RICE UNIVERSITY

RF Shot Noise Measurements in Au Atomic-scale Junctions

by

Ruoyu Chen

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE
Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

[Signatures]

Douglas Natelson, Chair
Professor of Physics and Astronomy

Ruirui Du
Professor of Physics and Astronomy

Kevin Kelly
Associate Professor of Electrical and Computer Engineering

Houston, Texas

January, 2016
ABSTRACT

RF Shot Noise Measurements in Au Atomic-scale Junctions

by

Ruoyu Chen

Conduction electrons are responsible for many physical or chemical phenomena in condensed matter systems, and their behavior can be directly studied by electronic transport measurements. In conventional transport measurements, conductance or resistance is usually the focus. Such a measurement can be as simple as a quick two-terminal DC check by a multi-meter, or a more sophisticated lock-in measurement of multiple higher harmonic signals synchronized to different frequencies. Conductance carries direct information about the quasi-particle density of states and the local electronic distributions, which are usually Fermi-Dirac distribution. Conductance is modified or dominated by scattering from defects or interfaces, and could also reflect the spin-spin exchange interactions or inelastic couplings with phonons and photons.

Naturally one can ask the question: is there anything else we can measure electronically, which carries extra information that a conductance measurement does not provide? One answer to this question is the electronic noise. While the conductance reflects the average charge conduction ability of a system, noise describes how the physical quantities fluctuate around their average values. Some of the fluctuations carry information about their physical origins. This thesis will focus on one particular type of the electronic noise shot noise, but other types of noise will also be introduced and discussed.
We choose to measure the radio frequency component of shot noise, combining with a modulated lock-in detection technique, which provides a method to largely get rid of other unwanted low-frequency noise signals. Au atomic-scale junctions are the systems we studied here. Au is relatively well understood and will not generate too many complications, so it’s ideal as the first platform for us to understand both shot noise itself and our RF technique. On the other hand, the atomic scale raises fundamental questions about electronic transport and local energy exchange and dissipation, which make our measurements fundamentally interesting.

We employed two different types of mechanical controlled Au break junctions: the Scanning Tunneling Microscope (STM)-style Au break junctions, and the mechanically-bending Au break junctions. We studied shot noise behaviors of individual configurations or ensemble averages over all the accessible configurations. Measurements were conducted at both room temperature and liquid He temperature. High quality shot noise measurements were demonstrated. New phenomena like anomalous excess noise enhancement at high bias voltages and non-zero shot noise variance below 1\(G_0\) were seen. We also found shot noise to be surprisingly insensitive to temperatures between 4.2K and 100K, and can be well described by the non-interacting approximation.
Six years is never short, and a Ph.D degree in a completely different country is no easy. In the past years I regret that I haven’t worked hard or smart enough in all my time. But I also truly know that I am fortunate enough to meet many nice people, receiving their helps and learning a lot from them. Without them I can not go that far.

Firstly I would like to thank my advisor Dr. Doug Natelson, who is the most ideal advisor I can imagine. He owns broad knowledge and sharp intuitions. He is patient and thoughtful. He pushes me forward but also gives me space and freedom to improve myself and conduct research at my pace. I particularly want to thank him for the private discussions and lessons lasting one semester in my third year, starting from which I learnt the way to consider real physical questions but not only know how to get an “A” in exams. I also greatly appreciate his strong support and recommendations to help me finding a postdoc position.

I also want to thank the other committee members: Dr. Ruirui Du, Dr. Patrick Hartigan and Dr. Kevin Kelly. Dr. Du has been in my committee for the whole six years. I am very grateful to his kindness and support for that many years. Dr. Hartigan used to be in my master defense committee. I particularly appreciate him that he did not let me easily pass my Ph.D oral qualification exam, which made me realize my problem and eventually I was able to overcome it. I also thank Dr. Kelly to join my Ph.D committee and witness my defense.

I thank all my great labmates, both past and present. They are all great friends and I keep learning from all of them. I particularly want to thank Dr. Jiang Wei, who is the second advisor in reality and also my best friend outside the lab. He teaches me
a lot about physics and techniques and encourages me to continue scientific research after my graduation. Will Hardy is the second person I must thank a lot, who nicely trained me the use of nearly every equipment and the discussions with him are always fruitful. Dr. Pavlo Zolotavin is the person I can really "fight" against, using math and physics. I enjoy debating with him and I can never overestimate how beneficial those discussions are. Dr. Patrick Wheeler is the person led me into the lab and taught me the basics. Dr. Gavin Scott and Dr. Joseph Herzog are always friendly and helpful. Dr. Heng Ji and Yajing Li are the long-term good friends and also helped me a lot. Benjamin Huber, Charlotte Evans, Loah Stevens and Panpan Zhou are young and energetic and bring laughs into the lab.

At last I want to thank my family. This six years is also not easy for my parents. But they never stop encouraging me moving forward to pursue my own goal. Without their support I can not be here. The biggest thank will be given to my wife Yongrong. She brings happiness and comfort to our home and my life becomes much easier and more organized. Without her understanding and encouragement, I purely can not keep concentrated on science in the past and can not make the decision (try) to continue in the future.
# Contents

Abstract ................................................................. ii
Acknowledgement ......................................................... iv
List of Illustrations ....................................................... ix

1 Introduction to quantum transport .................................... 1
   1.1 Length Scales and Transport: Tunneling, Point Contacts and
       Diffusive wires .................................................... 2
   1.2 Conductance Quantization or not? ............................... 8
   1.3 Quantum coherent transport ..................................... 10
   1.4 Electron-phonon Interactions ................................... 16
   1.5 Local heating ................................................... 20

2 Shot noise ............................................................... 27
   2.1 Basics of noise .................................................. 27
   2.2 Common Electronic Noise Types ................................. 30
   2.3 Shot Noise ...................................................... 35
       2.3.1 Non-interacting Shot noise: single barrier ............... 36
       2.3.2 Non-interacting Shot noise: beyond single barrier ...... 39
       2.3.3 Single-particle shot noise with interactions ............. 42
       2.3.4 Shot noise beyond single-particle limit ................. 46

3 Techniques .............................................................. 52
   3.1 Shot Noise Measurement Technique .............................. 52
       3.1.1 Common DC Measurement Techniques ..................... 53
### 4 Shot noise in STM-style Au break junctions at room temperature

4.1 Motivation and technical details ........................................... 72
4.2 Ensemble average at moderate biases .................................... 75
  4.2.1 Conductance histogram and shot noise suppression ............ 75
  4.2.2 Linear bias dependence of shot noise ............................... 77
  4.2.3 Non-linear bias dependence of shot noise ......................... 81
  4.2.4 Toy model calculation ................................................. 83
4.3 Enhanced ensemble average at high biases .............................. 85
  4.3.1 Possibilities to be excluded: Flicker noise, reservoir heating
  and electron-phonon interactions ......................................... 87
  4.3.2 Possible explanation: Local electronic heating .................. 92
4.4 Beyond ensemble average .................................................. 98
  4.4.1 Generation of 2D density plots .................................... 99
  4.4.2 Simulation and discussion ........................................... 105

### 5 Shot noise in Conventional Au MCBJs at multiple cryogenic temperature

5.1 Device fabrications and experimental methods .......................... 115
5.2 Noise on metal stripes ..................................................... 120
5.3 Bias dependence of shot noise at multiple temperatures ............ 123
5.4 Non-linear bias dependence close to $1G_0$ ............................ 129
5.5 Conclusion ...................................................................... 133
Bibliography 135
## Illustrations

<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Dimensionality and different transport regions</td>
<td>4</td>
</tr>
<tr>
<td>1.2</td>
<td>Conductance quantization study</td>
<td>9</td>
</tr>
<tr>
<td>1.3</td>
<td>Coherent transport</td>
<td>12</td>
</tr>
<tr>
<td>1.4</td>
<td>Weak localization</td>
<td>14</td>
</tr>
<tr>
<td>1.5</td>
<td>Universal conductance fluctuation</td>
<td>15</td>
</tr>
<tr>
<td>1.6</td>
<td>Electron-phonon interaction</td>
<td>18</td>
</tr>
<tr>
<td>1.7</td>
<td>Physical picture of local heating</td>
<td>21</td>
</tr>
<tr>
<td>1.8</td>
<td>Local ionic and electronic heating</td>
<td>25</td>
</tr>
<tr>
<td>2.1</td>
<td>Thermal noise of a complex impedance</td>
<td>31</td>
</tr>
<tr>
<td>2.2</td>
<td>1/f noise</td>
<td>34</td>
</tr>
<tr>
<td>2.3</td>
<td>Non-interacting ballistic shot noise</td>
<td>39</td>
</tr>
<tr>
<td>2.4</td>
<td>Inelastic effects to shot noise</td>
<td>44</td>
</tr>
<tr>
<td>2.5</td>
<td>Shot noise in correlated systems</td>
<td>49</td>
</tr>
<tr>
<td>3.1</td>
<td>DC current fluctuation detection technique</td>
<td>54</td>
</tr>
<tr>
<td>3.2</td>
<td>RF excess noise measurement circuit</td>
<td>56</td>
</tr>
<tr>
<td>3.3</td>
<td>Equivalent circuit</td>
<td>58</td>
</tr>
<tr>
<td>3.4</td>
<td>Gain-bandwidth product measurement</td>
<td>61</td>
</tr>
<tr>
<td>3.5</td>
<td>Reflection measurement</td>
<td>62</td>
</tr>
<tr>
<td>3.6</td>
<td>Reflection vs structure area</td>
<td>63</td>
</tr>
<tr>
<td>3.7</td>
<td>Break junctions</td>
<td>68</td>
</tr>
</tbody>
</table>
Chapter 1

Introduction to quantum transport

Shot noise, after its first discovery by Schottky in 1918[1], has been demonstrated to carry extra transport information inaccessible by purely measuring conductance[2]. This makes shot noise fundamentally interesting and attracts a reasonable amount of theoretical attention. On the other hand, the fact that strong electron-phonon interactions in the bulk completely suppress shot noise[3, 4] limits the size of any sample to be tested and results in challenges of fabrication. The very weak signal strength and calibration requirements further make shot noise’s detection nontrivial. A relatively small amount of experimental efforts have been invested into this field compare to transport. Much success has still been achieved though, in the past twenty years, in many different systems such as atomic-scaled junctions[5, 6], 2d electron gas (2DEG)[7, 8], Normal-Superconductor interfaces[9, 10] and fractional quantum hall states[11, 12, 13]. Different measurement techniques exist, most of which are in the low frequency regime. They suffer from 1/f noise or other low frequency noise sources, and have relatively low measurement rates. RF measurement techniques avoid those problems and should be considered as an alternative method, though these also have their own drawbacks.

Followed by earlier pioneers[7, 14], we have also built up our RF measurement technique and chose atomic-scaled Au junctions as the first platform to study and test this technique. We systemetically studied shot noise at different types of Au break junctions, varying conductance, bias or temperature. We also tried different ways to
analyze the data. Our RF technique is demonstrated to be successful and many new discoveries are made, though shot noise measurements in similar Au junctions in DC limit are not new.

This thesis will focus on all our measurement efforts at Au atomic junctions. The physical background will be introduced in the first two chapters, while the related techniques will be presented in the third. Chapter 4 and 5 are our measurement results realizing by two different mechanical break junction techniques: the STM-style break junction technique and the conventional mechanical bending break junction technique. In this chapter, I will focus on the general mesoscopic transport nature and discuss what’s special at atomic scale. Shot noise will be introduced in Chapter 2.

1.1 Length Scales and Transport: Tunneling, Point Contacts and Diffusive wires

In a probably over-simplified language, macroscopic electronic transport in a single crystal material is well described by its band structure. With its Fermi level located in the conduction band, the material is a metal; Otherwise, it is a semiconductor or insulator. Physics becomes more complicated when the size is reduced and approaches the mesoscopic limit[15]. In mesoscopic metallic systems, length scale matters for at least three aspects. Length scales determine the dimensionality, the importance of different scattering mechanisms, and the importance of quantum corrections[16, 17] to quantized conduction.

The Fermi wavelength $\lambda_F$ sets the dimensionality of the electronic structure. Macroscopic band structures assumes a periodic potential and a infinitely large system size, resulting in quasi-continuous bands. With the size largely reduced in one
direction, the infinite periodicity condition does not hold in that direction and the boundary conditions modify the bulk band structure. As a result, along the restricted direction, the allowed momentum and energy values are limited and electronic gaps tend to be opened. When the size is comparable to $\lambda_F$, the full quantum limit is reached where the energy bands are readily split into transverse modes and only one or a few subbands are occupied. Thus electrically the sample’s dimension is reduced by one. When more dimensions have sizes comparable to $\lambda_F$, the sample has more reduced dimensions. In two and one dimensions, the bulk bands split into subbands. And in zero dimension, only discrete energy levels exist. The density of states DOS(E) is changed from proportional to $\sqrt{E}$ in 3D to constant in 2D, inversely proportional to $\sqrt{E}$ in 1D and $\delta$ functions in 0D, which is shown in Fig. 1.1(a). In metals the high carrier density means very small $\lambda_F$; hence atomic scale is required to enter this full quantum limit. In the GaAs/AlGaAs heterostructures, the 2DEG carrier density is low and its $\lambda_F$ is several tens of nms, accessible by the nano fabrication technique even twenty years ago.

The elastic mean free path $l_e$ is another important length scale. With the device size smaller than $l_e$ the electrons transport ballistically, meaning that no defect scattering is expected (but boundary scattering is still allowed). In the limit that the size $L \gg l_e$, many scattering events occur and electrons transport diffusively. Figure 1.1(b) is a schematic illustration in 1D case, which is close to our device geometry. Note the concepts of ballistic or diffusive transport do not require 1D. An intermediate regime called the quasi-ballistic regime is also defined in Figure 1.1(b), where the constriction width W is smaller than $l_e$ but length L is allowed to be larger. The (quasi)1D constriction in the ballistic or quasi-ballistic regime is the so-called quantum point contact. The transport natures are dramatically different between ballistic
and diffusive systems. Many other more specific length scales exist and varying the
sample size between them can also selectively turn on and off some particular interac-
tions. However, none of those length scales can have such effects and we will discuss
them in the later parts of this thesis.

Figure 1.1 : a) From left to right: DOS(E) of 1D, 2D, 3D systems. Adapted from
[18]. b) From top to bottom: diffusive conductors, quasi-ballistic contacts, ballistic
contacts. Adapted from [16]

“Diffusive” is not equivalent to “classical”, however, as elastic scatterings do not
destroy phase coherence. A third important length scale coherence length $l_\phi$ can
be defined, meaning that the quantum mechanical phase information is maintained
when electrons transport over shorter distances than this $l_\phi$. As the main mechanism
destroying the phase is inelastic scattering, it is usually not bad to regard \( l_\phi \) as close to the inelastic mean free path \( l_{in} \). Thus a diffusive conductor with size smaller than the \( l_\phi \) still keeps its phase information and can be regarded as mesoscopic. If the sample size exceeds \( l_\phi \) at any directions, then the system is macroscopic (but it still does not mean quantum mechanics play no role, as clearly the bulk band structure is also a result of quantum mechanics).

The diffusive nature makes Ohm’s law hold, meaning that from the conductance \( G \), a conductivity can be defined as

\[
\sigma = (L/W)G \tag{1.1}
\]

The semi-classical Drude model can qualitatively describe the physics well, where the current induced by electric field needs to be balanced with the diffusion current in thermal equilibrium, thus the carrier density gradient reaches its equilibrated value. The result is the famous Einstein relation for degenerate systems,

\[
\sigma = e^2 \rho(E_F)D \tag{1.2}
\]

Here \( \rho(E) \) is the electronic density of states, and \( D \) is the diffusion constant. With a finite bias applied to the conductor, the same math yields [16]

\[
j = -\nabla E_F e^{-2} \int_0^\infty dE \sigma(E, 0) \frac{df}{dE_F} \tag{1.3}
\]

The bias is included in \( \nabla E_F \). This result means that no matter how big the bias is applied to the conductor, only the density of state within \( k_B T \) energy range around the Fermi energy matters and a linear I-V curvature is expected. Physically it means that within a distance of \( l_{in} \) the potential gradient built up by external bias is negligible, where electrons lose their original quantum numbers and are relaxed to establish local thermal equilibrium.
The ballistic transport is completely different, where the lack of scatterings makes the constriction length $L$ irrelevant to the conduction ability and the concept of conductivity loses its physical meaning. The classical consideration gives the Maxwell conductance\cite{19}.

$$G_M = \frac{2a}{\rho}$$

(1.4)

where $a$ is the orifice radius and $\rho$ is the resistivity. This is clearly not a real ballistic result as the resistivity still plays a role here, but the length $L$ has already disappeared from the formula and only the cross section matters. The semiclassical approximation was given by Sharvin\cite{20}.

$$G_S = \frac{2e^2}{h} \left( \frac{k_Fa}{2} \right)^2$$

(1.5)

In large point contacts ($a \gg \lambda_F$) the semiclassical solution already captures the physical natures, and Wexler\cite{21, 22, 23} gives the expression of its resistance $R = R_S + \gamma(K)R_M$, with $\gamma$ a slowly varying parameter with its value close to 0.69 and $K = l/a$ is the Knudsen number. Approaching the full quantum limit ($a \sim \lambda_F$), Landauer formula gives a more complete quantum-mechanical treatment:

$$G = \frac{2e^2}{h} \sum \tau_i$$

(1.6)

$0 \leq \tau_i \leq 1$ represents the transmittance of the $i$’th eigenchannel (transverse mode). Each eigenchannel at most contribute $\frac{2e^2}{h} \equiv 1 G_0$ conductance. In atomic-scaled junctions, the eigenchannel originates from valence electronic orbitals of the atoms\cite{24}. Though the transport mechanism is different, tunneling can also be described in the same language, without the summation over quantum channels and with the transmittance $\tau \to 0$.

This formula has two counter-intuitive features. Firstly, it indicates that the exact energy dispersion relation of each eigenchannel does not affect its conduction
ability. Secondly, even a perfectly-transmitted channel still has finite conductance. The latter can be understood as a contact resistance, which is caused by the channel mismatch between the constriction and the leads\cite{25}. The first can also be understood through a few lines’ algebra. With a finite bias applied between source and drain, the corresponding Fermi levels are differed by $eV$. The total current is simply a summation over all the available quantum states in the center ballistic constriction. In single mode case with $\tau = 1$,

$$I = \frac{e}{L} \sum_k v_k (f_L(\epsilon_k) - f_R(\epsilon_k)) = \frac{e}{L} \sum_k \frac{1}{\hbar} \frac{\partial E}{\partial k} (f_L(\epsilon_k) - f_R(\epsilon_k))$$ (1.7)

with $\sum_k \rightarrow 2 \times \frac{L}{2\pi} \int_{-\infty}^{+\infty} dk$,

$$I = \frac{2e}{\hbar} \int_{-\infty}^{+\infty} (f_L(\epsilon_k) - f_R(\epsilon_k)) dE$$ (1.8)

The integral generates $eV$, and the corresponding $G = \frac{I}{V} = \frac{2e^2}{\hbar} = 1G_0$. With multiple modes and partially transmitted channels, (1.6) can be recovered.

Later on Büttiker generalized the Landauer formula to the multi-terminal cases\cite{26} and this formula is then well-known with the name Landauer-Büttiker formula. This is still an idealized treatment, though. It only considers coherent transport with electrons remaining in their initial transverse modes. No interaction is considered, meaning that it’s still of single-particle model. Furthermore, the adjacent areas connecting to the center constriction also contribute to the electronic transport, which in atomic-scale junctions further introduce quantum correction from backscattering or electron-phonon interaction and so on. Thus deviation from the pure Landauer-Büttiker transport can occur and the atomic-sized metal junctions appear to be more complicated and fundamentally interesting.
1.2 Conductance Quantization or not?

One direct result of Landauer formula is the expectation of conductance quantization when the constriction size is approaching $\lambda_F$. It requires that the transmittance $\tau_i$ of each quantum channel can only be 0 or 1, then the conductance can be quantized to integer multiples of $G_0$. Such a quantization was indeed observed in conventional GaAs/AlGaAs 2DEG in 1988, by Van Wees et al.[27] and Wharam et al.[28] separately. The ability of electrostatically varying constriction width continuously and the relative ease of fabricating disorderless devices over large scales enable this observation. The natural question is whether in atomic-scale metal junctions similar physics can be reproduced or not.

Lacking a way to vary the constriction width electrically, in atomic-sized contacts a mechanical approach is applied, either by STM[31, 32, 33, 34] or mechanically controlled break junction technique (MCBJ)[29, 30, 35, 36, 24, 37, 38]. In those studies a wider metal contact is formed first and then is mechanically elongated and narrowed down until its width is comparable to a few atoms’ scale. Such a movement will be reversed after its breaking and a new contact will be reformed. Statistical information can be achieved after large amount of repetitions and the so-called conductance histogram could be generated. One such example is shown in Fig. 1.2(b). The simultaneously recorded conductance trace presents quantization-like steps, as shown in Fig.1.2(a). Though some studies even gave very perfect quantization results[39], most such steps are not located at integer conductance values. The conductance histogram studies the ensemble-averaged information and gives a result closer to the expected conductance quantization behavior, where the conductance peaks can be very close to integer values in many metals and seems to support conductance quantization.

Now it’s already widely accepted that neither the conductance steps nor the con-
Figure 1.2: a) Step-like features in conductance recorded when a metal wire is mechanically breaking. Adapted from [29]. b) An example of conductance histograms in different metals: sodium and copper [30].

ductance histograms are necessarily signs of conductance quantization. They only mark the geometrically stable atomic configurations. In atomic metal junctions the accessible constriction geometries are clearly limited by atomic sizes and stabilities, as real atomic rearrangement is going on. AFM measurements do show that the abrupt changes of conductance are always companioned by atomic rearrangements, through the detection of forces[40].

A key result was revealed by E. Scheer et.al[36, 24]. They found the number of quantum channels is related to the number of valence electrons of the atoms. Thus in
noble metal like Au or alkali metal like Na, the single atom chain only contains one quantum channel, while in other metals multiple channels are involved in conduction in single atom chains. The conductance histogram of the trivalence metal Al still has peaks very close to 1, 2, 3$G_0$[41] even though multiple channels contribute, meaning that the conductance histogram only reflects the stability information. It is indeed shown that the Au single-atom chain has single nearly fully opened quantum channel, regardless the length of the chain[34], while the semi-metal Sb has the last plateau at only $\sim 1M\Omega$[42]. Fully transmitted channels and the integer conductance values are not special at all.

