Lancaster University
Department of Physics

EXPERIMENTAL CHARACTERISATION OF MOLECULAR JUNCTIONS VIA SCANNING PROBE MICROSCOPY

Sam Harley

This thesis is submitted for the degree of Doctor of Philosophy
September 2024

ABSTRACT

The experimental characterisation of molecular junctions is perhaps the most important class of measurement within molecular electronics; as interest in the field has grown throughout the 21st century, SPM and break-junction methods for characterising electrical and thermal transport through molecular junctions have become increasingly advanced, providing insight into a diverse range of quantum transport phenomena. This work details the development of several bespoke instrumentation and data analysis systems to facilitate precise, repeatable measurement of the electrical and thermoelectric properties of molecular junctions. Novel fabrication strategies and careful design were employed to attain state-of-the-art performance at low-cost.

A bespoke STM break junction (STMBJ) system for single-molecule conductance measurements is presented, featuring a novel approach to the fabrication of an SPM head using steel-reinforced epoxy granite (SREG) composite. In order to extract additional information beyond the most-probable conductance from the large datasets produced by STMBJ, a detailed investigation of autoencoder-based feature extraction and clustering algorithms was conducted. Notably, the role of feature disentanglement in STMBJ data segmentation was studied for the first time.

Additionally, a high-precision thermovoltage AFM (ThAFM) module was developed which enables characterisation of the thermopower of multi-molecule junctions using a commercial AFM. The key challenges of ThAFM are outlined, and corresponding instrumentation design principles are identified and implemented to ensure optimal performance. The performance is demonstrated by measurements of the Seebeck coefficients of C_{60} and ODT monolayers with precision exceeding the best reported measurements to date. Finally, the system is used to produce the first reported measurement of the Seebeck coefficient of a ZnPc monolayer.

CONTENTS

DECLARATION

I confirm that this thesis is my own work and has not been previously submitted for the award of a degree elsewhere. All work contained herein is my own, except where indicated otherwise by reference or acknowledgement.

LIST OF FIGURES

1. INTRODUCTION

1.1 Molecular Electronics

Molecular electronics is an interdisciplinary field, at the interface between Physics and Chemistry, primarily concerned with the study and manipulation of electron transport in organic molecules [4] towards the development of electronic devices built from individual molecules rather than from bulk materials [5]. The motivations for pursuing molecular electronics research are twofold:

From a perspective of fundamental scientific curiosity, molecular devices provide a fascinating platform for the theoretical and experimental study of quantum transport in a highly customisable nanostructure. The vast freedom in molecular design afforded by modern synthetic chemistry provides far more control over the properties of the device than is available in other devices of comparable dimensions, such as inorganic quantum dots [5,6]. This flexibility has facilitated access to a rich assortment of quantum phenomena, and molecular electronics experiments continue to provide valuable insight into quantum transport regimes which are not yet well understood theoretically [7].

In addition to basic research, there is a strong technological motivation for pursuing molecular electronics. The majority of conventional electronic devices are built from bulk inorganic semiconductor materials using a 'top down' approach in which devices start from bulk materials, and structures are engineered subtractively. This approach is limited by the properties of the available materials and the manufacturing process (for example electron beam lithography is fundamentally limited in resolution by the achievable beam diameter and by scattering effects in the resist material [8]). In contrast, molecular electronics devices could use a 'bottom up' approach to fabrication, where the components are built from individual molecules which are specifically designed via synthetic chemistry to have the desired properties [5]. Nanostructured molecular scale devices may also leverage phase-coherent quantum transport effects, such as quantum interference, which are not available to conventional electronics. This could lead to the realisation of highly efficient devices with unparalleled performance, such as chemical sensors with single-molecule detection thresholds [9–11].

The concept of building devices in a 'bottom up' manner from individual molecules

originated in a rudimentary form, along with many of the foundational ideas of nanoscience, in the late 1950s. Richard Feynman's now-famous 1959 talk entitled 'there's plenty of room at the bottom' [12] briefly alluded to the synthesis of bespoke organic molecules as a means of building nanoscale 'machines'. In 1956 Von Hippel proposed creating new materials for electronic devices using a 'molecular engineering' approach where, rather than searching for bulk materials with the required properties, novel materials were to be designed at the molecular scale [13]. In 1959 this led to a joint R&D project between Westinghouse Electric and the US air force to develop robust miniaturised electronics based on the 'molecular engineering' paradigm to meet the increasing requirements of airborne systems [14]. However, by 1960 the program had already largely moved away from the 'molecular engineering' approach, and the program ended entirely in 1962, although Westinghouse Electric continued to use the term 'molecular electronics' to refer to conventional integrated circuit (IC) devices in subsequent years [15].

The field of molecular electronics began to emerge in a recognisable form in 1974 when the concept of single molecules functioning as analogues of conventional electronic components was introduced by Aviram and Ratner, who proposed that a single molecule could function as a rectifier [16]. This led to a surge of interest in molecular electronics, but progress was hindered by the limited theoretical understanding of transport in single molecules and the experimental challenges of measuring transport properties of molecular junctions [17].

The debut of scanning probe microscopy (SPM), beginning with the invention of the Topografiner in 1972 [18] and the scanning tunnelling microscope (STM) in 1981 [19], and the resulting development of atomic force microscopy (AFM) in 1986 [20], provided unprecedented access to matter at the molecular scale; for the first time molecules could be imaged in real space [21, 22], and even addressed individually to form molecular junctions. Experimental progress in the study of molecular junctions continued through the 1990s and into the 21st century, including the advent of single-molecule conductance measurements using break-junction techniques [23, 24], which remain one of the most powerful approaches for the characterisation of molecular junctions. Simultaneously, significant progress was made in developing theoretical descriptions of transport through single molecules [17, 25].

At the time of writing, molecular electronics presents as many opportunities and challenges as ever: activity in the field has grown significantly since the early 2000s (as illustrated in figure 1.1), and enumerable important developments have advanced the state-of-the art in almost all aspects including device fabrication, characterisation and theoretical modelling. Scanning probe microscopy and break-junction methods have continued to play an important role in experimental characterisation efforts, and myriad adaptations have been developed to extend these methods

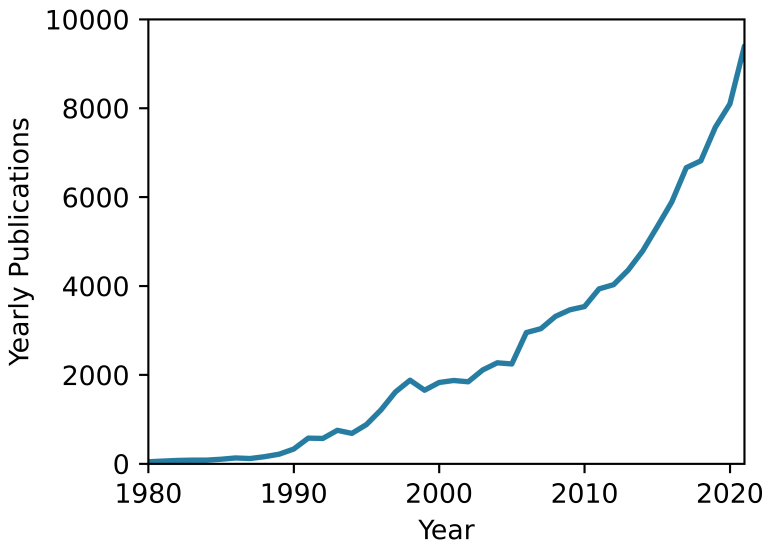


Fig. 1.1: A plot showing the growth in number of publications for 'Molecular Electronics' by year (from 1980 to 2021). Data from Web of Science [1]

beyond conductance measurements towards characterisation of various properties including thermopower [3] and thermal conductance [26–28]. However, even as measurements of molecular junctions become increasingly complex, accessing ever more parameters and introducing gating and modulation, implementing reliable and reproducible measurements remains challenging. Additionally, modern molecular junction experiments produce large amounts of complex multivariate data, which traditional ensemble analysis approaches are not well equipped to handle. The next generation of molecular electronics experiments will require sophisticated data analysis techniques to extract maximum insight from the resulting datasets.

1.2 Scientific Objectives & Overview of the Thesis

The objective of the work presented here was to develop enhanced SPM measurement techniques for the characterisation of the electrical properties of molecular junctions, with an emphasis on precision, reliability and ease-of-use. It is hoped that these improvements to existing measurement techniques will allow more accurate measurements of key parameters of molecular junctions, with higher throughput.

1.2.1 Extensible Break Junction Platform for Advanced Characterisation of Single-Molecule Junctions

The first objective of this work was to develop a low cost STM, optimised for STMBJ measurements, and accompanying software for automated data analysis. This in-

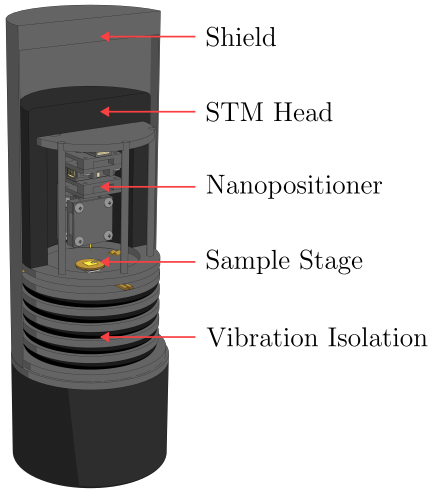


Fig. 1.2: A render of the STMBJ developed in this work, with the shield and composite casing partially cut away to expose the scan unit

strument is intended to provide an extensible platform for the exploration of advanced STMBJ techniques in future.

STMBJ

Break junction techniques form nanoscale junctions by repeatedly forming and breaking metallic contacts between a pair of electrodes; when performed in the presence of a target molecule, molecular junctions may be formed due to molecules binding between the electrodes during the breaking process. Some of the earliest experimental studies of the conductance of single-molecule junctions were performed using break-junctions [29], yet the technique remains highly relevant today, with sophisticated break-junction experiments providing insight into a wide range of phenomena beyond electrical conductance [28], such as room-temperature quantum interference [30] and thermoelectricity [26].

The STM break junction (STMBJ) combines the break junction technique with scanning tunnelling microscopy (STM). In its most basic form, STMBJ utilises the fine positioning capabilities of STM to form and break nanoscale metallic junctions between the STM tip and sample, while measuring the conductance of the junction as a function of the tip-sample separation. Since the introduction of STMBJ in 2003 [24] many groups have adapted standard STM instruments for STMBJ, while others have built dedicated STMBJ instruments optimised for the formation of stable junctions, achieving significant performance enhancements [31].

Automated Data Processing Using Variational Bayesian Methods

The formation of molecular junctions via STMBJ is a stochastic process, and a complete STMBJ measurement consists of an ensemble of data from thousands of junction formation and breaking cycles. Typically, statistical analysis is used to extract the ensemble average conductance and breaking length of the molecular junction. However, this approach can result in the loss of information. Recently, several groups have begun to use more advanced data analysis to extract additional information from STMBJ datasets. Among the most powerful of these are clustering techniques which identify sub-populations of similar traces within the dataset, providing insight beyond the average conductance. Due to the high dimensionality of STMBJ data, applying clustering algorithms directly to the dataset typically yields poor performance; a 'feature extractor' which encodes the salient features of the data into a representation with reduced dimensionality is therefore required.

One of the objectives of this work was an exploration of autoencoder neural networks as feature extractors for STMBJ, using a combination of procedurally generated and experimental datasets to identify strategies to improve downstream clustering accuracy and interpretability. This work is discussed in chapter 5.

1.2.2 Thermoelectric Characterisation of Multi-Molecule Junctions

Thermovoltage AFM (ThAFM) is an extension of AFM which allows the measurement of the thermoelectric properties of molecular thin films. Brief introductions to thermoelectricity and ThAFM are contained within this section; more detail can be found in chapter 6. One of the primary aims of this work was to identify limitations of existing ThAFM implementations and develop an inexpensive, reliable and stable ThAFM module to augment the capabilities of a commercial AFM. The resulting implementation of the ThAFM method provides unprecedented performance and ease-of-use in a self-contained module which can be built for under £100 (orders of magnitude less expensive than a typical ThAFM setup). This aspect of the work is discussed in chapter 6.

Thermoelectric Effect

The term thermoelectric effect refers to the generation of a potential difference (thermovoltage) due to a temperature gradient across a junction; conversely, a temperature gradient can be created by applying an external potential. There exists a linear relationship between the open-circuit thermovoltage ΔV and temperature difference ΔT, the slope of which is defined as the Seebeck coefficient S (also known as the 'thermopower') [32].

$$S = -\frac{\Delta V}{\Delta T} \tag{1.1}$$

Unlike electrical conductance, the Seebeck coefficient is an intrinsic property. Hence, neglecting the effect of interactions between molecules, the Seebeck coefficient of a multi-molecule junction should be equal to that of a single-molecule junction.

Devices which leverage the thermoelectric effect for energy generation or heating/cooling are of interest since, among other reasons, no moving parts are required and there is a high potential for miniaturisation and integration into other systems. Especially promising potential applications of future thermoelectric devices include miniaturised solid-state thermal management solutions and large scale harvesting of so called 'low grade' waste heat which cannot be recovered by most conversion methods [33]. Unfortunately, present thermoelectric devices suffer from low efficiency and power output compared to other heat-electricity conversion methods such as steam-turbines. In addition, many of the materials used in present inorganic thermoelectric devices are toxic, scarce or otherwise harmful to the environment, limiting their potential for widespread implementation. Despite these limitations, thermoelectric devices are currently in use in a number of specialised applications due to their unique advantages. The thermoelectric effect is also of scientific interest as it can provide additional information about the charge transport characteristics of a junction not encoded in other commonly measured properties such as the electrical conductance [34].

The efficiency of a thermoelectric device is quantified by the dimensionless 'figure of merit', ZT [35]:

$$ZT = -\frac{S^2 G T}{\kappa} \tag{1.2}$$

where T is the average temperature across the junction and G and κ are the electrical and thermal conductances, respectively. Current thermoelectric devices are mostly based on bulk inorganic materials (e.g. Bismuth-Telluride alloys), with typical room-temperature figure of merit values of $ZT \sim 1$ [11]. A commonly used guideline for the threshold at which thermoelectric devices would be competitive with other forms of heat-electricity conversion is $ZT \geq 3$ [36]. Over the past several decades efforts to achieve this target with bulk inorganic materials have been unsuccessful.

The theoretical efficiency limit for a thermoelectric device is the Carnot limit, since the processes underlying the thermoelectric effect are thermodynamically reversible. From equation 1.2 it can be inferred that in order to maximise the efficiency of a thermoelectric device at a given operating temperature the Seebeck coefficient S and electrical conductance G must be maximised while the thermal conductance κ is minimised. However, this is complicated by the fact that the thermal conductance is the sum of contributions from charge carriers and phonons, κ_e and κ_p respectively.

$$\kappa = \kappa_e + \kappa_p \tag{1.3}$$

The electrical conductance G and the electronic contribution to the thermal conductance κ_p are proportional to each other according to the Wiedemann-Franz law. Note that while the Wiedemann-Franz law was formulated for metals, a modified version is expected to hold for single-molecule junctions [37]. Hence efforts to optimise the figure of merit ZT by reducing the thermal conductance should aim to reduce primarily the phonon contribution while maintaining the electronic contribution in order to avoid inadvertently reducing the electrical conductance (and therefore ZT). The electrical conductance is also important in determining the power that can be generated by a thermoelectric generator; a high efficiency is of limited practical utility if the maximum power that can be generated is very low, except in niche applications where the power demands are low and the available heat energy is very limited (for example body heat harvesting). [33]

Nanostructured inorganic thermoelectric devices have enabled a number of new strategies for the tuning of thermoelectric properties, including independent tuning of κ_e and κ_p, resulting in enhanced ZT over their bulk counterparts [38]. Molecular thermoelectric devices offer additional avenues to improve the figure of merit using quantum interference effects, for example to enhance the Seebeck coefficient or suppress phonon transmission [39, 40]. Further discussion of the quantum transport concepts underpinning the thermoelectric effect in molecular junctions can be found in section 2.3.

ThAFM

Atomic force microscopy (AFM) is perhaps the most widely used variant of SPM, due in part to its flexibility; AFM is not limited to conductive samples, and the ability to measure and control the tip-sample interaction force opens possibilities for countless specialised contact and non-contact modes of operation for specific characterisation tasks. The most important variant of AFM for the study of molecular junctions is CAFM (conductive atomic force microscopy), in which a conductive probe is used to measure electrical properties of the sample [41]. Thermovoltage AFM (ThAFM) is a variant of CAFM introduced by Tan, Sadat and Reddy in 2010 [3], which measures the Seebeck coefficient of multi-molecule junctions formed by contacting the surface of a molecular thin film with a conductive AFM probe.

The fundamental working principle of ThAFM is as follows: A molecular junction is formed by a molecular thin film adsorbed on a conductive substrate (which acts as the drain electrode) contacted by a conductive AFM probe (which acts as the source electrode). A temperature gradient across the junction is established by heating

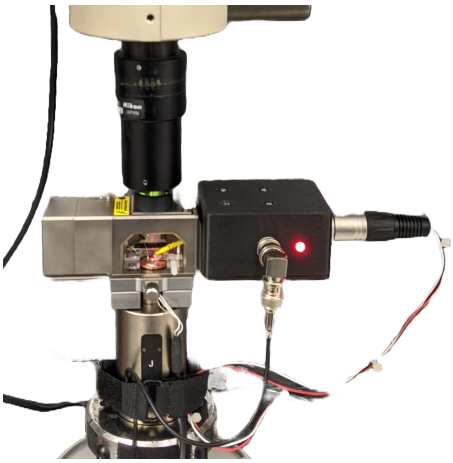


Fig. 1.3: A photograph of the ThAFM module installed on a Bruker multimode AFM

or cooling the sample, while a precision voltmeter connected between the tip and substrate measures the thermovoltage. Recording the thermovoltage as a function of the temperature gradient allows the Seebeck coefficient to be extracted. [3]

While simple in principle, ThAFM measurements are challenging in practice; inconsistent electrical contacts and unstable thermovoltage signals increase the uncertainty in the measured Seebeck coefficient and limit the throughput of the experiment. Chapter 6 details the results of an effort to improve the precision and reliability of an existing ThAFM implementation, resulting in the development of a new self-contained ThAFM module, pictured in figure 1.3.

2. PRINCIPLES OF QUANTUM TRANSPORT IN MOLECULAR JUNCTIONS

The primary focus of this work is the development of improved instrumentation for the characterisation of the electrical conductance and thermopower of molecular junctions: this chapter presents a qualitative introduction to the key concepts in quantum transport which underpin these properties.

Molecular junctions, as studied experimentally in this work, consist of organic molecules connected between two metallic electrodes, or 'leads', denoted source (s) and drain (d). A realistic junction may be a multi-molecule junction (MMJ), containing many molecules in parallel, or a single-molecule junction (SMJ) containing just one molecule; for the sake of simplicity the theoretical descriptions presented in this section will be restricted to the single-molecule case, illustrated in figure 2.1. The conductance and thermopower of these junctions can be understood in terms of the transport of electrons between the source and drain electrodes via the molecular junction. However, due to the nanoscale dimensions of the junction, classical models of electron transport are not sufficient and the quantum mechanical nature of the electron must be considered [42].

The electrodes can be modelled as quasi-infinite crystalline metallic structures, which are considered to be open systems possessing a continuum of energy levels described by the Fermi-Dirac distribution function [11]:

$$f(E) = \frac{1}{e^{\frac{E-\mu}{k_B T}} + 1} \tag{2.1}$$

where $T > 0$ is the temperature of the electrode, μ is the chemical potential and k_B is the Boltzmann constant. The highest occupied energy level in the electrode is the Fermi level, denoted by E_F , which at zero temperature is identical to the chemical potential. Note that in general the source and drain may have different temperatures $T_s \neq T_d$ and chemical potentials $\mu_s \neq \mu_d$ and therefore different Fermi distributions, $f_s(E) \neq f_d(E)$.

The molecule, however, has nanoscale dimensions and when in isolation is a closed system with finite energy levels corresponding to a discrete set of 'molecular orbitals', each of which can be occupied by a maximum of two electrons due to the Pauli exclusion principle. The orbitals are filled with pairs of electrons up to the

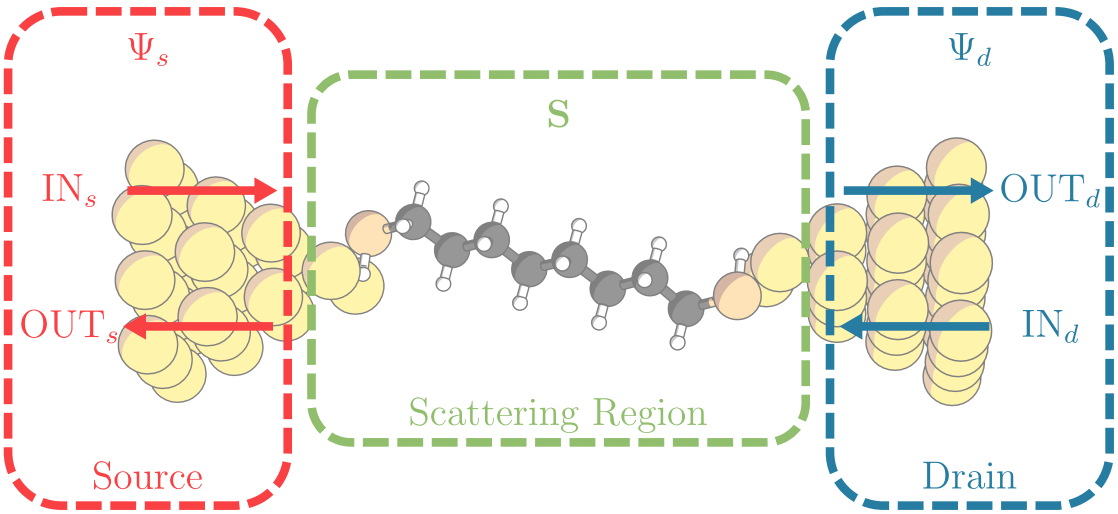


Fig. 2.1: An illustration of a single-molecule junction, showing how treating transport as a scattering problem allows the junction to be described in terms of a scattering matrix which relates the amplitudes of incoming and outgoing electron wavefunctions.

so-called highest occupied molecular orbital (HOMO). The adjacent unfilled orbital is called the lowest unoccupied molecular orbital (LUMO), and the energy difference between them is termed the 'HOMO-LUMO gap'.

Once the molecule is 'wired' into the junction, it can no longer be considered a closed system, as the hybridisation between the quantised molecular orbitals and the continuous levels found in the electrodes results in broadened resonances, depending on the strength of the electrode-molecule coupling [6, 33]. Additionally, interactions between the molecule and electrodes will result in reorganisation of the energy levels such that the Fermi energy will typically be between the HOMO and LUMO energies [6], and as a result the dominant electron transport mechanism will be 'off-resonance' tunnelling through the HOMO-LUMO gap. The Landauer formalism provides an intuitive, and often computationally efficient, approach to the calculation of transport properties such as conductance in the phase-coherent tunnelling regime [42].

2.1 The Landauer Formalism: Quantum Transport as a Scattering Problem

In the Landauer (also known as Landauer-Büttiker) approach, coherent quantum transport is modelled as a scattering problem, where the transport properties of

the junction are related to the transmission probability of a De Broglie wave representing an incident electron. For instance, the conductance G of a junction with n transmission channels is given by the Landauer equation:

$$G = G_0 \sum_{i}^{n} \tau_i \tag{2.2}$$

where τ_i, $i \in \{\mathbb{Z}, i \leq n\}$ is the transmission probability of the i_{th} conductance channel. The conductance quantum $G_0 = \frac{2e^2}{h}$ is the electrical conductance of a single perfect conduction channel [42] (see section 2.2.1).

In order to demonstrate the scattering approach, consider a simple time-independent one-dimensional situation, as illustrated in figure 2.3. The junction is modelled as a 1D 'scattering region', and the electrons in the source and drain electrodes are free quantum states represented by their wavefunctions, Ψ_s and Ψ_d respectively, which are solutions to the Schrödinger equation:

$$\mathbf{H}\,|\Psi\rangle = \mathbf{E}\,|\Psi\rangle \tag{2.3}$$

H and **E** represent the Hamiltonian and the energy operator, respectively. Using the superposition principle, the wavefunctions can be represented as a sum of plane-wave components propagating in opposite directions, representing the electrons entering and leaving the scattering region.

$$\begin{aligned} \Psi_s &= \mathrm{IN}_s e^{ikx} + \mathrm{OUT}_s e^{-ikx} \\ \Psi_d &= \mathrm{IN}_d e^{-ikx} + \mathrm{OUT}_d e^{ikx} \end{aligned} \tag{2.4}$$

The wavenumber k is related to the momentum p of the electron by $k = p\hbar$. The amplitudes IN_s and OUT_s represent incoming and outgoing electrons from the scattering region on the source side. Likewise, IN_d and OUT_d represent incoming and outgoing electrons from the scattering region on the drain side. These amplitudes can be grouped into incoming and outgoing states, written as column vectors of eigenvalues in the orthonormal basis formed by the plane wave eigenvectors $\hat{\mathbf{e}} \in \{(e^{ikx}, 0), (0, e^{-ikx})\}$ of the Schrödinger equation [43].

$$|in\rangle = \begin{pmatrix} \mathrm{IN}_s \\ \mathrm{IN}_d \end{pmatrix}, |out\rangle = \begin{pmatrix} \mathrm{OUT}_s \\ \mathrm{OUT}_d \end{pmatrix} \tag{2.5}$$

The relationship between the amplitudes of the incoming and outgoing states is described by the scattering matrix **S** as follows [33, 42]:

$$|out\rangle = \mathbf{S}\,|in\rangle \tag{2.6}$$

where $\mathbf{S}$ is a square, unitary matrix with size equal to the number of states, which encodes all measurable information about transport through the junction [44]. With two incoming and two outgoing states, the scattering matrix can be written as a 2×2 matrix.

$$\mathbf{S} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \tag{2.7}$$

Note that, in general, each of the elements of $\mathbf{S}$ is a function of energy [33].

The transmission probability τ_{sd} for an electron incident on the scattering region from the source electrode depends on the ratio between the outgoing amplitude at the drain and the incoming amplitude at the source.

$$\tau_{sd} = \left| \frac{\mathrm{OUT}_d}{\mathrm{IN}_s} \right|^2 \tag{2.8}$$

In the case where the electron starts in the source, moving towards the drain, the amplitudes of the incoming states are $\mathrm{IN}_s = 1$ and $\mathrm{IN}_d = 0$. The scattering matrix relates the amplitudes of the outgoing states from the amplitudes of the incoming states.

$$\begin{pmatrix} \mathrm{OUT}_s \\ \mathrm{OUT}_d \end{pmatrix} = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \tag{2.9}$$

The resulting transmission probability is $\tau_{sd} = |s_{21}|^2$. Applying the same process to the opposite scenario where the electron starts in the drain, moving towards the source, gives $\tau_{ds} = |s_{12}|^2$. Via the same process it can be inferred that the reflection probabilities are $R_{sd} = |s_{11}|^2$ and $R_{ds} = |s_{22}|^2$. This provides a physical interpretation for the elements of the scattering matrix in terms of transmission and reflection probabilities τ and R.

$$\mathbf{S} = \begin{pmatrix} \pm\sqrt{R_{sd}} & \pm\sqrt{\tau_{ds}} \\ \pm\sqrt{\tau_{sd}} & \pm\sqrt{R_{ds}} \end{pmatrix} \tag{2.10}$$

This gives a clear physical motivation for the unitarity of the scattering matrix: conservation of current requires that, for a given incident electron, the summation of all the transmission and reflection probabilities must be unity [33]. Recall that if a matrix is unitary, the Hermitian adjoint is equivalent to the inverse of the matrix. This condition can be expressed as follows

$$\mathbf{S}^\dagger \mathbf{S} = \mathbf{S}\mathbf{S}^\dagger = \mathbf{I} \tag{2.11}$$

where $\mathbf{I}$ is a square identity matrix of the same dimension as $\mathbf{S}$. By substitution

from equation 2.10, current conservation conditions are shown to hold where **S** is unitary.

$$\begin{aligned} R_{sd} + T_{ds} &= 1 \\ R_{ds} + T_{sd} &= 1 \end{aligned} \tag{2.12}$$

The transmission function $\tau(E)$, encoded in the scattering matrix, can be used to determine the electrical conductance and thermopower of the junction.

2.2 Electrical Conductance

The electrical conductance G of a junction is defined by the ratio of the net current flow I through the junction and the applied bias voltage V.

$$G = \frac{I}{V} \tag{2.13}$$

Recall that near equilibrium, the energies of the electrons in the leads are distributed according to the Fermi-Dirac distribution function (equation 2.1) [11]; Since source and drain may have different chemical potentials, for example due to the presence of a small external bias voltage, their electrons have different energy distributions $f_s(E)$ and $f_d(E)$. If the transmission function of the junction is $\tau(E)$, application of the Landauer formalism yields equation 2.14 for the current I_{sd} due to electrons passing from source to drain [45].

$$I_{sd} = \frac{2e}{h} \int_{-\infty}^{+\infty} \tau(E)\big(f_s(E) - f_d(E)\big)dE \tag{2.14}$$

In the case of an applied bias voltage V across the junction, the chemical potentials of each electrode are shifted by $\frac{1}{2}eV$, and the expression for the conductance is given by the average of the transmission function over an energy range eV around the Fermi energy [11]:

$$G = \frac{1}{V}\frac{2e}{h} \int_{E_F - \frac{1}{2}eV}^{E_F + \frac{1}{2}eV} \tau(E)dE \tag{2.15}$$

in the zero-temperature limit.

2.2.1 Conductance Quantisation

Equation 2.2 implies that a single conduction channel has a finite maximum conductance G_0, known as the 'conductance quantum'. The value of G_0 can be derived by considering the simple case of an ideal 1D wire, length L, between large metallic

leads denoted source (s) and drain (d), illustrated in figure 2.2. The bias voltage V applied to the junction can be written in terms of the difference between the chemical potentials μ_s and μ_d of each lead and the electron charge e.

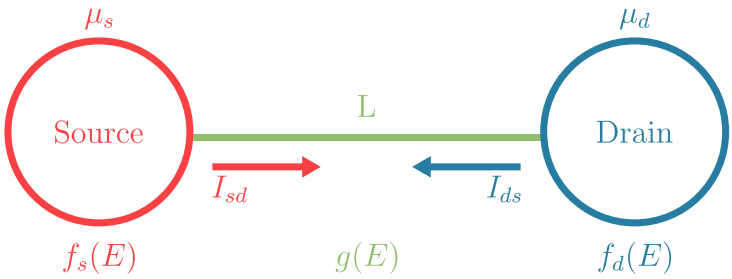


Fig. 2.2: Illustration of a ballistic 1D wire, length L, connected between 'source' and 'drain' leads.

$$V = \frac{\mu_s - \mu_d}{e} \tag{2.16}$$

Within the ideal wire, charge transport is ballistic (with no scattering). The contacts between the leads and the wire are also assumed to be ideal, with no reflections of incident electrons, corresponding to perfect transmission. This can be understood as a special case of the more general situation outlined in section 2.1, where $\tau = 1$. In the absence of scattering or reflections, the bias voltage inside the wire is continuous across its length [46].

Recall that electrons entering from the leads may possess a continuum of energy levels up to the chemical potential μ, described by the Fermi-Dirac distribution (equation 2.1). In the presence of an external bias, the leads have different chemical potentials (μ_s and μ_d) and therefore different electron energy distributions f_s and f_d.

As defined by equation 2.13, the electrical conductance is the ratio of the net current flow I through the junction and the applied bias voltage V (given by equation 2.16). Electrical current is defined as the rate of flow of charge Q through the wire ($I = \frac{\mathrm{d}Q}{\mathrm{d}t}$). The net current I through the wire is the difference between the current flowing from source to drain (I_{sd}) and the opposing current flowing from drain to source (I_{ds}).

$$I = I_{sd} - I_{ds} \tag{2.17}$$

The ideal 1D wire supports a single ballistic conductance channel [47]; the current in each direction can be obtained by integrating $\rho\nu(k)f(k)g(k)$ over the reciprocal k-space, where ρ is the charge density contributed by each electron, $\nu(k)$ is the group velocity, $f(k)$ is the Fermi distribution in the lead (as a function of $k = \sqrt{2mE}/\hbar$)

and $g(k)$ is the density of states in the wire [46].

$$\begin{aligned} I_{sd} &= \int \rho\nu(k) f_s(k) g(k) dk \\ I_{ds} &= \int \rho\nu(k) f_d(k) g(k) dk \end{aligned} \tag{2.18}$$

The charge density for a single electron travelling through the wire is $\rho = e/L$ (e is the electron charge and L is the length of the wire) [46].

The group velocity for a plane-wave electron with wavenumber k is $\nu(k) = \frac{1}{\hbar}\frac{\mathrm{d}E}{\mathrm{d}k}$ (where m is the electron mass) [6].

The density of states, $g(k) = \frac{\mathrm{d}n}{\mathrm{d}k}$ can be found by imposing periodic boundary conditions on the plane wave solution, such that the wavenumber is constrained to integer multiples of $2\pi/L$, i.e. $k = 2\pi n/L$, $n \in \mathbb{Z}$. An additional factor of 2 is also included in the density of states to account for spin degeneracy, resulting in $g(k) = L/\pi$.

Substituting the above equations for ρ, $\nu(k)$ and $g(k)$ into equation 2.18 results in an integral over energy up to the chemical potential in each lead [47].

$$\begin{aligned} I_{sd} &= \frac{2e}{h}\int_0^{\mu_s} f_s(E) dE \\ I_{ds} &= \frac{2e}{h}\int_0^{\mu_d} f_d(E) dE \end{aligned} \tag{2.19}$$

Evaluating these integrals and applying equation 2.17 to obtain the net current results in $I = \frac{2e}{h}(\mu_s - \mu_d)$.

The conductance $G = \frac{I}{V}$ of the 1D ballistic wire can then be obtained by reference to equation 2.16, resulting in the quantum of conductance G_0 [46]:

$$G_0 = \frac{2e^2}{h} \tag{2.20}$$

which represents the upper constraint of the conductance for a single channel in the Landauer formalism.

2.3 Thermopower

Recall from section 1.2.2 that due to the thermoelectric effect, a potential difference (thermovoltage) arises in the presence of a temperature gradient across a junction (and vice versa, application of an external potential establishes a temperature gradient). The Seebeck coefficient S defines the relationship between the open-circuit thermovoltage ΔV and the temperature gradient ΔT (see equation 1.1).

The thermoelectric effect arises due to energy-dependent charge transport, such as is present in semiconductor and molecular junctions. This can be understood in simple terms by considering a two-level model of electron transmission through a molecular junction between two metallic leads (source and drain) at temperatures T_s and T_d respectively, where $T_\mathrm{s} > T_\mathrm{d}$. In the two-level model only the frontier (HOMO and LUMO) orbitals of the molecule are considered. In this case, using the Landauer formalism, the transmission function can be approximated as the sum of two resonances described by Breit-Wigner distributions (a continuous probability distribution function with a single symmetrical peak, also known as the Cauchy distribution) with peaks at the HOMO and LUMO energies. The width of these resonances depends on the strength of the coupling to the leads: a stronger coupling will result in broader distributions [33]. The energy distribution of the electrons in each lead can be modelled by the Fermi-Dirac distribution, and the Fermi energy E_F lies within the HOMO-LUMO gap. This situation is illustrated in figure 2.3. Recall that the majority of electron transport occurs within $\sim k_B T$ of the Fermi energy E_F, and so the origin of the thermovoltage can be understood intuitively by considering the shape of the transmission function T_E in the vicinity of E_F; in the example shown in figure 2.3, E_F is nearest to the HOMO resonance, and the transmission function is therefore increasing in the region below E_F and decreasing above. Hence in this case electrons with lower energies are more likely to pass through the junction, while electrons with higher energies are more likely to be reflected. Since the hotter lead has a greater proportion of higher energy electrons than the colder lead, there is a net current of electrons from drain towards source. In the open circuit case, the steady-state condition is met once an opposing thermovoltage ΔV is established as a result of this transfer of negative charge from drain to source. From equation 1.1, given that $\Delta T = T_\mathrm{s} - T_\mathrm{d} > 0$ and $\Delta V = V_\mathrm{s} - V_\mathrm{d} < 0$, the sign of the Seebeck coefficient can be established as $S > 0$. Similarly, when E_F is nearest to the LUMO, the inverse applies and the Seebeck coefficient has the opposite sign, $S < 0$. [33]

In the low temperature limit, where the transmission function is approximately linear with small variations of energy ($\sim k_B T$) around the Fermi energy, the Landauer formalism can be used to obtain equation 2.21 for the Seebeck coefficient [33, 45]. This low temperature condition is met when the HOMO-LUMO gap is much greater than $k_B T$; typically, organic molecules possess a HOMO-LUMO gap of several eV [11], hence equation 2.21 can be assumed to be valid at room temperature for most molecular junctions.

$$S(E_F) = -\frac{\pi^2 k_B T}{3|e|}\left(\frac{\partial \ln(\tau(E))}{\partial E}\right)|_{E=E_F} \tag{2.21}$$

From equation 2.21 it can be seen that the Seebeck coefficient depends on the

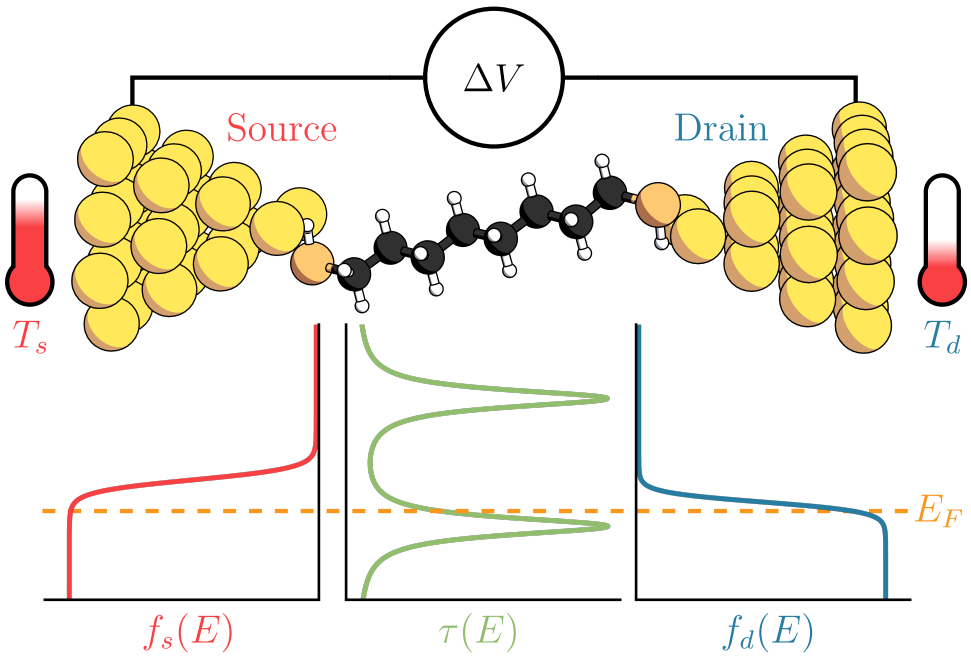


Fig. 2.3: An illustration showing the Landauer formalism description of the thermoelectric effect in a molecular junction with metallic leads in the presence of a temperature gradient.

slope of the transmission function at the Fermi energy. This suggests a promising strategy for the enhancement of the thermoelectric properties of a molecular junction; if the transmission function of the junction can be tuned such that the transmission function has a steeper slope around the Fermi energy, the magnitude of the Seebeck coefficient will be increased accordingly. One way to engineer the transmission function of a molecular junction is through quantum interference effects, introduced in section 2.4.

2.4 Quantum Interference

Perhaps the most intriguing and useful phenomenon associated with charge transport in molecular junctions is quantum interference (QI): in the case of phase-coherent transport, interference effects exhibited by De Broglie electron waves can influence the transmission function of the junction [48], with striking effects on measurable transport properties such as conductance and thermopower. Much like classical interference, for example in acoustic pressure waves, QI can be constructive or destructive: constructive quantum interference (CQI) results in a resonance in the transmission function corresponding to enhanced transmission for electrons in some energy range while destructive quantum interference (DQI) results in an antiresonance and reduced transmission [11].

Molecular junctions typically possess nanoscale dimensions and a wide HOMO-

LUMO gap (compared to thermal variations at room temperature) [33]. As a result, transport remains phase-coherent and thus molecular junctions can exhibit QI effects even at room temperature, prompting significant interest in potential technological applications of leveraging QI to engineer the transmission function and enhance the electrical and thermoelectric properties of molecular devices [48–50]. Additionally, the potential to manipulate the QI behaviour, for example switching from constructive to destructive interference [51], by external electrical, chemical, optical or mechanical gating opens up exciting possibilities for efficient, high-performance molecular devices such as switches [52–55], logic components [56, 57], memory elements [56, 58] and sensors [10, 11].

3. SAMPLE PREPARATION

In order to test and validate the experimental characterisation methods developed in this work, a number of samples were prepared using widely studied materials:

- Au (111)

  Widely used as a control and calibration measurement for thermoelectric SPM measurements, and for calibration of STMBJ via measurements of conductance quantisation.

- 1,8 Octanedithiol (ODT) monolayer on Au

  Used as a benchmark molecule for STMBJ conductance and thermopower measurements in a number of publications in the literature

- C_{60} monolayer on Au

  Established measurements of conductance and thermopower available in literature

- Zinc phthalocyanine (ZnPc) monolayer on Au

  Widely studied and technologically relevant; no studies of the thermopower of a ZnPc monolayer have been reported in the literature to date

Monolayer films of ODT, C_{60} and ZnPc on Au were prepared; the use of a single monolayer allows for a comparison between single-molecule break-junction measurements and contact mode AFM-based measurements of junctions which contain many single-molecule junctions in parallel.

