, (`https://doi.org/10.1038/ncomms9312`) (2015).
68. G. Czack, G. Kirschstein, W. Kurtz, F. Stein, in *W Tungsten: Supplement Volume A4 Surface Properties. Electron Emission*, ed. by W. Huisl, W. Kurtz, F. Stein (Springer Berlin Heidelberg, Berlin, Heidelberg, 1993), pp. 82–106, ISBN: 978-3-662-10154-4, (`https://doi.org/10.1007/978-3-662-10154-4_3`).
69. J. Barth, *Surface Science Reports* **40**, 75–149, ISSN: 0167-5729, (`https://www.sciencedirect.com/science/article/pii/S0167572900000029`) (2000).
70. D. Allan, T. Caswell, N. Keim, C. van der Wel, ISSN: 17446848, (`https://soft-matter.github.io/trackpy/v0.3.0/`) (2015).
71. X. Tong *et al.*, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* **12**, 4062–4081 (2019).
72. L. G. Brown, *ACM Comput. Surv.* **24**, 325–376, ISSN: 0360-0300, (`https://doi.org/10.1145/146370.146374`) (Dec. 1992).
73. M.-K. Hu, *IRE Transactions on Information Theory* **8**, 179–187 (1962).

74. A. P. Murani, *Journal of Magnetism and Magnetic Materials* **22**, 271–281, ISSN: 03048853 (1981).

75. F. Krzakala, J. Kurchan, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **76**, 1–13, ISSN: 15393755, arXiv: `0702546 (cond-mat)` (2007).

76. M. Bernaschi, A. Billoire, A. Maiorano, G. Parisi, F. Ricci-Tersenghi, *Proceedings of the National Academy of Sciences of the United States of America* **117**, 17522–17527, ISSN: 10916490 (2020).

77. D. A. Kessler, E. Barkai, arXiv: `arXiv:2107.09526v1` (2021).

78. W. Van Megen, S. M. Underwood, P. N. Pusey, *Physical Review Letters* **67**, 1586–1589, ISSN: 00319007 (1991).

79. P. Charbonneau, J. Kurchan, G. Parisi, P. Urbani, F. Zamponi, *Nature Communications* **5**, 1–6, ISSN: 20411723, arXiv: `1404.6809` (2014).

80. S. Kirkpatrick, C. D. Gelatt, M. P. Vecchi, *Science* **220**, 671–680, ISSN: 00368075 (1983).

81. G. Bel, E. Barkai, *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **73**, 1–14, ISSN: 15393755, arXiv: `0506338 (cond-mat)` (2006).

82. E. W. Montroll, *Journal of Mathematical Physics* **6**, 167–181 (1965).

83. R. Kutner, J. Masoliver, *European Physical Journal B* **90**, ISSN: 14346036, arXiv: `1612.02221` (2017).

84. M. F. Shlesinger, *European Physical Journal B* **90**, ISSN: 14346036 (2017).

85. F. Kindermann *et al.*, *Nature Physics* **13**, 137–141, ISSN: 17452481, arXiv: `1601.06663` (2017).

86. Y. LeCun, Y. Bengio, G. Hinton, *Nature* **521**, 436–444, (`https://doi.org/10.1038/nature14539`) (2015).

87. F. Besenbacher, E. Lægsgaard, I. Stensgaard, *Materials Today* **8**, 26–30, ISSN: 13697021, (`http://dx.doi.org/10.1016/S1369-7021(05)00843-6`) (2005).

88. M. Magdziarz, A. Weron, K. Burnecki, J. Klafter, *Physical Review Letters* **103**, 1–4, ISSN: 00319007 (2009).

89. M. Schwarzl, A. Godec, R. Metzler, *Scientific Reports* **7**, 1–18, ISSN: 20452322 (2017).

90. F. Rusciano, R. Pastore, F. Greco, *Physical Review Letters* **128**, 168001, ISSN: 0031-9007, (`https://doi.org/10.1103/PhysRevLett.128.168001`) (2022).

91. B. Wang, S. M. Anthony, C. B. Sung, S. Granick, *Proceedings of the National Academy of Sciences of the United States of America* **106**, 15160–15164, ISSN: 00278424 (2009).

92. T. F. Middleton, D. J. Wales, *Physical Review B - Condensed Matter and Materials Physics* **64**, 1–20, ISSN: 1550235X (2001).

93. J. Jung, A. M. Dasilva, A. H. Macdonald, S. Adam, *Nature Communications* **6**, 1–11, ISSN: 20411723, arXiv: `1403.0496` (2015).

