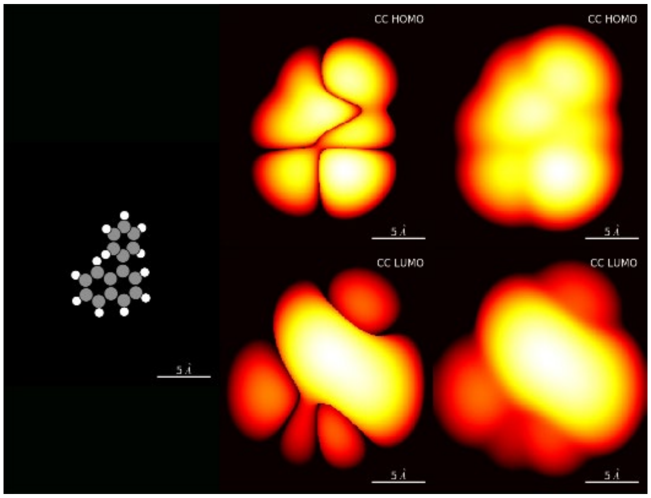

AN EXPERIMENTALIST'S GUIDE TO STM THEORY (AND THEORISTS)

Welcome to "An Experimentalist's Guide to STM Theory (and Theorists)." We made this guide to bridge the gap between experimentalists working with Scanning Tunneling Microscopy (STM) and the world of theoretical calculations. The guide is also an experiment in using a more informal approach to science communication --we hope you enjoy it!

Mads Engelund, Espeem

Introduction

Welcome to "An Experimentalist's Guide to STM Theory (and Theorists)." We made this guide to bridge the gap between experimentalists working with Scanning Tunneling Microscopy (STM) and the world of theoretical calculations. The guide is also an experiment in using a more informal approach to science communication --we hope you enjoy it!

This guide aims to provide experimentalists with a practical understanding of STM theory. It does not explain the STM simulation theories themselves but offers advice on their practical application. We'll delve into scenarios where straightforward calculations are helpful, such as using metal tips and samples, maintaining sharp and rigid apexes, and working with flat molecules.

However, even the most straightforward systems can present challenges. For instance, basis sets with cut-offs or insufficient vacuum are common pitfalls that can lead to misleading results. We will discuss these issues in detail, providing examples and solutions to help you avoid the most frequent mistakes encountered in STM calculations.

As your experiments evolve, so do the theoretical models required to interpret them. Complex situations involving specific tip apex configurations, semiconductor samples, or dynamic structure switching often call for the expertise of a theorist. Finally, this guide offers practical advice on understanding a theorist's perspective and fostering efficient communication.

Note on Language

For simplicity, in this guide, we will assume that the system under investigation is a molecule on a surface. There are naturally other systems, such as adatoms, surface defects, and clusters, which can be thought of as molecules that interact strongly with the surface.

Chapter 1: "Standard" STM Calculations

In this chapter, we examine what we will call "standard" STM simulations—some simple techniques that are widely used.

The Tersoff-Hamann Approximation

When you investigate STM simulations, you will quickly realize that one dominant approximation is ubiquitous—the Tersoff-Hamann (TH) approximation, which was discovered soon after STM. This approximation is almost synonymous with "STM simulation," it is the base of the STM simulations in most atomistic simulation packages, such as Siesta and QuantumEspresso.

The Tersoff-Hamann approximation is a collection of three assumptions:

1. The tip density of states is constant.
2. The tip frontier orbital is spherical.
3. The tip-induced electrostatic field can be ignored.

With these three assumptions, Tersoff and Hamann showed that the current flowing between the sample and tip would be proportional to the sample Local Density Of States (LDOS) integrated over the bias window at the center of the (spherical) tip frontier orbital.

The best case for the TH approximation is when both the tip and sample are metallic, and the tip apex is an add-atom or small metallic cluster. The structure under investigation should also be in good electronic contact with the sample. Some examples would be small molecules, flat molecules, or metallic clusters.