3.1 Self-Assembly

For all the samples characterised in this work, molecules were deposited onto Au(111) substrates by 'self-assembly', a process whereby molecules adsorbed onto a substrate self-organise into stable, ordered configurations which correspond to minima in the free energy associated with non-covalent intermolecular forces [59, 60]. The ODT sample was prepared in the liquid phase by immersion of the substrate in solution (see section 3.3), while the C_{60} and ZnPc samples were prepared by thermal

sublimation in UHV (see sections 3.4 and 3.5) using a borosilicate glass deposition cell with a tantalum wire heating element and a built-in thermocouple for precise temperature control.

After self-assembly, the presence of a monolayer was confirmed via AFM using a 'nano-scratch' procedure as follows: first, a large area (typically 1 μm^2) height map of the surface was captured in tapping mode (to avoid lateral forces which may disturb the layer) in order to ascertain the quality and coverage of the layer. Then, a contact mode scan was performed over a smaller inset region (typically 250 nm^2) with a large contact force setpoint, such that any deposited material was removed from this region by the high lateral forces exerted by the probe. Finally, another large area tapping mode height map was performed, revealing a depression in the surface in the scratch region. By comparing the average height inside and outside the scratch region (excluding the raised walls of the scratch which are higher than average due to the excess material 'piling up' during the scratching process) the thickness of the layer can be calculated.

3.2 Au(111) Substrates

All of the samples prepared for this work used crystalline (111) orientation Au thin films as a substrate for molecular deposition.

Au is a useful substrate for studies on molecular thin films because it is electrically conductive, chemically inert, highly corrosion-resistant, and offers high binding affinities to a range of commonly used anchor groups such as thiols [61]. Epitaxial Au thin films can be deposited on various substrates via thermal evaporation, but the resulting Au surface is typically too rough for the growth of a well-ordered SAM [61], and further processing is required. For samples prepared by self-assembly in solution, a template-stripping procedure was used to produce ultra-flat Au thin films on Si. For samples prepared by thermal sublimation in UHV, a sputtering and annealing procedure was used to preprocess commercially obtained Au thin films on mica. The template stripping and UHV sputtering and annealing processes are detailed in sections 3.2.1 and 3.2.2

3.2.1 Template Stripped Au

The template stripping process is used to transfer the flatness of a Si or mica template to a thermally evaporated Au layer, resulting in ultra-flat Au surfaces with sub-nanometre roughness over large areas [62, 63].

The template stripping process used in this work is as follows: a 200 nm thick Au film was deposited by evaporation onto a Si wafer, and an epoxy resin adhesive (EPO-TEK 353ND) was used to affix a second small piece of Si wafer to the top

of the Au layer. Immediately prior to use, a Au substrate was then prepared by removing the affixed Si, the process of which lifted away the deposited Au and revealed the pristine surface from the original Si-Au interface. An AFM height map of a typical Au substrate prepared by this method can be found in figure 3.1.

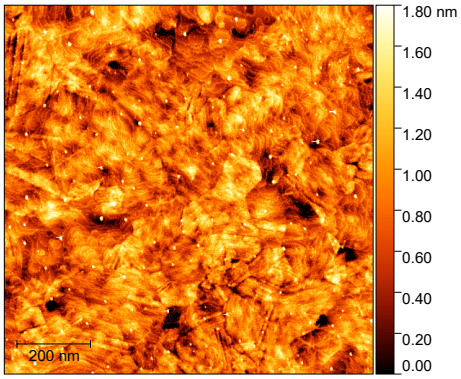


Fig. 3.1: AFM height map of a Au(111) substrate prepared by the template stripping process. Image obtained by Charlie Wells.

3.2.2 Au on Mica

Au(111) on mica substrates were purchased commercially, and pre-processed in UHV immediately prior to use by sputtering and annealing to clean and flatten the surface.

The specific pre-processing procedure used for Au on mica throughout this work is as follows: two cycles of sputtering with an energy of 0.6 KeV under an Ar gas pressure of 5×10^{-6} mbar were performed for 20 minutes each. The substrate was then thermally annealed at 500°C for 30 minutes, before cooling at a controlled rate of $1°Cs^{-1}$. An AFM height map of a typical Au on mica substrate prepared by this method can be found in figure 3.2.

The author would like to thank Charlie Wells for preparing the Au on mica samples.

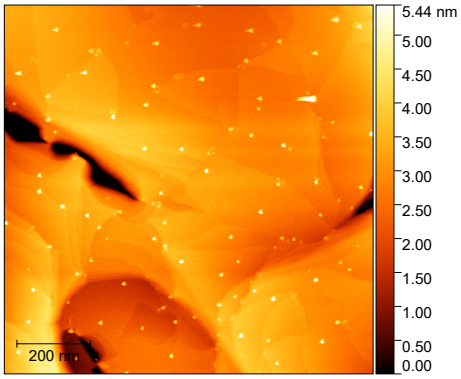


Fig. 3.2: AFM height map of a Au(111) on mica substrate. Image obtained by Charlie Wells.

3.3 1,8 Octanedithiol Monolayer on Au

1,8 Octanedithiol (ODT), illustrated in figure 3.3, is an alkanedithiol molecule which has been widely studied in molecular junction contexts as a 'molecular wire' [25,31, 64–71].

ODT is especially convenient for STMBJ measurements due to the thiol anchor groups at each end of the molecule which readily bind to Au electrodes, and in fact the molecule was among the first studied using the technique [24]. Typical single-molecule conductances around $4 \times 10^{-5} G_0$ have been reported in the literature [67], however, a range of other conductance levels have been identified; the cause of these differing conductance values is not fully understood [69], and a number of mechanisms have been proposed, including variations in the conformation of the molecule or the coupling to the electrodes [68,69,72]. Transport through ODT single-molecule junctions is HOMO dominated, with a small positive Seebeck coefficient of around 5 $\mu V K^{-1}$ reported for STMBJ measurements [73]. AFM-based thermopower measurements of multi-molecule junctions with an ODT SAM have not previously been reported.

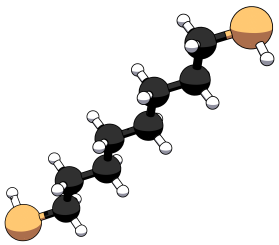


Fig. 3.3: An illustration of the chemical structure of 1,8 octanedithiol

In this work, a monolayer was formed by self-assembly in solution: a template stripped Au substrate was immersed in a 5 μM solution of ODT in 99.9% THF for 24 hours, before rinsing in THF and acetone and drying under nitrogen flow. Samples were prepared immediately prior to use, and stored in a vacuum desiccator between measurements.

An AFM nanoscratch test was used to confirm the formation of a monolayer. AFM height maps of the surface before and after the scratch are provided in figure 3.4. Histograms of the measured height inside and outside the scratch region were extracted from figure 3.4**b**. These can be seen in figure 3.5. The measured height of the layer is 1.1 ± 0.6 nm. This suggests that the ODT molecules are in an 'upright' configuration with only one thiol group bound to the Au; this is consistent with the typical monolayer thickness of 1.4 nm for such a configuration reported in the literature [74].

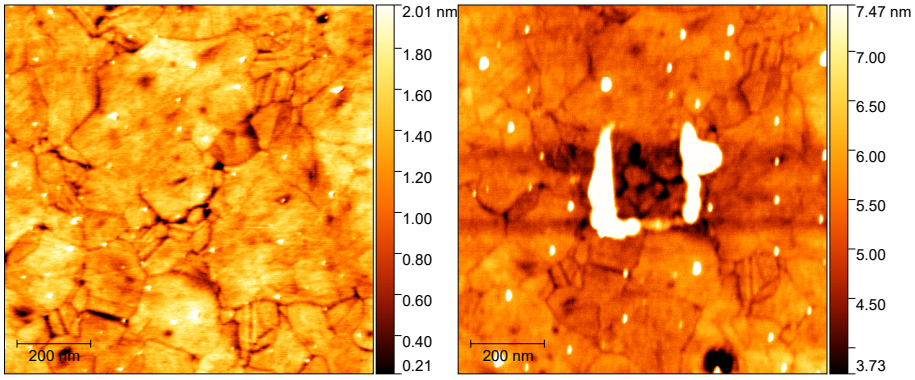


Fig. 3.4: AFM height maps of the prepared ODT layer **(a)** prior to nano-scratch **(b)** after nano-scratch.

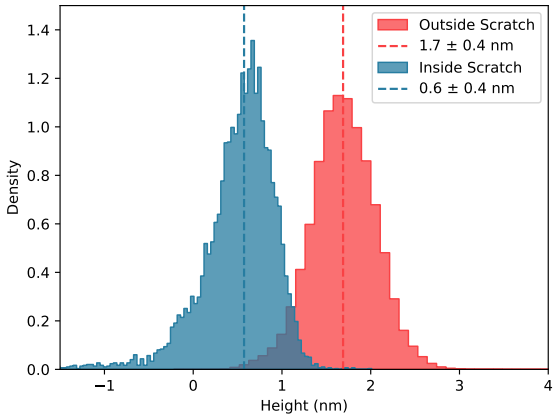


Fig. 3.5: Height distributions inside and outside the nanoscratch region for the ODT sample, indicating an average thickness of 1.1 ± 0.6 nm for the deposited layer.

3.4 C_{60} Monolayer on Au

C_{60}, or 'Buckminsterfullerene', is a spherical molecule consisting of 60 carbon atoms bound by sp^2 hybridisation into a 'cage' structure comprised of 5 and 6 membered rings [75] (illustrated in figure 3.6). Since it was identified in 1985 [76], C_{60} has received significant research attention, and molecular junctions based on C_{60} have been extensively studied both theoretically and experimentally. Multiple groups have measured the thermopower and electrical conductance of various configurations of C_{60} junctions [77–80]. Transport in C_{60} junctions is understood to be LUMO dominated, and these were among the first reported molecular junctions to exhibit a negative Seebeck coefficient [78], with typical reported values in the range -9 μVK^{-1} to -18 μVK^{-1} [77, 78]. Single C_{60} junctions also have a relatively high conductance, with typical reported values of around 0.1 G_0 [77].

The C_{60} sample used in this work was a monolayer on a Au on mica substrate, prepared in UHV as follows: first the substrate was prepared according to the pro-

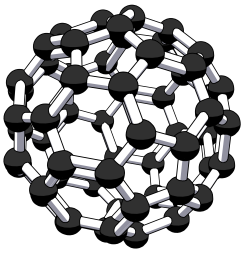
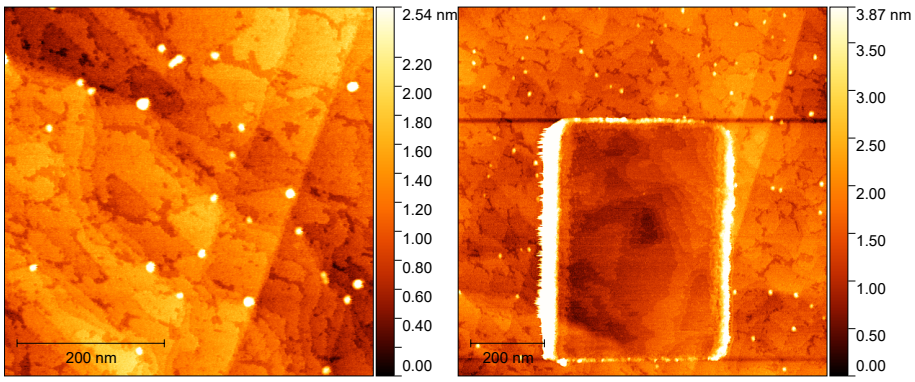


Fig. 3.6: An illustration of the chemical structure of C_{60}

cedure outlined in section 3.2.2. Then C_{60} was deposited by thermal sublimation, with a deposition cell temperature of 290°C sustained for 10 minutes. Samples were prepared immediately prior to use, and stored in a vacuum desiccator between measurements.

An AFM nanoscratch test was used to confirm the formation of a monolayer. AFM height maps of the surface before and after the scratch are provided in figure 3.7. Histograms of the measured height inside and outside the scratch region were extracted from figure 3.7**b**. These can be seen in figure 3.8. The measured height of the layer is 0.6 ± 0.3 nm, consistent with the value of 0.8 nm typically reported in the literature [81, 82]. The author would like to thank Charlie Wells for preparing the C_{60} sample.



(a) *(b)*

Fig. 3.7: AFM height maps of the prepared C_{60} layer **(a)** prior to nano-scratch **(b)** after nano-scratch. Images obtained by Charlie Wells.

3.5 ZnPc Monolayer on Au

Zinc phthalocyanine (ZnPc) is a metal complex consisting of a phthalocyanine macrocycle with a zinc coordination centre (illustrated in figure 3.9). Phthalocyanine metal complexes have shown potential in organic optoelectronics [83] and

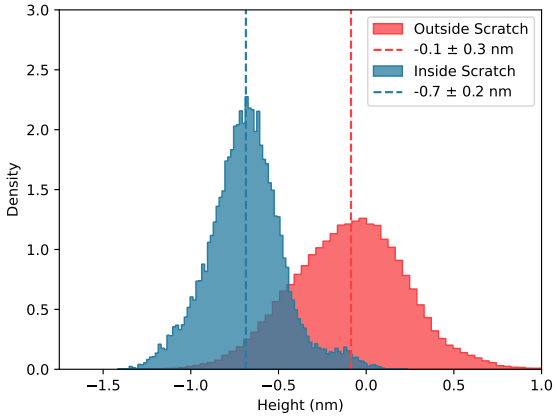


Fig. 3.8: Height distributions inside and outside the nanoscratch region for the C_{60} sample, indicating an average thickness of 0.6 ± 0.3 nm for the deposited layer.

molecular electronics [84] applications, both in thin-film and single-molecule configurations [85].

ZnPc has been studied experimentally in the literature, both as part of a monolayer using SPM [86] and in a 'molecular wire' configuration using break-junction experiments [85]. However, no experimental characterisation of the thermopower of ZnPc at either the single-molecule or monolayer level was found in the literature. Theoretical studies predict HOMO mediated transport [85, 87]; hence a positive Seebeck coefficient is expected.

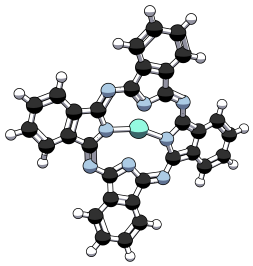


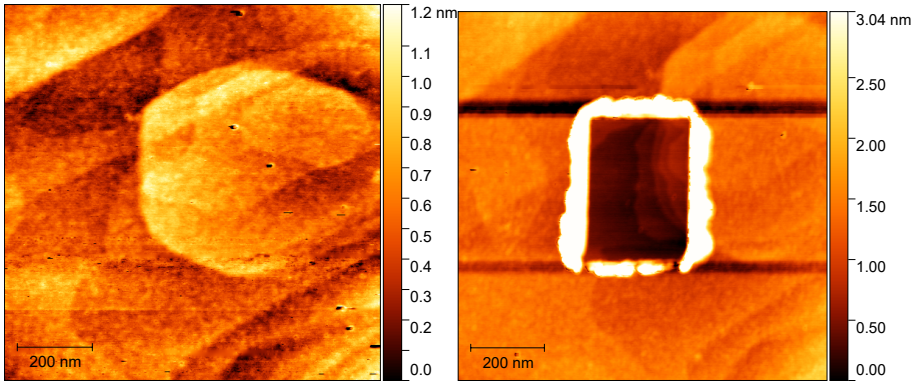
Fig. 3.9: An illustration of the chemical structure of zinc phthalocyanine

The ZnPc monolayer was prepared as follows:

First, the Au on mica substrate was prepared according to the procedure outline in section 3.2.2. A thick multilayer of ZnPc was deposited by thermal sublimation in UHV, with a deposition cell temperature of 320°C sustained for 60 minutes. The sample was then heated to 400°C in UHV for 12 hours to sublimate all but the base monolayer. Samples were prepared immediately prior to use, and stored in a vacuum desiccator between measurements.

An AFM nanoscratch test was used to confirm the formation of a monolayer. AFM height maps of the surface before and after the scratch are provided in figure 3.10. Histograms of the measured height inside and outside the scratch region were

extracted from figure 3.10**b**. These can be seen in figure 3.11. The measured height of the layer is 1.0 ± 0.2 nm. A monolayer thickness of 1.46 ± 0.25 nm has been reported in the literature for the case where the molecules are oriented in a flat configuration against the surface [88]. The author would like to thank Charlie Wells for preparing the ZnPc sample.



(a) *(b)*

Fig. 3.10: AFM height maps of the prepared ZnPc layer **(a)** prior to nano-scratch **(b)** after nano-scratch. Images obtained by Charlie Wells.

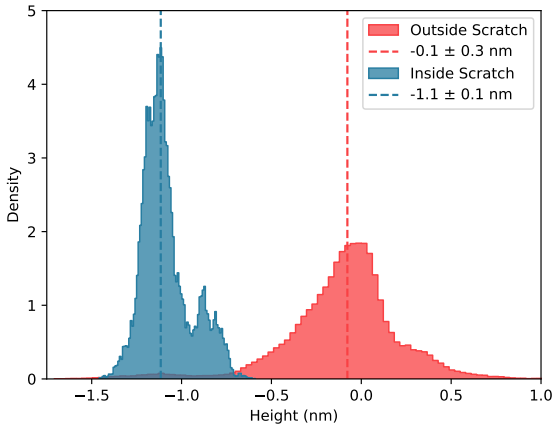


Fig. 3.11: Height distributions inside and outside the nanoscratch region for the ZnPc sample, indicating an average thickness of 1.0 ± 0.2 nm for the deposited layer.

4. A CUSTOM BREAK JUNCTION FOR THE CHARACTERISATION OF SINGLE-MOLECULE JUNCTIONS

From its inception by Binnig in Rohrer in 1981, the potential of STM not just as a microscope but as a versatile local probe for spectroscopy was recognised; in fact, the STM was originally conceptualised as an instrument for local tunnelling spectroscopy at the nanoscale, with its potential for topographic imaging realised only later in the development process [19]. In the decades since the introduction of STM, myriad advanced measurement modes have been developed, extending the applications of the technique far beyond those originally anticipated. Of particular interest to the field of molecular electronics is the STM break-junction (STMBJ) technique, in which individual molecules are 'wired' between the metallic STM tip and sample substrate, thereby facilitating the experimental study of single-molecule junctions.

The STMBJ technique, introduced in 2003 by Xu and Tao [24], leverages the fine-positioning capabilities of the STM to study the formation and breaking of single-molecule junctions. In its most basic implementation an STMBJ measurement cycle consists of a controlled contact, or 'crash', between the tip and substrate followed by withdrawal at a steady speed (on the order of 1-100nms^{-1}) and constant bias V until the contact is broken, while the current I flowing between the tip and substrate is recorded. The conductance $G = I/V$ as a function of Z during the withdrawal can then be calculated from the current data. Features in the resulting $G - Z$ trace encode a wealth of information about the junction formation and breaking process. As the junction breaks, the contact reduces from a bulk metallic contact to a few-atom junction, corresponding to a rapid decrease in the measured conductance. As the metallic junction is further stretched the number of atoms in the junction decreases. In the few-atom regime, each atom can be approximated as a discrete near-perfect conductance channel with conductance $G = G_0$ (see section 2.1); plateaus in the $G - Z$ trace at integer multiples of G_0 can be observed as the number of atoms in the junction decreases, until the junction is reduced to a single atom, corresponding to a conductance plateau at $G = 1G_0$. Beyond this point, in the absence of molecules, the junction will break entirely and the $G - Z$ trace will exhibit a rapid exponential decay corresponding to the through-space tunnelling regime. If some target molecule (which will typically have end groups chosen

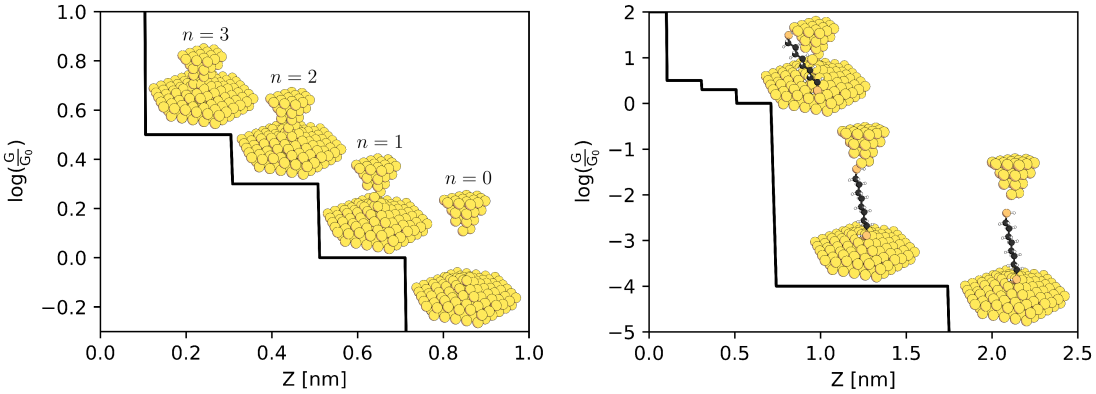


Fig. 4.1: Illustrations of an idealised STMBJ conductance-distance (G-Z) trace for a metallic tip and substrate: **(a)** in the absence of a molecule **(b)** in the presence of a molecule

for affinity to the electrode material) is present, either in solution or adsorbed to the substrate, then a molecular junction may be formed after the breaking of the metallic junction. This can be detected by the presence of additional features in the $G-Z$ trace in the region $G < G_0$. A conductance plateau corresponding to a single molecule junction provides information about the conductance and breaking length of the molecular junction.

Figures 4.1a and 4.1b illustrate the anatomy of a typical STMBJ $G-Z$ withdrawal trace with and without a molecule present in the junction.

A bespoke STM, optimised for STMBJ measurements, has been developed as a part of this work and is intended to serve as a versatile and extensible platform for the exploration of advanced STMBJ techniques. A number of novel fabrication approaches, such as the use of steel reinforced epoxy granite in the construction of the STM head, were implemented to achieve high performance at low cost. This instrument, coupled with machine-learning augmented data processing techniques, allows the study of electrical conductance in a single-molecule junction throughout the process of formation, stretching and eventual breaking. This chapter will detail the design, construction and operation of the STM break-junction (STMBJ), followed by the development and benchmarking of a data analysis workflow based on unsupervised clustering in chapter 5.

4.1 Design Motivations and Specifications

The primary motivation for building the STMBJ was to provide a platform for single-molecule conductance measurements, with the possibility to extend the electronics and software to support more sophisticated multivariate measurement modes as

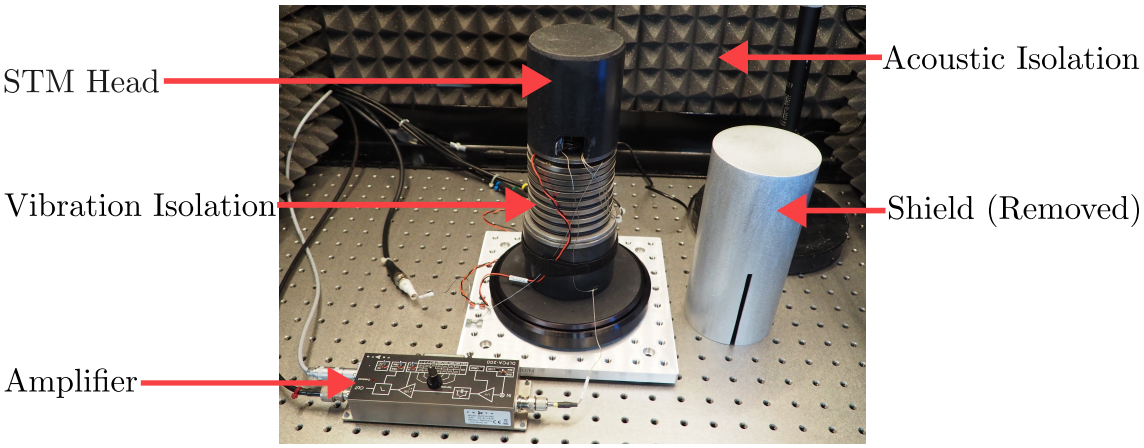


Fig. 4.2: A photograph of the STMBJ with the EMI shield removed

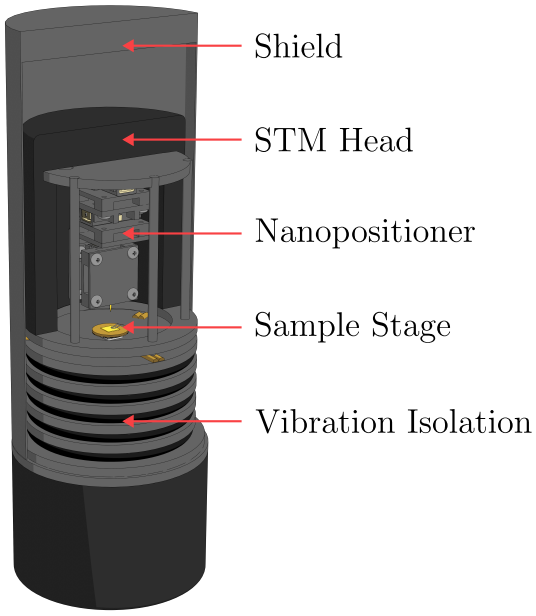


Fig. 4.3: A labelled diagram of the STMBJ, partially cut away to reveal the internal structure

required. At the outset of this work, the labs of the Ben Robinson group at Lancaster University were well equipped for characterisation of multi-molecular junctions with AFM, but capabilities for single-molecule characterisation were more limited and no dedicated STMBJ instrument was available. The instrument developed in this work will serve to fulfil this role in future, and the flexible nature of the electronics and hardware will allow the instrument to grow with the demands of future research directions explored by the group. It is hoped that the design strategies identified to reduce the cost and fabrication effort required to build the instrument will prove useful to those in other labs hoping to extend their capabilities in the same way.

Before beginning the development of the instrument, several important requirements were identified which guided the entire design and manufacture process. These specifications are outlined below:

4.1.1 Stability

Many STM imaging modes operate in a closed-loop 'constant-current' mode, where a feedback process adjusts the z position of the probe such that the tunnelling current is held constant. This controls the tip-sample separation during measurement, providing some resilience to instabilities caused by drift or external perturbations. In the withdrawal phase of an STMBJ I-Z measurement (after initial contact to the sample is established), the system operates in an open-loop mode, where feedback cannot be used since the tunnelling current is the measured parameter. Therefore, it is essential that external mechanical disturbances such as drift and vibrations are minimised: the length scale of a typical molecular junction is on the order of 1 nm, and small (0.1 nm) perturbations in the tip-sample separation can disturb or even prematurely break the junction.

Higher frequency mechanical disturbances from the environment, such as building vibrations and acoustic noise, can be effectively mitigated by proper environmental isolation. Low frequency 'drift', caused primarily by non-ideal characteristics of piezoelectric positioners and thermal expansion of the STM itself, is much more challenging. One approach to avoid drift from influencing the junction is to ensure the withdrawal rate is high compared to the rate of drift, such that the total displacement during any given measurement is small. However, this typically precludes more advanced measurement modes [31] which often require pausing the withdrawal at intervals throughout the trace to characterise signal fluctuations [89] or modulate external parameters [73].

Due to the above considerations, mechanical stability was carefully considered at each stage of the design process. Inspiration was drawn from the design of ultra-high precision machine tools [90, 91] and existing high-stability SPM designs reported in the literature [31, 92].

4.1.2 Ease of Manufacture

The STMBJ was designed to be easy to manufacture without the use of specialised tools, so that it can be reproduced and modified by others in future. This requirement led to optimisations in the design of an existing open source nanopositioner module (see section 4.2.6), and to the development of a novel STM head which can be cast from epoxy granite using 3D printed moulds (see section 4.2.3).

4.1.3 Cost

Commercially available STM systems, or components thereof (such as nanopositioners) typically represent a significant financial investment for most research groups operating with often limited budgets. The STMBJ hardware here was designed to be inexpensive to construct, without compromising on performance, so that the setup can be replicated and modified for minimal cost. Excluding the supporting electronics and other external components, the STMBJ cost less than £500 to build.

4.1.4 Flexibility

The STMBJ was intended to be highly extensible, such that new modes of operation could be added to suit future characterisation requirements. This consideration is represented in several design aspects of the instrument:

- The real-time control, modulation and sensing requirements of complex STMBJ experiments may exceed the capabilities of the existing control electronics. Hence the electronic control system was built from a number of off-the-shelf subsystems, each of which can be upgraded or replaced individually as required for more sophisticated measurements.

- The possibility of studying thermopower and thermal transport in future was also considered, and a variable temperature sample stage was implemented to facilitate such measurements.

- Most STMBJ modes require only one degree of freedom in the nanopositioner. However, the instrument may be adapted to perform more conventional STM imaging modes in future, which require 3D positioning. Therefore a long-range 3-axis nanopositioner was built, rather than the typical single-axis positioner used for STMBJ. This also provides coarse positioning in 3D space, potentially enabling automated sampling at multiple positions, or even automated sample changes.

- The control software was built in an extensible, modular manner, and new measurement routines can easily be defined without rewriting any of the lower

level control software. The software is also hardware agnostic: changing the SPM controller/DAQ hardware only requires a new wrapper for the STM control functions to be defined.

4.2 Hardware

This section details the design and construction of the hardware portion of the STMBJ, which consists of the environmental isolation systems, the STM itself and the electronic measurement and control systems.

4.2.1 Environmental Isolation

The STM is a very sensitive instrument, operating at nanoscale spatial precision while measuring small tunnelling currents in the pA to nA range. As explained in section 4.1.1, the quality of an STMBJ measurement is even more sensitive to the mechanical stability of the junction than a typical STM measurement since STMBJ operates open-loop. Proper isolation from environmental disturbances, both mechanical and electromagnetic, is essential for adequate performance.

In this context, the most important environmental disturbances consist of the following:

1. Low frequency (< 0.1 kHz) mechanical vibrations, typically coupled through the structure of the building

2. Mid-high frequency ($0.1 - 10$ kHz) mechanical vibrations coupled acoustically through the air

3. Mechanical drift (typically following a 1/f 'flicker noise' frequency dependence) from various sources including piezo creep and thermal expansion due to temperature fluctuations

4. Electrical interference from various sources including 1/f and digitisation noise in the acquisition signal chain, and electromagnetic pickup at the junction and cables leading to the amplifier

5. Noise inherent in the junction, for example 1/f noise originating from fluctuations in the junction configuration. These noise sources cannot be eliminated at room temperature, and actually provide a valuable additional source of information [89, 93].

Low Noise Laboratory and External Isolation

In order to ensure a stable, low-noise environment, the STMBJ was situated in one of three isolated measurement 'pods' in the Lancaster university Isolab facility, a specialised low-noise laboratory which features acoustic and EMI shielding, filtered power supplies, and high thermal stability due to the enclosed nature of the pod. The basement of each pod houses a 50 ton concrete block suspended on air springs, upon which the optical tables supporting instrumentation are mounted. Any other equipment stands on a 'floating floor', separate from the isolation block, and the instrumentation is operated from outside the pod to minimise disturbances [94]. In addition to the isolation afforded by design of the Isolab, the STM itself was placed within an acoustically and thermally insulated enclosure to further reduce acoustic noise pickup and thermal drift.

Within the enclosure, the instrument rested on a viscoelastic isolator situated atop a small aluminium breadboard (Thorlabs MR2020U/M) which sat on three Nitrile O-rings on the surface of the main optical table (a simple strategy to ensure self-levelling, damping and decoupling from the tabletop). All cables connecting to the instrument were firmly clamped at several points (the tabletop, the breadboard, the instrument base and the interior of the scan head) before reaching the junction to minimise coupling of vibrations along their length.

Integrated Vibration Isolation

In addition to the measures described in section 4.2.1, the design of the STMBJ itself included integrated environmental isolation measures, intended to provide an additional layer of isolation and enable operation outside of a specialised low-noise lab if required in future.

The integrated vibration isolation system consisted of an adjustable stacked viscoelastic damper made from a stack of up to 9 laser-cut 316 stainless steel discs, each weighing 220 g, separated by nitrile O-rings as damping elements. Two different sets of O-rings with different thickness (5 mm and 2 mm) were used, to produce a two-stage system with two different transfer functions, reducing overall transmission. Such isolators have been studied experimentally in the literature and found to be highly effective for vibrations above 10 Hz [95].

Integrated EMI Isolation

The entire STMBJ, including the viscoelastic stack, was fully encapsulated by a Faraday shield which could be connected to the amplifier ground to minimise EMI pickup. A 2 mm wide slot in the shield allowed cables to be routed to the internals. The shield was made from an aluminium tube (89 mm outer diameter, 1.6 mm wall

thickness) sealed at one end by a press-fit aluminium insert. The entire instrument rested on a plinth made from cast epoxy granite (see section 4.2.4), with a stainless steel upper surface integrated during the casting process. The stainless steel part was made from two laser-cut discs of different diameters, creating a 'lip' at the edge of the top surface - this served as an interface to securely mount the shield during operation, and provided a good electrical connection between the base and the shield for full encapsulation of the instrument. The steel internal structure of the STM head, the aluminium frame of the scan unit and the baseplate of the instrument were also electrically connected together and to ground. As well as providing EMI shielding, this is a safety measure as it prevents any of the external metal surfaces from accidentally becoming 'live' with the high-voltage piezo drive signals in the event of a loose connection.

In addition to the integrated shielding, a set of custom flexible (stiff cables can couple additional vibrations into the instrument) shielded cables were fabricated to carry the nanopositioner driving voltages, sample bias and tunnelling current. The cables were made from single-core brass wire in a CuNi shield, with an overall diameter of 0.65 mm, sourced from CMR-Direct (product code 02-32-051). Cables were directly soldered to the corresponding components within the STM, and the free ends were terminated with crimped MMCX coaxial connectors for interfacing with the control and acquisition electronics.

4.2.2 Baseplate and Kinematic Mount

The baseplate of the instrument, upon which the STM head rests, was made from a laser-cut 316 stainless steel disc 5 mm thick and 85 mm in diameter, weighing 220g. The temperature-controlled sample stage, details of which can be found in section 4.2.7, was affixed directly to the centre of this plate using cyanoacrylate adhesive.

A simple kinematic coupling scheme was implemented to provide a stable, reproducible means of mounting the STM head on the baseplate. The kinematic mount consisted of three pairs of brass half cylinders mounted to the baseplate in a triangular configuration, which interface with three ball bearings on the bottom of the STM head to provide a fully constrained coupling.

During operation the baseplate sits on top of the viscoelastic vibration isolation stack described in section 4.2.1. Figure 4.4 shows the base plate in-situ.

4.2.3 Steel-Reinforced Epoxy Granite Scanning Tunnelling Microscope Head

The STM head, shown in figure 4.7, holds the tip and scan unit in place above the sample holder during operation. In order to maintain a stable tip-sample junction, great care must be taken in the design and construction of the STM head to

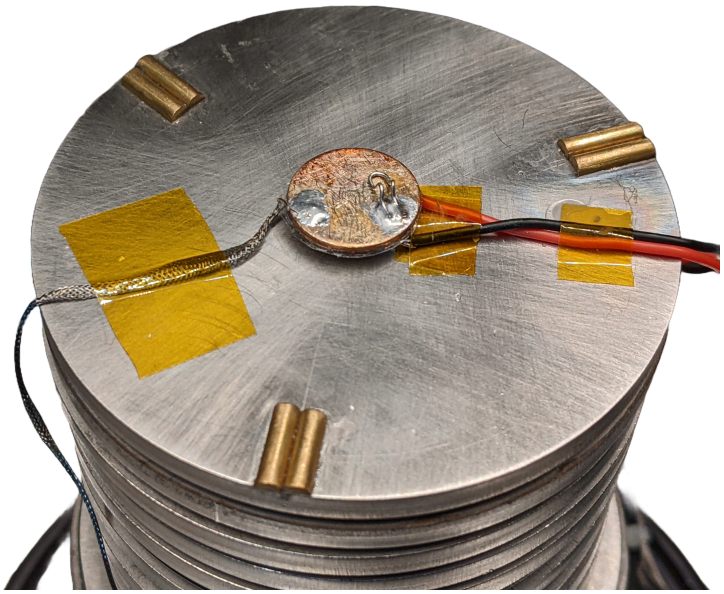


Fig. 4.4: A photograph of the STMBJ base plate in-situ atop the isolation stack, with the sample holder attached in the centre and the kinematic mounting features visible at the perimeter.

minimise the effect of external disturbances (such as mechanical vibrations and temperature variations) on the position of the scan unit. Thermal expansion, response to mechanical vibrations and the stability of the coupling between the head and base plate must all be carefully considered to achieve optimal performance [96]. These properties of the STM head depend both on the design of its structure and on the properties of the materials used in its construction.

In selecting the materials for the STM head, inspiration was drawn from the design of ultra-high precision machine tools, where significant resources have been invested to develop advanced materials for dimensionally stable, rigid, well damped structures [91]. Among the most important material properties for a precision machine are high stiffness, high damping capacity, and low coefficient of thermal expansion [97].

Stainless steel alloys are a popular choice for the construction of both commercial and lab-built STM heads, since they are inexpensive, easy to machine and relatively rigid. However, stainless steel offers poor damping performance, resulting in a tendency for 'ringing' behaviour in response to mechanical disturbances around the resonance modes of the structure [98]. Alternatively, cast iron has traditionally been favoured in the manufacture of precision machine tools, in part due to the damping provided by the presence of graphite structures within the material [97,99]. Cast iron has seldom been utilised in the construction of an STM head, however, likely due to fabrication challenges: producing cast iron parts is an involved process

which requires the fabrication of expensive moulds, making it impractical for small one-off parts [100]. Natural granite offers superior rigidity, damping and thermal stability to most metals, including cast iron, and is often used in ultra-high precision machining and metrology applications [91, 97], but making complex machined parts from granite is challenging and often prohibitively expensive for many labs.

Epoxy granite (EG) is a composite material consisting of granite aggregate bound in an epoxy polymer resin; EG provides superior damping performance to both cast iron and natural granite, but typically has only around $\frac{1}{3}$ of the static stiffness of cast iron [101]. Steel reinforced epoxy granite (SREG) aims to combine the stiffness of steel with the desirable damping properties of EG by integrating a steel reinforcement structure into the cast EG part [100]. With careful design, SREG can achieve stiffness exceeding that of an equivalent cast iron structure [101, 102]. Importantly, SREG is much more accessible for in-house fabrication by individual laboratories than cast iron or natural granite; parts can be cast at room temperature, allowing the use of 3D printed moulds, and the raw materials are inexpensive and readily available. Complex geometries can easily be obtained without the use of advanced machine tools, and additional components such as mounting hardware or threaded inserts can be inserted into the mould to produce highly functional parts with minimal manufacturing effort. The main anticipated limitation of SREG for STM applications is relatively poor thermal stability, due to mismatched coefficients of thermal expansion between the EG and the steel reinforcement structure [103]. In this work this issue was mitigated by ensuring that no components which were likely to introduce a thermal gradient were mounted to the STM head, and by operating the STMBJ in an enclosure with a stable ambient temperature. In future, multiple avenues for thermal drift compensation are available, such as integrating carbon-fibre composites with negative coefficients of thermal expansion [103].

The SREG STM head developed in this work is composed of a stainless steel internal frame encapsulated in cast epoxy granite. A review of the relevant literature revealed no previously reported use of SREG in an SPM; this instrument therefore serves to demonstrate the feasibility of the SREG approach for the construction of a high performance SPM at low cost, using only simple tools.

The steel inner structure of the STM head consists of two laser-cut 316 stainless steel discs separated by steel rods. A lathe was used to add a countersink to one end of three of the rods to hold a steel ball bearing (retained with cyanoacrylate adhesive): these ball bearings align with the kinematic mating surface on the baseplate described in section 4.2.2. The mounting holes for the positioner assembly (see section 4.2.6) were hand drilled in the upper plate and then tapped with M2.6 threads: hand-drilling was used instead of laser-cutting for these holes because the heating from laser-cutting can harden the steel edges around small holes and make thread

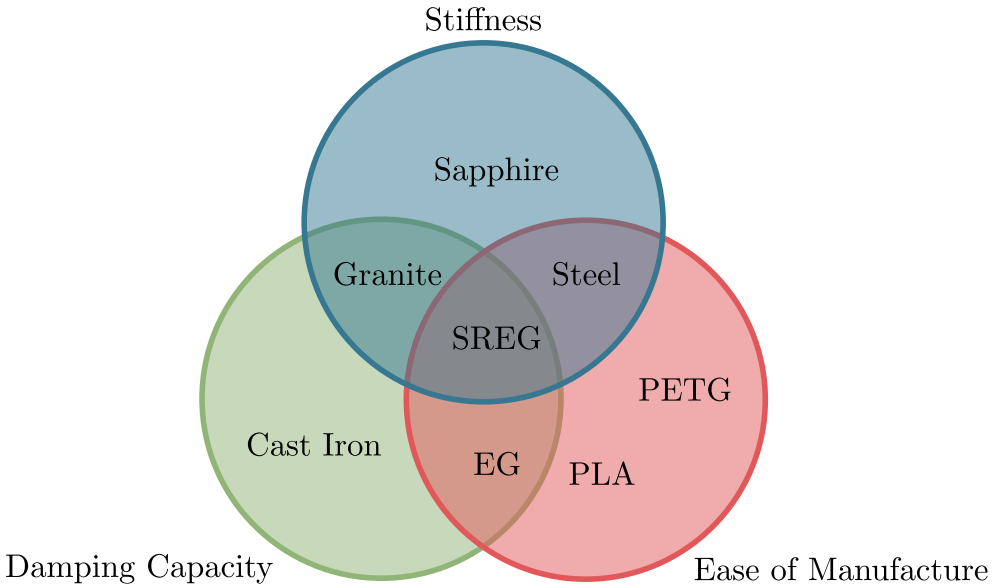


Fig. 4.5: A Venn diagram illustrating the suitability of SREG for STMBJ compared to a range of other available materials. SREG provides an ideal compromise between stiffness, damping capacity and ease of manufacture.

tapping challenging. A photograph of the inner frame before it was embedded in epoxy granite is shown in figure 4.6b.