94. M. Reguzzoni, A. Fasolino, E. Molinari, M. C. Righi, *Physical Review B - Condensed Matter and Materials Physics* **86**, 1–7, ISSN: 10980121 (2012).

95. W. J. Anderson, *Continuous-time Markov chains: an applications-oriented approach.* (Springer, 1991), p. 64, ISBN: 9781461277729.

96. N. Granik *et al.*, *Biophysical Journal* **117**, 185–192, ISSN: 15420086, (`https://doi.org/10.1016/j.bpj.2019.06.015`) (2019).

97. G. Muñoz-Gil *et al.*, *Nature Communications* **12**, ISSN: 20411723, arXiv: `2105.06766` (2021).

98. T. Dieker, *MSc theses, University of Twente, Amsterdam, The ...*, 77, (`http://www.wstein.org/home/wstein/www/home/simuw/simuw08/refs/fractal/dieker-2004-thesis-simulation_of_fractional_Brownian_motion.pdf`) (2004).

99. D. E. Rumelhart, G. E. Hintont, *Nature* **323**, 533–536 (1986).

100. S. Kullback, R. A. Leibler, *The Annals of Mathematical Statistics* **22**, 79–86, (`https://doi.org/10.1214/aoms/1177729694`) (1951).

101. D. P. Kingma, J. Ba, *Adam: A Method for Stochastic Optimization*, 2014, (`https://arxiv.org/abs/1412.6980`).

102. L. McInnes, J. Healy, J. Melville, arXiv: `1802.03426`, (`http://arxiv.org/abs/1802.03426`) (2018).

103. W. Chen, V. Madhavan, T. Jamneala, M. F. Crommie, *Phys. Rev. Lett.* **80**, 1469–1472, (`https://link.aps.org/doi/10.1103/PhysRevLett.80.1469`) (7 Feb. 1998).

104. S. Ernst, S. Wirth, M. Rams, V. Dolocan, F. Steglich, *Science and Technology of Advanced Materials* **8**, 347–351, eprint: `https://doi.org/10.1016/j.stam.2007.05.008`, (`https://doi.org/10.1016/j.stam.2007.05.008`) (2007).

105. M. Salmeron, in *Characterization of Materials* (John Wiley and Sons, Ltd, 2002), ISBN: 9780471266969, eprint: `https://onlinelibrary.wiley.com/doi/pdf/10.1002/0471266965.com084`, (`https://onlinelibrary.wiley.com/doi/abs/10.1002/0471266965.com084`).

106. M. J. Vasile, D. Grigg, J. E. Griffith, E. Fitzgerald, P. E. Russell, *Journal of Vacuum Science and Technology B: Microelectronics and Nanometer Structures Processing, Measurement, and Phenomena* **9**, 3569–3572, eprint: `https://avs.scitation.org/doi/pdf/10.1116/1.585846`, (`https://avs.scitation.org/doi/abs/10.1116/1.585846`) (1991).

107. K. Akiyama *et al.*, *Review of Scientific Instruments* **76**, 033705, eprint: `https://doi.org/10.1063/1.1865812`, (`https://doi.org/10.1063/1.1865812`) (2005).

108. W. Bao *et al.*, *Science* **338**, 1317–1321, eprint: `https://www.science.org/doi/pdf/10.1126/science.1227977`, (`https://www.science.org/doi/abs/10.1126/science.1227977`) (2012).

109. P. Hohenberg, W. Kohn, *Phys. Rev.* **136**, B864–B871, (`https://link.aps.org/doi/10.1103/PhysRev.136.B864`) (3B Nov. 1964).

110. W. Kohn, L. J. Sham, *Phys. Rev.* **140**, A1133–A1138, (`https://link.aps.org/doi/10.1103/PhysRev.140.A1133`) (4A Nov. 1965).

111. D. S. Sholl, J. A. Steckel (John Wiley and Sons, Inc., 2009), ISBN: 9781420045451.

112. G. H. Vineyard, *Journal of Physics and Chemistry of Solids* **3**, 121–127, ISSN: 0022-3697, (`https://www.sciencedirect.com/science/article/pii/0022369757900598`) (1957).

113. H. Eyring, *The Journal of Chemical Physics* **3**, 107–115, eprint: `https://doi.org/10.1063/1.1749604`, (`https://doi.org/10.1063/1.1749604`) (1935).

114. G. Henkelman, H. Jonsson, *The Journal of Chemical Physics* **113**, 9978–9985, eprint: `https://doi.org/10.1063/1.1323224`, (`https://doi.org/10.1063/1.1323224`) (2000).