TH can still be surprisingly helpful even when the tip and sample don't fulfill all conditions. Even when the tip DOS is not constant, if a molecular cluster is absorbed, then sharp peaks in the LDOS spectrum will still occur in the same place. If the tip frontier orbital is not precisely spherical, it will still often be laterally symmetric, which is enough to reproduce nodal features in the electron density of the sample. If your sample is semi-conducting, then some voltage drop in the system would occur in the sample, giving the simple effect of having a smaller effective voltage over the structure of interest.

Moving beyond the TH approximation to general Bardeen Theory immediately becomes a lot more cumbersome, first, since Bardeen Theory is much less readily available in familiar toolkits, and secondly, because you need an explicit model for the tip when using this level of theory. If your experimental setup does not have fine-grained control over the tip, the TH approximation is excellent because it has the most non-specific assumptions about the tip structure.

The reasons that the TH approximation is so dominant are:

1. It is easy to apply - the toolkit you are using will often be able to calculate LDOS anyway.
2. It gives good results in most cases - even when some of the assumptions break.
3. Going beyond the TH approximation requires a lot of careful theoretical work.
4. Going beyond the TH approximation typically only makes sense when you achieve fine-grained control over your tip apex.

Getting started with these calculations is relatively easy. If starting from scratch, you can use QuantumEspresso and the post-processing tool included in the QuantumEspresso package.

Gas-Phase Simulations

You can often ignore the substrate when simulating STM on molecules - this is called "gas-phase simulations." By simulating only the molecule, you can simplify the system and focus on intrinsic electronic properties without the complicating effects of substrate interactions. Simulating only the molecules is also much less expensive in computational time.

To use gas-phase simulations correctly, you should be aware of when they fail to give sensible results:

- **Strong Interaction:** If the sample-molecule interaction is strong enough to form bonds between the two, gas-phase simulations will be wholly inappropriate. The molecule will have the wrong configuration, and the bond with the surface will also affect the molecular orbitals.STR
- **Flexible Molecule:** Even if the sample-molecule interactions are weak, a problem might still exist if the molecule is very flexible. A flexible aromatic or linear carbon-chain molecule will be deformed even by weak dispersive forces.
- **Molecular Orientation:** With no surface in the simulation, the orientation you choose for the molecule is entirely based on your intuition. There is no issue when the molecule is flat, but the molecular orientation is not straightforward for non-planar molecules, such as sterically hindered polycarbonates.
- **Dominant Substrate:** In the previous points, we assumed that the molecule would dominate the image, but that is not always true, and when it does not, then gas-phase simulations are inappropriate. Dangling-bond defects on the Si(001):H substrate, for instance, have very localized and directional electronic surface states, which can influence the imaging of flat molecules on top.

Awareness of the limitations is the key to confidence in using the gas-phase approximation when appropriate. The gas-phase simulations should be your first try when molecules interact weakly with the substrate. And even when these molecule-only simulations are insufficient, the fact that they fail will give you information about the molecule and substrate. If adequate, you will save a lot of time compared to setting up and running simulations that include the substrate.

Conclusion on "Standard" STM Calculations

The Tersoff-Hamann (TH) approximation is a foundational tool for STM simulations, favored for its simplicity and effectiveness in most scenarios. It's well-integrated into popular simulation packages like Siesta and QuantumEspresso and can yield valuable insights even when ideal conditions are unmet. While moving beyond TH to theories like Bardeen can provide more precision, it comes with increased complexity and the need for detailed tip modeling.

Gas-phase simulations simplify STM studies by focusing on the molecule alone, which is computationally efficient and effective for weakly interacting systems. However, they have limitations, particularly with strong substrate interactions or flexible molecules.

By starting with TH and gas-phase simulations, you can quickly gain insights and identify when more detailed approaches are necessary, balancing ease of use with the need for accuracy.