In order to ensure good bonding between the metal inner frame and the cast epoxy granite, the metal parts were prepared as follows: all parts were sanded/roughed to 60 grit, and a checkerboard pattern of approximately 1 mm deep grooves were cut into any large flat surfaces (such as the faces of the steel plates) with an angle grinder to provide a mechanical key for the epoxy granite. The parts were ultrasonically degreased and cleaned in acetone immediately prior to assembly. A small amount of cyanoacrylate adhesive was used on each of the push-fit mating surfaces between the end plates and the rods to ensure that they remained in place during the curing process.

After assembly and preparation, the metal frame was encased in a 3D-printed mould consisting of three interlocking outer segments and an inner plug. In order to ensure a smooth surface finish on the cast EG the inner faces of the mould were prepared by filling any imperfections with polyfill and then sanding all surfaces to 600 grit. Figure 4.6b shows the metal frame installed in the mould, before the outer parts were clamped together around the frame using cable-ties.

Epoxy granite (EG), formulated according to the process outlined in section 4.2.4, was then gradually added to the mould under agitation until no more could be pressed in. The full mould was agitated for 1 hour, and then left to cure for 48 hours. The mould was then opened and the cast part was removed and washed in acetone before the post-cure, which consisted of 5 hours in an oven at 60°C.

(a) *(b)*

(c) *(d)*

Fig. 4.6: **(a)** The steel inner structure of the STM head prior to embedding in epoxy granite **(b)** The steel inner structure of the STM head partially installed in the 3D printed mould for the epoxy granite outer frame **(c)** The sealed mould for the SREG STMBJ head, prior to epoxy pouring **(d)** The cast SREG STMBJ head curing inside the mould

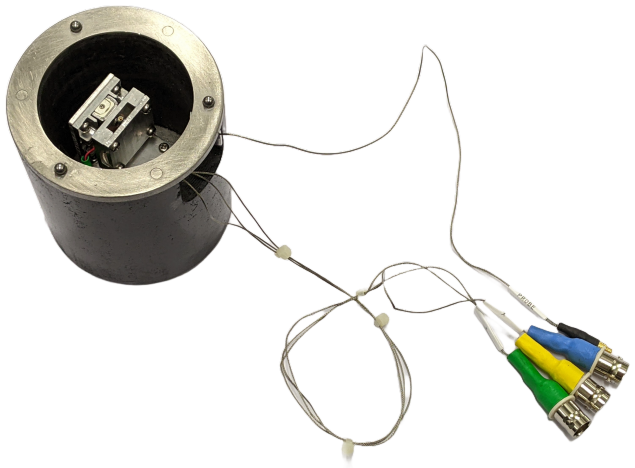


Fig. 4.7: A photo of the assembled SREG STMBJ head, with scan unit and cabling installed.

The EG surfaces of the assembled SREG STM head were then finished by sanding to 120 grit on a belt sander and coating with acrylic lacquer.

The finished SREG frame is shown in figure 4.7.

4.2.4 Epoxy Granite Preparation

The formulation of the epoxy granite used in the fabrication of the STM is detailed in table 4.1.

Proportion (By Volume)	Material
70%	Sharp Sand
10%	Black Mica
15%	Epoxy Resin (EL2, Easy Composites Ltd)
5%	Slow Hardener (AT30, Easy Composites Ltd)

Tab. 4.1: The formulation of the epoxy granite composite used in this work.

The process used to produce cast epoxy granite parts using this formulation was as follows: a mould to produce the required cast part geometry was designed using 3D CAD. Supports for any additional components to be cast into the part (such as the steel reinforcement frames used in the STM head, described in section 4.2.3) were built into the mould, which was comprised of multiple interlocking segments. The moulds were then 3D printed in PLA on an FDM printer (Creality Ender 3 V2). The internal surfaces of the printed moulds were post-processed by filling any small imperfections with a polyfiller and sanding to 600 grit.

The internal volume of the mould was determined from the CAD model and used to calculate the required volume of each component of the epoxy granite mix according to the formulation specified in table 4.1. The sharp sand (selected for its large, angular grains and higher granite/quartz content compared to soft sand) was measured out by volume and dried in an oven at 200°C for 3 hours. The sand was then mixed with black mica. The epoxy was mixed with slow hardener in a 100:30 ratio in a separate container before pouring in the sand/mica and stirring for 10 minutes to ensure an even mixture.

The epoxy granite mixture was gradually poured and compacted into the assembled mould with intermittent agitation (using vibrations from a hand sander held against the mould) to ensure no voids remained. The filled mould was agitated for one hour, rotating intermittently, and then left for 48 hours to cure. The mould was then opened and the part removed and washed in acetone. The part was then placed in an oven for a 5 hour post-cure at 60°C. The cured part was finished by sanding to 120 grit with a belt sander, cleaning with acetone, and finally spraying with clear acrylic lacquer.

4.2.5 Probe Holder

The probe holder for the STMBJ was made from a beryllium copper pin-socket (Samtec SS-101-T-2), which held a corresponding pin (Samtec TS-101-G-A) to which the probe was soldered. This was affixed to a CNC milled aluminium plate for mounting to the scan unit (see section 4.2.6). Connection to the probe was made via a shielded cable (the details of which can be found in section 4.2.1), taped down in a 'U' shape to the aluminium plate for strain relief and to prevent the cable from applying force to the probe holder. The aluminium plate was connected to ground, and when installed onto the scan unit this resulted in a grounded enclosure surrounding the probe holder for additional EM shielding. In order to prevent current leakage from the probe to the grounded plate, a sapphire pad (Thorlabs P25SK2) was used as a high resistance standoff.

4.2.6 Low-Cost 3-Axis Inertial Nanopositioner

Ultra-high resolution positioning of the probe relative to the sample surface is essential for most measurement modes of an SPM system; however, a vast dynamic range in the positioning capability is also necessary for practical operation. The probe and sample must be separated by a macroscopic distance (at least a few mm) prior to the start of a measurement, for sample/probe exchange. The positioning system must therefore be able to traverse this distance, while maintaining the ability to position the probe with sub-nm resolution in 3D space during the measurement. This

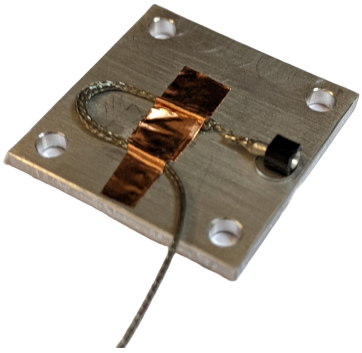


Fig. 4.8: A photograph of the STMBJ probe holder

represents a dynamic range (the ratio between the smallest step size and the largest motion range needed) on the order of 1×10^8. Typically, this obstacle is surmounted using two separate motion regimes, denoted 'coarse' and 'fine' respectively. The coarse motion is capable of long range traversal with lower resolution, while the fine motion has more limited range but higher resolution. The transition between these two domains is handled automatically by a 'coarse approach' algorithm, outlined in section 4.3.2.

Often, SPM systems use a separate actuator (for example a motor-controlled micrometer screw) for coarse motion, and coarse positioning is usually limited to the Z axis. An alternative approach is to combine the coarse and fine positioners into a 'friction-inertia' actuator, which interfaces a piezo actuator with a sliding linear stage via a friction coupling. The friction-inertia positioner leverages static and dynamic friction to enable two modes of operation:

The coarse motion mode is characterised by a two-phase 'stick-slip' stepping motion. In the 'stick' phase, the piezo actuator performs a slow extension, moving the stage forwards via the friction coupling. This is followed by the 'slip' phase, where the piezo actuator retracts rapidly enough to break static friction, so that the stage does not move as far in the slip phase as it did in the stick phase. The end result is a net forwards motion of the stage. A stick-slip cycle constitutes a single coarse step, and repeating this cycle allows the stage to be moved an arbitrary distance, limited only by the maximum travel of the linear stage. Reversing the direction of the piezo displacement during the slip and stick phases will reverse the direction of the stage. The relationship between piezo drive signal and stage displacement over several coarse steps is depicted in figure 4.9.

In the fine motion mode, the actuator is operated entirely within the static friction 'stick' mode, such that the motion of the piezo is directly coupled to the stage.

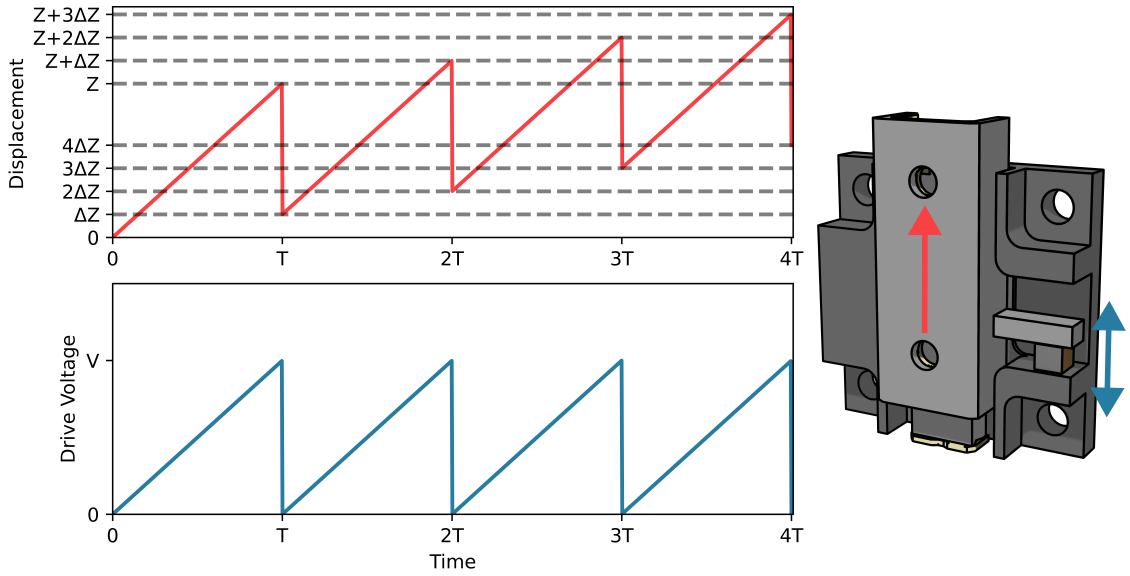


Fig. 4.9: An illustration of the piezo drive voltage and stage displacement over several coarse steps of a friction-inertia nanopositioner

In 2022 Liao et al [104] introduced an open source design for a low cost friction-inertia piezoelectric nanopositioner, capable of both long range stepwise movement for coarse positioning and the fine sub-nm movement required for SPM measurements. A bespoke 3-axis nanopositioner based on an adaptation of this design was created for use in the STMBJ. A number of modifications were made to the design to reduce cost and increase ease of manufacture. For example, in the simplified design presented here all three axes use identical parts, while in the original positioner the Z axis uses a different design from the X and Y axes. In addition, all parts were redesigned to only require machining on one face, eliminating additional repositioning operations which would otherwise be required when machining the parts on a 3-axis CNC mill: this can be a considerable factor in the manufacturability of a part, as care must be taken to ensure that the workpiece is re-aligned properly after repositioning between each machining step.

The actuator consists of a stationary frame which holds a precision linear slide (IKO Nippon Thompson BSP1025SL), and a moving stage mounted to the slide table. The frame features a mounting surface for a piezo stack actuator (Thorlabs PA4DGW) which is coupled to the slide by a magnet affixed to the piezo using cyanoacrylate adhesive. Figure 4.10 shows the individual components and assembly process for a single axis. Three stages were manufactured and stacked orthogonally to produce a 3-axis actuator, shown in figure 4.11.

The total parts cost of the 3-axis actuator was £150.

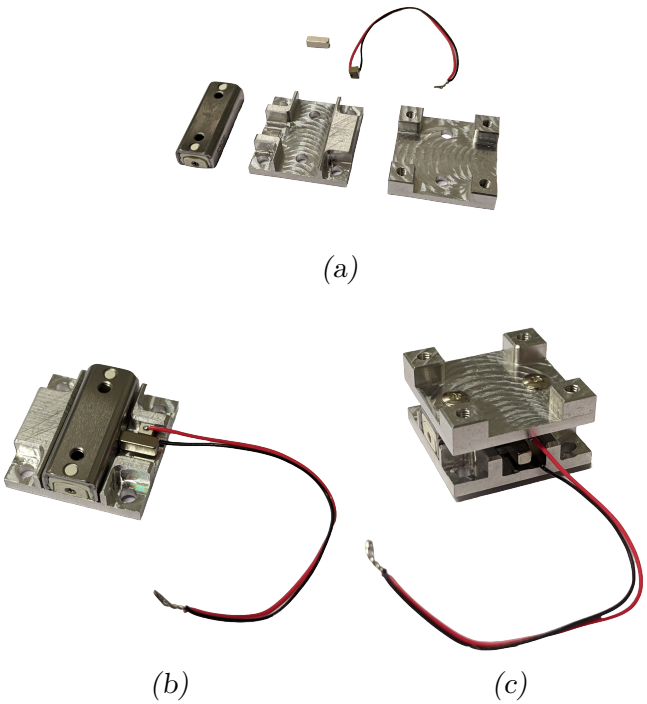


Fig. 4.10: **(a)** The individual components of one stage of the nanopositioner, before assembly. **(b)** The partially assembled stage, with piezo stack and magnetic coupling installed **(c)** The fully assembled stage

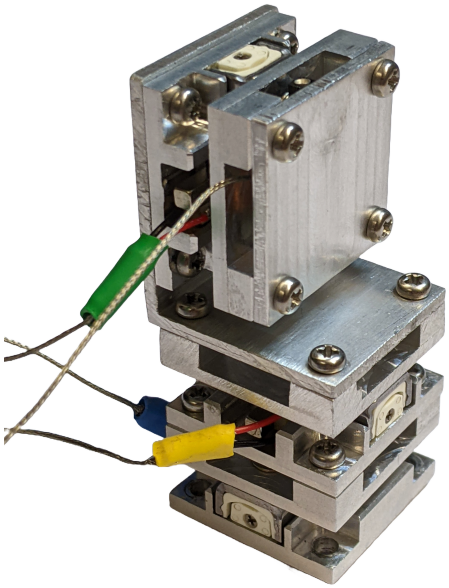


Fig. 4.11: The assembled 3 axis nanopositioner prior to installation in the STMBJ head

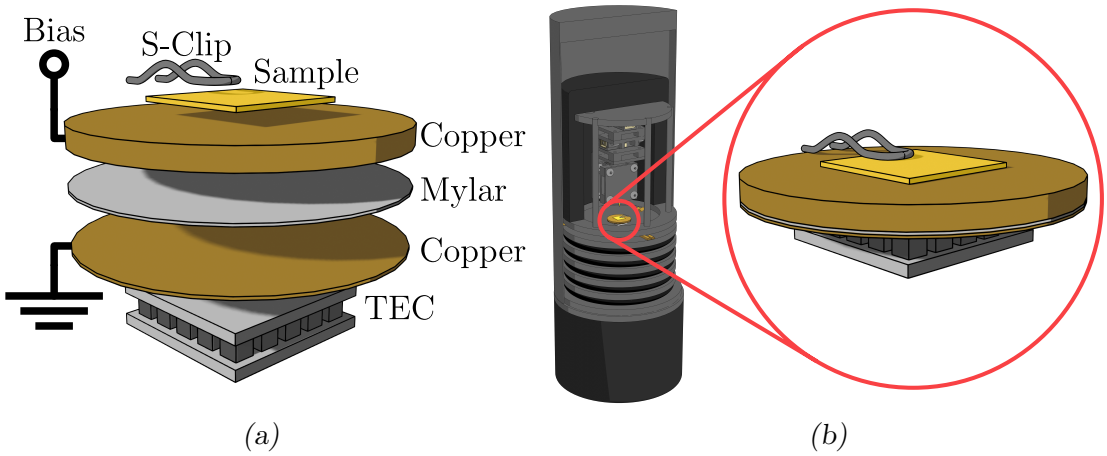


Fig. 4.12: Illustrations of the temperature-controlled sample stage assembly built for the STMBJ: **(a)** exploded view **(b)** assembled view

4.2.7 Variable-Temperature Sample Stage

The STMBJ was equipped with a custom sample stage with an integrated thermoelectric cooler (TEC) module (CUI CP1881-222) to allow the temperature of the sample to be varied. This functionality is essential for measurements of single-molecule thermopower [73] which may be performed using this instrument in future.

The construction of the sample stage is illustrated in figure 4.12. The TEC was separated from the sample plate by a grounded copper shield followed by a thin Mylar layer (chosen for its high resistance to minimise leakage currents). All parts of the stack were affixed with cyanoacrylate adhesive. The sample plate itself consisted of a copper disc 20mm in diameter, to which a shielded bias cable was directly soldered. An S-clip made from tinned copper-plated steel, soldered to the surface of the copper plate, held samples in place and provided an electrical connection to the sample surface. The sample stage can be seen in the photograph in figure 4.4.

The TEC was powered by a benchtop variable power supply (RS PRO RS-3005P), so that the temperature of the sample could be adjusted by changing the output voltage of the power supply. Details of the calibration of the stage are provided in section 4.5.2.

4.2.8 Electronics

The electronic control and data acquisition system for the STMBJ is comprised of a number of off-the-shelf subsystems, configured according to the block system diagram presented in figure 4.13.

DAQ

Signal generation and measurement is facilitated by a National Instruments USB-6363 Data Acquisition System (DAQ). The DAQ features 32 analogue input channels with sample rates up to 2 MHz and 4 analogue output channels with a maximum update rate of 2.86 MHz. All input and output channels are capable of sampling up to 16-bit resolution. Digitized input and output signals are transferred via USB between the control PC and FIFO buffers onboard the DAQ. While this arrangement lacks the fast real-time capabilities of more expensive dedicated SPM controllers, the buffered I/O stream and flexible onboard clock-routing allows the generation of predefined waveforms and synchronous measurement of signals at the full speed of the DAQ converters without USB communication overheads. This is sufficient for STMBJ measurements, which typically do not require the same high speed real-time feedback necessary for most imaging-based SPM modes.

Piezo Driver

The piezo stack actuators (Thorlabs PA4DGW) have a maximum displacement of 2.3 μm over a drive voltage range of $0-150$ V, corresponding to a sensitivity of 15.3 nm/V. Since a piezo actuator is a capacitive element, current is required to change the drive voltage (and therefore position):

Recall that the charge stored in a capacitance C is given by $Q = CV$ where V is the potential difference across the capacitor, and that the definition of current is $I = \frac{\mathrm{d}Q}{\mathrm{d}t}$. Therefore the current needed to drive a piezo element with capacitance C with a time-varying drive voltage $\frac{\mathrm{d}V}{\mathrm{d}T}$ is given by equation 4.1.

$$I = C\frac{\mathrm{d}V}{\mathrm{d}T} \tag{4.1}$$

The capacitance of the PA4DGW piezo stack used in the STMBJ is 50 nF. In the 'slip' phase of a coarse positioning step (see section 4.2.6), the piezo is driven with a linear ramp of 45 V and frequency 1 kHz. From equation 4.1 this requires a drive current of 2.3 mA (and therefore a peak power of 100mW).

The analogue signals generated by the DAQ are not sufficiently high voltage to drive the piezo elements without amplification. A Thorlabs MDT693B 3-channel piezo driver was used to amplify the drive signals with a gain of 15 and a maximum output current of 60mA [105]. The RMS output noise of the MDT693B is 1.5 mV, less than 1 LSB at 16-bit. Hence the resolution of the DAQ drive signal is preserved.

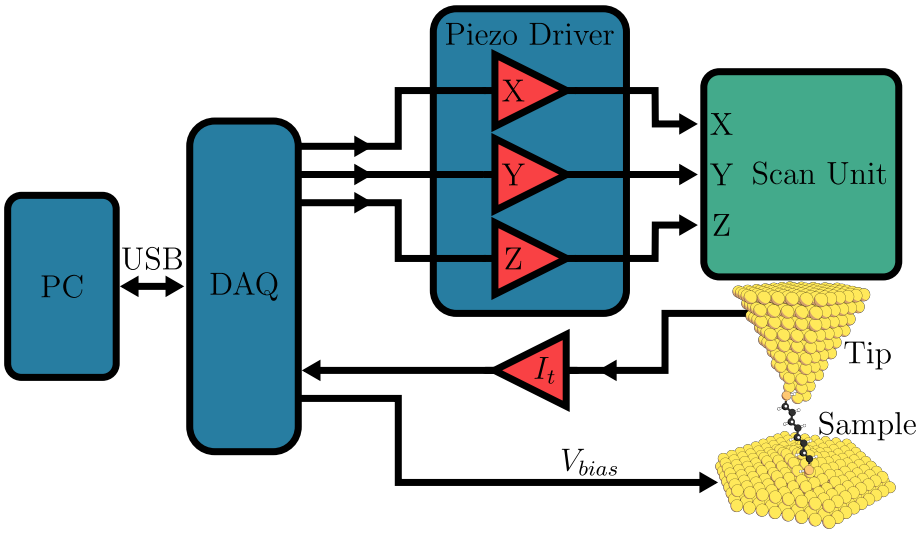


Fig. 4.13: A block diagram illustrating the control and data acquisition system of the STMBJ.

Transimpedance Amplifier

The STMBJ used a Femto DLPCA-200 low-noise current amplifier connected to the probe by a short length of shielded cable to amplify the junction current. The DLPCA-200 offers a selectable gain from 1×10^3 to 1×10^{11} nAV^{-1}, and an input bias current of 1 pA.

The amplifier was configured in the 'low-noise' mode with the input filter disabled ('full bandwidth' mode). The bandwidth of the amplifier depends on the gain setting used: at 1×10^8 nAV^{-1} (the setting used in most of the experiments presented here) the bandwidth is 7 kHz in low-noise mode [106]. The gain setting was selected to optimise the dynamic range of the amplifier for measurements of single-molecule conductance at the chosen bias voltage (a bias of 400 mV was used throughout this work).

4.3 Control Software

Custom software, comprised of a simple graphical user interface (GUI) and a corresponding backend written in Python, was developed to operate the STMBJ. The software was designed to be readily extensible; additional modes of operation can be added with minimal effort, and the software can be adapted to interface with alternative hardware (for example different SPM controllers) without major changes to the rest of the backend due to the modular design of the software.

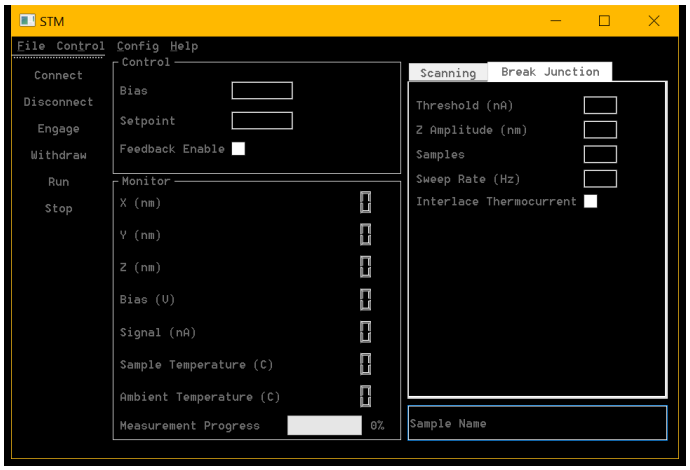


Fig. 4.14: The main window of the STMBJ control software UI

4.3.1 User Interface

The graphical user interface (GUI) was implemented using the QT 5 framework with PyQt5 Python bindings.

Main Window

The main window, pictured in figure 4.14, is split into sections which can be dynamically rearranged during use:

1. The drop-down toolbar at the top of the window provides access to the other windows of the UI
2. The quick-action bar on the left side of the window contains reconfigurable shortcuts to launch predefined automated behaviours (for example the coarse approach or measurement routines)
3. The control panel allows adjustment of general measurement parameters
4. The monitor panel shows measurement parameters (such as drive signals and percentage completion of the measurement routine) in real time
5. The experiment panel on the right hand side of the window contains the available measurement modes and presets, organised in a tab structure. An arbitrary number of tabs can be added to contain user-defined routines.

In addition to the main window, the UI has a number of dialogue windows to provide additional functionality. Most of these additional windows contain the more rarely used configuration settings and will not be discussed in full here. The most frequently used dialog window is the coarse positioning window (figure 4.15), which allows the user to control the 3-axis positioner in real-time. This may be used

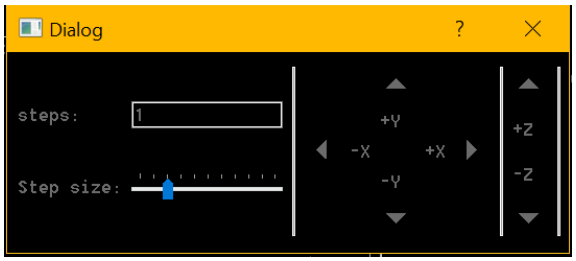


Fig. 4.15: The coarse positioning window of the STMBJ UI, from which the 3D nanopositioner can be directly controlled in real time by the user

alongside the live camera view from a USB microscope to move the probe to the desired region of the sample and bring the apex close to the surface prior to engaging the coarse approach routine (see section 4.4).

4.3.2 Measurement Routines

The control software features a number of automated control/measurement routines that can be engaged by the user to perform different procedures and experiments. Due to the extensible design of the software, new measurement routines can easily be defined to extend the capabilities of the instrument in the future. The main control and measurement routines included in the software are detailed in this section.

Feedback Loop

Feedback loops are control loops which compare a selected property of a dynamical system to some target setpoint, and modify the parameters of the system to maintain the selected property close to the setpoint [107]. In SPM, feedback loops are used to maintain a constant probe-sample separation by holding some distance-dependent measurement parameter constant. In STM this parameter is the tunnelling current, which follows an exponential dependence on the junction separation. While STMBJ measurements typically operate open-loop, without feedback engaged, feedback is still required to maintain the junction separation before and after measurements in order to protect the tip and sample. The STMBJ control software engages the feedback loop automatically whenever the tip position is not actively being controlled by an open-loop measurement mode.

The most common form of feedback loop used in SPM is a proportional-integral (PI) controller, where the control signal Z routed to the piezo at time t is given by:

$$Z(t) = K_p E(t) + K_i \int E(t)dt \tag{4.2}$$

where K_p is the proportional gain, K_i is the integral gain and $E(t)$ is the error (the difference between the tunnelling current and the setpoint at time t) [108].

In the feedback loop implemented for the STMBJ, the error was defined as the difference between the log of the setpoint and the log of the measured signal, in order to linearise the response of the feedback loop to displacement.

The proportional and integral terms were tuned iteratively to obtain stable performance: if the gains are too small, the signal will deviate too far from the setpoint and respond slowly to disturbances. If the gains are too high, the signal may overshoot the setpoint and/or exhibit unstable 'ringing' behaviour.

Coarse Approach Routine

SPM systems operate in two positioning regimes: the 'coarse' motion scheme provides the long range motion required for tip/sample exchange, while a high resolution 'fine' scheme with limited range is used for nanoscale scanning during the measurement. The coarse approach routine controls the transition between these positioning regimes at the beginning of a measurement.

The coarse approach algorithm implemented for this instrument is illustrated by the flowchart in figure 4.16. In summary, the algorithm is as follows:

First, the system measures the junction current I, and compares it to a user defined setpoint I_s. If the current is less than the setpoint then the system performs a fine extension of the Z piezo. If the Z limit (the maximum fine extension, by default set to half the fine travel range to ensure enough dynamic range for the feedback algorithm to engage when the setpoint is reached) is exceeded, the system performs a coarse step. Note that in order to avoid accidental tip crashes, the coarse step size must not exceed the Z limit. Once the current setpoint is exceeded, the system holds the probe stationary and measures the average current over 100ms (to avoid false triggering caused by transient increases in current). If the average current exceeds the setpoint then the approach is considered complete and the feedback routine is engaged to control the probe-sample separation.

G-Z Break Junction Routine

The most basic form of STMBJ experiment is the G-Z measurement, in which a metallic contact is formed between the tip and surface, before withdrawing the tip until the junction breaks while measuring the junction conductance.

The G-Z measurement routine defined in this work was as follows: at the start of the measurement, the feedback system is engaged and the tip is at some small controlled distance from the sample. The feedback loop is disengaged and the tip is advanced towards the sample using fine positioning until the junction conductance exceeds a user-defined setpoint corresponding to the formation of a metallic junction. At this point the tip is advanced a further 0.2 nm to ensure a robust contact, before

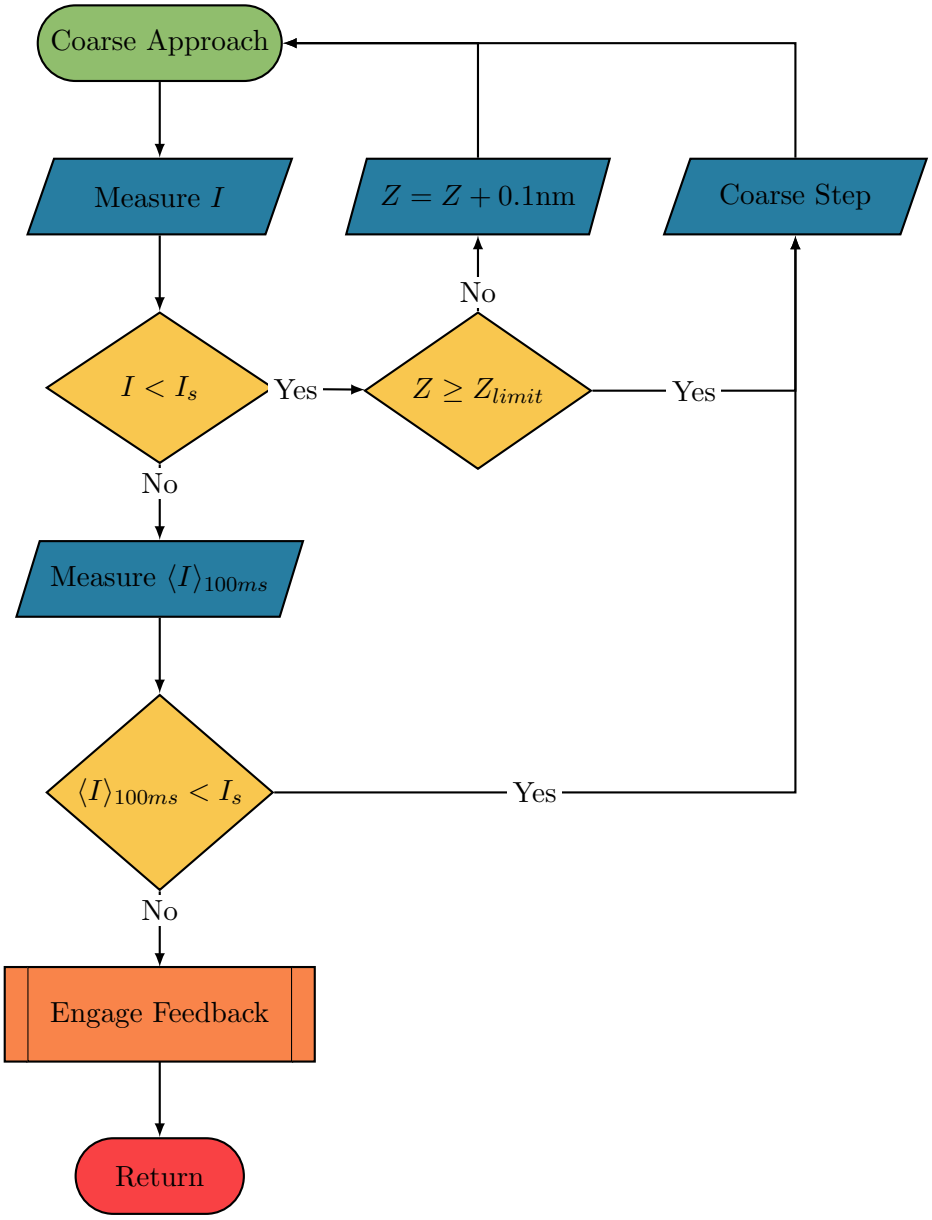


Fig. 4.16: A flowchart illustrating the coarse approach routine used in the STMBJ

withdrawing the probe a predefined distance at a constant rate. The output clock of the DAQ is routed to the input clock source during the withdrawal, such that the motion of the tip is synchronised with the conductance sampling. After the withdrawal is completed, feedback is re-engaged until a stable tunnelling current is attained, at which point the measurement repeats automatically.

4.4 Operating Procedures

The procedures used to operate the STMBJ within this work are detailed in this section.

4.4.1 Probe Preparation

Probe quality is of vital importance for STMBJ measurements. Cleanliness of the apex is especially critical, as contamination can alter the tunnelling decay factor [68] and even affect the coupling of molecules to the tip when forming molecular junctions.

Probes for the STM break junction were prepared as follows: a 1 cm length of Au wire (Goodfellow, 0.25 mm diameter, 99.99 + % purity) was soldered to a Au-plated pin (Samtec TS-101-G-A, plastic housing removed prior to soldering). The entire probe and pin were then sonicated in THF for 10 minutes to remove any flux or other contamination from the soldering process. The assembled probes were then plasma cleaned in air for 10 minutes and stored in a vacuum desiccator until needed. Immediately prior to installation in the STM each probe was rinsed in IPA and dried under nitrogen flow. The probe was installed in the STM by inserting the Au plated pin into the matching socket on the nanopositioner z-stage (see section 4.2.6). After installation the tip of the wire was cut in a pulling motion with a clean pair of wire cutters prior to each set of measurements to provide a sharp, uncontaminated tip apex. A careful coarse approach was then essential to maintain the quality of the tip: for details of the coarse approach procedure see section 4.4.2.

4.4.2 Coarse Approach Procedure

The coarse approach procedure is responsible for bringing the tip and sample into close enough proximity for tunnelling to take place, at which point the STM feedback loop can precisely maintain the junction separation. To achieve this, the tip must move from a position several mm away from the sample, to a few nm, without contacting the sample ('crashing'). A proper coarse approach routine is essential for ensuring the condition of the tip apex and sample at the measurement site: this phase of the measurement process carries the greatest risk of damage to the tip and

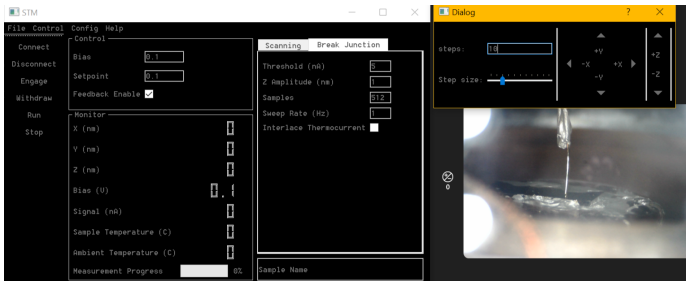


Fig. 4.17: A screenshot showing the controls and live view of the junction available to the operator during the manual phase of the coarse approach procedure

sample due to the relatively large and fast translations involved.

The first step of the coarse approach procedure is manually controlled: a USB optical microscope is used to observe the tip-sample separation, and the Z axis of the piezo-inertia actuator is operated in the coarse stepping mode to slowly move the tip closer to the sample. If a reflective substrate (such as Au) is used, the reflection of the tip on the sample surface can be used as a guide. Once the tip is within approximately 100-250 μm of the sample, the coarse motion is stopped and the automated phase of the approach is initialised.

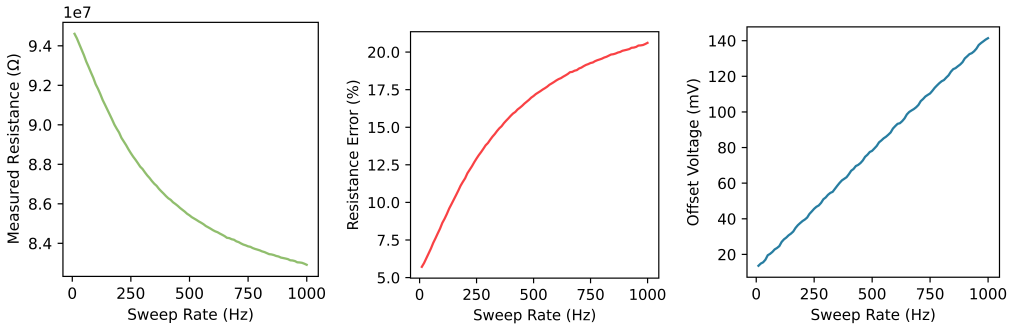
The automated approach phase takes a user-defined current threshold as a parameter to determine a successful coarse approach. This is set in software at the start of the measurement and must be selected based on the details of the sample, as well as the current amplifier gain and bias voltage. The software-controlled approach loop is described in detail in section 4.3.2. In practice it is sufficient to understand that the software executes alternate coarse and fine extensions of the tip towards the sample until the measured tip-sample current exceeds the threshold. At this point the STM feedback loop (described in section 4.3.2) engages automatically and maintains this setpoint until the user triggers an experiment routine.

4.5 Calibration and Testing

4.5.1 Electronics Calibration

Gain and Offset

In order to calibrate the total gain and offset of the measurement system, the STM tip was replaced with a 100 MΩ ±5% resistor wired directly between the sample plate and probe holder to simulate a junction with a known conductance of $log(\frac{G}{G_0}) = -3.9$. The conductance was measured over a range of bias values by performing an I-V sweep. The error and bias offset voltage (found from the zero-crossing point of the I-V sweep) were plotted as a function of sweep rate (figure 4.18). The error and offset were both found to increase with sweep rate (perhaps due to the settling



(a) *(b)* *(c)*

Fig. 4.18: **(a)** Measured resistance of a 100 MΩ resistor wired directly across the STM junction as a function of bias sweep rate. **(b)** Percentage error in measured resistance as a function of bias sweep rate. Note that the resistor had a specified tolerance of 5%. **(c)** Zero-offset in bias sweeps as a function of bias sweep rate.

time of the DAQ ADC, stray capacitance or the bandwidth of the current amplifier), however near DC the error and offset were found to be below 5% (the tolerance of the resistor itself) and 20 mV respectively. If bias modulation is required for future experiments, a more detailed investigation into the error observed at higher sweep rates should be carried out.

4.5.2 Sample Temperature Calibration

The variable-temperature sample stage (detailed in section 4.2.7) had to be calibrated to ensure accurate control of the sample temperature.

The relationship between the temperature of the sample stage and the voltage applied to the TEC element was determined as follows: a K-type thermocouple was affixed to the sample stage, and four different voltages were applied to the TEC; the corresponding temperature change for each voltage (after a 180 second thermalisation period) was recorded. This was repeated on three different days, noting the temperature in the STM enclosure on each day in order to account for variations in Peltier efficiency associated with changes in ambient temperature. Figure 4.19 shows the results of these calibration measurements. The uncertainty in the temperature of the stage associated with typical day-to-day variations in room temperature was estimated from these measurements to be approximately ±1°C. This could be improved in future by using a closed-loop TEC controller such as the TED200C (Thorlabs, Inc.) which specifies an accuracy of ±0.1°C.

4.5.3 Conductance Quantization

The conductance quantization predicted by the Landauer equation (described in section 2.1) provides an ideal benchmark for validating the conductance values measured by the STMBJ in I-Z mode. For I-Z measurements using a Au tip and un-

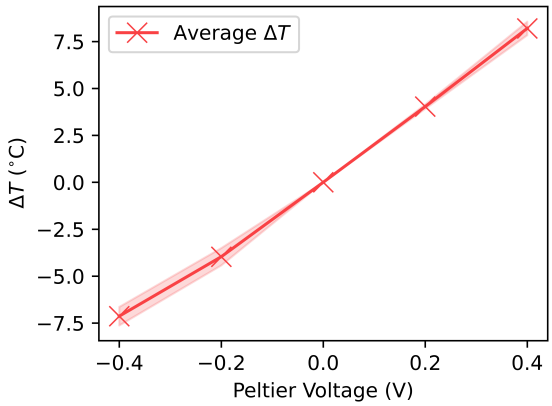


Fig. 4.19: Calibration of the variable-temperature sample stage.

coated Au substrate, a Au junction is formed, which is gradually stretched during the retraction phase until the contact is only a few atoms wide at its thinnest point. In this region, each atom can be considered as a single perfect conductance channel with conductance G_0. As the junction is stretched further and the number of atoms (and therefore conductance channels) in the junction reduces one by one, discrete conductance steps at integer multiples of G_0 should be observed in the conductance trace until the narrowest part of junction contains only one Au atom, corresponding to a conductance of 1 G_0. After this threshold is reached the contact will break, and the conductance trace will exhibit an exponential decay consistent with the through-space tunnelling regime. The 1D conductance histogram is therefore expected to exhibit peaks at integer values of G_0. Since G_0 is a physical constant, this measurement is independent of the specific details of the junction and can be used as a reference to confirm that the STMBJ correctly measures the value of conductance plateaus.

Figure 4.20 is a 1D conductance histogram for 1000 I-Z measurements performed with the STMBJ on a AU substrate using a Au tip. Clear peaks at 1 G_0, 2 G_0 and 3 G_0 are observed, corresponding to 1, 2 and 3 Au atoms in the tip-sample contact respectively. This serves as a validation of the conductance values measured by the STMBJ.