1.3 Quantum coherent transport

As already mentioned above, when electrons travel shorter distance than the quantum coherence length $l_\phi$, the phase information is maintained and as a direct result quantum coherent transport becomes relevant. In many textbooks or reviews people choose to discuss universal conductance fluctuation (UCF), weak localization and Aharonov-Bohm effect together as they are all direct result of quantum mechanical phase maintained in coherent transport. The atomic scale is small enough that even at room temperature quantum coherence is still maintained. The small junction size makes it hard to affect quantum coherence via magnetic field, and the ballistic nature further suppresses quantum corrections due to the lack of scatterings. On the other hand UCF tuned by bias voltage is still presented in atomic scale contacts as defect scatterings in electrodes are non-negligible. Assuming electrons start from the same initial state and propagate through two different paths to the same destination with quantum coherence maintained, during which different phase $\phi_1$ and $\phi_2$ are accumulated. Thus the probability to find one
electron at the destination is

$$A \propto |e^{i\phi_1} + e^{i\phi_2}|^2$$

$$= |e^{i\phi_1}|^2 (1 + e^{i(\phi_2 - \phi_1)})^2$$

$$= |e^{i\phi_1}|^2 (2 + 2\cos(\phi_2 - \phi_1)) \quad (1.9)$$

Two identical wavefunctions ($\phi_1 = \phi_2$) superposed with each other, which is the fully constructive limit, generates 4 times larger probability but not the classical result. If two wavefunctions have the phases differed by $\pi$, then the propagating probability is zero, which is the fully destructive limit. Consider the plane wave case, the phase satisfies $\phi = k \cdot r - Et/\hbar$ without magnetic field. Magnetic field further generates an extra phase of $\int A \, dl$ along the path and $2\pi \Phi/\Phi_0$ for a closed loop, where $\Phi$ is the magnetic flux threading the loop and $\Phi_0 \equiv \hbar/e$ is the flux quantum. The famous example is Aharonov-Bohm effect, which is shown in Fig. 1.3 (c) but will not be discussed in detail here. In diffusive conductors, many scattering events occur during electrons’ transport. Within the distance of $l_\phi$ all the scatterings are elastic. The magnitude of the relevant wave vector $k$ remains the same and only the moving directions are altered. Thus without the field the accumulated phase depends on the magnitude of $k$ and the total travel distance. The latter is determined by the position of defects which is relatively fixed in the conductor. Beyond a distance of $l_\phi$, inelastic scatterings take place, which mess up $E$ and $k$, and hence make the interference term $\cos(\phi_2 - \phi_1)$ totally randomized and washed out by average. This is the so-called decoherence. Higher temperatures raise the inelastic scattering rate and reduce the coherence length, making the long paths decoherent. On the other hand finite temperature broadens the available energy range by $\sim k_B T$. As a result $\Delta E t/\hbar$ starts to have non-zero contribution to the phase, where $\Delta E \sim k_B T$. Thus the new time
scale is $\sim \hbar/k_B T$, and the corresponding length scale is $L_T = (Dt)^{1/2} = (\hbar D/k_B T)^{1/2}$ with the name of thermal length. Beyond this length scale thermal excitations make transport appear decoherent.

![Figure 1.3](image)

Figure 1.3 : a) Weak localization. b) Universal conductance fluctuations. c) Aharonov-Bohm effect. (All three panels taken from [18])

When one considers the backscatter probability with time reversal symmetry, as shown in Fig. 1.3 (a), at zero temperature the phase accumulated through clockwise and counterclockwise loops are identical ($\mathbf{k} \cdot \mathbf{r} = -\mathbf{k} \cdot \mathbf{r}$). As a result the backscattering is always enhanced, and the resistance is raised higher than the classical value. This is "weak localization". As weak localization is geometry-independent, it's pretty robust
in large scale ensemble averages. One way to suppress weak localization is to raise
the temperature, where quantum coherence is destroyed and the resistance is reduced
back to its classical value, as shown in Fig. 1.4(a). Magnetic fields also suppress weak
localization, as this introduces extra phase varying from loop to loop and randomizes
the backscattering probabilities for different big loops, and with a field above $B_c$ the
classical value is recovered (Fig. 1.4(b)). In the 1D case due to the limited areas of
the loops, a much larger $B_c$ or a broader weak localization peak is expected compare
with the 2D case.

When one considers the transmission problems, electrons can pass along differ-
ent paths also determined by defect positions and also have interference-originated
corrections to its two terminal conductance, as shown in Fig. 1.3 (b). Similarly the
path and field dependent phase introduces quantum mechanical corrections to the
conductance. Varing the field and bias alter the phase and cause fluctuation-like os-
cillations of conductance, which is the universal conductance fluctuation (UCF). Bias
could alter the phase because it varies the wave vectors of the injected electrons hence
changing the $\mathbf{k} \cdot \mathbf{r}$ part. If the defect distribution is fixed, UCF retraces itself, with its
rms variation $\delta G_{rms} \sim e^2 / h$ at zero temperature. The value of this $\delta G_{rms}$ is universal
and not dependent on particular sample details provided that the quantum coherence
is maintained, though the exact fluctuation pattern varies from sample to sample.
For a diffusive conductor with its size much larger than the coherence length $l_\phi$, UCF
is suppressed ($\propto (l_\phi / L)^{3/2}$, L is the size,[16]) due to the ensemble average between un-
correlated regions while ensemble averages do not kill weak localization. At non-zero
temperature, both $l_\phi(T)$ and $L_T$ limits UCF. The thermal length matters only when
$L_T \ll l_\phi(T)$ holds, where a suppression of $(l_T l_\phi^{1/2} / L^{3/2})$ is expected[45]. Decoherence
can also be caused by high bias, where strong electron-electron interaction driven by
Figure 1.4: a) Temperature dependence of resistance in GaAs/AlGaAs 2DEG strips with different width. Adapted from [43] b) Field dependence at various cryogenic temperatures showing weak localization peaks in Si(111) MOSFET. Adapted from [44]

the non-equilibrium state reduces the coherence length. A cross-over between low bias enhanced UCF and largely suppressed UCF at high biases is expected[46, 47, 48].

The above discussion of coherent transport is all in the diffusive regime where the condition $l_e < L < l_\phi$ is valid. In ballistic regime, especially in atomic contacts, where we have $L \ll l_e$, the physics are different as no scattering is expected. As a result UCF is predicted to be strongly suppressed by a factor of $a/l_e$[51, 52]. It's not even weaker because the ballistic region has to connect with the electrodes, and the
Figure 1.5: a) Reproducible conductance fluctuation as a function of field. Adapted from [49]. (b) Non-equilibrium UCF as a function of two terminal bias voltage. Adapted from [48] (c) Ensemble averaged standard deviation of $\partial G/\partial V$. Adapted from [50].

defects in the electrodes within $l_\phi$ distance to the constrictions and the two contacts themselves still serve as scatterers. Small fraction of electrons are scattered back and forth through the ballistic constrictions multiple times and superpose with each other. The transparency of the constriction plays a role. Taking fully transmitted channels as an example, no electrons can be reflected back and UCF is suppressed. As a result, the standard deviation of $\partial G/\partial V$ over ensembles is minimized by fully transmitted channels[50, 53]:
\[
\sigma_{GV} \propto \sqrt{\sum_{i} \tau_i^2(1 - \tau_i)} \tag{1.10}
\]

where \( \sigma_{GV} \equiv \sqrt{< (\partial G/\partial V)^2 > - < \partial G / \partial V >^2} \) and \( \tau_i \) represents the transmittance of the i’th channel. The physical interpretation of Eq. 1.10 is stated below. Part of \( \partial G/\partial V \) has UCF origin. Only this part is sensitive to the atomic arrangement in the electrodes, which will be fully randomized in different ensembles. Thus the standard deviation of \( \partial G/\partial V \) only reflects the UCF, which is suppressed in fully transmitted or fully reflected channels.

### 1.4 Electron-phonon Interactions

The atomic contact itself is scatter-free. The elastic scatterings in the electrodes result in universal conductance fluctuations, as discussed in the previous section. Electron-phonon interactions are also possible, and these can affect transport. While we only focus on charge transport in the previous sections, another fundamental question is raised: how is the energy exchanged and dissipated during the charge transport through nano-scale junctions driven out of equilibrium? This will be mainly discussed here and in the next Section. 1.5, and will also be revisited during the discussion of inelastic effects on shot noise.

Energy is exchanged through (inelastic)scatterings. Through countless scatterings electrons and bulk phonon modes achieve thermal equilibriums in large conductors, like source and drain electrodes in typical atomic scale junctions. With a finite bias being applied, The Fermi energy of electrons in source and drain differ by eV. For an electron coming from the source, most of this extra eV energy is relaxed in the bulk electrode of the drain side, due to the ballistic nature of atomic contacts where the size
is much smaller than the electron-phonon mean free path $l_{e-ph}$. A small fraction of the energy is exchanged with the local vibrational modes in the constriction through electron-phonon interactions. The self-energy $\Sigma_{e-ph}$ broadens the electronic spectral function $A(\epsilon, k)$, from whose linewidth a finite life time or rate of electron-phonon scatterings is defined. The higher order terms of $\Sigma_{e-ph}$ are the process involving multiple phonons, which are usually neglected, while the lower order term includes both phonon emission and absorption processes. At zero temperature limit, the Fermi distribution $f \rightarrow 1$ below the Fermi level and the phonon population (with energy $\hbar \omega$) $n_\omega = 0$ based on the Bose-Einstein distribution $n_B = 1/(e^{\beta \hbar \omega} - 1)$. Only phonon emission processes are allowed if electrons have enough energy ($eV > \hbar \omega$) to excite a particular phonon (vibrational) mode, due to the natural requirement of energy and momentum conservation. This energy exchange mechanism transfers energy from excited electrons to atomic vibrations and causes heating, which will be discussed in more details in Section 1.5. Meanwhile the electrons are scattered, so that a correction of many transport properties are expected.

In high-transmission contacts the differential conductance is reduced by electron-phonon interactions and in tunneling regimes the conductance is generally enhanced. Both can be detected by very similar measurement techniques. The former is used to map out the phonon density of states, with the name of point contact spectroscopy[54, 55]. The latter is applied in tunneling through molecules to charaterize vibrational features in molecules, well known as inelastic electron tunneling spectroscopy (IETS)[56, 57]. Both show abrupt changes in differential conductance at a voltage exactly matching the phonon energy $\hbar \omega$, and therefore in second derivative $d^2I/dV^2$ spectrum peak or dip features appear. Regarding the sign change problem, multiple theories predict that the inelastic correction to conductance is negative if its conductance is larger.
than $0.5G_0$ and positive otherwise, in the single channel limit assuming symmetric couplings to electrodes\cite{58, 59, 60}. Such a crossover was indeed observed by Oren Tal \textit{et al.} in Pt/H$_2$O junctions\cite{61}. This sign change is suggested to be the result of a competition between enhanced backscattering in high-transmission case and enhanced forward scattering in the low-transmission case.

![Diagram](image1)

Figure 1.6 : a) IETS measurement on C$_2$H$_2$. Adapted from [57]. (b) Point contact spectroscopy on single atom Au chains. Adapted from [54] (c) Cross-over from positive modified conductance to negative modified conductance in H$_2$O/Pt junctions. Adapted from [61].

Point contact spectroscopy has been measured in larger metal wires\cite{62}. The second derivative measurements give\cite{17, 63}
\[
\frac{d^2I}{dV^2} \propto a^3 \alpha^2(\epsilon) F(\epsilon)
\] (1.11)

where the contact radius is represented by a, \(\alpha(\epsilon)\) is electron-phonon coupling strength, and \(F(\epsilon)\) represents the phonon density. This is seen in single atom junctions early this century\cite{55, 54, 64}. It’s clear that the phonon signal depends on the volume and suppose to be very weak at single-atom limit. Thus UCF is supposed to be stronger in most cases, except in contacts with conductance close to \(1G_0\), where UCF is suppressed and the phonon signal remains. Besides the weak signal strength, another feature of point contact spectroscopy at the atomic scale is its sensitivity of atomic configurations. It has been demonstrated that the phonon energies are shifted by applying strain\cite{54}, and the chain geometries also affect the electron-phonon coupling strength. For example, in zigzag Au single-atom chains, which in some cases are favourable under strains\cite{65}, electrons can couple with the transverse optical phonon mode while in linear chains this is suppressed\cite{58}. (In big point contacts the opposite is true, such that in all the noble metals the coupling with transverse phonon modes are stronger than the longitudinal modes\cite{66}, and the details of chain geometry are not important.)

Further complication is presented if the phonon population \(n_\omega\) is also driven out of equilibrium by bias, meaning that the electron-phonon coupling \(\lambda\) is strong and the corresponding electron-phonon scattering rate is much larger than \(\eta\), the dissipation rate from local vibrational modes to the bulk phonons. In other words, local ionic heating takes place. The elevated phonon population at least plays a role in determining the above-mentioned self-energy \(\Sigma_{e-ph}\), by increasing the phonon emission rate\cite{58, 67} and turning on the phonon absorption processes\cite{68}. Further corrections to conductance (and shot noise also, which will be discussed in detail lat-
er) are therefore expected, and may already have been observed in experiments[69].

A more complete microscopic understanding of both $\lambda$ and $\eta$ is still lacking and in many computations they are still pre-set parameters. It seems to be agreed that the local ionic (phonon) heating is relatively weak in Au contacts while stronger in molecules[68, 70, 71, 72].

1.5 Local heating

To address the local heating problem more clearly, let’s review the basic physical pictures of thermal equilibriums in a typical atomic contacts. Fig. 1.7 shows some of the relevant length scales to consider in this problem. A ballistic constriction bridges between macroscopic source and drain electrodes. With no bias or temperature gradient applied, the electronic distributions in the source and drain and throughout the junction region are described locally as Fermi-Dirac distributions, characterized by some electrochemical potential (Fermi energy) and a spread in energies given by $k_B T_0$, the equilibrium temperature of the system. Likewise, the local phonon populations (including optical phonons) are characterized by the same temperature, $T_0$. Finally, any local ionic excitations such as the ubiquitous tunneling two-level systems (TL-S) undergo dynamics determined by their coupling to the equilibrium phonon and electronic distributions.

When a voltage bias is applied and the system is allowed to reach steady-state, the situation is considerably more complicated, as shown in Fig. 1.7. Far from the junction (region A), the widths in energy of the electronic and vibrational distributions remain described by $k_B T_0$, presuming that heat is transported away into the bulk effectively. However, the Fermi energy of the source is raised over that of the drain by an amount $eV$. In the conventional treatment of such junctions[15, 73], the local
Figure 1.7: The physical picture of energy dissipation and local heating during electronic transport in a nano junction driven out of equilibrium.

electronic distribution functions within regions C and D are no longer simple Fermi-Dirac distributions. In the drain, for example, right-moving carriers that originated in the source (with energies $\sim eV$ above the Fermi level of the drain) and transmitted through the ballistic region are present. However, it is generally assumed that those transmitted carriers are drawn from a Fermi-Dirac distribution with a temperature $T_e \geq T_0$. As no thermal equilibrium is reached even between electrons in this region, the $T_e$ is not well-defined and can only be regarded as an effective electronic temperature. We note that dynamical TLS and mobile atoms in region C that couple
to electrons are sensitive to the full range of available electronic energies, and hence can respond\cite{74, 75} as if coupled to an effective temperature at most comparable to $eV$, generally much larger than $T_0$. On a length scale set by the electron-electron inelastic mean free path ($l_{ee}$), the electrons can exchange energy among themselves (B region), evolving their distribution toward a Fermi-Dirac distribution characterized by a position-dependent, elevated electronic temperature $T_e > T_0$. This is the electrode heating, which is trivial. On a still longer distance scale, the electron-phonon inelastic mean free path ($l_{e-ph}$), the electrons and phonons can both reach approximately thermal distributions (A region) given by position-dependent temperatures $T_e = T_{ph} > T_0$, which eventually converges to the $T_0$ deep into the bulk electrodes discussed in the very beginning. (The hierarchy of these length scales depends on relative inelastic scattering rates\cite{76}. The assumption $l_{ee} < l_{e-ph}$ is reasonable in Au at room temperature at a bias of 0.1 V\cite{77}, and $l_{ee}$ is likely to be further reduced in a junction’s confined geometry\cite{78}.)

By allowing electrons to exchange some amount of energy between each other in C and D region, part of the elevated electric potential energy $eV$ can be converted into thermal energy and some redistribution of electrons might be expected. Though fully thermal distribution is not reached, some elevation of local effective electronic temperature $T_e$ should be possible, which is the basic idea of local electronic heating. Indeed 200mV is equivalent to 2400K in degrees. To cause a significant heating effect one does not need a significant percentage of the electric energy converting into local heating. On the other hand, before the full thermal equilibrium between subensembles of electrons and phonons is reached, i.e. in the cross-over region between A and B (or even part of C) region, $T_e > T_{ph}$, but some energy exchange between electrons and phonons through electron-phonon interactions still occurs. This is the
so-called local ionic heating already partly discussed in the previous section. Local
electronic and ionic heating take energy from the same energy source (the biased
electronic distribution) and can be regarded as a cooling mechanism of the each
others[72].

As we already pointed out above, it is difficult to define “temperature” for a sys-
tem driven out of equilibrium and indeed several definitions can be advanced [79]. In
the above, the electronic and vibrational distribution functions in regions A and B
are well described as having the thermal functional form, with the effective temper-
ature defined through that description. In regions C and D, the electronic and ionic
populations are not thermally distributed. However, it is still possible to parametrize
such a nonthermal distribution with a single effective temperature. For example,
one can consider allowing local interactions with a large auxiliary “bath” (a thermal
probe), and determining $T_{\text{eff}}$ as the temperature of the bath such that there is no net
steady-state energy flux between the bath and the system [79].

In the conventional treatment of these systems[15, 73], the electronic distribu-
tions impinging on the ballistic constriction (region D) are well approximated as
Fermi-Dirac distributions with relative Fermi energies shifted by $eV$. There can be
corrections to this due to back-scattering from disorder in the diffusive source and
drain[50, 53, 55], and these are modeled as corrections to the overall transmittance.
In the standard picture, the effective electronic temperature $T_e$ of those FD dis-
tributions is established by energy redistribution among the electrons via inelastic
electron-electron scattering on the scale of $l_{ee}$ in region B, while local heating in
region C and D are not considered. By balancing the influx of energy from the bias-
driven nonequilibrium distributions and the outflow of energy by thermal conduction
of the electrons and eventually the phonons. The essential work by Henny et al.[80]
shows that at cryogenic temperatures it is easily possible to get significantly elevated electron temperatures (that is, $T_e - T_0 > T_0$) varying over tens of nanometers.

There have been a growing number of experiments to characterize the energy distributions of the electronic and ionic degrees of freedom in such nanoscale structures. In the case of ionic degrees of freedom, vibrational occupancies can be inferred directly by highly local Raman spectroscopy measurements[81, 82, 83, 84]. Estimates of effective ionic temperatures may be obtained through measurements of rupture strength of atomic-scale junctions[85, 70, 72], two-level fluctuations[86, 71] or black-body radiations[87]. Ionic temperatures may also be determined by measuring the response of an auxiliary system designed to exchange energy with the ions, such as a resistive thermometer[88] or a nanoscale thermocouple[89], though these typically indicate effective temperatures tens of nanometers away from the junction.

It is difficult to access the electronic distributions directly, particularly the non-thermal distributions in regions C and D. These nonthermal electronic distributions have been inferred quantitatively in some elegant experiments in metals[90, 91] and carbon nanotubes[92, 93] via local tunneling measurements. Noise measurements can provide comparatively direct access to the electronic distribution functions impinging on the constriction (the Fermi-Dirac distributions that go into the nonequilibrium distribution in region D, which dominates the transport). In equilibrium, the relationship between fluctuations and dissipation leads to the Johnson-Nyquist contribution,[94, 95] current noise that is white over a broad frequency range with a noise power proportional to the absolute (electronic) temperature and the electrical conductance, $S_I = 4k_B T_e G A^2$/Hz. Nyquist-Johnson noise is well established as a primary thermometer.[96] Additional “excess” or “shot” noise can appear when a nanostructure is biased out of equilibrium, due to the discrete nature of the electron-
Figure 1.8: a) Local ionic heating extracted by bond rupture strength measurements. The drop of high bias ionic temperatures indirectly probes electronic heating effects. Adapted from [72]. b) The inferred local electronic temperature elevation extracted from the anti-stokes background over single molecules. Adapted from [84].

ic charge. Previous shot noise experiments[80] in the diffusive limit, where a metal nanostructure is large compared to the mean free paths for both elastic and electron-electron inelastic scattering ($l_{ee}$), have found that the Nyquist-Johnson noise and shot noise clearly indicate elevated, bias-driven, position-dependent electronic temperatures $T_e(V) > T_0$ in mesoscopic wires at cryogenic temperatures. Those cryogenic experiments, accompanied by modeling, demonstrate that this heating is a result of the comparatively large thermal resistances of the reservoirs (regions A, B). We must
be mindful of this reservoir heating in any such experiment.

Beyond this quasi-equilibration of the electron subsystem at some effective $T_e$ that takes place in region B, it is an open question whether electron-electron interactions elevate the effective electronic temperature in the immediate vicinity of the junction (regions C and D). In other words, is the energetic width of the (approximate) FD distributions impinging on the ballistic region D the same as that at the nearby edge of region B, or are there local processes that further broaden those distributions (corresponding to an effective local increase in $T_e$ that is not the reservoir heating described above)? Theoretical predictions exist regarding local increases of the electronic temperature[97, 79] as a function of bias, but experimental probes have thus far been indirect[72, 84]. For instance, the bias evolution of a continuum background in anti-Stokes Raman scattering[84] ascribed to the electrons has implied effective electronic temperatures elevated by as much as several hundred Kelvin at biases of 0.4 V. Rupture strengths of molecule-containing junctions, accessing the electrons through their effects on the stability of ionic bonding, were reported[72] to be consistent with modeling of electronic heating in a hydrodynamic treatment of the electronic fluid[97, 79]. From the anomalous enhanced ensemble-averaged shot noise[98] at above 200mV in STM-style atomic-scale junctions, we are able to extract the effective electronic temperature relatively more directly. Significant local electronic heating is inferred and this work will be discussed in Section. 4.3.
Chapter 2

Shot noise

Some physical basics to understand mesoscopic transport phenomena, especially in atomic scale devices, are described in the previous chapter. Conductance, in other words the time-averaged electric responses in a stationary state, is considered there. However, current or voltage are always fluctuating and at each instance the accurate value of current or voltage is not predictable. Such a stochastic phenomenon is called (electronic) noise, which has its own physical origins and carries different information. Noise describes the deviations from the averaged value of the interested physical quantity and needs to be described in a different language, which is the main topic in this chapter and also most of this thesis.

The mathematical basics will be firstly introduced in Section. 2.1. After briefly discussing Johnson-Nyquist noise and flicker noise in Section. 2.2, the rest of this chapter will be spent on our main topic, shot noise.

2.1 Basics of noise

Math governs human’s thought, which is frustrating to say as a physicist but seems to be a truth. People rely on the language of averaging to understand the world, but encounter difficulties facing the true stochastic process, as the real-time response seems to be completely out of control. Fortunately, some forms of averaging, which don’t really need a lot more math beyond high school level, provide us a way to
incompletely describe some noise properties, which turns out to be very useful. As a result, we require a steady condition held to study noise.

In a steady state, meaning that the averaged properties of a system are not changing with time, we can consider a fluctuating variable $u(t) = u(t) + \delta u(t) \equiv u_0 + \delta u(t)$. In a realistic measurement $u(t)$ could be current or voltage. The time-averaged quantity $u_0 \equiv u(t)$ has to be independent of time as a natural requirement of a steady process. Equivalently $\langle \delta u(t) \rangle = 0$ has to hold. But $\langle (\delta u(t))^2 \rangle = \langle u^2(t) \rangle - \langle u(t) \rangle^2$ is non-zero, which is used to describe noise.