Appendix A

Focused Ion Beam Probe Sharpening for Local Probe Microscopy

Achieving atomic resolution in local probe microscopy critically relies on the radius of curvature of the probe. Tersoff showed theoretically that a larger of radius of curvature effectively averages the local density of states over a larger area, reducing pointwise resolution.(*57*) Typically, the probe is conditioned *in situ* on a metal surface by repetitively poking the surface with the probe until the probe has an atomically sharp apex.(*104*, *105*) However, long periods of operation can render probes unusable even after repeated rounds of tip conditioning, possibly due to contamination during scanning or a perturbation of the tip geometry. It is then necessary to use more perturbative methods to resharpen the probe.

Focused ion beam (FIB) milling has been successful for creating and reconditioning probes without requiring the use of wet chemistry.(*106–108*) We can make AFM/STM probes using $\sim$ 1mm Tungsten wire. We use a wire cutter to make the probes which leave the probe macroscopically sharp, but microscopically blunt. An example of how sharp an untreated wire at this stage is shown in Figure A.1. In order to sharpen the probe, we first align the probe such that it is collinear with the with focused ion beam column. The reason for this is we have found that the patterning selectivity seems to be proportional to the patterning cross section. This means that patterning is maximized when the surface normal faces the beam direction and minimized when it is nearly perpendicular. The consequence of this is that by aligning the beam with the probe, local maxima of the uncut probe are further sharpened, their aspect ratios narrowed, while blunt portions of the probe are more quickly milled. For milling we use the following parameters.

1. Beam energy = 30kV
2. Beam current = 65nA
3. Milling depth = 1 micron
4. Dwell time = 1 microsecond

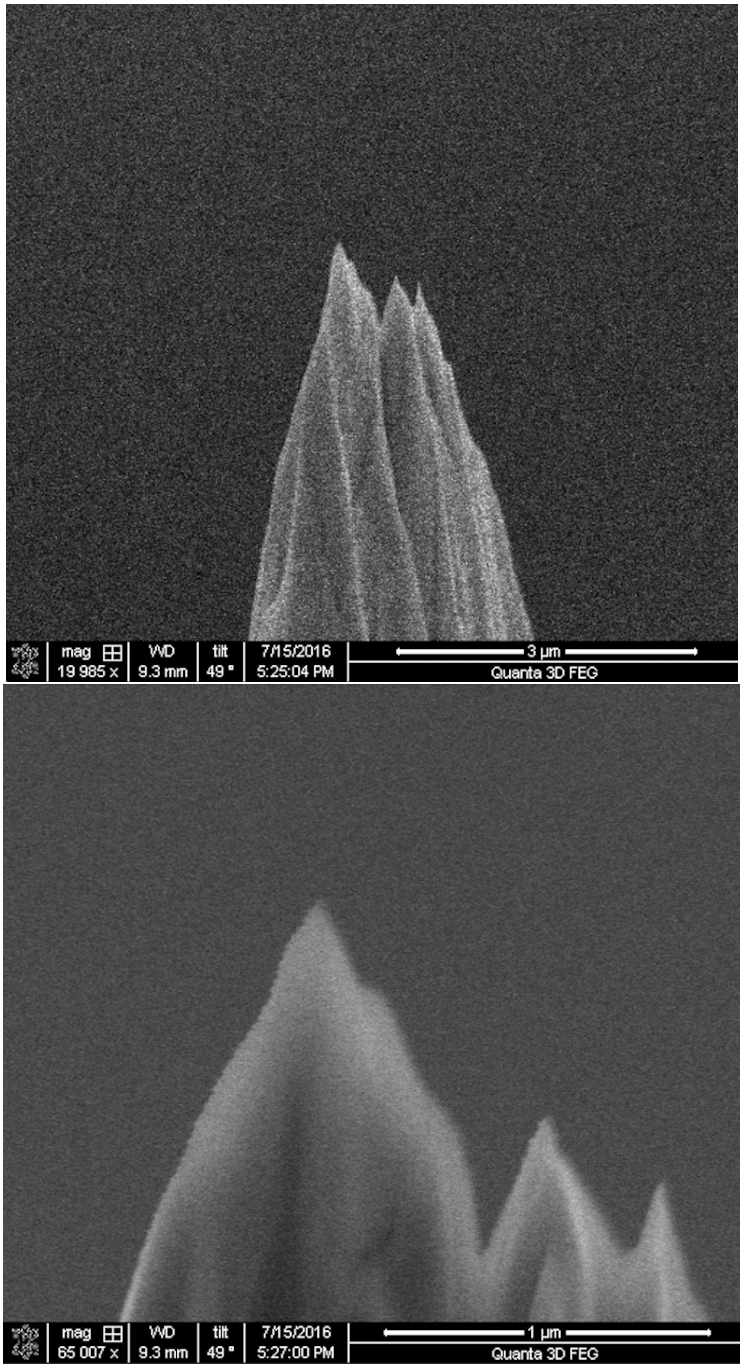

Figure A.1: Untreated tungsten probe. Scanning electron microscope images of tungsten wire probe prior to treatment. Analysis of the highest apex shows that the radius of curvature for the probe is on the order of $\sim$ 100nm.