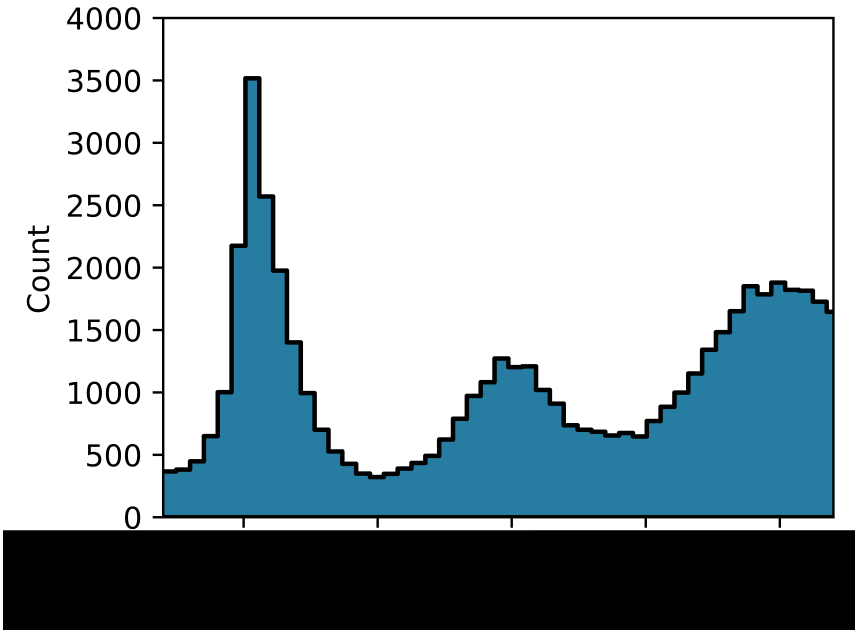


Fig. 4.20: Conductance histogram for 1000 I-Z measurements on a Au substrate, exhibiting peaks at integer multiples of G_0 consistent with the conductance quantisation predicted by the Landauer equation.

5. UNSUPERVISED CLUSTERING OF BREAK JUNCTION MEASUREMENTS

STMBJ measurements on single molecules are stochastic in nature: each measurement iteration entails the formation of a new molecular junction, the specific properties of which can have a significant effect on the measured signal [67]. For example, variations in the coupling between the molecule and the STM tip/substrate can change the conductance of the junction [66]. A typical STMBJ experiment will therefore consist of many (on the order of several thousand) individual measurements, from which a statistical description of the charge-transport properties of a particular molecular junction can be derived [109]. This chapter details the implementation and testing of an unsupervised feature extraction and clustering process for the analysis of these large datasets.

A number of different data analysis methodologies have been applied to the analysis of the large datasets produced by STMBJ conductance measurements. The oldest, and still most commonly employed, approach is to collapse the entire dataset into a single conductance-distance histogram. The most probable conductance and breaking length of the molecular junction can usually be obtained via this method. Information about less frequent junction configurations with different average conductances may be obtained also, provided that they occur frequently enough to provide a sufficient signal-to-noise ratio for detection [72, 109].

The aforementioned approach is simple and intuitive, but has a number of significant limitations: less frequently occurring conductance plateaus and subtleties such as variations in the plateau shape are often not detected. This additional information, inaccessible to all-data histogram analysis, can provide insights into the evolution and transport properties of the molecular junction [110]. In severe cases an 'all-data' histogram analysis can fail to reliably distinguish any of the molecular plateaus present in the data from the noise floor, obfuscating even the most probable molecular conductance value [31].

In an effort to overcome the limitations of 'all-data' analysis techniques, some form of data selection is often implemented to separate specific event classes from the background by application of certain selection/filtering criteria [68, 111–113]. These so-called supervised selection methods, which filter the data based on presupposed criteria specific to the experiment, can improve signal-to-noise ratio and potentially

reveal less abundant event classes within the dataset [68]. However, the efficacy of the data selection is dependent on the selection criteria used: assumptions must be made about the data in order to choose sensible selection criteria (for example, an assumption about the shape of conductance plateaus is often employed to eliminate seemingly noisy or featureless traces [68]). This limits the ability of supervised data selection to detect unexpected event classes. Additionally, the choice of selection criteria introduces a potential source of bias in the interpretation of STMBJ data [67] by discarding signal classes not anticipated by the researcher and emphasising traces which match the selection criteria.

More recently, progress has been made in applying unsupervised and semi-supervised machine-learning based techniques to the task of separating the different signal classes present in STMBJ datasets [110, 111, 114, 115].

The unsupervised clustering procedure developed in this work for STMBJ G-Z data consists of four distinct steps:

1. Preprocessing

   Data scaling, alignment and normalisation

2. Feature Extraction

   Extraction of meaningful low-dimensional features from the raw data

3. Clustering

   Organisation of data into clusters with similar features

4. Interpretation

   Once clusters of similar traces have been obtained, they must be inspected manually in order to obtain a physical interpretation of the data based on specific domain knowledge

These steps are discussed in depth in sections 5.1 to 5.4. The optimisation and testing process and results are detailed in section 5.6.

5.1 Preprocessing

Prior to feature extraction, some simple preprocessing steps are applied to the raw data:

First, the raw ADC input readings are rescaled to the actual measured current by dividing each value by the gain of the current amplifier (see section 4.2.8). The raw Z DAC output is scaled to the actual piezo displacement by multiplying by a scale factor of 230 nmV^{-1}, derived from the sensitivity of the piezo stack element

used in the actuator (15.3 nm/V) and the voltage gain of the piezo amplifier (15 V/V).

After scaling the data channels to the correct units, an alignment step was applied. As described in section 4.3.2, *I-Z* traces are acquired by first extending the *Z* piezo until a certain current threshold representing a tip-sample contact (or 'crash') is reached and then withdrawing some pre-set distance while recording the current at fixed intervals. Due to the stochastic nature of the initial tip-substrate contact, the distance (*Z*-offset) that the tip must withdraw before the initial contact is broken and the *I-Z* trace moves into the tunnelling regime can vary significantly between traces. In order to compare different *I-Z* traces, the *Z*-offsets of the individual traces must be aligned to some common reference. This is accomplished by selecting a current threshold and aligning the *Z* origin of the individual *I-Z* traces to the last point at which the current crosses below the threshold. The length of each aligned trace was truncated from 4096 to 392 samples to ensure that all the traces contained the same number of samples.

After alignment, the current (*I-Z*) traces were converted to conductance (*G-Z*) traces by dividing the current values by the applied bias voltage. The conductance values were then rescaled by dividing by the quantum of conductance G_0 and a log scale was applied to obtain $log(\frac{G}{G_0})$-Z traces.

Finally, the traces were normalised such that all values ranged between 0 and 1 (a prerequisite for the feature extraction network described in section 5.2) as follows: upper and lower boundaries of $log(\frac{G}{G_0})$, G_{max} and G_{min}, corresponding to the dynamic range of the amplifier were selected and the traces were cropped within this range. The samples of the cropped traces were then normalised by applying the following equation:

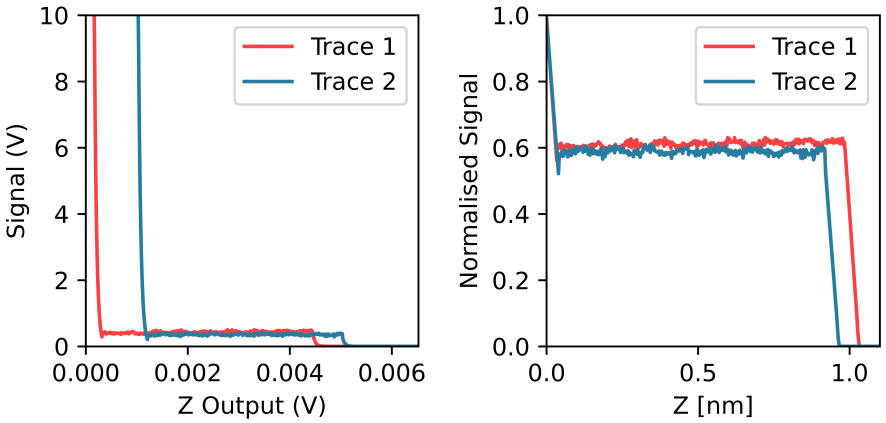
$$g_i' = \left| \frac{g_i - |G_{min}|}{|G_{min}| - |G_{max}|} \right| \tag{5.1}$$

where g_i and g_i', $i \in \{\mathbb{Z}, i \leq 392\}$, are the elements of the trace $\mathbf{g}$ before and after normalisation, respectively.

Figure 5.1a and figure 5.1b show a before-and-after example of the alignment and preprocessing applied to two synthetic I-Z traces.

5.2 Feature Extraction

Each STMBJ measurement trace, after preprocessing (see section 5.1), consists of 392 individual samples. For the purpose of identifying clusters of similar measurements, these traces could be considered vectors in a 392-dimensional space. However, this direct approach suffers from the so-called 'curse of dimensionality': identifying



(a) (b)

Fig. 5.1: **(a)** Two synthetic I-Z traces with varying z-offsets prior to preprocessing **(b)** The same traces shown in (a), after preprocessing

clusters of similar vectors in a high-dimensional space can become increasingly challenging as the dimensionality of the space increases, and clustering algorithms tend to under-perform on high dimensional data. This is because all points in the space tend to appear equidistant from each other as the dimensionality increases [116,117]. This somewhat counter intuitive phenomenon can be demonstrated through the following example:

For a k dimensional Euclidean space $\mathbb{R}^k$, Latin hypercube sampling was used to generate a set χ of n random points: $\chi = \{x_i \in \mathbb{R}^k | i \in \mathbb{Z} \wedge 1 \leq i \leq n\}$. The Euclidean distance d between all pairs of points in χ was then calculated in order to obtain the maximum, minimum and average pairwise distances, denoted d_{max}, d_{min} and $\overline{d}$ respectively. A simple measure of the 'contrast' C of the distance metric (i.e. the ability to identify differences in the distance between pairs of points) was defined.

$$C = \frac{d_{max} - d_{min}}{\overline{d}} \tag{5.2}$$

A smaller value of C means that the total range of pairwise distances is compressed around the average distance. A plot of C against k for 1000 points, presented in figure 5.2a, reveals that as the dimensionality k of the space increases the distances between pairs of points become increasingly similar. Figure 5.2b shows $d_{max}/\overline{d}$ and $d_{min}/\overline{d}$ separately to illustrate the concentration of distances around the mean. [118]

It has been suggested that the use of fractional distance metrics, in place of the most commonly used L2 (Euclidean) metric, can ameliorate the contraction of distances in high dimensional space [116, 119]. However, a recent study found no practical improvement in clustering performance in tests of various alternative

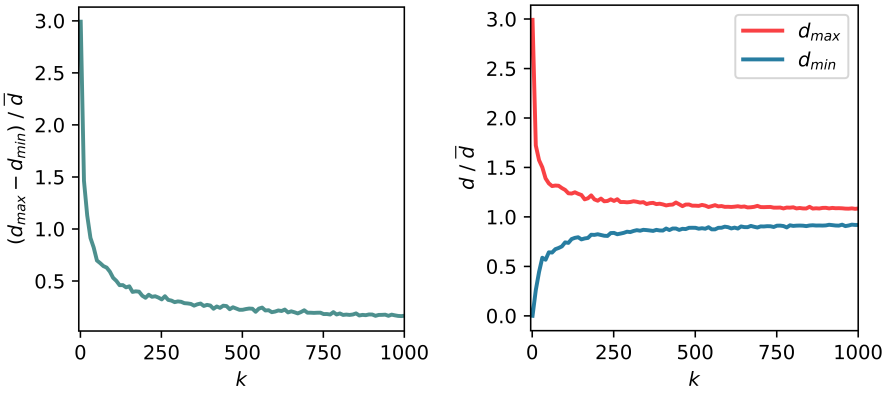


Fig. 5.2: Plots showing the contraction of distances in Euclidean spaces with high dimensionality k. d_{max}, d_{min} and $\overline{d}$ are the maximum, minimum and mean pairwise Euclidean distances over 1000 points randomly sampled by Latin hypercube sampling from a k dimensional space.

distance metrics [117].

Feature extraction techniques, which create a projection from high dimensional raw data to a lower dimensional 'feature space' (also known as the 'latent space') can be used to avoid these difficulties by applying the clustering algorithm to the lower dimensional features rather than to the data itself. Hence, clustering algorithms are typically preceded by a feature extraction step in data analysis workflows [120].

Myriad approaches to feature extraction exist, each with different advantages and limitations [111, 121]. An ideal feature extraction process for STMBJ should identify a mapping with the following properties:

1. Continuity: the mapping between the data and the extracted features should be continuous, such that similar traces are close together in feature space.

2. Disentanglement: each dimension of the feature space should correspond to a single salient characteristic of the data. This encourages the extraction of physically meaningful features, facilitating straightforward interpretation of downstream clustering results.

3. Robustness: the extracted features should be independent of random corruptions (e.g. the addition of white noise) of the input data. This is because the extracted features should correspond to physically meaningful properties, which remain unchanged in the presence of non-meaningful corruptions

[122]

These properties, or 'meta-priors' [123], are identified based on assumptions about the underlying distributions of the physical features of experimental STMBJ G-Z traces.

In the work presented here, autoencoder-based feature extraction techniques were explored, including strategies to enforce the meta-priors identified above. A review of the relevant literature has revealed no other studies considering the disentanglement of learned representations of STMBJ data.

5.2.1 Autoencoders

The autoencoder (AE) is a type of unsupervised machine-learning model which encodes an input into a lower dimensional representation and then attempts to accurately reconstruct the original input from the encoded representation.

The original [124,125] (and perhaps still the most popular) implementation of the autoencoder is the feedforward autoencoder, which consists of a multilayer neural network split into an 'encoder' section and a 'decoder' section separated by a 'latent layer' which is small compared to the dimensionality of the input data [126]. The architecture of a typical autoencoder is illustrated in figure 5.3.

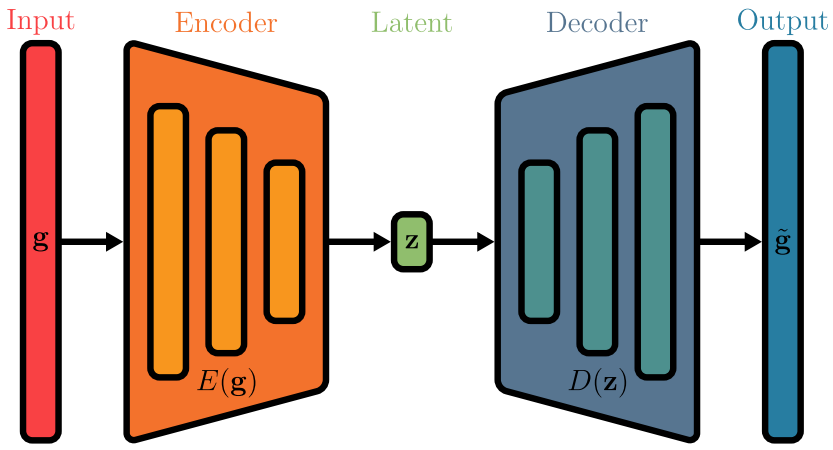


Fig. 5.3: Illustration of the architecture of a typical autoencoder network

The autoencoder network is trained to reconstruct the input of the network at the output layer: since the architecture contains an 'information bottleneck' at the latent layer, the encoder layers are encouraged to encode meaningful features of the data in the latent layer, to enable accurate reconstruction by the decoder layers. Therefore the encoder network can be used as a feature extractor to convert input data to a lower dimensional representation containing its most important features. The decoder network is not used for feature extraction, but it plays an essential role in the training of the autoencoder:

The feedforward networks in the encoder and decoder portions of the autoencoder can be considered universal function approximators [127]. The encoder approximates an arbitrary function E which takes input data $\mathbf{g}$ and maps it to $\mathbf{z}$ (the latent representation).

$$E(\mathbf{g}) = \mathbf{z} \tag{5.3}$$

Likewise, the decoder is an arbitrary function D which maps $\mathbf{z}$ to an output $\tilde{\mathbf{g}}$.

$$D(\mathbf{z}) = \tilde{\mathbf{g}} \tag{5.4}$$

Since the goal of the autoencoder is to produce a reconstruction of the input data, E and D should be found such that $\mathbf{g} \approx \tilde{\mathbf{g}}$ [126]. This is the purpose of the 'training' process.

In order to qualitatively assess the similarity between the input and output, it is necessary to define a reconstruction error function $\Delta(\mathbf{g}, \tilde{\mathbf{g}})$, which maps the tensors $\mathbf{g}$ and $\tilde{\mathbf{g}}$ to a scalar corresponding to the difference between them. Several error functions exist, with different applications. In this work the mean-squared-error (MSE) function, which calculates the average of the squared error across the entire input dataset, is used throughout. If χ represents a dataset of n observations $\mathbf{g}_i$, $i \in \{\mathbb{Z}, i \leq n\}$, the MSE function Δ_{MSE} can be written:

$$\Delta_{MSE}(\chi, \tilde{\chi}) = \frac{1}{N} \sum_{i=1}^{N} |\mathbf{g_i} - \tilde{\mathbf{g}}_\mathbf{i}|^2 \tag{5.5}$$

.

Neural networks are trained using stochastic gradient-based optimisation algorithms [128] to adjust the weights $\mathbf{w}$ of the connections within the network such that the objective function $\mathcal{L}(\mathbf{w})$ (also known as the loss function or cost function) is minimised [129]. Note that in general it is not possible to determine if the algorithm has converged to a global or local minimum. In its most basic configuration, the objective function of an autoencoder can simply be the reconstruction error.

$$\mathcal{L}(\mathbf{w}) = \Delta_{MSE}(\chi, \tilde{\chi}) \tag{5.6}$$

Regularisation terms can also be added to the objective function to encourage the learned representation to have specific properties, beyond simply minimising the reconstruction error. The bottleneck in the network structure is itself a form of regularisation; if the latent layer had the same dimension of the input, the encoder and decoder could simply learn the identity function to achieve perfect reconstruction, without extracting any meaningful features [126]. Regularisation of some form is often necessary to avoid overfitting, especially as the depth of the encoder and decoder networks increases. If the network has sufficient depth, the network could achieve optimal reconstruction for any arbitrary dataset with only a single dimension in the latent layer, by simply assigning an index to each training example and effectively 'memorising' the entire dataset. Clearly such a network would be completely unable

to generalise to new data not present in the training dataset, and would not result in the extraction of meaningful features.

Another drawback of an unregularised autoencoder is that the latent space is not encouraged to be smooth/continuous or disentangled, since this has no effect on reconstruction accuracy over the training dataset. This results in an inability to generalise to examples outside of the training dataset. The lack of continuity and disentanglement is also expected to lead to less interpretable results in downstream clustering.

Despite these limitations, standard autoencoders have been applied as feature extractors for STMBJ with success [114], although a tendency towards overfitting and poor performance has been noted in some cases, prompting some to seek alternative implementations such as combining autoencoders with additional pre-trained feature extractors based on transfer learning [115].

5.2.2 Variational Autoencoders

The variational autoencoder (VAE), introduced in 2013 by Kingma and Welling [130], is a class of variational Bayesian methods which uses an autoencoder-style architecture for variational inference. The VAE is often intuitively understood as a standard AE with added regularisation, although the mathematical motivations of the two models are entirely different, and in fact the resemblance is largely limited to the use of an encoder-decoder structure in training both models [131]. Nonetheless, the VAE can be used for feature extraction in much the same way as the standard AE (see 5.2.1), and offers a number of advantages in learning useful reduced-dimensionality representations [132]. In particular, VAEs are well suited to extracting disentangled features, and many of the best performing models in this regard are based on the VAE architecture [133–136].

The operating principle of the VAE is based on variational Bayesian inference. Consider the situation introduced in section 5.2.1, in which n individual STMBJ measurements, $\mathbf{g}_i$, comprise a dataset $\chi = \{\mathbf{g}_i\}$, $i \in \{\mathbb{Z}, i \leq n\}$. In Bayesian inference the data is considered to be fully described by a latent variable $\mathbf{z}$ which encodes all physical properties influencing the measured data, and is distributed according to a probability distribution function $p(\mathbf{z})$ which represents the random process from which the data emerges [130]. Of course, in a real-world measurement both $\mathbf{z}$ and $p(\mathbf{z})$ are unknown and can only be approximated *a posteriori* based on observations (χ). Bayesian inference provides the mathematical framework for this process [137]. The VAE provides a versatile, efficient and scalable approach to Bayesian inference [130]. The encoder E of the VAE maps the input data to the parameters of a distribution function, and the decoder then samples a latent variable $\mathbf{z}$ from this parametrised distribution and generates output data. Figure

5.4 illustrates this architecture.

The objective function $\mathcal{L}(\mathbf{w})$ of a VAE is the sum of the reconstruction error Δ (introduced in section 5.2.1) and the Kullback–Leibler (KL) divergence D_{KL}:

$$\mathcal{L}(\mathbf{w}) = \Delta + D_{KL} \tag{5.7}$$

.

The KL divergence quantifies the level of dissimilarity between the encoded distribution and an 'uninformative prior' distribution, which represents a weak assumption about the distribution of the generative factors and is typically selected to be a Gaussian with zero mean ($\mu = 0$) and unit standard deviation ($\sigma = 1$) [138]. The KL divergence term in the loss function encourages the mapping of a continuous latent space, from which feature vectors $\mathbf{z}$ can be extracted for use in clustering.

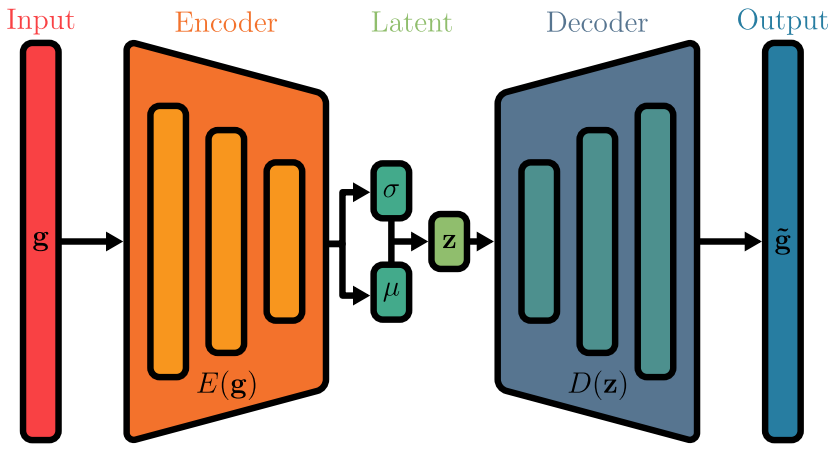


Fig. 5.4: Illustration of the architecture of a typical variational autoencoder network

5.2.3 Enforcing Meta-Priors Via Regularisation

Three assumptions, or 'meta-priors' [123], regarding the properties of meaningful features for STMBJ data are identified at the start of section 5.2. For completeness, the identified meta-priors are repeated here:

1. Continuity
2. Disentanglement
3. Robustness

A number of regularisation strategies to enforce these meta-priors during the training stage of the model have been proposed in the literature [123, 134, 135, 139, 140]. Regularisation refers to the process of making adjustments to the loss function

or training process of the model in order to encourage specific properties of the latent representation.

βVAE

The βVAE framework, introduced by Higgins et al. in 2017 [134], modifies the VAE objective function (equation 5.8) by introducing an additional hyperparameter β which can be tuned to enhance the disentanglement of the extracted features. The β factor adjusts the weight of the KL divergence term in the objective function.

$$\mathcal{L}(\mathbf{w}) = \Delta + \beta D_{KL} \tag{5.8}$$

When $\beta = 1$, the standard VAE objective function is recovered, and when $\beta = 0$ the objective function is equivalent to the standard autoencoder (equation 5.6). For $\beta > 1$, the objective function more strongly encourages minimisation of the KL divergence which increases the strength of the constraints on the learned latent representations, in turn reducing the effective capacity of the information bottleneck. This encourages the encoding of better disentangled features in the latent space [133].

Due to the constriction of the information bottleneck, a trade-off is expected between disentanglement and reconstruction accuracy as β increases [138]. A study of βVAE in data classification applications [141] showed that this trade-off may apply not just to reconstruction accuracy but also to downstream tasks such as clustering. Notably however, at small values of $0 < \beta < 1$ the classifier was observed to achieve higher classification accuracy than both the standard AE ($\beta = 0$) and VAE ($\beta = 1$) cases. This suggests that small non-zero β values could be used to provide regularisation and prevent over-fitting while approximating the autoencoder case.

The aforementioned compromise between β and reconstruction accuracy renders the βVAE vulnerable to an issue known as 'posterior collapse': this is a local optimum in the objective where the posterior approximates the uninformative prior regardless of the latent codes, such that the reconstructed output is unchanged with varying input data [142–144]. In this situation the extracted features become useless for downstream clustering.

One possible approach to circumvent the trade-off between disentanglement and reconstruction accuracy, as well as to prevent posterior collapse, is dynamic adjustment of the capacity of the information bottleneck during training [138, 144, 145]. This process is sometimes referred to as 'KL annealing', and typically involves adjusting the value of β during training.

Perhaps the simplest form of KL annealing is 'linear annealing': β increases linearly with each training epoch until some maximum value is reached, at which

point β remains constant for the remainder of the training process [145]. In 2019, Fu et al. [144] presented an improved 'cyclical annealing' schedule in which the β factor is continuously modulated throughout training.

The experiments with KL annealing presented in [144] and [145] were confined to $0 \leq \beta \leq 1$, where the maximum value of β during training corresponded to a standard VAE. No examples of KL annealing applied to an STMBJ feature extraction context could be found in the literature. Section 5.6.4 discusses the results of a comparison of disentanglement and clustering accuracy for a VAE trained with a cyclical annealing schedule compared to a standard VAE and AE, for synthetic STMBJ data.

5.3 Clustering

In machine learning, clustering refers to the task of partitioning the elements of a dataset into groups, or 'clusters', according to similarity [146]. While this task is easily understood from an intuitive perspective, a rigorous definition of what clustering means in a machine learning context is more challenging to formulate and remains the subject of some debate.

Typical criteria for clustering are:

1. Similarity between elements of the same cluster should be maximised
2. Similarity between elements of different clusters should be minimised

[147]

Selecting the correct algorithm for a given clustering task is of vital importance. A vast range of algorithms and variations are available, and a comprehensive review of the various techniques is beyond the scope of this work (surveys of existing clustering algorithms are available in the literature: [147–149]).

In this work a Dirichlet process Gaussian mixture model (DPGMM) clustering algorithm was utilised, which uses Bayesian inference with a nonparametric Dirichlet process prior to optimise the assignment of samples to clusters. One notable advantage of this method is that the number of clusters does not need to be selected manually, as the clustering algorithm can dynamically add, remove or merge clusters during the optimisation process. The algorithm used was sourced from the BNPy python package [150, 151].

5.4 Interpretation of Break Junction G-Z Traces

Once an STMBJ dataset has been partitioned by the clustering algorithm, a physical interpretation of the clusters must be obtained by manual inspection. Interpretation

of an STMBJ trace, or a 2D histogram comprising many similar traces, typically relies upon the identification of certain key features in the trace shape. These may include:

- The presence of a conductance plateau

  A plateau in the G-Z trace often indicates the formation of a molecular junction

- The average conductance of the plateau

  The average conductance at which a plateau occurs provides information about the most probable conductance of the molecular junction. This may depend on a number of factors, such as the binding configuration between the electrodes.

- The length of the plateau

  The length of the plateau indicates how far the junction may be stretched before breaking. This may encode information about the length and binding geometry of the molecule, as well as the mechanical stability of the junction.

- The slope of the plateau

  If the conductance of the molecular junction changes with the mechanical stretching caused by increasing electrode separation (Z) the plateau in the G-Z trace will possess a slope.

- Switching, steps or other sharp conductance transitions

  Sharp transitions in the conductance of the plateau, or the presence of multiple distinct plateaus, can be indicative of some dynamic process in the junction, such as reconfiguration of the binding geometry or supramolecular interactions [152].

- Noise

  The analysis of conductance noise in a G-Z trace can provide insight into the electrode-molecule coupling [93] and the stability of the junction.

There exist a vast number of physical factors which can influence the shape of the measured G-Z trace in an STMBJ measurement, and physical interpretation beyond the most probable conductance is non-trivial. This section presents an overview of some important factors which may affect the shape of STMBJ G-Z traces for alkanedithiol molecules between Au electrodes (as studied in this work).

5.4.1 Binding Sites on Electrodes

The anchor group of the molecule may bind to the electrodes in a number of different configurations, corresponding to different sites on the metallic electrode surface. For instance, the thiol anchor group of an alkanethiol molecule may bind to a Au(111) electrode in three different configurations [152]:

1. Atop

   Situated over a single Au atom.

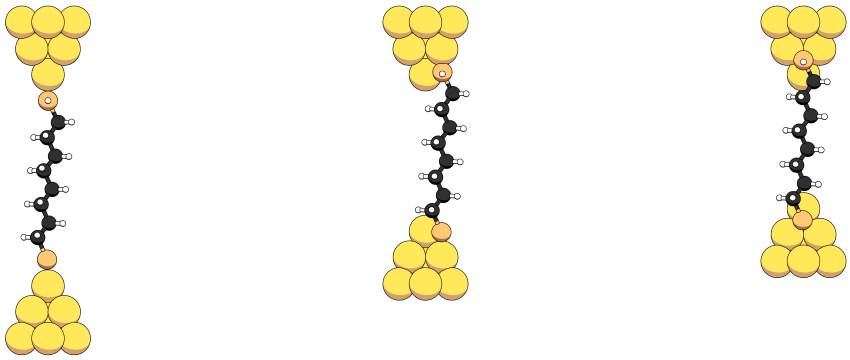
2. Bridge

   Situated between two Au atoms.

3. Hollow

   Situated between three Au atoms.

The binding configuration may affect the stability, breaking length and conductance of the junction [152, 153]. Figure 5.5 illustrates the different binding sites for an alkanedithiol between Au electrodes.



(a) *(b)* *(c)*

Fig. 5.5: Illustrations showing three possible binding configurations of the molecule to the Au electrodes: **a** Atop **(b)** Bridge **(c)** Hollow

5.4.2 Binding Strength

The strength of the coupling between the molecule and the metallic electrodes can depend on various factors beyond the geometry of the binding site (discussed in section 5.4.1). For example, thiolate molecules can form different types of bond with the gold atoms in the electrodes, with different binding strengths, depending on the loss or retention of the H atom in the SH anchor group upon bond formation [154]. If the H atom is lost, a covalent S-Au bond will be formed [155, 156] with a breaking

strength of around 2 nN [71]. Conversely, if the H atom is retained (a so-called 'protonated' state) the S-Au bond will be a weaker coordination bond [154] with a breaking strength of 0.6nN reported in the literature [71]. Figure 5.6 illustrates these two types of bond for alkanedithiol molecules between Au electrodes.

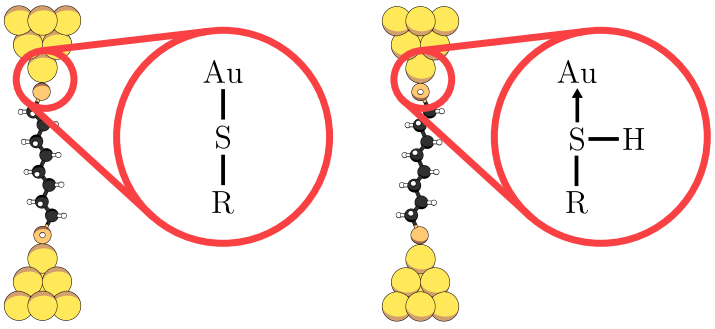


Fig. 5.6: Illustrations showing two different types of bonding between the Au electrode and the thiol/thiolate anchor group of an alkanedithiol: **(a)** Covalent bond (S-Au)**(b)** Coordination bond (SH-Au)

The binding strength can influence the shape of a G-Z trace in a number of ways. In the case of Au-alkanedithiol-Au junctions, the S-Au bond has been associated with higher conductance, lower conductance noise and a steeper gradient in the conductance plateau compared to the SH-Au coordination bond [155].

5.4.3 Deformation of Electrodes

As the Au-Molecule-Au junction is stretched during the retraction phase of the G-Z measurement, tensile force is exerted on the binding sites of the molecule and the Au electrodes. In the case of a covalent S-Au bond for a thiolate molecule, this force may be as high as 2 nN (the breaking force of a S-Au bond). This exceeds the breaking force of a Au-Au bond (1.6 nN) [71]; as such, considerable reconfiguration and deformation of the geometry of the Au electrodes may occur before the breaking of the junction. For example, Au atoms may be 'extruded' from the electrode, forming an atomic chain between the S-Au binding site and the bulk of the electrode (illustrated in figure 5.7). This may extend the measured breaking length of the molecular junction [152]. Reconfiguration of the electrode may also alter the binding configuration (see section 5.4.1) and therefore change the conductance of the junction, resulting in steps, spikes, or other sharp switching features in the conductance plateau [152].

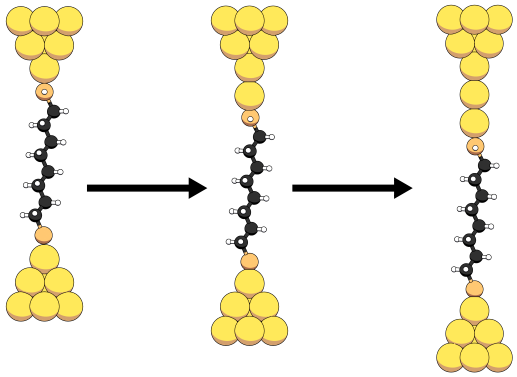


Fig. 5.7: An illustration showing the extrusion of gold atoms from an STMBJ tip during the retraction phase of a G-Z measurement.

5.5 Evaluation of Clustering Methods

5.5.1 Performance Metrics

In order to compare the performance of feature extraction and clustering algorithms, metrics must be defined by which key performance parameters can be quantitatively evaluated. In this work, metrics for unsupervised accuracy and disentanglement were selected based on a review of the suitability of existing metrics in the literature.

Accuracy

The clustering performance of the model was evaluated using the unsupervised clustering accuracy metric (denoted Acc), a popular performance metric in the evaluation of unsupervised clustering models, defined as:

$$\text{Acc} = \max_{m \in \mathbb{M}} \frac{1}{N} \sum_{i=1}^{N} \delta_{l_i, m(c_i)} \tag{5.9}$$

where N is the number of samples, and l_i and c_i are the ground truth label and assigned label for the i_{th} sample, respectively. $\mathbb{M}$ is the set of mappings between the assigned labels and the ground truth labels; m is the optimal mapping between the assigned labels and the ground truth labels, found using the Jonker-Volgenant algorithm [157] implemented in the SciPy Python package [158]. Acc can take values between 0 and 1: a value of Acc = 1 would indicate perfect clustering [159–161].

Disentanglement

A disentangled representation is generally understood to possess the following 3 properties:

1. **Modularity**: features are uncorrelated and causally independent from each other

2. **Compactness**: each generative factor of the data is encoded in a single feature of the representation

3. **Explicitness**: the features encode all the important generative factors of the data

[162]

Perhaps the most commonly used disentanglement metric is the mutual information gap (MIG), in which the mutual information between all factors and features is calculated, and the 'gap' between the highest and second highest mutual information values is returned as the MIG [136]. The MIG provides a simple, robust way to evaluate the compactness of a representation, but does not account for modularity [162]. In this work, disentanglement performance was evaluated using the JEMMIG (joint entropy minus mutual information gap) metric, an extension to the MIG which evaluates both modularity and compactness [162, 163]. JEMMIG can take a value between 0 and 1, where 1 indicates perfect disentanglement.

The use of JEMMIG was also supplemented by qualitative evaluation of the feature space to support conclusions drawn regarding the disentanglement performance of the models evaluated here.

5.5.2 Procedurally Generated Evaluation Datasets

In order to develop a robust and unbiased data analysis process and evaluate the performance of different approaches, labelled datasets containing a controlled set of defined event types was required. Procedurally generated 'synthetic' data was used for this purpose since experimental STMBJ data is inherently unlabelled and could contain unexpected event types (for example due to sample contamination). In addition, since the true generative factors of experimental data are innumerable and generally unknown, the application of disentanglement metrics (see section 5.5.1) to experimental STMBJ data is not possible in practice.

Two synthetic datasets were generated for use in this work: dataset I, a simple dataset completely defined by 5 generative factors, and dataset II, a more complex dataset more accurately representing real STMBJ data. Dataset I was used for the tuning and development of the model, as well as for evaluating the disentanglement performance, while dataset II was used as a final test to evaluate the accuracy of the model on realistic data before application to real experimental datasets. Both datasets were designed to exhibit features corresponding to real physical processes observed in STMBJ data in the literature, as discussed in section 5.4.

Dataset I: Known Generative Factors

Dataset I contains simple emulations of STMBJ G-Z traces, all of which contain single conductance plateaus with added sinusoidal noise. Each trace is completely defined by 5 continuous random variables, or 'generative factors', denoted f_i, $i \in \{\mathbb{N}, 1 \leq i \leq 5\}$, which correspond to:

- f_1: conductance at the start of the plateau
- f_2: conductance at the end of the plateau
- f_3: length of the plateau
- f_4: amplitude of sinusoidal noise
- f_5: period (in nm) of sinusoidal noise

Each factor can take a value between 0 and 1, which maps to a predefined range of values for the relevant parameters. Randomly sampling these factors allows a unique synthetic trace to be generated.

Dataset I consists of 10,000 individual traces, evenly distributed amongst 10 classes. An example trace from each class is shown in figure 5.8. Classes were created by defining normal distributions with selected means and standard distributions for each factor, and the traces within the class were generated by randomly sampling from these distributions. The mean and standard deviation for each factor in the different classes is presented in figure 5.9. Figure 5.10 is a principal component analysis (PCA) projection of the generative factors of the dataset into 2D representation, showing the distribution of the data in the factor space.

Despite its apparent simplicity, this dataset in fact presents a significant classification challenge. Since all traces contain plateaus, the only differences between classes are in the shape of the plateau and noise properties of the trace. These subtle variations are largely impossible to detect with all-data histogram based methods, and are still much more difficult for clustering-based methods than datasets with larger variations such as dataset II (see section 5.5.2) or even typical experimentally gathered data. Figures 5.8 and 5.9 show that even by eye, it is easier in some cases to distinguish the classes by their underlying generative features than by the G-Z traces accessible to the model. Hence this dataset is expected to be a good test of a model's ability to extract meaningful features of STMBJ data.

Dataset II: Unknown Generative Factors

Dataset II is intended to more closely represent experimental STMBJ data than dataset I, with increased noise and complexity and a wider range of distinct trace

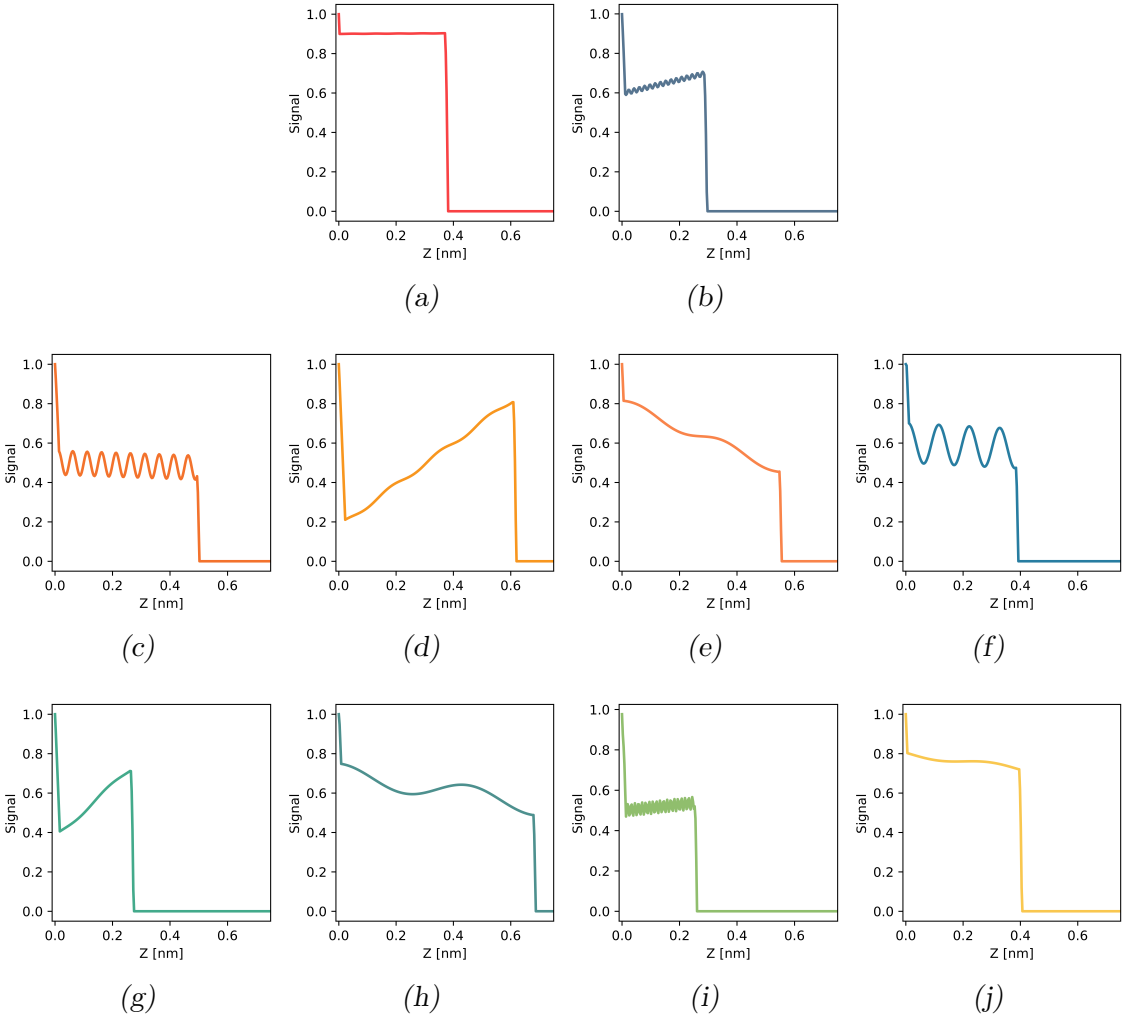


Fig. 5.8: Representative traces from each class in synthetic dataset I

types (such as traces with no plateau, or with multiple plateaus). As such, the data cannot be readily summarised by a small number of generative factors, and disentanglement metrics (see 5.5.1) are not compatible with this dataset. Dataset II also features uneven proportions of the different event classes, while in dataset I the data is evenly distributed across all 10 classes.