Mathematically an auto-correlation function of any $y(t)$ is defined to be

$$\rho(\theta) \equiv \langle y(t)y(t + \theta) \rangle \quad (2.1)$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y(t)y(t + \theta)dt$$

with $\theta$ representing any time difference. Auto-correlation functions measure how strongly the value of a variable is correlated to its own earlier values. Similarly a cross-correlation function between two variables $y_1$ and $y_2$ could be defined as

$$\rho_{12}(\theta) \equiv \langle y_1(t)y_2(t + \theta) \rangle \quad (2.2)$$

$$= \lim_{T \to \infty} \frac{1}{2T} \int_{-T}^{T} y_1(t)y_2(t + \theta)dt$$

The correlation between two different variables is measured. In frequency domain the Fourier transforms of both auto-correlation and cross-correlation functions are:

$$S(f) = \int_{-\infty}^{\infty} \rho(\theta)e^{-j2\pi f\theta}d\theta \quad (2.3)$$

$$S_{12}(f) = \int_{-\infty}^{\infty} \rho_{12}(\theta)e^{-j2\pi f\theta}d\theta \quad (2.4)$$

And inversely,

$$\rho(\theta) = \int_{-\infty}^{\infty} S(f)e^{j2\pi f\theta}df \quad (2.5)$$
\[ \rho_{12}(\theta) = \int_{-\infty}^{\infty} S_{12}(f) e^{j2\pi f \theta} df \]  

Take \( \theta = 0 \) and \( y(t) = \delta u(t) \), we have \( \langle (\delta u(t))^2 \rangle = \rho(0) = \int_{-\infty}^{\infty} S(f) df \). Thus the spectral density \( S(f) \) is introduced to describe noise properties in the frequency domain, and in a realistic noise measurement the frequency domain is usually what people really care about. On the other hand, cross-correlation method is applied to extract the same-origin part from two signals, which is usually used to suppress the background noise in amplifier chains. We can also easily see this from a few lines’ algebra. Let \( y_1 = u_0(t) + u_1(t), y_2 = u_0(t) + u_2(t) \), where \( u_1(t) \) and \( u_2(t) \) are completely independent and represent the background noise in two independent (but identical) amplifier chains. Thus

\[
\int_{-\infty}^{\infty} S_{12}(f) e^{j2\pi f \theta} df = \rho_{12}(\theta) = \langle (u_0(t) + u_1(t))(u_0(t + \theta) + u_2(t + \theta)) \rangle \\
= \langle u_0(t)u_0(t + \theta) \rangle + \langle u_0(t)u_2(t + \theta) \rangle + \langle u_1(t)u_0(t + \theta) \rangle + \langle u_1(t)u_2(t + \theta) \rangle \quad (2.7) \\
= \rho^{(0)}(\theta) = \int_{-\infty}^{\infty} S^{(0)}(f) e^{j2\pi f \theta} df
\]

As a result a cross-correlation measurement is equivalent to measure the spectral density of the common part shared in both \( y_1(t) \) and \( y_2(t) \) and the extra noise in the circuit can be suppressed.

Noise can be fluctuations of any physical quantities, such as current, voltage or resistance. In electronic systems, current fluctuation \( \langle (\delta I)^2 \rangle \) or voltage fluctuation \( \langle (\delta U)^2 \rangle \) are usually studied, as they can be directly measured. And the corresponding spectral densities are represented in \( S_I \) or \( S_V \). Sometimes \( S_R \) or \( S_G \) are also discussed. Note though many times people also call them current power or voltage power, the spectral density is not power really. The unit is \( A^2/Hz \) or \( V^2/Hz \), for
current fluctuation and voltage fluctuation respectively.

### 2.2 Common Electronic Noise Types

Depending on the specific physical origin, different types of electric noise exist. The most common electronic noise sources are Johnson-Nyquist (thermal) noise\cite{94, 95}, shot noise\cite{2} and 1/f noise\cite{99, 100}. This by no means includes all the fluctuation phenomena, as one can imagine any possible origins of noise.

Johnson-Nyquist noise or thermal noise can be observed everywhere and is well understood for decades. At frequencies below $k_B T / h$, its spectral density obeys

$$S_I = 4k_B TG$$  \hspace{1cm} (2.8)

$$S_V = 4k_B TR$$  \hspace{1cm} (2.9)

From Eq. 2.8 and 2.9 it’s clear that thermal noise is an equilibrium noise, meaning that this noise does not depend on external bias voltage or current. Only the temperature and the system resistance/conductance determine the noise spectral density. And it’s also clear that thermal noise is white, meaning that its spectral density is a constant at different frequencies. It’s true when the interested frequency $f \ll k_B T / h$, otherwise thermal noise will encounter a large roll off and quickly suppressed to zero at higher frequencies.

For a system with complex impedance, only the real part of the impedance $Z$ or admittance $Y$ generates thermal noise. As shown in Fig. 2.1, assuming a pure resistance $R'$ and a complex impedance $Z(f)$ are connected together, both maintain the same temperature $T$. The circuit is assumed to be perfect, meaning zero signal loss and zero noise generated (other than $R'$ and $Z(f)$). The allowed bandwidth is
\[ \Delta f \text{ and the transmission in the bandwidth is perfect. } W_u(f) \text{ represents the voltage spectral density of the thermal noise generated on the impedance } Z(f). \text{ The thermal equilibrium requires that the thermal power flow } P', \text{ from } R' \text{ to } Z(f), \text{ needs to be canceled by the thermal power flow } P \text{ with the opposite direction.} \]

\[ P' = Re\{4k_B T R' \Delta f \cdot \left( \frac{Z}{R' + Z} \right)^2 \cdot \frac{1}{Z} \} \]

\[ = \frac{4k_B T R' \Delta f}{|R' + Z|^2} \cdot Re\{Z\} \quad (2.10) \]

\[ P = W_u \Delta f \cdot \frac{R'}{|Z + R'|^2} \]

\[ (2.11) \]
Let $P' = P$ we can obtain $W_u = 4k_B T \cdot Re\{Z\}$. Similarly one can derive $W_i = 4k_B T \cdot Re\{Y\}$. It’s equivalent to say a pure capacitor or inductor does not generate thermal noise. Furthermore, applying the same method one can also easily verify the fact that thermal noise $S_T$ has to be proportional to $Re\{Z\}$ of its own.

With external biases, shot noise appears, if sample size is smaller than $l_m$. Shot noise is a nonequilibrium noise originated from the discreteness of charge carriers. And it’s also white with the frequency below $k_B T/\hbar, eV/\hbar$. Shot noise is the main topic of the whole thesis and will not be discussed here in detail. Both Johnson-Nyquist noise and shot noise in electronic systems are intrinsic fluctuation of a perfect steady Fermionic system and the physics behind them is unified. The only difference between the two lies in where the randomness comes from. R. Landauer has some very pretty and instructive works[102, 103] discussing both equilibrium noise and nonequilibrium noise using exactly the same wave-packet language. At finite temperature and zero bias case, the Fermi distribution introduces an averaged occupation $\bar{f}$ between 0 and 1 and generates Johnson-Nyquist noise due to the random unbalance of occupation numbers between the left and right electrodes. Clearly only electrons within the $k_B T$ window around the Fermi energy have $\bar{f} \neq 0, 1$ and generate thermal noise. Instead at zero temperature with finite bias, the electron occupation within the bias window are either 1 or 0 at source and drain, hence no flutuation from the occupation number is expected. But the finite probability (the transmittance $\tau_i \neq 0, 1$) to travel from source to drain generates the randomness between transmission and reflection and enables the shot noise. At both finite temperature and bias, both Fermi distribution around source and drain Fermi level and the barrier transmittance determines the noise. And the total noise (again when inelastic scatterings are absent), can not be simply distinguished into Johnson-Nyquist noise part and shot noise part. Based on
the same spirit, considering a system where nonzero transmittance is only available in a very narrow energy range around the Fermi level, people even predict an odd situation that the noise is reduced by finite bias[104].

Not the same as Johnson-Nyquist noise or shot noise, 1/f noise or flicker noise does not exist in the most simplified models. The biggest feature is that its spectral density is roughly inversely proportional to the frequency. 1/f noise is commonly observed in very diverse systems, ranging from the nonequilibrium low-frequency current fluctuation in resistive nano devices to traffic flows on highways. It’s clearly not rational to anticipate a single physical mechanism to explain all those loosely-correlated phenomena. The universality is not fully understood yet. The noise behavior varies from sample to sample, not only depending on the exact type of material but also the treatment experienced by the samples or substrates to choose and so on. Thus it’s largely believed to be an extrinsic noise, related to some dynamic process. In electronic systems, 1/f noise is a resistance fluctuation, which could be a result of atom motions and two-level fluctuations (TLF)[105], charge trapping[106] or mobility fluctuations[107, 108] and so on. Many of those suspected explanations have thermal-excitation origins and are expected to have temperature dependence, which is also the case being observed in experiments[105, 109, 110, 111, 112, 113]. Two-level fluctuations of any origins are usually assumed to be the most likely source of 1/f noise, where the 1/f spectrum is a result of superposition from many independent two-level fluctuators each obeying Lorentzian type spectrums, based on the McWhorter model[114].

The nature of resistance fluctuation means a quadratic dependence on bias current or voltage, because $< (\delta V)^2 > = < (\delta R)^2 > \cdot I^2$. Thus 1/f noise is a nonequilibrium noise, which is not generated by bias but revealed by bias. Empirically Hooge’s
law[116] has achieved big successes in describing 1/f noise, though it’s widely accepted that it has no fundamental origin.

\[ S_R(f)/R^2 = \frac{\alpha}{Nf} \]  \hspace{1cm} (2.12)

The parameter \( \alpha \) is material-dependent and \( N \) represents the number of fluctuators, which is not an accurate statement as the exact understanding of those fluctuators is absent. In bulk conductors this \( N \) is usually chose to be the total number of electrons. Note \( N \) is volume-related hence not independent from resistance \( R \), and \( S_R(f) \) does
not scale as $R^2$. Such a scaling rule was studied in break junctions and is different depending on the diffusive or ballistic nature of the particular devices[115].

2.3 Shot Noise

As already mentioned above, shot noise is a non-equilibrium noise originated from the discreteness of charge carriers. It was firstly discovered in vacuum diodes by Schottky in 1918[1]. He gives the very famous result of the classical limit:

$$S_I = 2e^* < I >$$

Eq. 2.13

The $S_I$ represents the spectral density of current fluctuation as usual, $e^*$ is the effective charge and $< I >$ the averaged bias current. This very neat formula reflects the non-equilibrium white noise nature and is valid in non-interacting tunneling processes, where the transmission is so weak that electrons cannot ”see” each other and pass the barrier independently. The uncorrelated arrival time intervals between electrons obey Poissonian distribution $P(\Delta t) = \tau^{-1} \exp(-\Delta t/\tau)$[2]. So shot noise expressed by Eq. 2.13 is also called the Poisson value. Correlations between electrons drive $P(\Delta t)$ away from the Poissonian distribution and reduces the shot noise spectral density (If the effective charge stays the same).

The unified understanding of noise has been partly mentioned in the previous section. The equilibrium noise (Johnson-Nyquist noise) originates from the uncertainty on occupation numbers $2f(E)(1-f(E))$. The zero-temperature non-equilibrium noise instead has pure partition origin, which is proportional to $T(E)(1-T(E))$. And the Poisson limit is recovered if $T(E) \to 0$. The most complicated situation is reached when both temperature and bias are finite, in which case the total noise should be related to $T(E)f_L(E)(1-f_R(E))(1-T(E)f_L(E)(1-f_R(E)))$ ($f_L(E)$ and $f_R(E)$ are
shifted by $eV$). A transparent system ($T(E) = 1$) recovers the equilibrium noise as there is no “partition”. Beyond the simplified physical picture, a more accurate description requires more sophisticated theoretical methods such as scattering approach, but all above conclusions still hold.

The non-Poissonian shot noise can be characterized by the Fano factor, simply defined as $S_I/S_P$, where $S_P$ represents the Poisson value. Note $S_P$ itself is proportional to the effective charge, which makes the Fano factor carry independent information from the effective charge and only reflects the partition nature. As a result, Fano factors are particularly useful in (but not only in) ballistic systems to determine the single-particle transmittances of multiple quantum channels[5, 6, 117]. Due to its sensitivity to correlations, the noise also reflects interactions between electrons or from phonons[118, 119]. Furthermore, in interacting systems shot noise provides a direct measurement of the effective charge $e^*$, which may be different from the free electron charge $e$. The detected effective charge could be a result of fundamental many-body quasi-particles like Cooper pairs[120, 121, 9] and Composite Fermions in fractional quantum Hall states[122, 11, 12, 13] or an average between multiple-electron processes[123]. To give a more complete discussion of shot noise, the following subsections will be organized in the order from simpler cases to more complicated cases.

### 2.3.1 Non-interacting Shot noise: single barrier

In this subsection let’s start the discussion from the simplest case. Here we ignore any type of interactions, but only focus on the partition nature of shot noise. And to make the topic simple enough we assume there is only one barrier existing in between two macroscopic metallic leads. This physical assumptions made here are simplified and instructive, but not invalid in practice. In reality it’s already sufficient to describe
the atomic scale junctions or 2DEG with one pair of split gates very well.

In elementary quantum mechanics, every student is required to calculate the transmittance of a particle with finite energy to tunnel through a barrier high compare with the particle’s energy. Unlike the classical physics, quantum mechanics allow the particle to have some finite probability to pass, if the barrier itself is not wide infinitely. Regardless of the exact microscopic nature of such barrier, we can borrow the concept of transmittance to describe the probability of a certain electrons to pass through the interested structures. Such non-transparency could be from a real tunneling or elastic backscatterings from disorders and so on. At zero temperature, as there are no Fermionic occupation number fluctuations on both leads, the only source for electric noise is from the partition noise generated at the barrier. As a result there is purely shot noise, which could be expressed in the language of transmittance $\{\tau_i\}$ of the system,

$$S_I = 2eVG_0 \sum_{i}^{N} \tau_i(1 - \tau_i)$$  \hspace{1cm} (2.14)

Combining the Landauer formula Eq. 1.6, it’s easy to see that the corresponding Fano factor is

$$F = \frac{\sum_{i}^{N} \tau_i(1 - \tau_i)}{\sum_{i}^{N} \tau_i}$$  \hspace{1cm} (2.15)

With the transmittance $\tau_i = 0$ or 1, electrons from the particular quantum channel are noiseless. In the former case $\tau_i \to 0$, current and noise are both reduced and the Fano factor is not minimized. If there is only one quantum channel then $F = 1 - \tau$, which approaches the Poisson limit $F \to 1$ in tunneling region when $\tau \to 0$. In the transparent limit $\tau_i = 1$, this fully transmitted channel presents zero noise but maximized conductance, thus minimized Fano factors are expected. Such quantum suppression of shot noise was unambiguously verified in quantum point contacts defined in 2DEG systems[7, 8], and later on in atomic scale metal junctions[5] and
molecular junctions[6]. The observation of quantum suppressions of shot noise (see Fig. 2.3 (a)(b)) confirms the role of fully transmitted channels in Au atomic contacts with conductance close to integer multiples of $1G_0$. On the other hand the same shot noise measurements in Al atomic contacts give much larger noise power[124, 17], which is consistent with the fact that single-atom Al contacts contain three quantum channels. The quantum suppressions demonstrate shot noise’s ability of determining the transmittances of quantum channels, as now shot noise and conductance provide two separate equations. If only one or two quantum channels are dominant in transport, the only transmittances could be accurately calculated[61, 125, 126, 127]; Even in the case of more than two quantum channels mathematically allowed ranges of the transmittances are still calculable[117].

At finite temperatures, as we already mentioned before, the Johnson-Nyquist noise and shot noise are mixed together. In that case the total current noise is expressed as

$$S_I = 4k_B T G_0 \sum \tau_i^2 + 2eV G_0 \coth \left( \frac{eV}{2k_BT} \right) \sum \tau_i (1 - \tau_i)$$  \hspace{1cm} (2.16)

With temperature $T \to 0$, the zero temperature shot noise result Eq. 2.14 is reproduced. With bias $V \to 0$ and combine Landauer formula Eq. 1.6 again the equilibrium Johnson-Nyquist noise expresion Eq. 2.8 is obtained. At finite temperatures, people are used to studying the excess noise($S_I(V) - S_I(V = 0)$):

$$S_I(V) - S_I(0) = 4k_B T G_0 \left( \frac{eV}{2k_BT} G_0 \coth \left( \frac{eV}{2k_BT} \right) - 1 \right) \sum \tau_i (1 - \tau_i)$$  \hspace{1cm} (2.17)

At high biases $eV \gg 2k_BT$, the excess noise scales linearly with the bias voltage just like its zero-temperature limit. But at low biases $eV < 2k_BT$, the hyperbolic cotangent term suppresses the excess noise and generates a low-bias curve, as shown in Fig. 2.3(c). Making use of this temperature dependence, shot noise thermometry is
2.3.2 Non-interacting Shot noise: beyond single barrier

A slightly more complicated physical problem is to consider the two barrier case. An idealized disorderless conductor, a 1D channel for example, is separated from both source and drain by two tunneling barriers. The transmittances through the two barriers $T_L, T_R \ll 1$, indicating very weak tunneling rates or couplings $\Gamma_L/\hbar, \Gamma_R/\hbar$. Possible, according to [128]. The main point is that the low bias curve is only related to $eV/2k_B T$ thus a calibration of absolute magnitude of shot noise is not needed.
Now we have two barriers generating noise. This simplified picture is directly related to molecular junctions or quantum dots with weak coupling to electrodes.

The resonance tunneling case requires full quantum coherence and could be effectively regarded as a special single-barrier problem. The overall transmittance of the whole system is

\[
T(E) = \frac{T_LT_R}{1 + (1 - T_L)(1 - T_R) - 2(1 - T_L)^{1/2}(1 - T_R)^{1/2}\cos\phi(E)}
\] (2.18)

The variation of the phase \(\phi(E)\) generates periodical sharp resonance tunneling features. Expanding the phase around the resonance tunneling energy values the Lorentzian shape Transmission functions appear,

\[
T(E) = \frac{\Gamma_L\Gamma_R}{\Gamma_L + \Gamma_R} \sum_n \frac{\Gamma}{(E - E_n)^2 + (\Gamma/2)^2} \quad (\Gamma \equiv \Gamma_L + \Gamma_R)
\] (2.19)

Index \(n\) marks the \(n\)'th resonance peak. The tunneling rates \(\Gamma_L,R/\hbar\) is directly related to the barrier transmittance \(T_{L,R}\) by \(\Gamma_{L,R}/\hbar = (v_g/w)T_{L,R}\), with \(v_g\) the group velocity of electrons and \(w\) the width of barrier. The expression Eq. 2.19 gives a series narrow peaks in \(T(E)\) with the peak value equals to \(4\Gamma_L\Gamma_R/\Gamma^2\). More opaque barriers(smaller \(T_{L,R}\), larger \(w\)) generates smaller couplings \(\Gamma_{L,R}\) and narrows the resonance features, but with the peak value remains unchanged if the ratio \(\Gamma_L/\Gamma_R\) is not altered.

Like the single-barrier case, integrations (assuming \(V > \Gamma\)) over \(T(E)\) and \(T(E)(1-T(E))\) gives the total current

\[
I = \frac{e}{\hbar} \sum_n \frac{\Gamma_{Ln}\Gamma_{Rn}}{\Gamma_n}
\] (2.20)

and zero temperature shot noise

\[
S_I = \frac{2e^2}{\hbar} \sum_n \frac{\Gamma_{Ln}\Gamma_{Rn}(\Gamma_{Ln}^2 + \Gamma_{Rn}^2)}{\Gamma_n^3}
\] (2.21)
The corresponding Fano factor therefore obeys following expression

\[ F = \sum_{n} \left( \frac{\Gamma_{Ln}^2 + \Gamma_{Rn}^2}{\Gamma_n^2} \right) \]  

(2.22)

At the symmetric coupling limit \( \Gamma_{Ln} = \Gamma_{Rn} \), the Fano factor is suppressed to its minimized value \( 1/2 \). This suppression of Fano factor was observed experimentally\[129, 130, 131\]. At the opposite limit \( \Gamma_{Ln} \gg \Gamma_{Rn} \) or inverse) the Fano factor approaches back to 1, due to the fact that only the limiting barrier is generating noise and the Poissonian value is recovered. A third limit is \( \Gamma_{Ln,Rn} \to +\infty \), meaning that \( T_{Ln,Rn} \to 1 \). In that case, the big \( \Gamma \) broadens \( T(E) \) and \( eV > \Gamma \) does not hold, thus Eq. 2.20 and 2.21 are both not valid. Instead, in the bias window \( T(E) = 1 \) (symmetric case) and no noise is expected, which recovers the result of fully transmitted channels in single-barrier case. Although above discussions are results of fully quantum coherence, surprisingly the Fano factor described by Eq. 2.22 is insensitive to dephasing and remains the same in sequential tunneling. A more instructive result was obtained when de Jong and Beenakker generalized the two-barrier sequential tunneling problem to \( n \)-barrier case\[132, 133\]

\[ F = \frac{1}{3} \left[ 1 + \frac{n(1 - T)^2(2 + T) - T^3}{[T + n(1 - T)]^3} \right] \]  

(2.23)

where they assume all the barriers share the same transmittance \( T \). This is surely an over-simplified assumption but it does link the two barrier ballistic systems to the diffusive conductors. With \( n = 1 \), \( F = 1 - T \), which is the single-barrier single-channel result; With \( n = 2 \) and \( T \ll 1 \), \( F = 1/2 \) and the symmetric two-barrier result is reproduced. And with \( n \to \infty \), we have \( F = 1/3 \), which is the universal Fano factor in the diffusive limit, without inelastic electron-electron and electron-phonon interactions. In such diffusive limit, multiple theories\[3, 134, 135, 132\], under classical or quantum mechanical assumptions, all agree with the universal 1/3 Fano
factor. The word “universal” means this $1/3$ Fano factors do not depend on the details of the particular diffusive conductor. The robustness of this universal value might indicate the equivalence between ensemble average and classical result.

Another interesting structure similar to the two-barrier system is the chaotic cavity realized in 2DEG. Such a cavity is also connected to source and drain through a pair of split gates, whose conductances are tuned to integer multiples of $1G_0$ hence present zero noise. Electrons spend long enough time in the cavity and the interference generates shot noise (or it’s the classical limit with no noise), which also maintains a universal Fano factor of $1/4$ in full quantum limit. I will stop here and not make the subsection too long, but refer the readers to several previous theoretical and experimental works[136, 137, 138, 139].

2.3.3 Single-particle shot noise with interactions

No interactions are discussed in above sections. When the size of the interested system is comparable to mean free paths of any inelastic effects, even weak inelastic scatterings could potentially complicate the physics and modify shot noise. Inelastic electron-electron interactions and electron-phonon interactions are the candidates. Here we will discuss the inelastic effects on shot noise, without counting manybody effects.

It’s well known that strong electron-phonon interactions completely suppress shot noise. Electron-phonon interactions also cause momentum change of electrons and affect electronic conduction ability. Indeed it’s the main reason metal has resistance at room temperature. Hence it’s easy to understand the fact that electron-phonon interactions can modify shot noise. On the other hand, inelastic electron-electron interactions do not change the total momentum of electrons and have no effect on
conductance. It’s still detectable by shot noise due to the fact that electron-electron interactions alter the energy distribution of electrons to which shot noise is sensitive. In non-interacting diffusive metal wires the Fano factor is of the universal value 1/3. Without the inelastic electron-electron interactions, the electronic distribution in the metal conductor is an average of the Fermi distributions on both source and drain, which is a position-dependent two-step non-Fermionic distribution function converging to Fermi distribution on the boundaries. When inelastic electron-electron interactions are turned on but electron-phonon interactions are still rare, energy exchange between electrons smears the two-step functions and make it Fermionic with temperatures varying locally in the limit of strong electron-electron energy exchange. Theories[140, 141, 142] predict the Fano factor at this limit is enhanced to $\sqrt{3}/4$, as a result of this local heating effect, which is also observed experimentally[118]. Similar effect was also observed in chaotic cavity where the cavity is large enough that electrons stay longer time than $\tau_{e-e}$ and enhancement of shot noise was seen[138].

