Towards Single Molecule Imaging using Nanoelectromechanical Systems

Thesis by

Adam Patrick Neumann

In Partial Fulfillment of the Requirements for

the degree of

Doctor of Philosophy

CALIFORNIA INSTITUTE OF TECHNOLOGY

Pasadena, California

2020

(Defended April 16, 2020)

Adam Patrick Neumann
ORCID: 0000-0002-2961-7640

ACKNOWLEDGEMENTS

Doctoral theses like this one list a single author, but are the result of extensive collaborations and generous knowledge-sharing from many individuals. Because of this, I took care to mention specific contributions as frequently as feasible in the main text, but sometimes direct attribution is not possible or accidently overlooked, and I apologize to those for whom this applies.

Primary thanks are due to my thesis advisor, Professor Michael Roukes, who helped see my work through both during times of great progress and excitement and also through stretches of glacial progress. Such is the nature of scientific research, and what I appreciated the most was an unfettered chance to engage in this field at a high level, including all of its multifaceted challenges and opportunities. His support and guidance made this dissertation possible.

I am grateful that Michael put me in position to collaborate on a regular basis with Dr. Alexander Makarov and his R&D team (Drs. Maria Reinhardt-Szyba and Dmitri Boll) at Thermo Fisher in Bremen, Germany. They provided extensive resources, expertise, and encouragement that was critical for the success of the hybrid instrument and my research. I am also thankful for Dr. Makarov's invitation for me to visit the mass spectrometry factory in Bremen and obtain detailed training on the instrument operation, as well as helping provide some funding for my research.

A second critical component was my introduction by Michael to Professor Albert Heck and his lab in Utrecht, the Netherlands. I want to thank Drs. Joost Snijder and Michiel Van de Waterbeemd, Anja Boumeester, and Dr. Kyle Fort for welcoming me there and offering detailed training for native mass spectrometry.

A third critical collaboration was with Professor John Sader, who provided great training and encouragement for my interest in nanoscale sensor theory. I want to thank him for inviting me to Melbourne to do some AFM experiments, helping me put inertial imaging theory into practice with the data I collected at Caltech, and helping me refine a new theory I conceived for imaging molecules.

Here in the Roukes group at Caltech, I wanted to thank Drs. Warren Fon and Eric Sage for their extensive help with the hybrid system design and operation, Professor Selim Hanay for helping introduce me to NEMS-MS measurements, as well as Jarvis Li for his assistance in the lab. I also want to thank my SURF students, Marec Serlin and Suyee Win. Additionally, I wanted to thank Rachel Galimidi from Professor Bjorkman's lab, and Yi-Ju Chen and Stephanie Barnes from Professor Phillip's lab for giving me the space and training to purify ribosomes and bacteriophages for NEMS-MS experiments.

TA funding was essential for my studies, and I would like to thank Professor Niles Pierce and the ACM department for providing me the opportunity to serve as TA, and later Head TA for ACM 95/100. It was a unique experience to lead recitation sessions, and later review sessions for the entire class, for so many brilliant and motivated students.

I wish to thank Professors Robert Phillips, Jack Beauchamp, and Bill Clemons from my candidacy committee for their guidance and advice on narrowing down a research topic.

I also wanted to thank Linda Scott and Su Naing from the administrative staff, Tess Legaspi from the registrar, and Professor Doug Rees, Dr. Kate McAnulty, and Natalie Gilmore from the Dean's office for their extensive support.

I want to thank Professor Chinlin Guo for supporting my initial research at Caltech, as well as Professors Saul Teukolsky and Jonathan Butcher for providing excellent

undergraduate and master's research opportunities for me at Cornell. My interest in math and science started at a young age, and Mr. Byrd from AP Physics stands out to this day as an enthusiastic, inspirational teacher.

Finally, I am eternally grateful to my family for their love and support.

ABSTRACT

We incorporate nanoelectromechanical systems (NEMS) into a state-of-the-art commercial mass spectrometer (Q Exactive Plus with Orbitrap detection). This unique hybrid instrument is capable of ionizing molecules up to 4.5 MDa in their intact native state, isolating molecules of interest according to their mass-to-charge ratio, performing high resolution mass spectrometry (MS), and delivering those molecules to the NEMS. We use NEMS optimized for detecting the inertial mass of adsorbed species directly, which contrasts with indirect measurements of the mass-to-charge ratio performed with typical instruments. This unique form of mass spectrometry, NEMS-MS, with its single-molecule sensitivity, has promising applications to the fields of proteomics and native mass spectrometry, including deep proteomic profiling, single-cell proteomics, mass spectrometry-based imaging, or identifying viruses in their *in vivo* state.

We analyze intact *E. coli* GroEL chaperonin, a noncovalent 801 kDa complex consisting of 14 identical subunits. GroEL was sent to NEMS operated with the first two vibrational modes monitored in real time. Molecules physisorbing to the NEMS cause an abrupt shift in its resonance frequencies. The change in resonance frequencies is used to calculate the mass of each molecule. A mass spectrum is compiled with a main peak of 846 kDa, close to the expected value, and a secondary peak resolved near twice the mass of GroEL.

Measurements are then performed operating the first three modes simultaneously. Using a technique called inertial imaging, frequency shifts are used to calculate the first three mass moments: mass, position, and variance (size). This is used to distinguish between adsorbates arriving in a single, point-like distribution or a more extended distribution, thus demonstrating a rudimentary form of molecular imaging.

Two new theories are presented for analyzing frequency-shift data. The first approach offers a more streamlined approach for calculating the mass moments. This approach is used to improve the mass spectrum of the GroEL calculated using three-mode data, producing a main peak almost fully resolved at 805 kDa. An entirely different approach is presented that allows

for obtaining the mass density distribution of an adsorbed molecule (i.e., imaging) with a higher number of modes.

PUBLISHED CONTENT AND CONTRIBUTIONS

Sader J.E., Hanay M.S., **Neumann A.P.**, and Roukes M.L. "Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation." In: Nano Letters 2018; 18(3): 1608-14. doi: 10.1021/acs.nanolett.7b04301.

APN assisted with the theoretical formulation and participated in the writing of the manuscript.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACRONYMS AND ABBREVIATIONS

1D	One-dimensional
3D	Three-dimensional
AC	Alternating current
ADP	Adenosine diphosphate
AFM	Atomic force microscope
ag	Attogram (10^{-18} g)
AlN	Aluminum nitride
amol	Attomole (10^{-18} mole)
ATP	Adenosine triphosphate
CF	Conflat
CMOS	Complimentary metal-oxide-semiconductor
CSD	Charge state distribution
CsI	Cesium-iodide
Da	Dalton, unit of mass (1/12th of the mass of a C12 atom)
dB	Decibel
DC	Direct current
DR	Dynamic range
EMR	Extended mass range
ESI	Electrospray ionization
fg	Femtogram (10^{-15} g)
FTICR	Fourier transform ion cyclotron resonance
FWHM	Full width at half maximum
HCD	Higher energy collisional dissociation
IMS	Ion mobility spectrometry
K	Kelvin
KE	Kinetic energy
LC	Liquid chromatography

LIA	Lock-in amplifier
LPF	Low-pass filter
MALDI	Matrix-assisted laser desorption ionization
MPC	Multiproteoform complex
MS	Mass spectrometry
NEMS	Nanoelectromechanical systems
OFHC	Oxygen-free high thermal conductivity
PCB	Printed circuit board
PDF	Probability density function
PE	Potential energy
PEEK	Polyether ether ketone
PID	Proportional-integral-derivative
PLL	Phase-locked loop
ppm	Parts per million
PS	Power splitter
QP	Quadrupole
RF	Radiofrequency
RNA	Ribonucleic acid
SEM	Scanning electron microscope
SMA	Subminiature version A
S/N	Signal-to-noise ratio
TEM	Transmission electron microscope
ToF	Time of flight
UT85	Uniform Tubes, Inc. 0.0865” outer diameter
VLSI	Very large scale integration
zg	Zeptogram (10^{-21} g)

SYMBOLS

$\langle x^2 \rangle$	Variance of mass density distribution (second mass moment)
$\mathcal{A}$	Agreement between solution and data
a	Position of landing of adsorbed molecule (first mass moment)
$\mathcal{B}$	Smoothness of solution
c_n	Normalized fractional frequency shift for mode n
E	Young's modulus
f_n	Resonance frequency for mode n
I	Second moment of area (with respect to device length)
k_n	Eigenvalue for beam normalized to length 1 for mode n
L	Device length
M	Total mass of NEMS resonator
M_n^{eff}	Effective modal mass for mode n
m/z	Mass-to-charge ratio
N	Number of modes
Q	Quality factor
q	Ion charge
S	Entropy functional of unknown mass density distribution
S_{rel}	Entropy functional of unknown mass density distribution relative to bias
t	Device thickness
V_b	Bias voltage of resonator
V_d	Drive (gate) voltage of resonator
x	Dimension along device length
w	Device width
α_n	Numerical constant for finding effective modal mass for mode n
$\delta(x)$	Dirac delta function
δ_{nm}	Kronecker delta

δf_n	Change in resonance frequency for mode n
δm	Added mass of adsorbed molecule (zeroth mass moment)
λ	Regularization parameter
λ_n	Eigenvalue for mode n
$\mu(x)$	Normalized mass density distribution of adsorbed molecule
$\mu_{ad}(x)$	Mass density distribution of adsorbed molecule
$\Pi(x)$	Prior estimate (or bias) for unknown mass density distribution
σ	Standard deviation
$\sigma(\tau)$	Allan deviation for time τ
σ_m	Mass resolution
Φ	Electrical potential
$\phi_n(x)$	Mode shape (eigenfunction) for mode n
Ω	Domain of integration
ω_n	Angular resonance frequency for mode n

Chapter 1

INTRODUCTION

This thesis advances development of a platform for measuring the mass of single molecules, as well as other physical attributes, such as their charge and the image they project on a flat surface, using nanoelectromechanical systems (NEMS). In NEMS-based mass spectrometry, particles are directed at nanometer scale mechanical structures that have been etched into silicon wafers. As the particles land one at a time on a device, the resonance frequency of each vibrational mode abruptly changes. A model is then used to calculate the mass or image of each particle that would induce such a set of frequency shifts. The devices are engineered to operate both as actuators and sensors through transduction between the electrical and mechanical domains: electrical energy is used to drive device motion mechanically, and mechanical motion in turn produces an electrical signal which can be monitored. NEMS-based sensors have captivated significant research attention due to their extreme responsivity and the possibility of co-opting very large scale integration (VLSI) technology in use by the semiconductor industry to achieve mass production.

Fully realizing the potential of NEMS devices in the domain of mass sensing will eclipse the ultimate limits of several figures of merit of traditional mass spectrometry technology, including mass resolution, mass range, linear dynamic range, and limit of detection. Traditional mass spectrometry works by separating or identifying particles based on their mass to charge ratio and is a cornerstone of analytical laboratories studying a wide variety of compounds, including biomolecules. Surpassing the state of the art with a fundamentally new technology, such as with NEMS transducers, will enable the detection of

increasingly minor changes to large molecules, such as individual proteins that have undergone post-translational modification, or the rapid fingerprinting or cataloging of increasingly complex mixtures, such as rare biomarkers in the human plasma proteome, or the analysis of an increasingly small sample volume, such as the contents that could be extracted from a single cell. This chapter will review traditional approaches to mass spectrometry and their limitations, with a special emphasis on biological applications, such as proteomics and structural biology, as well as operation of NEMS-based mass sensors and how they might exceed those limitations.

1.1 Mass spectrometry

Conventional mass spectrometry is a technique able to infer the mass of small particles to identify or analyze an unknown sample. This is accomplished by charging a molecular species, manipulating the ions with the use of electromagnetic fields, and detecting their motion. Instruments have been constructed that can distinguish species with different mass to charge ratio, or *m/z*, from one another based on different spatial paths they take (sector instruments) [19], different times to travel a path (time of flight) [23], stability or instability through an ion trap (3D quadrupole trap linear or quadrupole filter) [4], or frequency of continuous motion in an ion trap (Fourier transform ion cyclotron resonance or Orbitrap) [11, 24]. Ion signal versus *m/z* is reported with the result termed a mass spectrum, although "*m/z* spectrum" might be a more accurate descriptor. This thesis will include data involving both *m/z* spectra and true inertial mass spectra, with the former generated with conventional mass spectrometry techniques, and the latter generated using NEMS devices. However, the term mass spectrum will be used to refer to both types of data in order to keep with historical convention, since the nature of the data will always be unambiguous.

Common instrument configurations. A brief discussion of the major *m/z* separation methods is given below.

Time of flight (TOF) – ions are accelerated with an electric field. The time required for an ion to traverse a drift tube to the detector is proportional to $(m/z)^{-1/2}$, with the constant of proportionality dependent on instrument geometry and electric potential, and determined through calibration.

Aside from time of flight mass spectrometry or sector instruments, other instruments typically rely on ion trapping. According to Earnshaw's theorem (1842), a DC electric field is not sufficient to keep a charged body in stable equilibrium. Ion trapping makes use of: (1) a DC and an RF electric field, used by the quadrupole trap, (2) a DC electric field and an orbiting particle, used by the Orbitrap, or (3) a DC magnetic field, an electric field, and an orbiting particle, used by Fourier transform ion cyclotron resonance (FTICR). Methods (1) and (2), relying only on electrical potential set on some conducting surfaces in the instrument $\Phi = \Phi_1, \Phi_2, ...$, will have Φ satisfy Laplace's equation $\nabla^2\Phi = 0$ everywhere else inside the trap.

3D quadrupole trap – this utilizes an RF electric field with field strength proportional to distance from a straight line in order to focus the ions toward the center of the trap (Figure 1a) [4]. The potential for such a configuration works out to:

$$\Phi = \frac{U + V\cos\omega t}{a^2}\frac{(x^2 - y^2)}{2}, \qquad (1)$$

where U and V are the DC and RF voltages, and x and y are coordinates perpendicular to the focusing axis. Based on the choice of DC voltage and RF voltage and frequency, ions within a certain range of *m/z* will either maintain a stable trajectory through the trap and reach a detector, or become unstable and fail to reach the detector. Physically, the Paul trap is constructed using a hyperbolic ring electrode and a pair of hyperbolic end electrodes as shown in Figure 1b. Equipotential surfaces have a saddle shape as shown in Figure 1c. A mechanical analog of an ion in this potential would be a ball that falls down if the saddle is stationary, but could remain in motion without falling if the saddle is rotated at the correct frequency.

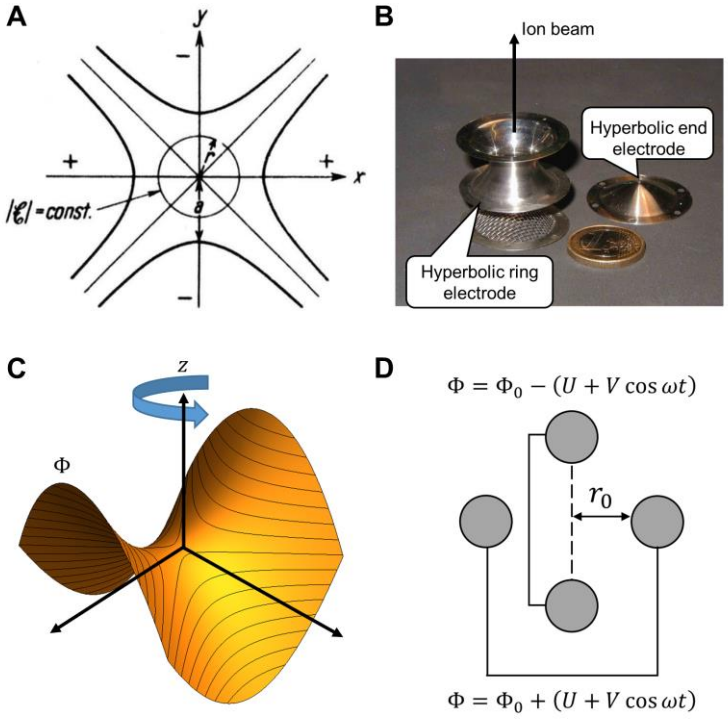


Figure 1. 3D quadrupole trap and 2D quadrupole filter. (A) A quadrupole trap uses an RF electric field whose magnitude $|\xi|$ is proportional to some distance r from a straight line. From [4]. **(B)** It is constructed using a hyperbolic ring electrode and two hyperbolic end electrodes. Adapted from [10]. **(C)** Mathematica plot showing that equipotential surfaces of the Paul trap are shaped like a saddle. **(D)** A linear quadrupole filter can be constructed using two pairs of circular rods. The DC potential across these rods U and the RF potential V and frequency ω determine if ions will remain radially stable. Adapted from [16].

2D quadrupole filter – this is constructed using two pairs of circular rods, with the two pairs operated with a DC and RF voltage U and V added to or subtracted from an overall DC potential Φ_0 as shown in Figure 1d. A more ideal quadrupole filter will use hyperbolic shaped rods, similar to the hyperbolic geometry used in the 3D Paul trap. However, circular rods produce a potential similar enough to that shown in Equation 1 [25], but are much simpler to construct. As discussed in Appendix A, transmission of ions through these quadrupoles will occur for a range of *m/z* values based on the values of U, V, and the RF frequency ω.

Orbitrap – ions orbit in a spindle-shaped trap with an internal potential given by

$$\Phi(r,z) = \frac{k}{2}\left(z^2 - \frac{r^2}{2}\right) + \frac{k}{2}R_m^2 \ln\left(\frac{r}{R_m}\right) + C, \tag{2}$$

where z and r are the axial and radial cylindrical coordinates, k is the field curvature, R_m is the characteristic radius, and C is a constant [11]. Equation 2 represents the sum of a quadrupolar potential that confines the ions axially and a new logarithmic potential that serves to provide orbital trapping. The internal potential follows Equation 2 from Laplace's equation if electrodes are constructed with the geometry matching equipotential surfaces $\Phi = \Phi_1, \Phi_2$. Such surfaces are shown in Figure 2a along with the final Orbitrap geometry in Figure 2b. Ions entering the trap have a rotating elliptical orbit in the radial direction, and harmonic motion in the axial direction given by

$$\mathrm{z(t)} = z_0 \cos\omega t + \sqrt{2E_z/k}\sin\omega t\,; \qquad \omega = \sqrt{k(q/m)}\,, \tag{3}$$

where E_z is an axial kinetic energy term and m/q is the mass to charge ratio of the ion. The axial frequency ω depends only on mass to charge ratio and a constant related to trap geometry. Orbiting ions will create an image current on the outer electrodes which are engineered to only collect this axial motion contribution. The detection scheme is shown in Figure 2c.

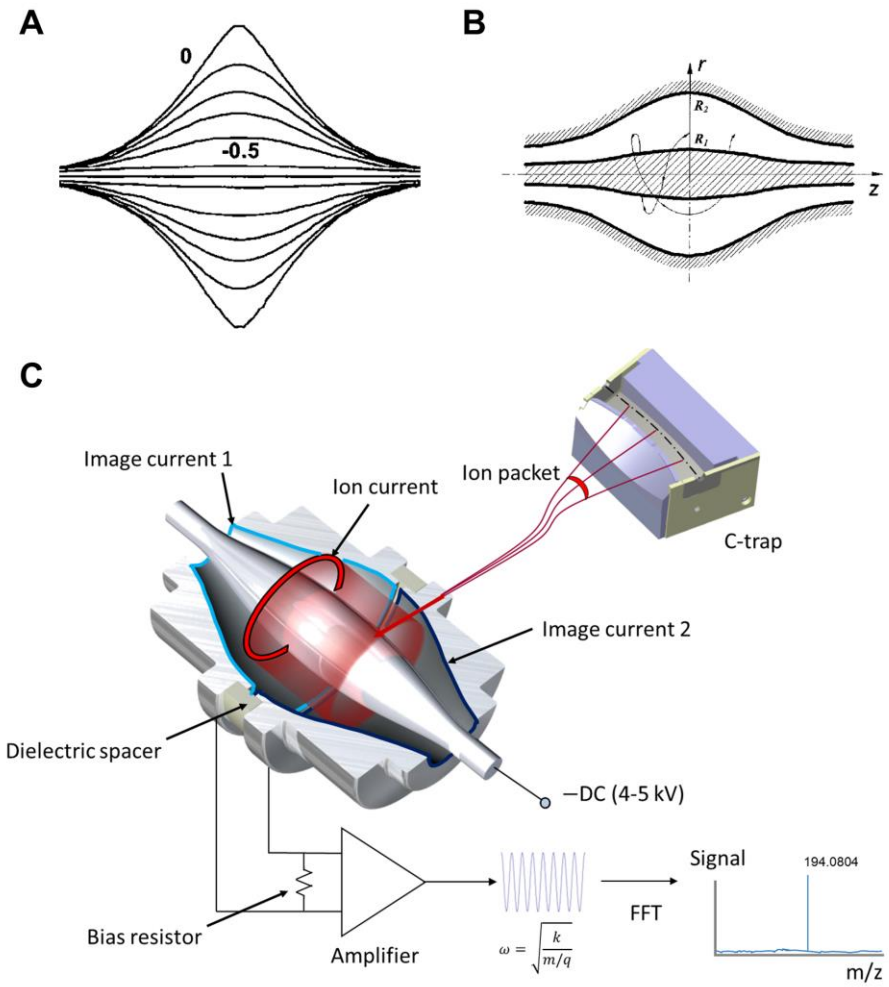


Figure 2. Orbitrap design and detection scheme. (A) Equipotential surfaces of the quadro-logarithmic potential given in Equation 2. From [3]. **(B)** Oribtrap design with inner and outer electrodes engineered to follow the geometry of two of the equipotential surfaces. From [11]. **(C)** Detection scheme for Orbitrap instruments. Once an ion packet is injected into the Orbitrap, it orbits the central electrode with elliptical radial motion and harmonic axial motion. The frequency of axial motion of the ion packet depends only on mass to charge ratio and is detected based on the differential image current formed between the two outer electrodes. The Fourier transform of this detected current is used to construct the mass spectrum. Adapted from [18].

FTICR – Ions are trapped using a large, static magnetic field with a small quadrupolar DC electric field for axial confinement. The ions undergo cyclotron motion with the frequency proportional to $(m/z)^{-1}$. A Fourier transform of the detected image current is used to produce a mass spectrum [24].

Ionization methods. Biomolecular ions are charged with the ionization techniques electrospray ionization (ESI) [26] or matrix-assisted laser desorption ionization (MALDI) [23, 27]. These techniques, which ionize peptides or proteins without fragmenting them, are referred to as soft ionization. In ESI, a syringe pump pushes the sample through a glass needle coated with conductive material. A large voltage (1−2 kV) is applied to the needle, causing it to emit a jet of charged liquid droplets that evaporate, leading to a process that results in charged ions. The molecule takes on multiple charges during this process, which is a beneficial feature for analyzing large molecules like biomolecules, as it lowers the *m/z* value, making it more accessible for typical instrumentation. It can also produce a variable number of charges, known as a charge state distribution (CSD). This variable charging effect, combined with the mass variability naturally associated with isotopes, produces multiple *m/z* peaks even for pure samples. In MALDI, the sample is mixed with an organic molecule (the matrix), which absorbs radiation from a laser beam. The matrix transfers the energy to the sample, typically forming positive ions with a single charge.

Deconvolution. Since the charge state of each peak in a mass spectrum is unknown, a procedure termed deconvolution must be performed to reconstruct mass. If peaks unambiguously belong to a single species, the mass and charges can be determined using a

system of equations [28]. However, assignment becomes difficult when sample heterogeneity causes several charge state distributions to overlap in *m/z*; in this case, Bayesian approaches requiring a priori knowledge of the sample have been proposed [29].

Sample pre-fractionation and tandem MS. In order to reduce sample complexity, additional methods for pre-stratifying molecules (*i.e.*, separation orthogonal to *m/z*) may be used such as liquid chromatography and ion mobility spectrometry. Fragmentation methods such as electron transfer dissociation (ETD), electron capture dissociation (ECD), collision-induced dissociation (CID), UV photodissociation, and higher energy collisional dissociation (HCD) can be included between *m/z* separation methods as well in order to simplify the final spectrum enough to identify the sample components; this is known as tandem MS [30]. Popular combinations include triple quadrupole and quadrupole-ToF hybrids; alternatively, some analyzers such as the Paul trap can perform multiple rounds of MS [18].

1.2 Ion mobility spectrometry

Ion mobility spectrometry (IMS) is an analytical technique used to separate and identify ionized molecules in the gas phase based on their mobility in a carrier buffer gas [31]. The separation mechanism, involving collisions with the background gas, differentiates analytes by the rotationally-averaged collisional cross-sections. When coupled with mass spectrometric analysis, IMS provides an additional level of information that enhances the characterization of biomolecular species – especially in the case of high-mass macro-molecular complexes.

For example, ligand binding to biomolecular complexes can sometimes induce conformational changes that affect the collision cross section of the complex, and these changes can be monitored via ion mobility measurements [32]. IMS can also be used to

monitor changes in the stability of a complex upon ligand binding. In order to do this, ions are stored in an ion trapping region where they undergo ion activation (e.g. heating) which will destabilize the complex and increase the ion mobility drift time; ligand binding can reduce this effect as a result of increased stability [33]. In both cases, IMS is useful for monitoring or discovering complex-ligand interactions that occur when the complex is in its native state; such information could not be obtained otherwise.

IMS has less orthogonality compared to chromatography methods due to the strong correlation between analyte mass and size [34], but remains useful as a high-speed technique that increases the maximum number of resolvable analytes (referred to as peak capacity [35]), in addition to providing better signal-to-noise and increased dynamic range. However, this increased analytic power comes at the cost of ion transfer loss, reducing the limit of detection for rare sample components.

1.3 Mass spectrometry performance metrics

Different aspects of a mass spectrometer's performance are quantified in their figures of merit, many of which are a function of *m/z*:

Mass range – upper, and sometimes lower, limit of *m/z* ratios amenable to analysis

Resolving power – $M/\Delta M$, where ΔM is the difference in *m/z* between adjacent peaks according to a specified overlap criterion, and M is the average *m/z* ratio of those two peaks; sometimes referred to as mass resolution

Mass accuracy – ratio of *m/z* measurement error divided by true *m/z* and stated in ppm

Dynamic range – range, *i.e.* number of molecules, over which high mass accuracy (typically 5 ppm) can be achieved in a single scan.

Detection limit – lowest quantity of analyte that can be distinguished from the system noise

Sensitivity – The signal response to a particular quantity of analyte normalized to the amount of analyte giving rise to the response

Throughput – mass spectra collected per second; sometimes referred to as scan rate.

Provided the mass range can cover the sample of interest, the resolving power is the most important metric, as it must be high enough to resolve adjacent peaks. Once that has been accomplished, mass accuracy must be high enough for sample identification. Often, increasing the resolving power comes with a tradeoff or other constraints. For example, the resolving power of FTICR increases with magnetic field strength, with the highest using an extremely costly 21-tesla magnet [36]. Similarly, the resolving power of ToF can be increased with longer flight path, with the longest attempt currently at 20 ft [37]. Both FTICR and the Orbitrap obtain a mass spectrum by Fourier transforming the detected current; longer acquisition time therefore increases the resolving power but at the cost of longer measurement time. Certain applications require a minimum throughput in order to continuously measure the output of a liquid chromatography column; in this case, resolving power must sometimes be decreased even if a higher capability is available [37].

Orbitrap technology, used in this thesis, has become a popular alternative FTICR for many applications because it can be realized as a cost effective benchtop instrument compatible with common tandem MS configurations [30]; its resolving power as a function of m/z compares favorably with FTICR as shown in Figure 3. There are some limitations that come with the convenience of the Orbitrap. Compared with FTICR, which is capable of maintaining ion cyclotron motion for several minutes to achieve extremely high resolving

power [24], Orbitrap measurements are limited to a few seconds due to orbital decay or decoherence due to dissociation or collisions with gas molecules or other ions [18]; this problem is especially severe for large molecules, limiting the upper mass range [38]. The dynamic range is also limited by Coloumb repulsion effects present when large numbers of ions are stored in the C-trap prior to injection into the Orbitrap [14].

Two major applications of mass spectrometers involving biomolecules include proteomics (characterizing the proteins encoded by a genome) and native mass spectrometry (characterizing proteins or protein complexes in their intact state). The most challenging aspects of these applications will be discussed in the following sections in order to highlight opportunities for disruption with NEMS devices. In order to compare instruments built for similar tasks, specifications for instruments used for proteomics is shown in Table 1.

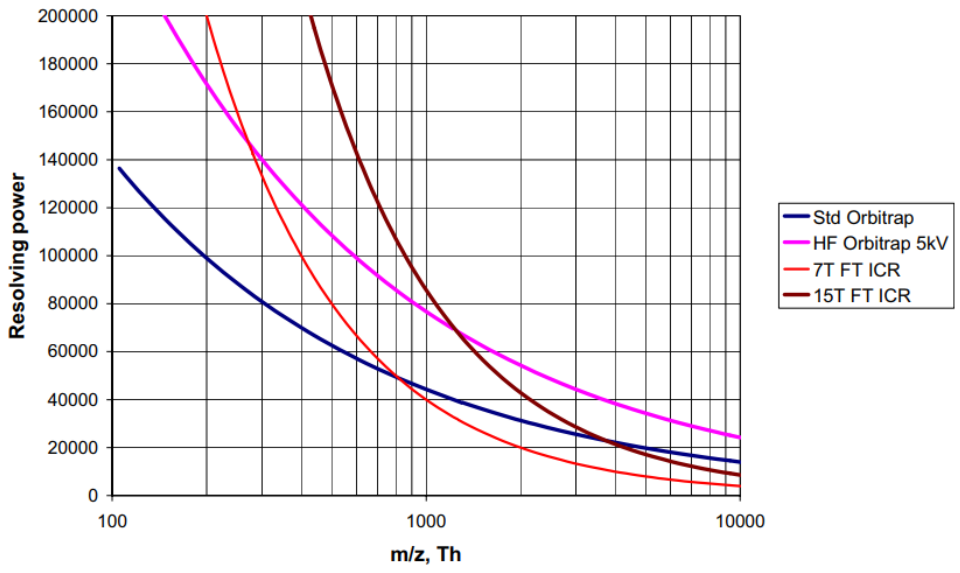


Figure 3. Resolving power as a function of m/z for some Orbitrap- and FTICR-based instruments. Higher resolving power enables instruments to resolve peaks that are closer and closer together in a mass spectrum. This figure-of-merit typically diminishes with higher m/z (or Thomson, Th). Shown are the resolving power for a standard and high field (HF) Orbitrap, which have a resolving power proportional to $(m/z)^{-1/2}$, and for 7T and 15T magnetic field strength FTICR instruments, which have a resolving power proportional to $(m/z)^{-1}$. Adapted from [9].

Table 1. Typical figures of merit for modern high resolution mass spectrometry instruments used for proteomics. Mass range, resolving power, mass accuracy, and detection limit are reported using company product literature for representative instruments for each technology. Here, ToF refers to Waters Vion IMS QToF, FTICR refers to Bruker solariX 15T, and Orbitrap refers to Thermo Q Exactive or Q Exactive EMR. Detection limit is reported in terms of a quantity added to an online LC column that feeds into an ESI unit, or, equivalently, the number of molecules. Dynamic range values are representative for the general class of technology.

Figure of Merit	ToF	FTICR	Orbitrap
Mass range	*m/z* 20–64,000	*m/z* 100–10,000	Standard: *m/z* 50–6,000 EMR: *m/z* 350–20,000
Resolving power	50,000 at *m/z* 1000	$>10^7$ at *m/z* 100 10,000 at *m/z* 10,000 [9]	>240,000 at *m/z* 200 17,500 at *m/z* 10,000
Mass accuracy	< 1 ppm at *m/z* 500	250 ppb	Standard: < 3 ppm EMR: < 10 ppm
Detection limit	5 fg verapamil (7:1 S/N) = ~10^7	100 amol = 6×10^7	500 fg buspirone (100:1 S/N) = ~10^8
Dynamic range	Few hundred [12-14]	Few thousand [21, 22]	5000 [14]

1.4 Proteomic mass spectrometry

The development of mass spectrometers combining high resolving power, high mass accuracy, and high throughput ushered in the era of large-scale proteomics studies. The term proteome refers to the complete set of proteins expressed by a genome. In contrast with the genome, which remains relatively stable, the proteome is highly dynamic, with the relative number of proteins changing as a result of a cell's regulation of gene expression in response to its environment. Biological functions at the cellular and molecular level are carried out by intact proteoforms and multiproteoform complexes. Proteoforms are the wide variety of different molecular forms that protein products of a single gene can assume, such as genetic variation, alternative splicing of RNA transcripts, and post-translational modifications

(PTMs) [39]. Individual proteoforms are subsequently assembled by covalent and non-covalent interactions to form multiproteoform complexes (MPCs) that enable the immense spectrum of cellular functions. Therefore, an inventory of each proteoform, MPC, and their concentration and interaction network in a given time frame under normal and abnormal conditions will lead to a detailed understanding of cellular function and disease. Unfortunately, undertaking such a task invites considerable complexity – while humans have 30,000 to 40,000 genes encoding up to 40,000 proteins, the number of distinct proteoforms is more than 2,000,000 [40].

The majority of proteomic analyses utilize a "bottom-up" approach for protein identification, in which proteins are enzymatic digested into peptide fragments small enough to be unambiguously identified by the mass spectrometer, then bioinformatics is used to deduce the original intact molecules. Identification of small peptides is possible using the *m/z* measurement obtained in a mass spectrometer, assuming its resolving power and mass accuracy is high enough. This is because small fragments are singly charged and only consist of a limited number of possible elements. Smaller amino acid fragments are significantly easier to identify, but harder to piece together using bioinformatics. The Journal of the American Society for Mass Spectrometry author's guidelines requires a mass accuracy of 34 ppm at 118 m/z but an accuracy of 0.018 ppm at 750 m/z for unambiguous identification of an unknown amino acid fragment composed of the elements C, H, O, and N. As shown previously in Table 1, modern instruments are capable of around 1 ppm accuracy in this mass range. Fragmenting the millions of proteins averaging 400 amino acids long, at 100 Da per amino acid [41], however, would lead to considerable peak overlap in this mass range. Pre-fractionation techniques such as 2D electrophoresis or online liquid chromatography are

often necessary, at the cost of sample dilution, which can render relatively rare components undetectable.