The synthetic data was generated using an exponential decay function to model the current-distance dependence in a tunnelling junction. At certain predefined conductance levels a plateau could be inserted to imitate the presence of a molecule in the junction. Small random variations in the conductance and length of the plateau were introduced between individual traces to simulate variations in experimental data. A random z offset was also added to the entire trace to emulate real STMBJ measurements, where the point at which the metallic junction breaks varies between measurements. Various sources of noise were added to each trace to more closely model experimental data. Two sinusoidal noise signals were superposed onto the trace, one at 50 Hz frequency to simulate mains electrical noise and another at 0.2

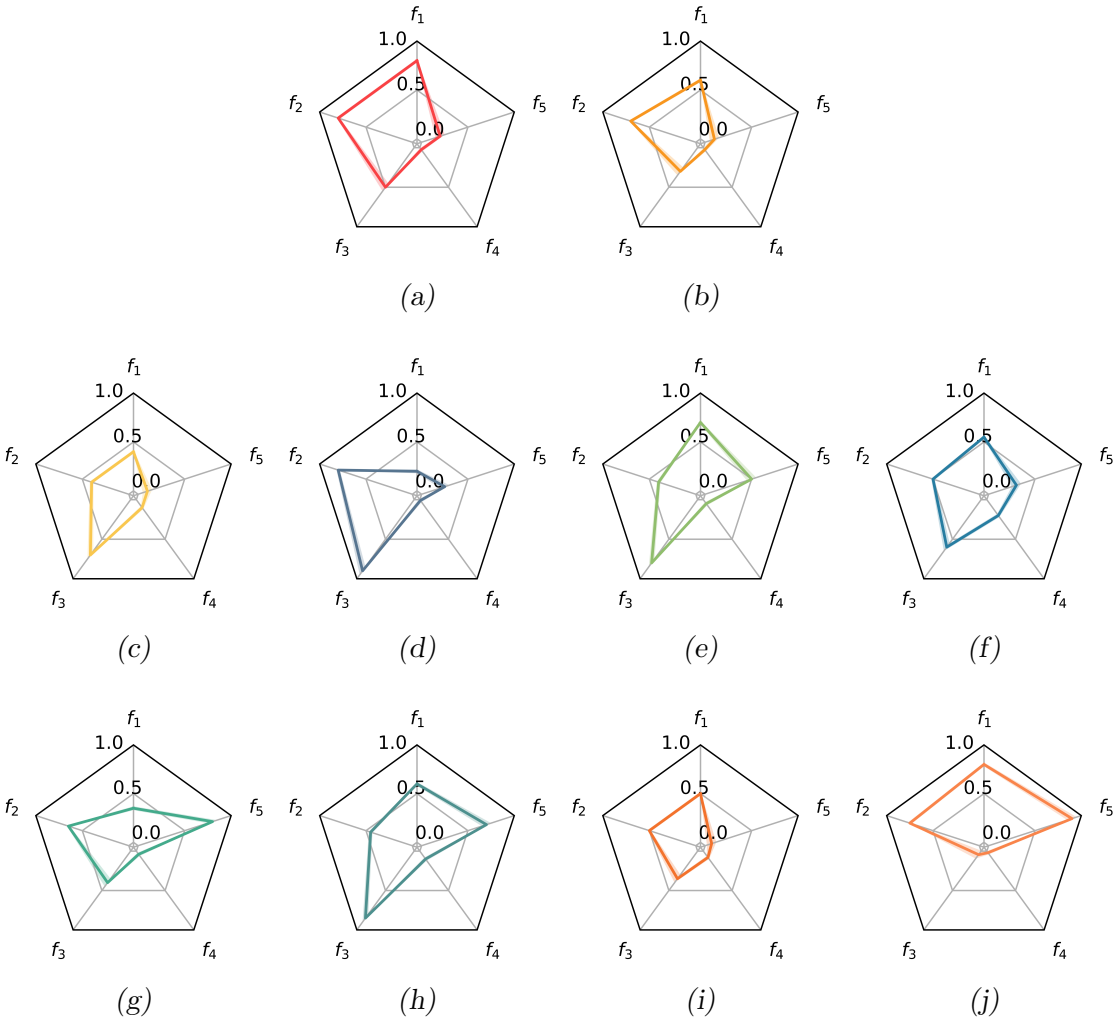


Fig. 5.9: Radar diagrams showing the distribution of factors for each class of traces in synthetic dataset I

Hz to simulate the low frequency 'flicker noise' present in physical measurements. The phase of these periodic noise signals was randomly varied between different traces. Additionally, non-periodic noise based on a Gaussian distribution was added to each sample to emulate the quantisation noise of the ADC.

Dataset II contains 10,000 traces, distributed across 10 classes in the proportions shown in figure 5.11. The characteristics of each class were chosen to test the capability of the classification algorithm to distinguish between some different event types which may occur in an STMBJ measurement (see section 5.4); in particular, the dataset was designed to contain events which cannot be identified using all-data histogram-based approaches. The dataset was also designed to closely match the measurement parameters of the actual STMBJ measurements conducted in this work: each trace contained 4096 samples distributed over a z range of 11.5 nm, and the conductance range was matched with the dynamic range of the current amplifier used in the STMBJ (see section 4.2.8). Figure 5.12 shows example traces from each

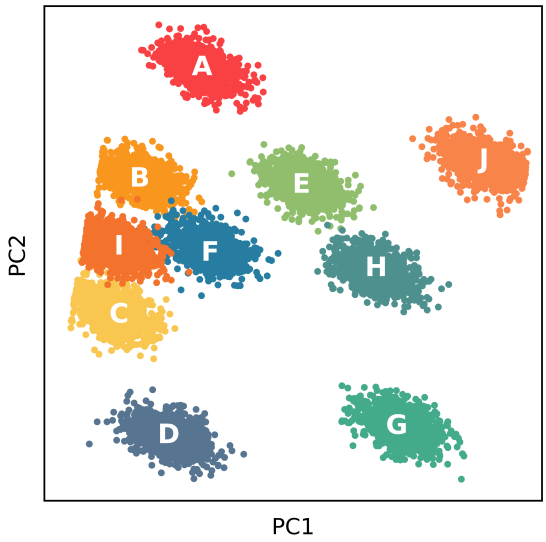


Fig. 5.10: 2D PCA projection of the space of generative factors for dataset I

class in dataset II.

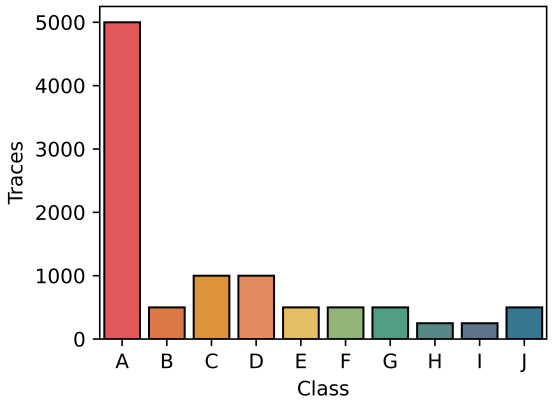


Fig. 5.11: A bar chart showing the number of traces in each class of synthetic dataset II

Baseline Metrics

In order to provide a basis of comparison for later tests of the feature extraction and clustering models, a baseline accuracy was obtained for an uninformative clustering of synthetic dataset I. Cluster labels were randomly assigned to each trace in the dataset, and the unsupervised clustering accuracy (Acc) was evaluated. This was repeated 1000 times while varying the number of clusters in order to obtain a distribution. The results, shown in figure 5.13, reveal that randomly assigning cluster labels cannot achieve an Acc score higher than 0.13 for synthetic dataset I.

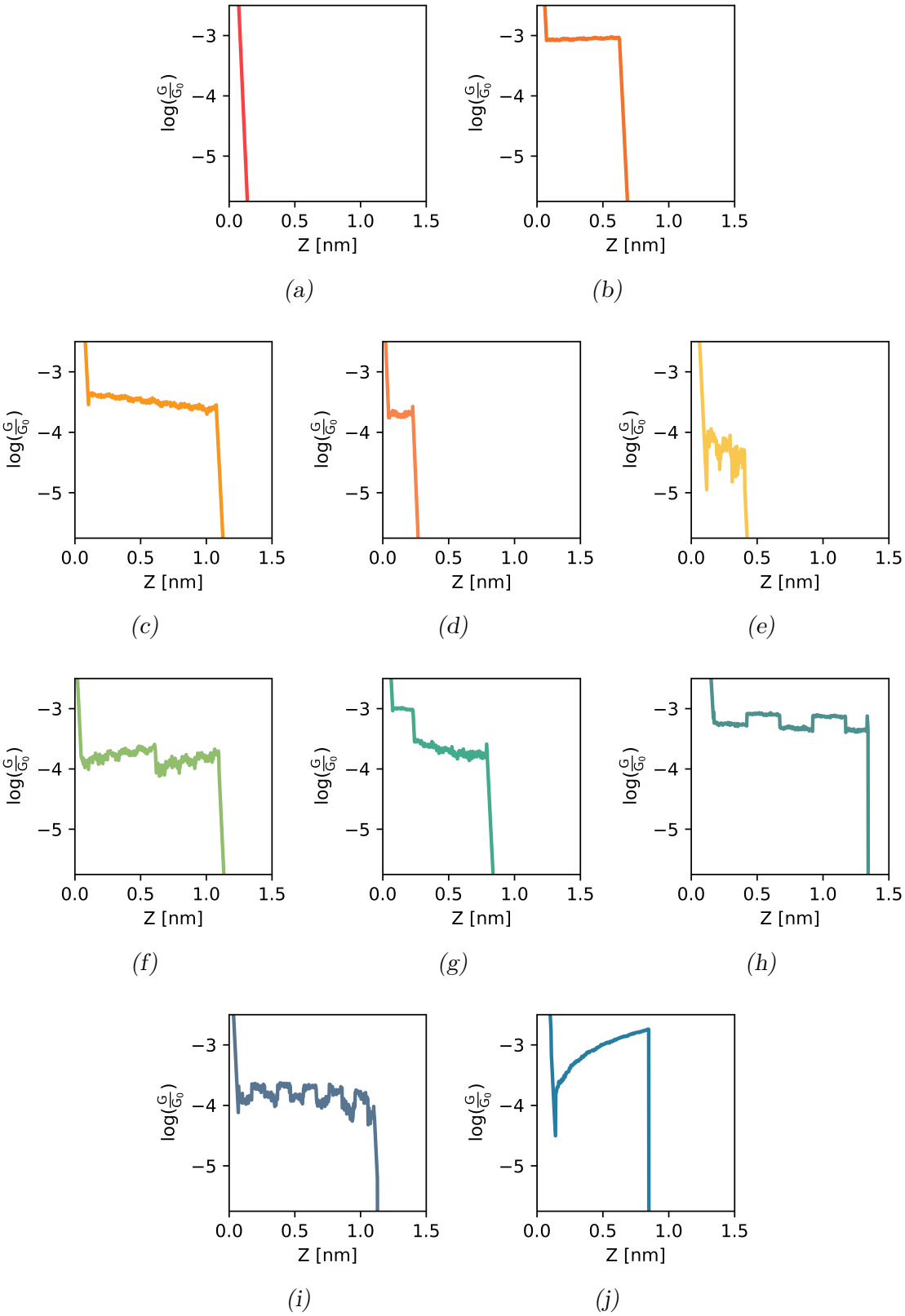
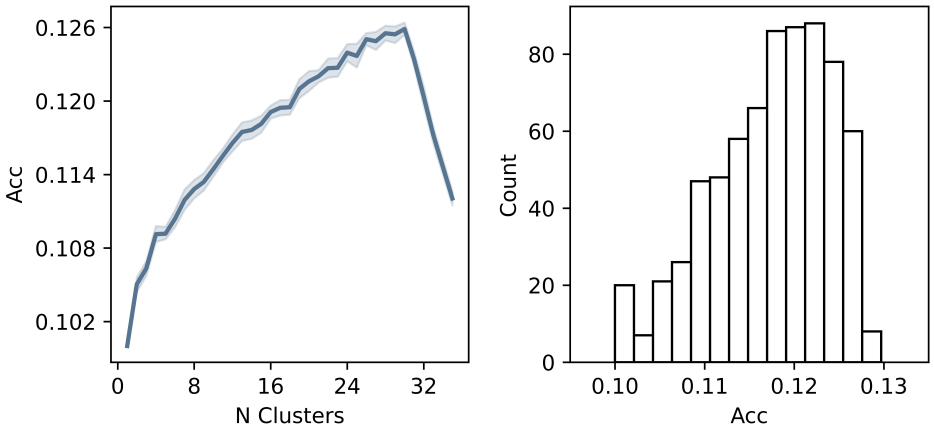


Fig. 5.12: Representative traces from each of the 10 event classes in synthetic dataset II



(a) *(b)*

Fig. 5.13: Baseline plots of unsupervised clustering accuracy (Acc) for randomly assigned clusters on synthetic dataset I. **(a)** Acc plotted against different numbers of randomly assigned clusters **(b)** A histogram of all the baseline Acc values across all tested cluster numbers

5.6 Testing and Optimisation

5.6.1 Base Model Configuration

The feature extractor investigated in this work was based on a βVAE implemented in PyTorch. The initial configuration of the model was as follows: the encoder consisted of an input layer with 392 nodes (to match the preprocessed STMBJ traces used in this work), followed by a single 100 node hidden layer and finally the latent layer, which initially contained 8 nodes (resulting in an 8-dimensional feature space). The decoder was a symmetrical copy of the encoder, and ReLU activation functions were used between all layers. The output layer was conditioned with a sigmoid activation function to ensure that the reconstructed traces were normalised, since the ReLU function is unbound in the positive direction.

This base configuration was gradually optimised through a series of experiments outlined in the following sections.

5.6.2 Optimiser Tuning

The AdamW optimiser [164], was selected for training the models studied in this work. The main hyperparameters of AdamW are the learning rate, which controls the coarseness of the steps taken when adjusting network weights, and the weight decay, which adjusts the rate at which weights of the network decay with each training step (a regularisation technique used to avoid overfitting) [165].

Due to constraints on processing time, a full hyperparameter optimisation could not be performed, but in order to select adequate parameters a coarse grid search

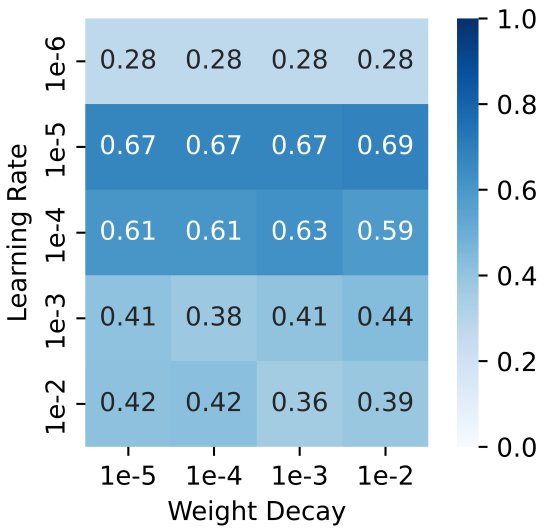


Fig. 5.14: A heatmap showing unsupervised clustering accuracy (Acc) of the base configuration (see section 5.6.1) for synthetic dataset I over different hyperparameter values for the AdamW optimiser

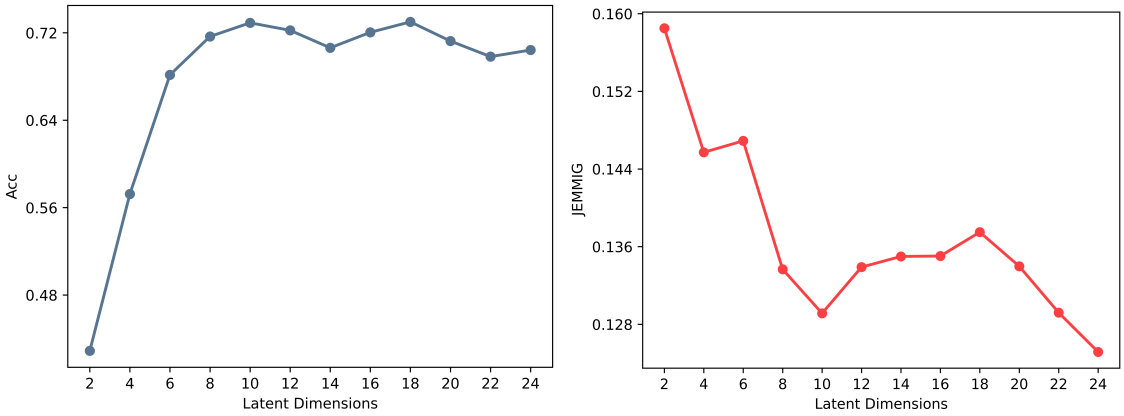
over 4 values of weight decay and 5 values of learning rate was performed, using the initial model described in section 5.6.1 and synthetic dataset I, described in section 5.5.2. Clustering was applied to the extracted features (see section 5.3 for details on clustering), and the unsupervised clustering accuracy, Acc (defined in section 5.5.1), was calculated for each set of parameters. A heatmap of the results is presented in figure 5.14. The results show that the downstream clustering performance strongly depends on the learning rate, but has no clear dependence on weight decay. Note that it is likely that the optimal parameters vary with different training duration and model hyperparameters; a full exploration of this would require significant additional computation time and is beyond the scope of this work.

Based on the grid search, the learning rate was set at 1×10^{-5} and the weight decay was set at 1×10^{-4} for all tests presented in this work.

5.6.3 Latent Layer Capacity

The capacity of the information bottleneck in VAE is directly connected to the dimensionality of the latent layer(s). The network described in section 5.6.1 was trained from scratch for 512 epochs on dataset I (see section 5.5.2) 12 times, with an increased number of latent dimensions each time. The performance metrics for the accuracy and disentanglement of downstream clustering applied to the features extracted with each number of latent dimensions are presented in figure 5.15.

The results reveal that the disentanglement score tends to be higher for features extracted using lower dimensionality latent layers. This result may be understood intuitively using similar arguments to those used to motivate βVAE (discussed in



(a) *(b)*

Fig. 5.15: Tests of the effect of VAE latent layer dimensionality on accuracy and disentanglement metrics for the base configuration on synthetic dataset I. **(a)** Acc plotted against number of latent dimensions **(b)** JEMMIG plotted against number of latent dimensions

section 5.2.3): as the information capacity of the latent bottleneck is restricted, the network is encouraged to learn only the most important generative factors (i.e. the factors which have the greatest impact on the log-likelihood) in a disentangled representation in order to make efficient use of the limited capacity [138]. Between 2 and 10 latent dimensions, the disentanglement score (JEMMIG) decreases as the dimensionality of the latent bottleneck is increased, as expected from the information capacity argument. However, between 10 and 18 latent dimensions the disentanglement begins to increase again before reaching a local maximum at 18 latent dimensions and falling sharply as the dimensionality increases through the range 20 to 24. This could be tentatively explained by the VAE beginning to encode more of the generative factors into the increased capacity, until at some critical point the information capacity of the latent representation is high enough that the VAE is able to encode representations which minimise both reconstruction and KL loss without disentanglement pressure. A more rigorous examination of the relationship between disentanglement, information bottleneck capacity and latent layer dimensionality may be informative. For example, a grid search comparing different β parameters and latent layer sizes may elucidate any similarities between these two forms of information capacity restriction.

The accuracy metric supports the notion that the feature extractor is 'ignoring' the less significant generative factors for lower dimensionality; while the dimensionality is lower than the number of generative factors (5) in the dataset, the accuracy rapidly increases with each increase in dimensionality, suggesting that the VAE may be prioritising encoding one factor per dimension in this region. Once the number of dimensions exceeds the number of factors, the accuracy remains largely constant

as the dimensionality increases.

Based on these tests, the dimensionality of the latent layer of the VAE was set to 18 for all further work.

5.6.4 β Annealing

The effects of cyclical β annealing (introduced in section 5.2.3) on the performance of the VAE model were studied, and both the standard VAE and the cyclically annealed βVAE (denoted 'CyclicalVAE' here) were compared to a standard AE model (obtained by setting $\beta = 0$).

Cyclical annealing was introduced by modulating the β parameter with a sawtooth waveform during training. The modulation was bound within the range $0 \leq \beta \leq 1$ with a period of 64 epochs. Each model was trained for 512 epochs, and figure 5.16 shows the evolution of the losses during training for each model, alongside the modulation of the β value.

Clustering accuracy and disentanglement were evaluated over 10 different random seeds to obtain average performance metrics. The results are presented in figure 5.17. A clear enhancement in clustering accuracy for both VAE-based models over the standard AE feature extractor was demonstrated, with the cyclically annealed βVAE (CyclicalVAE) obtaining the highest peak Acc score. The JEMMIG scores showed a less pronounced enhancement: the standard VAE was found to have by far the highest peak JEMMIG score, but the average score was similar to that of the standard AE. The cyclically annealed βVAE obtained the highest JEMMIG score on average.

In order to better understand the JEMMIG results obtained for each model (in particular the surprising similarity between the evaluated performance for the VAE and AE models), a more qualitative investigation of the feature space was performed:

First, 18×18 correlation matrices were plotted for the extracted features (i.e. the correlation between each pair of features was calculated and plotted as a heatmap). These are presented in figure 5.18. Note that the high correlation values on the rising diagonal represent the perfect correlation between any feature and itself, while high values in off-diagonal elements indicate correlations between different features. Clearly, a representation with good modularity (see section 5.5.1) could be expected to possess lower off-diagonal values in the correlation matrix. Figure 5.18 clearly shows a substantial improvement in modularity in the learned representation for the VAE over the AE model, while the cyclicalVAE appears to produce slightly higher modularity than the VAE.

Next, 5×18 correlation matrices were plotted between the extracted features and the generative factors of the dataset. These are presented in figure 5.19. Representations with high compactness (see section 5.5.1) are expected to distribute

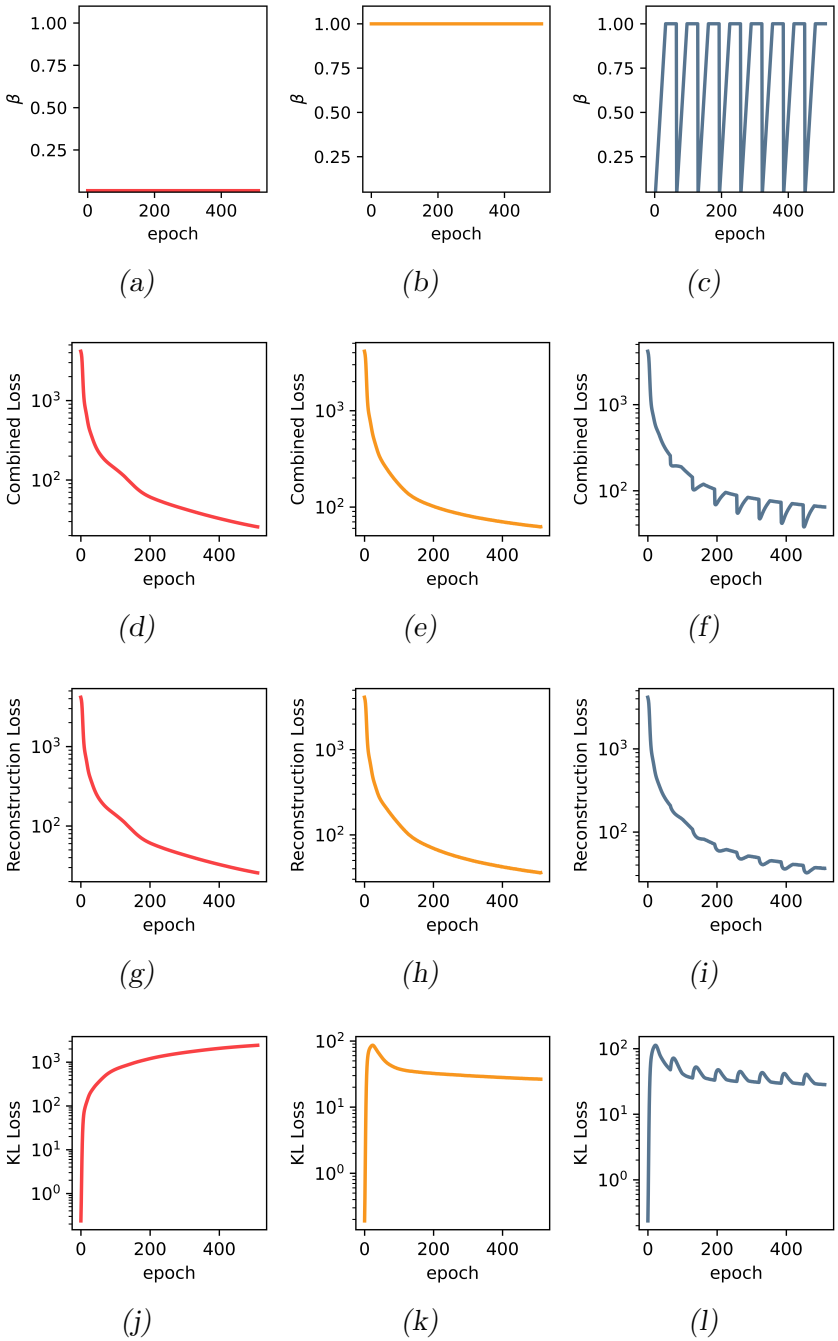
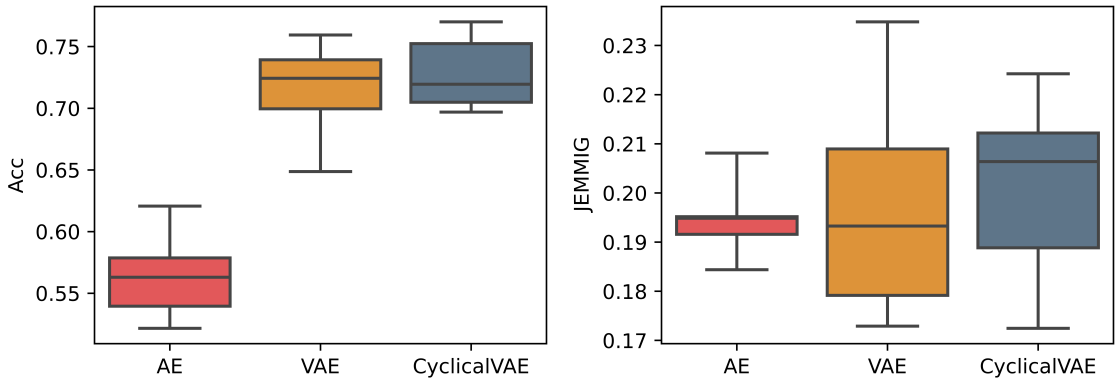
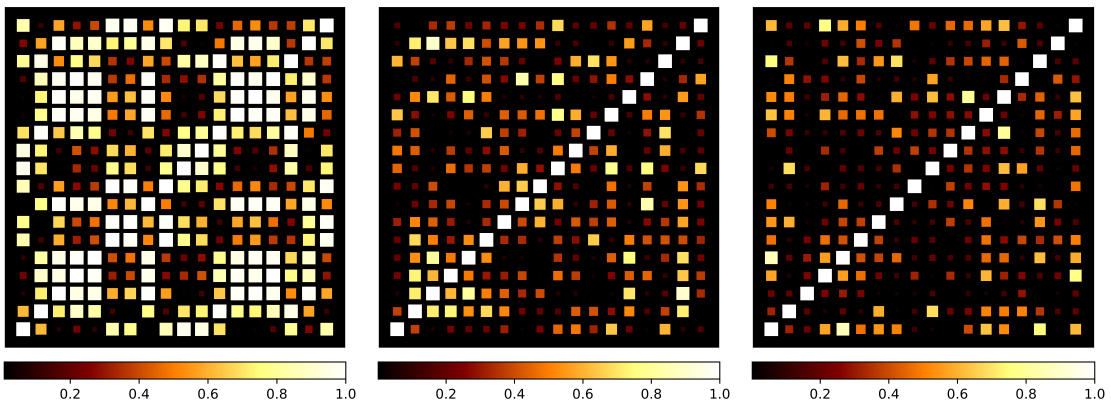


Fig. 5.16: Losses during training on synthetic dataset I for each of the feature extraction models tested. **(a)** β factor versus epoch for AE **(b)** β factor versus epoch for VAE **(c)** β factor versus epoch for CyclicalVAE **(d)** loss function versus epoch for AE **(e)** loss function versus epoch for VAE **(f)** loss function versus epoch for CyclicalVAE **(g)** reconstruction loss term versus epoch for AE **(h)** reconstruction loss term versus epoch for VAE **(i)** reconstruction loss term versus epoch for CyclicalVAE **(j)** KL loss term versus epoch for AE **(k)** KL loss term versus epoch for VAE **(l)** KL loss term versus epoch for CyclicalVAE



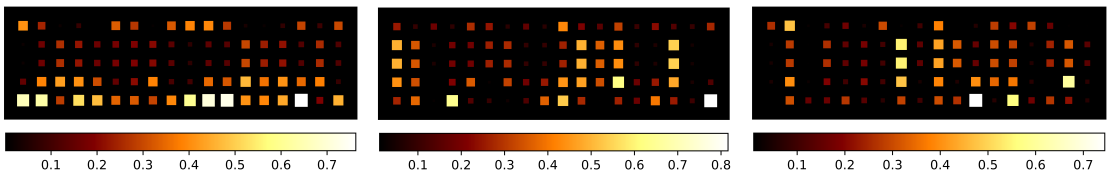
(a) *(b)*

Fig. 5.17: Comparison of performance metrics for a standard AE, standard VAE and a βVAE with cyclical β annealing for synthetic dataset I. **(a)** Unsupervised clustering accuracy (Acc) **(b)** Disentanglement score (JEMMIG)



(a) *(b)* *(c)*

Fig. 5.18: Correlation matrices for the features extracted from synthetic dataset I by **(a)** standard AE **(b)** standard VAE **(c)** βVAE with cyclical β annealing



(a) *(b)* *(c)*

Fig. 5.19: Correlation matrices between the extracted features and the generative factors of dataset I by **(a)** standard AE **(b)** standard VAE **(c)** βVAE with cyclical β annealing

information about the factors over as few feature dimensions as possible. Hence a higher compactness should correspond to a matrix with only a few (or ideally one) high correlation values in each row. Again, the standard AE appears to be outperformed by both VAE models in this regard, while the CyclicalVAE demonstrates the highest performance. However, the difference between models is less pronounced when compared to the modularity comparisons in figure 5.18. This provides a possible explanation for the similar JEMMIG performance between the VAE and the AE, suggesting that JEMMIG score is dominated by compactness, rather than modularity.

In summary, the VAE model investigated here has shown a significant increase in the modularity of extracted features compared to a standard AE, resulting in a corresponding increase in downstream clustering accuracy. Additionally, introducing cyclical annealing to the β term has been shown to further improve performance in both clustering accuracy and disentanglement.

5.6.5 Application to Dataset II

All previous tuning and tests were performed using synthetic dataset I. If the feature extraction model is to be used for real-world data, it is important to verify that the performance attained can be generalised to different STMBJ datasets without re-tuning the parameters for the specific dataset. Hence the classification test presented in section 5.6.4, in which the performance of AE, VAE and cyclicalVAE models were compared, was repeated using synthetic dataset II. Dataset II contains many event types (such as double plateaus) and noise sources (such as wideband white noise and simulated AC pickup) which are not present in dataset I, and therefore provides a good test of each model's ability to extract features from real data without specific re-tuning. Details of synthetic datasets I and II can be found in section 5.5.2.

As in section 5.6.4, each model was trained for 512 epochs; the loss curves from the training process are shown in figure 5.20.

After training for 512 epochs, each model was used to extract features from every

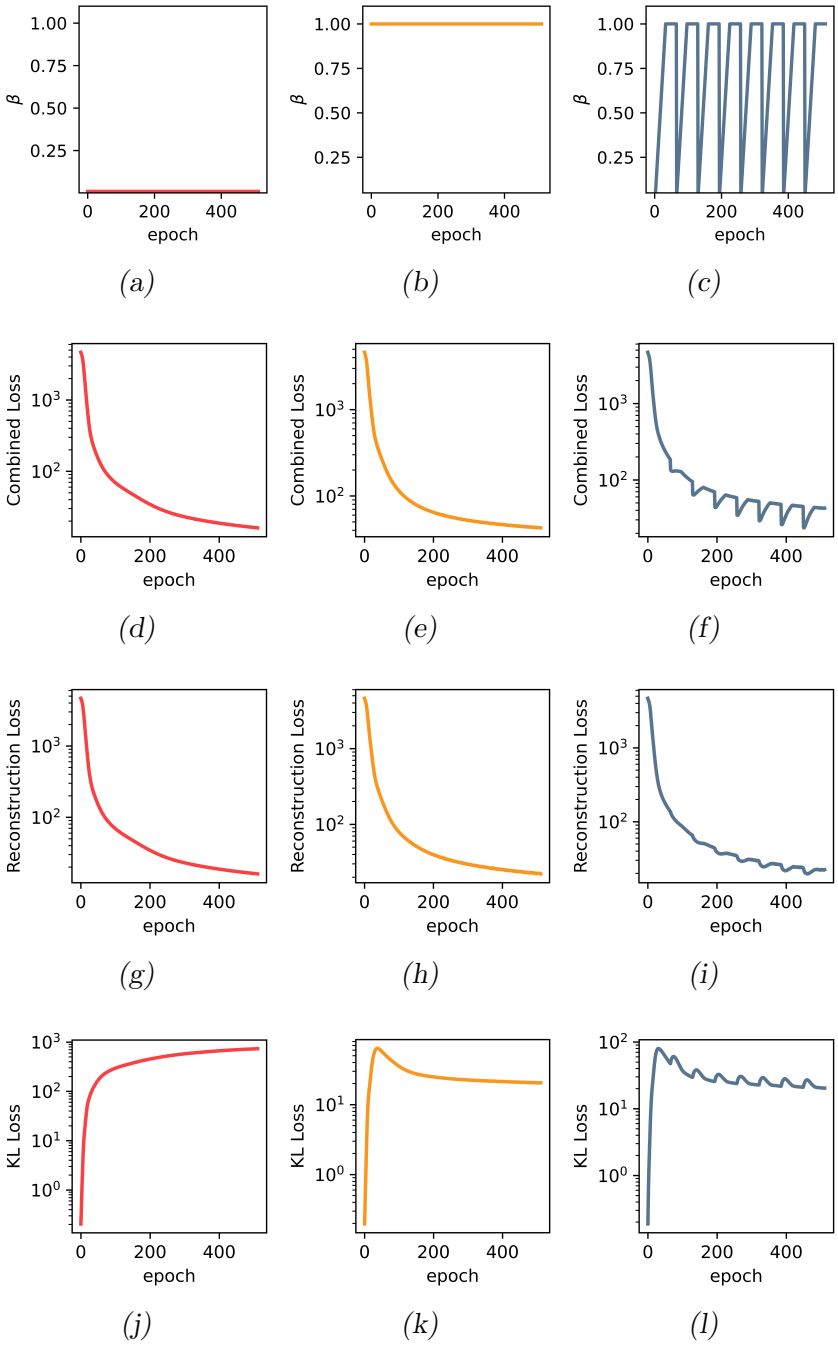


Fig. 5.20: Losses during training on synthetic dataset II for each of the feature extraction models tested. **(a)** β factor versus epoch for AE **(b)** β factor versus epoch for VAE **(c)** β factor versus epoch for CyclicalVAE **(d)** loss function versus epoch for AE **(e)** loss function versus epoch for VAE **(f)** loss function versus epoch for CyclicalVAE **(g)** reconstruction loss term versus epoch for AE **(h)** reconstruction loss term versus epoch for VAE **(i)** reconstruction loss term versus epoch for CyclicalVAE **(j)** KL loss term versus epoch for AE **(k)** KL loss term versus epoch for VAE **(l)** KL loss term versus epoch for CyclicalVAE

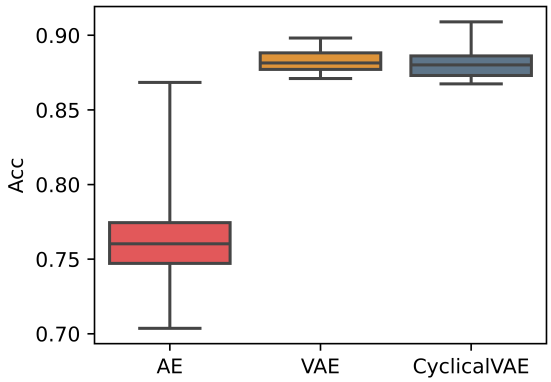


Fig. 5.21: Comparison of unsupervised clustering accuracy (Acc) for a standard AE, standard VAE and a βVAE with cyclical β annealing for synthetic dataset II

element of dataset II. These 18-dimensional features were then clustered using the DPMM-based algorithm discussed in section 5.3. This was repeated 10 times for each model, with a different random seed each time. A box plot of the unsupervised clustering accuracy (Acc) metric for all tests is shown in figure 5.21. The results are similar to those found for dataset I, with the VAE and CyclicalVAE models outperforming the standard AE. The VAE attained higher Acc scores on average than the CyclicalVAE by a small margin, however the CyclicalVAE obtained the highest peak Acc score. Interestingly, the average accuracy achieved by all models for dataset II was higher than that for dataset I, despite the increased complexity of the dataset. This supports the argument in section 5.5.2 that despite the apparent simplicity of dataset I compared to real-world data, the dataset presents a challenging benchmark test for clustering algorithms due to the qualitative similarity of all the event types contained within.

Unlike dataset I, the generative factors of dataset II are unknown due to the complexity of the procedural generation process used; hence the JEMMIG disentanglement metric is not applicable to this test. Instead, the correlation matrix approach introduced in section 5.6.4 is used. Figure 5.22 shows 18-dimensional correlation matrices for the features extracted by each model. By comparison with the equivalent result for dataset I (figure 5.18), it can be inferred that the enhanced disentanglement observed for the VAE and CyclicalVAE persists for dataset II, showing that the VAE models are capable of learning disentangled representations even for data with many unknown generative factors.

5.7 Application to Experimental Data

The real-world capabilities of the feature extraction and clustering algorithm presented here were evaluated for an experimental STMBJ dataset for Au-ODT-Au

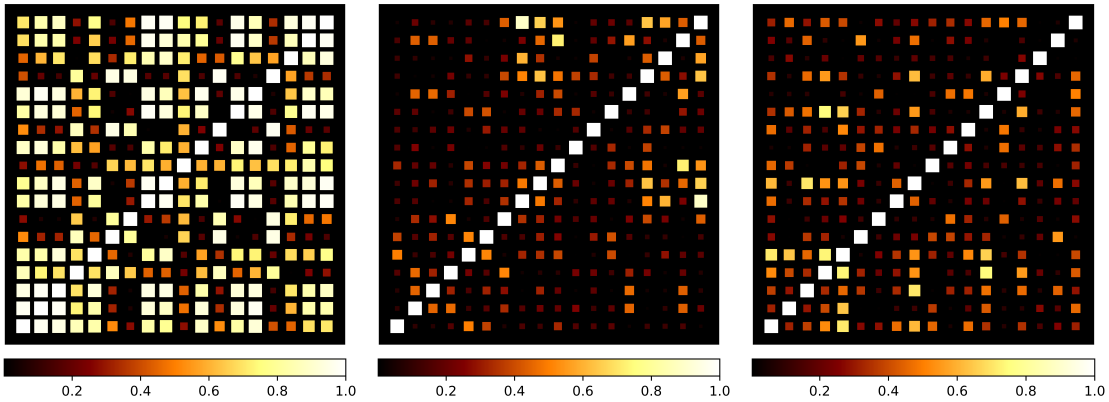


Fig. 5.22: Correlation matrices for the features extracted from synthetic dataset II by **(a)** standard AE **(b)** standard VAE **(c)** βVAE with cyclical β annealing

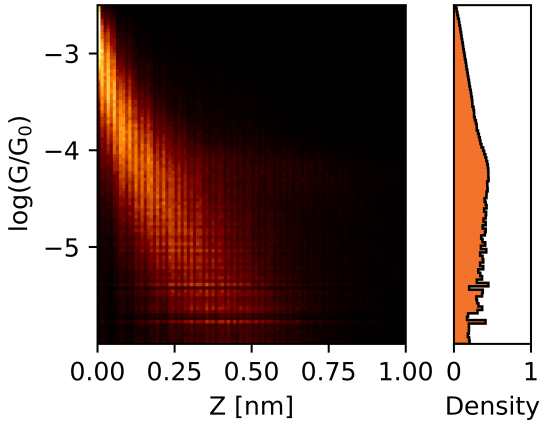


Fig. 5.23: All-data 2D and 1D histograms for the ODT dataset

junctions captured using the instrument developed in chapter 4. The dataset contained 7232 I-Z traces, each obtained with a bias voltage of 0.4 V, a sample rate of 2 K Hz and a retraction rate of 5 nm/s.