With the size of the diffusive conductor continues to increase and strong electron-phonon interactions start to play a role, the shot noise decays with length and vanishes in macroscopic systems. It’s widely known for long and has been discussed by many different theoretical approaches[4, 3]. As a simplified physical picture, a fictitious voltage probe can be inserted (virtually) into the conductor to generate dephasing, and Beenakker[3] found shot noise is cut by half for a single voltage probe. Hence the noise is naturally expected to vanish with countless electron-phonon inelastic events presented. Thus by increasing the length of a diffusive metal wire, shot noise is firstly enhanced due to local heating and then suppressed to zero by electron-phonon interactions, as shown in Fig. 2.4(a).

Even in as small as molecular scale systems where the sample size is a lot small-
er than any inelastic mean free paths, weak interactions can still matter. Super-

Figure 2.4 : a) Shot noise expected in diffusive metal wires with different length at cryogenic temperature. Adapted from [118]. b) The signs of computed inelastic corrections to conductance and shot noise in atomic junctions, where local ionic heating is assumed to be negligible. Red color indicates an enhancement and blue suppression. $\alpha$ represents the symmetry of coupling to both source and drain electrodes ($\alpha = 1$ means $\Gamma_L = \Gamma_R$). The top panel is of conductance which gives the already known sign change occur at 0.5$G_0$. The bottom panel is of shot noise. Adapted from [143].

Poissonian[144] noise due to inelastic cotunneling process in carbon nanotubes[145, 146] or due to interacting localized states in resonant tunneling structures[129] were observed, and suppressed noise in Coulomb blockade (which is a little different as the smaller size enhances the Coulomb repulsion) region was also reported[131].
In atomic or molecular junctions, though the role of electron-phonon interactions is still an open question, theories do predict observable effects on shot noise. It’s true that up to moderate biases, the non-interacting predictions of shot noise agree well with experimental observations in break junctions at both cryogenic conditions and room temperature\cite{7, 8, 5, 6, 125, 61, 147, 148, 117}. However it’s also not difficult to imagine the existence of small inelastic corrections given that conductance can also be modified in similar systems. At zero temperature, non-interacting shot noise is always linear to the bias, if the Fano factor is independent of bias. Similar as what discussed in section 1.4, at zero temperature electrons still can emit phonons if the bias is at or larger than the electron-phonon interaction excitation voltage ($V = \hbar \omega / e$). Therefore the bias-dependent features in shot noise at bias larger than the phonon voltages are believed the key to understand this inelastic effects.

The existing theoretical predictions\cite{149, 150, 143, 151, 152, 153} all consider the coupling between electrons of single quantum channel and a single local phonon mode. The main physical differences between the models lies in the basic question of how to properly describe the phonon population of that bosonic mode. Thermally equilibrated phonon populations, thermally nonequilibrated phonon populations, and nonequilibrated phonon populations (driven into nonthermal steady state by the electronic current) corrected by their fluctuated dynamics are discussed. The relative importance between electron-phonon coupling rate $\Gamma_{e-ph}$ and the relaxation rate $\eta$ at which the local vibrational mode loses energy to the bulk phonons in realistic devices determines which regime is experimentally relevant. These theories all predict that a modification of the Fano factor that turns on when the bias voltage across the junction exceeds energy of the local vibrational mode, but disagree with each other about the exact modification of the Fano factor. Theories con-
sidering the coupling between electrons and thermally nonequilibrated vibrational modes ($\Gamma_{e-ph} \gg \eta$)[151, 152, 153] all predict that the bias dependence of $S_I$ has components that scale as $V^2$ or faster. Theories considering only thermally equili-

brated vibrational populations ($\Gamma_{e-ph} \ll \eta$)[143, 151, 150] instead predicts an abrupt change of $F$ at electron-phonon interaction onset voltage. (One result[143] is shown in Fig. 2.4(b)) Depending on the model, some complicated temperature dependence is predicted. Kumar[119] in 2012 observed kink-like features in the bias dependence of shot noise in many atomic-scale Au junctions at 4.2K, which is consistent with the predictions based on thermal-equilibrated phonon populations. These kinks are usually observed at a voltage between 10 to 20mV, consistent with the known energy of the optical phonon in Au[58, 54] and also their measured bias-dependent features in the differential conductance.

If the inelastic contribution to the noise is detectable, a temperature dependen-
t study is favored, as the phonon population plays an important role. In the last Chapter of this thesis we will have a detailed discussion about this topic. Our mea-

surements were conducted in atomic Au junctions with temperatures varying between 4.2K to 100K. The result seemly shows no sign of corrections from electron-phonon interactions, possibly because the transmission in Au atomic junctions is so easy that electrons spend too short time in the contact and the electron-phonon interactions are not strong enough to produce large effects.

2.3.4 Shot noise beyond single-particle limit

Under single-particle approximation, most correlations suppress shot noise below its Poissonian value. Only rarely in some special cases, noise is enhanced[2, 144]. And some of those enhancements are not purely shot noise: the enhancement from strong
electron-electron interactions in diffusive conductors and chaotic cavities are local heating effects; the enhancement due to interacting localized states in resonant tunneling regime has random telegraph noise origin[129] where the electronic conduction is randomly turned on and off. All those modifications of shot noise are included in Fano factors. Beyond the single-particle limit, many-body effects might drive the effective charge deviating from the value of free electron charge, which does not affect Fano factors but provides a separate mechanism to modify shot noise. As a result shot noise directly measures the effective charge, if the tested structure is designed to have a well-controlled Fano factor. Such a system can be a tunneling junction, diffusive conductor or quantum point contact and so on.

In superconductors, the supercurrent is carried by Cooper pairs, which is dissipationless and noiseless. However, hybrid structures combining both normal metals and superconductors can be designed to have finite current fluctuations, like NS interfaces or SNS junctions. The electronic transport through N-S interface is described by Blonder-Tinkham-Klapwijk (BTK) theory[154]. A NS interface can be modeled as a normal metal region connecting with a superconductor region, with a finite barrier (characterized by a dimensionless barrier strength $Z$) in between. With a bias small compare with the superconducting gap ($eV < \Delta$), Andreev reflection[155, 156] is the only way to transport charges: consider an electron with an energy $E_F + \epsilon$ is injected from the N side into a perfect ($Z = 0$) NS interface, a hole with an energy of $E_F - \epsilon$ is reflected back to the N side and a cooper pair at $E_F$ is transmitted in the S side and the conductance is doubled as doubled charges are passing. In reality the interface is not perfect and a barrier exists ($Z \neq 0$) so some electrons are reflected normally before reaching the NS interface hence the interface conductance is reduced. With even higher bias $eV > \Delta$, electrons reaching the NS interface even have finite probability
transmitting normally given that the quasi-particle excitations are also allowed in superconductors. This uncertainty (normal transmission or Andreev reflection) also generates noise[157, 2]. With lower bias \( (eV < \Delta) \), the situation is relatively simpler, that the normal side (including the barrier \( Z \)) itself generates noise and the Andreev reflection converts the normal current fluctuation into supercurrent fluctuation.

For a normal region which can be described by transmittances, de Jong and Beenakker[120] calculated the general results (assuming zero temperature again) of NS interface with arbitrary combination of transmittances \( \{\tau_i\} \)

\[
G_{NS} = G_0 \sum_i^N \frac{2\tau_i^2}{(2 - \tau_i)^2} \tag{2.24}
\]

and

\[
P_{NS} = P_0 \sum_i^N \frac{16\tau_i^2(1 - \tau_i)}{(2 - \tau_i)^4} \tag{2.25}
\]

where \( P_0 \equiv 2eG_0|V| \). It’s clear that both conductance and shot noise are deviated from the Landauer type results and there’s no easy way to compare Eq. 2.25 with the non-superconducting result Eq. 2.14. At the tunneling limit where the only transmittance \( \tau \ll 1 \), \( P_{NS} = 4eI \), with the Poisson value doubled. Similarly if the normal region is in the diffusive limit, shot noise will be \( 2/3eI \), which also doubles the normal result. Doubled shot noise is experimentally observed in both limits[158, 159, 9]. In SNS junctions which are long enough to prevent Josephson type cooper pair tunneling but still short enough to maintain phase coherence, electrons and holes bounce back and forth between the two NS interfaces and multiple Andreev reflections occur before the quasiparticles eventually move out of the bias window. A \( N \)-time Andreev reflection transports \( N + 1 \) charges and cause divergence of effective charge at low bias \( e^* = (1 + \frac{2A}{eN})e[160] \), which can be detected by shot noise[121, 10, 124]. At low transmission limit Cuevas et al. predicts discrete effective charges though
the sub-gap signals become small[124]. At high transmission limit multiple Andreev reflection feature is pronounced and but no discrete divergence of effective charges is expected[121].

![Figure 2.5](image)

Figure 2.5: a) Doubled shot noise in a diffusive normal metal region contacting a superconductor region. Adapted from [9]. b) The detection of Laughlin quasi-particle at 1/3 fractional quantum Hall states. Adapted from [12]. c) The shot noise measurements in quantum dots at and away from the Kondo resonance peak. Adapted from [161].

Another well-known success of detecting fundamental charge by shot noise is surely in fractional quantum Hall states, which is simpler due to the lack of a special transport mechanism like Andreev reflection. The effective fractional charge arises at
high magnetic field where vortices are attached to charges and generate a new quasi-particle with the name of Composite Fermions[162, 163, 164]. The fractional quantum Hall effect is thus reduced to the non-interacting integer quantum Hall effect of the Composite Fermions. The 1D Luttinger liquid nature makes quasi-particle transport through the edge states not the same with Landauer type ballistic transport. Luckily theories[122, 165, 166] still found in the narrow constriction of a quantum point contact, weak scattering between the edge states has exact the same form as classical Poisson formula

$$S_I = 2e^* I_B$$  \hspace{1cm} (2.26)

with $I_B$ is the backscattering current. Based on this predictions, fractional charge was experimentally observed at 1/3 state in 1997[12, 11] and later on at 2/5 states, further demonstrating shot noise measurements are not duplicating conductance information[13].

Kondo effect is also predicted to carry $5/3e$ effective charge (in full Kondo limit with charge energy $U \to \infty$ and symmetric couplings) revealed by shot noise[167, 123, 168, 169, 170] as a function of backscattering current when the transmission is close to unity. Unlike the case of superconductors or fractional quantum Hall systems this $5/3e$ effective charge is not related to fundamental quasi-particles but reflects an average between different physical processes. The existence of the inelastic scattering between two quasi-particles raises the averaged effective charge above one unit of free electron charge and below two units, which is the origin of the number $5/3$. The main feature of Kondo effect is the zero-energy peak in differential conductance due to the creation of a new many-body ground state pinned at Fermi energy as a result of full screening of the net spin at temperature $T < T_K[171]$. In quantum dots defined in conventional GaAs/AlGaAs 2DEG structures shot noise measurements firstly verified
this 5/3e effective charge coexisting with the Kondo peak, at some limited gate voltage range in 2008[161]. Away from that narrow gate region the Kondo resonance disappears and the effective charge drops back to 1. More careful and better-controlled experiments[172] were conducted later also in similar 2DEG structures, which further achieved temperature and gate dependence of the effective charge, consistent with the theories. Though inconsistency between two supposedly equivalent definitions of effective charge is also reported, which might be a consequence of contributions from other physical process to the current.

The intuitive understanding of why shot noise can potentially reflect the effective charges in many diversified correlated systems is easily achieved in the language of quasi-particles: charged particles bound to each other and as a whole behaves like a new version non-interacting particle carrying an effective charge. However, the correlated nature in those many-body systems make the general predictions of shot noise not that trivial and prevent people extracting effective charges from random structures containing those many-body systems. The successes people achieved here all make use of the weakly transmitted or weakly backscattered structures, where the transmitted or backscattered events are rare and the correlations are not important and the poisson type shot noise is expected.
Chapter 3

Techniques

3.1 Shot Noise Measurement Technique

Multiple techniques exist to measure shot noise. Based on the fact that shot noise is white up to a cut-off frequency $f \sim eV/h, k_B T/h$, which is very high in normal measurement conditions, both DC and RF measurement techniques are allowed. The DC measurements are technically simpler as this avoids the complicated RF signal reflection and radiation issues and their calibration could be simpler. On the other hand, for shot noise measurements, the $1/f$ noise always dominates the low frequency spectrum, which makes DC measurements painful to properly subtract the background not related to shot noise. In addition, for some quickly-varying systems the measurement of noise has to be rapid and the low frequency bandwidth is not suitable.

In this section, some common shot noise measurement techniques will be introduced. They all have their own advantages and drawbacks and all have been used by many research groups and achieved successes historically. Most of them are DC techniques, which will be introduced first, in subsection 3.1.1. The technique we use focuses on the RF bandwidth of noise, followed by other pioneers’ work[7, 14]. It will be introduced in more details in subsection 3.1.2. The DC techniques are suitable for detections of any type noise in the available bandwidth, while the RF techniques are only good for excess noise.
3.1.1 Common DC Measurement Techniques

A spectrum analyzer is the equipment most DC techniques used to measure noise, which gives the spectrum of the input voltage fluctuations within the allowed bandwidth. In a few works people also conduct high-speed measures of the voltage signals and calculate the corresponding fast Fourier Transform (FFT) spectra, which is equivalent to the spectrum analyzer measurements. The use of a spectrum analyzer is straightforward and convenient, and the real challenge lies in how to properly amplify the interested signal and suppress the unwanted background.

In a realistic noise measurement, the sample serves as a noise source, simultaneously generating Johnson-Nyquist noise at finite temperature, shot noise and 1/f noise at finite bias. The amplifier chains also generate current noise and voltage noise, within which the current noise further induces voltage fluctuations on the sample while the voltage noise is directly included in the output. Similarly, other resistive components in the circuit all generate their Johnson-Nyquist noise, part of which will also drop on the sample. Thus the total spectral density to be measured is (assuming the voltage fluctuation is measured)

\[
S_V = S_V^{(Johnson)} + S_V^{(shot\ noise)} + S_V^{(1/f)} + S_I^{(Amp)} R_{sample}^2 + S_V^{(circuit)} \left(\frac{R_{sample}}{R_\Sigma}\right)^2 \tag{3.1}
\]

where \(R_{sample}\) represents the sample resistance, and \(\left(\frac{R_{sample}}{R_\Sigma}\right)^2\) is the ratio other resistive components in circuit transferring their thermal noise to the sample. The above expression assumes a two terminal circuit geometry, and is subject to changes depending on the exact equivalent circuit. Technically there is no way to distinguish these five terms through a simple measurement, as they are real voltage fluctuations occurring at the amplifier inputs. So to measure shot noise more accurately what people can do is to reduce the temperature, to choose higher measurement frequency,
to choose low-noise amplifiers and to reduce the total resistance in the circuit.

Figure 3.1: a) The equivalent circuit of a cross-correlation spectrum measurement. Adapted from [5]. b) A Johnson-Nyquist noise measurement as a function of conductance, using the LCR circuit technique. The offset is the background of amplifiers. The schematics of the circuit is shown in the inset. Adapted from [12]. c) Current noise measurement circuit using SQUID. Adapted from [173].

Other instrumental background noise can be suppressed by choosing the right technique. The cross-correlation spectrum [8, 174] is one method commonly used, whose principle has been briefly discussed in section 2.1. This method (as shown in Fig. 3.1(a)) requires two identical amplifier chains measuring the same voltage fluctuations on the sample and assumes the background noises in the two amplifier
chains are purely independent. The main idea is that only the signals from the same origin correlate with each other and can survive after calculating the cross-correlation spectrum, thus all the background noises lie in the two amplifier chains are removed.

People can also choose not to suppress the background in amplifier chains but amplify the noise more before reaching the voltage amplifier. LCR circuit[12] is used based on this consideration. In this case, a two terminal geometry feeds the current passing the sample into the LCR circuit, as was shown in the inset of Fig. 3.1(b). The current fluctuations are selectively amplified and converted into voltage fluctuations at the resonance frequencies, which is further amplified and read by the spectrum analyzer.

Another special method to measure current fluctuation needs superconducting quantum interference device (SQUID)[173]. Fig. 3.1(c) gives the schematics of a typical measurement circuit, where the sample is represented by $R_x$. Small current fluctuation fed through the coil is measured by the SQUID, through the detection of the small magnetic flux variations caused by the current fluctuation. The bridge structure tunes the averaged current across the coil to be vanishingly small thus the SQUID could be set at its best sensitivity.

3.1.2 Our RF Measurement Techniques

Apart from DC measurement techniques, RF techniques also achieved considerable success in shot noise detections. The resistance fluctuation in the samples or external background, coming from the amplifiers for example, are all 1/f type or other low frequency effects. Hence naturally the RF measurement technique is less sensitive to these noise sources, which is the main advantage. Lacking a method equivalent to the cross-correlation spectrum to suppress the background, the RF technique actually
suffers more from the background noise and often has worse noise to signal ratio. A combination of modulated bias and lock-in detection is good for overcoming this issue, which means the RF technique is particularly good for excess noise $S_I(V) - S_I(V = 0)$ but not suitable for Johnson-Nyquist noise at temperature lower than the effective temperature of the background floor.

Figure 3.2 : Schematics of the RF excess noise measurement circuit. Used in [148]

Our technique follows previous pioneers’ work described in [7, 14]. The schematics of the circuit is shown in Fig. 3.2. A function generator (Stanford Research DS-345 or Rigol DG1022) is used to apply square wave bias switching between zero voltage and
a finited voltage across the sample and all the other series resistance in the circuit, if there is any. At the zero-voltage off state, the sample serves as a noise source of only equilibrium noise, while at the finite-voltage on state, the sample serves as a noise source of both equilibrium noise and nonequilibrium noise. Bias-tees are employed to separate the DC and RF signals coming from the junction. On the DC side, a current preamplifier (SR570) measures the current and its output is recorded electronically both directly by NI Daq and one lock-in amplifier synchronized to the square wave, giving a measure of the junction’s conductance. At the same time, the power of the RF signal is amplified and measured by a power detector and converted into voltage output. The power detector is in principle a diode, integrating the total power within the bandwidth limited by bandpass filters, which in our noise measurement is roughly 250-580 MHz. The other lock-in amplifier also synchronized to the square wave detects the difference between the power detector’s voltage outputs with and without bias applied to the junction; this is finally calculated back to the excess noise power. All the background carried by the active components in the circuit are assumed to be bias-independent thus supposedly removed in our measurements, which might not be true in all the cases but works pretty well till now. The power detector itself has not small voltage fluctuation at the output side but not correlated to the modulation frequency. All those together means our technique naturally only measures the nonequilibrium noise but a small background is still expected in practice. A long time constant set in the lock-in amplifiers compared with the modulated frequency provided by the function generator is definitely helpful to reduce the background. One must make sure the device to be stable enough during the time period. The best background that can be reproducibly achieved in our setup using a long enough time constant is $\sim 7 \times 10^{-27} A^2/Hz$ in lock-in R measurement and $\sim 1 \times 10^{-27} A^2/Hz$
in lock-in $x$ measurement. This measurement requires good RF shielding from outside EM waves and use commercial SMA cables with 50 $\Omega$ characteristic impedance as transmission lines, to minimize the RF loss and capacitive filtering of high bias signal to be transmitted. All the electric components are of 50$\Omega$ impedance type, to minimize power reflections between components.

![Equivalent circuit diagram](image)

Figure 3.3 : Equivalent circuit, adapted from the supplemental material of [147].

To deduce the correct value of interested current fluctuation spectral density from the power signal being measured, a detailed circuit analysis and systematic calibration are need. Here I will repeat the equivalent circuit analysis described in former
student Patrick Wheeler’s publication and thesis[147, 175] to make this section complete. In the equivalent circuit shown in Fig. 3.3, the sample with finite size together with all the structures before reaching the first 50 Ω component, which is the bias-tee in our circuit, is treated as an ideal current source $i_s$ parallel with an finite impedance $Z_s$. As the circuit is symmetric by using the same type bias-tee on both sides, we use $Z_0$ to represent the impedance of the bias-tee, which is 50 Ω. It’s an idealized equivalent circuit. One must note that the finite size and the complicated RF environment surrounding the sample will introduce further RF reflection, signal loss or radiation and absorption on the sample and make the result deviated from our calculation presented here. Indeed there’s no way to fabricate samples also in a nicely manufactured coaxial components.

Under the ideal condition, in our equivalent circuit the current source $i_s$ has two paths to complete a complete loop: $i_1$ through $Z_s$ and $i_2$ through $2Z_0$. And $i_2$ is the signal being measured. Thus we have

$$i_s = i_1 + i_2$$ (3.2)

$$Z_s i_1 = 2Z_0 i_2$$ (3.3)

Then we have

$$i_2 = \frac{Z_s}{2Z_0 + Z_s} i$$ (3.4)

And the power measured by the power detector

$$P = i_2^2 Z_0 G\Delta f = \left| \frac{Z_s}{2Z_0 + Z_s} \right|^2 i^2 Z_0 G\Delta f$$ (3.5)

$G\Delta f$ represents gain-bandwidth product, which is usually an integration if gain depends on the frequency but can be calibrated as a number if the noise is white. On the other hand, the reflection $|\Gamma|^2$ looking from the 50 Ω component into the sample
could be expressed as

$$|\Gamma|^2 = \left| \frac{(Z_0 + Z_s) - Z_0}{Z_0 + Z_s + Z_0} \right|^2 = \left| \frac{Z_s}{Z_s + 2Z_0} \right|^2$$  \hspace{1cm} (3.6)$$

As a result we arrive in the final relation between the current spectral density and the measured power

$$<i^2> = \frac{P}{|\Gamma|^2 Z_0 G \Delta f}$$  \hspace{1cm} (3.7)$$

The gain-bandwidth product $G \Delta f$ is measured by the spectrum analyzer (Agilent 89410A) as shown in Fig. 3.4, which is a systematic constant only related to the circuit and only measured once. And reflection coefficients, which instead is sample dependent, are measured using the circuit shown in Fig. 3.5. However, in the measurements involved in this thesis, we focus on the devices with resistances $\sim 10k\Omega$, where $|\Gamma|^2 \rightarrow 1$ hence the reflections are almost unchanged. A commercial white noise source is used to provide wide-band white noise across the RF bandwidth of interest. The amplitude of this noise is modulated by a subsequent RF switch that is turned on and off at the same (acoustic) frequency used for the square wave voltage bias we applied in excess noise measurement. Part of the white noise will be reflected at the boundary between the gold junction and the transmission line to the bias tee. The reflected power goes through a directional coupler as well as the same amplifier chain used in the excess noise measurements, and is registered by the logarithmic power detector. Besides the method shown in Fig. 3.5 one can also use spectrum analyzer to measure reflections[175], which contains the whole spectrum hence is nicer. But it’s a slow measurement and needs further integration to pull out the number, so it is not as quick and convenient as what’s used here.