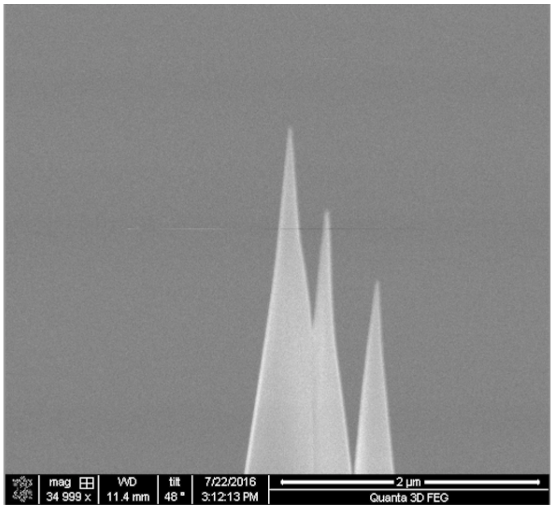

Figure A.2: FIB treated probe. Scanning electron microscope images of tungsten wire probe after FIB treatment showing a radius of curvature < 10nm.

The result of this milling procedure on the probes shown in Figure A.1 are shown in Figure A.2

Appendix B

Density Functional Theory for Modeling Diffusion

B.1 Density Functional Theory Principles

Density functional theory (DFT) is a modeling method for estimating the electronic structure of materials. The modeling framework rests on two fundamental theorems from Kohn and Hohenberg and the Kohn Sham equations which prove that the ground state energy of a many body system is a unique functional of the electron density and that the minimal energy functional corresponds to the ground state solution to the Schrodinger equation.(*109–111*) The electron density of a system is collectively defined in terms of single-electron wave functions $\psi_i(\mathbf{r})$, where the electron density is then.

$$n(\mathbf{r}) = 2\sum_i \psi_i^*(\mathbf{r})\psi_i(\mathbf{r}) \tag{B.1}$$

The energy functional of a system is written as:

$$E[\{\psi_i\}] = E_{\text{known}}[\{\psi_i\}] + E_{XC}[\{\psi_i\}] \tag{B.2}$$

where $E_{\text{known}}[\{\psi_i\}]$ includes electron kinetic energies and Coulomb interactions between electrons and nuclei, which mathematically is:

$$E_{\text{known}}[\{\psi_i\}] = \frac{h^2}{m}\sum_i \int \psi_i^* \nabla^2 \psi_i d^3r + \int V(\mathbf{r})n(\mathbf{r})d^3r + \frac{e^2}{2}\int\int \frac{n(\mathbf{r})n(\mathbf{r}')}{\mathbf{r}-\mathbf{r}'}d^3r d^3r' + E_{\text{ion}} \tag{B.3}$$

while $E_{XC}[\{\psi_i\}]$ encompasses all terms which cannot be written down analytically and is called the exchange-correlation energy. Consequently, the exact ground state energy functional is almost never found in practice and instead the variational principle is used to estimate the ground state energy while the exchange-correlation energy must be approximated. Kohn and Sham showed that solving for the ground state energy reduced to solving the Kohn-Sham equation, which is in terms of a single particle wavefunction instead of all of the particle wavefunctions:

$$\left[\frac{h^2}{2m}\nabla^2 + V(\mathbf{r}) + V_H(\mathbf{r}) + V_{XC}(\mathbf{r})\right]\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}) \tag{B.4}$$

where the electron-electron Coulomb repulsion is given by the Hartree potential, $V_H(\mathbf{r})$:

$$V_H(\mathbf{r}) = e^2\int\frac{n(\mathbf{r}')}{\mathbf{r}-\mathbf{r}'}d^3r' \tag{B.5}$$

and the exchange-correlation contributions are defined in terms of a functional derivative:

$$V_{XC}(\mathbf{r}) = \frac{\delta E_{XC}(\mathbf{r})}{\delta n(\mathbf{r})} \tag{B.6}$$

For crystalline systems, typically a plane wave basis is chosen for the trial wavefunctions since solutions must obey periodic boundary conditions. Since the Hartree potential requires knowing the electron density, the quantity of interest, this equation is frequently solved self-consistently in an iterative fashion, typically using the following algorithm.