Chapter2: Pitfalls in "Standard" STM Calculations

While the Tersoff-Hamann (TH) approximation offers a robust and straightforward method for STM simulations, several common pitfalls can influence the accuracy and reliability of your results. Errors arising from these pitfalls are common, and we have observed them all in peer-reviewed literature.

Insufficient Tip Distance

A widespread problem for simulations in the scientific literature is that the tip is too close to the sample in simulation. Actual experiments typically have at least a 5-7 Angstrom distance from the sample, but some simulations put the tip as close as 1 Angstrom. The problem exists for both atomic-centered orbitals and plane-wave codes, but for different reasons:

- **Atomic-Centered Orbitals**: In simulations using atomic-centered orbitals, the orbitals often do not accurately describe the electron density far from the atoms. This limited range can lead to inaccuracies in the simulated STM images, especially at large tip-sample distances.
- **Plane-Wave Codes**: Using sufficient vacuum can be computationally expensive for plane-wave codes. Insufficient vacuum regions can lead to artificial interactions with periodic images, skewing the results.

The problem is that when the tip distance is too small, the STM image might have the correct nodal planes, but the lateral size of the image will be way too small.

One solution is to use a plane-wave code and include enough vacuum (20-30 Angstrom) despite the high computational cost. Beyond the computational cost, having enough computer memory may also be a concern. Something that makes the solution more manageable is that you do not need to relax the molecule with this large vacuum—it is sufficient to perform a one-shot simulation with an additional vacuum after the relaxation.

The fix for the atomic-centered orbitals is more complex, and you would likely need to use established post-processing tools. We have encountered two tools that address this problem.

1. Extend the atomic orbitals as is done in the Probe Particle project.
2. Evaluate wave functions close to the surface and extend them into the vacuum using the Paz/Soler method, as we do in the CalcTroll project. We prefer this method because it can also be used on plane-wave codes, giving us flexibility in code use.

The simplest way to circumvent the problem is to use a plane-wave code and a substantial vacuum but to avoid the high computational cost; you should look into either the Probe-Particle project or the CalcTroll project.

Incorrect Application of Lateral Tip Uncertainty

Applying spatial broadening to the simulated images can be crucial when comparing experimental and simulated constant-current images. The spatial broadening arises in experiments from a combination of two effects:

1. Lateral drift in the tip position.
2. Contributions to tunneling from secondary tip atoms.

If one atom dominates the tunneling from the tip apex, you can effectively model these two effects as a Gaussian broadening.

There is some difficulty in applying the broadening correctly. You cannot use the included STM post-processing tools in your ab initio toolkit for this since they typically only produce the zero-noise image. One way to deal with this discrepancy between the simulated and experimental is to ignore it. In scientific literature, you often see simulated images with no broadening next to experimental images. The authors then leave it up to the reader's imagination to apply the broadening to the simulated image. While this approach can be sufficient in simple cases, it tends to fail when molecules are non-planar or become more complex.

If you find ignoring the broadening problem unsatisfactory, we couldn't agree more! However, when applying Gaussian broadening, there is a common pitfall to be aware of. To see this problem, let us take a look at how you would construct a constant-current image in a two-step process:

1. Construct the 3D map of the current given the tip position (in TH given by the LDOS in the bias window).
2. Find an iso-current surface and map the height to a 2D image.

The pitfall is that you might be tempted to apply the Gaussian broadening to the final image. To see why this is incorrect, consider the experiment we are simulating. In the experiment, you record the current at a position with a lateral uncertainty and then adjust the tip height until the current is correct. To model this, we must apply the lateral uncertainty to the current 3D map before finding the iso-surface.

Finding a way to use this approach within the common ab initio toolkits is not that simple since the STM tools native to the package typically do not have a broadening function. The only way to deal with this problem is to construct your STM utility. Output a 3D map of the LDOS in the bias window you are interested in, apply the 2D Gaussian kernel, and then build the iso-current surface.