Recently, using these techniques, a draft map of the human proteome was compiled [42], using mass spectra corresponding to proteins encoded by ~17,000 genes, covering ~84% of the human genome. The proteome of human HeLa cancer cells was also quantified [43], identifying 10,000 distinct protein groups with a concentration dynamic range spanning up to seven orders of magnitude. Such deep preotemic coverage of HeLa cells was possible using large scale experiments with isolated cell clones occurring over weeks. Ordinarily, low abundance proteins may not be detected as their isotope distributions are obscured by components present with higher prevalence [44]. The human plasma proteome consists of proteins with concentrations spanning 10 orders of magnitude as illustrated in Figure 4; only a few dozen high abundance proteins are used for routine clinical diagnostics due, in part, to

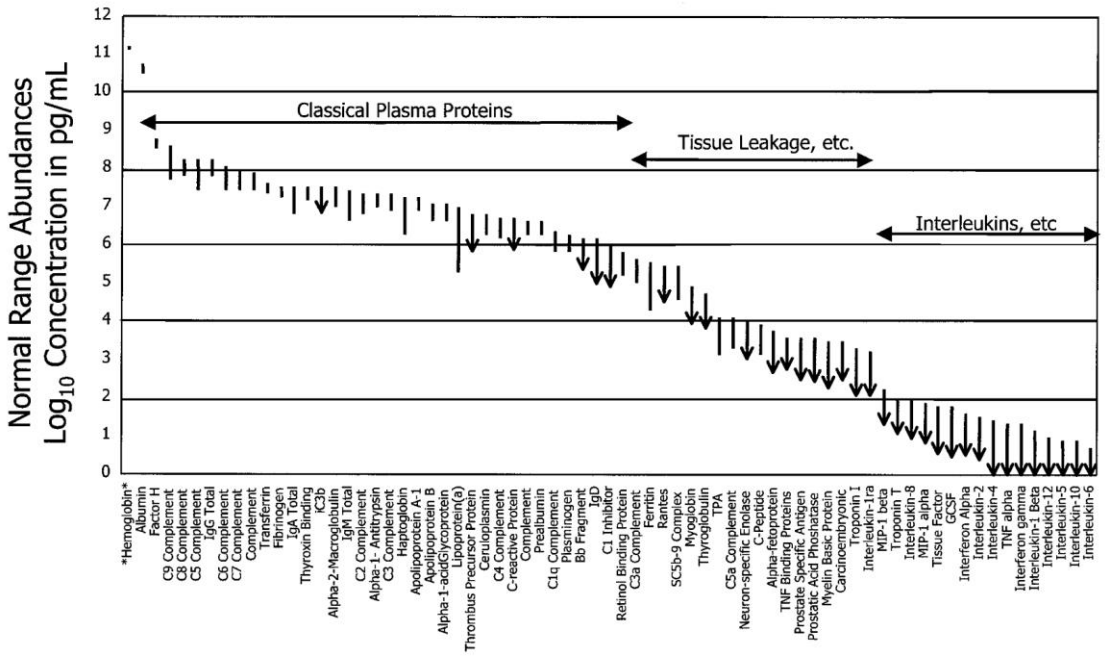


Figure 4. Concentration dynamic range of the human plasma proteome. Human plasma is a common clinical specimen and represents the deepest accessible version of the human proteome. Ten orders of magnitude in concentration separate albumin from the rarest proteins, but only a few dozen are used clinically. From [5].

the demand for consistent detection in a realistic setting [5].

These studies demonstrate that bottom-up proteomics studies can effectively create an inventory of expressed proteins, but cannot directly reveal the nature of intact proteoforms and MPCs. Although they enable critical structural characteristics and protein-protein interactions enabling proteoform and MPC function to be hypothesized, these hypotheses cannot be directly confirmed. Accordingly, "top-down" analyses have emerged [45]; these begin with intact proteoforms and MPCs. However, present top-down MS methods are severely challenged by the complexity of MPC heterogeneity, and addressing this requires laborious and often problematic preparatory protocols to avoid "averaging" over species. These protocols can completely remove the rarest analytes – often the ones that are especially significant.

1.5 Native mass spectrometry

Recently, it has been shown that noncovalent interactions, including higher order protein structure and weak binding between molecules, can be preserved as analytes are transported from the native cellular environment to the gas phase and then onto subsequent MS analysis. This enables direct elucidation of intact protein complexes and protein assemblies and, thereby, information on subunit stoichiometry, binding partners, protein complex topology, protein dynamics, and even binding affinities. Cumulatively, this method has been termed native mass spectrometry [46]. Carrying out native MS analysis involves exchanging the native buffer solution to one with similar ionic strength and pH that preserves the original molecular configuration, but which is volatile. This allows for evaporation and minimizing the formation of salt adducts in order to be compatible with MS. The solution is ionized using nano-ESI (ESI through a 1-10 micron diameter needle tip at nanoliter/min flow

rates), and instrument electronics and ion optics are modified to handle large *m/z* analytes. Achieving the goals of native MS is not straightforward, as protein complexes assemble with a high proportion of labile, noncovalent interactions. Despite this challenge, membrane proteins [47], ribosomes [48, 49], and 18 MDa viruses [7] have been successfully elucidated.

Mass spectra obtained with native MS feature a complex charge state distribution due to a variable number of charges being incorporated into the molecule during nano-ESI (Figure 5). Determining mass requires deconvolving this series of peaks in the *m/z* spectrum, so each charge state peak must therefore be distinct or the analysis will fail. This in turns requires both high resolving power at high *m/z* values, as well as a relatively pure (homogeneous) sample [50-52]. Typically, this is accomplished with a targeted protein of interest that has been expressed at high levels in a genetically engineered culture, where a

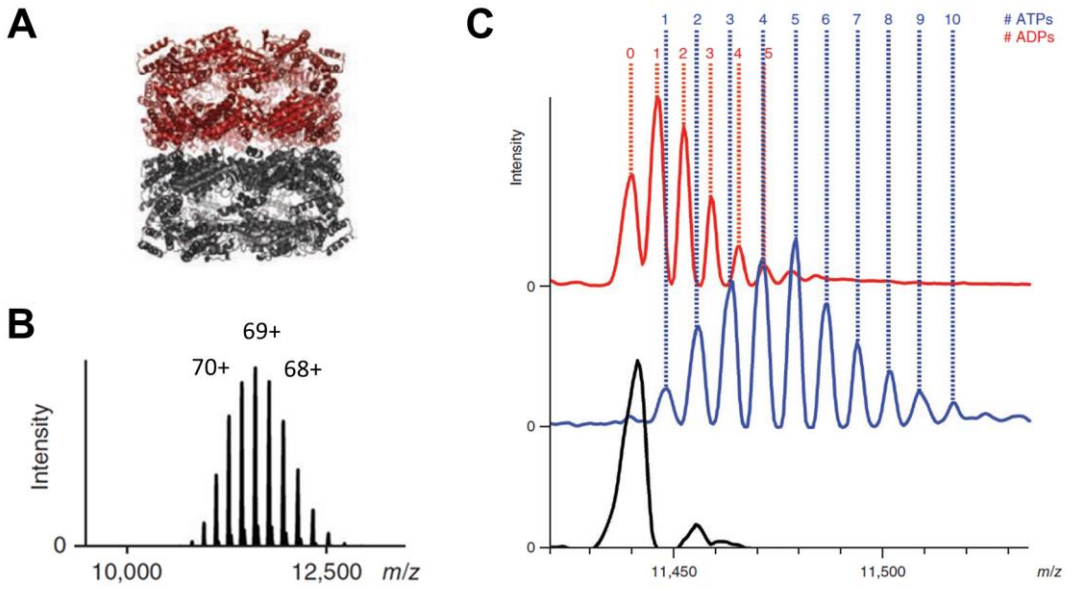


Figure 5. Native mass spectrometry. (A) Crystal structure of E. coli GroEL, an 800 kDa protein assembly consisting of 14 identical monomers. **(B)** GroEL produces a mass spectrum with a charge state distribution produced from the same molecule associated with a variable number of charges. Each assigned charge state (68+, 69+, 70+) multiplied by the *m/z* value for each peak recovers the intact mass of 800 kDa. **(C)** High resolution instruments can distinguish ADP and ATP molecules (~500 Da) co-incubated with GroEL, providing information on binding dynamics. Adapted from [6].

large amount of sample can be produced in order to survive multiple filtration and purification stages. The latter often must be developed *de novo* for each new molecule of interest. This process thus does not lend itself to large-scale experimental analysis and, especially, the discovery of rare molecular complexes in an unknown, complex mixture such as typically obtained via a proteomic workflow.

The applications of Native MS are vast. For example, screening millions of interacting compounds to target specific proteins is critical to development of new pharmaceuticals. These can involve antibody isoforms (proteins encoded from the same gene with alternative splicings or post-transcriptional modifications) with masses exceeding 1 MDa. Native MS can also provide information about topology and dynamics that is important for structural biology; for example, the locations and nature of ligand binding sites [53-56], and information about membrane protein complexes [47]. By contrast, top-down MS, which involves molecular dissociation and subsequent analysis of the resulting molecular fragments, can generate sequence and identity information for monomeric proteins, but provides limited applicability for analyzing large, noncovalent protein complexes. In part, this limitation is instrumental: measurement of large (greater than 100 kDa) protein complexes is difficult with conventional MS systems. However, existing MS systems also typically do not have sufficient sensitivity to resolve individual protein complexes, so purification to create a homogeneous population of analytes becomes essential. Otherwise, the analysis of heterogeneous populations of protein complexes in more typical experiments will yield an "average" over differently assembled complexes – conflating, rather than stratifying, the variety of species present within the sample.

Despite these advances, native mass spectrometry faces several important limitations, the most important being the need for homogeneous samples. As discussed previously, intrinsic to the technique of mass spectrometry is the requirement to calculate mass via deconvolution of a series of peaks formed from individual charge states. Mass assignment then fails if sample heterogeneity causes peaks to broaden enough that they overlap. Heterogeneity can arise in samples with a variable number of water molecules or salt adducts due to incomplete solvent evaporation, as well as heterogeneous mixtures with small mass variations or substoichiometric components. Each of these cases will be discussed briefly.

Adduct formation due to incomplete desolvation. Nano-ESI forms smaller droplets that evaporate more efficiently, allowing for the use of aqueous solvents that help preserve structure. However, this process of evaporation becomes less efficient for huge particles with more sites to interact with water molecules or ammonium ions. Therefore, for very large particles, even extremely pure samples will still carry a variable number of water molecules or salt ions, resulting in broader peaks [50]. Currently, desolvation limits the largest detectable complex in native mass spectrometry to around 18 MDa [7] (Figure 6a).

Heterogeneous mixtures. Separation of large heterogeneous mixtures proves more challenging. Peaks for the 30S small ribosomal subunit can be resolved, but not for the 50S large ribosomal subunit (Figure 6b). This is due to loss of the L10/$(L7/L12)_4$ stalk complex from some of the 50S large ribosomal subunits, so the 50S units that are detected are too heterogeneous for the peaks to be distinguishable [15]. In possibly a more extreme case of heterogeneity, small RNAs of unknown, random length are incorporated during assembly of the Hepatitis B virus, preventing charge state and mass assignment for the entire, intact virus (Figure 6c) [20]. Researchers could only estimate the mass at between 5,000 – 6,500 kDa,

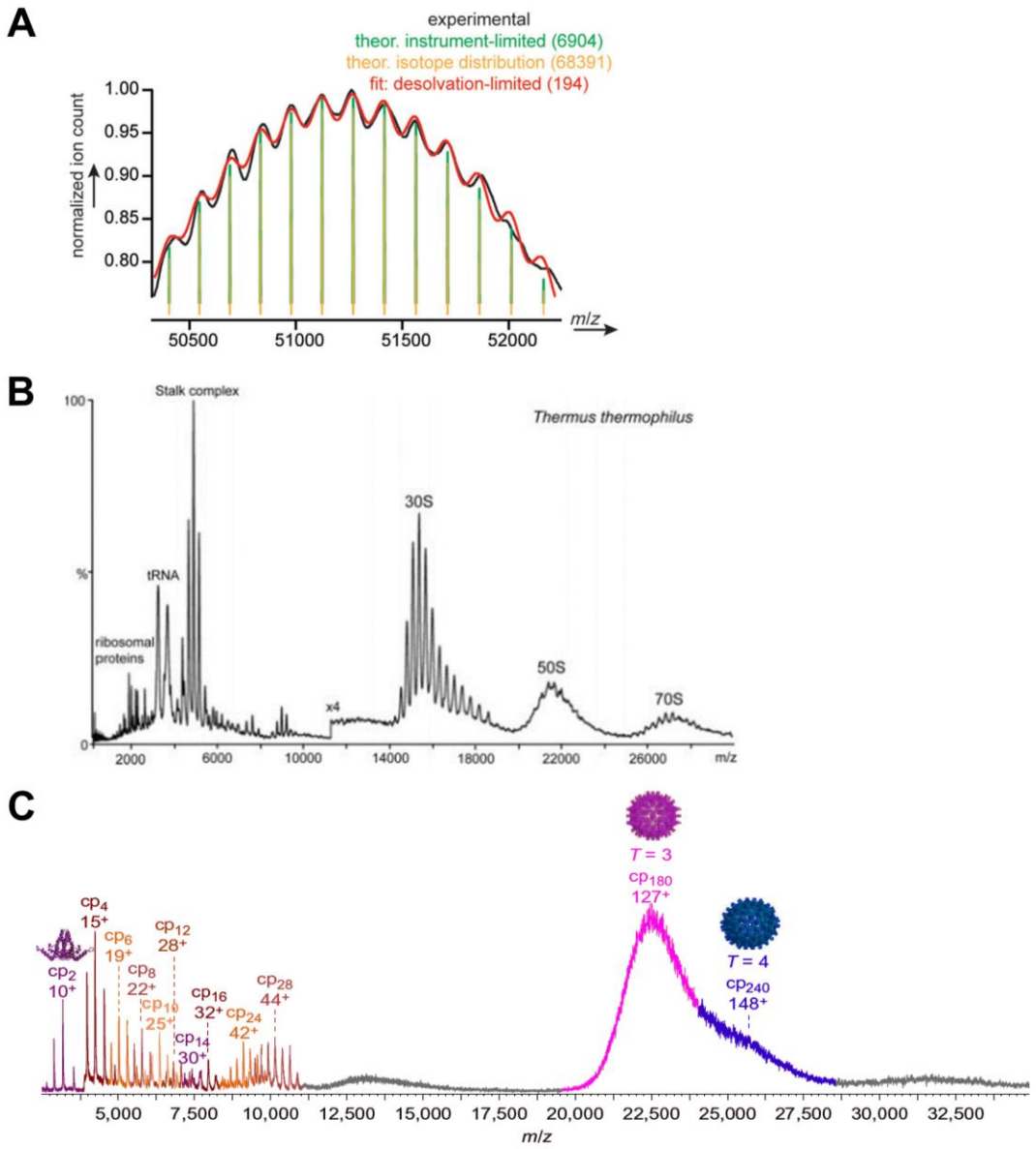


Figure 6. Limitations of native mass spectrometry. Heterogeneity broadens peaks, which can prevent charge state assignment and deconvolution to determine mass. Heterogeneity can be observed in samples with **(A)** incomplete desolvation, resulting in complexes with variable water and ammonium ions [7], **(B)** loss of the stalk complex from some of the 50S ribosomal subunits [15], and **(C)** incorporation of random small RNAs from the cellular environment during assembly of Hepatitis B virus [20]. Oligomers of up to 28-mers (~440kDa) and intact capsids (~2.5-4 MDa) are resolved. The number of monomers and the charge of the most abundant charge state is indicated above the oligomer peaks. The different charge-state distributions are color-coded. Intact HBV capsids (magenta/purple traces) do not show charge-state resolution due to low activation energies and sample heterogeneity; the sample contains complete capsid assemblies and those lacking one or two dimers.

demonstrating that, without charge state assignment, mass values are highly uncertain. In contrast, the capsid shell proteins engineered to self-assemble in vitro form two homogeneous structures with clearly resolvable peaks.

*Qualitative nature of data***.** Finally, it should be noted that peak size in native mass spectrometry is related to ionization, transmission, and detection efficiency, which can differ between different species; because of this, quantitative information is difficult to obtain [57].

1.6 Overview of single molecule NEMS-MS and inertial imaging

Over the past two decades, mass measurements using nanomechanical devices have systematically improved to the point where they now offer the capability for a new form of mass spectrometry. Upon adsorption onto the surface of a NEMS resonator, an individual molecule will downshift the device's resonant frequency by a small, but measurable, amount. The small size of NEMS makes them extremely sensitive to the added mass of adsorbed particles [58-63], and this has enabled mass detection of individual proteins [8, 64], nanoparticles [65], large biomolecules [66, 67], and individual atoms [68-71].

Details on the operation of NEMS devices for real-time single-molecule mass detection have been described elsewhere [8, 64, 72]; critical aspects will be discussed here and the next few sections for convenient reference. Upon its adsorption onto a nanomechanical resonator, typically a cantilever or doubly-clamped beam, a point analyte with mass δm – a single molecule, molecular complex, or nanoparticle – induces a downshift in resonant frequency of the resonator expressed by the formula

$$\frac{\delta f_n}{f_n} = -\frac{\delta m}{2M_n^{eff}}\phi_n(a)^2. \qquad (4)$$

The resulting fractional frequency shift, $\delta f_n/f_n$, is proportional to the fractional mass change, $\delta m/M_n^{eff}$. Here, f_n is the resonant frequency of the NEMS resonator's n^{th} mode, M_n^{eff} is the resonator's effective modal mass, δf_n is the mode's frequency shift, ϕ_n denotes the mode shape, and a is the position-of-adsorption of the molecule upon the beam (normalized to unitary beam length). The effective mass can be written $M_n^{eff} = M\alpha_n$ where M is the device mass and the numerical constant α_n, which depends on mode number, is of order unity [8].

In the experimental approach developed in the Roukes group, a NEMS device or array of devices is placed in a vacuum chamber, cooled below ambient temperature, and its frequencies are continuously tracked with a sensitive electronic (phase-locked) control loop. Using methods from conventional mass spectrometry, biomolecules are delivered sequentially to the NEMS device, and the induced frequency shifts arising from single-

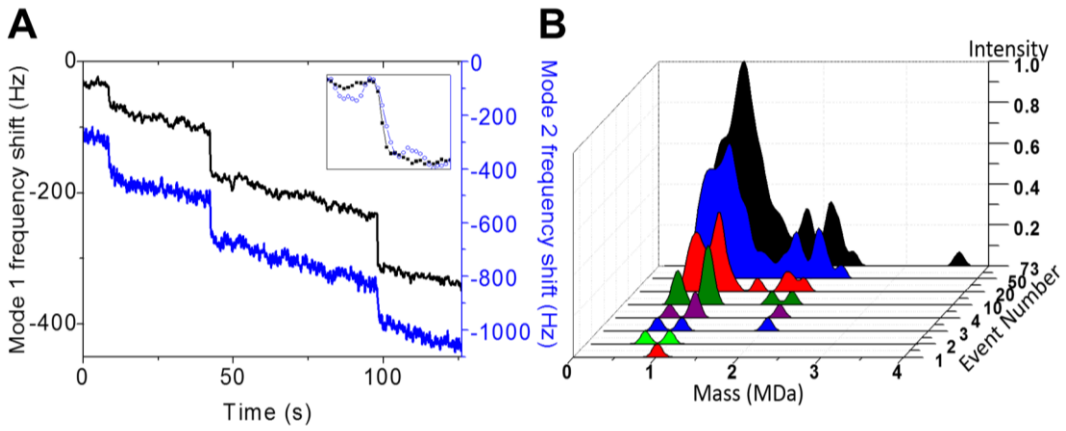


Figure 7. Detection of human IgM antibodies by NEMS-MS. (A) Individual IgM particles landing on a doubly-clamped nanomechanical beam resonator produce abrupt shifts in the first (black) and second (blue) mechanical modes. The frequency offset is 25 (67) MHz for the first (second) mode. Inset: magnification of a typical frequency shift as tracked by the feedback loop. **(B)** Molecule-by-molecule acquisition of the mass spectra for human IgM proteins. Different molecular isoforms accumulate at their respective mass values. The sum spectrum (black) combined from all 74 events has readily identifiable peaks corresponding to major isoforms of IgM typically found in human serum. From [8].

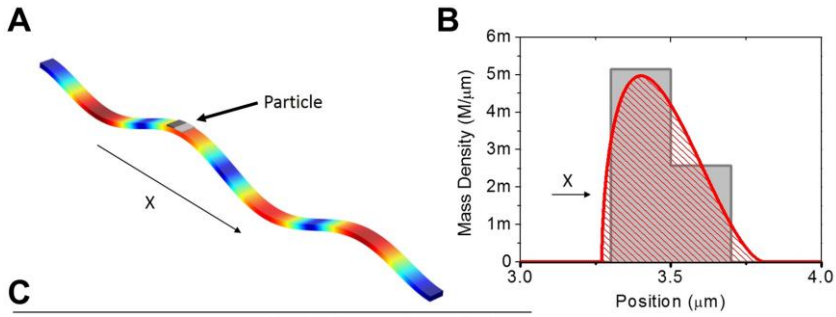
molecule events are measured for two modes and used to deduce the adsorbing analyte's mass and position [8]. Cooling the NEMS enhances non-specific physisorption of the arriving analytes on the surface of the device(s). Figure 7a shows example raw data of time-correlated frequency shifts induced in the first two displacement mechanical modes of a NEMS resonator by single-molecule events.

By individually measuring the mass of sequentially arriving particles, a mass spectrum representing the entire heterogeneous sample can be constructed, as seen in Figure 7b. Here, each IgM molecule landing on the device appears as a Gaussian-like mass distribution, with the distribution representing the uncertainty in deduced mass (in comparison, peak height in conventional mass spectrometry represents the number of species detected, with peak width the result of heterogeneity). As subsequent molecules land on the device, the mass spectrum for each molecule can be added together if desired, to form a composite spectrum representing the entire sample (Figure 7b, black curve) – but NEMS-MS resolves its intrinsic components.

Over the past decade, the Roukes group has constructed several experimental systems for performing NEMS-based mass spectrometry. One tabletop system, used to acquire the IgM data shown in Figure 7, employs ESI and ion optics to guide individual analyte ions onto the NEMS sensor(s). The setup consists of an ESI system to launch protein ions into three successive, differentially pumped vacuum chambers. The analytes are transported along their trajectory by hexapole ion guides, and ultimately delivered to the NEMS analysis stage. A flow cryostat is used to cool this stage to stabilize analyte physisorption onto the NEMS sensors. Complete system details can be found in [8].

Inertial Imaging. Inertial imaging is somewhat analogous to IMS, but provides the

enhanced capabilities of single-molecule analysis without rotational averaging. Inertial imaging is a new NEMS-based technique that provides the spatial distribution of mass within an individual analyte – in real time and with molecular-scale resolution – when it adsorbs onto a nanomechanical resonator [17]. By continuously monitoring multiple vibrational modes of a nanomechanical device, the spatial moments of mass distribution can be deduced for individual analytes, one-by-one, as they adsorb. This new method for inertial imaging has been validated with finite-element simulations to permit analysis of the inertial mass, position-of-adsorption, and the molecular shape of individual analytes. Details of the mathematical formalism underlying inertial imaging can be found in [17]. In brief, when an analyte lands on a nanomechanical resonator, each of its vibrational modes frequency shifts differently in response to the attached load. An ensemble of these distinct modal frequency shifts can then be used to yield moments of the analyte's mass density distribution; to deduce



C

Parameter	Simulated Value	Calculated Value
Mass (x 10^{-3} M)	1.545	1.540 ± 0.005
Mean Position (nm)	3,467	3,467 ± 2.5
Std. Deviation (nm)	111	113 ± 4.8
Skewness	0.439	0.424 ± 0.075
Kurtosis	2.12	2.41 ± 0.29

Figure 8. Inertial imaging FEM simulations. (A) A rectangular test particle (gray) is placed along a doubly-clamped beam. A snapshot of displacement for the fourth out-of-plane flexural mode of the beam is shown. Colors represent displacement magnitude. The physical characteristics of the beam are the same as those in [8]. **(B)** The simulated step-wise density profile of the test particle in the x-direction (gray). The spatial moments derived using inertial imaging are used to recreate an approximation of the original distribution using Pearson's method (red). The frequency shifts in the first five out-of-plane displacement beam modes were used. **(C)** The simulated and calculated values of the first five spatial moments of the particle using the frequency shifts. From [17].

N moments requires measuring induced shifts in a minimum of $N + 1$ modes. The new method is termed inertial imaging as it enables reconstruction of the analyte's spatial mass density from the deduced moments – *e.g.*, using a Pearson distribution [17] (Figure 8).

1.7 Mass adsorption to an Euler-Bernoulli beam

NEMS-MS efforts to date have made use of flexural beams such as cantilevers or doubly-clamped beams that are simple to fabricate and have dimensions much longer than their width or thickness. Small deflections of such devices are governed by Euler-Bernoulli beam theory, which postulates that beam cross-sections remain perpendicular to the neutral axis (the geometric centroid for these devices), and transverse deflections are a function of x (along the beam's length) only. An illustration of a beam is shown in Figure 9 along with variable conventions used in this text.

For a homogeneous beam with deflection function $u(x, t)$, the kinetic energy per unit length is

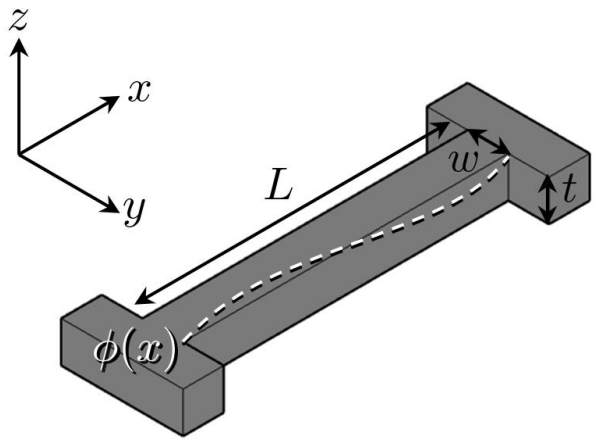


Figure 9. Doubly-clamped beam geometry. Variable conventions for axes and dimensions used in the text are shown for a doubly-clamped beam. For in-plane motion, beam displacement is in the y direction and a function of x only. The white dashed line shows the second resonant mode $\phi_2(x)$.

$$T = \frac{1}{2}\mu_0 \left(\frac{\partial u}{\partial t}\right)^2, \tag{5}$$

where μ_0 is the linear mass density of the beam, and the potential energy (strain energy) per unit length is

$$V = \frac{1}{2}EI \left(\frac{\partial^2 u}{\partial x^2}\right)^2, \tag{6}$$

where E is the Young's modulus and I is the second moment of area. In the language of modern physics, we can write the Lagrangian density as

$$\mathcal{L}(x, t, u, \dot{u}, u_{xx}) = T - V = \frac{1}{2}\mu_0 \dot{u}^2 - \frac{1}{2}EIu_{xx}^2, \tag{7}$$

where we have assumed no external loads or forces, and no damping. From here, we can derive the equation of motion:

$$\frac{\partial}{\partial t}\frac{\partial \mathcal{L}}{\partial \dot{u}} - \frac{\partial^2}{\partial x^2}\frac{\partial \mathcal{L}}{\partial u_{xx}} - \frac{\partial^2}{\partial t \partial x}\frac{\partial \mathcal{L}}{\frac{\partial \dot{u}}{\partial x}} = 0$$

$$EI\frac{\partial^4 u(x,t)}{\partial x^4} + \mu_0 \ddot{u}(x,t) = 0. \tag{8}$$

Given harmonic motion $u(x,t) \mapsto u(x)e^{i\omega t}$, we can write

$$EI\frac{d^4 u(x)}{\partial x^4} - \mu_0 \omega^2 u(x) = 0. \tag{9}$$

The associated boundary conditions for a cantilever anchored at $x = 0$ and free at $x = L$ are:

$$u(0) = 0, \quad u'(0) = 0, \quad u''(L) = 0, \quad u'''(L) = 0 \tag{10a}$$

and for a doubly-clamped beam anchored at $x = 0$ and $x = L$ are:

$$u(0) = 0, \quad u'(0) = 0, \quad u(L) = 0, \quad u'(L) = 0. \tag{10b}$$

Equations 9 and 10 together constitute a regular fourth-order Sturm-Liouville problem [73], which guarantees an infinite number of solutions for ω and associated displacement

functions $u(x)$. These solutions are the resonant frequencies ω_n and mode shapes or eigenfunctions $\phi_n(x)$, which (according to Sturm-Liouville theory) are orthogonal:

$$\frac{1}{L}\int_0^L \phi_m(x)\phi_n(x)\,dx = \delta_{mn} = \begin{cases} 0, & m \neq n \\ 1, & m = n \end{cases} . \tag{11}$$

To solve for ω_n and ϕ_n, we first write the general solution to Equation 9:

$$\phi_n(x) = a_n \cos\frac{\lambda_n}{L}x + b_n \sin\frac{\lambda_n}{L}x + c_n \cosh\frac{\lambda_n}{L}x + d_n \sinh\frac{\lambda_n}{L}x \tag{12}$$

where

$$\frac{\lambda_n}{L} = \left(\frac{\mu_0\omega_n^2}{EI}\right)^{1/4} . \tag{13}$$

Given $I = tw^3/12$ for beams with rectangular cross-section, and writing the linear mass density of the beam in terms of its density as $\mu_0 = \rho_0 wt$, the resonance frequencies can be written in terms of the eigenvalues as

$$f_n = \frac{\omega_n}{2\pi} = \frac{w}{4\pi}\sqrt{\frac{E}{3\rho_0}}\left(\frac{\lambda_n}{L}\right)^2 . \tag{14}$$

Applying the four boundary conditions in Equation 10a or 10b to Equation 12 is sufficient to solve for the eigenvalues λ_n and the eigenfunctions $\phi_n(x)$ up to some scaling factor. Writing $k_n = \lambda_n/L$, we obtain, for a cantilever:

$$[\sinh^2 k_n - \sin^2 k_n] - [\cosh k_n + \cos k_n]^2 = 0 \tag{15a}$$

$$\phi_n(x) = A_n\left([\sinh k_n x - \sin k_n x] - \frac{\sinh k_n + \sin k_n}{\cosh k_n + \cos k_n}[\cosh k_n x - \cos k_n x]^2\right) \tag{16a}$$

and for a doubly-clamped beam:

$$[\sinh^2 k_n - \sin^2 k_n] - [\cosh k_n - \cos k_n]^2 = 0 \tag{15b}$$

$$\phi_n(x) = B_n\left([\sinh k_n x - \sin k_n x] - \frac{\sinh k_n - \sin k_n}{\cosh k_n - \cos k_n}[\cosh k_n x - \cos k_n x]^2\right). \tag{16b}$$

Table 2. Eigenvalues for first five resonant modes of a cantilever and doubly-clamped beam. Shown are values for k_n solving Equations 15a and 15b, where $k_n = \lambda_n/L$ are the eigenvalues for a beam normalized to length 1.

Device	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Cantilever	1.8751	4.6941	7.8548	10.9955	14.1372
Doubly-clamped beam	4.7300	7.8532	10.9956	14.1372	17.2788

Equations 15a and 15b can be solved numerically for k_n and the results for the first five modes are shown in Table 2. The constants A_n and B_n for the mode shapes $\phi_n(x)$ can be set to satisfy the normalization condition in Equation 11, but conventionally they are set such that the maximum value of the function is 1. The first five mode shapes are plotted for a cantilever and a doubly-clamped beam in Figure 10.

Upon adding a flat adsorbate (such as a molecule) with linear mass density $\mu_{ad}(x)$, for a beam undergoing a deflection $u(x,t)$, the kinetic energy of the system increases by

$$\Delta T = \frac{1}{2}\mu_{ad}\left(\frac{\partial u}{\partial t}\right)^2, \tag{17}$$

where we have assumed the adsorbate is thin and flat relative to the device motion, allowing

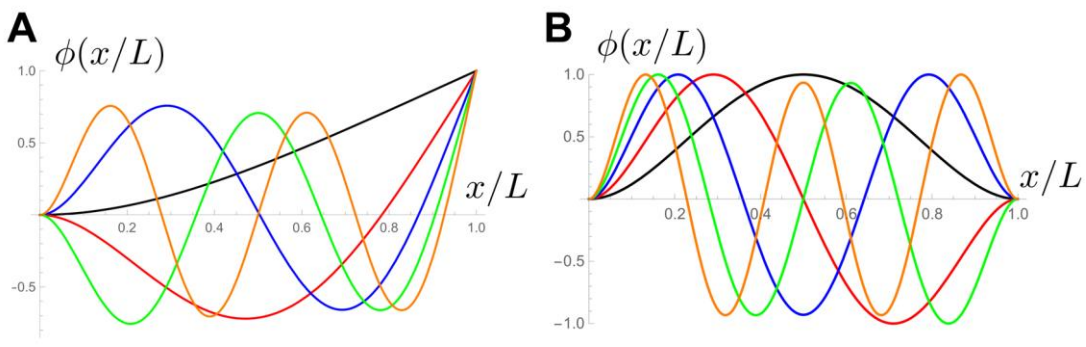


Figure 10. Device mode shapes. Mode shapes (eigenfunctions) normalized for a beam of length 1 are shown for **(A)** a cantilever and **(B)** a doubly-clamped beam. The nth mode crosses the x-axis $n-1$ times.

us to neglect any contribution due to rotational inertia. To determine the change in resonance frequency to first order, we can first use Rayleigh's energy method, which states that the energy in a dynamical system is conserved:

$$\frac{d}{dt}(\mathrm{KE}+\mathrm{PE}) = \frac{d}{dt}\left(\frac{1}{2}\int_0^L \mu_0\left(\frac{\partial u}{\partial t}\right)^2 dx + \frac{1}{2}\int_0^L EI\left(\frac{\partial^2 u}{\partial x^2}\right)^2 dx\right) = 0 \qquad (18)$$

$$\omega_n = \sqrt{\frac{\int_0^L EI\left(\frac{\partial^2 \phi_n}{\partial x^2}\right)^2 dx}{\int_0^L \mu_0 \phi_n^2(x)\, dx}} = \sqrt{\frac{K_n^{eff}}{M_n^{eff}}}\,. \qquad (19)$$

To go from Equation 18 to 19, harmonic motion $u(x,t) \mapsto u(x)e^{i\omega t}$ was used and the resulting deflection function at resonance relabeled in terms of the mode shapes ϕ_n. The resonance frequency written in Equation 19 can be understood in terms of the beam's effective modal stiffness K_n^{eff} and effective modal mass M_n^{eff}. Assuming the beam stiffness and the mode shapes do not change upon adding the adsorbate, the same method can be used to determine the new frequency:

$$\omega_n' = \sqrt{\frac{K_n^{eff}}{M_n^{eff} + \int_0^L \mu_{ad}\phi_n^2(x)\, dx}}\,. \qquad (20)$$

Assuming the mass of the adsorbate is small relative to the beam, the relative frequency shift can be written:

$$\frac{\Delta\omega_n}{\omega_n} = \frac{\omega_n' - \omega_n}{\omega_n} \approx -\frac{1}{2M_n^{eff}}\int_0^L \mu_{ad}\phi_n^2(x)\, dx, \qquad (21)$$

where $M_n^{eff} = M\alpha_n$, with M denoting total beam mass. Calculated values for $\alpha_n = \int_0^L \phi_n^2(x)\, dx$ are given in Table 3.