With both gain-bandwidth product and reflection coefficient known, from the measured power at every instant the current noise generated on the sample can be
Figure 3.4: Adapted from the supplemental material of [147]. a) Gain-bandwidth product measurement using spectrum analyzer. b) Background measurement.

calculated based on Eq. 3.7. And in our modulated RF measurements, \( P(V) - P(V = 0) \) is proportional to the interested non-equilibrium current noise on the sample. The only thing slightly non-trivial is that the power detector we use is logarithmic, as a result the lock-in measurement of the voltages outputted from the power detector itself in not proportional to the excess noise and the averaged voltage level is also required to calculate the accurate \( S_I \). The details could be found in Patrick Wheeler’s thesis[175].

In our measurements, the idealized formula Eq. 3.7 works pretty well and the
Figure 3.5: Schematics of reflection measurement circuit. Used in [148].

obtained results are proportional to the real unit excess noise. However the exact knowledge of the impedance network on the sample is far not trivial to know, especially at high frequencies and in highly impedance-mismatched cases, even with the help of a network analyzer. As a result, an unknown damping ratio which not only related to the measurement setup but may also varied from sample to sample is always presented. The absolute values of the excess noise always require a normalization by the measurements in the tunneling regime where we are sure the classical shot noise is measured. The systematic correcting factor existing in the measurement
Figure 3.6: Reflections from devices with different pad areas. Inset: the equivalent circuit including capacitive substrate.

setups, such as noise loss in the transmission lines, should be a constant and very trivial to measure. We can repeat the reflection measurements with no sample in. In that case, if the cabling is lossless, all the RF signal must be reflected and the measured reflection coefficient is 1. Any values smaller than 1 are the result of noise loss during the transmission and a systematic damping ratio can be obtained. Note by using this method, the RF signal travels twice the distance in the transmission lines compare with a realistic noise measurement, so instead the square root of the resulted damping ratio is what we want.
On the other hand, the sample dependent noise loss is not fully understood. Surely the device geometry and wiring can carry extra capacitive or inductive impedances. But assuming small size of the devices and good low-resistance ohmic contacts, which is usually the case, it’s not irrational to assume these factors are not the most important. At this moment, I tend to believe the substrates, which could carry non-negligible capacitance and allow part of RF signal pass, dominate the sample-dependent correction (reducing) factors in many cases.

The corresponding simple model is described in the inset of Fig. 3.6, where the only change relative to the idealized equivalent circuit shown in Fig. 3.3 is the extra red capacitor supposedly carried by the substrate. Based on this model, the substrate carries an impedance of $Z_{\text{sub}} = 1/(j\omega C)$, and part of the RF signal flows through this $Z_{\text{sub}}$ back to the ground and is not measured. This could generate an extra damping ratio of $|Z_{\text{sub}}/(Z_{\text{sub}} + Z_0)|^2$. Nano-fabricated devices have gold pads and leads used as contacts. Larger pad areas correspond to smaller $Z_{\text{sub}}$ and larger damping ratio due to the fact that the parallel-plate capacitor is proportional to the plate area. Indeed on lithographically-patterned devices, it’s common to have smaller noise values measured with larger contact areas. On the other hand, if the reflection $|\Gamma|^2$ is measured, a minimized reflection should be observed when $Z_{\text{sub}} = 50$ Ω. Thus with too large pads where $Z_{\text{sub}} < 50$ Ω, $|\Gamma|^2$ will instead increase with larger area. Such a test was done and shown in Fig. 3.6. A minimum of reflection is indeed seen, confirming the role of substrate impedance and its pad-area dependence.

The magnitude of the capacitance carried by the substrate could be estimated. Assume a 1μm thick polyimide film ($\epsilon_r \sim 3$ at DC limit, which is pretty much a constant below GHz range) on top of a stainless steel substrate and a pad area 400μm × 300μm, which is very close to the case of the lowest-area point shown in
Fig. 3.6, the simple parallel-plate capacitance is $\sim 3.2pF$ and the corresponding impedance at a frequency of 350MHz is $Z_{sub} \sim 142\Omega$. It already means a damping more than one half. One can further estimate the inductance carried by the thin gold wire connecting the pad and chip carrier, that for a 2cm long 0.1mm thin Au wire the inductance could be $\sim 23nH$ and the corresponding impedance $Z_{wire} = j\omega L \sim 50\Omega$. Even based on that simple estimations, the damping of RF noise at 350MHz at the sample is $|Z_{sub}/(Z_{sub} + Z_{wire} + Z_0)|^2 \sim 0.34$ ($\sim 2.9$ times) and the reflection $|(Z_{sub} + Z_{wire} - Z_0)/(Z_{sub} + Z_{wire} + Z_0)|^2 \sim 0.34$. Both are somewhat consistent with the measured total noise damping $\sim 13$ times (including the cabling and system damping ratio) and the reflection 0.33 measured as the lowest-area point shown in Fig. 3.6.

The minimum in Fig. 3.6 means that the gold wire can’t carry an impedance more than 50$\Omega$ thus our above magnitude estimations are not accurate. The weak increase of reflection after aggressively increasing the pad area means that the impedance carried by the gold wire is not too far from 50$\Omega$ so it’s reasonably close to our estimation. Clearly the exact impedance network on the sample can not be that simple and the broad-bandwidth measurement procedure further requires the knowledge of the whole frequency dependence, which make the mathematical estimation of the systematic damping ratio very difficult and normalization is not avoidable if we want to know the absolute value of our measurement. If we believe we understand the physical origins of the sample-dependent noise damping, we can be more confident that we are doing a correct job.

Further tests also support the scenario that RF signal can leak through the substrates by significant amount. On the same chip with polyimide coated stainless steel substrate, if injected RF power within the same 250-500MHz bandwidth into one
pad, a significant fraction (depending on the relative position between the pads) of the signal can be detected from many other isolated pads where the only path RF signal can flow through is the polyimide layer. On the other hand not small gold leads on the chip carrier (if completely not touching any part of the chip) pick up nearly zero signal. It at least means in our measurements where the device size is still small, radiation is not a big factor, which is surely what we expected. Similar tests are reproduced well on doped silicon substrate with $2\mu m$ SiO$_2$ layer but the RF leakage is better probably due to the thicker insulating layers. Pavlo zolotavin in our lab reported a larger noise detected on his device using macroscopically thick sapphire substrates. And another student Panpan Zhou observed significantly larger RF noise at temperatures below 30K in tunneling junctions with doped silicon substrates and unknown thickness SiO$_2$ layers in between. At similar temperature regions the dopants were indeed observed to be frozen, which also highly suggests a capacitive origin of RF signal damping in his observation. Such a temperature dependence is not seen in stainless steel as metal is conducting intrinsically.

To conclude the whole section full of technical discussions, I want to highlight a few suggestions in the RF noise measurements based on my experience. They may not be all correct but I found them usually helpful.

- Use highly insulating substrates like thick sapphire, if possible.
- Minimize the total areas of pads, leads and so on.
- Always maintain good RF shielding. Turn off active components (like heater, temperature controller and so on) if possible. If significant amount of external background is picked up, it will complicates everything and don’t expect the resulted background to be bias-independent. Shielding is easy. A piece of Al
foil is enough, and don’t over-tighten the SMA connections.

- Be careful with gates. Firstly they introduce further noise damping as they are in principle capacitors. Secondly with charges accumulated on the sample, shot noise is predicted not to be white any more[2].

- Do normalization. But be careful when the normalization factor is smaller than 1.

- Though broad bandwidth help to increase the resolution in our measurements, a lower frequency (also means narrower bandwidth) is worth a try.

3.2 Break Junction Techniques

Break junctions are usually employed to study electronic transport phenomena in atomic or molecular scale. The challenge is that the normal nano-fabrication techniques are not sufficient to reproducibly generate devices in that small length scales and special techniques are needed. STM[176, 177], mechanical controlled break junction (MCBJ) technique[29, 178, 179, 180] and electromigration[181, 182] are the common methods to achieve those small devices. While true STM is not accessible in our lab, MCBJ and electromigrated (EM) junctions are used a lot in our experiments and will be introduced in this section.

3.2.1 Electromigrated Junctions

Electromigrated junctions were firstly employed to form molecular junctions by Park in 1999[181], (later than the invention of MCBJ techniques) though electromigration had already been known for long by then to be the main reason causing microelectronic failure[183]. Electromigration describes the physical phenomenon that the current
passage moves metal atoms, eventually weakens metal constrictions, and results in failure in the end. This is caused by the direct momentum transfer from conduction electrons to metal atoms, but not as a pure thermal effect (melting).

Figure 3.7: a) A SEM image of a typical bowtie junction after electromigration. The gap is very clear, though the narrowest point is beyond the resolution of SEM and can not be captured visibly in the photo. b) The typical design of a conventional MCBJ setup. Adapted from [178].

This technique can reproducibly generate nm-scale lateral nano-gaps and has no special equipment requirements, hence is widely used in the fabrications of molecular junctions. It has a clear disadvantage of lacking fine control and a continuous variation of conductance/resistance is not practical. In our experiments, large ensembles
of atomic configurations are of interest and the control is important, thus electromigration is usually not time-efficient. In the work presented in the Chapter 5, we still employ this technique to narrow down the Au constrictions before the mechanical breaking process, which saves time and travel distance and makes the breaking easier.

In a typical procedure of electromigration, a current constantly sweeping from zero to a finite value in a sawtooth shape is applied across the target device while its resistance is watching closely meanwhile. The highest currents trigger electromigrations while the low bias I-V curvatures are used to calculate the resistance. The control of the bias current/voltage determines whether the break down procedure is slow and smooth or quick and abrupt. I believe the smoother EM procedure creates mechanically more stable and more controllable junctions which are nicer in the following mechanical breaking process. The geometrical size and mechanical property of the device matters in an EM procedure. A wider, hence more conductive metal constriction, needs a lot higher current to start EM, and overheating at the moment starting the EM procedure might cause quick failures. For typical Au bowtie junctions with \( \sim 800\text{nm} \) long and \( \sim 150\text{nm} \) wide, I always choose a thickness less than 20nm where the EM procedure is still easy but an overloading of the current preamplifier (SR570, lowest gain) is already not avoidable. Such a device is typically of \( \sim 70\ \Omega \) at room temperature and \( \sim 40\ \Omega \) at 4.2K. An annealing occurs first and junction resistance will decrease before an EM starts. Usually before reaching 1k\(\Omega\) the EM procedure is easy. Above 1k\(\Omega\) EM is poorly controlled and results in random higher resistances or even failures. Fig. 3.7(a) shows a SEM image of one typical junction after EM procedure. Note that though the gap in the image is pretty clear, the dominant resistance comes from the narrowest point which is not visible by SEM.
3.2.2 Mechanical Controlled Break Junctions

Compare with the EM break junction technique, MCBJ technique can achieve sub-angstrom precision and fine control of atomic configurations. The disadvantage is that MCBJs always require finite volumes and specific setups.

A typical design of the conventional MCBJ setup is shown in Fig. 3.7(b), which is adapted from the pioneer work done by Muller[178]. A suspended metal wire is attached on a flexible substrate, with a piezo actuator pushing in the middle and two counter supports pressing against the substrate on its two ends. In the earliest versions, the suspended metal wires were usually macroscopically thick (1mm diameter for example) with a cut in the middle. Microfabricated suspended devices were also employed[179, 184] shortly after the invention of MCBJ technique. In Fig. 3.7(b), $l_s$ represents the distance between the two supports and $u$ is the length of the suspended part of the device. Based on this geometry, one can obtained the displacement ratio

\[ r_d = \frac{3ut}{l_s^2} \] (3.8)

$u$ and $l_s$ are the length of the suspended part in the device and the distance between the two counter supports, as defined in Fig. 3.7(b), and $t$ is the thickness of the substrate. This displacement ratio reflects how effectively the vertical pushing displacement caused by the piezo can result in horizontal displacement on the suspended part of the device. The smaller the ratio is, the more precise the position control can be. This ratio can be easily reduced as it’s a pure geometric factor. Another advantage of the small displacement ratio is that it provides high stability, but on the other hand it also slows down the bending procedure and limits the total travel distance on the device.

Besides the conventional design, it’s also possible to make a STM-like vertical
version, which has become very popular in the study of molecular conduction[185, 186]. In those STM-style MCBJs, the piezo controls a tip moving towards or apart from the substrate, and the displacement ratio $r_d$ is roughly 1. Such a simple setup won’t have good position control and the stability is poor, but it’s fast and very suitable for ensemble average measurements. Our work described in Chapter. 4 all use this type MCBJ technique.
Chapter 4

Shot noise in STM-style Au break junctions at room temperature

4.1 Motivation and technical details

In this chapter, we consider the noise properties of ensembles of atomic-scale metal point contacts from tunneling to the multi-channel \((G \sim 6G_0)\) regime, at room temperature, realized by a scanning tunneling microscope (STM)-style break junction. At this temperature inelastic processes involving phonons should be considerably more important than in the cryogenic limit. Even at 4.2K inelastic effects have already been observed in nanotubes[187], bilayer graphene[188], and in atomic-scale Au junctions[119], though the particular changes in \(F\) depend in detail on channel transmission. So here at room temperature of particular interest are the accuracy and utility of Eq. 2.16, over a broad range of applied bias, and the relative contributions of other noise mechanisms, such as conductance fluctuations[100]. The related works discussed in Section. 4.2, 4.3, 4.4 were published in [148, 98, 189] respectively.

In all the works mentioned in this chapter, a junction is repeatedly made and broken in ambient conditions between a 50 nm-thick gold film evaporated on an oxidized silicon substrate, and a cut gold wire. This gold wire was originally 0.5 \(mm\) thick and cut before experiments to expose clean surfaces. No other special treatments are required to further sharpen the tip, as single-atom or a-few-atom contacts are always reached during the breaking procedure. A computer-controlled piezo actuator
is used to form and break the junction typically several times per second. A photo of the setup is shown in the inset of Fig. 4.1. The noise measurement approach is already discussed in the previous chapter. The desire to examine the ensemble-averaged noise leads to the choice of the STM break junction method; the need for rapid measurement of the noise during the junction breaking process necessitates the use of our high bandwidth radio frequency (RF) technique. Due to the same reason the modulation frequency also needs to be quick and we choose it $\sim 10kHz$.

Figure 4.1 : Ensemble-averaged reflection measurements over STM-style Au break junctions

As an added complication compared to a fixed device configuration, the tip’s re-
peatedly vertical motions introduce the extra complexity of a strongly time-varying DC conductance into the junction’s RF properties. To make sure nothing dramatic occurs regarding the impedance during the STM-style motions, a reflection measurement was applied. Given that the DC conductance of the junction changes by orders of magnitude on millisecond timescales, it is not possible to measure all of the relevant RF parameters in real time. As an approximation to this, we instead measured the reflection properties as a function of conductance *averaged* over the ensemble of junction configurations.

The reflection measurement technique is standard and the same as what has been described in Chapter 3.1.2. Simultaneously, the junction’s cyclical STM-style motion is executed, with an applied DC bias across the junction to allow the simultaneous acquisition of conductance. The resulted reflection averaged over the ensemble of junction configurations is around 22%, shown in Fig. 4.1. There is some systematic variation with conductance, but this is small, less than 2% over $10G_0$. Note as what already discussed in Chapter 3.1.2, this reflection measurement actually gives a total ratio combining not only reflection but also other impedance-related RF correction factors. The reflection coefficients $|\Gamma|^2$ itself has to be very close to 1 even at as low as 1kΩ resistance.

Since the impedance mismatch between the junction and the measurement circuit does not vary dramatically on average over the range of junction configurations, there should be a scale factor (approximately constant across the conductance range) between the measured power and the true current noise across the junction. In all the work involved in this chapter, we attempt to find this factor by using the knowledge that the Fano factor in the tunneling regime $G << G_0$ approaches one in the limit that a single poorly transmitting channel dominates the conductance, as apparent
from Eq. 2.14.

4.2 Ensemble average at moderate biases

4.2.1 Conductance histogram and shot noise suppression

At every single bias, the STM style motion of the gold tip repeats hundreds times to generate a histogram of conductance, as well as a plot of the ensemble averaged excess noise power vs. conductance. An example is provided in Fig. 4.2.

Figure 4.2: An example of conductance quantization (a) and noise suppression (b) at room temperature. (a) A conductance histogram acquired with 180 mV amplitude for the biasing square wave. (b) Averaged power spectral density vs. conductance. The bin size in both plots is 0.01 $G_0$. 

As has been seen in many previous experiments in atomic-scale metal contacts[17], peaks are observed in the conductance histograms, signifying preferred junction configurations with specific values of conductance. Peaks are observed at $1 \, G_0$, and near other integer multiples of $G_0$, consistent with past results on Au junctions at room temperature[38]. Cryogenic experiments involving shot noise[5] and subgap structure in superconducting contacts[190] have demonstrated that the $1 \, G_0$ peak in Au junctions is dominated by configurations with a single highly transmitting channel ($\tau \to 1$). In our structures usually the first three conductance peaks are readily resolvable in the histogram. The related ensemble-averaged noise power measurement is also shown. As is clear from the figure, the ensemble-averaged noise power is clearly suppressed near conductance values where the conductance histogram is peaked.

A series of square wave bias voltages from a few millivolts to several hundreds of millivolts are used to acquire conductance histograms and ensemble-averaged noise. An example of such ensemble-averaged noise data for eight different bias levels is shown in Fig. 4.3. All the data shown in this figure were taken continuously in one day to ensure an identical experimental environment (lab temperature, any stray RF background) and nearly identical conductance histograms are generated at all the eight biases. In our system no clear excess noise can be detected at bias voltages below about 25 mV, corresponding to $k_B T$ at $T \approx 300$ K. Noise suppression at conductances corresponding to the first three peaks in the conductance histograms are very clear, and the magnitude of the detected noise power is monotonously increasing with bias as expected. Note that the suppressions are not complete. This indicates that at least some of the junction configurations corresponding to peak conductance values result from a mixture of multiple, partially transmitting quantum channels.
4.2.2 Linear bias dependence of shot noise

We are interested in the scaling of the noise with bias, as this reveals the Fano factor. Working with the data sets in Fig. 4.3, we can specify a particular conductance value, and for each applied square wave bias voltage we can compute the actual voltage drop, $V$, across the junction.

Starting from Eq. 2.16, we can see that plotting the detected power spectral density as a function of $X = 4k_B T G[(eV/2k_B T) \coth(eV/2k_B T) - 1]$ should produce
a linear graph, with the slope giving the (zero temperature) Fano factor, $F$. Here $G$

![Graph](image)

Figure 4.4 : Noise as a function of scaled bias. a) Data below $1\ G_0$. b) Data around $4\ G_0$, $5\ G_0$ and $6\ G_0$.

is the conductance of the junction and $T$ is assumed to be the ambient temperature, $\approx 300$ K. In Fig. 4.4(a), we plot data for three closely spaced values of conductance, all below $G_0$. The figure shows that the noise is in fact linear when plotted as a function of $X$, and the Fano factor is decreasing as the conductance approaches $1\ G_0$. 
These are consistent with the measured signal originating from shot noise, and the suppressions shown in Fig. 4.3 originating from the saturated channel mechanism of Eqs.2.14, 2.16. Qualitatively similar declines in Fano factor are also observed as the conductance approaches the values associated with the other two peaks in conductance histograms/suppressions in measured noise power.

We now consider the situation at higher conductances, when the number of channels involved in transport is large compared to one. As discussed above, we assume that the Fano factor in the low conductance regime is close to one, as expected if transport is dominated by a single poorly transmitting channel. We perform a linear fit to the noise vs. $X$ data at $G \approx 0.1 \, G_0$, and assume that corresponds to a true Fano factor close to one (though a true single channel device would have $F \approx 0.9$ near that conductance). In the particular data set shown in Fig. 4.4(a), the lowest conductance data around $G = 0.45 \, G_0$ are still well described by a similar Fano factor, as shown by the blue line. This suggests that, when looked over the full ensemble average, the conductance in this regime results from several poorly transmitting channels, rather than being dominated by a single channel with a transmittance of 0.45. This indicates that a large contribution to the conductance histogram results from junctions with comparatively blunt tips. Fig. 4.4(b) shows, on the same plot as that low conductance data, the noise data vs. $X$ at around 4,5 and 6 $G_0$. Identically scaled linear fits show slopes of 0.39, 0.34, and 0.31, respectively. These relative slopes are comparatively insensitive to the choice of the low conductance fits used to find the scaling factor, provided $G < \sim 0.5 \, G_0$. At even higher conductance values not shown in the figure, the inferred Fano factors also do not vary by much, decreasing to around 0.25 near 10 $G_0$.

These observations are roughly consistent with expectations for a crossover toward
diffusive conduction. In diffusive conductors that are small compared to the inelastic scattering length for both electron-electron (\(\ell_{e-e}\)) and electron-phonon (\(\ell_{e-ph}\)) interactions, theoretical calculations\[3, 4, 191, 132\] and experiments on different samples at helium temperature\[80\] agree that \(F \to 1/3\) in this regime. In the limit that the constriction length \(L\) exceeds \(\ell_{e-e}\), the expected reduction factor is predicted\[140\] to be \(\sqrt{3}/4 \approx 0.433\). Since both longitudinal and transverse dimensions in our junctions are comparable to the atomic scale, it seems likely that \(L \ll \ell_{e-e}, \ell_{e-ph}\), and our data show consistency with the prediction of 1/3. This consistency with a diffusive picture is a bit surprising, given that the length and transverse dimensions of the point contact should still be around 1 nm even when \(G \approx 10 G_0\). Experiments involving Pb junctions\[192\] have shown a similar transition to the diffusive regime even at relatively low conductances, though in that work this is ascribed to the importance of multiple orbitals per atom contributing to the conductance, leading to enhanced channel mixing. While this multiorbital mechanism should not apply for Au junctions, the lack of clear peaks in the conductance histograms (and corresponding suppressions in the noise) above \(3 G_0\) are consistent with enhanced channel mixing in our structures relative to that seen in other Au experiments\[53, 38\]. This is further supported by the observation mentioned in the previous paragraph that conduction in the tunneling regime in this data set appears to involve multiple poorly transmitting channels. It has been pointed out before\[80\] that the prediction of a suppression factor of about 1/3 is surprisingly robust. Investigations with different materials and over a broader range of temperature and bias should shed light on the extent of universality in this suppression.
4.2.3 Non-linear bias dependence of shot noise

We also consider the scaling of the excess noise at conductance values where the shot noise contribution is expected to be maximally suppressed. We have taken data sets of the type shown in Fig. 4.3 on multiple occasions with different tips, film samples, and cleaning/annealing procedures. To best study the bias scaling of the remaining excess noise near $1G_0$, we examine a data set when the suppression at $1G_0$ was particularly well defined. The result is shown in Fig. 4.5(a). With the comparatively strong suppression of shot noise power in this data set, we find that the measured noise vs. $X$ is nonlinear. This nonlinearity is only clearly seen at conductances where the noise is relatively suppressed.