5.7.1 Raw Data

The entire ODT dataset, consisting of 7232 traces, was compiled into 2D and 1D conductance histograms (figure 5.23) to test the feasibility of a traditional 'all-data' approach. From this it is possible to observe that a plateau does occur in some traces, and that the most-probable conductance lies between 1×10^{-5} G_0 and 1×10^{-4} G_0, in line with commonly reported values in the literature [67, 70, 72, 166]. However, it is not possible to infer the presence of multiple conductance values, accurately determine the average breaking length, or observe any distance/time-dependent behaviours such as switching. Clearly, the analysis of this dataset could benefit from a clustering approach.

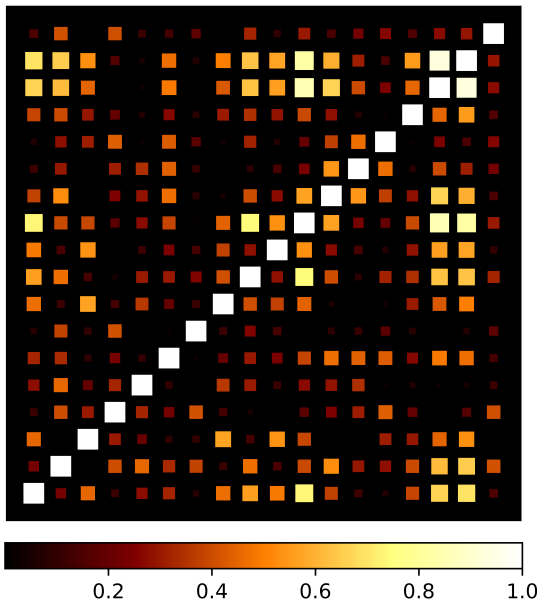


Fig. 5.24: Correlation matrix for the 18 features extracted for the ODT dataset

5.7.2 Extracted Features

After training, an 18-dimensional feature vector was extracted for each trace.

The correlation matrix for the features is shown in figure 5.24. While the generative factors are not available for real-world data, comparison with the results in section 5.6.4 suggests that this correlation matrix indicates reasonable disentanglement.

5.7.3 Clustering Results

The extracted features for the ODT dataset were clustered using the DPMM clustering algorithm described in section 5.3. A total of 9 distinct clusters, labelled A to I, were identified by the algorithm; a summary of how many traces were contained in each cluster is provided in figure 5.25.

2D $log(G/G_0)$-Z histograms, alongside collapsed 1D $log(G/G_0)$ histograms, for each cluster are presented in figure 5.26. Comparison to figure 5.23 demonstrates that the clustering has revealed a wealth of information not accessible using the all-data approach, including several distinct conductance plateaus. Discussion of the physical interpretation of this clustering result is presented in section 5.7.4

5.7.4 Interpretation

In section 5.7.3, the experimental Au-ODT-Au STMBJ dataset was partitioned into nine clusters. Five of these clusters (A, B, C, D and E) contain either no plateau, or very short unstable plateaus which cannot be observed in the 2D/1D histograms in

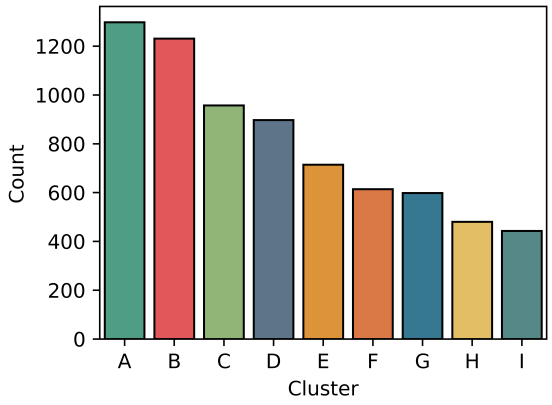


Fig. 5.25: A bar chart showing the number of traces in each of the 9 clusters identified for the ODT dataset

figure 5.26, suggesting that these clusters correspond to G-Z traces where no stable molecular junction was formed. The remaining four clusters (F, G, H and I) each show evidence of molecular junction formation, but with different plateau shapes, breaking lengths and average conductance values. The most probable conductance and breaking lengths were determined for each group (where a conductance plateau was present). These are summarised in table 5.1.

Cluster	$log(\frac{G}{G_0})$	Breaking Length (nm)
A	No plateau	No plateau
B	No plateau	No plateau
C	No Plateau	No plateau
D	No Plateau	No Plateau
E	No plateau	No plateau
F	-4.1 ± 0.5	0.5 ± 0.1
G	-5.1 ± 0.5	0.5 ± 0.1
H	-4.9 ± 0.9	0.4 ± 0.1
I	-4.4 ± 0.4	0.8 ± 0.2

Tab. 5.1: A summary of the conductance and breaking length for each cluster.

Clusters F and I contain plateaus with similar conductance, $log(G/G_0) = -4.1$ and $log(G/G_0) = -4.4$ respectively. Cluster G contains a lower conductance plateau ($log(G/G_0) = -5.1$). Cluster H contains a plateau with similar average conductance ($log(G/G_0) = -4.9$) to that of cluster G, but with a wider range of conductance values in the plateau. All the plateaus have similar breaking lengths around 0.5 nm, with the exception of the plateau in cluster I which has an average breaking length of 0.8 nm.

Inspection of the 2D histograms (5.26) reveals differences in the plateau shapes for each of these clusters. Cluster I contains a 'flat' plateau, with no discernible

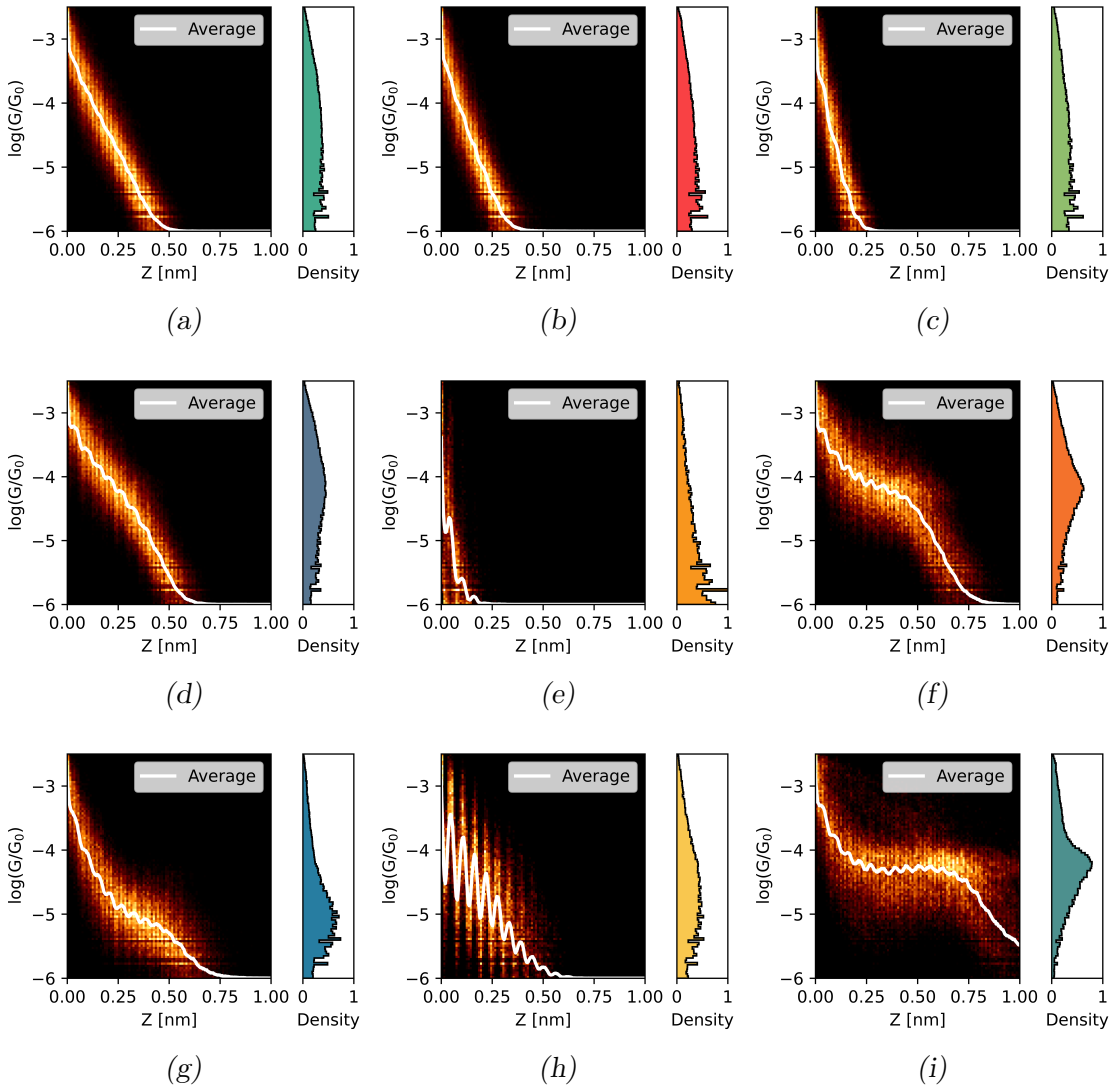


Fig. 5.26: 2D G-Z histograms of the 9 clusters identified for ODT

slope, while clusters F, G and H contain negatively sloped plateaus. Additionally, cluster H exhibits large-amplitude conductance modulation, with a spatial period of around 0.05 nm, corresponding to a frequency of around 115 Hz; this may simply be attributed to a harmonic of AC line interference from some intermittent source which did not affect all traces, or may correspond to some physical process in the junction causing a periodic modulation of the conductance. The signal contributed by the individual traces appears to be in phase (otherwise the periodicity of the modulation would not be visible in the 2D histogram and average trace), which supports the idea that this is a spatially dependent modulation of the junction conductance and not an external time-varying noise source.

The distinctions between clusters A, B, C, D, and E are more difficult to decipher. From the 2D histograms, these may be interpreted to be traces without molecular plateaus, with different exponential decay factors. Upon manual inspection of the

individual traces within cluster D, evidence of very short unstable plateaus was observed. None of these clusters feature a pronounced enough plateau in the conductance histograms for a statistically meaningful interpretation, and these clusters are therefore excluded from the analysis as 'empty' clusters.

In order to further explore the differences between the identified clusters, plots of the mean power spectral density (PSD) of the raw signal were produced for each cluster (figure 5.27). It is immediately apparent that the clusters which most clearly possess molecular junctions (F, G, H and I) tend to have higher 1/f noise, and indeed higher noise across the frequency spectrum than the other clusters. All clusters appear to exhibit a peak in the 100-200 Hz range, possibly corresponding to the conductance modulation observed in the 2D histogram for cluster H. This peak appears sharper and more pronounced in the clusters which most clearly exhibit molecular plateaus, although the presence of similar peaks in the 'empty' plots suggests that the observed conductance modulation may not originate within the molecular junction. In future, access to lower frequencies in the PSD could be obtained by pausing the probe motion periodically during withdrawal to obtain stable measurements over longer periods. This may provide additional insight into the behaviour of the junction, including how the charge transport mechanism changes throughout the G-Z measurement [93].

The single-molecule conductance of ODT has been widely measured in the literature, with typical $log(G/G_0)$ values around -4.4 [67, 72], comparable to clusters F and I. Lower conductance states, such as that observed in clusters G and H, have also been reported in studies of ODT and other alkanedithiols, and a number of mechanisms (typically related to variations in the Au-molecule contacts) have been suggested for the presence of multiple conductance states [66, 68–70, 72, 167, 168].

One such study [72] has reported observation of three distinct conductance states for Au-ODT-Au junctions formed by STMBJ: a 'high' conductance state at $log(G/G_0) = -3.6$, a 'medium' conductance state at $log(G/G_0) = -4.2$, and a 'low' conductance state at $log(G/G_0) = -5.0$. The high and medium states were attributed to different binding geometries of the molecule to the Au electrodes (as discussed in section 5.4): the high conductance state was assigned to the bridge configuration (in which the binding site is between two Au atoms) and the medium conductance state to the atop configuration (in which the binding site is located directly on a single Au atom). The low state was attributed to a gauche defect in the junction. The average conductance values for clusters F and I are in agreement with the medium conductance state, while clusters G and H are consistent with the low conductance state. No cluster corresponding to the high conductance state was obtained in this work.

There is also a small difference between the $log(G/G_0)$ values obtained for clus-

ters F (-4.1) and I (-4.4), suggesting the existence of additional conductance levels within the 'medium' conductance state. Indeed, a recent study [155] of Au-alkanedithiol-Au junctions demonstrated the existence of two different molecule-electrode binding regimes, with different conductance levels: the SH anchor group of the molecule may form a covalent S-Au bond, or a weaker SH-Au coordination bond (this is discussed in section 5.4.2). Break-junction measurements of Au-ODT-Au junctions obtained a value of $log(G/G_0) = -4.0$ for S-Au binding, and $log(G/G_0) = -4.3$ for SH-Au binding. The S-Au binding case was also found to produce a negatively sloped plateau, whereas the SH-Au binding case produced a flatter plateau. Additionally the measured conductance noise was reported to be lower for S-Au binding than SH-Au binding. These observations are consistent with the conductance values (table 5.1), plateau slopes (figure 5.26) and noise levels (5.27) for clusters F and I: hence cluster F may correspond to S-Au binding while cluster I may be attributed to the SH-Au binding case. The aforementioned study [155] only observed the S-Au binding case for in-solution STMBJ measurements, and conversely only observed the SH-Au binding for 'dry' STMBJ measurements performed on a SAM (the configuration used in this work). However, since an 'all-data' analysis approach was used, and the two conductance levels are close together, the presence of S-Au binding traces in the latter measurement may not have been detected. Hence the clustering result presented here may provide evidence that both the S-Au and SH-Au binding configurations can occur for STMBJ measurements on a 'dry' SAM.

In summary, a feature extraction and clustering process has been developed and applied to an experimental dataset of G-Z traces. While the all-data analysis was only able to extract a single average conductance value, the clustered data has revealed several different conductance levels, consistent with observations of Au-alkanedithiol-Au junctions in the literature.

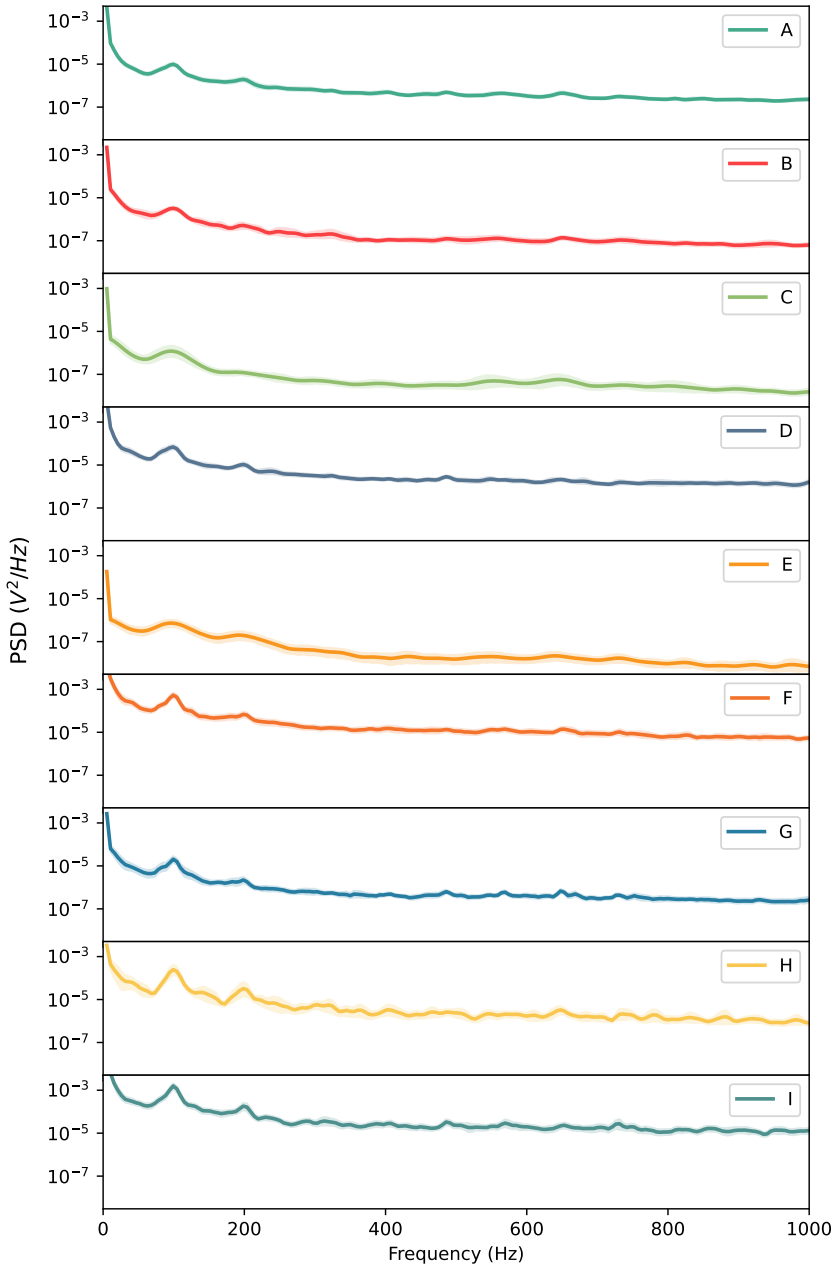


Fig. 5.27: Average PSD plots for each of the 9 clusters identified for ODT (signal converted from conductance to raw DAQ input voltage)

6. THERMOELECTRIC CHARACTERISATION OF MOLECULAR JUNCTIONS BY ATOMIC FORCE MICROSCOPY

Atomic force microscopy (AFM) is a powerful tool for studying multi-molecule junctions (MMJs). Like STM, which is discussed in chapter 4, the AFM is a type of scanning probe microscope, in which a nanoscale probe moves across the sample surface and measures some interaction between the probe and sample in order to obtain a topographic map of the sample surface. While STM measures the tunnelling current between an ultra-sharp tip and the sample, and is therefore typically restricted to the imaging of conductive surfaces [169], the probe in AFM consists of a microfabricated cantilever with a sharp protrusion at one end, the deflection of which is used to measure a distance-dependent force acting between the probe apex and sample surface. The net tip-sample force is a superposition of a number of interactions, including van der Waals (due to noncovalent chemical interactions between atoms/molecules in the tip and sample) and electrostatic forces (due to variations in surface potential); the resulting relationship between force and tip-sample separation is complex, as different contributions may dominate at different ranges [6, 169, 170].

An ideal companion to STM-based single-molecule junction experiments, AFM offers a larger contact area (typically in the range 1-1000nm^2 [41]) and controllable contact force. Since its introduction in 1986 by Binnig, Quate and Gerber [20], AFM has been extended far beyond simple topographic imaging, with numerous advanced AFM modes allowing the measurement of various properties of samples including electrical and thermal conductivity. Of particular interest for the study of thermoelectricity in MMJs is ThAFM (thermovoltage AFM), a technique introduced in 2010 by Reddy et al. [3], which facilitates the measurement of the Seebeck coefficient of a MMJ by contact-mode CAFM. The Seebeck coefficient, unlike conductance, is an intrinsic property independent of the number of molecules in parallel in the junction; hence thermoelectric measurements on MMJs can be compared to equivalent measurements by other methods on single-molecule junctions, potentially elucidating important differences between the behaviour of molecules in ordered films/monolayers and their isolated single-molecule counterparts.

ThAFM is a variant of CAFM, which uses a conductive probe and sample stage to provide electrical connections to both sides of the tip-sample junction. Unlike

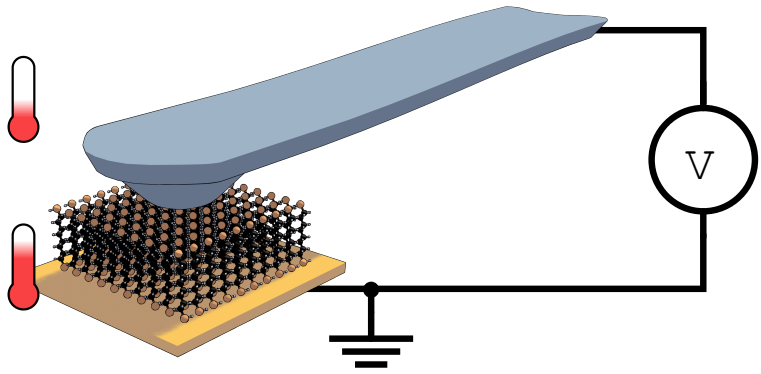


Fig. 6.1: An illustration showing the working principle of ThAFM.

CAFM, in which a bias is applied to the sample stage and the resulting current through the junction is measured by a transimpedance amplifier connected to the probe, in ThAFM the sample is grounded and a differential voltage amplifier is connected across the junction (i.e. between the probe and the sample) to measure the open-circuit thermovoltage. The sample stage temperature is then varied to measure the variation in thermovoltage as a function of the temperature difference across the junction; from this the Seebeck coefficient of the junction can be inferred [3].

The ThAFM technique is much less widespread than CAFM, and no turnkey solutions exist for commercial AFM instruments: hence the necessary modifications must be implemented in-house by individual labs, and a range of experimental setups with varying levels of performance have been developed by different research groups.

This chapter details the development of a bespoke ThAFM module for use with a commercially available Bruker multimode 8 AFM, intended to improve upon existing setups in a number of key areas including ease-of-use, reliability, measurement stability and noise rejection. Additionally, the module is inexpensive and easy to reproduce with off-the-shelf components: hence the setup detailed here could easily be adopted by other labs, potentially helping to standardise the measurement and simplify comparisons of results generated by different groups. In section 6.1 the main experimental challenges limiting the performance of ThAFM measurements are described. In section 6.2 the design, development and calibration of the apparatus are discussed, and finally the full experimental setup is tested on a number of widely-studied materials in section 6.3.

6.1 Experimental Challenges

Although ThAFM is similar to CAFM in many ways, it presents a number of unique experimental challenges. These are discussed in this section, and strategies to mitigate them are outlined.

One of the most important differences in instrumentation between CAFM and

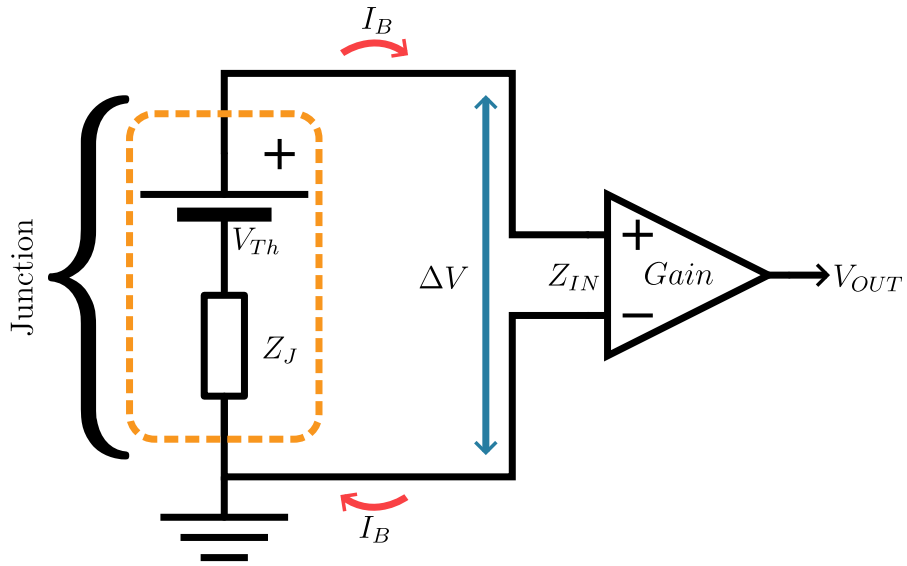


Fig. 6.2: A simplified circuit diagram of a thermovoltage measurement.

ThAFM is that CAFM is a current measurement while ThAFM is a voltage measurement. The molecular junction typically presents a high impedance, due to the nanoscale contact area of the junction and the low conductance of most organic molecules. For a high source impedance, voltage measurements are typically more challenging than current measurements. Any non-ideal operational amplifier has a finite input impedance which forms a potential divider with the source impedance, resulting in a reduction in the measured voltage. For an amplifier with input impedance Z_{IN}, the voltage V_m measured from a junction with impedance Z_J generating thermovoltage V_{th} is given by:

$$V_m = V_{th}\frac{Z_{IN}}{Z_{IN} + Z_J} \tag{6.1}$$

Figure 6.3 illustrates the consequences of increasing source impedance on the measured voltage for different amplifier input impedances. Clearly an amplifier with an input impedance much greater than the impedance of the junction must be used if the thermovoltage (and therefore Seebeck coefficient) is to be measured accurately.

Another important challenge when measuring voltage from a high impedance source is the offset voltage V_{OS} arising from the input bias current I_B of the amplifier flowing through the source impedance:

$$V_{OS} = I_B Z_J \tag{6.2}$$

This offset voltage is added to the measured thermovoltage, and since it is a function of the junction impedance it varies with stochastic fluctuations in the con-

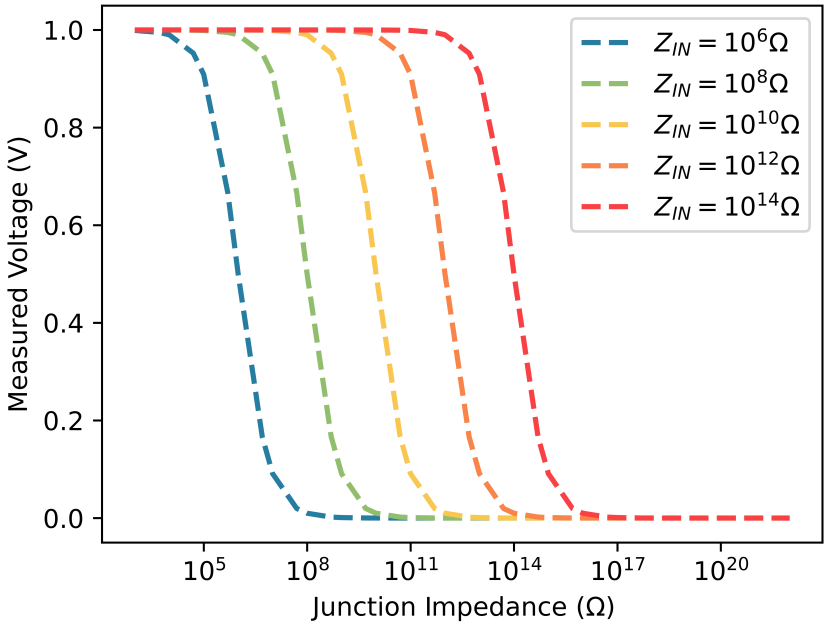


Fig. 6.3: An illustration of the effect of increasing source impedance on the measurement of a 1 V DC signal for a given amplifier input impedance Z_{IN}.

ductance caused by various factors including tip condition, junction geometry, variations in the number of molecules in the junction and the presence of water or other mobile contamination on the surface [171]. These variations in the junction conductance have a lesser effect on CAFM measurements because of the shorter time scale of the measurement (so enough measurements can be obtained that the conductance variations form part of a normal distribution around the most probable conductance), and additionally CAFM measurements can employ bias modulation and measure the resulting AC signal with a lock-in amplifier, eliminating the effect of DC offsets. The ThAFM technique is a DC measurement and operates over a longer time scale since the thermovoltage must be measured for multiple different probe-sample temperature gradients and the temperature cannot be modulated as quickly as the bias in AC CAFM measurements due to thermal inertia. Hence the offset voltages caused by the input bias current of the amplifier can have a significant effect on the stability of ThAFM measurements.

6.2 Thermovoltage Application Module Development

As part of this work a custom expansion module has been developed to extend the capabilities of a commercial AFM to include ThAFM measurements on molecular junctions. Every aspect of the system, from the sensitive analogue electronics to the physical integration with the AFM, has been carefully considered and optimised considering the key experimental challenges identified in section 6.1. The resulting

module is compact, simple to set-up and use, and offers performance competitive with the best existing ThAFM implementations reported in the literature. This section details the development and testing of the system.

6.2.1 Thermovoltage Amplifier

The thermovoltages measured in a typical ThAFM experiment are on the order of 10 μV, and amplification is required in order to attain an acceptable signal-to-noise ratio. As discussed in section 6.1, careful consideration must be given to the details of the amplifier circuit in order to attain optimum performance.

The first ThAFM setup introduced by Reddy et al. utilised a custom-built thermovoltage amplifier [3] based on a design introduced for STM-based 'scanning thermopower microscopy' measurements by Poler et al. in 1995 [2]. This amplifier was comprised of a buffer stage made from two DiFET operational amplifiers (Burr-Brown OPA128) configured as unity-gain voltage followers, followed by a simple differential amplifier stage with a voltage gain of 48 dB. The (no longer manufactured) Burr-Brown OPA128 op-amp was used for the input buffer due to its ultra-low input bias current and low offset drift. The amplifier design also included a driven guard to protect the sensitive buffer inputs from excess leakage current. A simple schematic of this amplifier design is provided in figure 6.4. Here U_1 and U_2 make up the unity-gain input buffer and U_3 is configured as a differential amplifier with voltage gain R_2/R_1. U_4 is an additional unity-gain buffer used to drive a guard which surrounds the probe input and protects it from leakage current.

Since the introduction of the technique in 2010, a number of research groups have adopted ThAFM, often using a commercially available voltage amplifier. The most commonly used amplifier for ThAFM is the Stanford Instruments SR560 low noise preamplifier in a differential configuration. The SR560 uses a more sophisticated signal path, consisting of a differential input buffer followed by two amplifier stages interspersed with filters, as illustrated in figure 6.5. Many of the parameters of the amplifier, such as total voltage gain, filter cutoffs and input configuration are reconfigurable by the end user. The input stage is of most interest for this work, as it largely defines the input bias current, impedance and offset drift of the amplifier. Rather than using operational amplifier integrated circuits, the SR560 uses a discrete differential amplifier constructed from individual matched JFET pairs (National Semiconductor NPD5564) in its input stage [172]. These FETs have a gate leakage current of around 3 pA, which defines the lower bound of the input bias current for the SR560 amplifier. This is an order of magnitude greater than the 100 fA input bias current claimed by Poler et al. for their custom built amplifier [2]. In addition, amplifiers built from discrete components typically have higher input offset voltage and drift than the equivalent integrated circuit, due to difficulty in precisely

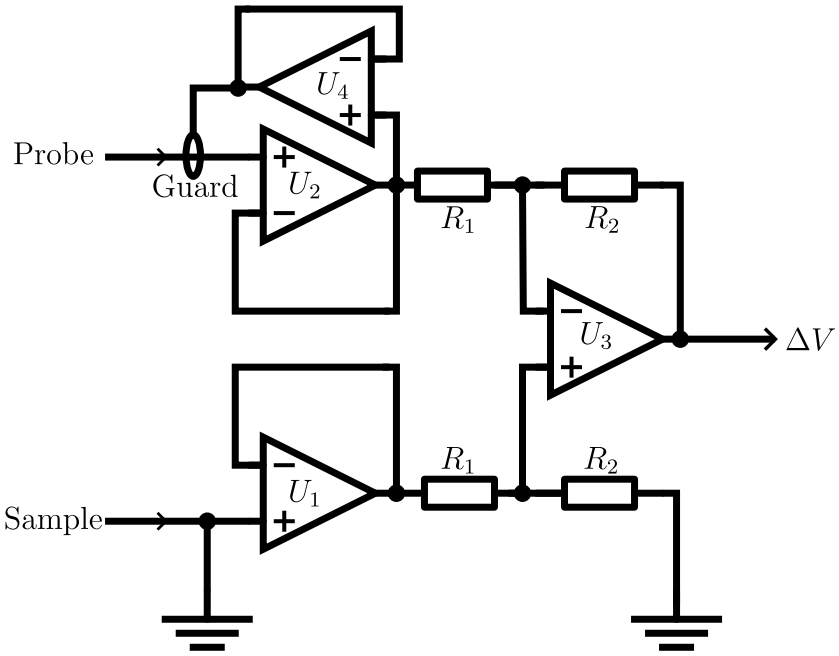


Fig. 6.4: A schematic diagram of the thermovoltage amplifier introduced by Poler et al. in 1995 [2] and subsequently used by Reddy et al. in the first implementation of ThAFM [3].

matching the values of discrete components and greater thermal gradients over the larger circuit footprint, among other reasons. As a result, the SR560 requires manual offset nulling, while the OPA128 is specified for a maximum offset of 500 μV owing to laser-trimming of the thin-film resistors in the IC. The SR560 input stage is likely designed to optimise noise and bandwidth performance for higher frequency measurements, making it a sub-optimal choice for DC thermovoltage measurements.

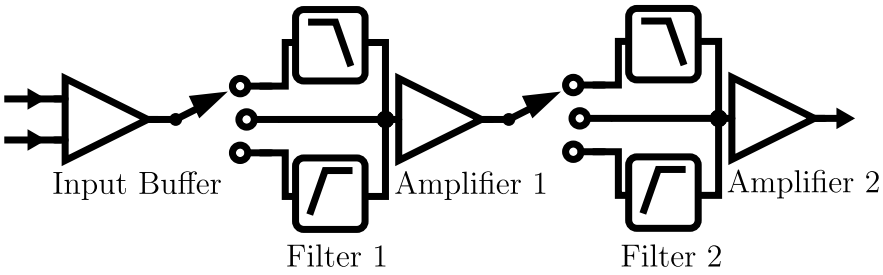


Fig. 6.5: A simplified representation of the signal path of the Stanford Instruments SR560 low-noise voltage preamplifier.

For the ThAFM implementation developed in this work, an instrumentation amplifier topology was selected for the thermovoltage amplifier: the high input impedance, low drift and low input bias current offered by instrumentation amplifiers makes them ideal for this application. The instrumentation amplifier is a differential voltage amplifier, and its topology is similar to the amplifier used by Poler et al. [2] in that it is based on three operational amplifiers, configured as an input buffer followed

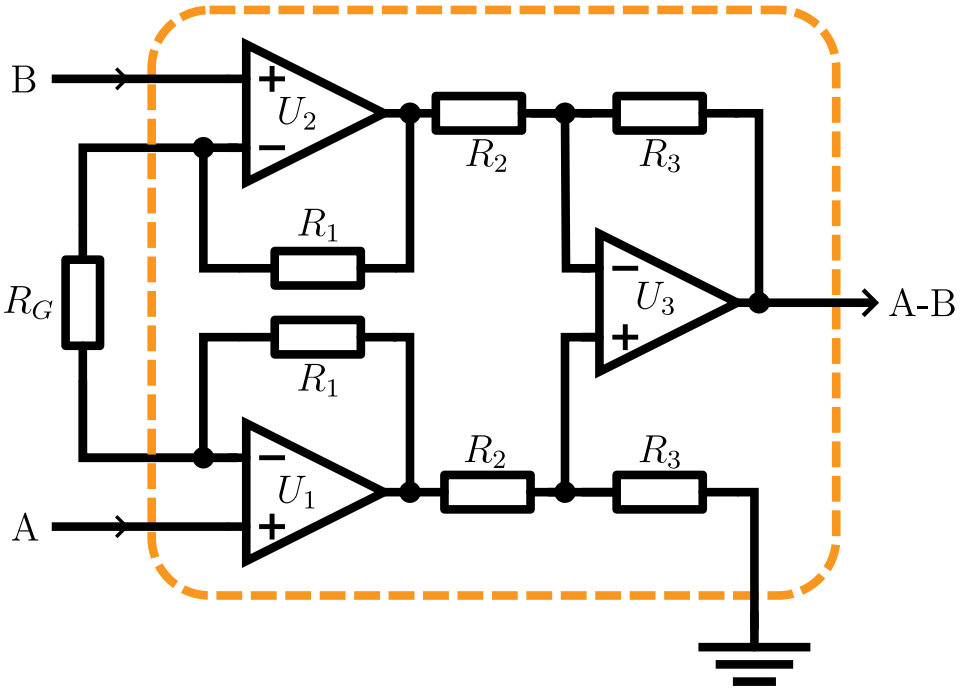


Fig. 6.6: A schematic diagram showing the topology of a typical instrumentation amplifier. The circuit elements inside the dotted line are usually contained within the IC, eliminating the need for precise matching of external components.

by a differential amplifier stage (see figure 6.4); the instrumentation amplifier differs in the addition of a resistor, R_G, between the inverting inputs of the two operational amplifiers that comprise the buffer stage, as well as matched resistors in the feedback paths of the buffer amplifiers. This configuration is illustrated in figure 6.6. This design offers a number of benefits over the simple buffered differential amplifier: the input buffer now has a differential gain factor of $1 + 2R_1/R_G$ but maintains unity gain for common-mode signals. The differential amplifier stage (U_3) applies additional gain R_3/R_2 to differential signals and attenuates common-mode signals. This means that the overall differential gain of the amplifier can be increased by reducing R_G without changing the common-mode gain, resulting in a common-mode rejection ratio (CMRR) which increases with overall gain [173]. This is a very desirable characteristic as it allows high differential gain without signal-to-noise ratio degradation or amplifier saturation caused by common-mode noise and offsets present on both inputs. This design can also achieve very low offset voltage and drift, especially when implemented as an integrated circuit (IC); the matched resistor pairs are typically included in the IC and matched to a very high tolerance by laser-trimming, while the single resistor R_G is located externally and can be used to change the gain without introducing excessive offsets and drift due to resistor mismatch and thermal variations.

For the sake of simplicity, and to ensure reproducibility between experiments, a

fixed-gain configuration was chosen for the thermovoltage amplifier. The gain was selected to give the best compromise between dynamic range and signal-to-noise ratio; the gain must be large enough to accurately measure the thermovoltage, but if the gain is too large the amplifier may saturate when measuring samples with high Seebeck coefficients. The ADC converters of the Bruker Nanoscope 6 AFM controller, used in this work to digitise the thermovoltage signal, offer 16 bit code resolution over their 20V input range, resulting in an LSB voltage of 305 μV; in order to detect a 10 μV change in thermovoltage the amplifier must have a voltage gain of at least 30 dB. However, testing showed that the actual effective resolution of the analogue input is limited by the input referred noise rather than the code resolution, and so a higher gain is necessary. In this work a voltage gain of 54 dB was selected in order to ensure that the signal level would be above the noise floor of the ADC.

The specific instrumentation amplifier IC used in the thermovoltage amplifier was carefully selected. From a wide number of options, three ICs were selected for testing by comparing several important specifications (sourced from the manufacturer's datasheets):

1. Input bias current
2. Input impedance
3. Offset drift
4. Cost

The specifications of the two established thermovoltage amplifiers previously discussed were used as a basis for comparison. In the cases where specifications were not available (for example the input bias current of the SR560), these parameters were estimated based on the specifications of the components used in the front end. The amplifier ICs selected for further testing were the INA821 (Texas Instruments), AD8230 (Analog Devices) and INA116 (Texas Instruments). The bar plots in figure 6.7 summarise the key specifications of these ICs compared to the SR560 and the custom amplifier presented by Poler et al. [2, 172, 174–176]

All of these amplifier ICs are instrumentation amplifiers, however each is implemented using a different process (bipolar, CMOS and JFET for the INA821, AD8230 and INA116 respectively), resulting in some important performance differences in key areas [177]:

The INA821 uses a fairly traditional instrumentation amplifier topology based on bipolar transistors [175]. However, while a typical bipolar amplifier might be expected to have a much larger input bias current compared to FET or CMOS

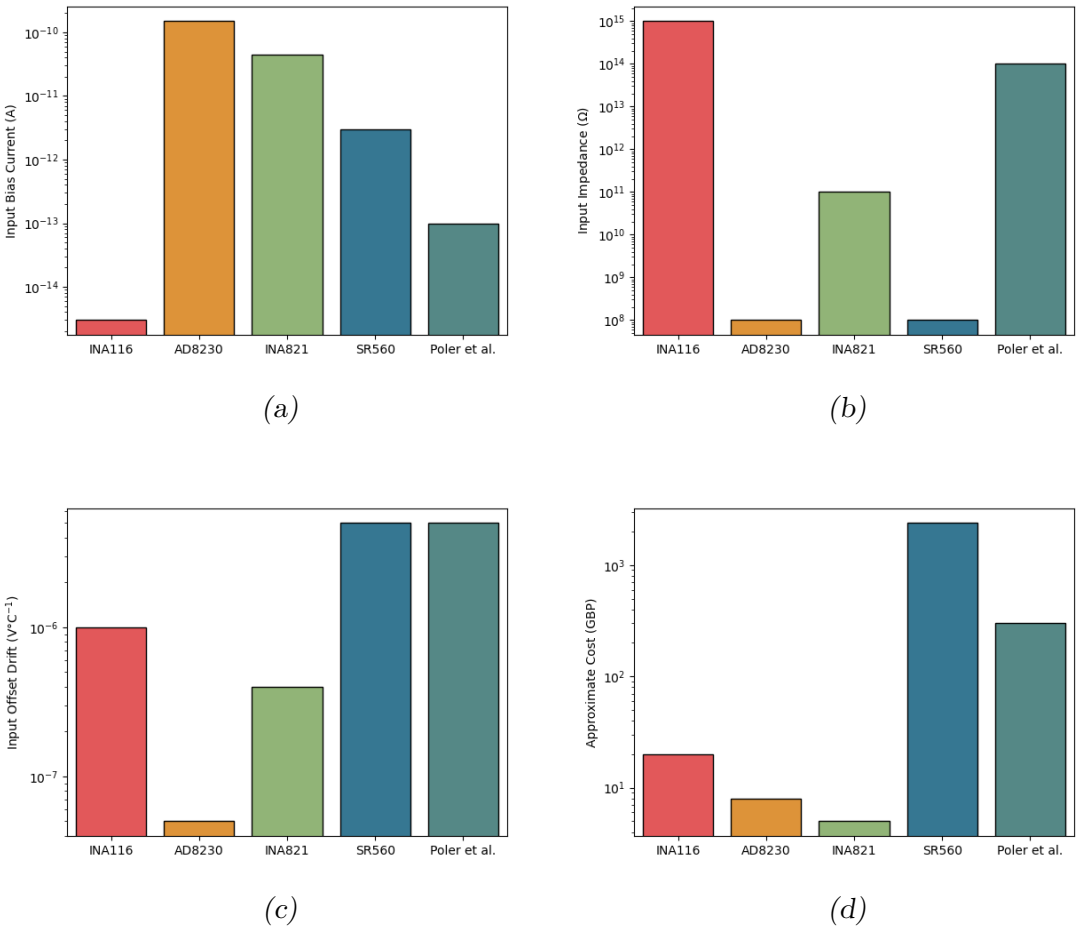


Fig. 6.7: Comparisons of important specifications for each of the amplifiers considered in this work: **(a)** Input bias current **(b)** Input impedance **(c)** Input offset voltage drift **(d)** Cost (estimated from bill of materials for the lab-built amplifiers, and quoted from the manufacturer in the case of the SR560)

technologies [178], the INA821 utilises so-called 'super-beta' transistors which have significantly greater current gain than typical transistors; this results in a lower input bias current, comparable to some FET input stages, while retaining the low offset drift and $1/f$ noise associated with bipolar amplifier technologies [175, 179].