In the case of a point particle, $\mu_{ad} = m_{ad}\delta(x-a)$, where δ is the Dirac delta

Table 3. Calculating effective modal mass. Effective modal mass used for calculating frequency shifts is given by $M_n^{eff} = M\alpha_n$, with values for α_n tabulated below.

Device	Mode 1	Mode 2	Mode 3	Mode 4	Mode 5
Cantilever	0.25	0.25	0.25	0.25	0.25
Doubly-clamped beam	0.3965	0.4390	0.4371	0.4372	0.4245

function, m_{ad} is the particle mass, and a is the particle's position, the formula simplifies:

$$\frac{\Delta\omega_n}{\omega_n} = \frac{\Delta f_n}{f_n} = -\frac{m_{ad}}{2M_n^{eff}}\phi_n^2(a)\,. \tag{22}$$

The two unknowns m_{ad} and a can be solved given the fractional frequency shifts of the first two resonant modes of a doubly-clamped beam [8]; for a cantilever, three modes are required. In the more general case where the particle has an arbitrary mass distribution $\mu_{ad}(x)$, the mass *moments* – total mass, position, standard deviation $\int_0^L \mu_{ad}(x)(x-\bar{x})^2\,dx$, skewness, etc., of the analyte, can be calculated using more than two fractional frequency shifts [17, 74].

Determination of the distribution itself $\mu_{ad}(x)$, i.e. imaging the particle, is an ill-posed inverse problem as discussed in Chapter 5. Additionally, these calculations assume no kinetic energy contribution due to rotational inertia. Certain classes of analytes, such as a nanoparticle with minimal contact area compared with its height, or a thin molecule undergoing in-plane motion as shown in Figure 9 would need a new theory to account for this effect. As mentioned previously, no change in device stiffness is assumed; other groups have explored classes of analytes which would induce this effect [75, 76]. Finally, the device mode shapes can be appreciably modified for certain classes of analytes; see Appendix B.

Ignoring position, the sensor's mass responsivity (sensitivity to added mass) is:

$$\mathcal{R}_{\mathrm{m}} = \frac{\Delta f_n}{\Delta m} = -\frac{f_n}{2M_n^{eff}} = -\frac{1}{8\pi\rho_0 tL}\sqrt{\frac{E}{3\rho_0}}\left(\frac{\lambda_n}{L}\right)^2 \propto \frac{1}{tL^3}. \tag{23}$$

Scaling linear dimensions by a factor of 10^3 from MEMS to NEMS produces twelve orders of magnitude in mass responsivity, demonstrating enormous potential for NEMS technology.

1.8 Device fabrication and operation

Devices used for measurement in this thesis were fabricated at CEA-LETI (Grenoble, France) under the framework of the Alliance for Nanosystems VLSI, a collaboration between Dr. Roukes and Leti-Minatec, a micro and nanotechnologies campus with clean rooms devoted to teaching and tech transfer. The devices are in-plane resonators for mass sensing applications designed with a top-down process compatible with VLSI CMOS on 200 mm silicon-on-insulator (SOI) of (100) orientation with a p^{++} doped top silicon layer ($9 \cdot 10^{18}$ cm^{-3} to $5 \cdot 10^{19}$ cm^{-3}) [77-79]. Colorized SEM images of cantilevers and doubly-clamped

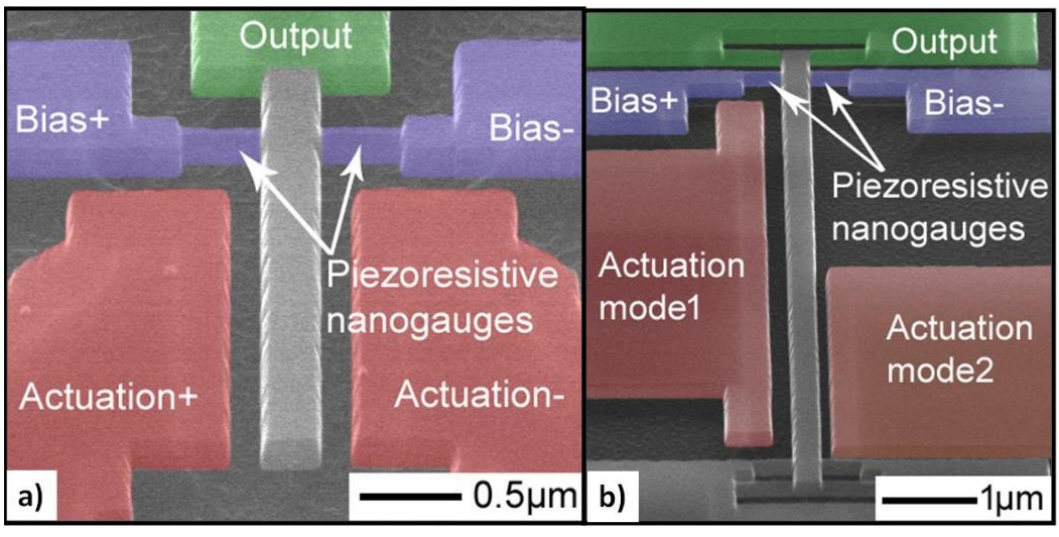


Figure 11. Colorized SEM images of devices used in this thesis. In-plane nanomechanical (a) cantilevers and (b) doubly-clamped beams fabricated at CEA-LETI under the Alliance for Nanosystems VLSI. Actuation pads in red, bias gauges in blue, and output in green. From [2].

beams are shown in Figure 11. Several fabrication runs were made, the most recent providing hundreds of devices per chip, including single devices and arrays of 20, 49, and 100 devices [80]. The devices used in this thesis are from the CAL3 wafer (referring to the third fabrication run of the Alliance). The typical device is 300 nm wide by 160 nm thick, and the chip includes devices with a range of lengths from 7.2-10 μm, with longest giving a resonance frequency of the first mode at 25 MHz. The beams are sent into motion by the capacitive force between actuation pads and the beam. The nanogauges alternately experience compressive and tensile strain as the beam moves; this is converted to an electrical signal through the piezoresistive effect. Details are provided or referenced in [2] and summarized below.

Actuation. Device motion occurs through capacitive force originating from a voltage difference between the actuation electrode and beam. Assuming small displacements relative to the gap between the two surfaces, this electrostatic force can be approximated as [2]:

$$F_{el} \approx -\frac{A\varepsilon_0}{2g^2}V_d^2\,, \tag{24}$$

where A is the capacitor surface area, g is the gap between the electrode and beam, and V_d is the actuation voltage. For an actuation voltage with DC and AC components $V_d = V_{DC} + V_{AC}\cos\omega t$, this expands to

$$F_{el} = -\frac{A\varepsilon_0}{2g^2}\left(V_{DC}^2 + \frac{V_{AC}^2}{2} + 2V_{DC}V_{AC}\cos\omega t + \frac{V_{AC}^2}{2}\cos 2\omega t\right), \tag{25}$$

which has components at DC, ω, and 2ω. This nonlinear transduction effect can be combined with the standard lock-in amplifier's capability of detection at the reference's higher harmonics to enable two detection modes. In 1f mode, $V_d = V_{DC} + V_{AC}\cos\omega t$ with $V_{DC} \neq 0$ and ω set to the device resonance frequency. In 2f mode, $V_d = V_{AC}\cos\omega t$ with ω set to

half the device resonance frequency and the lock-in amplifier set to detect at 2ω. The capacitive force is proportional to $V_{DC}V_{AC}$ for 1f mode and V_{AC}^2 for 2f mode. At higher frequencies, where it is more difficult to deliver higher voltages, 1f mode can provide a stronger force with the much more accessible DC voltage. On the other hand, 2f mode prevents detection of any unintended DC voltage contribution that might arise; Chapters 2 and 3 discuss the advantages of this approach when detecting charged analytes.

Detection. As the beam is driven into motion, one nanogauge undergoes tension and the other compression. This alternates as the beam completes each cycle of oscillation. Strain in the gauges at frequency ω is transduced to resistance variation $R \pm \Delta R \cos \omega t$ through the piezoresistive effect [77]:

$$\frac{\Delta R}{R} = \gamma\varepsilon(\omega) = \gamma \frac{F_g(\omega)}{2sE}, \tag{26}$$

where γ is the piezoresistive gauge factor (caused primarily by modification of the energy bands inside the semiconductor), s is the gauge cross-sectional area, E is the Young's modulus, and F_g is the force acting on the gauges. This force can be related to the electrostatic force on the beam given in Equation 25 by modelling the beam as a driven-damped oscillator:

$$F_g(\omega) \propto \frac{\omega_0^2}{\omega_0^2 - \omega^2 + i\omega\omega_0/Q} F_{el}, \tag{27}$$

where ω_0 is the resonance frequency and Q is the beam quality factor. The use of two opposing nanogauges allows for their operation in bridge configuration for background cancellation. At resonance frequency, which is typically in the MHz range, the electrical signal would be rapidly attenuated due to parasitic capacitance C_p. This includes cable capacitance (~100pF/m), the input impedance of the lock-in amplifier, the device pads for wirebonding, giving an overall C_p of ~125 pF, which combines with the electrical resistance

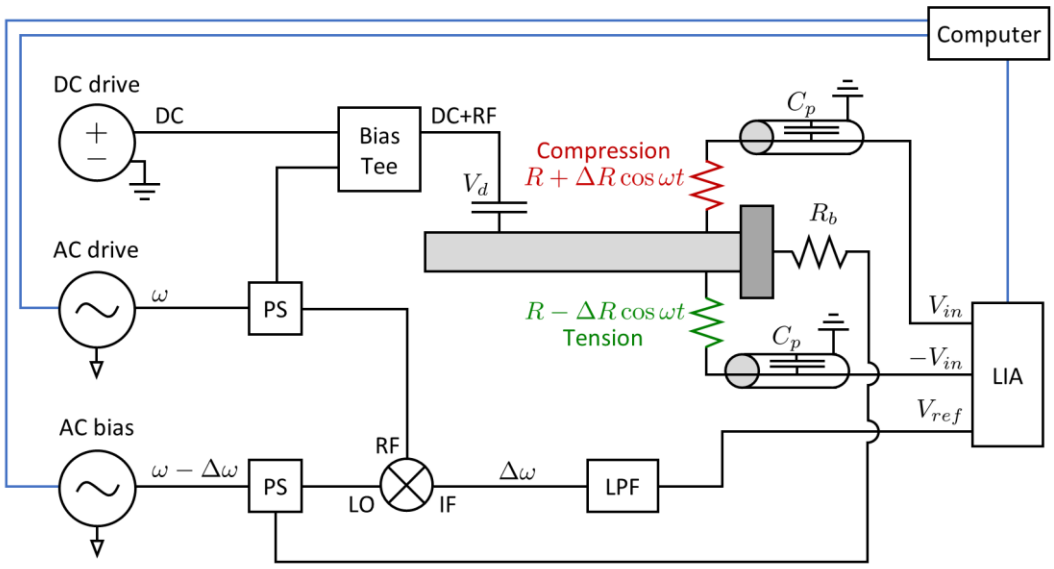


Figure 12. Circuit schematic for piezoresistive downmixing. Gate voltage $V_d \cos \omega t$ is used to drive the device using capacitive force, and device motion induces resistance variation in the nanogauges $R \pm \Delta R \cos \omega t$. A downmixing scheme is then used in which a bias voltage $V_b \cos[(\omega - \Delta\omega)t]$ is applied, leading to a signal proportional to $\Delta R/R$ at a lower frequency $\Delta\omega$ that can pass through the parasitic cable capacitance C_p. This output signal is differentially read into the lock-in amplifier. PS, LPF, LIA are power splitter, low-pass filter, lock-in amplifier.

of the device to effectively create a low pass filter at 120 kHz [77]. To overcome this, a downmixing technique [77, 81-83] is used in which the device is driven at its resonance ω, and a bias voltage is applied to the nanogauges at $\omega - \Delta\omega$ to read out resistance variation at a lower frequency $\Delta\omega$ (at <100 kHz). Three electrical ports are used with this approach, shown in Figure 11 as Bias+, Bias–, and Output.

A circuit schematic of a device operated with downmixing is shown in Figure 12. In the standard approach, a bias voltage is applied with 180 degree splitter to opposite ends of nanogauges $V_b \cos[(\omega - \Delta\omega)t]$ and the output is monitored at the "output" electrode labeled in Figure 11. Alternatively, a bias voltage $V_b \cos[(\omega - \Delta\omega)t]$ is applied to the "output" electrode, and the voltage between "Bias+" and "Bias–" monitored. This approach was consistently found to yield superior background cancelation during the course of this thesis,

so it is the one depicted in Figure 12. A circuit analysis shows this results in a differential current induced in nanogauges to leading order in $\Delta R/R$ given by

$$\Delta i_o = \frac{V_b \cos[(\omega - \Delta\omega)t]}{4R(R_b + R/2)} \Delta R \cos \omega t \ , \tag{28}$$

which is converted to a voltage through the input impedance of the lock-in amplifier Z_{in}. An output signal proportional to $\Delta R/R$ is produced at $\Delta\omega$ and $2\omega + \Delta\omega$, with the higher frequency signal removed by the lock-in's built-in low pass filter if it is not already attenuated by the parasitic capacitance. The lock-in amplifier extracts the component of the input signal in-phase and out-of-phase with the reference at $\Delta\omega$, which is output as the magnitude of the signal $V_o(\omega)$ and its phase θ relative to the reference for a given ω: $V_o(\omega) \propto \Delta i_o(\omega) \cos\theta$.

Open and closed loop operation. Equations 24-28 provide the overall signal transduction pathway between the input and output signals. The transfer function as extracted by the lock-in can be written:

$$V_o(\omega) \propto \frac{\omega_0^2}{\omega_0^2 - \omega^2 + i\omega\omega_0/Q} V_d^2(\omega) \ . \tag{29}$$

The device operates like a bandpass filter around its resonance frequency ω_0 with bandwidth $\Delta\omega$ related to its quality factor $Q = \omega_0/\Delta\omega$. The device also introduces a phase shift $-\pi/2$ at resonance. In open loop operation, the actuation frequency ω is swept around ω_0, and the magnitude of the output signal, given by Equation 29, is fitted to a Lorentzian function to obtain ω_0 and Q for the device. Once these parameters are known, closed-loop operation is used to operate the device continuously at its resonance frequency with the use of a phase-locked loop (PLL). The AC drive and bias frequency ω is continuously adjusted to target a phase shift between it and the signal detected by the lock-in amplifier of $-\pi/2$. This is achieved in a robust and stable manner, returning the device to resonance quickly for a large

perturbation (e.g., when an adsorbate lands), while filtering out high frequency disturbances (e.g., arising from device noise), through the use of H∞ loop shaping [84].

Measurement noise. Fluctuations in the resonance frequency tracked in the PLL can be characterized with a statistical tool called the Allan deviation [85]. It is analogous to a standard deviation calculation over (averaged) sequential fractional frequency data:

$$\sigma_y(\tau) = \sqrt{\frac{1}{2(N-1)}\sum_{n=1}^{N-1}(\bar{y}[n+1]-\bar{y}[n])^2}\ , \tag{30}$$

where τ is the PLL time, N is the number of samples, y is the averaged fractional frequency difference:

$$\bar{y}_n(i) = \frac{\bar{f}_n(i+1)-\bar{f}_n(i)}{\bar{f}_n}, \tag{31}$$

$\bar{f}_n(i)$ is the mean frequency at time i over some averaging time, typically a multiple of τ. The slope of Allan deviation versus time scale can be used to identify the dominant noise source driving frequency fluctuations as shown in Figure 13. Devices used in this thesis are

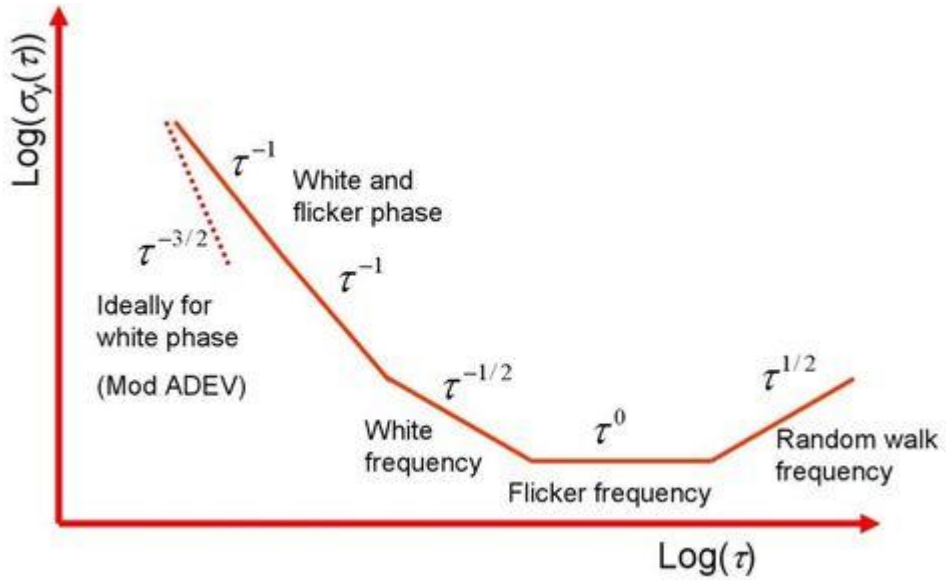


Figure 13. Noise source based on τ-dependence of Allan deviation. From [1].

typically limited by white noise for short measurement times. Lower noise can be achieved with longer measurement times (at the expense of reducing the rate of detection of molecular adsorption events), up until the device is dominated by its intrinsic 1/f mechanical domain noise (having an Allan deviation with τ^0 dependence), at which point further signal averaging does not provide any advantage. Typically, the best Allan deviation for CAL3 devices is achieved with a measurement time of around 100 ms.

The mass resolution at 1σ can be related to the Allan deviation using Equation 21:

$$\sigma_m = \sigma_y(\tau)\frac{2M_n^{eff}}{\phi_n^2(a)}\,. \tag{32}$$

A theoretical formula for Allan deviation for devices limited by thermomechanical noise has been developed [77, 86]:

$$\sigma_y(\tau) \approx \frac{10^{-\mathrm{DR}/20}}{\sqrt{Q}}\,, \tag{33}$$

where DR is the dynamic range. The in-plane piezoresistive devices described in this section have a DR of 100 dB and should have a theoretical Allan deviation of $1.5 \cdot 10^{-9}$ over an integration time of 1 second, corresponding to a mass resolution of 0.3 zg or 180 Da at room temperature [77]. Experimentally, Allan deviations of $6 \cdot 10^{-6}$ at room temperature appear to be the lower limit [77], corresponding to a mass resolution of 720 kDa. The large gap between experimental and expected mass resolution is an open question; progress in this area would clearly reap enormous benefits. Empirically, the performance improves with lower temperature. Experiments are typically performed at 80 K with the use of liquid nitrogen; typical performance gives a mass resolution of 50 kDa [8].

1.9 Contribution of NEMS to classical mass spectrometry

NEMS-MS promises several intriguing avenues to enhance the traditional field of mass spectrometry due to the unique measurement methodology and extreme sensor responsivity. Most notably, it is fundamentally a single molecule technology. This allows for accurate signal quantification over an arbitrary range, in principle providing unlimited dynamic range. As shown in Table 1, modern high resolution mass spectrometers have a limited dynamic range due to the inability to trap or focus large numbers of ions; mutual repulsion causes ions to affect each other's flight time or trajectory, impacting the measurement [14]. This issue is most problematic when attempting to detect a rare analyte in a mixture. As discussed in Section 1.4, the human proteome has a concentration dynamic range of $>10^{10}$ with the rarest proteoforms or MPCs often the most clinically significant. With NEMS-MS technology, detection of each molecule would not in principle be affected by prior measurements, allowing for much deeper proteomic profiling.

NEMS-MS provides a direct measurement of each analyte's inertial mass, rather than an indirect measurement provided with *m/z* instruments. Intact proteoforms or MPCs are detected as multiple peaks in traditional mass spectra, further complicating a sample likely to already be heterogeneous. NEMS-MS mass spectra do not require deconvolution algorithms that can fail if such spectra become too complex.

NEMS devices compatible with VLSI CMOS, such as discussed in Section 1.8, have the possibility of mass production and multiplexing. A field of 100,000 NEMS or more can be feasibly fabricated using VLSI technology, allowing for experiments to be completed faster, or to keep pace with a liquid chromatography column eluting at a fixed rate. As an example, for a human cell with 2 billion protein molecules [41], it would take a single NEMS

device operating with four modes around one year to compile the complete proteome using a measurement time of 10 ms. With a field of 100,000 NEMS, this could be reduced to about 10 minutes.

As discussed in Section 1.5, traditional mass spectrometry can handle very large biomolecules (>18 MDa), but these samples must be pure, with minimal heterogeneity. Obtaining such a sample often requires laborious sample preparation that must be tailored to each individual molecule prior to it producing a viable mass spectrum, and the resulting purified sample does not necessarily resemble its native *in vivo* state. In contrast, the upper mass limit of NEMS-MS is around 10% the mass of the beam (which is around 30 GDa for the LETI style devices). As NEMS-MS directly measures mass, no deconvolution of complicated mass spectra is necessary, so the sample could be measured from the *in vivo* environment without specialized purification. Additionally, the mass resolution (or resolving power) for typical mass spectrometers typically becomes much worse for larger *m/z* values (Figure 3); in contrast, the mass resolution for NEMS-MS is a function of the frequency fluctuations present while tracking its resonance frequency alone as given by Equation 32; this equation has no dependency on the analyte's mass.

Despite the promising applications of NEMS technology in the field of mass spectrometry, important technological barriers remain. The most pressing is the relatively low mass resolution realized by devices used in our group to date for detecting molecules. While carbon nanotubes have been used with reported mass resolution at the yoctogram level (~1 Dalton) [71], such devices are not readily compatible with CMOS electronics and VLSI, both of which are needed to form large arrays for proteomic or single-cell experiments; additionally these devices may be too narrow for adsorption of proteoform complexes. The

LETI fabricated devices we use meet both these requirements, but the mass resolution is 50 kDa. As discussed in Section 1.8, the noise of these devices is significantly higher than their theoretical limit, which should yield a resolution of 180 Da, indicating there is a substantial opportunity for improvement. Lower temperatures can also be used; plans are currently in motion to design devices that can be operated at milliKelvin temperatures. The ultimate limit for mass sensing with the use of mechanical devices has been conjectured, as dictated by quantum mechanics, to be at the microDalton level [87].

Mass resolution at the tens of kiloDalton level has only limited practical applications, while devices at the microDalton level would offer potential for a true paradigm shift in the field of mass spectrometry. Mass differences of less than a single electron could be detected in the latter case. Intermediate milestones might still provide some utility if the NEMS devices can be configured to analyze molecules *prior* to final measurement in an MS analyzer. Active efforts are under way in the Roukes group to desorb molecules from NEMS devices after they have been adsorbed and analyzed, while retaining their charge. Such a workflow would allow for pre-fractionation of molecules in a sample according to orthogonal physical attributes, including their mass, shape, charge [88], heat capacity [89], and so on, as determined by NEMS, prior to the ultimate *m/z* measurement, with the aim of simplifying a complex mass spectrum for a proteomic MS or native MS experiment. This multi-physical approach to single-molecule analysis could present a competitive alternative to chromatography methods with large sample dilution, poor resolution, and a limited ability to keep weakly bound protein complexes intact.

The small cross-sectional area of NEMS devices (~3 μm^2 for the LETI devices) would seem to severely limit the molecular detection rate as well as the detection limit

(number of molecules in the initial solution needed to produce a signal); the earliest ESI-based setup developed in the Roukes group detected an event every ~20 minutes [72]. Proteomic experiments with heterogeneous samples would then require long experiment times and enormous amounts of starting material to produce a useful mass spectrum. However, the flux rate is determined based on the NEMS capture cross section relative to the ion beam. In Chapter 3, we will show that ion optics can be used to focus the beam to a small spot size, such that the NEMS devices can then have an appreciable capture cross-section relative to the incoming ions. As a benefit, an appreciable fraction of ions in the original solution could be detected, with the only other limits being the ionization efficiency and transfer efficiency to the NEMS. The overall detection limit for the hybrid instrument can be estimated, as discussed in Chapter 6, and it could feasibly improve on LC-based separation methods by several orders of magnitude.

In summary, combining the benchtop high resolution Orbitrap instrument with the unlimited dynamic range and improved detection limit of NEMS-MS has the potential to open new horizons in proteomics and native mass spectrometry, especially as the resolution of NEMS devices improves. Other instrument combinations can enable new applications as well. Using the high upper mass limit of TOF systems in a TOF-NEMS instrument could allow for the measurement of large viruses [7] in their native state. Using the position-controlled laser in a MALDI-NEMS setup would enable MS-based imaging with similarly improved metrics.

1.10 Summary of work

This thesis is split into two parts, with Part I (Chapters 2 and 3) covering the experimental work and Part II (Chapters 4 and 5) dedicated to the theoretical work.

In Chapter 2, we will discuss the construction and calibration of a hybrid Orbitrap-NEMS instrument. This represents the first attempt to incorporate a nanoscale sensor into a high resolution, commercial mass spectrometer to improve its capabilities, and is a step towards realizing a tangible commercial application of NEMS technology. The ability of the Orbitrap to precisely isolate and detect molecules in their native state will be demonstrated, allowing for the identity of the molecules delivered to the NEMS devices to be known with great confidence. The system will be characterized in detail, including the overall ion transmission rate and the first NEMS adsorption events showing the detection of both mass and charge of the molecules. A method to operate the sensor to eliminate effect of charge on frequency-based mass measurements is demonstrated.

In Chapter 3, two- and three-mode measurements are performed of intact GroEL molecules adsorbing onto NEMS devices. Mass spectra are recovered, featuring peaks that agree with the known mass of GroEL to excellent precision. A rudimentary form of adsorbate imaging is demonstrated with the use of three modes.

Chapter 4 presents a new, streamlined approach to recover the mass moments from a set of frequency shifts induced by the physisorption of an analyte. This approach was utilized with great facility to analyze the three-mode data in Chapter 3.

Chapter 5 presents an entirely new theory to directly recover an adsorbate's image from the frequency-shift data, without the intermediate calculation of its mass moments. This approach is shown to have comparable, and in some cases, improved, accuracy to the approach in Chapter 4. It is also demonstrated experimentally using previously-published data involving droplets on microcantilevers.

1.11 Bibliography

1. NIST. *Time and Frequency from A to Z.* Available from: https://www.nist.gov/pml/time-and-frequency-division/popular-links/time-frequency-z.
2. Sage, E., *Nouveau concept de spectromètre de masse à base de réseaux de nanostructures résonantes (Doctoral dissertation).* 2013, Grenoble University.
3. Perry, R.H., R.G. Cooks, and R.J. Noll, *Orbitrap mass spectrometry: Instrumentation, ion motion and applications.* Mass Spectrometry Reviews, 2008. **27**(6): p. 661-699.
4. Paul, W. and H. Steinwedel, *Ein neues massenspektrometer ohne magnetfeld.* Zeitschrift für Naturforschung A, 1953. **8**(7): p. 448-450.
5. Anderson, N.L. and N.G. Anderson, *The human plasma proteome: History, character, and diagnostic prospects.* Molecular & Cellular Proteomics, 2002. **1**(11): p. 845-867.
6. Rose, R.J., et al., *High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies.* Nature Methods, 2012. **9**(11): p. 1084-1086.
7. Snijder, J., et al., *Studying 18 mega Dalton virus assemblies with native mass spectrometry.* Angewandte Chemie, 2013. **52**(14): p. 4020-4023.
8. Hanay, M.S., et al., *Single-protein nanomechanical mass spectrometry in real time.* Nature Nanotechnology, 2012. **7**(9): p. 602-608.
9. Makarov, A., E. Denisov, and O. Lange, *Performance evaluation of a high-field Orbitrap mass analyzer.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1391-1396.
10. March, R.E. and J.F. Todd, *Quadrupole ion trap mass spectrometry*. Vol. 165. 2005: John Wiley & Sons.
11. Makarov, A., *Electrostatic axially harmonic orbital trapping: A high-performance technique of mass analysis.* Analytical Chemistry, 2000. **72**(6): p. 1156-1162.
12. Blom, K.F., *Estimating the precision of exact mass measurements on an orthogonal time-of-flight mass spectrometer.* Analytical chemistry, 2001. **73**(3): p. 715-719.
13. Colombo, M., F.R. Sirtori, and V. Rizzo, *A fully automated method for accurate mass determination using high-performance liquid chromatography with a quadrupole/orthogonal acceleration time-of-flight mass spectrometer.* Rapid Communications in Mass Spectrometry, 2004. **18**(4): p. 511-517.
14. Makarov, A., et al., *Dynamic range of mass accuracy in LTQ Orbitrap hybrid mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2006. **17**(7): p. 977-982.
15. Gordiyenko, Y., et al., *Mass spectrometry defines the stoichiometry of ribosomal stalk complexes across the phylogenetic tree.* Molecular & Cellular Proteomics, 2010. **9**(8): p. 1774-1783.
16. Gross, J.H., *Mass spectrometry: A textbook.* Springer, 2004. **16**: p. 29.
17. Hanay, M.S., et al., *Inertial imaging with nanomechanical systems.* Nature nanotechnology, 2015. **10**(4): p. 339-344.
18. Zubarev, R.A. and A. Makarov, *Orbitrap mass spectrometry.* Analytical Chemistry, 2013. **85**(11): p. 5288-5296.

19. Yergey, A.L. and A.K. Yergey, *Preparative scale mass spectrometry: A brief history of the calutron.* Journal of the American Society for Mass Spectrometry, 1997. **8**(9): p. 943-953.
20. Uetrecht, C., et al., *High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids.* Proceedings of the National Academy of Sciences, 2008.
21. Limbach, P.A., P.B. Grosshans, and A.G. Marshall, *Experimental determination of the number of trapped ions, detection limit, and dynamic range in Fourier transform ion cyclotron resonance mass spectrometry.* Analytical Chemistry, 1993. **65**(2): p. 135-140.
22. Syka, J.E., et al., *Novel linear quadrupole ion trap/FT mass spectrometer: Performance characterization and use in the comparative analysis of histone H3 post-translational modifications.* Journal of Proteome Research, 2004. **3**(3): p. 621-626.
23. Tanaka, K., et al., *Protein and polymer analyses up to m/z 100 000 by laser ionization time-of-flight mass spectrometry.* Rapid Communications in Mass Spectrometry, 1988. **2**(8): p. 151-153.
24. Marshall, A.G., C.L. Hendrickson, and G.S. Jackson, *Fourier transform ion cyclotron resonance mass spectrometry: A primer.* Mass Spectrometry Reviews, 1998. **17**(1): p. 1-35.
25. Douglas, D. and N. Konenkov, *Mass resolution of linear quadrupole ion traps with round rods.* Rapid Communications in Mass Spectrometry, 2014. **28**(21): p. 2252-2258.
26. Fenn, J.B., et al., *Electrospray ionization for mass spectrometry of large biomolecules.* Science, 1989. **246**(4926): p. 64-71.
27. Karas, M., et al., *Matrix-assisted ultraviolet laser desorption of non-volatile compounds.* International Journal of Mass Spectrometry and Ion Processes, 1987. **78**: p. 53-68.
28. Mann, M., C.K. Meng, and J.B. Fenn, *Interpreting mass spectra of multiply charged ions.* Analytical Chemistry, 1989. **61**(15): p. 1702-1708.
29. Marty, M.T., et al., *Bayesian deconvolution of mass and ion mobility spectra: from binary interactions to polydisperse ensembles.* Analytical chemistry, 2015. **87**(8): p. 4370-4376.
30. Toby, T.K., L. Fornelli, and N.L. Kelleher, *Progress in top-down proteomics and the analysis of proteoforms.* Annual Review of Analytical Chemistry, 2016. **9**(1): p. 499-519.
31. Eiceman, G.A., Z. Karpas, and H.H. Hill Jr, *Ion mobility spectrometry*. 2013: CRC press.
32. Marcoux, J. and Carol V. Robinson, *Twenty years of gas phase structural biology.* Structure, 2013. **21**(9): p. 1541-1550.
33. A. Loo, J., S. A. Benchaar, and J. Zhang, *Integrating native mass spectrometry and top-down MS for defining protein interactions important in biology and medicine.* Mass Spectrometry, 2013. **2**(Spec Iss): p. S0013.
34. Ruotolo, B.T., et al., *Peak capacity of ion mobility mass spectrometry:: Separation of peptides in helium buffer gas.* Journal of Chromatography B, 2002. **782**(1-2): p. 385-392.