A natural explanation for this is a contribution to the measured noise from conductance fluctuations, commonly termed “flicker” noise, which often has (in macroscale systems) a $1/f$ frequency dependence[115]. Such flicker noise, originating with fluctuations in the actual junction resistance, should scale quadratically with voltage across the junction at a given conductance, when recast as voltage fluctuations. Fig. 4.5(b) plots the excess noise spectral density vs bias voltage at the first two suppressions, with the noise power converted to mean squared voltage per unit frequency. Quadratic fits of the form $S_V = AV^2 + BV + C$ describe the data well, with a comparatively small linear term, $B$, as well as an even more tiny residual intercepts, $C$. With this fitting procedure we find $A \approx 4 \times 10^{-15}$ Hz$^{-1}$. We note that such a fitting procedure overestimates the size of $A$, since the hyperbolic cotangent term in Eq. 2.16 does contribute some nonlinearity in such a plot even when the only noise is finite temperature shot noise. Instead if we fit a quadratic in voltage term in addition to the finite temperature shot noise expectation of Eq. 2.16, we find $A \approx 1.4 \times 10^{-15}$ Hz$^{-1}$, smaller
but of the same order of magnitude. Using the phenomenological Hooge’s law[100],

\[ S_R(f)/R^2 = \frac{\alpha}{N f^\nu} \]  \hspace{1cm} (4.1)

with power \( \nu \) ranging between 1 \( \sim \) 2, we compared our inferred magnitude of \( A \) with values observed at lower frequencies and higher junction conductances in other gold point contacts by Wu et al.[115]. The result depends strongly on the assumed value of \( \nu \), which is expected to fall between 1 (traditional 1/f noise) and 2 (expected for
a single two-level fluctuator[115]). We find a flicker noise amplitude smaller than the Wu et al. values extrapolated to 1 $G_0$ with the fitted $\nu$ value in their case; however, if $\nu \approx 1.7$ over frequency from their sub-100 kHz measurements to our RF scale, our value for $A$ is compatible with their results. We can reasonably conclude that much of the nonlinearity in the noise vs. bias at conductances when shot noise is suppressed results from flicker noise/conductance fluctuations. The relatively small magnitude of this noise suggests a comparatively rapid decay of flicker noise with frequency in the RF range in atomic-scale junctions. It would be interesting to consider the temperature variation of the flicker noise magnitude over this frequency range, which would help clarify whether thermal activation of defect motion is relevant here.

4.2.4 Toy model calculation

It is worth considering what kind of mixing of quantum channels is required to produce the evolution of Fano factor with conductance that we infer from our data, within the noninteracting picture of Eq. 2.14. A simple “toy model” simulation of possible Fano factor evolution is provided, which depends on the choice of how the channels $\{\tau_i\}$ open as conductance is increased. The result is shown in Fig. 4.6 (a). We assume for simplicity a set of uniformly separated hyperbolic tangent functions for the form of $\{\tau_i\}$ as a function of $G$, which is shown on the bottom. Each hyperbolic tangent is specified by a central conductance at which $\tau_i = 0.5$, and a width in $G$ over which the channel “turns on”. The resulting Fano factors have been plotted on the top. Multiple colors/linestyles represent different choices of the central conductances and turn-on widths. In top graph, the uppermost line (blue) shows no clear suppression of noise vs. $G$ due to the heavy overlaps between channels. In contrast, the lower most curve (black) is in the limit of well-separated channels that do not mix. Channels’ mixture
Figure 4.6: (a) The expected Fano factor for the particular channel mixing shown in the second panel. (b) The measured Fano factor, and a possible channel openings shown in the bottom panel.

tends to smear out suppressions and raise the Fano factor, which is the qualitative reason that in a larger-sized diffusive conductor only a roughly constant Fano factor results. Fig. 4.6(b) shows an example (nonunique), of a set of transmission channels that agree reasonably well with the measurements. Additional constraints (such as those provided by subgap conductance in superconducting junctions[24, 190]) are necessary to better constrain the specific channels and their evolution.
We note that we do not see clear signatures of bias-dependent changes in the Fano factor in these ensemble averaged measurements. As the low temperature data in Kumar et al. shows[119], changes in the apparent Fano factor due to excitation of optical phonons can be of either sign in single-channel devices, with a crossover between enhancement and suppression of $F$ taking place at a particular transmittance, around 0.95. With respect to the Au optical phonon at $\sim 17$ meV, all of the data in this paper are in the high bias regime above the voltage threshold. Our measurements show that in Au junctions at room temperature, there are no clear inelastic thresholds above this energy that survive ensemble averaging over junction configurations. Other analysis approaches that look at subensembles may be more revealing[193]. (All the work in this section was published in [148])

4.3 Enhanced ensemble average at high biases

The non-interacting prediction of shot noise Eq. 2.16 is demonstrated working well in the previous section. After further increasing bias to higher levels, as shown in Fig. 4.7(b), significant nonlinearities as a function of scaled bias are present at all the conductance values, well above the extrapolated linear dependence found at low bias.(This work was published in [98]) As a comparison, Fig. 4.7(a) shows low bias linear fits at the same conductances away from the peaks in the conductance histograms. Different markers represent different independent data sets while each color indicates a particular conductance value. The solid lines are the low bias linear fits. In all the data sets, the nonlinearity is concave upward and independent of the order in which the voltages were applied (high to low; low to high; or interleaved). This indicates that the nonlinearity is an intrinsic function of the bias, not a result of irreversible changes to the junctions during the sequence of histogram acquisition
Figure 4.7: Noise as a function of scaled bias. a) At moderate bias $V < 180mV$. b) At high bias $V < 450mV$.

at different applied biases.

The high bias nonlinearity clearly deviates from the non-interacting prediction Eq. 2.16 and we must consider possible explanations for this trend in the noise at high bias. Possible contributors include: flicker noise; significant heating of the electronic reservoirs (region B in Fig. 1.7); electron-phonon inelastic corrections to the noise; and local electronic heating (in region C of Fig. 1.7).
4.3.1 Possibilities to be excluded: Flicker noise, reservoir heating and electron-phonon interactions

Flicker noise is produced by fluctuations in the device resistance due to scattering of the electrons by dynamical defects[116, 99, 100]. At low bias, such that the dynamical defects remain effectively at $T_0$, flicker noise manifests as a voltage noise power that scales quadratically in the applied dc bias. At higher biases, because of the availability of carriers as much as $eV$ above the local equilibrium Fermi level, the dynamical defects can have effective temperatures (of their ionic degrees of freedom) that are considerably elevated[74]. The result of this ionic heating is increased flicker noise as more defects are able to participate, and a bias dependence of the voltage noise that is super-quadratic[105, 109]. The operative question is whether the increased noise we see at high bias is indicative of this kind of enhanced flicker noise contribution.

We assess the role of flicker noise in multiple ways: considering the overall contribution of flicker noise to the total noise signal; examining the scaling of noise response with RF frequency range; and examining the expected dependence of noise on bias and junction conductance. The noise measurements in this work are broadband and in the rf regime, a frequency range (250 MHz to 600 MHz) considerably higher than that in which flicker noise is typically measured. The presence of strong noise suppression at $1 \, G_0$ at the highest biases establishes that the bulk of the measured noise likely results from shot noise of the type described at low bias by Eq. 2.16. We have also performed noise measurements with a different RF filter set (limiting the band to between 400 MHz and 800 MHz) at a variety of biases. While environmental backgrounds proved more annoying over that higher frequency band, the relative magnitude of the noise suppression near values of quantized conductance is essentially unchanged. This again is consistent with the significant majority of the total
measured noise being shot noise, since one would expect overall reduced flicker noise at the higher frequencies.

We look at the bias dependence of the noise. In the last section, when looking at $1 \, G_0$, where the shot noise contribution is maximally (but not completely) suppressed, we found residual nonlinearity with bias in the ensemble-averaged noise power at low bias that was roughly consistent with a flicker noise contribution. We repeat such an analysis here. Note that the superlinear behavior in Fig. 4.7(b) is as a function of scaled bias, not just $V$. Fig. 4.8 shows the measured noise (converted to voltage noise, $S_V$) as a function of bias across the junction, $V$, for $G = 3G_0$. The data are well described as a linear + quadratic term. There is no clear evidence of any superquadratic dependence of the type expected for high bias flicker noise.

Finally, we consider the expected dependence of noise with the junction conductance. We can be overly conservative and ascribe all of the nonlinearity in $S_V$ as a function of $V$ to flicker noise. This knowingly overestimates the magnitude of flicker noise at $1 \, G_0$, since the true finite-temperature expression for shot noise (Eq. 2.16) has some nonlinearity due to the coth($eV/2k_BT$) factor. Previous experimental work at lower frequencies (< 400 kHz)[115] used lithographically defined mechanical break junctions at cryogenic temperatures to examine flicker noise down to conductances comparable to $G_0$ in individual junction configurations. In those experiments, for a fixed bandwidth, the investigators found that the ensemble-averaged flicker noise scales like $S_V/V^2 \propto G^{-1.5}$ as junctions approach the ballistic regime, though there is variation within the ensemble. In Fig. 4.8, we compare the bias-dependent noise at $3 \, G_0$ with that at $1 \, G_0$ (dashed line). Contrary to the ensemble-average scaling expectation for flicker noise, we find that the observed nonlinearity at $3 \, G_0$ is nearly two times larger than the overestimate of the $1 \, G_0$ flicker noise. While large varia-
Figure 4.8: Excess noise at 3 $G_0$ converted into voltage noise power $S_V$ as a function of bias. The solid blue line is the best quadratic fit to the experimental data, while the dotted blue line is a systematic overestimate (see text) of the contribution of flicker noise at 1 $G_0$. Any flicker noise contribution at 3 $G_0$ is expected to be even smaller. Thus, flicker noise is unlikely to be a compatible explanation for the observed increase in noise as a function of $V$.

A second possible explanation for the observed data that must be considered...
is heating of the “reservoir” electrodes due to current flow. In prior experiments by Henny et al.\cite{80} considering shot noise at cryogenic temperatures in a diffusive wire connected to thick film reservoirs, the heating of those reservoirs was definitely non-negligible. In the present experiment, the situation is considerably different: an atomic-scale junction in what is usually considered the ballistic regime, at room temperature, with most of the dissipation happening in effectively bulk reservoirs. The Au tip has a measured electrical resistance of less than 1 Ω, and the sheet resistance of the Au film is similarly around 2 Ω. The question is whether power dissipation near the junction could, through the finite thermal resistance of the electrodes, lead to significant heating of the electrodes (both electrons and phonons) near the junction. To be very conservative, assume that only the electronic component of the thermal conductivity is relevant. Furthermore, assume that 100% of the dissipated power is deposited in either one of the electrodes proximal to the junction. For 300 mV bias across a 3 kΩ junction, the power $P = 3 \times 10^{-5}$ W, and this must correspond to $(T_{res}^2 - T_0^2)LG_{res}/2$, where $T_{res}$ is the local temperature of the reservoir electrode near the junction, $T_0$ is the environmental temperature, $L = 2.44 \times 10^{-8}$ W Ω/K$^2$ is the Lorenz number, and $G_{res}$ is the conductance of the reservoir, in this case one siemens. For $T_0 = 300$ K, we find $T_{res} = 304$ K. In the real situation, when roughly half of the power will be dissipated in the other electrode\cite{89} and phonons will contribute significantly to the thermal path in both electrodes, the actual local temperature rise of the bulk electrodes (region B in Fig. 1.7) is expected to be even smaller. Note that this simple estimate based on the electronic thermal conduction is independent of the detailed electrode geometry. Therefore, we conclude that this kind of reservoir heating, while definitely important in the cryogenic experiments of Henny et al.\cite{80}, is unlikely to be a significant contributor to the high bias noise increase in the present
work.

At the biases considered in this work, and at room temperature, it is important to consider the possible effects of electron-vibrational coupling on the measured noise. As mentioned previously, there are many theoretical predictions describing electronic transport in small, ballistic junctions coupled to a local vibrational mode. [194, 149, 195, 143, 151, 152, 153] In an individual junction one expects a modification in the Fano factor of the shot noise as the bias exceeds the characteristic energy of the vibrational mode. The change in $F$ depends sensitively on the particular channel transmittance[195, 143, 153, 119], as well as the coupling of the channels to the phonon degree of freedom.

A detailed comparison with particular theoretical descriptions is difficult: Our data is in the ensemble average, generally with multiple contributing channels; if one expects the Au optical phonon to be a dominant contribution, our data is acquired at the limit $k_B T > \hbar \omega$ so that finite thermal phonon population is possible. A detailed consideration of ensemble averaging over channel distributions would have to match the observed increase in enhanced noise at higher conductances. However, the observed power law dependence of the excess noise, $\sim V^2$, experimentally constrains possible phonon-based mechanisms for the increase in noise at high bias. Descriptions that assume strong pumping of the local phonon population by the electrons predict voltage dependences with $V^3$ and $V^4$ contributions[152, 153]. For vibrational effects to be responsible for the enhanced noise, the ensemble-averaged system must be in the “equilibrated phonon” limit, such that the local phonon population $\bar{N}$ remains bias independent[153]. This suggests an experimental test for this mechanism. By performing such measurements as substrate temperatures are decreased down to the cryogenic regime, one would expect decreases in the local phonon population, and
therefore decreases in such nonlinear corrections to the noise.

4.3.2 Possible explanation: Local electronic heating

Given that temperature gradients within the bulk of the electrodes are small, an additional possible explanation for the observed trends in the measured noise would be some amount of effective electronic heating local to the junction (that is, electronic distribution functions impinging locally on the ballistic region (region D) that evolve nontrivially with bias for reasons other than bulk heating of the reservoirs (region B)). For example, the single-particle scattering formalism neglects electronic heating due to the electronic viscosity, inherently an effect of electron-electron interactions.[97, 78] Before addressing this specific proposal, we attempt to assess whether the observed trend could be consistent with some broadened electronic distribution function (the idea that \( T_e \) describing the components of the nonequilibrium distribution function in region C of Fig. 1.7 is bias-dependent). To do this, we parametrize the true nonthermal carrier distribution at the junction location through an effective electronic temperature, \( T(V) \), that depends on the bias, appearing in Eq. 2.16. Obviously, a general nonequilibrium distribution does not have to be parametrized by a single effective temperature; however, some kind of parametrization is necessary to perform an analysis, and this is closely related to the approach of Henny et al. when considering “traditional” electronic heating. We employ the Landauer transmittance expression and the resulting Eq. 2.16, but assume that \( T = T(V) = T_0 + \delta T(V) \), where \( T_0 \equiv T(V = 0) \) is the ambient temperature, \( \approx 300 \) K. To further simplify the analysis, we leverage the fact that \( eV \gg 2k_B T_0 \) in the high bias regime of interest. We can then move forward approximating \( \coth \frac{eV}{2k_B T} \to 1 \), which we will check for self-consistency (to make sure that this continues to hold for the inferred \( T(V) \)).
Thus Eq. 2.16 for a single junction may be reduced to

\[ S_I(V) = G_0 [4k_B(T_0 + \delta T) \sum_i \tau_i^2 + 2eV \sum_i \tau_i(1 - \tau_i)] \] (4.2)

And at zero bias, we already know that

\[ S_I(V = 0) = G_0 [4k_B T_0 \sum_i \tau_i^2 + 4k_B T_0 \sum_i \tau_i(1 - \tau_i)] \] (4.3)

which is purely Johnson-Nyquist noise. Solving for \( \delta T(V) \),

\[ \delta T(V) = \frac{S_I(V) - S_I(V = 0)}{4k_BG} - \left( \frac{eV}{2k_B} - T_0 \right) F \sum_i \tau_i \sum_i \tau_i^2 \] (4.4)

where \( S_I(V) - S_I(V = 0) \) is exactly the excess noise we measure. The ensemble-averaged Fano factor, \( F \), can be determined from the low bias data described above.

The only missing information is the ratio \( \sum_i \tau_i / \sum_i \tau_i^2 \). For our ensemble-averaged measurements, by specifying \( G \) we then need to find the ensemble average \( \langle \frac{G}{G_0 \sum_i \tau_i^2} \rangle \). While in a single junction with Fano factor \( F_0 \), one can write the parameter of interest as \( 1/(1 - F_0) \), note that in general one cannot assume that \( \langle \frac{G}{G_0 \sum_i \tau_i^2} \rangle = 1/(1 - F) \), where \( F \) is the ensemble average.

We can consider models of channel mixing to illuminate possible values, as shown in Fig. 4.9. The dotted line represents the Fano factor, the solid blue line \( \sum_i \tau_i^2 \), and the solid red line \( \sum_i \tau_i / \sum_i \tau_i^2 \). Fig. 4.9(a) represents one limiting case, no channel mixture at all, with each channel turning on linearly sequence as the conductance is increased. Fig. 4.9(b) assumes much stronger mixing of the channels, so strong that there would be little detectable shot noise suppression in this limit. In both limits, the ratio of interest approaches a constant (one) quickly after the conductance is increased above \( 1 \, G_0 \). As the diffusive limit is approached,[80] \( F \to 1/3 \) because of partially open channels, meaning \( \sum_i \tau_i / \sum_i \tau_i^2 \to 1.5 \). In either case, it is safe to approximate \( \sum_i \tau_i / \sum_i \tau_i^2 \) as a constant on the order of 1 when \( G \) is much above \( 2G_0 \) (though the precise value
introduces a systematic uncertainty in what follows). Within this approximation, we can evaluate Eq. (4.4), and the results are shown in Fig. 4.10.

The effective electronic temperature elevations in Fig. 4.10 are calculated from the data sets used to produce Fig. 4.7(b). In this figure, 13 conductance values are chosen, representative of conductances higher than $G_0$. The approximation $\coth \frac{eV}{2k_B T} \to 1$ is only valid outside the shaded region. At a bias of several hundred mV, within this
Figure 4.10: Within the model of a locally elevated electronic temperature at the ballistic junction, the inferred electronic temperature from Eq. (4.4) vs bias voltage. For data sets at 13 different conductances $3, 3.25, 3.5, 3.75, 4, 4.27, 4.5, 4.75, 5, 5.25, 5.51, 5.73, 6 \ G_0$, the colors/markers, respectively, are: blue point, red point, green point, purple point, khaki point, blue star, red star, green star, purple star, khaki star, blue open square, red open square, green open square.

model an effective electronic temperature elevation as high as $350 \text{ K}$ is inferred. The error bars are dominated by the uncertainty in $F$, as inferred from $\chi^2$ of the linear fits in Fig. 4.7(a). These local electronic temperature elevations are the same order as those inferred through completely independent experiments employing indirect optical methods[84] and bond rupture measurements[72].

We use a deliberately simplistic model to check the reasonableness of these $\delta T$ values. The inferred effective temperature elevations are approximately independent...
of conductance, which is consistent with a local electronic heating picture assuming a balance between local power generation and heat current carried by the electrons.[97] In this local heating picture, some fraction $\alpha$ of the total Joule heating is dissipated local to the junction, $P_{in} = \alpha IV = \alpha V^2 G$. (In the standard approach, $\alpha$ would be zero; given the general success of the standard approach, the expectation is that $\alpha$ should be small compared to one.) Since the electron-phonon inelastic mean free path is much longer than the junction size, the electrons must carry away the heat generated. In the limit of local quasithermal energy distributions (so that a local electronic temperature is well defined) for the outgoing carriers, the electronic thermal conductance is $LTG$. Balancing the Joule heating with the thermal current $P_{out} = (T^2 - T_0^2)LG/2$ leads to a predicted local temperature of $(T_0^2 + 2\alpha V^2/L)^{1/2}$. The observed data are roughly consistent with this for $\alpha \approx 0.02$, not bad considering the simplicity of the model, which assumes the Wiedemann-Franz law and no dependence of the $\tau_i$ on bias. The inferred $\delta T$ seems to increase more rapidly than this dependence at higher bias; this simple model may be inadequate, assuming for example that the transmittances $\tau_i$ are independent of bias (generally not a poor assumption for Au over this energy range[196, 89]).

We see that the observed noise trend could be explained by a relatively modest amount of electronic heating local to the junction (dissipating 2% of the total power in region C of Fig. 1.7, though in general one would expect $\alpha$ to depend on the conductance). We now consider whether there is a physical mechanism that could contribute this amount of local electronic heating. Indeed, similar to the local, nonequilibrium ionic heating due to electron-phonon interactions, the internal Coulomb friction of the electron liquid is predicted to create local electron heating [97]. Due to the large current densities in nanojunctions compared to their bulk counterparts the electron-
electron scattering rate increases locally in the junction [73]. The underlying Fermi sea thus “heats up” locally due to this increased scattering, via production of electron-hole pairs whose energy needs to be dissipated away from the junction. This would manifest itself in the noise just as an elevated electronic temperature does so in the regime of Henny et al. [80]. The local effect was first computed in Au nanojunctions [97] assuming no dependence of the electronic viscosity on the junction geometry, namely on the electron transmission. Using the viscosity of that work, namely $10^{-7}$ Pa·sec for the electron liquid of density comparable to gold ($r_s = 3$), we obtain an estimate of $\delta T$ that is about two orders of magnitude smaller than what is shown in Fig. 4.10.

The drawback of neglecting the dependence of viscosity on junction geometry was however understood and corrected in Roy et al. [78], where the zero frequency viscosity $\eta(0)$ was evaluated as the product of the high-frequency shear modulus $\mu_\infty$ and a junction-specific momentum relaxation time $\tau$. The actual scattering potential used in Ref. [78] to compute $\tau$ was chosen precisely to model Au quantum point contacts. By referring to Fig. 4 of Ref. [78] we see that $\eta(0) \sim 10^{-4}$ Pa·sec for the transmissions typical of these point contacts. Since the zero-frequency viscosity enters as the square root of its value in the effective electronic temperature (see Ref. [97]) we obtain a value of $\alpha = 0.05$, a factor 2.5 higher than the one fitted in Fig. 4.10, a very reasonable agreement considering that the actual geometry of the junction is not known.

This quantitative consistency between the modeled experiment and this theoretical treatment suggests that the electronic viscosity should definitely be considered as a potential contributor to the observed enhanced noise. Within this theoretical picture the effective electronic heating depends strongly on the transmittance, implying that such heating would be much less important in lower conductance structures such as molecular junctions, where it might be detected through inelastic electron tunneling.
spectroscopy linewidths. We further note that this electron-electron interaction effect should be comparatively independent of the background equilibrium temperature. Therefore, measurements of the scaling of noise at high bias at lower substrate temperatures should allow the discrimination between electron-vibrational enhancement of the noise discussed above and this local electronic heating mechanism.

4.4 Beyond ensemble average

In most research works regarding the atomic-scale and molecular junctions, including the previous sections in this chapter, statistical analyses are usually conducted to average out the microscopic variability of junction configurations. For instance, conductance histograms are a common tool to study the preferred conductance values, and hence preferred sets of discrete quantum channel transmittances, averaged over accessible atomic configurations. The obvious shortcoming is that information associated with each specific atomic arrangement is lost in the averaging.

To go beyond this ensemble-averaged analysis, considerable efforts have been made in recent years. Measurements of multiple Andreev reflection reveal the transmittance of each specific quantum channel in individual superconductor junctions. From ensembles, the range of mathematically allowed transmittances may be estimated. In the mean time, the development of new analysis techniques also enables more information to be extracted from the mountain of data over all the ensembles. These methods include “density plots”, where every experimental data point in the ensemble is counted. Examples are density plots of conductance vs. elongation, current-voltage characteristics, and mechanical stiffness vs. elongation. Two-dimensional (2D) cross-correlation methods also take advantage of entire ensembles, revealing detailed information about motifs of
junction formation [206, 193, 207].

Over the same time period, there have also been many advances in computational techniques, particularly the ability to combine MD with quantum transport to describe the inherent interplay between mechanical and electrical properties in these systems, which is crucial to establish a direct comparison with the experimental results [208, 209, 210, 211].