The AD8230 is a CMOS process amplifier; typically CMOS-based amplifiers exhibit higher input impedance and lower input bias current than bipolar amplifiers, at the cost of higher $1/f$ noise, offset and drift [177]. However, the CMOS process allows the implementation of an 'auto-zeroing' topology which reduces the offset drift (and by extension $1/f$ noise) of the AD8230 by more than an order of magnitude compared to any of the other amplifiers considered in this work. This topology is similar in application to the standard instrumentation amplifier in figure 6.6, but works in a very different way: the input stage of the AD8230 consists of a capacitive differential sample-and-hold stage, driven by an internal 6 kHz clock, in

place of the dual op-amp buffer in a traditional op-amp. This stage provides the high common mode rejection of the amplifier while minimizing drift. The second stage is a more conventional differential amplifier which provides gain to the signal. Both stages incorporate active auto-zeroing to minimize offset drift and $1/f$ noise. The main disadvantage of this topology for a ThAFM application is the relatively large, time varying input bias current due to the current needed to charge the sampling capacitors during the short sampling period. [176]

The INA116 has many important internal similarities to the amplifier created by Poler et al. [2] (see figure 6.4): in particular, it uses a dielectrically-isolated FET (DiFET) front end and includes driven guard pins to minimise input bias current and leakage. However, the INA116 benefits from being an integrated monolithic instrumentation amplifier rather than a buffered differential amplifier built from discrete op-amps. As a result, the INA116 offers a lower input bias current (3 fA compared to the 100 fA reported by Poler et al.), higher input impedance and lower offset drift [2,174]. As is typical of a FET-based amplifier, the INA116 has somewhat higher offset drift than the transistor-based INA821 and the auto-zeroing AD8230, but significantly lower input bias current and higher input impedance than both. The INA116 also outperforms the SR560 in all the metrics considered.

Prototype amplifiers were constructed using each of these ICs so that their performance could be compared in a real-world setting. A separate PCB was built for each amplifier, with the same mounting hole configuration so that they could be installed interchangeably within the shielded enclosure (described in section 6.2.3). Each amplifier was configured with a differential gain of 54 dB.

The noise and drift of each amplifier circuit was characterised by connecting both inputs to ground and recording the output signal for 80 minutes at a sample rate of 820 Hz. Power spectral density periodograms, shown in figure 6.8, were plotted from the resulting data and used to compare the noise and drift (1/f noise) performance of each system. The SR560 previously used in the lab for thermovoltage measurements was also included as a basis for comparison. Since ThAFM consists of a DC measurement, the low frequency 1/f noise is the most important region of the PSD to consider. The AD8230 is unique among the amplifiers tested in that it exhibits no observable 1/f noise, due to the auto-zero topology. However, while the INA821 does show a small 1/f noise trend, the actual noise power at low frequency remains lower than that of the AD8230. In fact, the INA821 offers the lowest noise across the frequency range tested of all the amplifiers tested here. The INA116 offers comparable noise to the INA821 at higher frequencies, but ultimately exhibits the highest 1/f noise of the three amplifier ICs. This is characteristic of a FET-input amplifier, and consistent with the specified drift in the INA116 datasheet [174] (see figure 6.7c). Nonetheless, the INA116 demonstrated two orders of magnitude lower

1/f noise than the SR560 during the test.

After consideration of the parameters identified in figure 6.7 and the measured noise performance in figure 6.8, the INA116 was selected for use in the ThAFM module. The INA116 exhibits slightly higher 1/f noise than the INA821 and AD8230, but vastly outperforms both in other key areas (input bias current and input impedance). The INA116 also exceeds the specifications of even the best amplifiers (such as that introduced by Poler et al. [2]) previously used for ThAFM in all metrics considered.

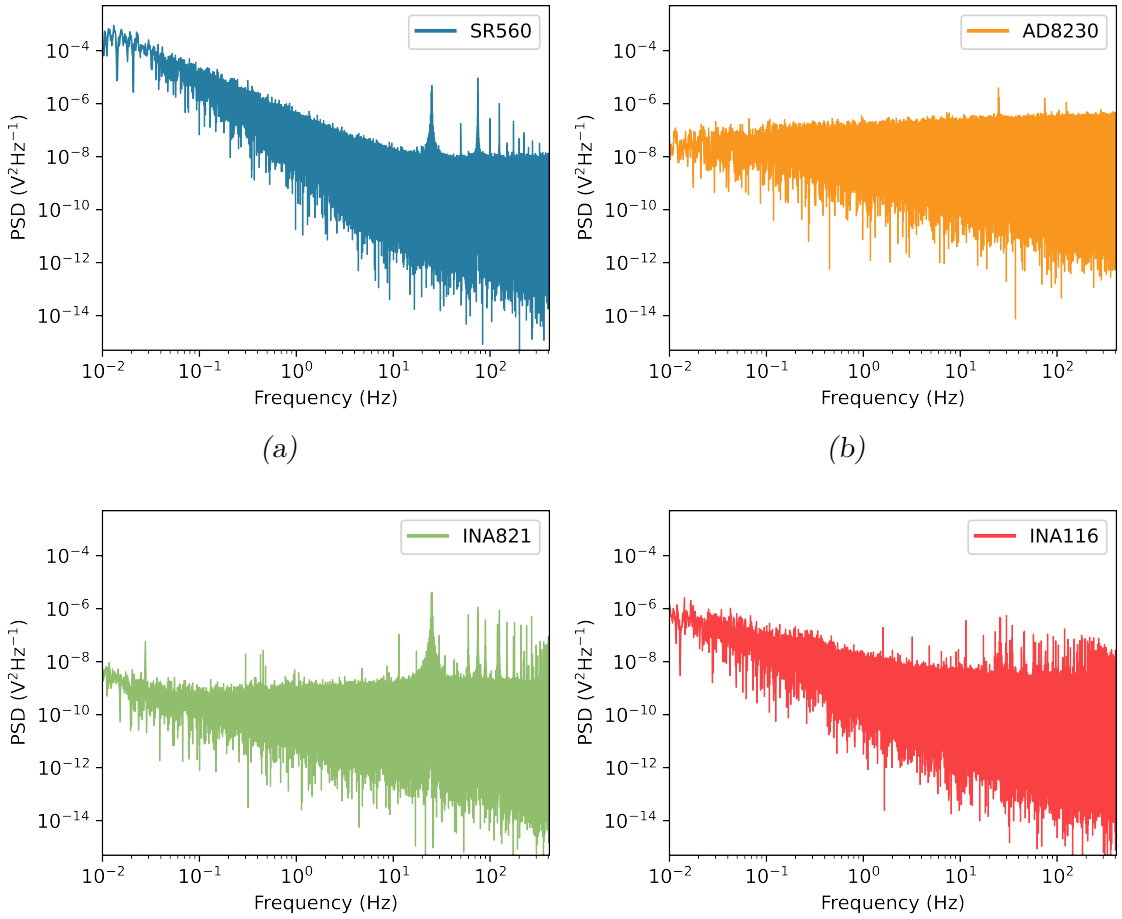


Fig. 6.8: Power spectral density plots of the output offset voltage between 0.01 Hz and 400 Hz for the amplifiers tested. Test was conducted over 80 minutes with inputs tied together to amplifier ground: **(a)** SR560 **(b)** AD8230 **(c)** INA821 **(d)** INA116

6.2.2 Low Noise Dual-Rail Power Supply

The thermovoltage amplifier (described in section 6.2.1) developed in this work operates on a dual rail power supply from ±4.5 V to ±18 V. Such power supplies can be readily obtained or constructed, but vary significantly in their output noise and stability [178]. A stable, low noise power supply is desirable for precision applications since noise in the power supply rails of the amplifier will appear at the output,

albeit attenuated according to the amplifier's power supply rejection ratio (PSRR). PSRR varies with gain and frequency: at 54 dB gain and $\leq$ 100 Hz frequency, the INA116 used in the thermovoltage amplifier specifies a PSRR of around 110 dB for the positive supply and 90 dB for the negative supply. This is sufficient to reduce the noise from typical switching power supplies, with ripple voltages on the order of a few 100 mV, below the noise floor of the measurement. However, in case of external factors limiting the PSRR, and in the interest of ensuring the best possible performance from the amplifier, a lower noise linear power supply was selected.

A custom $\pm$5 V PSU was constructed, based on an open frame linear power supply module (PSU 303, Lascar) with low ripple noise specified as $\leq$5 mV. Combined with the high PSRR of the amplifier, this should contribute ≤ 1 μV of noise to the output signal. The power supply module takes a mains AC input, and uses a step-down transformer and bridge rectifier with smoothing capacitors to produce positive and negative DC voltage rails. A pair of linear regulator ICs are then used to reject ripple noise and create stable $\pm$5 V outputs. A schematic of this configuration is provided in figure 6.9.

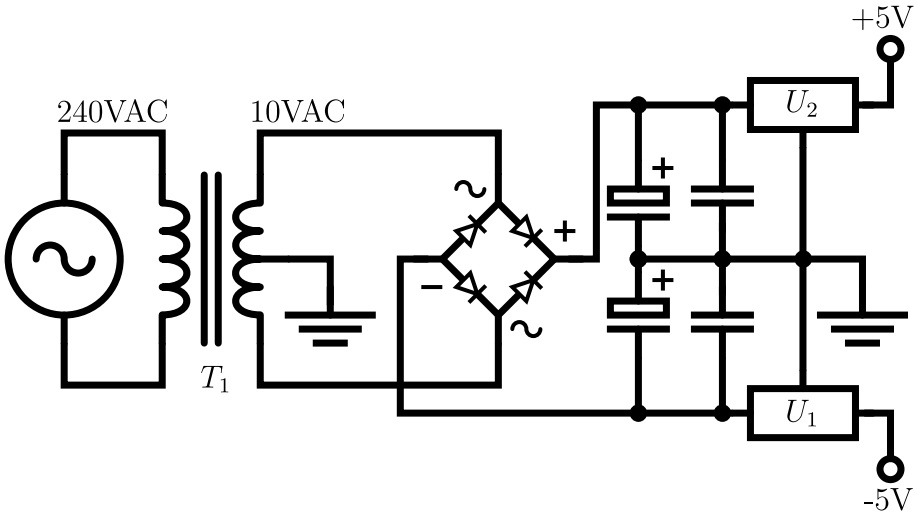
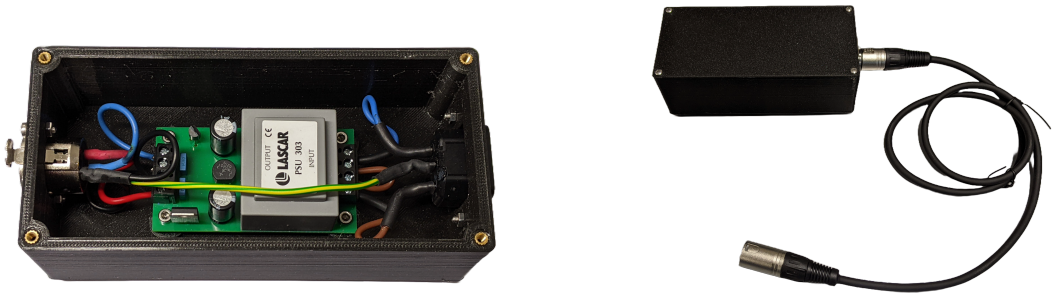


Fig. 6.9: A schematic showing the workings of the dual-rail low-noise PSU used to power the thermovoltage amplifier. From left to right is the mains AC input, followed by a step-down transformer T_1, a bridge rectifier with smoothing capacitors to provide dual DC rails, and finally a pair of linear regulator ICs (U_1 and U_2) to create the low-noise $\pm$5 V outputs.

A custom 3D-printed enclosure was created to house the PSU module and provide safe, convenient electrical connectors in place of the exposed screw terminals on the module itself. A standard 3-pin IEC connector was used for the mains input, and a 3-pin XLR connector was used for the output. For safety reasons all external metal surfaces (such as the XLR connector body) were connected to the earth pin of the IEC connector. Threaded heat-set inserts were used to secure the lid of the enclosure,

allowing for easy maintenance and inspection while remaining secure during normal use. The enclosure was also outfitted with self-adhesive polyurethane rubber 'feet' (3M SJ5302) for vibration damping. Photographs of the assembled PSU can be found in figure 6.10.



(a) *(b)*

Fig. 6.10: Photographs of the dual-rail power supply assembled for the ThAFM module **a** lid removed **b** lid closed

During testing, standard XLR cables were found to transfer excessive vibrations through the thermovoltage amplifier (mounted to the AFM head) into the junction. To reduce vibration coupling a more flexible, lightweight XLR cable was made using three separate silicone insulated wires with no outer jacket. This cable is shown in figure 6.11.

Fig. 6.11: A photograph of the custom flexible XLR cable made for the ThAFM PSU to minimise transmission of vibrations to the AFM junction

6.2.3 Implementation

The physical implementation of the ThAFM setup within the AFM system requires careful consideration: the performance of even a meticulously designed amplifier can

be severely degraded when installed in a suboptimal configuration.

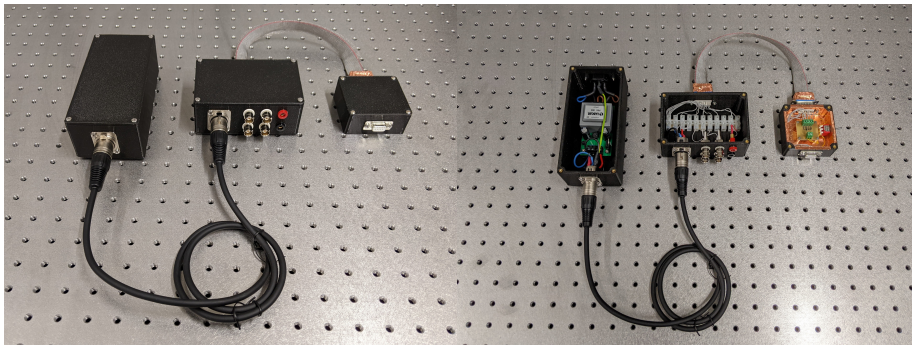
The physical layout of the ThAFM setup can have a significant effect on electrical noise pickup, mechanical stability, and leakage current. Some key issues and corresponding design guidelines have been identified, and are as follows:

- Long cables connecting the thermovoltage amplifier to the junction can introduce significant noise and electrical instability via EMI pickup and microphonics. Hence the amplifier should be located as close as possible to the junction in order to minimise the length of these cables. Where long cable runs are necessary, a balanced twisted pair configuration could be considered to improve noise rejection, however this not ideal due to leakage current considerations (see below).

- Any cables and electrical connections to the junction can introduce mechanical instability. All cables should be flexible and lightweight. Longer cables, such as the power supply cable for the amplifier, can be clamped down to the table or some other heavy object to effectively decouple different sections of the cable and reduce transmission of vibrations.

- In order to maintain the extreme low input bias current of the thermovoltage amplifier, the conductive path from the probe to the input must be protected from leakage currents arising from conduction pathways to other nearby conductors at different potentials (the sample connection is less critical since it is connected to ground directly). For example, if there is potential difference of 100 mV between the input node and a nearby conductor, the resistance between the two must exceed 10 TΩ to maintain a leakage current below 5 fA. The amplifier itself utilises careful layout and a driven guard trace around the inputs to effectively mitigate leakage, but the connection to the probe remains exceptionally vulnerable. A very short cable connected directly to the probe with at least 5mm air gap on all sides is preferred; if this is not possible a triaxial cable with the inner shield connected to the driven guard can be used, at the expense of reduced flexibility. The probe holder itself can also be a source of leakage. The Bruker MMCHAM TUNA-compatible probe holder is made from a non-conductive material and designed to isolate the probe from all other conductors to reduce leakage.

The development of the ThAFM module was an iterative process, and a number of different configurations were prototyped and evaluated before arriving at the final design; the use of FDM 3D printing to fabricate the bespoke enclosures and fittings facilitated inexpensive rapid prototyping, allowing different permutations to



(a) (b)

Fig. 6.12: Photographs of an early prototype of the ThAFM system. **(a)** The electronic modules which comprise the system, prior to installation on the AFM. **(b)** The modules with the enclosure lids removed to reveal the internal electronics and shielding.

be explored easily. An example of an earlier prototype of the ThAFM system can be seen in figure 6.12.

Of the various configurations tested, an 'application module' format was identified as providing the best balance of performance, ease-of-use and versatility: The Bruker Multimode 8 AFM head features a mounting shelf for turnkey application modules, available from Bruker, which can extend the capabilities of the instrument (for example the 'TUNA' module enables CAFM measurements). A custom 3D-printed enclosure, designed to be compatible with the application module mounting shelf, was created to house the thermovoltage amplifier. The module features an XLR connector for power input from the linear PSU (described in section 6.2.2), a BNC connector for the thermovoltage amplifier output, and a power indicator LED. The interior of the 3D-printed enclosure is coated with conductive copper tape, connected to the amplifier ground, to provide shielding and reduce noise pickup.

The probe is mounted in a Bruker MMCHAM conductive holder, which is made of a non-conductive material to minimize leakage and allows electrical connection between the probe and application modules via a Au-plated pin socket. The sample is mounted in a custom sample holder (illustrated in figure 6.13) which isolates the sample electrically from the built-in sample bias of the AFM using a Mylar insulating layer (chosen to minimise leakage). The sample is secured by a copper S-clip, which also provides the electrical contact to the top of the sample and allows direct connection to the ThAFM module.

Two short (<3 cm), flexible silicone-insulated wires connect the probe and sample to the inputs of the thermovoltage amplifier. Care was taken when routing these wires to minimize leakage: for example, the grounded copper tape used for shielding inside the module is omitted around the area where the wires enter the enclosure. Additionally, these input wires are directly routed to the input pads of the amplifier,

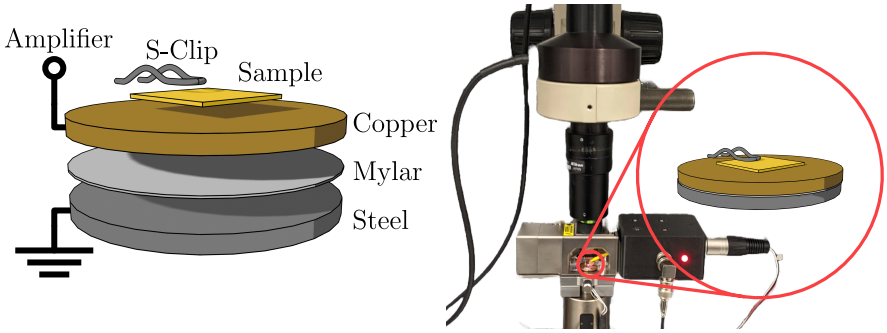


Fig. 6.13: Illustration of the sample stage used for ThAFM measurements **(a)** exploded view **(b)** assembled view (inset)

Fig. 6.14: Photographs of the finished ThAFM module **(a)** Alongside a Bruker TUNA CAFM application module for comparison **(b)** In-situ on the applications-ready multimode AFM head

inside the driven guard traces, without touching any other part of the amplifier PCB to further reduce leakage.

Photographs of the assembled ThAFM application module can be found in figure 6.14.

6.2.4 Temperature Control

In order to establish a temperature gradient across the tip-sample junction, the sample plate is heated while the probe apex is assumed to remain at ambient temperature, due to the high thermal conductivity of the Si cantilever substrate compared to the surrounding air and the large thermal mass of the probe holder and AFM head. This assumption has been shown to hold to within 5% error by Reddy et al. [3] using a mathematical model validated by a previous experimental study of heat transfer in nanoscale contacts [180].

Bruker provides a sample heater compatible with the multimode AFM used in this work, consisting of a resistively heated sample stage and a corresponding

'thermal applications controller' (TAC) which uses feedback control to maintain a constant temperature setpoint.

Heated Sample Stage Calibration

The TAC measures the internal temperature of the heated sample stage to provide a feedback signal and control the sample temperature. Since the sample is mounted in a separate sample holder which sits atop the heated stage, the actual temperature at the surface of the sample differs from the temperature setpoint of the TAC and must be measured using a separate temperature sensor located at the sample surface. This is not feasible during ThAFM measurements, due to the additional current leakage and mechanical noise introduced by positioning a temperature sensor close to the junction; instead, a calibration curve was measured such that the true sample temperature could be calculated from the TAC setpoint.

A precision bipolar junction temperature sensor IC (LMT84, Texas Instruments) was mounted to the top surface of a piece of template-stripped Au on Si (the same substrate used for the samples measured in this work, prepared as described in section 3.2.1) using a thin layer of cyanoacrylate adhesive. This was then installed in the custom sample holder, described in section 6.2.3, which was affixed magnetically to the heated sample stage. The output of the temperature sensor was recorded using one of the analogue input channels of the AFM controller while the temperature setpoint of the TAC was varied in 5 K increments over a 30 K range. At each temperature setpoint, the sample temperature was allowed to stabilise before recording 32,768 temperature readings over a 256 second period (twice the duration of a typical thermovoltage measurement).

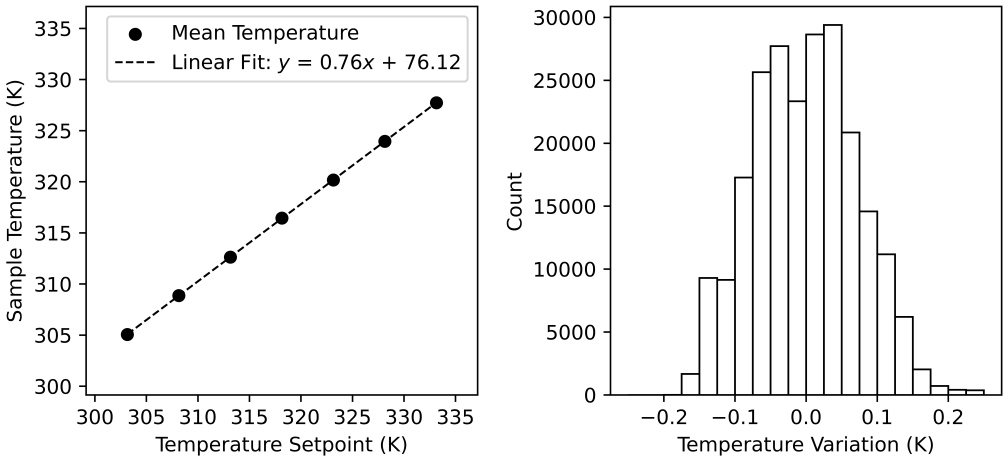
In order to determine the relationship between the TAC setpoint and the actual sample temperature, the mean value of the measured temperature was plotted as a function of setpoint (see figure 6.15a). A linear relationship was observed between setpoint temperature T_{Setpoint} and sample temperature T_{Sample} across the operating range of the TAC; a linear fit yielded the following calibration function (in Kelvin):

$$T_{\text{Sample}} = 0.76 T_{\text{Setpoint}} + 76.12 \tag{6.3}$$

Equation 6.3 can be used to determine the sample temperature for a given TAC setpoint within the operating range of the TAC.

The uncertainty in sample temperature was estimated from the distributions of measured temperatures across all setpoints: for each setpoint, the mean of the temperature measurements was subtracted from all the measurements to align the distribution around 0. All the measurements from all setpoints, after scaling, were then collated into a single histogram, presented in figure 6.3b. Assuming a normal

distribution, the uncertainty in the sample temperature was then estimated from the standard deviation as ±0.1 K.



(a) *(b)*

Fig. 6.15: **(a)** A plot of sample temperature against TAC setpoint temperature over 7 setpoint values separated by 5 K increments. Each point corresponds to the mean of 32,768 temperature readings over a 256 second period. A linear fit applied to the data gives a calibration function which can be used to estimate the sample temperature for a given TAC setpoint. **(b)** A histogram of all 229,376 individual temperature measurements, re-centred around 0 by subtracting the mean temperature from all measurements at each TAC setpoint. The standard deviation of the resulting distribution is ±0.1 K.

6.2.5 Data Processing

A typical thermopower measurement consists of measuring several thousand thermovoltage samples at four different temperature setpoints (specific details are provided in section 6.3). A least-squares fitting algorithm (implemented in SciPy [158]) is then applied to the resulting thermovoltage distributions to estimate the temperature dependence of the thermovoltage; the Seebeck coefficient is extracted from the gradient of the fit.

The heated sample stage takes several minutes to reach a stable temperature at each different setpoint, limiting the speed at which a measurement of the Seebeck coefficient can be completed (since the thermovoltage must be measured at several temperature setpoints). Obtaining enough complete measurements to obtain a robust estimate of the uncertainty from their standard deviation is therefore infeasible; instead, the uncertainty in each measurement is obtained from the covariance matrix of the least-squares linear fit, which takes into account the standard deviation of the distribution of thermovoltages measured at each temperature setpoint. Upon calculating the average Seebeck coefficient from multiple repeat measurements, the

final uncertainty is obtained by summing the uncertainty of each measurement in quadrature.

6.3 Testing

In order to calibrate the system and evaluate its performance, a series of test measurements using samples with known Seebeck coefficients was performed. All measurements were carried out with a Pt coated AFM probe (BudgetSensors ElectriMulti75-G) in contact mode with zero scan size (point-mode). Each complete Seebeck measurement consisted of four temperatures, with 16,384 thermovoltage samples recorded over 128 seconds at each temperature.

6.3.1 Au

As a control measurement, and to calibrate the offset due to the contribution of stray thermovoltages (for example a thermovoltage arises from the copper-solder joint at the sample stage when heated), the Seebeck coefficient of a bare Au(111) substrate (prepared by template stripping, as outlined in section 3.2.1) was measured. The result of this measurement is presented in figure 6.16. A small positive Seebeck coefficient of $S = 1.38 \pm 0.02\ \mu VK^{-1}$ was obtained for the Au-Pt junction. This value is subtracted from all future measurements in order to obtain the true Seebeck coefficient of each sample relative to a plain Au-Pt junction, accounting for constant measurement offsets inherent to the instrumentation.

6.3.2 C_{60}

The Seebeck coefficient of a C_{60} monolayer on a Au substrate, prepared by thermal sublimation in UHV by Charlie Wells (see section 3.4 for details on sample preparation) was measured using the ThAFM application module, and a value of $S = -11.5 \pm 1.2\ \mu VK^{-1}$ was obtained. The thermovoltage-temperature plot is shown in figure 6.17.

Measurements of the Seebeck coefficient of C_{60} using various SPM methods have been reported in the literature. The sign of S is universally agreed to be negative, corresponding to LUMO dominated transport. A value of $-8.9 \pm 2.2\ \mu VK^{-1}$ was reported [78] from STMBJ with a Pt probe and Au substrate, although it is unclear if the standard offset calibration step of subtracting the Seebeck coefficient of plain Au was included in the reported figure. Another STM study [77] using a Au probe and substrate measured an average S value of -18 μVK^{-1}, although a wide range of values from -40 $\mu VK^{-1} \leq S \leq 0\ \mu VK^{-1}$ were reportedly observed across repeat measurements, suggesting a large unspecified uncertainty.

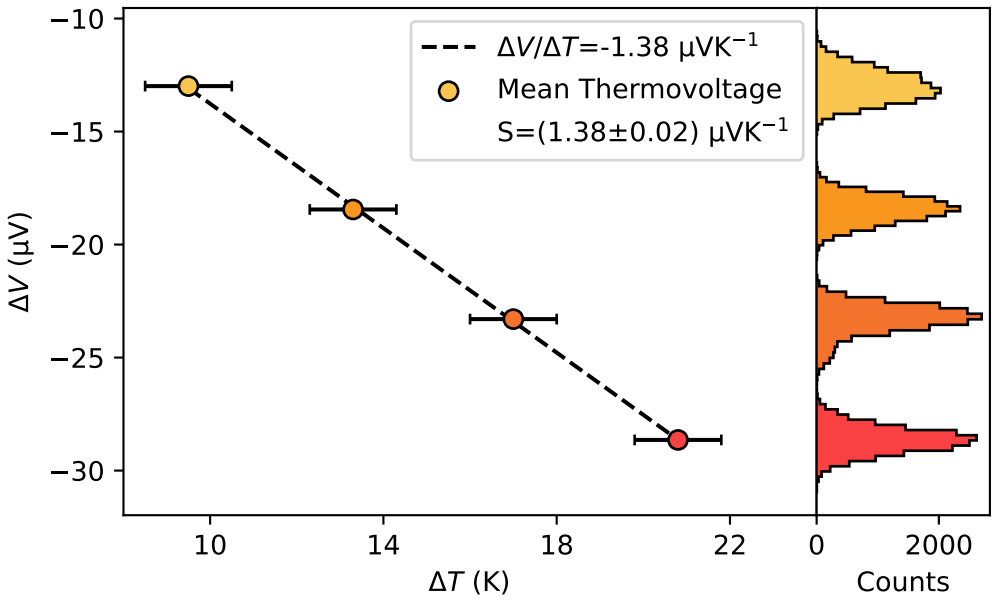


Fig. 6.16: A plot of measured thermovoltage ΔV as a function of tip-sample temperature difference ΔT for the control measurements performed on Au with a Pt probe. A histogram on the right shows the distribution of the measured thermovoltage.

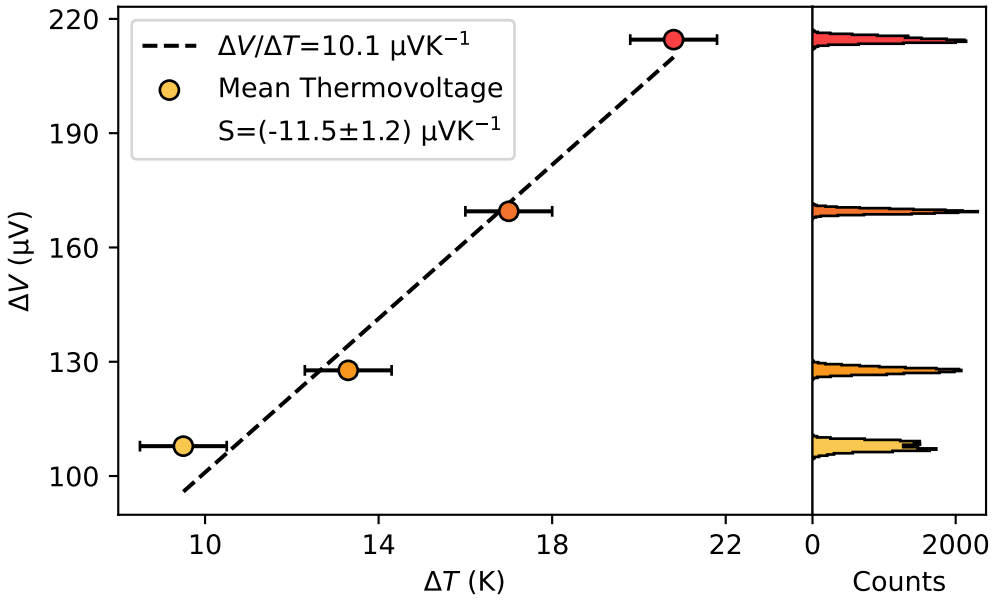


Fig. 6.17: A plot of measured thermovoltage ΔV as a function of tip-sample temperature difference ΔT for a C_{60} monolayer on Au with a Pt probe. A histogram on the right shows the distribution of the measured thermovoltage.

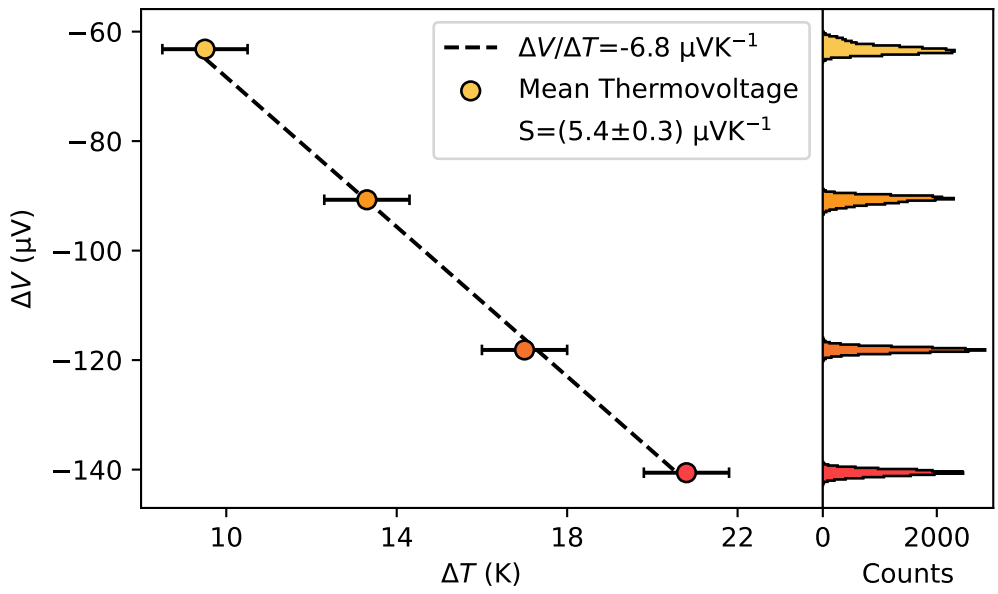


Fig. 6.18: A plot of measured thermovoltage ΔV as a function of tip-sample temperature difference ΔT for an ODT monolayer on Au with a Pt probe. A histogram on the right shows the distribution of the measured thermovoltage.

In summary, the value of $S = -11.5 \pm 1.2\ \mu V K^{-1}$ obtained here for a C_{60} monolayer on Au with a Pt probe is consistent with values reported in the literature, and appears to present a lower uncertainty than any previously reported measurement.

6.3.3 ODT

The Seebeck coefficient of a 1,8 octanedithiol (ODT) monolayer on a Au substrate, prepared by self assembly in solution (see section 3.3 for details on sample preparation) was measured with the ThAFM application module, and a value of $S = 5.4 \pm 0.3$ $\mu V K^{-1}$ was obtained. The thermovoltage-temperature plot is shown in figure 6.18.

Measurements of the thermopower of ODT are less frequently reported than for C_{60}, but transport is known to be HOMO dominated [181], resulting in a positive sign of S. In 2021 an STMBJ study with a Au probe and substrate reported a value of $5 \pm 2\ \mu V K^{-1}$ [73]. This is consistent with the value measured here, and again the ThAFM system introduced in this work appears to measure the Seebeck coefficient to a higher precision than previously reported measurements.

6.3.4 ZnPc

In order to test the application of the ThAFM system to an 'unknown' system, the Seebeck coefficient of a ZnPc monolayer on a Au substrate, prepared by thermal sublimation in UHV by Charlie Wells, was measured. The thermovoltage-temperature plot is shown in figure 6.19. A value of $S = 4.4 \pm 0.4\ \mu V K^{-1}$ was obtained.

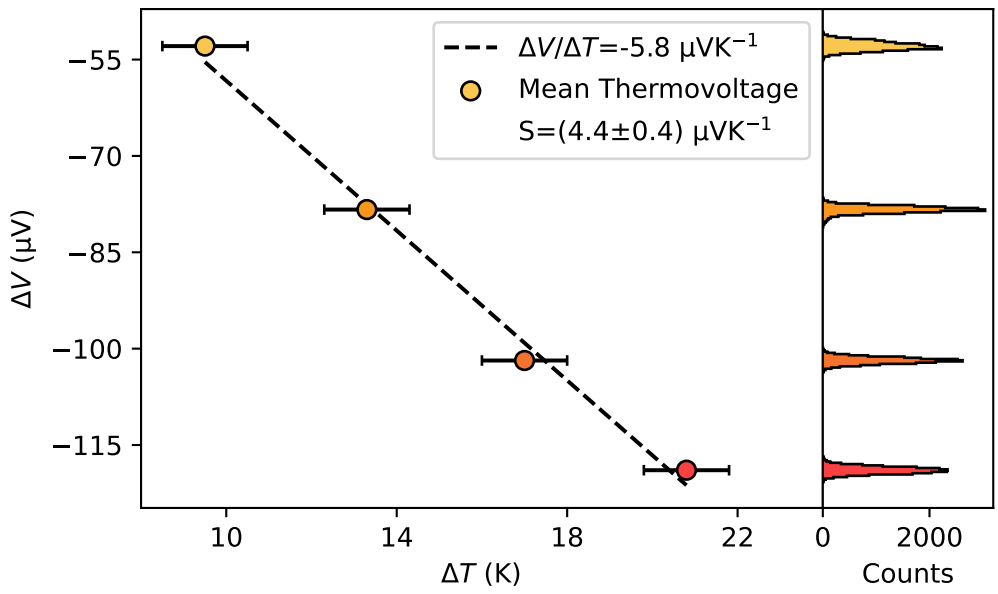


Fig. 6.19: A plot of measured thermovoltage ΔV as a function of tip-sample temperature difference ΔT for a ZnPc monolayer on Au with a Pt probe. A histogram on the right shows the distribution of the measured thermovoltage.

Measurements of the Seebeck coefficient for ZnPc monolayer-scale films have not been previously reported in the literature. However, a recent theoretical modelling study has predicted HOMO mediated transport [85], corresponding to a positive Seebeck coefficient, consistent with the measured result.

7. CONCLUSIONS

At the outset of this work, a clear necessity for a new generation of improved techniques for the characterisation of molecular junctions via SPM was identified. The development efforts presented here focused on the STMBJ technique for the formation of single-molecule junctions, and the ThAFM technique for thermoelectric measurements of multi-molecule junctions. The primary contributions of this work are outlined below, followed by a number of recommendations for future work in section 7.1.

An extensible bespoke STMBJ platform was developed in order to facilitate access to single-molecule junctions, a capability that had previously been missing within the research group. The development of this instrument resulted in a number of novel instrumentation contributions, including the first demonstration of the use of steel-reinforced epoxy granite (SREG) parts, produced by casting in 3D printed moulds, in the construction of an SPM. This approach was shown to effectively combine the ease and flexibility of FDM 3D printing with the mechanical performance of parts fabricated from alloys such as steel or cast-iron, providing a route towards inexpensive rapid prototyping of SPM instruments.

Due to the stochastic nature of molecular junction formation in STMBJ, large datasets containing measurements of many junctions are required in order to extract reproducible results. The application of feature-extraction algorithms to assist in the segmentation and analysis of such datasets has become increasingly relevant in recent years as a means of identifying patterns and sub-populations within the dataset which are inaccessible to typical 'all-data' statistical approaches. An investigation into autoencoders as feature extractors for STMBJ was conducted in this work, leading to the first implementation of a feature extraction algorithm based on a cyclically annealed variational autoencoder in an STMBJ data analysis context. The resulting data segmentation algorithm showed enhanced accuracy on synthetic datasets compared to an equivalent algorithm using a standard autoencoder. Additionally, the role of disentanglement in feature extraction for STMBJ data analysis was explored for the first time, and a new procedurally generated synthetic STMBJ dataset with a known distribution of generative factors was introduced to enable quantitative evaluation of the disentanglement of STMBJ feature extraction methods.

ThAFM, a variant of CAFM which enables the measurement of the thermopower of multi-molecule junctions, has long been one of the most important measurement techniques used within the research group. ThAFM has been found to be a challenging technique, with instabilities in the thermovoltage signal due to various instrumentation challenges resulting in unreliable performance and large uncertainties in the extracted Seebeck coefficients. As part of this work a bespoke self-contained ThAFM module was developed based on a review of the limitations of existing systems. The new module cost less than £100 in parts to construct, and significantly outperformed the setup previously used in the lab. Measurements of the Seebeck coefficient were demonstrated for two benchmark molecules (ODT and C_{60}) with precision exceeding that of existing measurements reported in the literature. Finally, a measurement of the Seebeck coefficient of a ZnPc monolayer was presented for the first time, supporting theoretical predictions of HOMO mediated transport in the literature.

Many of the key contributions of this work are also applicable to many other SPM techniques within molecular electronics and beyond. For instance, the demonstration of SREG as a viable composite material for the fabrication of SPM instruments could provide others with a simple, inexpensive means to build high-performance instrumentation for other applications with similarly demanding requirements for rigidity and damping.

7.1 Recommendations for Further Work

While the instruments and techniques developed and evaluated in this work have produced several novel contributions to knowledge, they largely represent incremental improvements to existing measurement techniques. Hence, future efforts should aim to extend the capabilities of the platforms introduced here, towards new multivariate characterisation techniques to explore previously inaccessible correlations between the diverse phenomena present in molecular junctions.