35. Giddings, J.C., *Two-dimensional separations: concept and promise.* Analytical chemistry, 1984. **56**(12): 1258A-1270A.
36. Shaw, J.B., et al., *21 Tesla Fourier transform ion cyclotron resonance mass spectrometer greatly expands mass spectrometry toolbox.* Journal of The American Society for Mass Spectrometry, 2016. **27**(12): p. 1929-1936.
37. Arnaud, C.H. and C. Washington, *High-res mass spec.* Chemical & Engineering News, 2010. **88**: p. 10-15.
38. Makarov, A. and E. Denisov, *Dynamics of ions of intact proteins in the Orbitrap mass analyzer.* Journal of the American Society for Mass Spectrometry, 2009. **20**(8): p. 1486-1495.
39. Smith, L.M., et al., *Proteoform: a single term describing protein complexity.* Nature Methods, 2013. **10**(3): p. 186.
40. Cho, W.C., *Proteomics technologies and challenges.* Genomics, Proteomics & Bioinformatics, 2007. **5**(2): p. 77-85.
41. Milo, R. and R. Phillips, *Cell biology by the numbers*. 2015: Garland Science.
42. Kim, M.-S., et al., *A draft map of the human proteome.* Nature, 2014. **509**(7502): p. 575-581.
43. Nagaraj, N., et al., *Deep proteome and transcriptome mapping of a human cancer cell line.* Molecular systems biology, 2011. **7**: p. 548-548.
44. Zhang, L., et al., *In situ metabolic analysis of single plant cells by capillary microsampling and electrospray ionization mass spectrometry with ion mobility separation.* Analyst, 2014. **139**(20): p. 5079-5085.
45. Michalski, A., et al., *Ultra high resolution linear ion trap Orbitrap mass spectrometer (Orbitrap Elite) facilitates top down LC MS/MS and versatile peptide fragmentation modes.* Molecular & Cellular Proteomics, 2012. **11**(3).
46. Leney, A.C. and A.J.R. Heck, *Native mass spectrometry: What is in the name?* Journal of The American Society for Mass Spectrometry, 2017. **28**(1): p. 5-13.
47. Barrera, N.P., M. Zhou, and C.V. Robinson, *The role of lipids in defining membrane protein interactions: insights from mass spectrometry.* Trends in cell biology, 2013. **23**(1): p. 1-8.
48. Rostom, A.A., et al., *Detection and selective dissociation of intact ribosomes in a mass spectrometer.* Proceedings of the National Academy of Sciences, 2000. **97**(10): p. 5185-5190.
49. van de Waterbeemd, M., et al., *High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles.* Nature Methods, 2017. **14**(3): p. 283-286.
50. McKay, A.R., et al., *Mass measurements of increased accuracy resolve heterogeneous populations of intact ribosomes.* Journal of the American Chemical Society, 2006. **128**(35): p. 11433-11442.
51. Uetrecht, C., et al., *High-resolution mass spectrometry of viral assemblies: Molecular composition and stability of dimorphic hepatitis B virus capsids.* Proceedings of the National Academy of Sciences, 2008. **105**(27): p. 9216-9220.
52. Snijder, J., et al., *Studying 18 MDa Virus Assemblies with Native Mass Spectrometry.* Angewandte Chemie-International Edition, 2013. **52**(14): p. 4020-4023.
53. Yin, S. and J.A. Loo, *Elucidating the site of protein-ATP binding by top-down mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2010. **21**(6): p. 899-907.

54. Zhang, H., et al., *Native electrospray and electron-capture dissociation in FTICR mass spectrometry provide top-down sequencing of a protein component in an intact protein assembly.* Journal of the American Society for Mass Spectrometry, 2010. **21**(12): p. 1966-1968.
55. Sinha, S., et al., *Lysine-specific molecular tweezers are broad-spectrum inhibitors of assembly and toxicity of amyloid proteins.* Journal of the American Chemical Society, 2011. **133**(42): p. 16958-16969.
56. Xie, Y., et al., *Top-down ESI-ECD-FT-ICR mass spectrometry localizes noncovalent protein-ligand binding sites.* Journal of the American Chemical Society, 2006. **128**(45): p. 14432-14433.
57. Hernandez, H. and C.V. Robinson, *Determining the stoichiometry and interactions of macromolecular assemblies from mass spectrometry.* Nat. Protocols, 2007. **2**(3): p. 715-726.
58. Ekinci, K.L., X.M.H. Huang, and M.L. Roukes, *Ultrasensitive nanoelectromechanical mass detection.* Applied Physics Letters, 2004. **84**(22): p. 4469-4471.
59. Ilic, B., et al., *Attogram detection using nanoelectromechanical oscillators.* Journal of Applied Physics, 2004. **95**(7): p. 3694-3703.
60. Yang, Y.T., et al., *Zeptogram-scale nanomechanical mass sensing.* Nano Letters, 2006. **6**(4): p. 583-586.
61. Li, M., H.X. Tang, and M.L. Roukes, *Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and very high-frequency applications.* Nature Nanotechnology, 2007. **2**(2): p. 114-120.
62. Gil-Santos, E., et al., *Nanomechanical mass sensing and stiffness spectrometry based on two-dimensional vibrations of resonant nanowires.* Nature Nanotechnology, 2010. **5**(9): p. 641-5.
63. Chen, C.Y., et al., *Performance of monolayer graphene nanomechanical resonators with electrical readout.* Nature Nanotechnology, 2009. **4**(12): p. 861-867.
64. Naik, A.K., et al., *Towards single-molecule nanomechanical mass spectrometry.* Nature Nanotechnology, 2009. **4**(7): p. 445-450.
65. Schmid, S., et al., *Real-time single airborne nanoparticle detection with nanomechanical resonant filter-fiber.* Scientific Reports, 2013. **3**: p. 1288.
66. Gupta, A., D. Akin, and R. Bashir, *Single virus particle mass detection using microresonators with nanoscale thickness.* Applied Physics Letters, 2004. **84**(11): p. 1976-1978.
67. Burg, T.P., et al., *Weighing of biomolecules, single cells and single nanoparticles in fluid.* Nature, 2007. **446**(7139): p. 1066-1069.
68. Jensen, K., K. Kim, and A. Zettl, *An atomic-resolution nanomechanical mass sensor.* Nature Nanotechnology, 2008. **3**(9): p. 533-7.
69. Chiu, H.Y., et al., *Atomic-scale mass sensing using carbon nanotube resonators.* Nano Letters, 2008. **8**(12): p. 4342-6.
70. Lassagne, B., et al., *Ultrasensitive mass sensing with a nanotube electromechanical resonator.* Nano Letters, 2008. **8**(11): p. 3735-8.
71. Chaste, J., et al., *A nanomechanical mass sensor with yoctogram resolution.* Nature Nanotechnology, 2012. **7**(5): p. 300-303.

72. Hanay, M.S., *Towards single-molecule nanomechanical mass spectrometry (Doctoral Dissertation)*. 2011, California Institute of Technology.
73. Greenberg, L. and M. Marletta, *The code SLEUTH for solving fourth order Sturm-Liouville problems.* ACM Transactions on Mathematical Software, 1997. **23**: p. 453-493.
74. Sader, J.E., et al., *Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation.* Nano Letters, 2018. **18**(3): p. 1608-1614.
75. Ruz, J., et al., *Physics of nanomechanical spectrometry of viruses.* Scientific Reports, 2014. **4**: p. 6051.
76. Malvar, O., et al., *Mass and stiffness spectrometry of nanoparticles and whole intact bacteria by multimode nanomechanical resonators.* Nature Communications, 2016. **7**: p. 13452.
77. Mile, E., et al., *In-plane nanoelectromechanical resonators based on silicon nanowire piezoresistive detection.* Nanotechnology, 2010. **21**(16): p. 165504.
78. Mile, E., *Systèmes électromécaniques nanométriques a base de nano-fils de silicium et nanotubes de carbone (Doctoral dissertation)*. 2010, Ecole Polytechnique.
79. Labarthe, S., *Conception et modélisation d'un nanocapteur de masse par détection piézorésistive (Doctoral dissertation)*. 2010, Joseph Fourier University.
80. Sage, E., et al., *Single-particle mass spectrometry with arrays of frequency-addressed nanomechanical resonators.* Nature Communications, 2018. **9**(1): p. 3283.
81. Bargatin, I., et al., *Sensitive detection of nanomechanical motion using piezoresistive signal downmixing.* Applied Physics Letters, 2005. **86**(13): p. 133109.
82. Bargatin, I., I. Kozinsky, and M. Roukes, *Efficient electrothermal actuation of multiple modes of high-frequency nanoelectromechanical resonators.* Applied Physics Letters, 2007. **90**(9): p. 093116.
83. Bargatin, I., *High-frequency nanomechanical resonators for sensor applications*. 2008, California Institute of Technology.
84. Kharrat, C., E. Colinet, and A. Voda. *H∞ Loop shaping control for PLL-based mechanical resonance tracking in NEMS resonant mass sensors*. in *Sensors, 2008 IEEE*. 2008. IEEE.
85. Allan, D.W., *Statistics of atomic frequency standards.* Proceedings of the IEEE, 1966. **54**(2): p. 221-230.
86. Ekinci, K., Y. Yang, and M. Roukes, *Ultimate limits to inertial mass sensing based upon nanoelectromechanical systems.* Journal of Applied Physics, 2004. **95**(5): p. 2682-2689.
87. Liu, J. and K. Di Zhu, *A micro dalton resolution mass sensor using optically cooled microdisks via short-time measurement.* arXiv preprint arXiv:1807.11531, 2018.
88. Cleland, A.N.R., M.L., *A nanometre-scale mechanical electrometer.* Nature, 1998. **392**: p. 160-162.
89. Fon, W.C., et al., *Nanoscale, phonon-coupled calorimetry with sub-attojoule/Kelvin resolution.* Nano Letters, 2005. **5**(10): p. 1968-71.

PART I: Experimental Work
Chapter 2

CONSTRUCTION OF A HYBRID ORBITRAP-NEMS SYSTEM

2.1 Introduction

Thermo Scientific, at the behest of Dr. Alexander Makarov, generously provided our lab with a custom Q Exactive Plus instrument modified with EMR (extended mass range) capabilities.[1] The instrument is outfitted with a special quadrupole for selecting species of interest up to 4.5 MDa, as well as an Orbitrap detector capable of analyzing the mass-to-charge ratio of species with extremely high resolving power (up to 140,000 for 200 *m/z*) [1]. A similar instrument was constructed in Professor Albert Heck's lab in collaboration with Dr. Makarov, and was capable of measuring intact IgG antibody at 149 kDa, yeast proteasome at 730 kDa, and GroEL at 801 kDa [2], up to viral nanoparticles at 4.5 MDa [4]. The acquisition of this unique instrument thus immediately brought our capabilities of transferring biomacromolecular species from solution to the gas phase, and selecting and analyzing species of interest, to the forefront of native mass spectrometry research. Modification of this instrument into a hybrid system also capable of delivering molecules to NEMS for the purpose of single molecule analysis is the primary subject of this chapter.

The design, construction, and operation of the hybrid system form the bulk of my experimental work for this thesis. The design of the instrument modifications was supervised by Professor Michael Roukes, Dr. Warren Fon (a senior staff scientist in the group), and Dr.

[1] The addition of a quadrupole to the Exactive Plus EMR, or conversely, modification of the Q Exactive Plus for EMR capability, is not yet available commercially. For brevity, in this thesis, the custom mass spectrometer will simply be referred to as the Q Exactive.

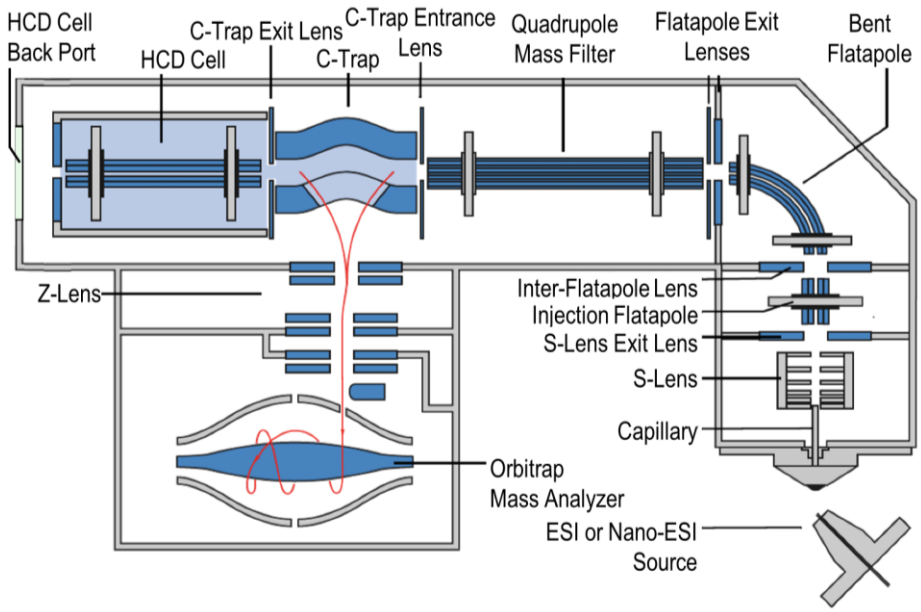


Figure 14. Diagram of Q Exactive. A Thermo Exactive Plus Extended Mass Range (EMR) instrument outfitted with an electrospray ionization (ESI) or nanoelectrospray ionization (nanoESI) source is used to introduce intact proteins or protein complexes from solution to the gas phase. Molecules with a specific mass-to-charge ratio can be selected with the quadrupole mass filter, then delivered to the Orbitrap for analysis. The back port of the HCD cell can be removed for instrument modification or extension. Reproduced from Q Exactive product literature.

Alexander Makarov and his team at Thermo Scientific (Dr. Maria Reinhardt-Szyba and Dr. Dmitri Boll). The construction and operation was carried out with extensive help from Dr. Eric Sage (a postdoc in the group). A schematic of the custom Q Exactive we obtained is shown in Figure 14.

Successful realization of the hybrid system required careful consideration of the design of the Q Exactive, and also the necessary operating conditions of both the Q Exactive and our NEMS devices. The Q Exactive includes multiple chambers for transferring or trapping ions. One such chamber, the higher energy collisional dissociation (HCD) cell, includes a removable back port purposely designed to allow for instrument modification or extension for special projects such as ours. The existence of this back port, as well as precise engineering schematics of the HCD cell in the form of STEP files provided to us by Thermo,

greatly facilitated the practical modification and construction of the hybrid instrument. Once properly adapted, this HCD cell back port serves as the boundary between the Q Exactive and the NEMS analysis chambers.

The Q Exactive is intended to operate under continuous high vacuum to prevent moisture or external contaminants from degrading the performance of the system. Although the instrument can be completely shut off and vented with nitrogen, restoring the instrument state following shut off requires an overnight bakeout, extensive calibration routines, and runs the risk of contamination with any interruption of the flow of nitrogen. NEMS devices, on the other hand, tend to fail fairly frequently, especially when operated under new conditions. Therefore, a key constraint on the design was the installation of a gate valve between the Q Exactive and the NEMS analysis chamber allowing for the easy exchange of NEMS devices in the NEMS chamber without breaking the vacuum of the Q Exactive.

Another necessary consideration for the hybrid system design includes maintaining a proper vacuum regime for both the Q Exactive and for the NEMS. As discussed in Section 2.5, the HCD cell in the Q Exactive acts as an extended trapping region for larger molecular weight complexes. Relatively high pressure (10^{-4} Torr) is used in this chamber both to trap incoming ions so they can be collected in the C-Trap prior to injection into the Orbitrap, and to remove excess water and salt molecules via collisions with gas molecules [2]. On the other hand, NEMS operation requires relatively low pressure ($<10^{-6}$ Torr). While dissipation of NEMS devices due to collisions with gas molecules becomes negligible at pressures less than a few mTorr [5], a larger concern was the possibility of significant adsorption of ambient gas while operating the devices at cryogenic temperatures. Such adsorption would lead to a constant drift in resonance frequency of the devices and

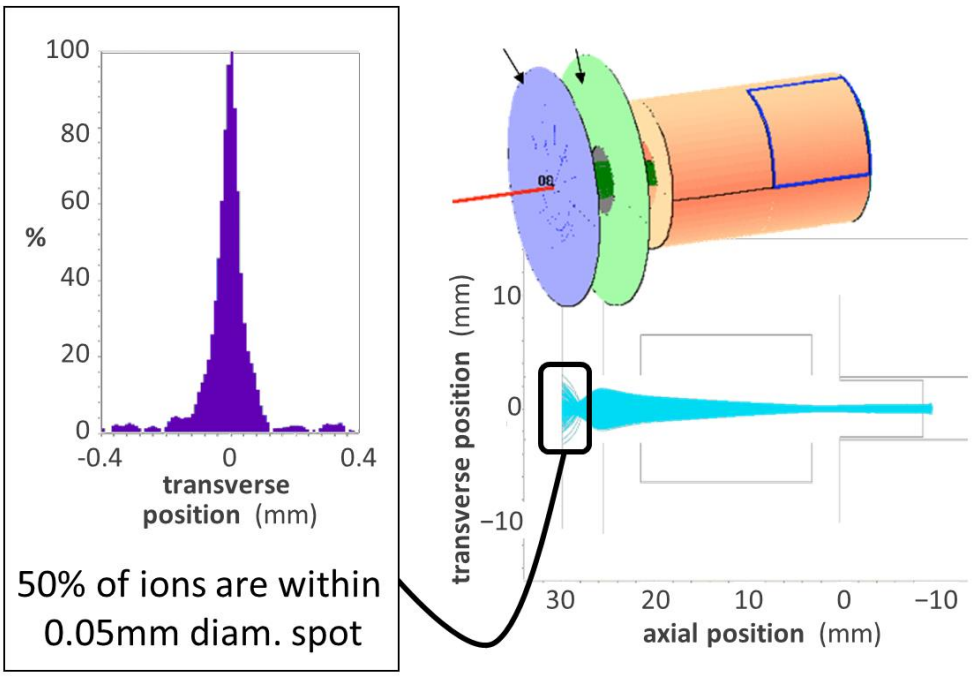


Figure 15. Ion optics design. Thermo designed a series of ion lenses which could be used to focus the ion beam exiting the HCD cell and continuing on through a series of transfer quadrupoles into a 50 micron spot (Figures and simulations provided by Thermo). Assuming precise positioning of the NEMS devices, this would guarantee sufficient flux rate to the NEMS, with a typical cross-sectional area of 0.3 μm^2, for ample data for the calculation of mass spectra.

enhancement of fluctuation processes that degrade their noise performance [6, 7]. Therefore, a differential pumping design was employed, using a separate NEMS inlet chamber with its own turbo pump in order to isolate the high pressure of the HCD cell from the low pressure of the NEMS analysis chamber. Pressure measurements of the NEMS inlet and analysis chambers under the full range of HCD cell pressure settings are later given in Section 2.4 and demonstrate that the pressure of the NEMS inlet chamber does rise significantly with increasing HCD gas pressure, validating this design decision. The HCD cell, on the other hand, continues to maintain a relatively high pressure despite the addition of the new vacuum chambers due to the presence of a circular disc with an orifice of only 2.5 mm diameter (not

shown). This disc serves to trap gas inside the HCD cell while allowing ions to continue past the HCD cell if necessary.

A final critical design criterion was achieving sufficient ion flux to the NEMS. The typical cross section for devices used for mass sensing is on the order of 10 microns by 0.3 microns, which represents an extremely small capture cross section for typical ion beam sizes on the order of a few mm^2. This previously presented a barrier to obtaining abundant data; a prior ESI-based instrument developed in the Roukes group achieved a maximum event rate of 1 per minute [8]. This limitation, however, could be circumvented with the addition of ion optics immediately prior to the NEMS. Thermo provided an optimal design for such ion optics appropriate for the Q Exactive hybrid system (Figure 15); it was up to us to implement the design and ensure that the NEMS could be placed in the center of the focused ion beam with adequate precision.

2.2 Hybrid system architecture

An overview of the hybrid system architecture satisfying the main design constraints is shown in Figure 16. Upon removing the back plate of the HCD cell, an additional vacuum chamber (NEMS Inlet Chamber) is attached, followed by a gate valve, and a second vacuum chamber (NEMS Analysis Chamber). Both NEMS chambers are outfitted with turbo pumps backed by diaphragm pumps that provide sequential differential pumping. The NEMS chip itself is mounted on a printed circuit board (PCB) that rests on a 3D translatable stage and is thermalized to a cryostat which can be cooled with liquid nitrogen or liquid helium.

Dr. Makarov advised that the new transfer quadrupoles should be collinear with the HCD cell axis to within 50 microns, and any end-to-end gaps between quadrupoles should be less than 5 mm to avoid excessive loss of ion transfer efficiency. To achieve this level of

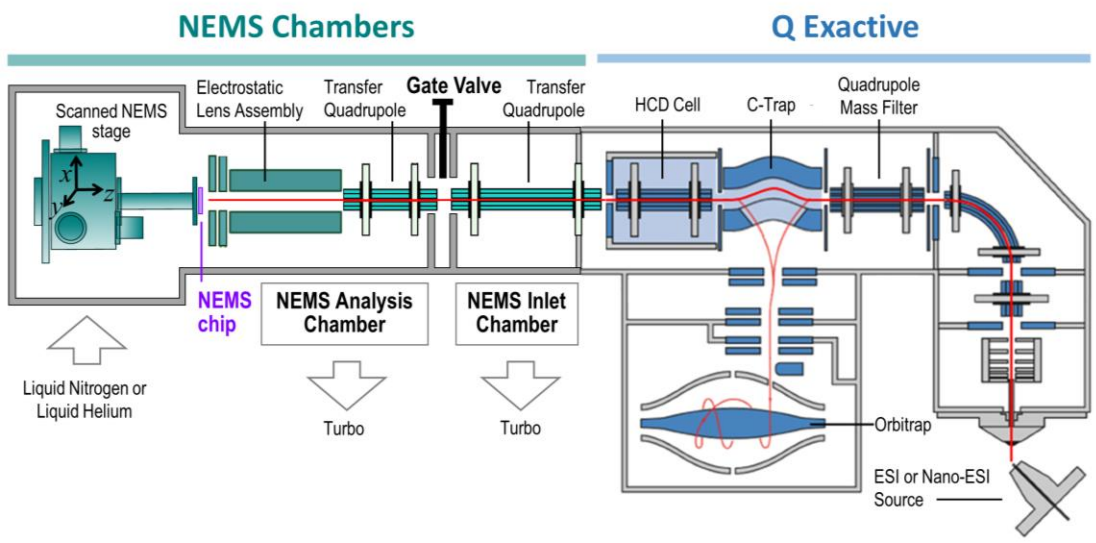


Figure 16. Architecture of the Hybrid Q Exactive-NEMS System. Once properly modified, the hybrid instrument is capable of delivering intact proteins or protein complexes either to the Orbitrap chamber for analysis of mass-to-charge ratio, or to the NEMS for single molecule analysis. Delivery to the NEMS is accomplished by removing the back plate of the HCD cell and the use of additional vacuum chambers, transfer quadrupoles, and ion lenses. The NEMS chip is thermalized to a cold finger cryostat and can be precisely positioned using a three dimensional translatable stage.

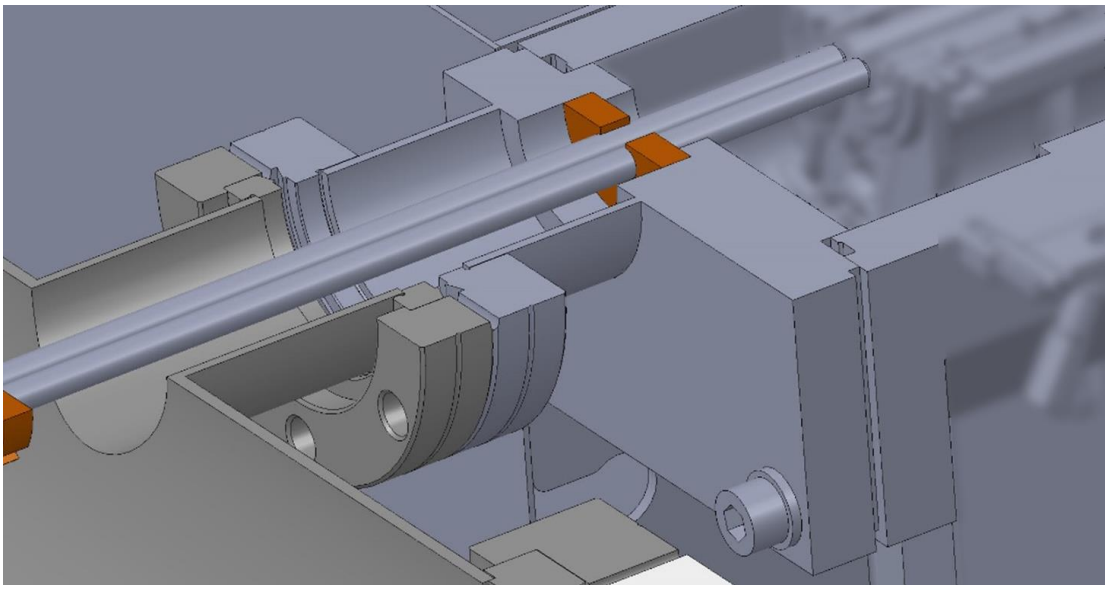


Figure 17. HCD adapter plate. The back plate of the HCD cell was replaced with an adapter plate that allows for precise positioning of the NEMS transfer quadrupole and converts the custom Thermo vacuum flange to a standard one. Details of the HCD cell are proprietary and have been intentionally blurred out.

precision, an HCD adapter plate was designed that would allow the NEMS inlet chamber quadrupole to slide in place precisely via a slip fit (Figure 17). The quadrupole rods are held in place via circular PEEK inserts (shown in orange); PEEK was chosen for its low outgassing properties. Finally, the HCD adapter converts from the custom Thermo vacuum flange to a standard vacuum flange (which was welded in place concentric to the HCD cell axis) and allows for the remaining design to consist of standard vacuum components to reduce the time and cost of construction.

The opposite end of the NEMS inlet chamber quadrupole ("QP1") was clamped into place using a "groove grabber" from Kimball Physics (Figure 18). These groove grabbers are seated into grooves in the interior of double sided CF flanges which also have an exterior groove for mounting brackets. The double-sided CF flanges also have offset screw holes so that one side can be tightened at a time, providing for straightforward assembly: QP1 is inserted into place, then tightened down with the groove grabber while one side of the double sided CF flange is attached to the inlet vacuum chamber. Then, once QP1 has been mounted,

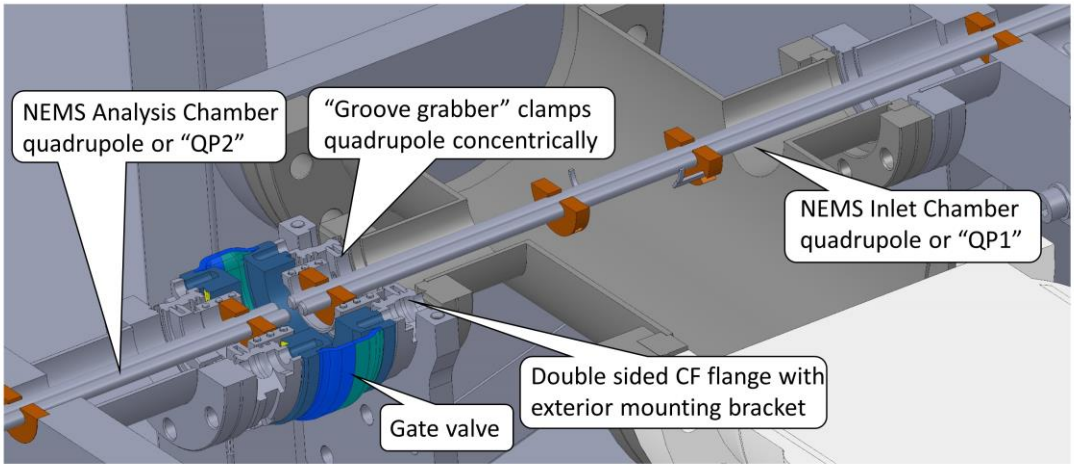


Figure 18. Quadrupole mounting near gate valve. The two transfer quadrupoles are mounted into place by fitting into a groove machined into a vacuum flange on one end, and clamped into place on the other end using "groove grabbers" (Kimball Physics).

the gate valve is attached using the other set of screw holes. The second transfer quadrupole corresponding to the NEMS analysis chamber ("QP2"), which is also visible in Figure 18, is assembled in a similar fashion. QP2 slides into place into a groove machined into a flange, then the other side is clamped down using the groove grabber with one side of the double sided CF flange having been assembled. Finally, the remaining side of the CF flange is screwed in to the gate valve.

Per Dr. Makarov's suggestion, the end-to-end gap between QP1 and QP2 was set to 5 mm. A miniature gate valve (VAT) was commercially available with a gate thickness of 4 mm, which made this possible. The lengths of the quadrupole rods and locations of the PEEK inserts via tapped threads needed to be machined to within 0.5 mm precision. The variable gap between CF flange components, even when fully tightened, made it challenging to achieve this precision. In practice, the quadrupole rods needed to first be machined slightly longer, then small amounts taken off the ends until the gate valve could close.

As mentioned in Section 2.1, an ion lens assembly was required to focus the spot size of the ion beam. Thermo provided us with a schematic including lens geometries and

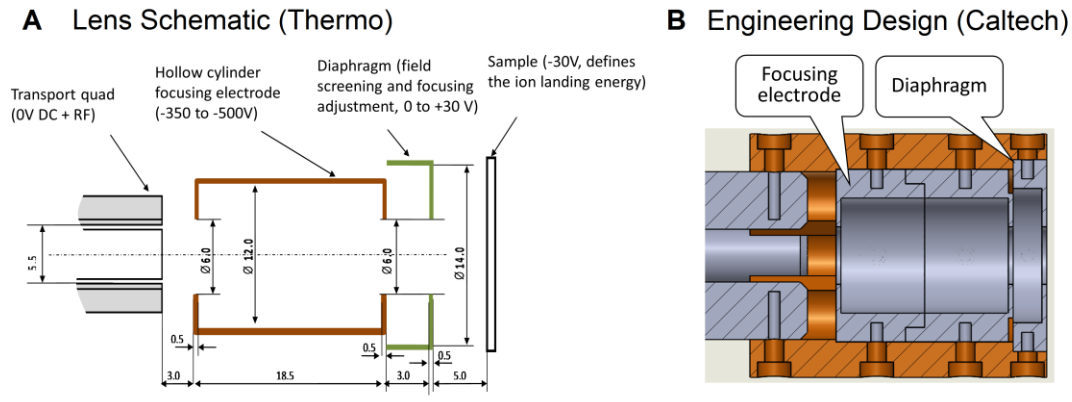


Figure 19. Ion lens design. (A) Thermo provided a schematic of an ion lens assembly that would enable the ion beam to be focused to a 50 micron spot according to proprietary simulations. **(B)** This was used to create an engineering design.

voltages that would result in the required spot size (Figure 19). For ease of assembly and alignment, the lenses were all mounted together and to QP2 using a single PEEK housing. Although a mere 0.5 mm gap separated two of the lenses separated by up to 350 V, no shorting between these components was observed provided the turbo pump for the NEMS analysis chamber was on (once, the voltages were accidentally turned on with only a rough vacuum in place, and this led to shorting and ablation of some of the lens surfaces, after which the components needed to be replaced). Similarly, QP2 was operated at up to 500 V_{pp} at 650 kHz despite 5 mm worth of PEEK material in contact with all four rods, and no shorting between these components was observed provided the turbo pump was on.

The final components of the Q Exactive hybrid system architecture not yet discussed include a 3D XYZ positioner (Attocube), the NEMS PCB, and a liquid helium cryostat

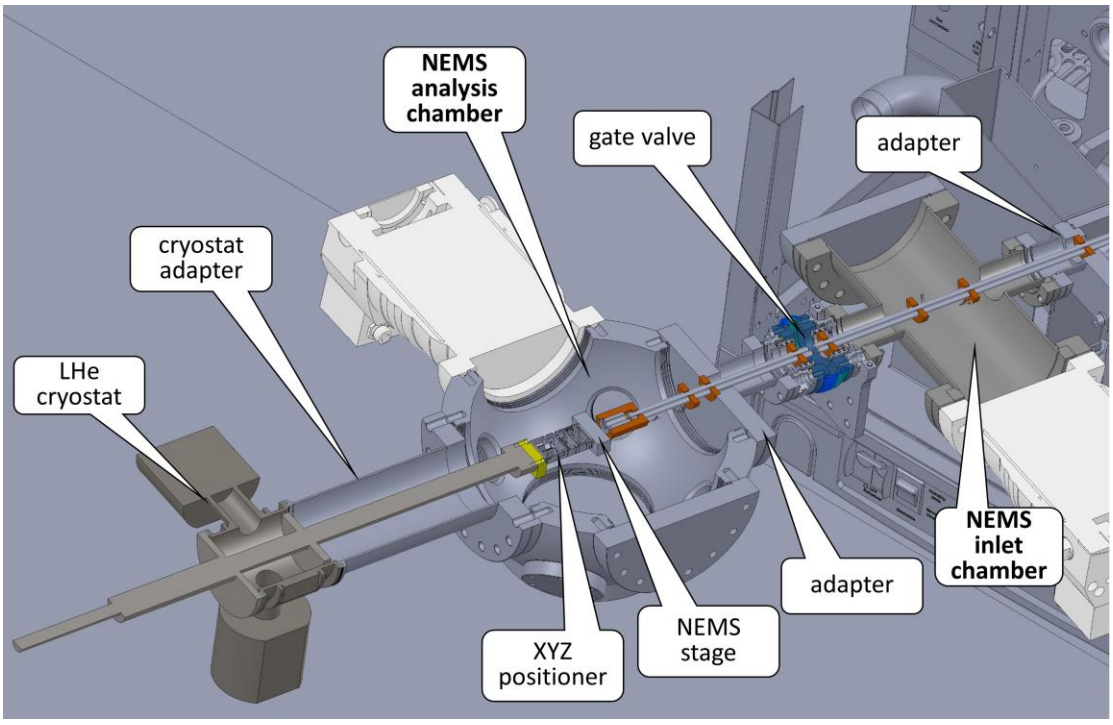


Figure 20. Q Exactive Hybrid System overview (cutout drawing). The NEMS chambers are connected together and to the Q Exactive through a few custom vacuum flange adapters and otherwise standard vacuum components. The transfer quadrupoles and ion lenses are concentric with the HCD cell. After closing the gate valve, the cryostat can be readily removed, and the NEMS replaced.