In this work (published in [189]), we analyse conductance and shot noise measurements over whole ensembles of STM-style gold break junctions at room temperature, expanding upon prior ensemble-averaged treatments [147, 148, 98]. When mapping out two dimensional density plots of shot noise vs. conductance, we are able to find the standard deviation (over the ensemble) in the noise at each conductance, $G$. A state-of-the-art calculation combining MD, an accurate tight-binding model, and Green’s function techniques to examine junction formation is compared to these experimental results. As we discuss below, surface contamination and disorder in the electrodes are likely responsible for many of the differences between the experiment and the calculations.

The theory part was done by M. Matt, F. Pauly, P. Nielaba from Department of Physics, University of Konstanz, D-78457 Konstanz, Germany, and J. C. Cuevas from Departamento de Física Teórica de la Materia Condensada and Condensed Matter Physics Center (IFIMAC), Universidad Autónoma de Madrid, E-28049 Madrid, Spain.

4.4.1 Generation of 2D density plots

The measurements are conducted across STM-style gold break junctions at room temperature in air. As described in the past sections, a gold tip attached to an end of one piezo actuator is electrically controlled, moving towards and away from
an evaporated thick gold film, allowing the cyclic formation and breaking of gold junctions. The gold tip is controlled to obtain approximately one trace per second. Conductance and noise power from lock-in outputs are acquired at $10^5$ samples/s, but with the time constant on each lock-in set to $100 \mu s$. In each trace, the time-averaged (with sub-millisecond averaging periods) shot noise power as a function of conductance is computed. Both noise power and conductance axes are binned. The junction conductance is calculated from the measured current and applied voltage bias, taking into account series resistances in the measurement circuit. The statistical uncertainty in the conductance data is less than 1%. The statistical uncertainty in the measurement of the rf power is at a similar level, though the background contributes a systematic uncertainty to individual measurements. The number of counts in each bin $(G, P)$ is used to construct a 2D trace density plot.

Typical traces have from five to more than ten thousand counts between $G = 4 G_0$ and $G = 0.01 G_0$. Traces with less than 1000 counts are not included as these indicate anomalously rapid breakage or a measurement problem; varying this cutoff from 500 counts to 2000 counts produces no noticeable change in the resulting plots or analysis. For clarity of color scale when plotting the trace density, at each conductance the bins are normalized to show the relative probability (between 0 and 1) of measuring a particular noise power.

In the analysis, a subtle technical issue arises regarding the rigorous extraction of the shot noise from raw data. The noise power detector has small random fluctuations in addition to the true target signal, causing a small but non-negligible positive background in the lock-in amplifier measurement of its amplitude. When performing a full ensemble average analysis, statistical methods [148] can be applied to remove this small background, as we have done previously. However, this background re-
moval method does not work rigorously at the single trace level; thus, in this work this background remains. Under this circumstance, there is therefore a slight systematic overestimate of the true shot noise spectral density, especially relevant when shot noise is small. The background is independent of $G$ and is equivalent to approximately $0.15 \times 10^{-24} \text{ A}^2/\text{Hz}$, and sets the “floor” of the data shown in Fig. 4.11, panels (c) and (d).

Figure 4.11: Noise and its variation across ensembles. (a) Conductance histogram acquired with 180 mV bias (816 traces). (b) Conductance histogram acquired with 120 mV bias (807 traces). (c, d) 2D density plots for the respective ensembles, with data at each conductance normalized to show the probability for finding a particular shot noise power value at any particular conductance. The black curves show the ensemble-averaged noise power. (e, f) The associated variance of shot noise at each conductance. The vertical black lines are guides to the eye to indicate the conductance for which the shot noise variance is a
Fig. 4.11 shows the resulting density plots obtained at biases of 180 mV and 120 mV, respectively. The top-most panels are the conductance histograms, commonly used to find the preferred conductance values. The middle panels are the trace density plots, where warmer colors indicate higher densities of traces. The black curve indicates the ensemble-averaged excess noise power as a function of conductance. From this panel, clearly the distribution of shot noise has some interesting features. The “envelope” of the traces is comparatively large away from the conductances where shot noise is suppressed, while that envelope becomes significantly concentrated at those suppression regions near integer multiples of $G_0$.

To quantify this distribution, the variance $\langle S_I^2 \rangle - \langle S_I \rangle^2$ at each conductance is computed and shown in panels (e, f). Note that the shot noise variance minima are coincident with the conductance peaks and noise suppressions. Ensembles acquired at many other biases, or with a different radio frequency bandwidth, reproduce the same features. According to Eq. 2.14 and 2.16, shot noise suppressions originate from the fully transmitting channels. It is also widely known that shot noise suppressions in gold junctions coincide with the conductance peaks [53], indicating a potential relation between the fully transmitted channels and one or a few preferred atomic arrangements existing at the conductance plateaus. There is no obvious theoretical explanation for these shot noise variance minima. We infer that somehow $\{\tau_i\}$ has limited variations at these preferred conductance values compared to higher or lower conductances. Either relatively fewer atomic configurations are allowed at the conductance plateaus, or the allowed atomic configurations have relatively similar transmittance sets $\{\tau_i\}$, or both.

The density plots also reveal information about the nature of the shot noise suppression at the conductances indicated by the peaks in the conductance histograms.
At cryogenic temperatures, theory and experiments agree very well that shot noise in gold junctions is suppressed down to near zero at conductance plateaus [50, 5], particularly when \( G \approx G_0 \). This indicates a single fully transmitted channel at \( 1G_0 \) and some limited channel mixture at the higher \( G \) plateaus. In our ensemble-averaged measurements, all the suppressions are only partial [148]. The density plot reveals the explanation: At \( 1G_0 \), most traces do show nearly complete suppression, while a small fraction relatively elevated noise. At other plateaus, though the lower count numbers and sparse distributions make the situation less obvious, the most populated shot noise values are also lower than the ensemble average. These observations show that the channel mixing responsible for affecting the depth of the noise suppression takes place only in a subset of the full ensemble. This is consistent with the channel mixing resulting from either particular realizations of disorder in the electrodes or surface contamination. The peaks in the conductance histograms are consistent with deviations from integer quantization in Au junctions previously ascribed to work hardening (and therefore disorder in the electrodes) [38].

The nonzero shot noise variance in the nominal tunneling regime also indicates a break-down of the single-channel picture below \( 1G_0 \) in these ambient condition experiments. In single channel transport, the summation in Eq. 2.14 and 2.16 reduces to \( \tau(1 - \tau) \) while total conductance becomes proportional to \( \tau \) of the only channel. The resulting Fano factor is \( 1 - \tau \), and at a given \( G \) its variance is zero. Thus a nonzero noise variance when \( G < G_0 \) clearly indicates the participation of multiple quantum channels. The diffusive-like features at high conductance reported in our former paper also indicate that the channel mixture in the high conductance region in our measurements is stronger than commonly expected [148].

A debatable question here is whether the variance \( \langle S_I^2 \rangle - \langle S_I \rangle^2 \) is a fair way to
Figure 4.12: Fractional variance of the noise power as a function of conductance, using the same data as in Fig. 4.11.

examine shot noise’s variation over ensembles. In Fig. 4.12 we plot the fractional variance \((\langle S^2 \rangle - \langle S \rangle^2)/\langle S \rangle^2\), using the same data as in Fig. 4.11. After this normalization, the fractional variance as a function of conductance tends to be relatively featureless or even have maxima at the preferred conductances. This suggests that the variance in the noise is approximately proportional to the square of the noise itself, though there is no obvious explanation for this. The nonzero noise variance
4.4.2 Simulation and discussion

This section is devoted to the description of the theoretical methods that we employed to shed some light on our experimental results for the shot noise of Au atomic-size contacts. The theory works are all done by our theoretical collaborators, and we discuss the results together.

For this purpose, we combined classical MD simulations of the formation of atomic-size contacts with a tight-binding description of the electronic structure to compute the transport properties with the help of Green’s function techniques. Our approach follows closely Refs. [208, 209, 212, 211], and we now proceed to explain it in some detail.

In these atomic wires there is an intricate relation between the mechanical and the transport properties. For this reason, and in order to establish a direct comparison with our experiments, it is necessary to describe the formation process of these atomic-size contacts. For this purpose, we carried out classical MD simulations using the open source program package LAMMPS [213, 214]. These simulations are based on the so-called embedded atom method and, in particular, we employed the empirical potentials from Ref. [215]. Let us emphasize that these potentials overcome several problems of two-body potentials like, for instance, the coordination independence of the bond strength. This is important for our calculations because in our atomic contacts we have, in particular, regions with low coordination number. To generate the geometrical configurations, we start with an ideal face-centered cubic lattice, where the crystal direction $\langle 100 \rangle$ lies parallel to the $z$ axis, which corresponds to the transport and elongation direction. For the MD calculations, we divided the
Figure 4.13: Example of our simulations of the stretching of a Au wire at 300 K. (a) Conductance, shot noise, and Fano factor as a function of the elongation of the wire. (b) Individual transmission coefficients as a function of the elongation. The upper panels show the wire geometries at different stages of the elongation process. The left geometry corresponds to the initial configuration of the simulated wire.

We divided the wire geometry into three parts: Two electrodes connected by a central wire (see Fig. 4.13). The electrodes consist of 661 atoms each and they are kept fixed during the MD calculations. On the other hand, the wire is made up of 563 atoms which follow the Newtonian equations of motion. We assume a canonical ensemble and use the velocity Verlet integration scheme [216]. The wire has an initial length of 0.83 nm and the starting velocities of the atoms in the wire were chosen randomly with a Gaussian distribution to yield the desired average temperature. As we discuss below, we performed simulations at room temperature ($T = 300$ K), but also at cryogenic temperatures ($T = 4$ K) to compare with previously published results. Because of the
randomness in the initial velocities, every elongation process evolves differently, while a Nose-Hoover thermostat keeps the temperature fixed [216]. To relax the system, the wire was equilibrated for 0.1 ns. Finally, the elongation process is simulated by separating the electrodes at a constant velocity of 0.4 m/s. During this process, every 10 ps the geometry is recorded. A stretching process needs a total simulation time of about 4.5 ns, until the contact breaks.

Once the geometries of the atomic wires were determined through the MD simulations, we used them to compute the conductance and the shot noise in the spirit of the Landauer-Büttiker formalism. As explained in the introduction, within this formalism the transport properties are fully determined by the set of transmission coefficients \( \{ \tau_i \} \). We computed these coefficients by combining an appropriate description of the electronic structure of the Au wires with non-equilibrium Green’s function techniques following a standard recipe that we have explained in detail elsewhere [208, 217, 218]. The electronic structure was described within the framework of the Slater-Koster tight-binding parameterization of Ref. [219], which has been quite successful in determining a variety of properties of these atomic wires [218]. To compute the transmission coefficients with the help of the Green’s function techniques, we divided the system into three regions as in the MD simulations, \( i.e. \) the upper and lower electrodes and the central wire. The electrodes were considered to be semi-infinite perfect crystals. Their surface Green’s functions were computed with the help of a decimation technique [220, 217], using the same tight-binding parameterization as for the central part to determine their electronic structure. It is worth stressing that we enforced charge neutrality for all the atoms of the wire, which is a condition that is typically fulfilled in metallic systems [221].

Finally, to accumulate sufficient statistics for our study, we carried out 100 MD
simulations of the breaking of the Au atomic wires and checked that this number is sufficient, especially, to converge the conductance histograms. We show in Fig. 4.13 a typical example of the simulation of stretching of a wire at room temperature. In particular, we show the different physical properties of interest in the last stages of the breaking of the wire, namely the conductance, shot noise, and Fano factor. Additionally, we show the individual transmission coefficients $\tau_i$ as a function of the elongation.

Our main results are summarized in Fig. 4.14, where we show the conductance histograms, density plots of the shot noise, and the noise variance as a function of the conductance for 4 and 300 K. The conductance and the noise were computed in the linear response regime, i.e. using the zero-bias transmission coefficients. Moreover, for the noise, we focus on the shot noise contribution to the excess noise by merely plotting the quantity $2eG_0V\sum_{i=1}^N \tau_i(1-\tau_i)$ in equation (2), which ignores the contribution of the thermal noise. This is justified as $k_BT$ is much smaller than $eV$ in our experiments. For comparison with the experiments we have set the bias voltage in the prefactor to 180 mV. As one can see in panels (a) and (b), the conductance histograms are dominated by a pronounced peak at $1G_0$, which is much higher at low temperatures. This is due to the fact that this peak partially originates from the formation of monoatomic chains [39, 34]. These chains exhibit a conductance very close to $1G_0$ and we find that their formation is clearly less favorable at room temperature. Notice also the appearance of another pronounced peak at $2G_0$ at low temperatures, which is to a large extent washed out at room temperature. Turning to the noise (Fig. 4.14(c,d)), the most noticeable feature is its suppression close to multiples of $G_0$, which is specially pronounced around $1G_0$ and at $T = 4$ K. This is consistent with the measurements reported above and supports the assumption that
the channel mixture is stronger outside the tunneling region at higher temperatures. On the other hand, the variance of the shot noise (Fig. 4.14(e,f)) generally exhibits maxima correlated with the maxima in the conductance histograms. While this is consistent with the low-temperature results reported by Vardimon et al., this is clearly at odds with our experimental results (Fig. 4.11(e,f)). Notice also that the variance of the noise does not vanish below $1G_0$, which clearly suggests that several channels are contributing in that region.

The origin of the results just described can be clarified by analysing both the distribution of transmission coefficients, $\{\tau_i\}$, and the Fano factor and its corresponding variance. This information is displayed in Fig. 4.15. The channel distributions shown in panels (a) and (b) indicate that the conduction channels open one by one as the conductance increases, as reported in Ref. [24] with the help of proximity-induced superconductivity in Au atomic-size contacts. This is a typical behavior of a monovalent metal. Focusing on the low temperature results, we find that although the conductance region $G < G_0$ is largely dominated by a single channel, a second one gives a sizeable contribution. This additional channel manifests as a kink in the Fano factor around $0.5G_0$ and as a pronounced maximum in the Fano factor variance at that position. It is worth stressing that the presence of this second channel does not “spoil” the conductance quantization and the histogram still exhibits a very pronounced peak at $1G_0$. For higher conductance, the Fano factor exhibits a partial suppression close to multiples of $G_0$, while its variance has maxima correlated with the maxima of the conductance or the minima of the shot noise. All these findings at 4 K are in very good agreement with the recent experimental results of Refs. [117] where shot noise was measured at cryogenic temperatures and used to extract the channel distribution of Au few-atom contacts. Turning now to the room temperature
results, one can still see a small contribution of a second channel for $G < G_0$, which explains the non-vanishing variance of both the shot noise and the Fano factor for these conductances. However, at 300 K the weight of this channel is not sufficient to produce a maximum in the variance of these quantities below $G_0$. For higher conductance, the variance of the Fano factor exhibits again shallow maxima correlated with the maxima of the conductance and the minima of the shot noise.

From the previous discussion, a natural question arises: What is the origin of the second channel that appears in some geometries in the tunnel regime? To answer it we analysed systematically those contact geometries where at least two channels contribute significantly to the transport properties in the conductance region $G < 0.5G_0$. Some representative examples are shown in Fig. 4.16. We found that these geometries can be grouped into two categories. In the first one, the electrodes contain several neighbouring atoms that are at similar distances to the atoms in the other electrode, see Fig. 4.16(b). At room temperature, practically all the geometries with two channels in the tunnel regime belong to this kind. On the other hand, at 4 K we also find geometries where the contact breaks at several points giving rise to parallel junctions, see Fig. 4.16(a). In both types of geometries the two channels originate from parallel paths with the peculiarity that in the second kind these paths are basically independent, i.e. they do not interfere. Interestingly, we find that the geometries with several parallel junctions are also responsible for the peak at $2G_0$ that appears in the low temperature conductance histogram.

Clearly our theoretical simulations have difficulties to reproduce the observed shot noise variance minima, and the robust quantization peaks and noise minima at $G > G_0$ at room temperature. The origins of these discrepancies are not clear at the moment. One possibility might be related to the unavoidable limitations in our MD
simulations. In particular, one has to bear in mind that the stretching velocity in these simulations is many orders of magnitudes higher than in the experiments. Moreover, because of computational reasons we have to constrain ourselves to relatively small contacts. Although we have investigated these issues systematically within our capabilities and have not found any significant difference in our results upon changes in the stretching velocity, we cannot rule out that this discrepancy is due to systematic deficiencies in our simulations. However, it is worth remarking that our low temperature simulations are in very good agreement with published data in UHV conditions [117] and similar simulations have been very successful in reproducing, for instance, the thermopower of different atomic wires [212] or the channel distribution of Al atomic-size contacts at very low temperatures [211].

Another possibility for this discrepancy might be the fact that these experiments are conducted in air. Though the related environmental parameters such as pressure, temperature or humidity should be relatively stable at each set of measurements, contamination from adsorbates (water molecules for instance) cannot be ruled out. Molecules randomly attached to the junction can surely alter the dynamics and stability of the junctions and the conductance as well, especially in the last stages of the breaking process. Both hydrogen [222, 223] and oxygen [224] are known to incorporate into junctions and alter the available stable atomic configurations. This possibility can be tested by further room temperature experiments in a UHV environment. On the other hand, the disorder of gold junctions may also be relevant [38], which is proven to have effect on peak positions in conductance histogram. Dislocations and other defects introduced by “work hardening” can in principle affect atomic mobility and therefore structural stability, physics very difficult to capture in computationally tractable simulations.
Figure 4.14: (a,b) Conductance histogram obtained from 100 MD simulations at 4 and 300 K, respectively. (c,d) The corresponding 2D density plots for the shot noise. The black solid lines indicate the ensemble-averaged values. (e,f) The associated variance of the shot noise as a function of the conductance. The bias voltage for the calculation of $S_I$ is set to 180 mV. The vertical lines are at the same position as in Fig. 4.11.
Figure 4.15: (a,b) Channel distributions corresponding to the results of Fig. 4.14 at temperatures of 4 and 300 K, respectively. (c,d) The corresponding 2D density plots for the Fano factor. The black solid lines indicate the ensemble-averaged values. (e,f) The associated variance of the Fano factor as a function of the conductance.
Figure 4.16: Two representative examples of contact geometries where two conduction channels give a significant contribution in the tunnel regime ($G < 0.5G_0$). The total conductance and the individual transmission coefficients in the different examples are (where we consider only channels with $\tau_i > 0.01$): (a) $G = 0.29G_0$, $\tau_1 = 0.28$, $\tau_2 = 0.11$, $\tau_3 = 0.01$ and (b) $G = 0.15G_0$, $\tau_1 = 0.09$, $\tau_2 = 0.05$, and $\tau_3 = 0.01$. 
Chapter 5

Shot noise in Conventional Au MCBJs at multiple cryogenic temperature

5.1 Device fabrications and experimental methods

To conduct measurements over MCBJs using our home-built apparatus, suspended nano-devices need to be fabricated on top of some flexible but insulating substrates. Our fabrication procedure closely follows the method developed by van Ruitenbeek et al. [184].

Stainless steel is chosen to be the substrate due to its flexibility, with polyimide (PI-2610) film coated as the insulating layer. Before the polyimide coating procedure, a rough polish is needed. An Allied MultiPrep polishing system is used and 1200-grit level roughness is good enough. Sub-micron suspensions could be used to achieve much finer and nicer surface, but it needs longer time and not necessary as the polyimide coating will further smoothen the surface. The polyimide film coating procedure has also been introduced in Leland Richardson’s thesis [225]. In his recipe VM651 is coated first to enhance the adhesion between polyimide and the substrate, which turns out to be not necessary. The current recipe is therefore simplified:

1. Basic cleaning with IPA and acetone.
2. Hot plate bake at 180° for 75 s.
3. Apply a few drops polyimide (PI-2610) and cover the whole chip. Spin coat
500 rpm for 10 s; then 4000 rpm for 1 min 25 s.

4. Hot plates bake at 90° for 90 s, then 150° for 90 s.

5. Repeat step 3 and 4 a few times to coat multiple layers of polyimide films.

6. Vacuum bake using the tube furnace. Temperature raise by 4°/min till 300°, then keep 30 min. Finally naturally cools down.

A critical step after polyimide coating, is to apply an Ar plasma treatment (high power level for 3-4 mins) to enhance the adhesion between polyimide and any metal layers, by a Harrick pdc-32g plasma cleaner. Without this there is no way to bond Au wire onto the contact pads without demaging the contacts. This step in principle can be applied before defining the big contact pads, but it will also etch away polyimide film at a rate ~ 50 nm/min and cause contact problems if there is any structure already defined. As a result I always conduct this step after polyimide coating, which makes the surface rougher but turns out not to be a big deal in the following steps.

Arrays of bowtie junctions are defined by e-beam lithography and e-beam evaporation (Ti/Au). The narrowest part of a typical bowtie junction is a Au bridge of ~150 nm wide and 20 nm thick, which typically has a resistance ~ 70 Ω at room temperature. Device thinner than 20 nm may suffer from severe blow up issues. (Devices thicker than 20 nm are harder to break by electromigration, so I normally just use 20 nm thickness.) After the lift-off step, thick Cr/Au pads and leads connecting the junctions are evaporated making use of a shadow mask. The adhesion between metal pads and polyimide has long been an issue and makes wire bonding normally impossible. Within all the tested metal adhesion layer selections, Cr provides the best performance. Ti is the second best, which could achieve ~ 10% success rate in wire bonding. Other tested metals never work. The combination of Ar plasma
treatment, adhesion layer material selection of Cr, and comparatively thick bonding pads together make wire bonding very easy on our chips (usually 100% success rate). Usually I choose to increase the thickness of Cr to $\sim 100 \text{ nm}$ and keep Au to be still $\sim 50 \text{ nm}$ thick, as this combination is cheaper.

Before the measurement, $O_2$ plasma etching ($> 3 \text{ min high level, etching rate }> 100 \text{ nm/min}$) is applied to suspend the narrow Au bridges, leaving the other metal regions slightly undercut. It turns out that the suspended Au bowtie junctions are extremely stable in the air and can be stored for months without problems. They even seem to be less vulnerable to the static charges and are therefore more robust. The SEM photo of a typical suspended junction is shown in Fig. 5.1(a).

The whole chip is placed on top of a 100-pin square chip carrier, where only 16 pins transmit signals to the external SMA transmission lines. Wire bonding is used to electrically connect the contact pads on the sample and the corresponding leads on the chip carrier. Spring-loaded contacts are installed as the counter supports and hold the sample steadily. A stepper motor, remotely controlled by the computer, is employed to generate pushing displacements of a screw and the bending of the flexible substrate is realized. The whole apparatus is home-built, and mostly done by Leland Richardson. Regarding more details of the design, I will refer to his thesis[225]. A silicon diode DT-670 and a resistive heater with $25\Omega$ and $25W$ power allowed to dissipate are installed near the sample holder to measure and control the chip temperatures. Lakeshore 331 is the temperature controller used to realize the above-mentioned functions. In reality, the heating power is not sufficient to efficiently heat up the sample, and the temperature controller itself introduces large background noise. Thus during the experiments at cryogenic temperature, I only occasionally turn on 331 to double check the temperature and the temperature control is through
Figure 5.1: (a) An electron micrograph of a suspended bowtie junction. (b) An example of linear scaled-bias dependence, with the red line the best linear fitting result. This example data is of a 2.4 $G_0$ device, taken at 45 K. In the bottom inset, if the wrong temperature is used in the scaled bias, the linear fit will generate a larger intercept (and the low bias data will deviate visibly from the fit). This panel is a self-consistency check of the validity of the linear fitting analysis. (c) A photo of the setup.

adjusting the He flow but not through the heater. Using this dumb method, it turns out that the sample temperature can be maintained very well and the variation is $\sim 5 \, K/day$.