The only post-processing tool we know that can handle this easily is the one we made in Espeem with the CalcTroll package, which can post-process SIESTA and QuantumEspresso calculations.

DFT Band Gap Problem

Another significant issue that experimentalists encounter when delving into STM simulations is the Density Functional Theory (DFT) gap problem. The DFT gap problem refers to the inaccurate prediction of the energy band gap by standard DFT calculations. The DFT gap problem arises due to the inherent nature of DFT, which tends to underestimate the band gap of insulating or semiconducting materials. This underestimation occurs due to the self-interaction error in standard DFT functionals, leading to a systematic error in predicting the electronic band gap.

Here are a few strategies you might apply to the problem:

1. **Work around it:** Although the energy an electronic state occurs at is not accurate in DFT, the wavefunction and the order that states appear in is. So, you can compare the experimental STM image at the voltage where the first peak occurs in STS with the simulated image of the first unoccupied state in the calculation.
2. **Better Theory:** Some flavors of DFT are substantially more accurate than standard DFT when it comes to band gap prediction. For example, hybrid DFT functionals, such as the HSE functional, incorporate a portion of exact exchange from Hartree-Fock theory, improving band gap calculation accuracy. Additionally, many-body perturbation theory methods, such as the GW approximation, can provide highly accurate band gap predictions, but they are computationally more demanding.
3. **Empirical Corrections:** Another strategy is to employ empirical corrections based on experimental data. This strategy involves fitting a correction function to the calculated band gap using a set of known experimental values. You can then apply this correction function to the DFT band gap for a more accurate estimate. However, this approach relies on reliable experimental data and must be fitted for both sample and molecule.

For an experimentalist, we recommend working around the issue. Navigating functional and empirical corrections is complicated and best left to the professionals. Practically, you will often get excellent agreement between the simulated and experimental images if you use the first voltage that gives a signal in both theory and experiment.

Conclusion on Pitfalls

In summary, while the Tersoff-Hamann (TH) approximation is a powerful and widely used method for STM simulations, several common pitfalls can compromise the accuracy and reliability of the results. Ensuring that the tip distance is sufficient is crucial, as insufficient tip distance can lead to misleading STM images, regardless of whether you use atomic-centered orbitals or plane-wave codes. Advanced post-processing tools, such as those in the Probe Particle or the CalcTroll project, can mitigate these inaccuracies for atomic-centered orbitals. Despite the computational cost, incorporating a substantial vacuum region is essential when using plane-wave codes.

Another area where caution is needed is the application of lateral tip uncertainty. Correctly modeling the experimental conditions requires applying Gaussian broadening to the current 3D map before constructing the iso-current surface. Ignoring this step can lead to discrepancies between experimental and simulated images, especially for complex or non-planar molecules.

Developing or utilizing post-processing tools that can handle this step, such as the CalcTroll package, is vital for accurate STM simulations.

Finally, the DFT band gap problem remains a significant challenge in STM simulations. While various strategies exist to address this issue, including using hybrid functionals and many-body perturbation theory, these methods can be complex and computationally intensive. For experimentalists, a practical workaround is to compare the experimental STM image at the voltage where the first peak appears with the simulated image of the first unoccupied state. This approach often yields satisfactory agreement between theory and experiment without needing advanced corrections.

By being aware of these pitfalls and implementing the recommended solutions, researchers can enhance the fidelity of their STM simulations, leading to more accurate and reliable interpretations of experimental data.

Chapter 3: When Things Get Difficult

As the complexity of your experimental setup or the observed phenomena increases, so does the need for more specialized theoretical tools and expertise. This chapter delves into those situations where the standard approximations fall short, and collaboration with a theorist becomes essential for accurate interpretation and understanding. We'll discuss the signs that indicate the need for theoretical support and provide practical guidance on fostering effective collaboration to maximize the impact of your research.