(Janis). The PCB and cryogenics will be discussed later in Section 2.6. For the positioner, we acquired the Attocube low temperature, high vacuum piezo positioner with 5 mm of travel range in three dimensions and sub-nanometer positioning capabilities. An overview of the entire system design is given in Figure 20.

2.3 Assembly of NEMS inlet chamber

After the instrument was unboxed and set up, Dr. Reinhardt-Szyba from Thermo performed the necessary instrument modifications and calibration routines for us to have an operational Exactive Plus EMR instrument outfitted with a quadrupole mass filter. Prior to continuing on with the planned modifications as depicted in Figure 16, several components that contribute to the normal functioning of the Q Exactive needed to be permanently removed from the end of the HCD cell in order to enable ion transmission through this compartment. One of these components was an electrometer sensor electrode used by the Q Exactive to monitor ion current. Before removing this component, electrospray ionization (ESI) was performed on a standard calibration mixture ("Calmix," or LTQ Velos ESI Positive Ion Calibration Solution, Pierce) in order to obtain a baseline ion current measurement for characterizing the ion transfer efficiency of each of the NEMS chambers.

Calmix is typically measured in the Q Exactive using voltage settings optimized for calibration as shown in Figure 21a. These settings include a gradually decreasing DC voltage up until the C-trap entrance lens, upon which the voltage suddenly increases. The reason for this voltage barrier is as follows: caffeine (one of the Calmix components) ionizes at higher efficiency than the other components of the calibration mixture and would typically dominate the mass spectrum, preventing calibration from completing. The presence of the C-trap entrance lens voltage barrier, on the other hand, partially blocks ion transmission, and does

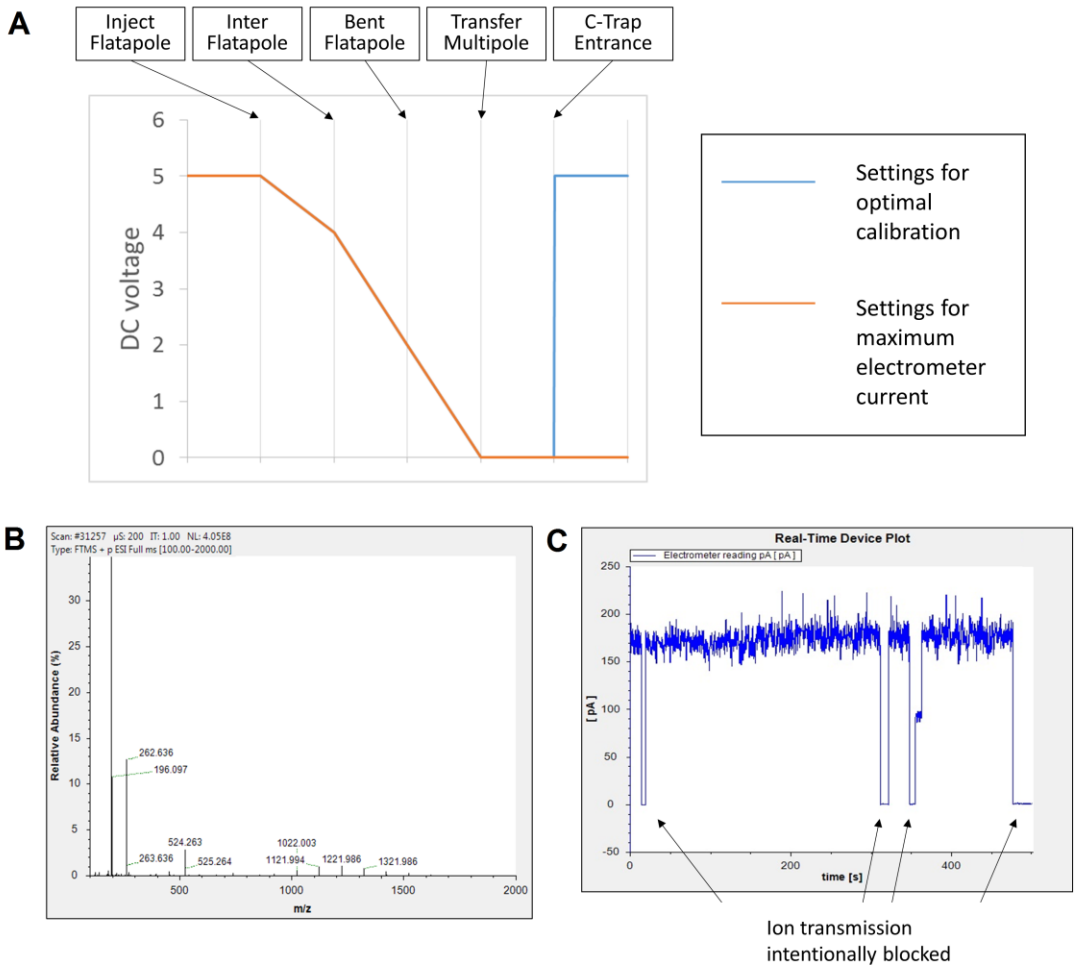


Figure 21. Calmix mass spectrum and ion current. (A) Voltage settings were modified from those typically used for calibration purposes in order to maximize total ion transmission to the electrometer. **(B)** Mass spectrum of standard calibration mixture ("Calmix") under these modified settings. **(C)** An ion current of 175 pA was measured via a built-in electrometer at the back of the HCD cell. Various lens settings were periodically adjusted to block Calmix transmission. These adjustments correctly resulted in zero current as shown in the plot.

so much more effectively for ions with lower kinetic energy (and hence *m/z*). For the purpose of determining ion transfer efficiency, however, these voltage settings were modified to maximize total current, rather than using the suggested settings for optimal calibration. This was possible simply by removing the C-trap entrance lens voltage barrier.

A mass spectrum of Calmix under these modified settings and electrometer measurements at the back of the HCD cell are shown in Figure 21b and Figure 21c.

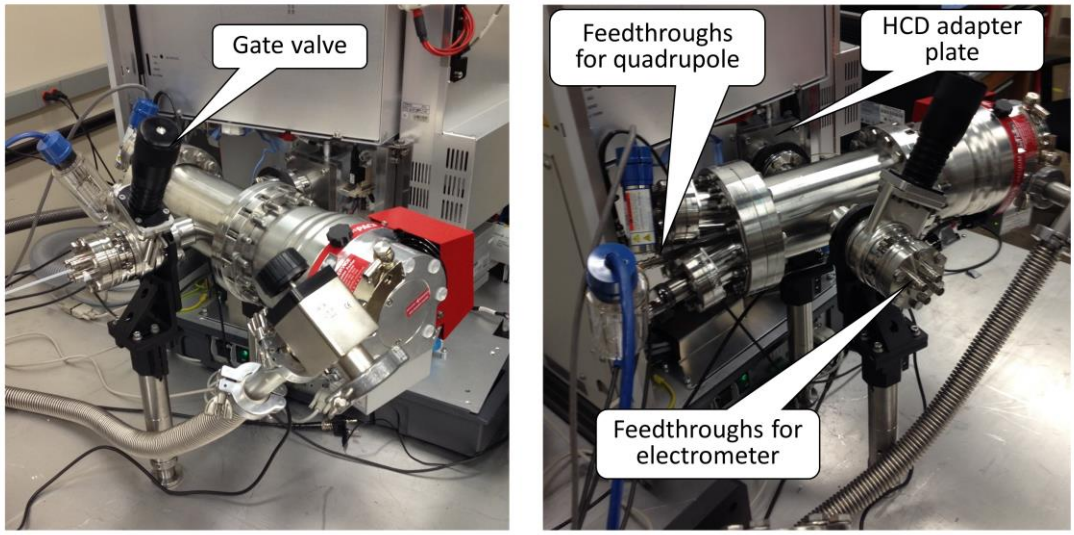


Figure 22. Assembly of first stage. Completed assembly of NEMS inlet chamber up to and including the gate valve. A custom electrometer was mounted after the gate valve to monitor ion transmission efficiency through this stage.

Approximately 175 pA of current was measured. Various ion lens voltages could be adjusted to intentionally block or allow ion transmission, demonstrating the monitored current was due to Calmix ions reaching the HCD cell.

After performing this initial measurement, the instrument was vented, the HCD back plate removed, the electrometer sensor electrode at the back of the HCD cell was removed, and then construction proceeded up to the gate valve. A picture of construction up to this intermediate stage is shown in Figure 22. Once the gate valve was installed, vacuum in the Q Exactive was restored, an overnight bakeout of the Q Exactive was performed, and the instrument was successfully recalibrated. At this point, rather than continue construction of the NEMS analysis chamber, it was decided to operate the partially constructed system to individually evaluate the ion transmission efficiency of the NEMS inlet chamber.

To evaluate ion transmission of the NEMS inlet chamber, a new electrometer sensor electrode was installed immediately after the gate valve. The Thermo electrometer electrode removed from the HCD cell was unable to be used for this purpose, so a new electrometer

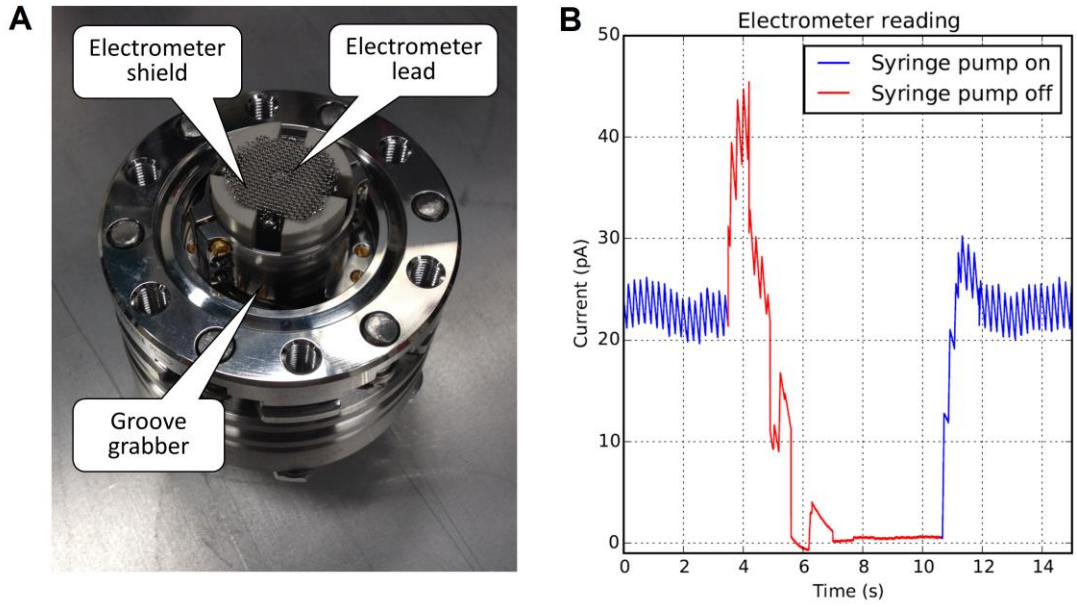


Figure 23. Electrometer reading past first stage. (A) A custom electrometer was mounted concentrically into a flange which was later attached to the gate valve. **(B)** This electrometer recorded a current of 23 pA for Calmix during the periods of time that the solution was continuously delivered through the ESI needle via a syringe pump.

electrode was constructed consisting of a screw embedded in a PEEK cylinder serving as the electrometer lead, shielded by a wire mesh grid (Figure 23a). This new construction was mounted concentrically with the HCD cell via a groove grabber, and installed immediately after the gate valve to monitor ion transmission. The mesh grid provides an equipotential surface while allowing ions to pass through [9]. This equipotential surface was floated to –30 V to achieve suitable ion landing potential as depicted in Figure 19a, and additionally served as a shield for the purpose of performing low noise current measurements of ions landing on the lead with an electrometer (Keithley 6514).

Calmix was introduced into the Q Exactive as before, except this time a setting was used (“beam mode”) to continuously send ions through the HCD cell, rather than the usual configuration, in which ions are trapped in the HCD cell, prior to sending them to the Orbitrap for analysis. Voltage settings for the HCD cell used in the operation of beam mode

are adjusted ahead of time in the control software, so that the only action needed to switch to the new modality of allowing ions to pass through the HCD cell continuously is a single mouse click.

Power was also supplied to QP1 to allow for ion transmission through the NEMS inlet chamber. As discussed in Appendix A, linear quadrupole ion guides consist of two pairs of circular rods operated with a DC and RF voltage added to or subtracted from an overall DC potential. Transmission of ions of a broad range of *m/z* values can be achieved by operating all four quadrupole rods at the same DC. Given an operating frequency of 650 kHz, the most stable trajectory for Calmix was predicted to occur with an RF voltage of 60V. The overall DC potential of all rods was set to –25V, halfway between the voltage at the back of the HCD cell (–20V) and the electrometer potential (–30V).

Calmix was sent to the electrometer using these initial voltage settings for QP1, then the RF voltage was tuned until the current was maximized. A maximum current of 23 pA was recorded by the electrometer with 50V RF, close to the predicted RF value (Figure 23b). Compared to the electrometer current of 175 pA measured at the back of the HCD cell, the ion transmission through the NEMS inlet chamber and across the gate valve was 13%.

2.4 Assembly of NEMS analysis chamber and adjustment of ion optics

The remaining portion of the hybrid instrument after the gate valve, the NEMS analysis chamber, was then assembled. Pictures of this PCB, the interior of the NEMS analysis chamber, and the exterior of the completed hybrid instrument are shown in Figure 24. The interior includes a new quadrupole which slides into a custom vacuum flange via a slip fit, then is clamped down via a groove grabber near the gate valve as shown in Figure

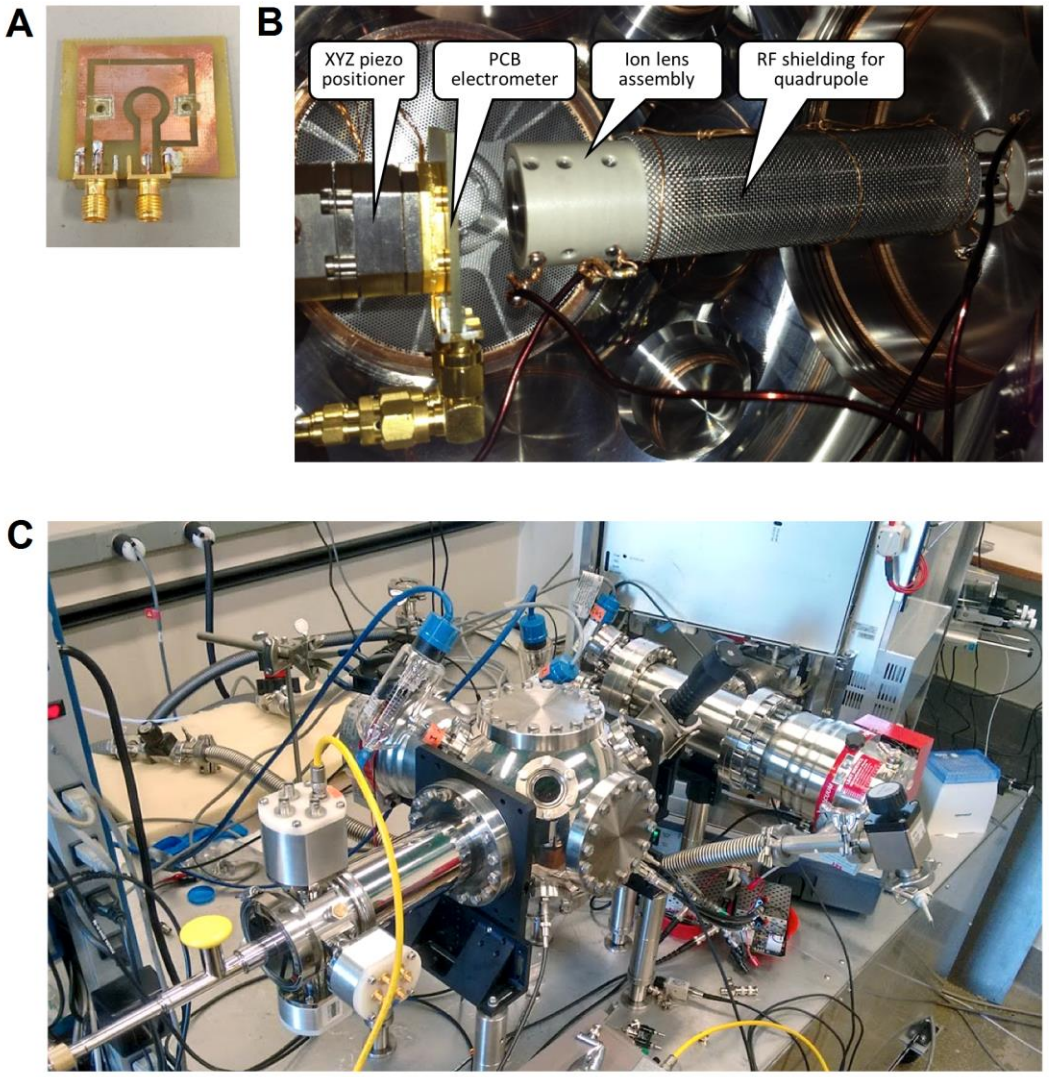


Figure 24. Hybrid instrument following assembly of NEMS analysis chamber. (A) Electrometer PCB for preliminary monitoring of ion current through chamber. **(B)** Interior view of chamber showing mounted PCB and quadrupole. **(C)** Exterior view of completed construction.

18. A custom electrometer sensor, fabricated as a PCB, was also temporarily mounted on the piezo positioner to monitor ion transmission through this new chamber.

Following assembly, the hybrid instrument was tested with a helium leak detector, then the NEMS analysis chamber was soft baked at 325 K for a few days, then the gate valve opened while the Q Exactive was in standby mode. In this configuration, the NEMS analysis

chamber had a base pressure of 2.6E–7 Torr, while the NEMS inlet chamber had a base pressure of 4.7E–7 Torr. As discussed in Section 2.1, during normal operation of the Q Exactive, nitrogen gas is introduced into the HCD cell in order to facilitate the trapping of large biomolecules, and a differential pumping design was used to isolate this higher pressure from the lower pressure required in the NEMS analysis chamber. To verify that this design decision was effective in practice, the trapping gas pressure of the HCD cell was increased in increments, up to the maximum setting enabled by the software, and the pressure of the two chambers was recorded after the system reached equilibrium. These pressure readings are given in Table 4 and demonstrate that the NEMS inlet chamber pressure increases substantially with the use of high trapping gas pressure, while the NEMS analysis chamber pressure does not increase substantially and can remain in a regime where adsorption of ambient gas to NEMS devices is minimized. This justifies the use of the differential pumping design despite the added cost and complexity to the overall instrument construction.

After measuring the vacuum chamber pressures, Calmix ions were sent to the NEMS

Trapping Gas Pressure Setting for HCD Cell (Q Exactive software)	NEMS Inlet Chamber Pressure (Torr)	NEMS Analysis Chamber Pressure (Torr)
(Q Exactive Off)	4.7E–7	2.6E–7
1	1.2E–6	2.7E–7
3	3.3E–6	2.8E–7
5	5.6E–6	2.9E–7
7	8.1E–6	2.9E–7
8	9.2E–6	3.0E–7

Table 4. Vacuum pressure under typical instrument operation. Pressure readings of the two vacuum chambers under a range of HCD trapping gas settings used for a typical native MS experiment. The NEMS inlet chamber pressure increases substantially with higher trapping gas pressure, but the effect on the NEMS analysis chamber is minimal.

analysis chamber using +30V for the initial value for the ion lens diaphragm (Figure 25a). The electrometer PCB was scanned over the full translational range using the piezo positioner, but no ions were detected. To verify the ions could cross the 5 mm gap between QP1 and QP2, QP2 was used as an electrometer during Calmix transmission and a current of 50 pA was measured. The lens diaphragm was then adjusted to a value of 0 V, the PCB scanned, and ions were detected with a maximum current of 18 pA. Once the position for maximum ion current was found with the lens diaphragm set at 0 V, the lens focusing cylinder and diaphragm voltages were adjusted over a wide range to find the settings maximizing ion transmission. The lens focusing cylinder was found to have minimal impact on ion transmission until it increased to higher than –10V, at which point ion transmission was blocked. Ion transmission was found to be more sensitive to the diaphragm setting, and the electrometer current as a function of diaphragm voltage is shown in Table 5. In general, ions were detected with a range between -60V and +10V, which differed from the expected range of 0V to +30V obtained by Thermo's proprietary ion simulations used for lens design.

Ion Lens Diaphragm Voltage	Electrometer Current on NEMS PCB
–60V	34 pA
–40V	30 pA
–20V	27 pA
0V	18 pA
+10V	10 pA
+20V	0 pA
+30V	0 pA

Table 5. Optimizing ion lens diaphragm voltage. Ion detection with the electrometer PCB required the correct diaphragm voltage setting which differed from the expected range of 0-30V.

To better understand the discrepancy between the expected and observed range of lens diaphragm voltages leading to efficient transfer of ions to the electrometer PCB, ion trajectories were calculated using SIMION, a standard ion optics simulator. The precise geometry and potential energies of each component of the ion lens assembly was input into

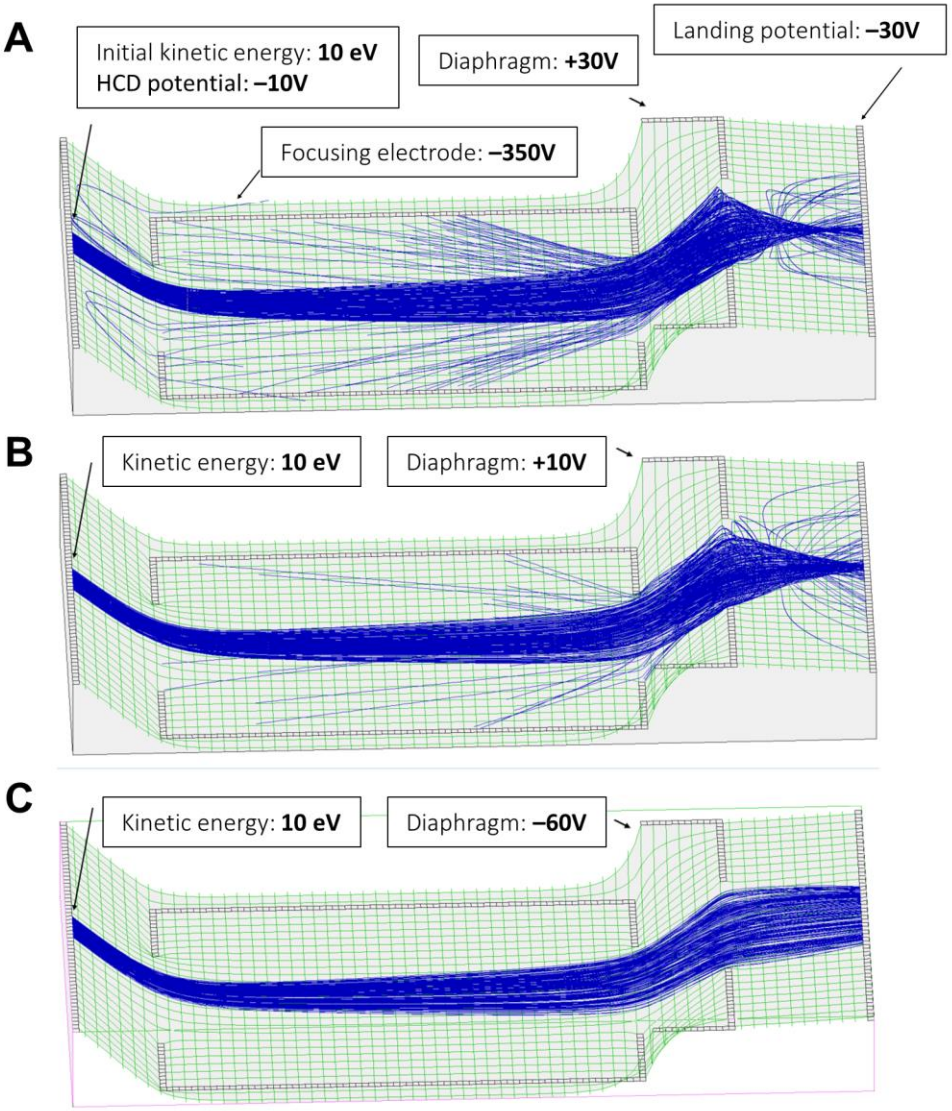


Figure 25. Ion trajectory simulations for lens assembly. The lens assembly geometry was modeled in SIMION, and particle trajectories were calculated and superimposed on potential energy surface plots for various choices of ion kinetic energy and lens diaphragm voltage. **(A)** Configuration using recommended voltages and estimated ion kinetic energy. **(B)** Focused spot size can be moved towards landing surface by reducing diaphragm voltage. **(C)** Further reducing diaphragm voltage maximizes transmission at the loss of a defocused ion beam.

the simulator; however, a key variable dictating ion trajectories are their initial ion kinetic energy, which is unknown. Discussions with Thermo led to an estimate of 10 eV for ions emerging at the back of the HCD cell. With this value, ion trajectory simulations were performed as shown in Figure 25a and demonstrated a focal point slightly in front of the landing surface, along with a small spot size of landed ions, which is qualitatively in agreement with the Thermo simulation shown in Figure 15. Decreasing the diaphragm voltage moved the focal point for the ions backwards toward the landing surface, reaching the surface at $+10\text{V}$ as shown in Figure 25b. Further reducing the diaphragm voltage results in an ion beam focused behind the landing surface as shown in Figure 25c, with total ion transmission increasing at the cost of a defocused ion beam. Similar to the Thermo simulation, the SIMION simulation predicted that ions should have been detected with a diaphragm voltage of $+30\text{V}$, but additionally revealed that ion flux could be increased with reduced diaphragm voltage. This general trend in the data of increased flux with reduced diaphragm voltage was observed in the data as shown in Table 5; the fact that ions were initially observed for a choice of $+10\text{V}$ rather than the expected $+30\text{V}$ could be due to the fact that the ions have less than 10 eV kinetic energy emerging from the HCD cell.

The SIMION simulations additionally confirmed the experimental finding that the ion trajectory was relatively insensitive to the choice of focusing electrode voltage. The choice of -30V for the landing potential is relatively fixed and determines the ion's landing energy, independent of ion beam focus:

$$E_{\text{land}} = E_k^0 + q(V_{\text{HCD}} - V_{\text{land}}), \tag{34}$$

where E_k^0 is the initial ion kinetic energy, $q = ez$ is the charge of the ion, and z is the integer number of elementary charges e where $e = 1.6 \times 10^{-19}$ Coulomb [10]. The choice of

landing potential for the Q Exactive hybrid instrument thus gives a landing energy of 30 eV per charge, and this value is in line with experiments demonstrating *soft landing* – that is, physisorption of intact species onto solid surfaces. Single-molecule NEMS-based analysis requires deposition of intact proteins and protein complexes onto surfaces while preserving their composition and charge. ESI-generated biomolecular ions have low translational kinetic energies (< 100 eV) – ranging from the hyperthermal down to the thermal regimes [3, 11] (Figure 26). To avoid both *dissociative landing* (protein / protein complex fragmentation) and *reactive landing* (formation of strong covalent or electro-static bonds with the surface), soft landing is essential [12-14].

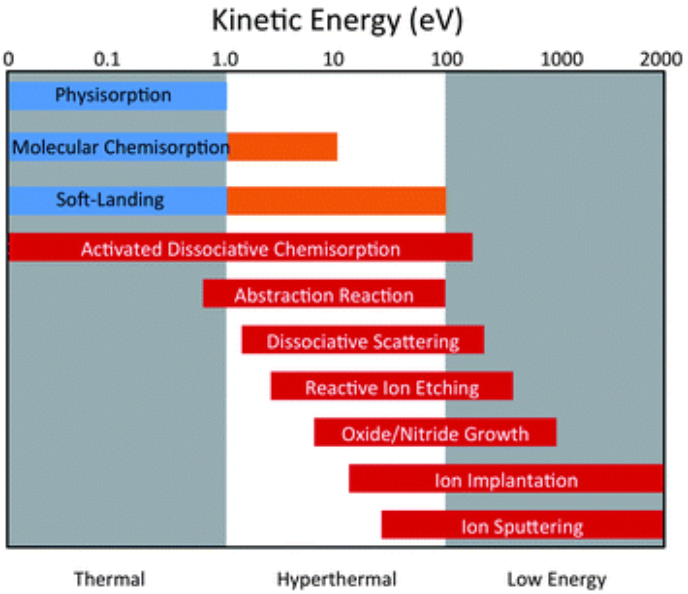


Figure 26. Energy regimes for ion-surface collisions. A variety of molecule-surface interactions can occur at various translational kinetic energies. From [3].

Soft landing of large protein complexes, such as GroEL and apoferritin (801 and 440 kDa, respectively), has been demonstrated using TEM to characterize deposited molecules [15]. It is hypothesized that these species quickly dissipate collisional energy internally given their myriad internal degrees-of-freedom. Preservation of the structure and infectivity of a virus after soft landing has been confirmed by TEM [16], while conservation of biological activity of proteins such as trypsin and lysine after soft landing has been verified with post-collection bioassays, SEM and AFM [17, 18]. In other studies [16, 19, 20], adsorbed peptides and chromophores have been analyzed using secondary ion mass spectroscopy; molecules

were shown to remain chemically active with up to 35 eV impact energy. The Q Exactive hybrid instrument employs similar ion optics and landing energies in these studies, but future work involving orthogonal analytic techniques (cryo-EM and nano-SIMS) will be necessary to validate soft landing and preservation of analyte composition and conformation.

2.5 Native MS measurements

As mentioned in Section 2.1, modification of the Q Exactive for the purpose of delivering intact proteins and protein complexes to the NEMS enables preparation and selection of molecules according to protocols standardized by native MS researchers. In order to have the capability of delivering appropriate samples to NEMS devices in the lab at Caltech, I visited Prof. Albert Heck's lab at Utrecht University and was provided with general training for native MS measurements, as well as specific training with analyzing three molecular complexes in particular: horse heart myoglobin, the *E. coli* chaperonin GroEL, and chaperonin complex GroEL-GroES, each in their intact or denatured state. This section discusses several measurements that were performed at Utrecht under the direction of Drs. Joost Snijder and Michiel Van de Waterbeemd and with assistance from Anja Boumeester that were later replicated with the assembled hybrid instrument at Caltech.

Measurements in the remainder of this chapter utilize nano-electrospray ionization (nano-ESI) [21] rather than the standard electrospray ionization used for the calibration type measurements in the previous sections. Nano-ESI ionization is achieved in the hybrid instrument simply by switching out the front end source (depicted in Figure 14) and using glass needles with a ~2 μm tip opening and coated with metal. Such needles are available from standard suppliers (NewObjective), but the majority of data in this thesis was collected using custom gold coated needles manufactured by Anja Boumeester and Joost Snijder and

generously provided to us. Nano-ESI refers to the flow rate (typically ~20 nl/min [21]) and is the most commonly used ionization method in Native MS due to its low sample use and improved ability to desolve salts and buffers from the sample [22]. All protein solutions were analyzed at a concentration of 5 μM in 50 mM aqueous ammonium acetate.

Myoglobin with heme (holo-myoglobin) is a ~17.5 kDa, 3.5 nm diameter protein complex [23] that can be readily prepared in abundance from a lyophilized powder (Sigma-Aldrich). The ease of preparation of myoglobin facilitated the numerous calibration experiments to be discussed in the remaining parts of this chapter, which were initially necessary to achieve molecular detection with NEMS devices in a well-characterized manner. Its molecular weight, however, is below the mass sensitivity threshold to achieve single-molecule detection with LETI NEMS devices as discussed in Chapter 1. The larger GroEL complex, consisting of 14 identical subunits of ~57 kDa each, for a total of 801 kDa [23], is above the sensitivity threshold and more suitable for this purpose. Preparation of GroEL for native MS measurements, however, is much more extensive than of myoglobin: it requires expression in a genetically engineered culture, as well as multiple rounds of filtration and purification [24]. Prof. Albert Heck generously provided us enough GroEL sample for a limited number of experiments. He also provided GroES, the co-chaperonin composed of 7 identical subunits of ~10 kDa, which combines with GroEL to form a 21-subunit complex in the presence of ATP or ADP [23]. GroEL is a stable complex that has previously been analyzed with native MS using the Q Exactive [2]. Crystallographic models of GroEL and the GroEL-GroES chaperonin complex have been published and are shown in Figure 27. The GroEL complex has a width of 13.7 nm and a height of 14.6 nm [25]. GroES has a diameter of 7.5 nm and a height of 3 nm [26].