This setup is built based on a variable temperature dewar dip-stick cryostat (CRY-O Industries SN4311) and can be cooled down to liquid He temperature. Before cooling down the chamber, the whole system needs to be pumped to remove any water or
gas which could potentially freeze and cause contamination or damage to the system. Then it’s lifted and flipped, and directly inserted into a liquid He dewar, as shown in Fig. 5.1(c). The He recovery system is always used. If the liquid He dewar is not full, one can insert the system halfway and cool the chamber down by the cold gas first, then insert it down to the bottom and reach 4.2K. This system does not have a perfect RF shielding and during the noise measurements a big cloth made from Al foil is further used to cover the whole system and all the external background can be completely removed (Don’t forget to turn off the temperature controller and unplug the corresponding cable from the apparatus).

In a typical measurement, the Au bridges are first narrowed down by electromigration, and then further broken by slowly bending the substrate, using the conventional MCBJ technique. Electromigration is usually easy below 1 kΩ, and will be even easier at higher temperatures, which is natural as the device will be less stable and easy to break. I usually only electromigrate the device till $\sim 600$ Ω before starting the mechanical bending. A lower start resistance increases the total bending distance and forces the device elongation exceeding its elastic limit and makes a reconnection possible after breaking the device.

We stop at any interested conductance values and record the bias current and the excess noise simultaneously. Reconnections to $\sim 10G_0$ are needed to effectively randomize the microscopic structures. The small size of the suspended part in the device and the mechanical bending geometry lead to greatly enhanced mechanical stability of the Au junctions when compared to STM-style junctions.

Most of the measurements are conducted on fixed stable atomic configurations with motor stopped, avoiding any transients associated with changes in junction geometry during a particular noise measurement. The RF modulated excess noise
measurement technique is the same as what we introduced before. A slower 1 kHz square waves bias and as long as 300 ms lock-in time constants are chosen to improve the measurement resolution. We don’t need a rapid measurements because now the device is stable and not subject to ms-scale variations.

We also tried the measurements while the motor is moving back and forth without stops, similar to what are discussed in the last chapter (but a lot slower). After the device is broken, we can easily back the motor and reconnect the device. A nicely controlled program with correct feedback can allow the device keeps being pulled apart and reconnecting between pre-set conductance values and ensembles of data could be measured. The inset of Fig. 5.2 shows the simultaneous measurements of conductance steps and the corresponding shot noise steps. In those measurements, the lock-in time constant has to be turned down and 3 ms is the usual choice. Short time constants means a higher background. From the data (the inset of Fig. 5.2 shows an example), a distribution of Fano factor vs conductance can be obtained, as shown in Fig. 5.2. It is qualitatively the same as the measurements over fixed geometries (see Fig. 5.4), meaning that the concerned transient errors are minimized by the slow motion of motor and the high stability on the device.

5.2 Noise on metal stripes

Shot noise should be killed completely by strong inelastic electron-phonon interaction. The bottom line check in our measurement is to measure the noise directly on the Au bridges without any electromigration, to see whether it is really zero. Samples with different sizes are tested. A: 800 nm × 150 nm × 20 nm, B: 800 nm × 150 nm × 14 nm, C: 5 μm × 150 nm × 14 nm. The result is surprising. Though in type A samples no excess noise can be measured at 300 K and only a little at 4.2 K, significant excess
Figure 5.2: Fano factors vs Conductance values. Data are taken at different biases. Blue: 191 mV; Green: 86 mV; Purple: 70 mV. All three overlap pretty well, indicating the Fano factors are unchanged at different bias. The red curve is assuming the case there is always only one channel to be partially transmitted.

Noise are seen in the thinner devices B and C, at both room temperature and liquid He temperature. The results on type B and C are similar and the longer devices C seem having stronger noise. This is even more confusing as the longer device should be more affected by inelastic electron-phonon scatterings and therefore tend to have less shot noise. The typical result of type C devices is shown in Fig. 5.3. Fig. 5.3(a) gives the temperature dependent excess noise measurements as a function of the bias current. The low bias curvature indicates background floor.

At lower temperatures the excess noise is stronger. This trend could be consistent
Figure 5.3: (a) Excess noise vs bias current at multiple temperatures ranging from 4.2 K to 300 K. (b) Calculated $\delta T$ at the bias maximum vs the environmental temperature, assuming all the excess noise is from heating effect. (c) Calculated Fano factor fitted from the slope vs the environmental temperature, assuming all the excess noise are shot noise.

with shot noise. If there’s some reason, the inelastic effects are not strong enough to suppress all the shot noise (indeed our devices are still very small systems), at higher temperature the electron-phonon interaction mean free path is smaller hence less shot noise survives. This trend could also be a two terminal heating effect: at lower temperatures, the heat dissipation is more difficult and the electronic heating should be stronger and the noise should be larger. To clarify the origin, we can assume all the measured excess noise are two terminal heating effect and compute
the corresponding temperature elevation $\delta T$, as shown in Fig. 5.3(b); or assume all of the observation are shot noise and compute the Fano factors, as shown in Fig. 5.3(c). It turns out that both effects are weak and it is not easy to rule out any of them. However, while the observation of finite shot noise at cryogenic temperatures could be imagined, it’s not straightforward to anticipate them existing in the same micron-meter scale devices at room temperature. The longer device shows stronger noise seems to be consistent with the heating picture more, since the suspended part are longer and thermal dissipation is more difficult. Tens of Kelvin’s heating is also a reasonable number. I tend to believe what we saw here are simply heating effects and no shot noise is presented, as what we expected originally.

5.3 Bias dependence of shot noise at multiple temperatures

At zero temperature, shot noise is always linear to the bias, if the Fano factor is independent of bias. At finite temperatures, as we discussed in the previous chapters, a linear scaling of the scaled bias $X$ is still used to extract Fano factors. In the presence of electron-phonon interactions, the Fano factor is expected to exhibit a bias dependence above the electron-phonon onset voltage, $V = \hbar \omega / e$. Deviations from linearity in plots of $S_I(V) - S_I(0)$ vs. $X$ can be indicative of these inelastic processes.

Kumar[119] in 2012 observed kink-like features in the bias dependence of shot noise in many atomic-scale Au junctions at 4.2K. These kinks are usually observed at a voltage between 10 to 20mV, consistent with the known energy of the optical phonon in Au[58, 54] and also their measured differential bias-dependent changes in the conductance. If the inelastic contribution to the noise is detectable, a temperature dependent study is favored, as the phonon population plays an important role.

In previous chapters we find Landauer-Buttiker expectations work pretty well at
biases below about 200 mV and no sign of inelastic effects are seen. And if the local heating picture is correct as what we expected, it also means the phonon cooling of electronic temperatures are somehow weak. However, since weak electron-phonon interactions should be small, and the coupling strength $\lambda$ between electrons and the vibrational modes depends on the microscopic junction details, ensemble averages might wash out modest modifications to the noise. Measurements of individual electromigrated junctions at reduced temperature have shown nonlinearities and asymmetries in the noise at 77 K. Detailed measurements over many individual stable microscopic configurations are needed to see whether temperature strongly affects the noise bias dependence in the moderate bias regime[226].

Measurements are conducted with sample temperatures between 4.2 K to 100 K. As the theories generally discuss single channel transport, and additional quantum channels complicate even the non-interacting interpretation, we mostly focus on junction configurations with conductances below $2 G_0$. Because the shot noise signal scales with current, the noise is most clearly resolved at bias voltages that are well in excess of $\sim 17$ mV, the expected threshold for the Au optical phonon. Thus the deduced Fano factors should already be modified by inelastic effects, if any are present. We restrict applied biases to less than 200 mV to maintain junction stability; hence the nonlinear increase in noise at higher biases[98] is not examined here.

We performed detailed measurements in different conductance regimes: the tunneling regime with $G < 0.1 G_0$, the single-atom contact regime with $G \sim 1 G_0$, and the few-channel point contact regime, $1 G_0 < G < 10 G_0$. Au junction stability does not favor the conductance values between 0.1 $G_0$ and much below 1 $G_0$, though repeatedly disconnecting and reconnecting the contacts sometimes results in configurations with sufficient stability for data collection. The scaled bias depen-
dence of the noise found in most junction configurations is linear with approximately vanishing intercepts. A typical example is shown in Fig. 5.3(b), where the environmental temperature is \( \sim 45 \) K. The exact value of the temperature matters in the scaled-bias fitting, as shown in the inset of Fig. 5.3(b); only the correct temperature minimizes the intercept in a linear fit, which is also the main basis of the shot noise thermometry[128]. Consistency between the measured sample temperature (via conventional thermometry) and the temperature parameter that linearizes (with zero intercept) noise vs. scaled bias indicates the high data quality and the dominance of shot noise.

The lack of clear nonlinearities means that no significant effects from thermally nonequilibrated vibrational modes are observed in most devices. In a small fraction of the measurements we do see small upwards nonlinearities close to the highest applied voltages, at all temperatures. This could be a result of unusually strong coupling with nonequilibrated vibrational modes, weak energy relaxation associated with some particular atomic configurations, or junction instability driven by bias. Given the challenge of interpreting these idiosyncratic configurations, we focus on the configurations that do not show these nonlinearities.

In typical, stable junctions, the Fano factors could be extracted by the linear fitting procedure, allowing us to generate a map of the distribution of Fano factors, shown in Fig. 5.4(a). Close to the \( 1 \, G_0 \) plateau, some recorded bias dependences have small nonlinearities that result in uncertainties when determining the Fano factors, but will not cause significant variations in the exact values. These features will be discussed in Section 5.4 and are believed to have non-interacting origins.

In Fig. 5.4(a) red circles represent measurements at 4.2 K, while blue squares and purple triangles are the data at 50 K and 100 K, respectively. The black curve
Figure 5.4: (a) The conductances and Fano factors measured on all the devices. The red circles represent measurements at 4.2 K, and the blue squares and purple triangles represent data at \( \sim 50 \) K and 100 K, respectively. The black curve is the non-interacting "forbidden region" boundary. (b) The deduced transmittances of the first three quantum channels from panel (a). Blue, red, green colors represent \( \tau_1 \), \( \tau_2 \) and \( \tau_3 \). Filled circles represent data at 4.2K, and the open squares and triangles represent the data at \( \sim 50 \) K and 100 K, respectively.

is the mathematically allowed smallest shot noise in non-interacting limit, reached when only one quantum channel has transmittance not equal to 1. Atomic-scale Au junctions are well approximated by this limit, especially when \( G < 1 \ G_0 \). The region below the black curve is forbidden in the Landauer-Büttiker model with spin degeneracy[126, 127]. Beyond experimental uncertainty, no measured points fall in this region. Most data are distributed close to the forbidden region boundary and
two pronounced quantum suppressions of Fano factors are seen, at 1 $G_0$ and close to 2 $G_0$. All these observations are consistent with previous experiments that show that clean Au atomic-scale junctions favor fully transmitted channels. Interestingly, the distribution of Fano factors does not seem to depend significantly on temperature. Especially when $G \rightarrow 0$, 1 $G_0$, the measured data at different temperatures overlap with each other very well. While the 4.2K data seems follow the same trace and have very small variation, a few outliers at higher temperatures do have higher Fano factors. This likely results from channel mixture and greater likelihood of junction contamination by adsorbates at higher temperature.

Over the temperature range 4.2 K to 100 K, the corresponding vibrational occupations $n_B = 1/(e^{\beta \hbar \omega} - 1)$ increase from $3 \times 10^{-21}$ (essentially zero) to 0.16, if taking $\hbar \omega = 0.17$ meV. Even as small as an occupancy of 0.16 might generate an additional noise larger or comparable to the total inelastic correction at zero temperature limit[151, 152] and should be detectable. The linear bias dependence in our measurements already indicates little inelastic contribution to the noise. The lack of phonon population dependence further indicates apparently weak e-ph coupling in our devices. Hence not only the difference caused by higher temperature but the whole inelastic contribution of noise might just be too small to detect. It’s also possible that the role of phonon population is somehow overestimated and 100 K is actually not enough to observe any change of the inelastic part in shot noise. On the other hand, the temperature insensitivity means the dynamics of Au atoms is not altered in that largely varied temperature range either, which again is surprising. To further clarify the effect of thermal phonon occupation measurements at higher temperatures might be favored, but the stability of atomic-sized devices naturally limits practical environmental temperatures and make such measurements challenging.
Regardless of temperature dependence, theories considering the interactions between electrons of single quantum channel and single thermally equilibrated vibrational mode predict that Fano factors are suppressed when the single channel transmittance is between $1/2 \pm 1/2\sqrt{2}$ and enhanced otherwise\cite{150, 143}, even at zero temperature. This is because electrons can always emit phonons regardless phonon population and the only requirement is to have enough energy. Thus in the predicted suppression region, if electron-phonon interactions cause a significant modification, we might have a chance to see some data fall below the non-interacting forbidden region, but this is not seen. Instead, compared with the data with $G \to 0, 1 \, G_0$, the measured Fano factors in between $1/2 \pm 1/2\sqrt{2} \, G_0$ are enhanced relative to the black curve. Such a positive deviation is likely to be a result of non-interacting channel mixture, which potentially hides any effects from electron-phonon interactions. One possible origin of such a channel mixture is adsorbate contamination. Indeed, geometrically Au devices do not favor conductance values between 0 and 1 $G_0$. The reason stable configurations in this conductance region sometimes still can be reached might be a result of contamination, which itself could also induce channel mixture. If that’s true, we are selectively measuring dirty gold contacts in that conductance region. The direct tunneling through contamination layer\cite{227, 228} is not likely to dominate though, as the relevant conductance ($0.5G_0$ for example) is too large. In the cases that we suspected molecules fell in the gap between Au electrodes, reconnections are very difficult, which is not the case here.

Assuming the noise results are well described by the Landauer-Büttiker model, combining the conductance and Fano factor measurements we are able to calculate the transmittances of the first three quantum channels below 2 $G_0$, as what we show in Fig. 2(b). We assume the third channel is not opened below 1 $G_0$ ($\tau_3=0$), and
the first channel is always nearly fully transmitted above 1 $G_0$ ($\tau_1 \to 1$). Thus at all the conductance values below 2 $G_0$ we always only have two variables and independent measurements of noise and conductance determine them completely. When $1 G_0 < G < 2 G_0$ we have to manually set the first channel’s transmittance $\tau_1$ to be a value slightly lower than 1 (but larger than 0.95), to force the calculated $\tau_2$ and $\tau_3$ to be real. This introduces some uncertainty but will be qualitatively valid. This implies that the first quantum channel is not fully opened near 1 $G_0$ in our devices, which is consistent with previous observations not showing perfect noise suppression[229, 117, 189]. As a self-consistency check, $\tau_2$ and $\tau_3$ are indeed small below 1 $G_0$ and 2 $G_0$ respectively. Simulations in a previous paper[189] based on molecular dynamics and the Green functions technique match these observations. At 4 K such a simulation predicted an enhancement of the ensemble-averaged Fano factors close to 0.5 $G_0$, as a result of maximized $\tau_2$ contribution; This is reproduced by the transmittance mapping extracted from the data. This good agreement favors the channel mixing interpretation and confirms that the intrinsic motions of Au atoms could also be a possible origin of the enhanced Fano factors between 0 and 1 $G_0$, as well as the non-saturated quantum channel near 1 $G_0$. On the other hand, the same simulation suggests that the second channel’s contribution below 1 $G_0$ will be largely washed out at room temperature, due to changes in configuration stability. Such a difference is not seen experimentally up to 100 K.

5.4 Non-linear bias dependence close to 1$G_0$

We also find that the bias dependence of shot noise close to the 1 $G_0$ plateau is sometimes slightly curved and seems to have reproducible deviations from linearity as a function of bias. Two examples are shown in Fig. 5.5(a-d) and 3(e-h). Data
points in both bottom panels 5.5(d) and 5.5(h) with error bars represent the measured excess noise spectral density as a function of bias. Top panels 5.5(a) and 5.5(e) are the simultaneously measured I-V curves. The example shown in left column 5.5(a-d) is a $\sim 0.89 \, G_0$ device with the second channel participation $\tau_2$ as small as 0.03. The measured noise clearly is concave upward. The right column 5.5(e-h) represents another device at $\sim 0.94 \, G_0$ but has a significant contribution(0.14) from $\tau_2$ and less nonlinearity with bias for both $G$ and $S_I$.

At each bias, the measured I-V curves, though still pretty linear, have small but highly reproducible deviations from linearity. Such small deviations are easier to see in panels 5.5(b) and (f), where the difference between the measured conductance data and the linear fittings (shown as solid straight lines in 5.5(a) and (e)) are plotted as a function of bias. This indicates a weak bias dependence of the conductance, or the transmittances $\{\tau_i\}$ of the single-atom constriction, as what has been shown in 5.5(c) and (g). As a result, based on the Eq. 2.15, Fano factors can vary as a function of bias and even the non-interacting shot noise can have more complicated nonlinear bias dependence, which could be easily calculated. Indeed when larger variations of conductance over bias are observed, the junctions with conductance close to 1 $G_0$ tend to be more curved, as shown in 5.5(a-d). In junctions whose conductance have smaller variations like the device in 5.5(e-h), the bias dependence tend to be linear and featureless.

Universal conductance fluctuations (UCF) provide a mechanism to explain such a weak bias dependence of the $\{\tau_i\}$. Though below 0.9 V there won’t be significant variations of the transmittances in single-atom Au constrictions[230] and the density of states of Au in the relevant energy range are relatively featureless, disorder and quantum interference can affect the $\{\tau_i\}$[50].
Figure 5.5: Two examples of the nonlinear bias dependence near $1 G_0$: (a-d) $\sim 0.89 G_0$ with the best-fitted $\tau_2 \sim 0.03$. The conductance has a relatively strong bias dependence. (e-h) $\sim 0.94 G_0$ with the best-fitted $\tau_2 \sim 0.14$. The conductance has weaker bias dependence. From top to bottom, (a)(e) I-V curve, (b)(f) deviation $\delta I \equiv I - I_{\text{linearity}}$ vs V, (c)(g) $G(V)$ vs V and (d)(h) the measured excess noise vs the scaled bias respectively. In the bottom panel (d)(h) the solid curve represents the expected shot noise after properly accounting for the bias dependence of device conductance and choosing $\tau_2$ to be the best-fitted value. The dotted line forces $\tau_2$ to be zero (single channel case), while the dashed curve in (d) forces $\tau_2 = 0.1$. Based on the best-fitted value of $\tau_2$, the $\tau_1$ values associated to the total conductance are also indicated in panels (c) and (g).

To verify the correlation between the observed nonlinear features in shot noise and the conductance fluctuation, we recalculate the expected shot noise, allowing the conductance to be weakly bias dependent. The single-atom Au constriction can be well approximated by a single quantum channel picture, and a more complete description allows the second channel to be weakly transmitted. Thus we have $G/G_0 = \tau_1 + \tau_2$, with $G \rightarrow 1 G_0$, $\tau_1 \gg \tau_2$. To model the slightly bias dependent conductance, we have $G = G(V) = <G> + \delta G(V)$. The ideal procedure to extract $G(V)$ is to densely
measure the differential conductance as a function of bias, and integrate the result from zero to the target voltage. Our apparatus did not allow simultaneous noise and differential conductance measurements; instead, we approximated this by taking $\Delta I/\Delta V$ at each bias from the I-V curve, though this sacrifices the voltage resolution.

By plugging $G(V)$ into Eq. 2.16, the expected shot noise spectral density $S_I = S_I(\tau_2, X)$ was calculated. A least square fitting to the data generates the best-fitted value of $\tau_2$ and the $S_I(\tau_2, X)$ is reduced to $S_I(X)$. The best-fitted results of both examples are plotted as the solid curves in Fig. 5.5(d) and (h), retracing the data reasonably well. This simple toy model calculation can be used to describe. This simple toy model calculation can be used to describe varying degrees of nonlinearity in $G(V)$, and was successfully applied to all other devices not shown here. Consistency between the independent measured $G(V)$ and $S_I(V)$ when analyzed this way strongly suggests that such a voltage dependent $\{\tau_i\}$ can explain both data sets. A similar analysis has been performed by others (unpublished, Vardimon and Tal et al. [231]).

It is striking that such small nonlinearities in I vs V can result in a readily apparent effect on shot noise, though this is mathematically natural. The shot noise carried by each individual quantum channel is proportional to $\tau_i(1 - \tau_i)$. Considering a small variation $\delta \tau_i$, the resulting shot noise is $(\tau_i + \delta \tau_i)(1 - \tau_i - \delta \tau_i)$. As the transmittance of the dominant channel $\tau_1 \to 1$, $\tau_1 + \delta \tau_1$ is almost unchanged but $\delta \tau_1$ can be comparable to $1 - \tau_1$ and leads to a change by a significant fraction. If $\tau_1 \to 1$ is not valid, shot noise becomes comparatively insensitive to the small $\delta \tau_1$ and the normal linear bias dependence of $S_I$ is recovered, as what we have seen away from the $1 G_0$ plateau. Similarly, if $\tau_2$ is non-negligible, even at the $1 G_0$ plateau $\tau_1$ (due to the fact that $\tau_1 = G/G_0 - \tau_2$) is not close to 1 and shot noise tends to be less sensitive to the conductance fluctuations.
In the Section 4.2.3, nonlinearities have also been seen in ensemble-averaged shot noise measurements at room temperature, also only observed close to quantum suppression. We interpreted that observation as a result of residual flicker or 1/f noise contributions. However, here the strong correlation between conductance and excess noise is not consistent with residual 1/f noise picture. In atomic scale junctions, flicker noise does scale with resistance due to different numbers of effective fluctuators[115]. But lacking the proportionality to $1 - \tau$, it should be completely insensitive to the $\sim 5\%$ conductance variations and present much more smooth quadratic curvatures. Furthermore, the origin of 1/f noise in atomic-scale metal junctions is the superposition of many thermally-activated two-level fluctuations caused by defect motions[105, 109], which should be suppressed at cryogenic temperatures. And such suppression of 1/f noise at lower temperatures were observed[105, 113].

Other mechanisms of the super-linear bias dependence beyond non-interacting limit might even be more unlikely, as there is no clear reason why $1G_0$ conductance is that special.

5.5 Conclusion

In this chapter we examined the bias dependence of shot noise at atomic-scale Au junctions, over a large bias, conductance, and temperature range. Even at the biases well above the energy of the Au optical phonons and varying the temperature from 4.2 K to 100 K, no clear temperature dependence is seen. This likely indicates weak inelastic contribution to the shot noise associated with very small electron-phonon coupling strength $\lambda$ and thermal phonon populations. Fano factors are observed to be enhanced near $0.5 \ G_0$, which either results from contamination or the intrinsic atomic configurations favored by Au junction formation. Near the $1 \ G_0$ plateau,
the bias dependence of shot noise shows geometry-dependent small nonlinearities, which are well described by a simple non-interacting scenario considering weak bias dependence of transmittances as expected from UCF.
Bibliography


“Quantum contact in gold nanostructures by scanning tunneling microscopy,”


[72] Z. Huang, F. Chen, R. D’agosta, P. A. Bennett, M. Di Ventra, and N. Tao,


noise thermometry,” *Metrologia*, vol. 33, no. 4, p. 325, 1996.