1. Start with a trial electron density $n(\mathbf{r})$
2. Solve Kohn-Sham equations using $n(\mathbf{r})$ to find $\psi_i(\mathbf{r})$
3. Calculate density from Kohn-Sham single particle solutions $n_{KS}(\mathbf{r}) = 2\psi_i^*(\mathbf{r})\psi_i(\mathbf{r})$
4. Compare $n_{KS}(\mathbf{r})$ with $n(\mathbf{r})$. If they are different, update $n(\mathbf{r})$

B.2 Modeling Diffusion Processes

It is possible through DFT to calculate rates of dynamical processes such as diffusion or chemical reactions. One approach of such a calculation is finding the energy surface of an adsorbate on a substrate. To calculate such an energy surface, one might place fix the adsorbate somewhere in (x, y) (which is parallel to the substrate) while allowing the z coordinate of the adsorbate to be optimized along with all of the substrate atoms. An example energy contour for Ag adsorption on Cu(100) is shown in Figure B.1

Motion is typically dominated by the lowest energy pathway between local minima which generally pass through the high symmetry, low energy saddle points in the energy surface which are called transition states. Transition state theory posits that the transition rate from local minimum states ($A \to B$) is given by:

$$k_{A\to B} = \nu\exp\left(\frac{E_T - E_A}{k_BT}\right) \tag{B.7}$$

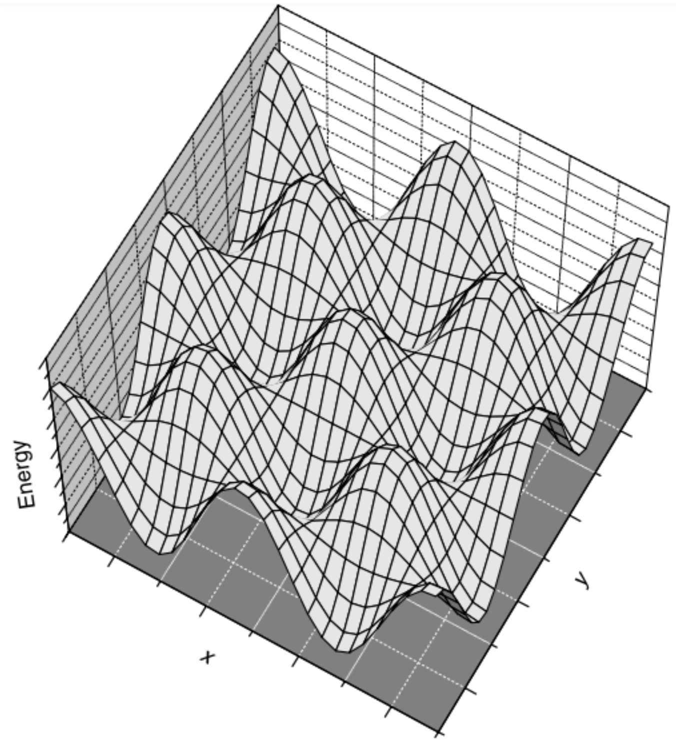

Figure B.1: Energy surface for Ag on Cu(100)(*111*)

where $k_{A \to B}$ is the transition rate, ν is the attempt frequency corresponding to the oscillation frequency of the adsorbate within the local minima, E_T and E_A are the transition and A state energies respectively. Harmonic transition state theory models the attempt frequency via the following:(*112*, *113*)

$$\nu = \frac{\prod_i^{3N} \nu_i^A}{\prod_i^{3N-1} \nu_i^T} \tag{B.8}$$

where ν_i^A and ν_i^T are the vibrational frequencies of A and the transition state respectively. Calculating the vibration frequency at a location can be as simple as doing a modeling the adsorbate as a simple harmonic oscillator in the neighborhood of any stationary point and taking numerical second derivative with respect to the position of the adsorbate using a finite difference method.(*111*) For example, in one dimension, for a single coordinate centered at 0:

$$\omega^2 = \frac{d^2E}{dx^2} \approx \frac{E(\Delta x) - 2E(0) + E(-\Delta x)}{\Delta x^2} \tag{B.9}$$

For more than 1 dimension, the Hessian would have to be estimated and diagonalized in order to compute all the normal modes and frequencies for the states of interest. There remains the issue of actually determining the transition states and reaction coordinates. One widely used technique is called nudged elastic band.(*114*) The exact description of this method is omitted, but at a high level, the technique uses some initial proposal for the pathway between two energy minima and iteratively minimizes energy of the path by using a force projection scheme such that forces from the potential energy act perpendicular to the path and spring forces work parallel to the path.