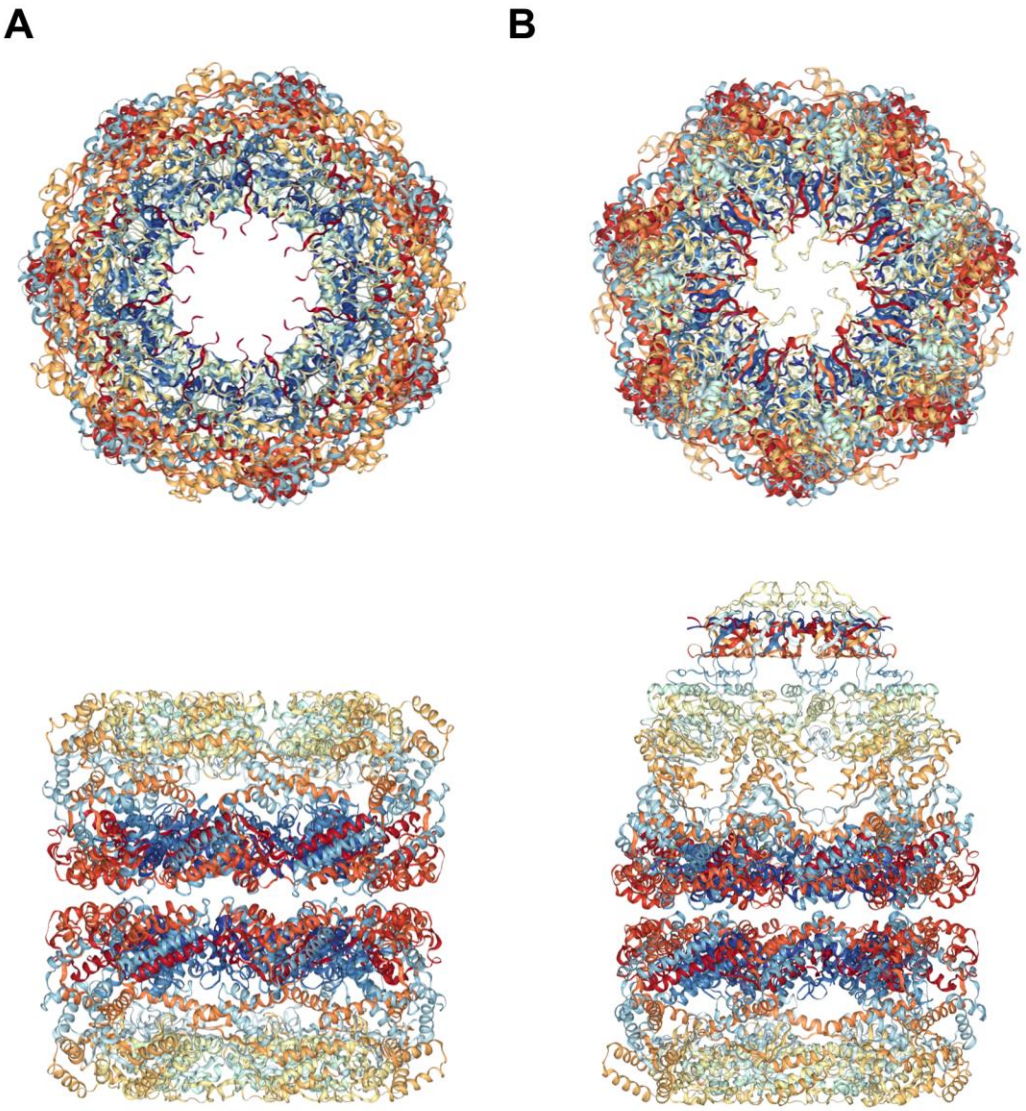


Figure 27. Representation of crystallographic models of GroEL and GroEL-GroES. *E. coli* chaperonin GroEL and chaperonin complex GroEL-GroES are shown in **(A)** and **(B)** with both top and side orientation. PDB accession codes 4V43 and 1AON were used.

Efficient transfer of ions in the Q Exactive is achieved using a combination of instrument settings including resolving power and measurement time, DC and RF voltages of ion optics components shown in Figure 14, and trapping gas pressure in the HCD cell. While every setting could be optimized for each different molecule, a number of general principles have been developed based on the physics of ion trajectories in electromagnetic fields as well as extensive testing by native MS researchers. These settings have been relayed

to me through discussions with researchers in Professor Albert Heck's lab (mentioned previously) as well as with the R&D team at Thermo (Drs. Reinhardt-Szyba, Boll, and Makarov); the most critical settings necessary to replicate the results in this thesis and transfer knowledge to future researchers will be included in this section despite the somewhat technical nature of the discussion. Some settings specific to GroEL, as well as the rationale for these settings, are also discussed in [2].

In general, different approaches are used depending on whether the molecule of interest is small (<20 kDa) or large (>150 kDa). For small molecules, the resolving power is set to the instrument maximum of 140,000 in order to achieve longer measurements with isotopic resolution. Beyond around 150 kDa, isotopic resolution is no longer possible due to the large number of adducts [27], and the resolving power is instead set to the instrument minimum of 8,750 in order to achieve quicker measurements and higher throughput. A longer measurement time is not necessary for larger molecules, as they have reduced orbital stability due to the higher number of collisions with gas molecules in the Orbitrap chamber; this is in turn due to the higher pressure in the Orbitrap chamber that develops as a result of the higher trapping gas pressure used for large molecules, as well as a small leak between the HCD cell and Orbitrap chamber.

The HCD cell trapping gas pressure is lower for smaller molecules and higher for large molecules. There is no direct pressure gauge in the HCD cell or precise calibration for instrument settings vs. pressure; the instrument setting ranges from 1-10, with a setting of 1 corresponding to roughly 3.5E-5 Torr. A setting of ~3 is used for small molecules and 8 for large molecules. The rationale is that larger molecules have more kinetic energy and require higher pressures to stop and trap them in the HCD cell prior to analysis in the Oribtrap.

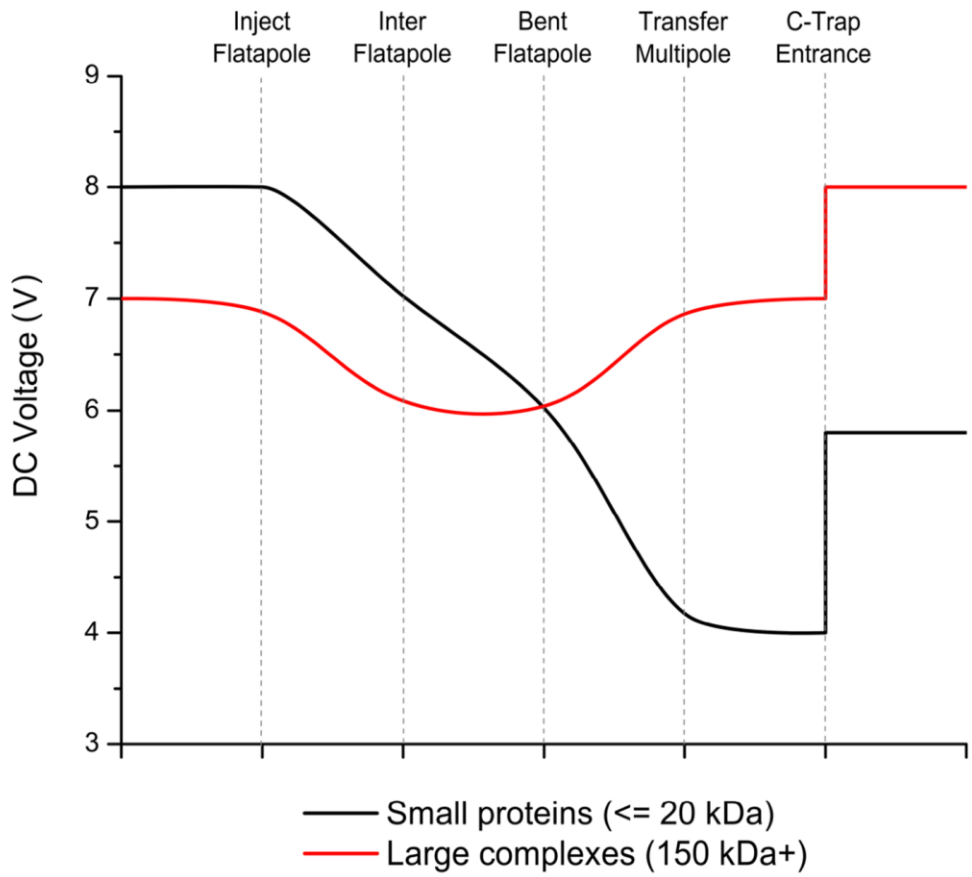


Figure 28. DC voltages used for Native MS experiments. Voltage settings optimizing ion transmission are shown as communicated from the Heck lab. A retarding voltage at the C-Trap entrance lens can be used to block smaller fragments [2].

DC ion optics voltages begin with the nanospray needle set to a spray voltage of 1.4 kV, an S-Lens DC voltage of 50V, and additional consequential DC voltages are set according to the schematic in Figure 28. Of particular note is the retarding voltage set with the C-trap entrance lens, to which ion transmission is particularly sensitive. This retarding voltage is used to block low *m/z* ions or fragments (< 3000 *m/z*) while allowing higher *m/z* ions through; more details are discussed in [2]. In addition to these DC ion optics, the RF and DC voltages of the Quadrupole mass filter are set automatically via the software in order to select ions of interest according to a particular range of *m/z* values. Additional voltages of

interest include the HCD collision energy (positive) voltage at the back of the HCD cell used to trap ions. A typical value is 100V, with higher voltages sometimes able to remove subunits from protein complexes.

A typical mass spectrum that I obtained for holo-myoglobin with the Q Exactive is shown in Figure 29a. The spectrum features a charge state distribution (CSD) of ion signals arising from a variable number of charges (*z*) transferred to the same molecular species. The phenomena of CSDs occurs during ESI and is due to the transfer of protons from the ammonium acetate buffer to chargeable amino acid residues of the protein of interest [28]. The main charge states for holo-myoglobin are labeled in Figure 29a as 7+, 8+, and 9+. Unambiguous determination of these charge states is possible due to the isotopic resolution present in the data inset shown in Figure 29a: each adjacent peak has the same charge state, but differs in mass by one neutron. Two such adjacent peaks are labeled with *m/z* of 2196.81 and 2196.94; these peaks which differ in *m/z* of 0.13 for *m* = 1 corresponds to *z* = 8+.

Alternatively, the charge states can be assigned by examining peaks corresponding to the same species, but differing in charge state [28]:

$$p_1 \times z_1 = M_{protein} + M_{proton} \times z_1 = M_{protein} + 1.0079 \times z_1 \quad (35)$$

$$p_2 \times (z_1 - 1) = M_{protein} + 1.0079 \times (z_1 - 1) \quad (36)$$

$$z_1 = (p_2 - 1.0079)/(p_2 - p_1) \quad (37)$$

where $p_1 = m/z$ peak 1, $p_2 = m/z$ peak 2, z_1 charge of peak at p_1. Using $p_1 = 2196.81$ and $p_2 = 2510.49$ shown in Figure 29a, we obtain $z_1 = 8+$, which matches the prior method. The assigned charge state and m/z peaks can be combined to calculate the mass of

the protein. Averaging the three peaks evident in Figure 29a, we obtain a mass of 17566.40 Da which is close to the theoretical mass of 17566.67 Da [28].

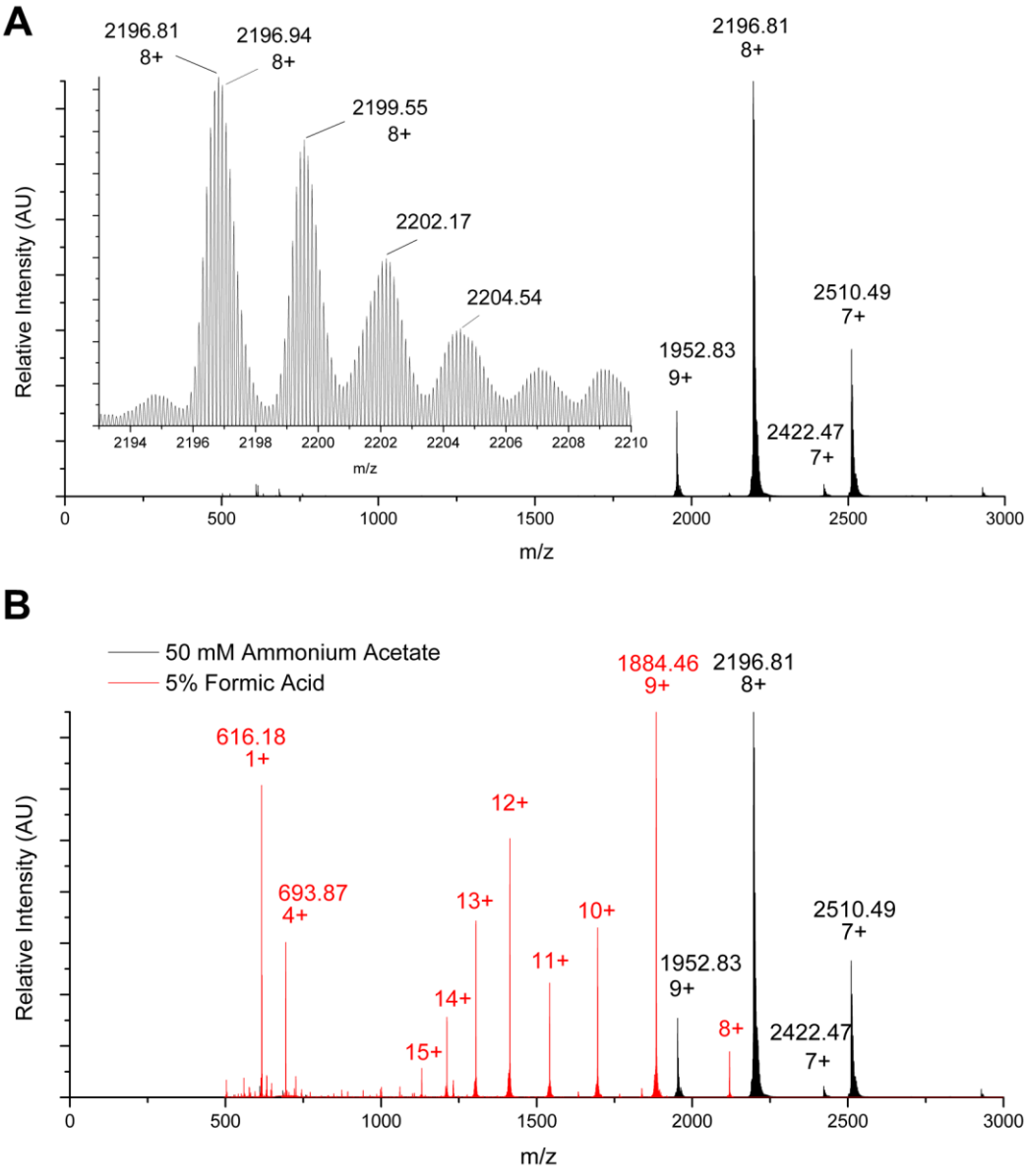


Figure 29. Myoglobin measurements using the Q Exactive. (A) Holo-myoglobin measured in its intact state shows a charge state distribution with the same species producing three ion peaks. The inset shows details of the measurement which features adducts (variable sodium atoms) as well as isotopes (variable neutrons). **(B)** A denaturing solvent can be used to measure apo-myoglobin and heme separated from each other (red). The denatured protein collects more charges compared with the intact protein.

The inset in Figure 29a details not only species differing in isotope, but also the presence of a variable number of adducts. These adduct peaks with *m/z* 2196.81, 2199.55, 2202.17, 2204.54 differ by a variable mass (on average, the difference corresponds to ~21 Da). The most reasonable identity of the main adduct is sodium, which at 22.99 Da is not an exact match, but the discrepancy can be explained by the fact that different isotopes will have slightly different ion transmission efficiencies.

To demonstrate measurement of a protein in its intact vs. denatured state, holo-myoglobin was incubated in a solution of 5% formic acid instead of the usual native buffer, 50 mM ammonium acetate. The mass spectrum is shown in Figure 29b in red and clearly shows apo-myoglobin with a mass of 16951.07 Da and heme with a mass of 615.18 Da (this totals to 17566.24, close to the intact mass calculated previously). Additionally, the denatured complex shows a much larger range of charge states, including higher charge states, a phenomena that can be attributed to the fact that the dissociated molecule has a larger surface area [29]. This simple dissociation experiment demonstrates the possibility of introducing molecules with different conformations through a simple sample preparation protocol prior to measurement with the Q Exactive, a capability that could enhance NEMS single-molecule measurements.

A mass spectrum for GroEL obtained with the Q Exactive is shown in Figure 30a. For this large molecule, the resolution of the measurement (i.e., width of each peak) is no longer limited by the instrument, but instead the variable number of adducts such as water and salt molecules [30]. Additionally, the number of adducts differs for each peak in the CSD shown in Figure 29a due to the fact that the more energetic species with higher charge states have higher kinetic energies, and therefore experience more efficient desolvation. Because

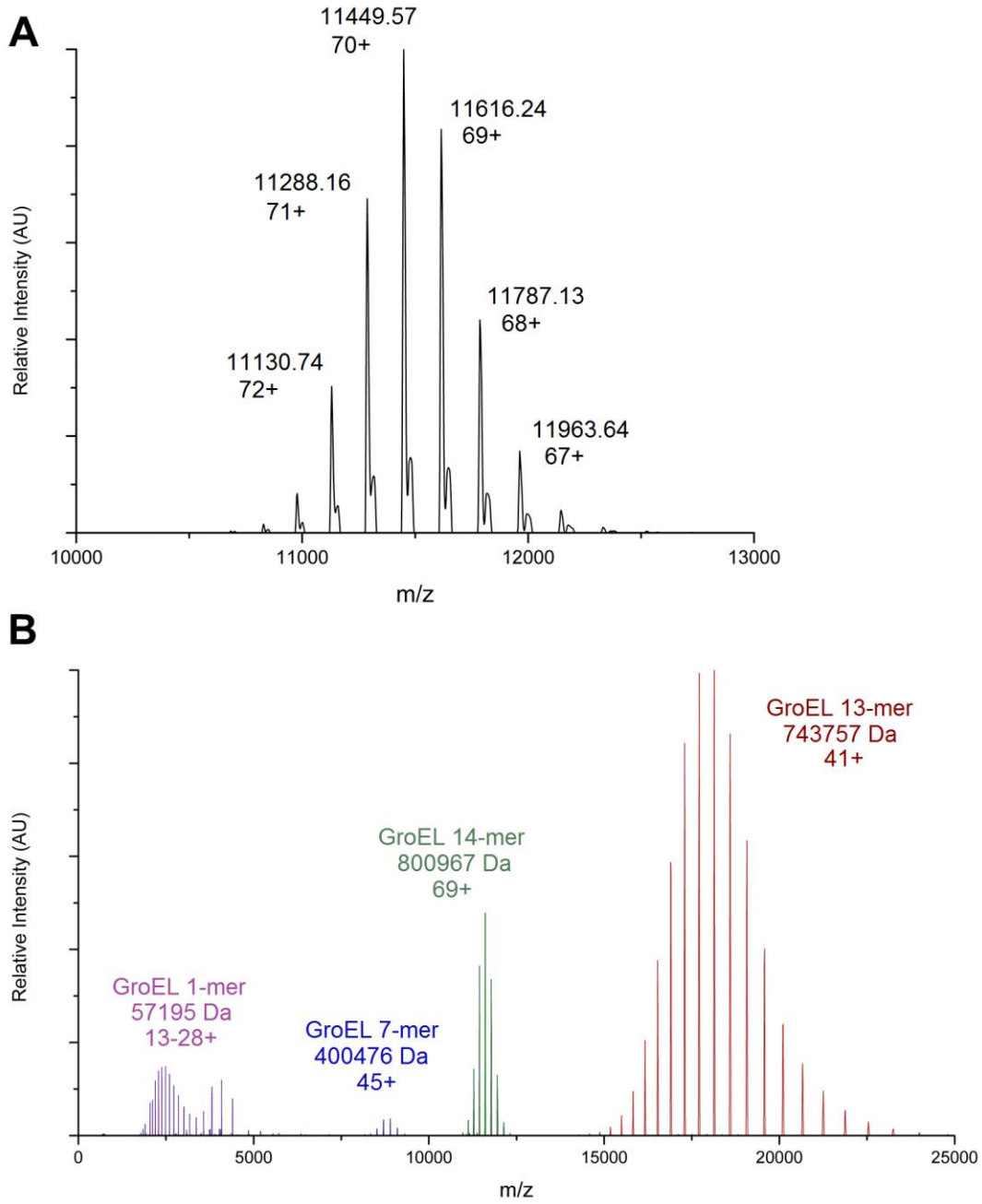


Figure 30. GroEL measurements using the Q Exactive. (A) Intact structure showing charge state assignment minimizing the standard deviation of calculated mass for each peak. **(B)** Mass spectrum recorded under higher collision energy, the effect of which removes a monomer from some of the measured species.

each species in the CSD is not identical, Equation 37 does not correctly assign charge states. Instead, Equation 35 or 36 could be used to calculate $M_{protein}$ for a *guessed* value of z for each peak (with adjacent peaks differing by one charge). The guess that minimizes the

standard deviation for $M_{protein}$ for all peaks is assumed to be the best choice [31, 32]. Using this procedure, the charge states for GroEL are assigned as demonstrated in Table 6 and shown along with the *m/z* peaks in Figure 30a. The mass of GroEL calculated using this procedure was 801.421 kDa, which is close to the expected value of 800.7664 kDa [23].

Increasing the collision energy to 200V as shown in Figure 30b removes a monomer from some of the 14-mer GroEL complexes, yielding the 1-mer GroEL, the 13-mer, the 14-mer, as well as a small amount of the 7-mer (evidently, at these energies, GroEL can be split in half yielding a 400 kDa complex as well). After charge state assignment, the detected 14-mer has a calculated mass of 800.967 kDa, which is closer to the expected molecular weight than the previous measurement at lower collision energy. This suggests that more adducts are present at the lower collision energies, but can be removed with higher collision energies. Finally, it can be noted that, when the 14-mer is split, charge is nearly evenly split between

Charge state guess (dominant peak)	GroEL calculated mass (standard deviation)	GroEL calculated mass (average)
72+	619.42 Da	824.498 kDa
71+	335.04 Da	812.960 kDa
70+	**51.38 Da**	**801.421 kDa**
69+	234.18 Da	789.884 kDa
68+	518.52 Da	778.345 kDa

Table 6. Charge state and mass assignment for GroEL. Larger species such as GroEL are not isotopically resolved with the Q Exactive due to the large number of adducts present. Charge state assignment is performed by selecting that which minimizes the standard deviation of the calculated mass of the species (shown in bold).

the 13-mer and the 1-mer, rather than being split proportional to mass.

Interrogation of protein complexes in the native state by the Q Exactive enables observation of assembly and stability of these complexes. As a demonstration, 160 μM Mg·ATP was added to 2 μM GroES, then this solution was added to 1 μM GroEL. This promotes the assembly of the 14/7 GroEL/GroES chaperonin complexes; to study the dynamics of assembly, the solution was incubated on ice and analyzed with the Q Exactive after 5 minutes and then again after 30 minutes as shown in Figure 31a and Figure 31b. After 5 minutes, the partially assembled complex with 7/7 GroEL/GroES was the dominant species; also present was the 14-mer GroEL and various partially assembled GroEL/GroES complexes with 14 GroEL subunits.

After 30 minutes incubation, the fully assembled 14/7 GroEL/GroES chaperonin complex becomes the dominant structure. This solution was then analyzed in the Q Exactive with high collision energy (200 V) as shown in Figure 31c. A monomer of GroES was present in the mass spectrum, along with 14/6 GroEL/GroES, indicating that a GroES subunit is the least strongly bound of the intact 14/7 GroEL/GroES chaperonin complex.

The color-coding of the mass spectrum and determination of the ratio of GroEL to GroES units in each of the experiments shown in Figure 31 is not unambiguous as indicated. Previously, a method was discussed to assign charge states, and therefore mass, based on minimizing the standard deviation of the calculated mass of each charge state belonging to the CSD of an intact complex. The inability to unambiguously rely on this calculation is demonstrated in Figure 32 in detail for the presumed 14/7 GroEL/GroES complex indicated in Figure 31b. Additional information from tandem MS/MS studies is necessary to fully resolve the identity of the mixture.

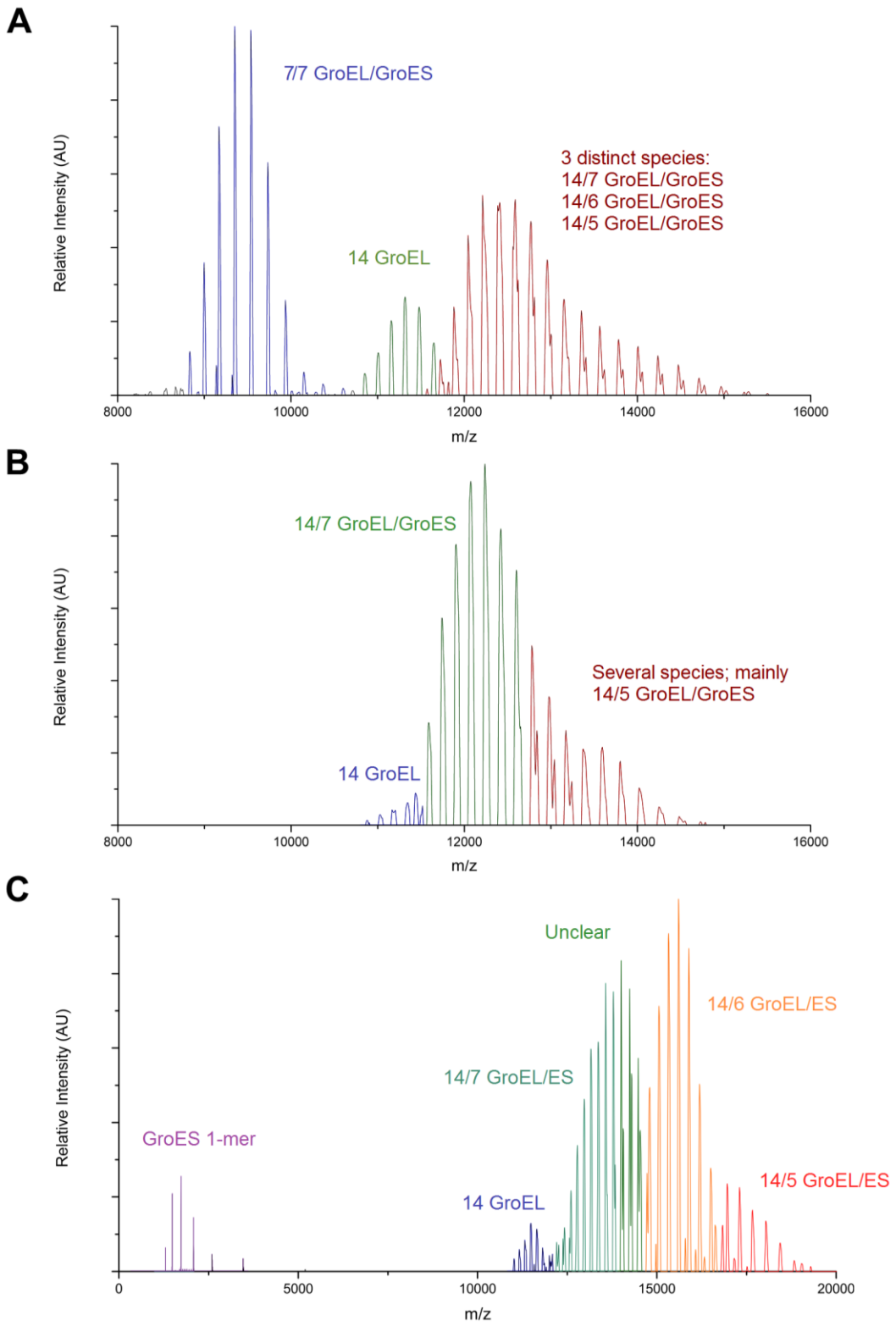


Figure 31. GroEL-GroES measurements using the Q Exactive. GroEL and GroES incubated with Mg and ATP on ice for **(A)** 5 minutes and **(B)** 30 minutes measured with low collision energy and **(C)** 30 minutes with high collision energy

Ambiguous charge state (and hence mass) assignment for heterogeneous complexes such as the 14/7 GroEL/GroES with a large, variable number of adducts as shown in Figure 32 is a common limitation of the techniques of Native MS. New methods to improve

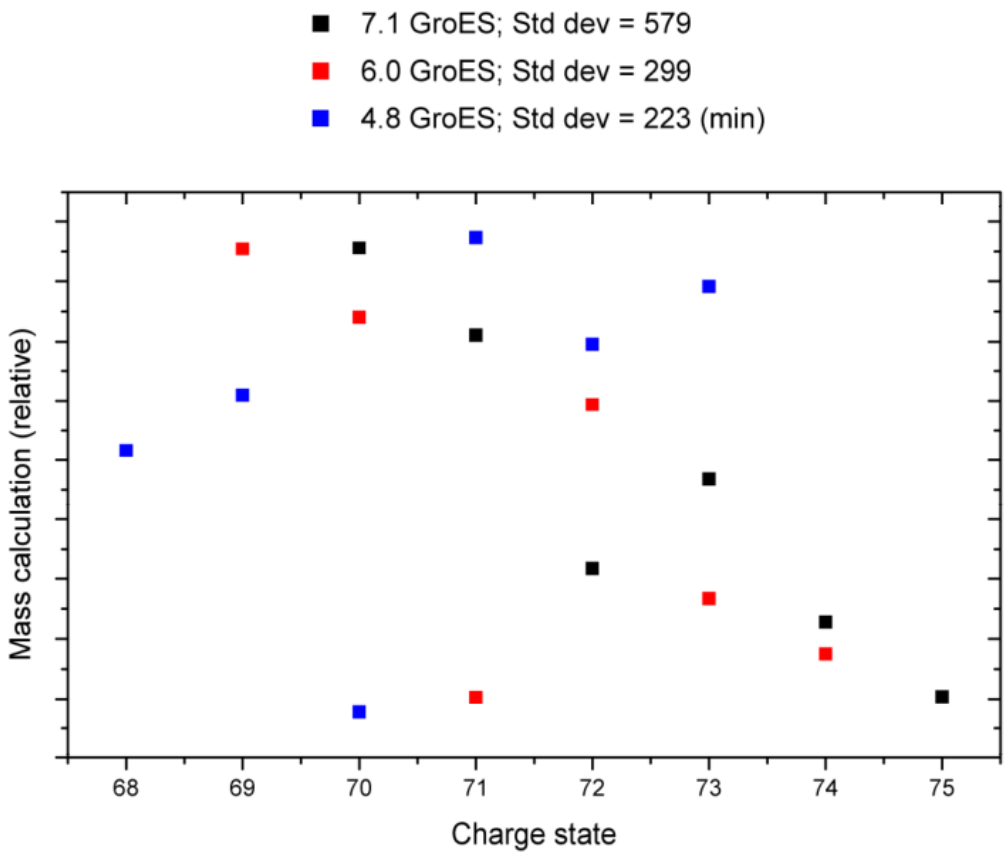


Figure 32. Charge state assignment for dominant GroEL/GroES complex shown in Figure 31b. Correct charge state assignment of the peaks identifying the dominant structure shown in Figure 17b as 14/7 GroEL/GroES based on the total calculated mass is shown in black. This charge state assignment does not minimize the standard deviation of calculated masses for each peak as shown in Table 3, but does show the expected trend of decreased mass with higher charge state (due to more efficient desolvation of higher kinetic energy complexes) and matches the known stable crystallographic data. Alternative contenders for charge state assignment are shown in blue, which minimizes the standard deviation of calculated mass, but shows an increase in mass with charge state, or red, which also has a low standard deviation and the correct trend of mass versus charge state, but incorrectly identifies the complex as 14/6 GroEL/GroES. The choice between identifying the complex as 14/6 in red and 14/7 in black could be considered ambiguous from this data alone; additional information such as tandem MS/MS is necessary to unambiguously identify the subunit architecture.

desolvation, such as the use of in-source trapping has been used to reduce heterogeneity and increase the mass range of the Q Exactive [33, 34], but this type of limitation is likely to persist due to the nature of inferring mass indirectly from a series of unknown charge states. Monitoring the absolute mass of partially assembled GroEL/GroES complexes one-by-one directly with NEMS devices to resolve the ambiguity could therefore present a new opportunity. Such an experiment would require NEMS mass resolution on the order of a single GroES subunit (10 kDa), which our group has recently exceeded (Jarvis Li, 2019, unpub.).

2.6 Molecular detection with NEMS operated at cryogenic temperatures

Previously, delivery of calibration ions to basic circuitry mounted on the cryo-positioner was detailed in Section 2.4, and introduction and detection of intact protein complexes using the Orbitrap was detailed in Section 2.5. Provision of additional electrical cabling and enhanced sample and NEMS stage thermalization was necessary to enable operation of NEMS devices at cryogenic temperatures, for which significantly reduced phase noise has been reported [35]. Low temperatures also promote physisorption of the analyte [8]. A Janis ST-100 continuous flow cryostat compatible with both liquid nitrogen and liquid helium was adapted for this purpose. The cryostat is outfitted with an electrical head consisting of SMA feedthroughs; custom UT-85 stainless steel cabling was manufactured to mate with these feedthroughs and carry external RF signals to PCBs mounted and thermalized to a cold stage consisting of gold plated copper. Due to the need to operate the XYZ positioner with minimal external torque from any attached cabling, highly flexible braided copper cables (Lakeshore) were constructed to continue the RF signal from these cooling PCBs further to a NEMS PCB mounted on the XYZ positioner, which is in turn

mounted to the cold stage. Care was taken to use low outgassing materials, such as Kapton insulating tape, Viton O-rings, OFHC copper, and PEEK for plastic. An overview schematic showing the NEMS mounting configuration is shown in Figure 33.