Identifying When to Call in a Theorist

As an experimentalist, recognizing the limits of standard STM simulation techniques and knowing when to seek the expertise of a theorist can be crucial for your research. While the methods discussed in previous chapters provide a solid foundation, specific experimental observations may necessitate a deeper theoretical understanding. Here are some key indicators that your experiment may warrant the involvement of a theorist:

1. **Non-Standard Tip Structures:** If your STM setup utilizes a tip with a well-defined and complex apex structure, such as a functionalized tip or a specific cluster arrangement, the Tersoff-Hamann approximation might not be sufficient.
2. **Semiconductor Substrates:** When working with semiconducting samples, the DFT band gap problem becomes more pronounced, and simple workarounds may not suffice.
3. **Small Tip-Sample Distances or High Field:** If your experiment requires moving beyond the standard Tersoff-Hamann approximation due to factors like strong tip-sample interactions or significant tip-induced electric fields, theoretical expertise becomes essential.
4. **Dynamic Features in STM Images:** If your STM images exhibit unexpected lines, streaking, or other dynamic features that suggest molecular movement or conformational changes, a theorist can help interpret these observations by modeling the underlying dynamics.

5. **Interpreting Complex or Unexpected Results:** When confronted with STM images that contain intricate patterns, unusual contrasts, or other features that defy simple explanations, a theorist can help unravel the underlying physics.

If you know when it is needed, you can engage a theorist early on and ensure that your experimental results are interpreted accurately and comprehensively.

Practical Tips for Working with a Theorist

Collaborating with a theorist can be rewarding but requires effective communication and a mutual understanding of expertise and limitations. Here are some practical tips to foster a productive working relationship:

1. **Communicate Clearly and Openly:** Clearly articulate your experimental goals, methods, and observations to the theorist. Share relevant data, images, and any specific questions you have. Be open to discussing the limitations of your experimental setup and potential sources of error.
2. **Speak the Same Language:** While you don't need to become a theoretical expert, familiarize yourself with the basic terminology and concepts relevant to your experiment (this guide is a good start).
3. **Set Realistic Expectations:** Remember that supporting your experiment is likely not their primary research focus, so managing expectations regarding time commitment is critical**.**
4. **Be Patient and Flexible:** Theoretical work often involves trial and error; finding suitable models or parameters can take time. Be patient with the process and be open to adapting your experimental plans based on the insights gained from the simulations.
5. **Give Credit Where It's Due:** Acknowledge the theorist's contribution appropriately when publishing your results. Remember that detailed theory-experiment collaborations may naturally lead to publications in specialized journals, where the theorist is naturally a co-author. But if you subsequently create an article highlighting the key findings targeting a high-impact journal, consider if the theorist should also included as a co-author here.
6. **Think Long-Term:** Fostering a long-term collaborative relationship can be mutually beneficial, allowing for deeper understanding and more impactful research outcomes.

By following these tips, you can build a solid and fruitful collaboration with a theorist, leveraging their expertise to enhance your research.

Summary: When to Call a Theorist

While "standard" STM calculations offer valuable insights into many experimental scenarios, recognizing when to seek collaboration with a theorist is crucial for advancing your research. Non-standard tip structures, semiconductor substrates, small tip-sample distances, dynamic features, or complex results indicate that specialized theoretical expertise is needed.

You can cultivate a productive and mutually beneficial working relationship by communicating, speaking the same language, setting realistic expectations, being patient and flexible, and acknowledging the theorist's contributions.

The End of the Guide

After having read it, you should now be able to:

- Understand the idea behind “standard” STM simulations.
- Know about some pitfalls when applying these simple simulations.
- Recognize indicators that you may need a theorist.
- Engage with a theorist in a way that fosters an excellent long-term relationship.

We hope this guide has provided a helpful foundation for your endeavors in bridging the gap between experimentalists and theorists. Keep experimenting, keep exploring, and keep pushing the boundaries of scientific understanding.