7.1.1 Towards Next-Generation Multivariate Break Junction Measurements

The STMBJ developed in this work demonstrated the capability to perform single-molecule conductance measurements. However, the instrument was intentionally designed with more advanced measurement techniques in mind, and extensibility of the mechanical, electronic and software systems was one of the key design specifications outlined in chapter 4. Future development of this instrument should focus on extending the measurement capabilities to include multivariate measurements, where multiple properties beyond conductance are measured simultaneously throughout the junction formation and breaking process. Such measurements have become

increasingly commonplace in recent years, with several groups routinely using simultaneous measurements of properties such as thermopower [73], mechanoresistive responses to piezo modulation [182] and flicker noise [93] to extract additional information from break-junction measurements.

Most STMBJ instruments are limited by the capabilities of the control and data acquisition system; typically either a modified commercial SPM controller or a general-purpose DAQ (as was the case in this work) is used. The main limitation of the former is the restricted ability to implement custom modes beyond those intended by the manufacturer, while the latter often does not possess the same multi-channel real-time DSP functionality of a dedicated SPM controller. Recently, an open-source SPM controller named 'Gwyscope', which offers support for adaptive scanning modes and other advanced capabilities, has been introduced [183]. The open-source and modifiable nature of the Gwyscope controller could enable the development of new STMBJ modes at the FPGA HDL level, affording unprecedented access to advanced modulation, signal routing and real-time DSP capabilities.

In order to process the multi-channel data produced by the upgraded STMBJ, further development of the unsupervised feature extraction and clustering algorithm will be necessary. Disentangled feature extraction will be especially useful in this context; encoding each generative factor into a single latent code simplifies the task of identifying which of the measured properties of the junction is responsible for a given clustering result. Future efforts should therefore focus on extending the cyclically annealed βVAE to handle multi-channel data. Additionally, different annealing schedules and values of β should be explored be as a possible route towards further enhanced disentanglement and clustering accuracy.

7.1.2 Improved Temperature Control and Monitoring for Thermovoltage Atomic Force Microscopy

The ThAFM measurement described in chapter 6 consists of thermovoltage measurements at 4 different sample temperatures (with the probe held at ambient temperature), from which the Seebeck coefficient can be obtained via a linear fit to the thermovoltage-temperature data. This method, as implemented here, has several drawbacks: firstly, the process of performing thermovoltage measurements at each temperature is time consuming, since the temperature of the sample stage takes time to stabilise. Secondly, the method assumes a linear relationship between thermovoltage and temperature gradient. Over large temperature ranges this is not necessarily true, as the Seebeck coefficient itself is expected to vary with average junction temperature [45, 184].

An improved temperature control system, with a wider temperature range (the Bruker TAC is limited to a maximum sample temperature of 60°C) and the ability to

measure sample temperature in real time would facilitate a different measurement procedure, in which the sample temperature is continuously ramped over a wide range and the thermovoltage and temperature gradient are recorded simultaneously. This would improve throughput, since there is no need to wait for the temperature to stabilise at multiple discrete setpoints, and allow the Seebeck coefficient to be measured as a function of temperature, by fitting a non-linear curve to the data and calculating the first derivative [184].

7.1.3 A Versatile Open-Source 'Multi-Tool' Module for Atomic Force Microscopy

The ThAFM expansion module developed here has achieved the stated objective of adding precision thermovoltage measurement capabilities to a commercial AFM. However, the current implementation is limited in that simultaneous electrical conductance measurements are not possible: presently conductance measurements are performed using the Bruker TUNA amplifier module, which cannot be installed at the same time as the ThAFM module. Future work could remedy this by integrating a current amplifier and precision bias voltage source alongside the thermovoltage amplifier in the module, with a multiplexed input to provide quasi-simultaneous measurements of thermopower and electrical conductance. This would allow the full thermoelectric 'power factor' (S^2G) to be characterised in a single measurement.

A wide dynamic range, high bandwidth amplifier for conductance measurements of molecular junctions has been demonstrated in the literature using a Wheatstone bridge front-end with an INA116 instrumentation amplifier IC [185]. Coincidentally, this is the same IC used in the ThAFM module developed here; hence a simple circuit to switch between the two different front-end configurations would allow thermopower and conductance measurements using a single amplifier IC. The main challenge is likely to be in developing a sufficiently low-leakage multiplexing circuit to switch between the input configurations without adversely affecting the performance of either measurement.

LIST OF ABBREVIATIONS

AC	Alternating Current.
ADC	Analogue to Digital Converter.
AE	Auto Encoder.
AFM	Atomic Force Microscope.
BNC	Bayonet Neill–Concelman.
CAD	Computer Aided Design.
CAFM	Conductive Atomic Force Microscope.
CMOS	Complementary Metal-Oxide-Semiconductor.
CMRR	Common Mode Rejection Ratio.
CNC	Computer Numerical Control.
CQI	Constructive Quantum Interference.
DAC	Digital to Analogue Converter.
DAQ	Data Acquisition System.
DC	Direct Current.
DiFET	Dielectrically Isolated Field Effect Transistor.
DPGMM	Dirichlet Process Gaussian Mixture Model.
DQI	Destructive Quantum Interference.
DSP	Digital Signal Processing.
EG	Epoxy Granite.
EMI	Electromagnetic Interference.
FDM	Fused Deposition Modelling.
FET	Field Effect Transistor.
FIFO	First-In First-Out.
FPGA	Field-Programmable Gate Array.
GUI	Graphical User Interface.

HDL	Hardware Description Language.
HOMO	Highest Occupied Molecular Orbital.
I/O	Input/Output.
IC	Integrated Circuit.
JEMMIG	Joint Entropy Minus Mutual Information Gap.
JFET	Junction Field Effect Transistor.
KL	Kullback–Leibler.
LED	Light-Emitting Diode.
LSB	Least-Significant Bit.
LUMO	Lowest Unoccupied Molecular Orbital.
MIG	Mutual Information Gap.
MMCX	Micro-Miniature Coaxial.
MMJ	Multi-Molecule Junction.
ODT	Octanedithiol.
PETG	Polyethylene Terephthalate.
PID	Proportional-Integral-Derivative.
PLA	Polylactic Acid.
PSD	Power Spectral Density.
PSRR	Power Supply Rejection Ratio.
PSU	Power Supply Unit.
QI	Quantum Interference.
ReLu	Rectified Linear Unit.
SMJ	Single Molecule Junction.
SPM	Scanning Probe Microscopy.
SREG	Steel Reinforced Epoxy Granite.
STM	Scanning Tunnelling Microscope.
STMBJ	Scanning Tunnelling Microscope Break Junction.
TAC	Thermal Applications Controller.

TEC	Thermoelectric Cooler.
ThAFM	Thermovoltage Atomic Force Microscopy.
TUNA	Tunnelling AFM.
UHV	Ultra-High Vacuum.
USB	Universal Serial Bus.
VAE	Variational Auto Encoder.

BIBLIOGRAPHY

[1] Certain data included herein are derived from clarivate (web of science). clarivate 2024. all rights reserved.

[2] J. C. Poler, R. M. Zimmermann, and E. C. Cox. Scanning thermopower microscopy of guanine monolayers. *ACS Langmuir*, 11, 1995.

[3] A. Tan, S. Sadat, and P. Reddy. Measurement of thermopower and current-voltage characteristics of molecular junctions to identify orbital alignment. *Appl. Phys. Lett. 96*, 2010.

[4] J. Tomfohr and O. F. Sankey. Theoretical analysis of electron transport through organic molecules. *Journal of Chemical Physics 120*, 2004.

[5] D. Xiang, X. Wang, C. Jia, T. Lee, and X. Guo. Molecular-scale electronics: From concept to function. *Chem. Rev. 116*, 2016.

[6] E. Scheer J. C. Cuevas. *Molecular Electronics: An Introduction to Theory and Experiment*. World Scientific, 2010.

[7] G. Cohen and M. Galperin. Green's function methods for single molecule junctions. *J. Chem. Phys. 152*, 2020.

[8] V. R. Manfrinato, J. Wen, L. Zhang, Y. Yang, R. G. Hobbs, B. Baker, D. Su, D. Zakharov, N. J. Zaluzec, D. J. Miller, E. A. Stach, and K. K. Berggren. Determining the resolution limits of electron-beam lithography: direct measurement of the point-spread function. *Nano Lett. 14*, 2014.

[9] M. Baghernejad, C. Van Dyck, J. Bergfield, D. R. Levine, A. Gubicza, J. D. Tovar, M. Calame, P. Broekmann, and W. Hong. Quantum interference enhanced chemical responsivity in single-molecule dithienoborepin junctions. *Chem. Eur. J. 25*, 2019.

[10] O. Sengul, J. Völkle, A. Valli, and R. Stadler. Enhancing the sensitivity and selectivity of pyrene-based sensors for detection of small gaseous molecules via destructive quantum interference. *Phys. Rev. B 105*, 105, 2022.

[11] C. Lambert. Basic concepts of quantum interference and electron transport in single-molecule electronics. *Chem Soc Rev, 44*, 2015.

[12] R. P. Feynman. There's plenty of room at the bottom. *APS Annual Meeting*, 1959.

[13] A. Von Hippel. Molecular engineering. *Science 123*, 1956.

[14] S. Marqués-González and P. J. Low. Molecular electronics: History and fundamentals. *Australian Journal of Chemistry 69*, 2016.

[15] E. A. Sack and D. A. Laws. Westinghouse: Microcircuit pioneer from molecular electronics to ics. *IEEE Annals of the History of Computing 34*, 2012.

[16] M. A. Ratner A. Aviram. Molecular rectifiers. *Chemical Physics Letters 29*, 1974.

[17] M. Ratner. A brief history of molecular electronics. *Nature Nanotechnology 8*, 2013.

[18] R. Young, J. Ward, and F. Scire. The topografiner: An instrument for measuring surface microtopography. *Rev. Sci. Instrum. 43*, 1972.

[19] G. Binnig and H. Rohrer. Scanning tunneling microscopy—from birth to adolescence. *Rev. Mod. Phys. 59*, 1987.

[20] G. Binnig, C. Quate, and C. Gerber. Atomic force microscope. *Phys. Rev. Lett. 56*, 1986.

[21] C. A. Mirkin and M. A. Ratner. Molecular electronics. *Annu. Rev. Phys. Chem 43*, 1992.

[22] R. Wiesendanger. *Scanning Probe Microscopy and Spectroscopy.* Cambridge University Press, 1994.

[23] M. A. Reed, C. Zhou, C. J. Muller, T. P. Burgin, and J. M. Tour. Conductance of a molecular junction. *Science 278*, 1997.

[24] B. Xu and N. J. Tao. Measurement of single-molecule resistance by repeated formation of molecular junctions. *Science, 301*, 2003.

[25] S. M. Lindsay and M. A. Ratner. Molecular transport junctions: Clearing mists. *Advanced Materials 19*, 2007.

[26] A. Gemma, F. Tabatabaei, U. Drechsler, A. Zulji, H. Dekkiche, N. Mosso, T. Niehaus, M. R. Bryce, S. Merabia, and B. Gotsmann. Full thermoelectric characterization of a single molecule. *Nature Communications 14*, 2023.

[27] S. V. Aradhya and L. Venkataraman. Single-molecule junctions beyond electronic transport. *Nature Nanotechnology 8*, 2013.

[28] C. Tang, R. T. Ayinla, and K. Wang. Beyond electrical conductance: progress and prospects in single-molecule junctions. *J. Mater. Chem. C 10*, 2022.

[29] C. J. Muller, B. J. Vleeming, M. A. Reed, J. J. S. Lamba, R. Hara, L. Jones, and J. M. Tour. Atomic probes: a search for conduction through a single molecule. *Nanotechnology 7*, 1996.

[30] P. Gehring, J. M. Thijssen, and H. S. J. van der Zant. Single-molecule quantum-transport phenomena in break junctions. *Nature Reviews Physics 1*, 2019.

[31] W. Lee and P. Reddy. Creation of stable molecular junctions with a custom-designed scanning tunneling microscope. *Nanotechnology 22*, 2011.

[32] L. Cui, R. Miao, C. Jiang, E. Meyhofer, , and P. Reddy. Perspective: Thermal and thermoelectric transport in molecular junctions. *The Journal of Chemical Physics, 146*, 2017.

[33] C. Lambert. *Quantum Transport in Nanostructures and Molecules*. IOP Publishing Ltd, 2021.

[34] D. Segal. Thermoelectric effect in molecular junctions: A tool for revealing transport mechanisms. *Physical Review B, 72*, 2005.

[35] A. F. Ioffe; L. S. Stil'bans, E. K. Iordanishvili, T. S. Stavitskaya, A. Gelbtuch, and G. Vineyard. Semiconductor thermoelements and thermoelectric cooling. *Physics Today 12 (5)*, 1959.

[36] J. He and T.M. Tritt. Advances in thermoelectric materials research: Looking back and moving forward. *Science, 357*, 2017.

[37] G.T. Craven and A. Nitzan. Wiedemann–franz law for molecular hopping transport. *Nano Lett. 20*, 2020.

[38] P. Pichanusakorn and P. Bandaru. Nanostructured thermoelectrics. *Materials Science and Engineering: R: Reports, 67*, 2010.

[39] H. Sadeghi. Quantum and phonon interference-enhanced molecular-scale thermoelectricity. *J. Phys. Chem. C 123*, 2019.

[40] A. K. Ismael and C. J. Lambert. Molecular-scale thermoelectricity: a worst-case scenario. *Nanoscale Horiz. 5*, 2020.

[41] M. Lanza. *Conductive atomic force microscopy : applications in nanomaterials*. Wiley, 2017.

[42] G. B. Lesovik and I. A. Sadovskyy. Scattering matrix approach to the description of quantum electron transport. *Physics-Uspekhi*, 2012.

[43] J. J. Sakurai. *Modern Quantum Mechanics (Revised Edition)*. Pearson, 1994.

[44] M. V. Moskalets. *Scattering Matrix Approach to Non-Stationary Quantum Transport*. World Scientific Publishing, 2011.

[45] K. Wang, E. Meyhofer, and P. Reddy. Thermal and thermoelectric properties of molecular junctions. *Adv. Funct. Mater. 30*, 2020.

[46] S. Datta. *Electronic Transport in Mesoscopic Systems*. Cambridge University Press, 1995.

[47] D. K. Ferry, S. M. Goodnick, and J. Bird. *Transport in Nanostructures*. Cambridge University Press, 2009.

[48] S. X. Liu, A. Ismael, A. Al-Jobory, and C. Lambert. Signatures of room-temperature quantum interference in molecular junctions. *Acc. Chem. Res. 56*, 2023.

[49] D. Z. Manrique, C. Huang, M. Baghernejad, X. Zhao, O. A. Al-Owaedi, H. Sadeghi, V. Kaliginedi, W. Hong, M. Gulcur, T. Wandlowski, M. R. Bryce, and C. J. Lambert. A quantum circuit rule for interference effects in single-molecule electrical junctions. *Nature Communications 6*, 6, 2015.

[50] R. Miao, H. Xu, M. Skripnik, L. Cui, K. Wang, K. G. L. Pedersen, M. Leijnse, F. Pauly, K. Wärnmark, E. Meyhofer, P. Reddy, and H. Linke. Influence of quantum interference on the thermoelectric properties of molecular junctions. *Nano Lett. 18*, 2018.

[51] J. Liu, X. Huang, F. Wang, and W. Hong. Quantum interference effects in charge transport through single-molecule junctions: Detection, manipulation, and application. *Acc. Chem. Res. 52*, 52:151–160, 2019.

[52] D. Q. Andrews, G. C. Solomon, R. P. Van Duyne, and M. A. Ratner. Single molecule electronics: Increasing dynamic range and switching speed using cross-conjugated species. *J. Am. Chem. Soc. 130*, 130:17309–17319, 2008.

[53] D. M. Cardamone, C. A. Stafford, and S. Mazumdar. Controlling quantum transport through a single molecule. *Nano Lett. 6*, 6:2422–2426, 2006.

[54] A. Saraiva-Souza, M. Smeu, L. Zhang, A. G. S. Filho, H. Guo, and M. A. Ratner. Molecular spintronics: Destructive quantum interference controlled by a gate. *J. Am. Chem. Soc. 42*, 2014.

[55] D. Nozaki, M. Lokamani, A. Santana-Bonilla, A. Dianat, R. Gutierrez, and G. Cuniberti. Switchable negative differential resistance induced by quantum interference effects in porphyrin-based molecular junctions. *J. Phys. Chem. Lett. 6*, 2015.

[56] S. Sangtarash, C. Huang, H. Sadeghi, G. Sorohhov, J. Hauser, T. Wandlowski, W. Hong, S. Decurtins, S. X. Liu, and C. J. Lambert. Searching the hearts of graphene-like molecules for simplicity, sensitivity, and logic. *J. Am. Chem. Soc. 137*, 2015.

[57] R. Stadler, S. Ami, C. Joachim, and M. Forshaw. Integrating logic functions inside a single molecule. *Nanotechnology 15*, 15:S115–S121, 2004.

[58] M. Carlotti, S. Soni, S. Kumar, Y. Ai, E. Sauter, M. Zharnikov, and R. C. Chiechi. Two-terminal molecular memory through reversible switching of quantum interference features in tunneling junctions. *Angew. Chem. Int. Ed. 57*, 130:15907–15911, 2018.

[59] R. Otero, J. M. Gallego, A. L. Vázquez de Parga, N. Martín, and R. Miranda. Molecular self-assembly at solid surfaces. *Advanced Materials 23*, 2011.

[60] G. M. Whitesides, J. P. Mathias, and C. T. Seto. Molecular self-assembly and nanochemistry: A chemical strategy for the synthesis of nanostructures. *Science 254*, 1991.

[61] T.M. Uehara, H.B. de Aguiar, K. Bergamaski, and P.B. Miranda. Adsorption of alkylthiol self-assembled monolayers on gold and the effect of substrate roughness: A comparative study using scanning tunneling microscopy, cyclic voltammetry, second-harmonic generation, and sum-frequency generation. *J. Phy. Chem. C*, 2014.

[62] N. Vogel, J. Zieleniecki, and I. Köper. As flat as it gets: ultrasmooth surfaces from template-stripping procedures. *Nanoscale, 4, 3820*, 2012.

[63] M. Hegner, P. Wagner, and G. Semenza. Ultralarge atomically flat template-stripped au surfaces for scanning probe microscopy. *Surface Science 291*, 1993.

[64] S. Y. Guan, Z. Y. Cai, Z. W. Ma, D. Y. Wu, and Z. Q. Tian. Binding structure, breaking forces and conductance of Au-Octanedithiol-Au molecular junction under stretching processes: a DFT-NEGF study. *Nanotechnology 34*, 2022.

[65] J. Li, J. K. Tomfohr, and O. F. Sankey. Electric field effects on the octanedithiol wire. *PSS B 239*, 2003.

[66] K. Wang, J. M. Hamill, J. Zhou, and B. Xu. Mapping the details of contact effect of modulated Au-Octanedithiol-Au break junction by force–conductance cross-correlation. *J. Am. Chem. Soc. 136*, 2014.

[67] S. Jang, P. Reddy, A. Majumdar, and R. Segalman. Interpretation of stochastic events in single molecule conductance measurements. *Nano Letters 6 (10)*, 2006.

[68] J. He, O. Sankey, M. H. Lee, Y. Tian, X. Li, and S. Lindsay. Measuring single molecule conductance with break junctions. *Faraday Discuss. 131*, 2006.

[69] F. H. van Veen, L. Ornago, H. S. J. van der Zant, and M. El Abbassi. Benchmark study of alkane molecular chains. *J. Phys. Chem. C 126*, 2022.

[70] J. Hihath, C. Bruot, and N. Tao. Electron-phonon interactions in single octanedithiol molecular junctions. *ACS Nano 4*, 2010.

[71] Y. Qi, J. Qin, G. Zhang, and T. Zhang. Breaking mechanism of single molecular junctions formed by octanedithiol molecules and Au electrodes. *J. Am. Chem. Soc. 131*, 2009.

[72] C. Li, I. Pobelov, T. Wandlowski, A. Bagrets, A. Arnold, and F. Evers. Charge transport in single Au / alkanedithiol / Au junctions: coordination geometries and conformational degrees of freedom. *J. Am. Chem. Soc. 130*, 2008.

[73] J. M. Hamill, C. Weaver, and T. Albrecht. Multivariate approach to single-molecule thermopower and electrical conductance measurements. *J. Phys. Chem. C 125*, 2021.

[74] Y. Jun, X.Y. Zhu, and J. W. P. Hsu. Formation of alkanethiol and alkanedithiol monolayers on GaAs(001). *Langmuir 22*, 2006.

[75] B. D. Mansuriya and Z. Altintas. *Fundamentals of Sensor Technology*, chapter Carbon nanomaterials for sensing applications. Woodhead Publishing, 2023.

[76] H. W. Kroto, J. R. Heath, S. C. O'Brien, R. F. Curl, and R. E. Smalley. C60: Buckminsterfullerene. *Nature 318*, 1985.

[77] C. Evangeli, K. Gillemot, E. Leary, M. Teresa González, G. Rubio-Bollinger, C. J. Lambert, and N. Agraït. Engineering the thermopower of C60 molecular junctions. *Nano Lett. 13*, 2013.

[78] S. K. Yee, J. A. Malen, A. Majumdar, and R. A. Segalman. Thermoelectricity in fullerene–metal heterojunctions. *Nano Lett. 11*, 11:4089–4094, 2011.

[79] Y. Komoto, Y. Isshiki, S. Fujii, T. Nishino, and M. Kiguchi. Evaluation of the electronic structure of single-molecule junctions based on current–voltage and thermopower measurements: Application to C60 single-molecule junction. *Chem. Asian J. 12*, 2017.

[80] Y. Kim, W. Jeong, K. Kim, W. Lee, and P. Reddy. Electrostatic control of thermoelectricity in molecular junctions. *Nature Nanotechnology 9*, 2014.

[81] T. R. Ohno, Y. Chen, S. E. Harvey, G. H. Kroll, J. H. Weaver, R. E. Haufler, and R. E. Smalley. C60 bonding and energy-level alignment on metal and semiconductor surfaces. *Phys. Rev. B*, 1991.

[82] X. Liu, Y. Zhan, S. Braun, F. Li, and M. Fahlman. Interfacial electronic properties of pentacene tuned by a molecular monolayer of C60. *Phys. Rev. B*, 2009.

[83] P. Das, K. Chakraborty, S. Chakrabarty, S. Ghosh, and T. Pal. Reduced graphene oxide - zinc phthalocyanine composites as fascinating material for optoelectronic and photocatalytic applications. *ChemistrySelect 2*, 2017.

[84] Z. Li, B. Li, J. Yang, and J. G. Hou. Single-molecule chemistry of metal phthalocyanine on noble metal surfaces. *Acc. Chem. Res. 43*, 2010.

[85] M. T. González, A. K. Ismael, M. García-Iglesias, E. Leary, G. Rubio-Bollinger, I. Grace, D. González-Rodríguez, T. Torres, C. J. Lambert, and N. Agraït. Interference controls conductance in phthalocyanine molecular junctions. *J. Phys. Chem. C 125*, 2021.

[86] P. Amsalem, L. Giovanelli, J. M. Themlin, and T. Angot. Electronic and vibrational properties at the ZnPc/Ag(110) interface. *Phys. Rev. B 79*, 2009.

[87] A. A. K. Al-mebir, M. D. Noori, and B. B. Kadhim. Tuning the electrical and thermoelectric properties of phthalocyanine and metallo-phthalocyanine molecular junction. *J. Phys. Conf. Ser. 1999*, 2021.

[88] Y. Acikbas, M. Erdogan, R. Capan, C. O. Erdogan, Y. Baygu, N. Kabay, Y. Gök, and G. Kucukyildiz. Preparation and characterization of the phthalocyanine–zinc(ii) complex-based nanothin films: optical and gas-sensing properties. *Appl. Nanosci. 13*, 2023.

[89] C. Tang, Y. Tang, Y. Ye, Z. Yan, Z. Chen, L. Chen, L. Zhang, J. Liu, J. Shi, H. Xia, and W. Hong. Identifying the conformational isomers of single-molecule cyclohexane at room temperature. *Chem 6*, 2020.

[90] M. K. Marichelvam, K. Kandakodeeswaran, and M. Geetha. Development of hybrid composite materials for machine tool structures. *Materials Today: Proceedings 47*, 2021.

[91] X. Luo, K. Cheng, D. Webb, and F. Wardle. Design of ultraprecision machine tools with applications to manufacture of miniature and micro components. *Journal of Materials Processing Technology 167*, 167:515–528, 2005.

[92] I. Battisti, G. Verdoes, K. van Oosten, K. M. Bastiaans, and M. P. Allan. Definition of design guidelines, construction, and performance of an ultra-stable scanning tunneling microscope for spectroscopic imaging. *Rev. Sci. Instrum. 89*, 2018.

[93] Z. Pan, L. Chen, C. Tang, Y. Hu, S. Yuan, T. Gao, J. Shi, J. Shi, Y. Yang, and W. Hong. The evolution of the charge transport mechanism in single-molecule break junctions revealed by flicker noise analysis. *Small 18*, 2022.

[94] `https://www.lancaster.ac.uk/physics/isolab/`.

[95] A. I. Oliva, M. Aguilar, and V. Sosa. Low- and high-frequency vibration isolation forscanning probe microscopy. *Meas. Sci. Technol. 9*, 1997.

[96] J. Stroscio and W. Kaiser. *Scanning Tunneling Microscopy*. Academic Press, Inc., 1993.

[97] H. Möhring, C. Brecher, E. Abele, J. Fleischer, and F. Bleicher. Materials in machine tool structures. *CIRP Annals 64*, 2015.

[98] J. H. Kindt, G. E. Fantner, J. A. Cutroni, and P. K. Hansma. Rigid design of fast scanning probe microscopes using finite element analysis. *Ultramicroscopy 100*, 2004.

[99] R. Adams. The damping characteristics of certain steels, cast irons and other metals. *Journal of Sound and Vibration 23*, 1972.

[100] P. Dhanabal, P. R. Thyla, M. Kalayarasan, and N. Mahendrakumar. The influence of steel reinforcements on static rigidity of epoxy granite spindle head for dynamic performance enhancement of machining center: Experiments and finite element simulation. 2023.

[101] P. R. Venugopal, M. Kalayarasan, P. R. Thyla, P. V. Mohanram, M. Nataraj, S. Mohanraj, and H. Sonawane. Structural investigation of steel-reinforced epoxy granite machine tool column by finite element analysis. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications. 233*, 2019.

[102] S. Chinnuraj, P. R. Thyla, S. Elango, P. R. Venugopal, P. V. Mohanram, M. Nataraj, S. Mohanraj, K. N. Manojkumar, and S. Ayyasamy. Static and dynamic behavior of steel-reinforced epoxy granite CNC lathe bed using finite element analysis. *Proceedings of the Institution of Mechanical Engineers, Part L: Journal of Materials: Design and Applications. 234*, 234:595–609, 2020.

[103] V. R. Mula, A. Ramachandran, and T. P. Ramasamy. A review on epoxy granite reinforced polymer composites in machine tool structures – static, dynamic and thermal characteristics. *Polym. Compos. 44(4)*, 2023.

[104] H.S. Liao, C. Werner, R. Slipets, P.E. Larsen, I.S. Hwang, T.J. Chang, H.U. Danzebrink, K.Y. Huang, and E.T. Hwu. Low-cost, open-source XYZ nanopositioner for high-precision analytical applications. *HardwareX*, 11, 2022.

[105] Thorlabs. *MDT693B and MDT694B Open-Loop Piezo Controllers User Guide, Rev F*, 2023.

[106] FEMTO Messtechnik GmbH. *DLPCA-200 Datasheet, Rev. 19*, 2019.

[107] J. Bechhoefer. Feedback for physicists: A tutorial essay on control. *Rev. Mod. Phys. 77*, 2005.

[108] J. Stirling. Control theory for scanning probe microscopy revisited. *Beilstein J. Nanotechnol. 5*, 2014.

[109] R. Quan, C. S. Pitler, M. A. Ratner, and M. G. Reuter. Quantitative interpretations of break junction conductance histograms in molecular electron transport. *ACS Nano 9*, 2015.

[110] M. Lemmer, M. Inkpen, K. Kornysheva, N. Long, and T. Albrecht. Unsupervised vector-based classification of single-molecule charge transport data. *Nat. Commun., 7*, 2016.

[111] W. Bro-Jørgensen, J. Hamill, R. Bro, and G. Solomon. Trusting our machines: validating machine learning models for single-molecule transport experiments. *Chem. Soc. Rev., 51*, 2022.

[112] V. Sacchetti, J. Ramos-Soriano, B. Illescas, M. T. González, D. Li, L. Palomino-Ruiz, I. R. Márquez, E. Leary, G. Rubio-Bollinger, F. Pauly, N. Agraït, and N. Martín. Effect of charge-assisted hydrogen bonds on single-molecule electron transport. *J. Phys. Chem. C 123*, 2019.

[113] L. Wang, Z. Gong, S. Y. Li, W. Hong, Y. W. Zhong, D. Wang, and L. Wan. Molecular conductance through a quadruple-hydrogen-bond-bridged supramolecular junction. *Angew. Chem. Int. Ed. 55*, 2016.

[114] F. Huang, R. Li, G. Wang, J. Zheng, Y. Tang, J. Liu, Y. Yang, Y. Yao, J. Shi, and W. Hong. Automatic classification of single-molecule charge transport data with an unsupervised machine-learning algorithm. *Phys. Chem. Chem. Phys., 22*, 2020.

[115] A. Vladyka and T. Albrecht. Unsupervised classification of single-molecule data with autoencoders and transfer learning. *Mach. Learn.: Sci. Technol. 1*, 2020.

[116] D. Francois, V. Wertz, and M. Verleysen. The concentration of fractional distances. *IEEE Transactions on Knowledge and Data Engineering 19*, 2007.

[117] E. M. Mirkes, J. Allohibi, and A. Gorban. Fractional norms and quasinorms do not help to overcome the curse of dimensionality. *Entropy 22*, 2020.

[118] S. Kumari and B. Jayaram. Measuring concentration of distances—an effective and efficient empirical index. *IEEE Transactions on Knowledge and Data Engineering 29*, 2017.

[119] C.C. Aggarwal, A. Hinneburg, and D.A. Keim. *Database Theory - ICDT*, chapter On the Surprising Behavior of Distance Metrics in High Dimensional Space. Springer, 2001.

[120] R. Xu and D. Wunsch. Survey of clustering algorithms. *IEEE Transactions on Neural Networks 16*, 2005.

[121] W. Jia, M. Sun, J. Lian, and S. Hou. Feature dimensionality reduction: a review. *Complex and Intelligent Systems 8*, 2022.

[122] Y. Bengio, A. Courville, and P. Vincent. Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence 35*, 2013.

[123] M. Tschannen, O. Bachem, and M. Lucic. Recent advances in autoencoder-based representation learning. *Third workshop on Bayesian Deep Learning (NeurIPS)*, 2018.

[124] D.E. Rumelhart, G.E. Hinton, , and R.J. Williams. *Parallel Distributed Processing. Vol 1: Foundations*, chapter Learning internal representations by error propagation. MIT Press, 1986.

[125] G. E. Hinton and R. R. Salakhutdinov. Reducing the dimensionality of data with neural networks. *Science 313*, 2006.

[126] U. Michelucci. An introduction to autoencoders. *arXiv:2201.03898*, 2022.

[127] K. Hornik, M. Stinchcombe, and H. White. Multilayer feedforward networks are universal approximators. *Neural Networks 2*, 1989.

[128] D. P. Kingma and J. Ba. Adam: A method for stochastic optimization. *3rd International Conference for Learning Representations*, 2015.

[129] K. Berahmand, F. Daneshfar, E. S. Salehi, Y. Li, and Y. Xu. Autoencoders and their applications in machine learning: a survey. *Artificial Intelligence Review*, 2024.

[130] D. P. Kingma and M. Welling. Auto-encoding variational bayes. *arXiv:1312.6114*, 2013.

[131] C. Doersch. Tutorial on variational autoencoders. *arXiv:1606.05908*, 2016.

[132] D. P. Kingma and M. Welling. An introduction to variational autoencoders. *Foundations and Trends in Machine Learning 12*, 2019.

[133] M. Rolínek, D. Zietlow, and G. Martius. Variational autoencoders pursue PCA directions (by accident). *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2019.

[134] I. Higgins, L. Matthey, A. Pal, C. Burgess, X. Glorot, M. Botvinick, S. Mohamed, and A. Lerchner. beta-VAE: Learning basic visual concepts with a constrained variational framework. *ICLR*, 2017.

[135] H. Kim and A. Mnih. Disentangling by factorising. *Proceedings of the 35th International Conference on Machine Learning, PMLR 80*, 2018.

[136] R. T. Q. Chen, X. Li, R. Grosse, and D. Duvenaud. Isolating sources of disentanglement in VAEs. *Proceedings of the 32nd International Conference on Neural Information Processing Systems (NIPS)*, 2018.

[137] M. Ding. The road from MLE to EM to VAE: A brief tutorial. *AI Open 3*, 2022.

[138] C. P. Burgess, I. Higgins, A. Pal, L. Matthey, N. Watters, G. Desjardins, and A. Lerchner. Understanding disentangling in β-VAE. *NIPS*, 2018.

[139] P. Vincent, H. Larochelle, Y. B. Bengioy, and P. A. Manzagol. Extracting and composing robust features with denoising autoencoders. *ICML 25*, 2008.

[140] A. Achille and S. Soatto. Information dropout: Learning optimal representations through noisy computation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016.

[141] P. Liò M. Peychev, P. Veličković. Quantifying the effects of enforcing disentanglement on variational autoencoders. *31st Annual Conference on Neural Information Processing Systems (NIPS)*, 2017.

[142] Y. Wang, D. M. Blei, and J. P. Cunningham. Posterior collapse and latent variable non-identifiability. *35th Conference on Neural Information Processing Systems (NIPS)*, 2021.

[143] J. He, D. Spokoyny, G. Neubig, and T. Berg-Kirkpatrick. Lagging inference networks and posterior collapse in variational autoencoders. *ICLR*, 2019.

[144] H. Fu, C. Li, X. Liu, J. Gao, A. Celikyilmaz, and L. Carin. Cyclical annealing schedule: A simple approach to mitigating KL vanishing. *NAACL*, 2019.

[145] S. R. Bowman, L. Vilnis, O. Vinyals, A. M. Dai, R. Jozefowicz, and S. Bengio. Generating sentences from a continuous space. *SIGNLL Conference on Computational Natural Language Learning*, 2016.

[146] A. Saxena, M. Prasad, A. Gupta, N. Bharill, O. P. Patel, A. Tiwari, M. J. Er, W. Ding, and C. T. Lin. A review of clustering techniques and developments. *Neurocomputing 267*, 2017.

[147] D. Xu and Y. Tian. A comprehensive survey of clustering algorithms. *Annals of Data Science 2*, 2015.

[148] S. Zhou, H. Xu, Z. Zheng, J. Chen, Z. Li, J. Bu, J. Wu, X. Wang, W. Zhu, and M. Ester. A comprehensive survey on deep clustering: Taxonomy, challenges, and future directions. *arXiv:2206.07579*, June 2022.

[149] A. E. Ezugwu, A. M. Ikotun, O. O. Oyelade, L. Abualigah, J. O. Agushaka, C. I. Eke, and A. A. Akinyelu. A comprehensive survey of clustering algorithms: State-of-the-art machine learning applications, taxonomy, challenges, and future research prospects. *Engineering Applications of Artificial Intelligence 110*, 2022.

[150] M. C. Hughes and E. B. Sudderth. Memoized online variational inference for dirichlet process mixture models. *Advances in Neural Information Processing Systems (NIPS) 26*, 2013.

[151] M. C. Hughes and E. B. Sudderth. BNPy: Reliable and scalable variational inference for bayesian nonparametric models. *Proceedings of the NIPS Probabilistic Programming Workshop*, 2014.

[152] E. Gorenskaia and P. J. Low. Methods for the analysis, interpretation, and prediction of single-molecule junction conductance behaviour. *Chem. Sci. 15*, 2024,.

[153] H. Rascón-Ramos, J. M. Artés, Y. Li, and J. Hihath. Binding configurations and intramolecular strain in single-molecule devices. *Nature Materials 14*, 2015.

[154] H. Häkkinen. The gold–sulfur interface at the nanoscale. *Nature Chemistry 4*, 2012.

[155] M. S. Inkpen, Z. F. Liu, H. Li, L. M. Campos, J. B. Neaton, and L. Venkataraman. Non-chemisorbed gold–sulfur binding prevails in self-assembled monolayers. *Nature Chemistry 11*, 2019.

[156] L. Kankate, A. Turchanin, and A. Gölzhäuser. On the release of hydrogen from the S-H groups in the formation of self-assembled monolayers of thiols. *Langmuir 25*, 2009.

[157] D. F. Crouse. On implementing 2d rectangular assignment algorithms. *IEEE Transactions on Aerospace and Electronic Systems, 52*, 2016.

[158] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors. SciPy 1.0: fundamental algorithms for scientific computing in python. *Nature Methods 17*, 2020.

[159] Y. Yang, D. Xu, F. Nie, S. Yan, and Y. Zhuang. Image clustering using local discriminant models and global integration. *IEEE Transactions on Image Processing, 19*, 19:2761–2773, 2010.

[160] A. Farhadi J. Xie, R. Girshick. Unsupervised deep embedding for clustering analysis. *PMLR, 48*, 2016.

[161] Z. Bing, Y. Meng, Y. Yun, H. Su, X. Su, K. Huang, and A. Knoll. DIVA: A dirichlet process mixtures based incremental deep clustering algorithm via variational auto-encoder. *https://doi.org/10.48550/arXiv.2305.14067*, 2023.

[162] M. A. Carbonneau, J. Zaïdi, J. Boilard, and G. Gagnon. Measuring disentanglement: A review of metrics. *IEEE Transactions on Neural Networks and Learning Systems (Early Access)*, 2022.

[163] K. Do and T. Tran. Theory and evaluation metrics for learning disentangled representations. *ICLR*, 2020.

[164] I. Loshchilov and F. Hutter. Decoupled weight decay regularization. *ICLR*, 2019.

[165] S. José Hanson and L. Y. Pratt. Comparing biases for minimal network construction with back-propagation. *Proceedings of the 1st International Conference on Neural Information Processing Systems (NIPS)*, 1988.

[166] M. T. González, S. Wu, R. Huber, S. J. van der Molen, C. Schönenberger, and M. Calame. Electrical conductance of molecular junctions by a robust statistical analysis. *Nano Lett. 6*, 2006.

[167] B. Xu. Modulating the conductance of a Au–octanedithiol–Au molecular junction. *Small 3*, 2007.

[168] Y. H. Kim, H. S. Kim, J. Lee, M. Tsutsui, and T. Kawai. Stretching-induced conductance variations as fingerprints of contact configurations in single-molecule junctions. *J. Am. Chem. Soc. 139*, 2017.

[169] F. J. Giessibl. Advances in atomic force microscopy. *Rev. Mod. Phys. 75*, 2003.

[170] G. Haugstad. *Atomic Force Microscopy: Understanding Basic Modes and Advanced Applications*. Wiley, 2012.

[171] L. Jiang, J. Weber, F. M. Puglisi, P. Pavan, L. Larcher, W. Frammelsberger, G. Benstetter, and M. Lanza. Understanding current instabilities in conductive atomic force microscopy. *Materials, 12*, 2019.

[172] Stanford Research Systems. *Model SR560 Low Noise Preamplifier, User Manual, Rev. 3.1*, 2023.

[173] C. Kitchin and L. Counts. *A Designer's Guide to Instrumentation Amplifiers, 3rd Edition.* Analog Devices, 2006.

[174] INA116 ultra low input bias current instrumentation amplifier. *Texas Instruments, SBOS034*, 1994.

[175] INA821 35μV offset, 7nV/$\sqrt{\text{Hz}}$ noise, low-power, precision instrumentation amplifier. *Texas Instruments, SBOS893D Rev. D*, 2020.

[176] AD8230 16V rail-to-rail, zero-drift, precision instrumentation amplifier. *Analog Devices, D05063-0-9/07, Rev. B*, 2007.

[177] M. Lis. Trade-offs between CMOS, JFET, and bipolar input stage technology. *Texas Instruments, SBOA355*, 2019.

[178] P. Horowitz and W. Hill. *The Art of Electronics, 3rd Edition.* Cambridge University Press, 2015.

[179] B. Hess, S. Bendaoud, and D. Terrazas. Super-beta input amplifiers: Features and benefits. *Texas Instruments, SBOA305*, 2020.

[180] L. Shi and A. Majumdar. Thermal transport mechanisms at nanoscale point contacts. *J. Heat Transfer 124*, 2002.

[181] H. Song, Y. Kim, Y. H. Jang, H. Jeong, M. A. Reed, and T. Lee. Observation of molecular orbital gating. *Nature 462*, 2009.

[182] A. Vezzoli. Mechanoresistive single-molecule junctions. *Nanoscale 14*, 2022.

[183] M. Valtr, P. Klapetek, J. Martinek, O. Novotný, Z. Jelínek, V. Hortvík, and D. Nečas. Scanning probe microscopy controller with advanced sampling support. *HardwareX 15*, 2023.

[184] J. Martin, T. Tritt, and C. Uher. High temperature seebeck coefficient metrology. *J. Appl. Phys. 108*, 2010.

[185] T. K. Johnson, J. A. Ivie, J. Jaruvang, and O. L. A. Monti. Fast sensitive amplifier for two-probe conductance measurements in single molecule break junctions. *Rev. Sci. Instrum. 88*, 2017.