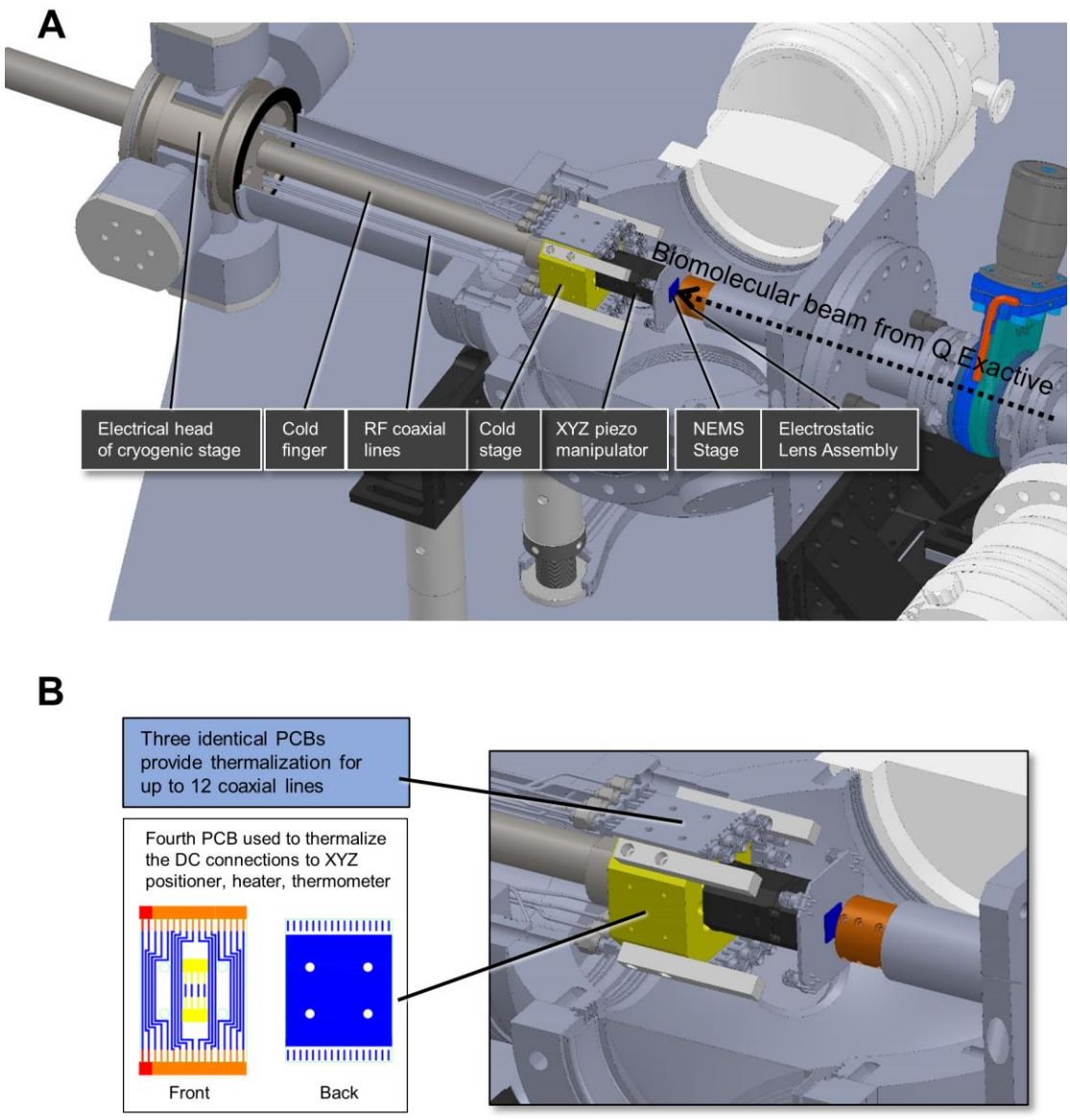


Figure 33. NEMS thermalization via cryogenic cabling and PCBs. (A) A Janis ST-100 continuous flow cryostat thermalizes a cold stage consisting of a copper block constructed such that the surface of the NEMS chip (shown in blue) is positioned 5 mm away from the ion lens assembly (shown in orange) at the center of the positioner's travel range. **(B)** Cooling PCBs help thermalize the inner conductor of up to 12 coaxial lines as well as the DC connections used to operate the XYZ positioner, heater, and thermometer.

DC leads include the 15 needed for the XYZ positioner as well as the 6 needed for the 4-wire thermometer and Ohmic heater. These were thermalized to the central shaft of the ST-100 using standard techniques [36], including the use of a thin layer of cigarette paper underneath the wires to prevent shorts, and cryogenic varnish (General Electric) to secure the leads in place (Figure 34a), as well as supplemental gold-coated OFHC copper bobbins pressed into the cold stage. The back side of the cooling PCBs were also gold coated, and a second blank PCB (with dielectric facing the cooling PCBs) was screwed in on top as a second layer in order for the cooling PCBs to have good physical contact with the cold stage (Figure 34b). Gold coated copper ribbons thermally anchor the top of the XYZ stage to the cold stage.

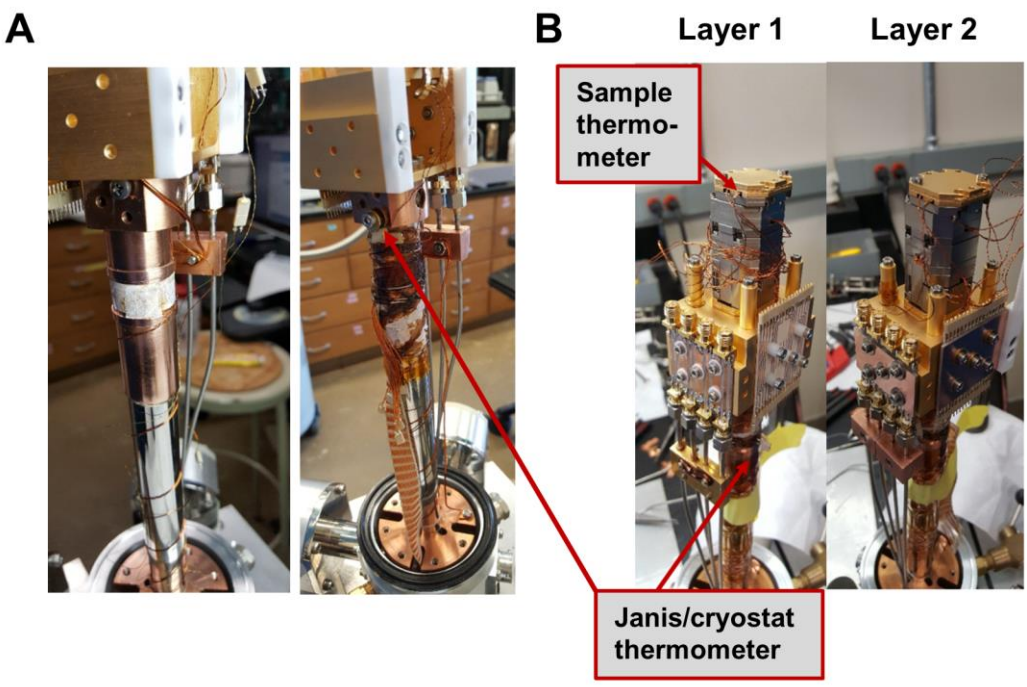


Figure 34. Cryogenic cabling and PCBs, as constructed. (A) DC wires are wrapped around the central cooling shaft of the ST-100 over thin paper and secured in place using cryogenic varnish. **(B)** Cooling PCBs are thermalized to the gold-plated cold stage in two layers. The first layer consists of a thin stripline PCB with gold coating on the back; the second layer is a thick, rigid, blank PCB with copper on the outside and dielectric facing the stripline PCB, with maximum torque applied to hex screws going through both PCB layers.

A custom PCB was constructed for the NEMS containing 132 through holes consisting of silver vias thermally connecting the NEMS chip to the cold stage (Figure 35). RF signals are carried to the NEMS with the use of wirebonds to silver pads, which connect to SMA jacks. Thermal contacts additionally anchor the PCB to the cold stage, and a radiation shield was built to optionally enclose the NEMS and XYZ positioner. Without the radiation shield, base temperature as measured by the sample thermometer immediately

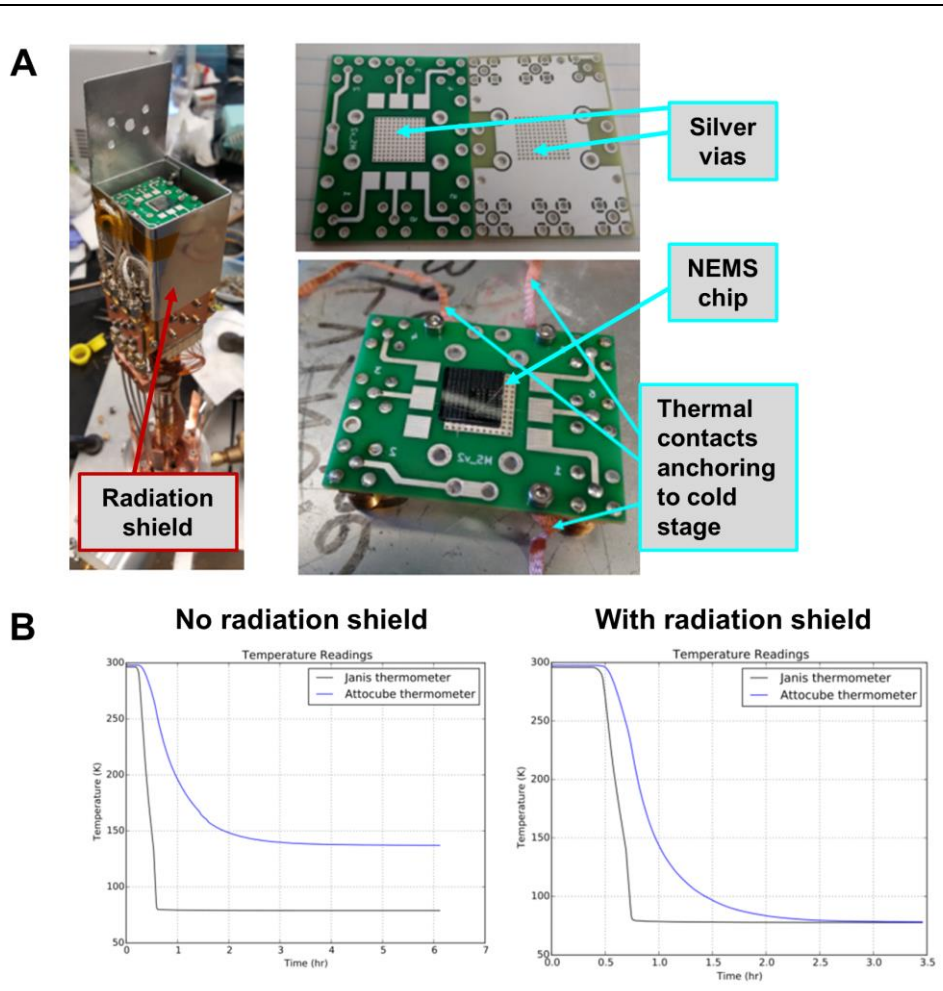


Figure 35. NEMS PCB, radiation shield, and base temperature readings. (A) The NEMS chip is mounted to a custom PCB using silver paste. The PCB features multiple silver vias to improve heat transfer from the NEMS chip to the cold stage. A radiation shield can optionally be installed to surround the NEMS and the XYZ positioner. **(B)** Temperature readings during system cooldown show that base temperature of the NEMS PCB ("Attocube Thermometer") is reached within 2 hours. Base temperature was 145K without the radiation shield, and 80K with the radiation shield.

under the PCB was 145 K with the use of liquid nitrogen. The radiation shield significantly improved this base temperature to 80 K with liquid nitrogen and 10.2 K with liquid helium.

The cryogenic cabling was tested using a LETI NEMS device (CAL3 PP0) actuated and detected under a closed loop PLL at resonance as discussed in Chapter 1. Allan deviation was measured versus integration time for a variety of temperatures from 300 K to 87 K using a Lakeshore temperature controller under PID control with liquid nitrogen cooling, allowing

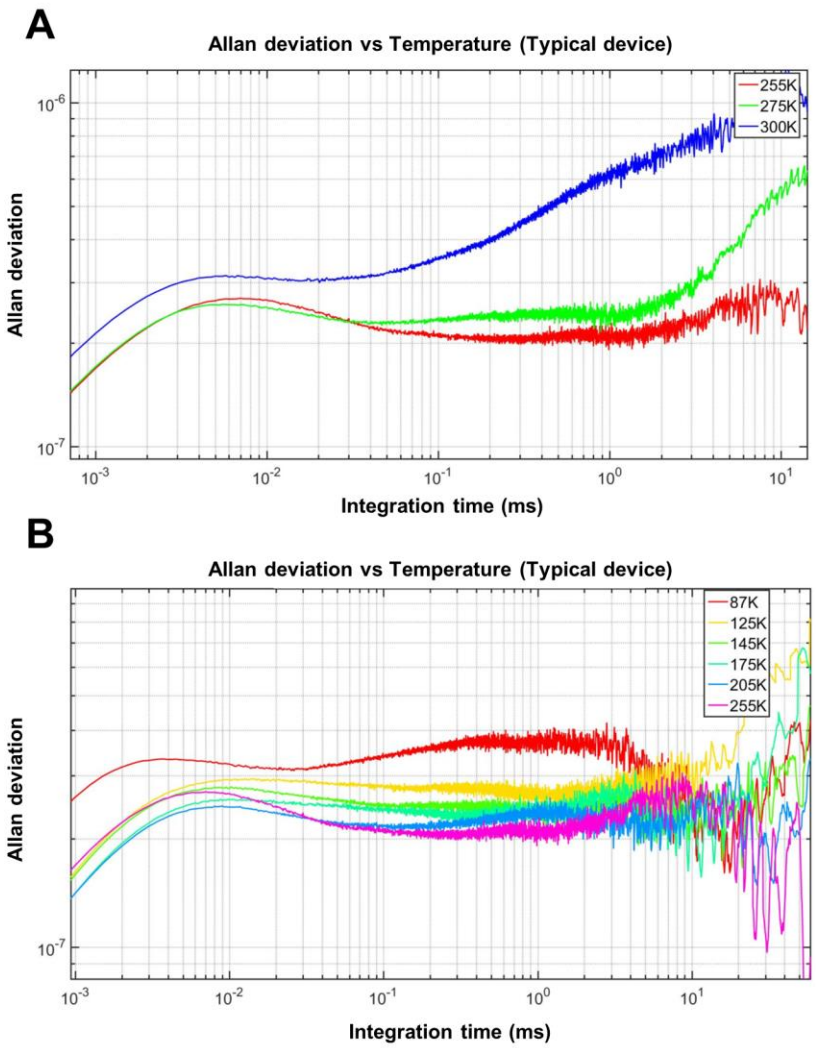


Figure 36. Allan deviation for a typical device versus stage temperature. A CAL3 PP0 LETI device was thermalized to the cryogenic stage and actuated and detected at resonance under a PLL. **(A)** Allan deviation vs integration time from 300 K to 255 K. **(B)** Allan deviation under further cooling below 255K.

at least 4 hours for the device to reach thermal equilibrium with each temperature change. A significant improvement in Allan deviation was observed cooling the device from 300 K to 275 K, with the lowest Allan deviation observed at 255 K. Further reductions in temperature began to show increased Allan deviation (Figure 36), in contrast with the expected trend [35]. This suggests that operation of the device may lead to device self-heating above the stage temperature. This has been subsequently confirmed by recent experiments in our group (Jarvis Li, 2019, unpublished). Replacing the piezosemiconductor-based nanogauges with piezometallic material is one possible way to mitigate this effect; alternatively, the device could be operated at lower bias voltages (see Chapter 3).

With the NEMS actuation and readout cabling in place, and the capability of introducing intact proteins into the Orbitrap verified, the electrometer PCB was mounted on the positioner, and the positioner scanned to check the ion current of myoglobin and GroEL and determine the approximate location of ion focusing of GroEL to narrow down the search region for later experiments using the NEMS (Figure 37).

The NEMS was operated according to the circuit schematic shown in Figure 12, except the entire device was floated to –15V with the use of bias tees or capacitors as appropriate in order to attract the positively charged ions with a lower DC potential than the Q Exactive, which is designed around having the C-trap at ground. After multiple days sending GroEL ions to a device with this arrangement, only a single adsorption event was detected (Figure 38a), even with the device centered at the location of peak ion flux recorded by the electrometer.

An alternative configuration was tested as shown in Figure 38b, where a large number of internal voltages of the Q Exactive were adjusted around having the C-trap at 30V

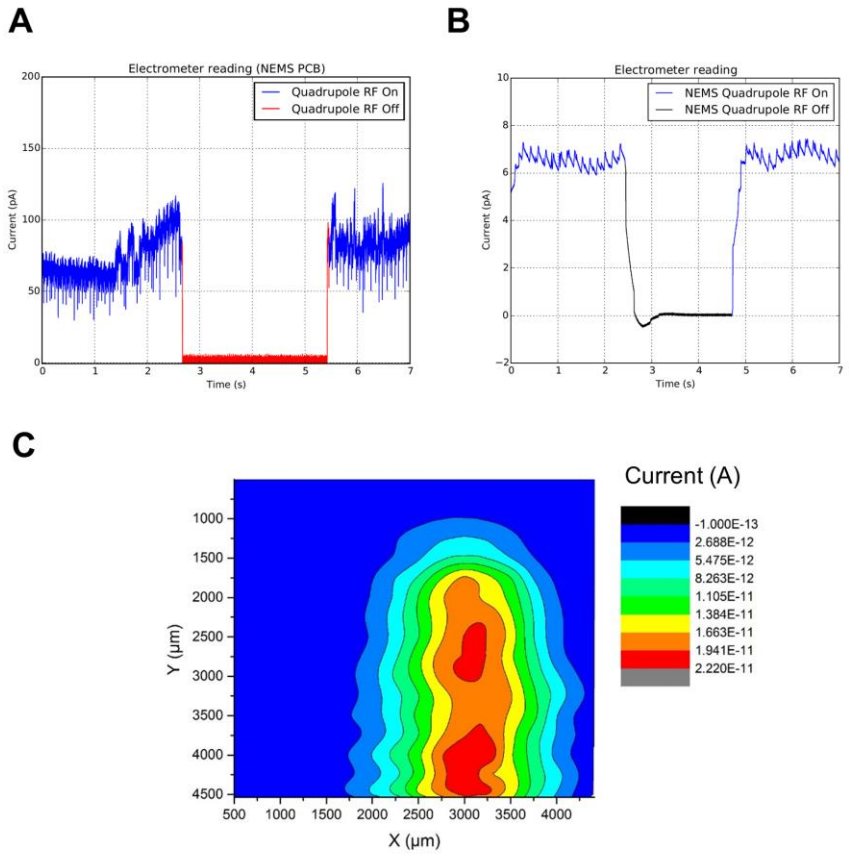


Figure 37. Detection of molecules on PCB electrometer sensor. (A) The PCB electrometer sensor was positioned until maximum ion flux was achieved with myoglobin. A current of 50-100 pA was detected. **(B)** Similarly, a current of 6-7 pA was detected for GroEL. **(C)** The X and Y axes of the positioner was scanned, showing ion current vs. positioner coordinate. The detected ion current mirrors the geometry of the electrometer electrode as shown in Figure 24a.

(thus allowing the NEMS to operate at ground). These internal voltages were chosen by Dr. Dmitry Boll of Thermo Fisher and were found to optimally transmit CsI ions in their instrument in Bremen. Once the instrument at Caltech was configured this way, a large event rate (a few per minute) of GroEL ions adsorbing to the NEMS could easily be detected, after making two important changes with some experimentation here; without these changes, very few ions can be transmitted: (1) the C-trap entrance lens inject is a sensitive parameter that must be set to 19 or 20V (it is shown as 14V in the figure above), and (2) the trapping gas

pressure must be set to 10. The adjustment of the C-trap entrance lens to 1V higher than the flatapole voltage is consistent with the choices typically used for large molecules as communicated by the Heck lab and shown in Figure 28. For the trapping gas pressure, typically 8 is used for GroEL, and higher values can be used to dissociate GroEL shown in Figure 30b. Here, the instrument maximum setting of 10 must be used, possibly due to the need to stop the higher kinetic energy of the ions, with dissociation no longer possible.

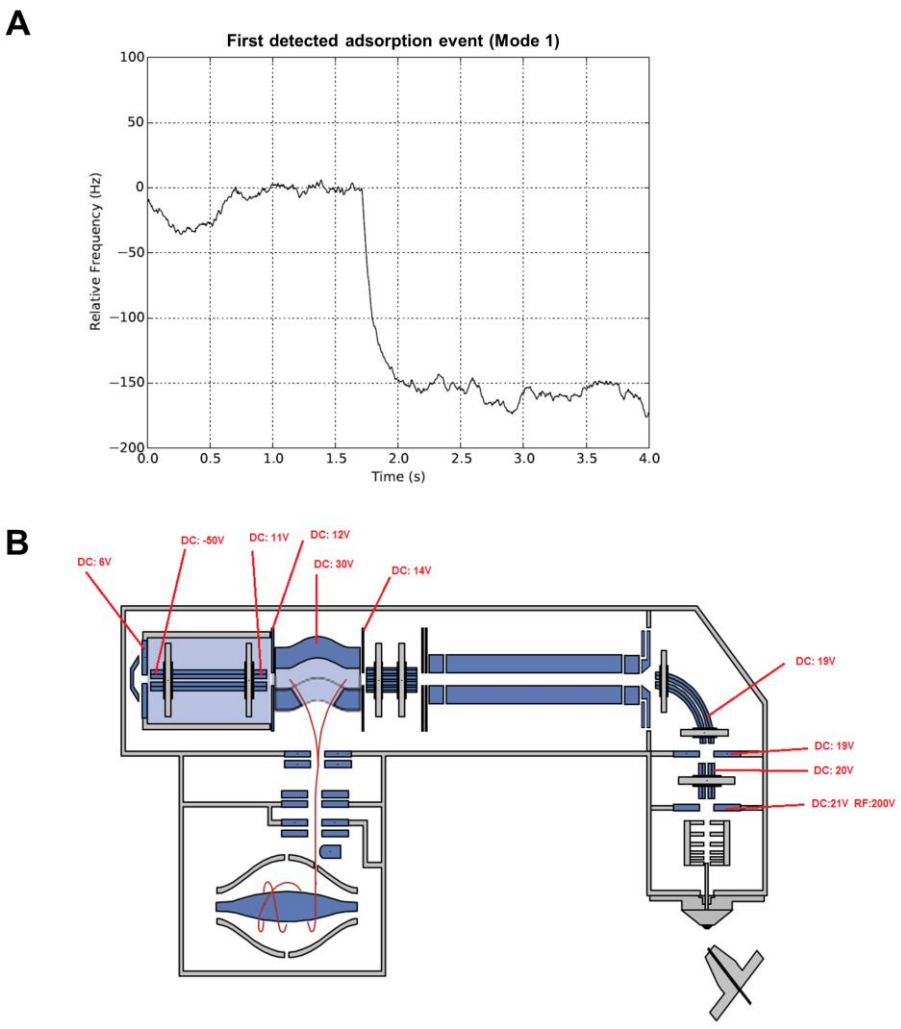


Figure 38. Single molecule detection with NEMS. (A) First molecular detection event of GroEL adsorbing onto a NEMS resonator in the hybrid Orbitrap system. **(B)** Voltages adjusted internally in Q Exactive to allow for operation of NEMS at ground. Voltages shown are optimized for CsI ion transmission; some changes (as shown) were needed for GroEL transmission as discussed in the text.

2.7 Charge detection and retention using NEMS

With the Q Exactive voltages raised as shown in Figure 38b and the NEMS operated at ground, a large number of GroEL adsorption events could easily be obtained. The first two resonant modes of a device were simultaneously tracked, in similar fashion to the schematic shown in Figure 12 except that an AC drive and bias is used for each mode and combined with power splitters, and the output signal is read by a separate lock-in amplifier for each mode. Initial data unexpectedly showed positive and negative frequency jumps upon molecular adsorption (Figure 39a). This is in contrast with only negative frequency jumps predicted by Equation 21, which assumes flat or point-like adsorbates increasing the kinetic energy of a device, and is observed in other NEMS-MS systems constructed in the Roukes group [8, 37] as well as other studies monitoring mass adsorption to nanomechanical

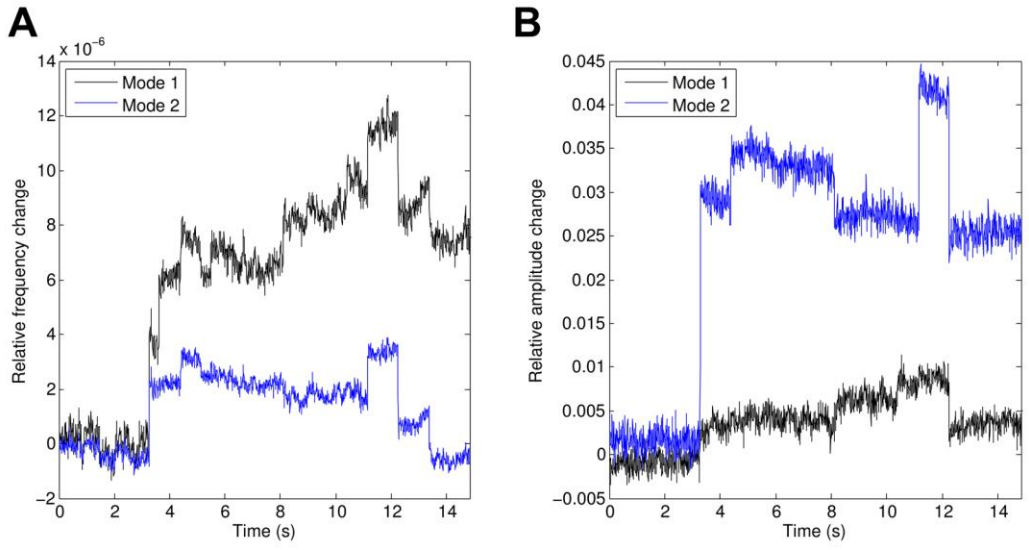


Figure 39. Real-time detection of GroEL using two simultaneous resonant modes. **(A)** As GroEL adsorbs onto a NEMS resonator, the resonance frequency of each tracked resonant mode abruptly shifts, and this shift in resonant frequency is observed in real time. **(B)** Initial GroEL adsorption events were also associated with large changes in device amplitude. Reported are the relative changes, i.e., $\Delta f/f$, $\Delta A/A$.

structures. Carefully inspecting the data, I noticed large swings in device amplitude as well (Figure 39b), which has not previously been reported or observed.

The large amplitude shifts that occur during adsorption are problematic due to the nonlinear nature of the devices. A demonstration of this nonlinearity is shown in Figure 40. The DC drive voltage was increased in 1V increments every few seconds while tracking the device's resonance frequency. Every 1V increase in the DC drive corresponds roughly to a relative amplitude change ($\Delta A/A$) of 3% and 5% for the first two modes, respectively, corresponding in turn to a relative frequency change ($\Delta f/f$) of –2.5E6 and 5E6. The devices have a mass sensitivity of 12 Hz/ag and 32 Hz/ag for the first two modes [38], or 20 Hz/MDa and 50 Hz/MDa, respectively. With resonance frequencies typically around 25 MHz and 67 MHz for the first two modes, such large amplitude changes could indirectly drive frequency shifts on the order of that expected from mass loading of single molecules of GroEL.

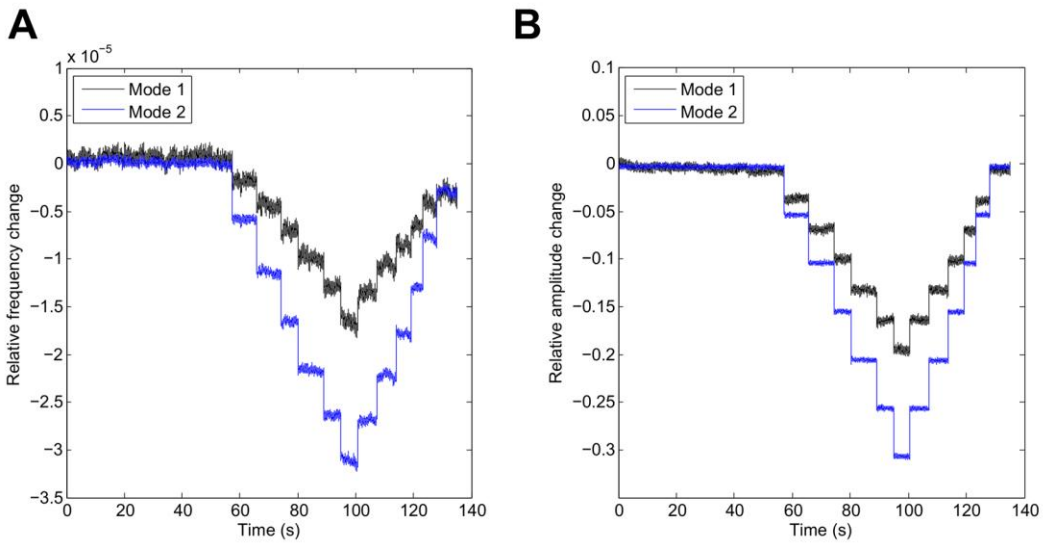


Figure 40. Frequency and amplitude changes due to applied DC gate voltage. The DC drive voltage was increased in 1V increments every few seconds while simultaneously tracking the resonance frequency of the device's first two modes. The changes in **(A)** resonance frequency and **(B)** amplitude are shown. Reported are the relative changes, i.e., $\Delta f/f$, $\Delta A/A$.

The large number of charges (~70 for each GroEL molecule) accumulating on the beam and electrode was hypothesized to modulate the DC potential between gate and beam. This would increase the actuation force (Equation 23) and, in turn, the resonant amplitude (Equations 24-26). The nonlinear effect transforming the increased amplitude to frequency could possibly be due to elastic hardening induced by the potential difference between the active area and bulk [39].

In order to more systematically measure the amplitude and frequency dependency of accumulated charged molecules on NEMS devices, myoglobin (8+, 17.5 kDa) was sent to a 20-device NEMS array positioned in the ion beam path. These arrays consist of 20 NEMS devices, each with the same width and thickness as the individual devices discussed so far, with the device lengths within the array varying from 7 μm to 9.2 μm. This separates their individual resonance frequencies so they can be tracked individually. Their operation is virtually identical to that used for single devices; details of their fabrication and operation are discussed in Chapter 3 and [38]. The device array used was pristine; no molecules had been sent prior to the experiment. Collection of data with arrays was performed with Dr. Eric Sage.

Myoglobin adsorbing to the array initially caused the resonance frequencies and amplitudes of each device to increase, until the amplitude increase reached a point of saturation (100-200s), after which the resonance frequencies of each device began to decrease continuously in a linear manner (Figure 41ab). Myoglobin is much smaller than the mass resolution of these individual NEMS devices, so discrete adsorption events could not be detected. However, the rate of decrease of the NEMS resonance frequency is proportional to the molecular flux rate observed by each device. The ion flux rate based on the known mass sensitivity of each mode and device differed for each device due to the non-uniform

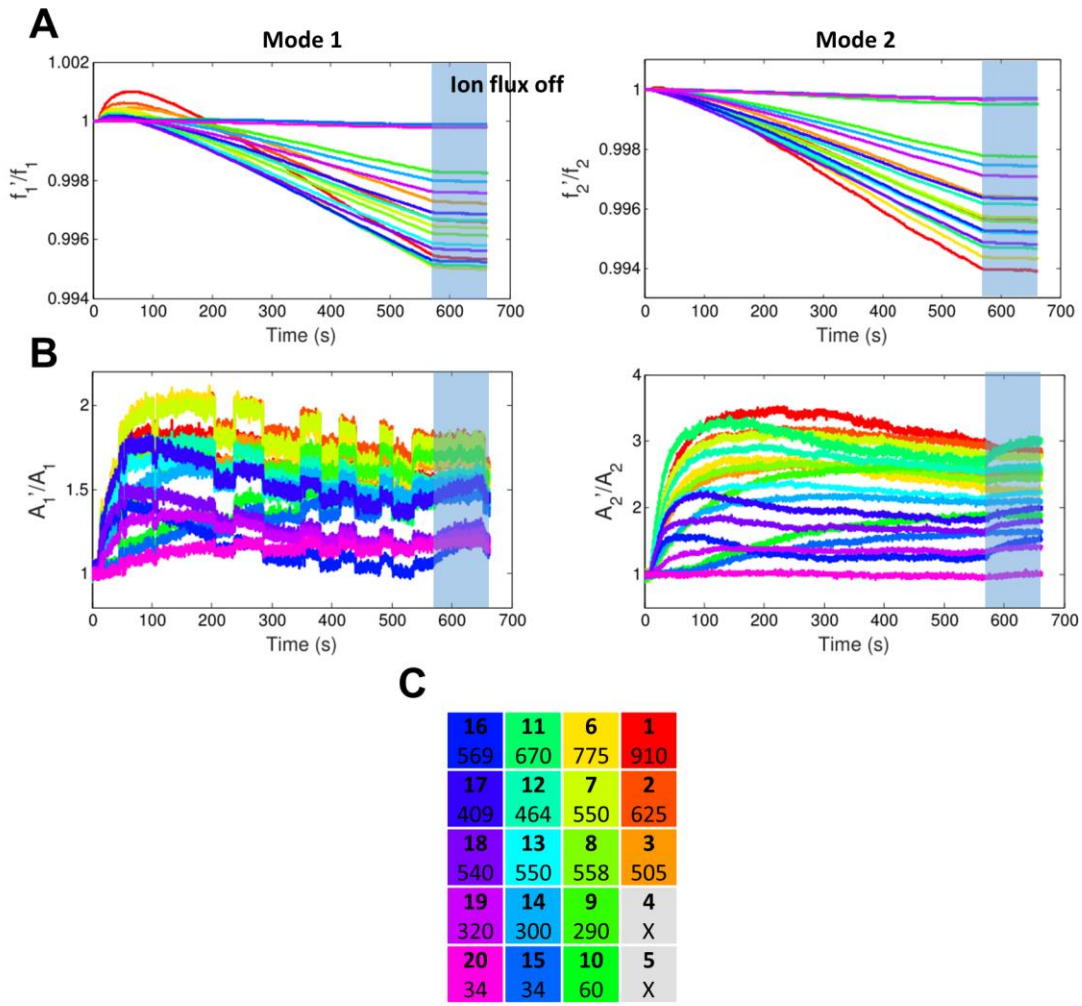


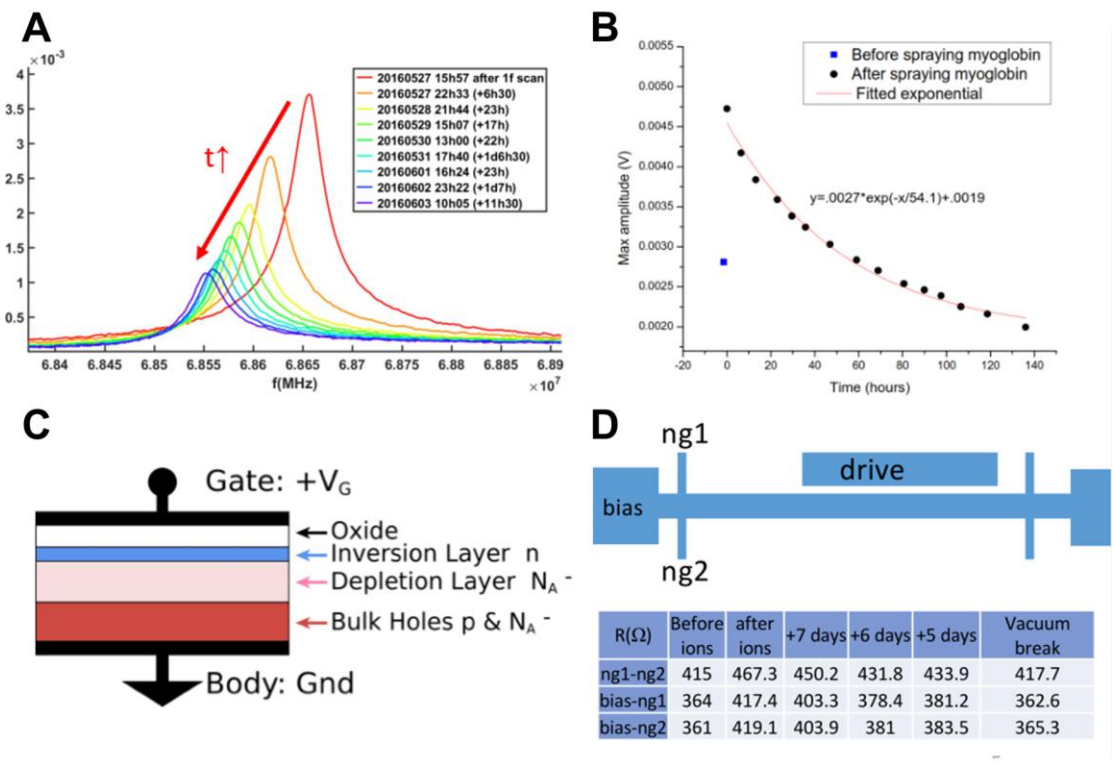
Figure 41. Resonant frequency and amplitude change during continuous myoglobin deposition to a 20-device array. The **(A)** frequency and **(B)** amplitude change of the first two modes is shown for 18 working devices in a 20-device array during continuous adsorption of charged molecules. A map of the devices in the array is shown in **(C)**, with the device number 1-20 shown in bold and the adsorption rate in ion/s shown below the device number. Device 1 corresponds to the longest device at 9.2 µm and device 20 corresponds to the shortest device at 7 µm. The colors match the plots in (A) and (B).

ion beam and is shown in Figure 41c. The amplitude increased significantly for the longer devices, with the longest device having an amplitude increase by a factor of 2 for the first mode and a factor of 3 for the second mode. The amplitude increase appears not to occur monotonically with device length and is independent of the ion flux rate; the effect likely depends on the specifics of the gate electrode and beam geometry and how it relates to mode

shape. Because of this complication, evaluation of the resonant amplitude increase and resulting nonlinear effect on frequency driven by adsorption of charged molecular species will be the subject of further research.

The time required for amplitude saturation combined with the calculated ion adsorption rate yields ~1E5 ions for saturation. Assuming that each myoglobin molecule covers 4 nm^2 surface area, ~1.5E5 molecules are required to form a single layer on top of the beam surface. Therefore, I hypothesize that the amplitude saturation effect may occur when a monolayer of myoglobin fully covers the beam and shields further incoming charges from modulating the DC potential between the beam and gate.

Once the resonant amplitude reached this saturation level, the new resonant amplitude was found to be very stable. The native oxide layer on the device likely prevents charge leakage into the bulk, therefore allowing for the preservation of molecular charge upon adsorption to the device. This is a potentially useful finding, because if the molecules can be removed from the NEMS devices following their measurement, the charges could potentially be preserved as well, allowing for the molecules to be measured in a mass spectrometer after they are monitored with NEMS devices. In this way, the NEMS devices could be used to pre-stratify like molecules based on some physical property independent from their mass spectra, which suggests the new approach to single-molecule proteomics discussed in Section 1.4. To monitor charge retention systematically, the resonant amplitude of each device was monitored for a week in the same array used in the myoglobin adsorption study above, as shown in Figure 42. Both the resonant amplitude and the electrical resistance of the nanogauges decayed with a half-life of around 50 hours, indicating that at least a



R(Ω)	Before ions	after ions	+7 days	+6 days	+5 days	Vacuum break
ng1-ng2	415	467.3	450.2	431.8	433.9	417.7
bias-ng1	364	417.4	403.3	378.4	381.2	362.6
bias-ng2	361	419.1	403.9	381	383.5	365.3

Figure 42. Retention of molecular charge on NEMS devices. (A) An open loop sweep of a single device in the array following adsorption shows gradually decreasing amplitude; **(B)** the amplitude decays with a half-life of 54 hours. **(C)** Positive charges applied to the surface of a p-doped conductive nanogauge generates a depletion layer and increases resistance. **(D)** The resistance across various electrodes connected to the device was monitored before and after ion adsorption; the resistance decay had a similar time constant.

monolayer worth of charge associated with adsorbed molecules is preserved for long periods of time.

In order to maintain device amplitude, a control loop could be used alongside the PLL used to maintain device frequency; however, a more direct solution providing for mass detection without amplitude changes is preferred. Referring to Equation 25 of Chapter 1:

$$F_{el} = -\frac{A\varepsilon_0}{2g^2}\left(V_{DC}^2 + \frac{V_{AC}^2}{2} + 2V_{DC}V_{AC}\cos\omega t + \frac{V_{AC}^2}{2}\cos 2\omega t\right), \qquad (25)$$

the electrical actuation force has a DC component when used in 1f mode, but has no DC component when used in 2f mode ($V_d = V_{AC}\cos\omega t$ with ω set to half the device resonance

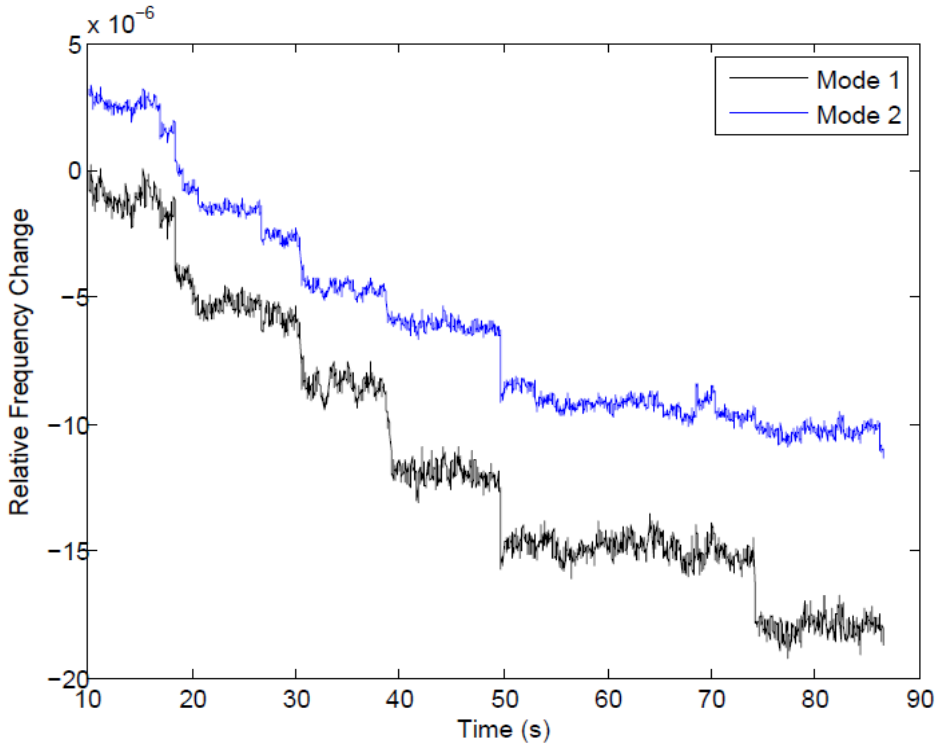


Figure 43. Detection of GroEL using 2f actuation. GroEL adsorption events are detected in real time for the first two modes with each event showing only negative frequency shifts; changes in resonant amplitude were not observed.

frequency and the lock-in amplifier set to detect at 2ω). Assuming the resonant amplitude increase is due to charge accumulation modulating the DC potential between the device and the beam, it could therefore be avoided with the 2f actuation technique. This was attempted with GroEL molecules sent to a single CAL3 device operated with the first two modes at resonance as shown in Figure 43. In contrast with the 1f actuation technique, the 2f actuation technique produced adsorption events with only negative frequency jumps as predicted, as well as no detectable change in resonant amplitude. The 2f actuation technique therefore avoids issues with charge adsorption affecting frequency in a nonlinear way, preventing the measurement of inertial mass. For this reason, the 2f technique was used to produce the remaining mass spectra in this thesis. Calculation of a mass spectrum for the data collected for Figure 43 is discussed in Chapter 3.

2.8 Summary

This chapter detailed the modification of a Q Exactive Plus, outfitted with extended mass range (EMR) capabilities, to acquire the capability to deliver intact, noncovalent macromolecules (up to 4.5 MDa) to nanomechanical devices operated at cryogenic temperatures. Such an undertaking required careful consideration of vacuum requirements, the construction and alignment of ion optics to maintain the motion of charged ions, electrical cabling thermalization, and thermal radiation shielding. The final ion lens needed to be calibrated to focus ions onto a small enough spot to achieve reasonable flux onto NEMS devices while achieving soft landing so the molecules would adsorb intact. In order for molecules to ultimately be detectable on the NEMS, the NEMS must operate at a lower potential than the Orbitrap's analysis chamber, which is typically at ground. We could not float the NEMS at large enough negative voltages without destroying the devices, so a large number of internal voltages to the Q Exactive needed to be raised in order to operate the NEMS at ground, and some pressure settings adjusted as well.

With the hybrid system built, a small series of mass spectra were recorded with myoglobin (17.5 kDa), GroEL (801 kDa), and 14/7 GroEL/ES (870 kDa) in their native state. These molecules could be isolated with the quadrupoles and characterized in the Q Exactive to deliver highly purified molecules to the NEMS devices with absolute certainty as to the molecule identity, a previously unavailable capability in our lab.

The cryogenic stage in the instrument could reach 10K base temperature, but the VLSI CMOS compatible devices we had on hand could only be cooled to 150 K due to device self-heating. At this temperature, single GroEL molecules could be detected adsorbing to the devices in real time. These adsorption events were initially associated with

large swings in device resonant amplitude, which interfered with the mass detection via measuring frequency shifts due to nonlinear coupling between device amplitude and frequency. This nonlinear coupling effect could be eliminated by using the 2f actuation technique, where no DC gate voltage is used and the devices are actuated only at AC at half their resonant frequency (with resonance achieved through the nonlinear electrostatic gate force). Up to a monolayer of molecular charge could be detected and is retained on the devices with a half-life of ~50 hours.

2.9 Bibliography

1. *Technical Information for Thermo Scientific Q Exactive Plus Orbitrap LC-MS/MS System*. 2018.
2. Rose, R.J., et al., *High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies.* Nature methods, 2012. **9**(11): p. 1084-1086.
3. Verbeck, G., W. Hoffmann, and B. Walton, *Soft-landing preparative mass spectrometry.* Analyst, 2012. **137**(19): p. 4393-4407.
4. Snijder, J., et al., *Defining the stoichiometry and cargo load of viral and bacterial nanoparticles by Orbitrap mass spectrometry.* Journal of the American Chemical Society, 2014.
5. Ekinci, K. and M. Roukes, *Nanoelectromechanical systems.* Review of scientific instruments, 2005. **76**(6): p. 061101.
6. Yang, Y.-T., *Phase noise of nanoelectromechancial systems*. 2006, California Institute of Technology.
7. Yang, Y.T., et al., *Surface adsorbate fluctuations and noise in nanoelectromechanical systems.* Nano Lett, 2011. **11**(4): p. 1753-9.
8. Hanay, M.S., *Towards Single-Molecule Nanomechanical Mass Spectrometry*. 2011, California Institute of Technology.
9. Manura, D.D.D., *SIMION (R) 8.0 User Manual*. 2008, Ringoes, NJ Scientific Instrument Services, Inc.
10. Mikhailov, V.A., et al., *Mass-Selective Soft-Landing of Protein Assemblies with Controlled Landing Energies.* Analytical Chemistry, 2014. **86**(16): p. 8321-8328.
11. Jacobs, D.C., *The role of internal energy and approach geometry in molecule/surface reactive scattering.* Journal of Physics: Condensed Matter, 1995. **7**(6): p. 1023.
12. Laskin, J., *Ion–surface collisions in mass spectrometry: Where analytical chemistry meets surface science.* International Journal of Mass Spectrometry, 2015. **377**: p. 188-200.
13. Johnson, G.E., D. Gunaratne, and J. Laskin, *Soft-and reactive landing of ions onto surfaces: concepts and applications.* Mass Spectrom. Rev, 2016. **35**: p. 439-479.

14. Cyriac, J., et al., *Low-energy ionic collisions at molecular solids.* Chemical reviews, 2012. **112**(10): p. 5356-5411.
15. Mikhailov, V.A., et al., *Mass-selective soft-landing of protein assemblies with controlled landing energies.* Analytical chemistry, 2014. **86**(16): p. 8321-8328.
16. Siuzdak, G., et al., *Mass spectrometry and viral analysis.* Chemistry & biology, 1996. **3**(1): p. 45-48.
17. Ouyang, Z., et al., *Preparing protein microarrays by soft-landing of mass-selected ions.* Science, 2003. **301**(5638): p. 1351-1354.
18. Volný, M., et al., *Preparative soft and reactive landing of multiply charged protein ions on a plasma-treated metal surface.* Analytical chemistry, 2005. **77**(15): p. 4890-4896.
19. Hadjar, O., J.H. Futrell, and J. Laskin, *First observation of charge reduction and desorption kinetics of multiply protonated peptides soft landed onto self-assembled monolayer surfaces.* The Journal of Physical Chemistry C, 2007. **111**(49): p. 18220-18225.
20. Nie, Z., et al., *In situ SIMS analysis and reactions of surfaces prepared by soft landing of mass-selected cations and anions using an ion trap mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2009. **20**(6): p. 949-956.
21. Wilm, M. and M. Mann, *Analytical properties of the nanoelectrospray ion source.* Analytical chemistry, 1996. **68**(1): p. 1-8.
22. Leney, A.C. and A.J.R. Heck, *Native Mass Spectrometry: What is in the Name?* Journal of The American Society for Mass Spectrometry, 2017. **28**(1): p. 5-13.
23. Consortium, T.U., *UniProt: a worldwide hub of protein knowledge.* Nucleic Acids Research, 2018. **47**(D1): p. D506-D515.
24. Quaite-Randall, E. and A. Joachimiak, *Purification of GroEL from an Overproducing E. coliStrain*, in *Chaperonin Protocols*. 2000, Springer. p. 29-39.
25. PDB ID: 4V43. Wang, J., *Structural and mechanistic basis for allostery in the bacterial chaperonin GroEL. DOI: 10.2210/pdb4v43/pdb*. 2002.
26. Xu, Z., A.L. Horwich, and P.B. Sigler, *The crystal structure of the asymmetric GroEL–GroES–(ADP)7 chaperonin complex.* Nature, 1997. **388**(6644): p. 741-750.
27. Lössl, P., J. Snijder, and A.J. Heck, *Boundaries of mass resolution in native mass spectrometry.* Journal of the American Society for Mass Spectrometry, 2014. **25**(6): p. 906-917.
28. Duijn, E.v., *New insights into chaperonin-assisted protein folding revealed by native mass spectrometry*. 2006, Utrecht University.
29. Konermann, L. and D. Douglas, *Unfolding of proteins monitored by electrospray ionization mass spectrometry: a comparison of positive and negative ion modes.* Journal of the American Society for Mass Spectrometry, 1998. **9**(12): p. 1248-1254.
30. Snijder, J., et al., *Studying 18 Mega Dalton Virus Assemblies with Native Mass Spectrometry().* Angewandte Chemie (International ed. in English), 2013. **52**(14): p. 4020-4023.
31. Rostom, A.A., et al., *Detection and selective dissociation of intact ribosomes in a mass spectrometer.* Proceedings of the National Academy of Sciences of the United States of America, 2000. **97**(10): p. 5185-5190.

32. Tito, M.A., et al., *Electrospray time-of-flight mass spectrometry of the intact MS2 virus capsid.* Journal of the American Chemical Society, 2000. **122**(14): p. 3550-3551.
33. van de Waterbeemd, M., et al., *High-fidelity mass analysis unveils heterogeneity in intact ribosomal particles.* Nat Meth, 2017. **14**(3): p. 283-286.
34. Fort, K.L., et al., *Expanding the structural analysis capabilities on an Orbitrap-based mass spectrometer for large macromolecular complexes.* Analyst, 2018. **143**(1): p. 100-105.
35. Fong, K.Y., W.H. Pernice, and H.X. Tang, *Frequency and phase noise of ultrahigh Q silicon nitride nanomechanical resonators.* Physical Review B, 2012. **85**(16): p. 161410.
36. Pobell, F., *Matter and methods at low temperatures*. Vol. 2. 2007: Springer.
37. Hanay, M.S., et al., *Single-protein nanomechanical mass spectrometry in real time.* Nat Nano, 2012. **7**(9): p. 602-608.
38. Sage, E., *Nouveau concept de spectromètre de masse à base de réseaux de nanostructures résonantes*. 2013, Grenoble.
39. Kozinsky, I., et al., *Tuning nonlinearity, dynamic range, and frequency of nanomechanical resonators.* Applied Physics Letters, 2006. **88**(25): p. 253101.

Chapter 3

MASS SPECTROMETRY AND INERTIAL IMAGING OF PROTEINS

3.1 Introduction

In this chapter, we used the hybrid Orbitrap-NEMS system to perform single-molecule analysis of GroEL (801 kDa). GroEL could be pre-selected using the quadrupoles of the Q Exactive, thus ensuring the identity of each molecule adsorbing to the NEMS. This is a new capability in our lab, which is documented extensively in Chapter 2. Data for the two-mode mass spectra were initially collected with Dr. Eric Sage, and then evaluated with calculations using techniques similar to those previously published by our group [2]. I then collected the first three-mode mass spectra of single proteins in real time using GroEL molecules. The intent was to deduce the first three mass moments of the 1D contact area of the molecule according to the method of inertial imaging [3, 4]. The experimental method I devised for this study is only a minor extension to the two-mode technique, but several important practical details needed to be investigated. Among these were whether the modes should be actuated simultaneously or sequentially. Additionally, the statistical framework needed to be upgraded to permit detecting events and then transforming the measured frequency jumps upon analyte adsorption into uncertainties for the first three mass moments. Such an experiment, if successful, would demonstrate the first form of molecular imaging based on the change in kinetic energy induced in nanomechanical structures; the technique is, in principle, capable of atomic-level resolution with today's smallest NEMS devices [3].

3.2 3D mapping of ion flux focused by ion lenses

The initial adsorption attempts discussed in Sections 2.6 and 2.7 each required significant time to locate the ion beam due to its tight focusing through our implementation of ion optics. Prior to collecting more MS data, we decided to systematically characterize in three dimensions the GroEL ion beam directed toward the NEMS. A 20-device CAL3 NEMS array was used for this purpose to increase the throughput of data collection. Details concerning the operation and fabrication of this array are provided in [1, 5] and summarized here for convenient reference.

The NEMS arrays, fabricated by our Alliance for Nanosystems VLSI with LETI, are either cantilevers or doubly-clamped beams. Actuation and detection are performed

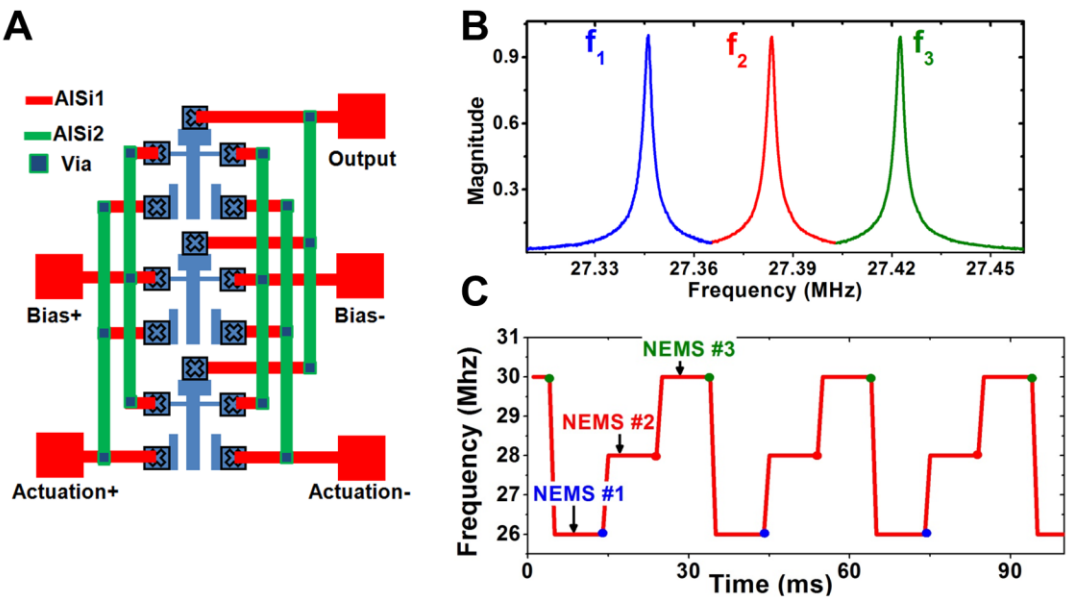


Figure 44. Frequency addressed NEMS array. **(A)** Devices in array are interconnected in parallel using two levels of metal. The input and output electrodes are the same as for single devices. **(B)** Different device lengths provide a different resonance frequency for each device. An open loop sweep reveals a succession of peaks. **(C)** Frequency tracking involves the sequential use of a PLL for each device in the array; the frequency of each device can be reconstructed by splicing together the data collected just prior to switching to the next device. Figures from [1].

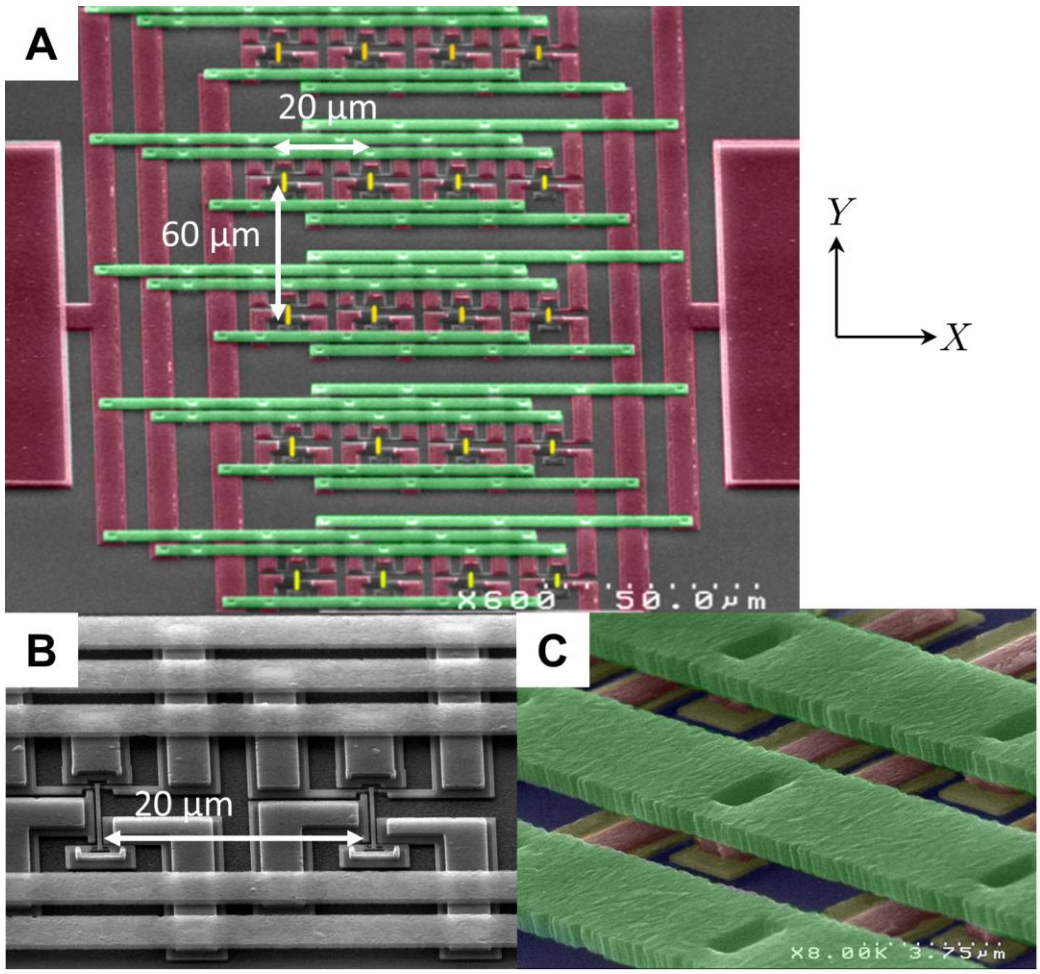


Figure 45. SEM images of NEMS array. (A) A 20-device array of doubly-clamped beams is shown with the first layer of metallization, AlSi1, in red, and the second layer of metallization, AlSi2, in green. The devices have a pitch of 20 µm in X and 60 µm in Y for a total field size of 60 × 240 µm. **(B)** Close up of two adjacent devices showing parallel interconnections. **(C)** Close up of metal interconnections and vias. From [1].

identically to the single devices discussed in Section 1.8. Each device possesses five electrical connections. They are interconnected in parallel using two additional metal layers. This ensures that the input signal is applied to all NEMS simultaneously, and the overall output is the sum of all devices' individual outputs. As their resonance frequencies are different, they are separately addressable. An example layout of three cantilevers is illustrated in Figure 44. Devices are 160 nm thick by 300 nm wide; for the doubly clamped beams, the lengths vary from 7–9.2 µm.

Fabrication details are discussed in [1] and colorized SEM images are shown in Figure 45. Simple arrays of 20, 49, and 100 devices have been fabricated; however, operation of arrays comprising more than 20 devices proved somewhat problematic beyond the second mode with some of the higher modes of the largest devices starting to overlap some of the lower modes of the smallest devices. The next generation design will circumvent this issue; these arrays were designed prior to our group's invention of multimodal inertial imaging.

Successful closed loop operation of a 20-device array is shown in Figure 46a. Myoglobin was directed toward the array, and the resonant frequency of each device was successfully tracked as shown in Figure 46b. The mass resolution of each device is not high enough to sufficiently resolve each adsorbed molecule, but the rate of change of the resonant

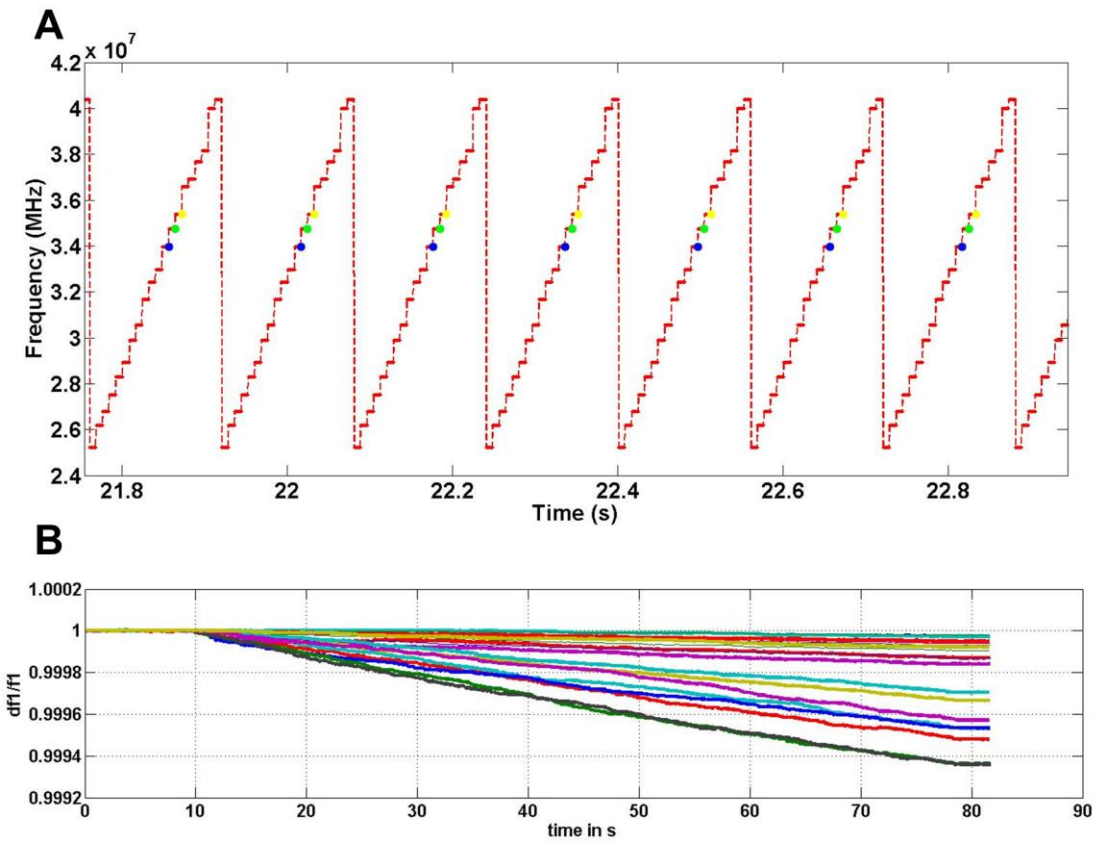
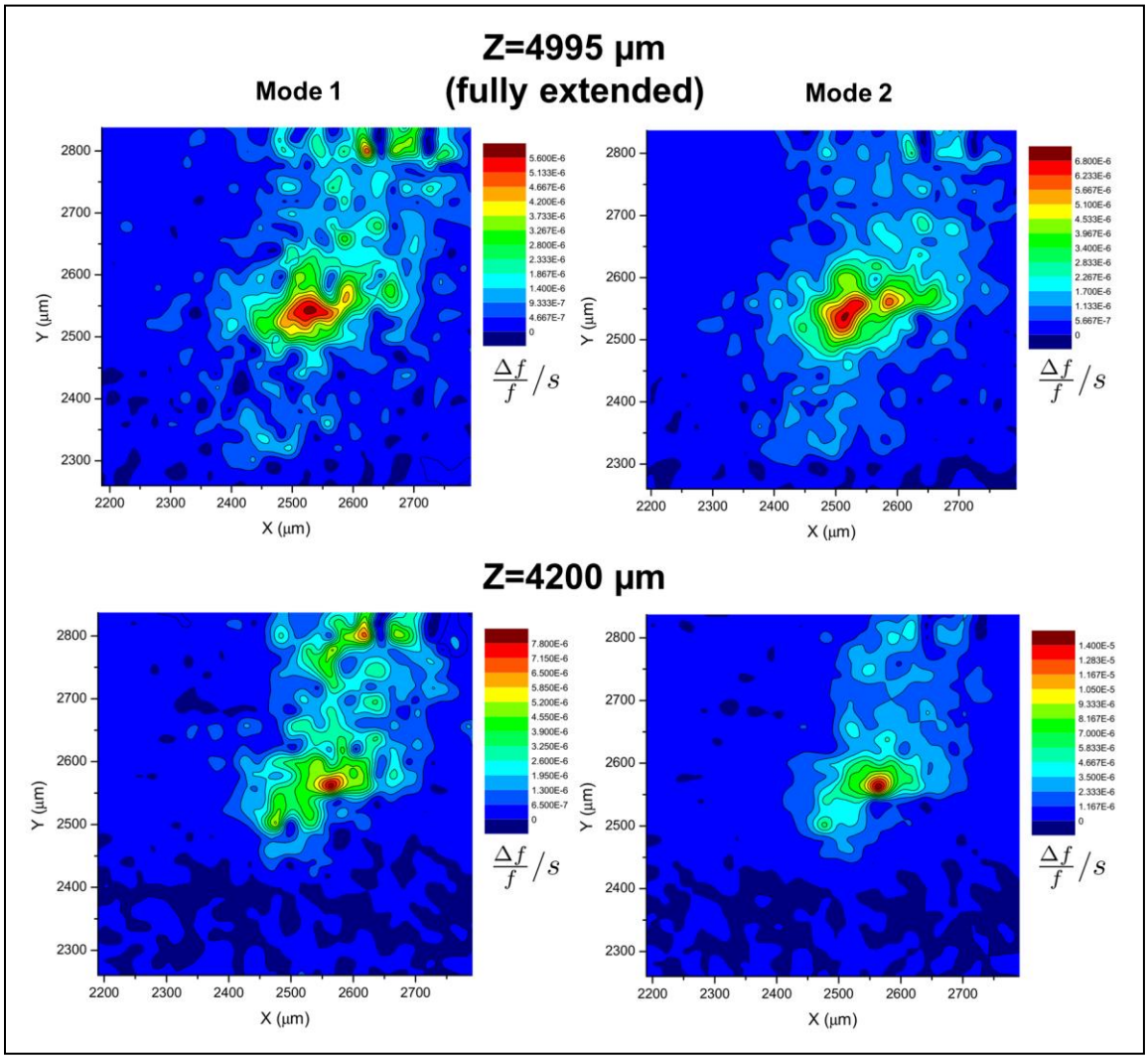


Figure 46. Operation of 20-device NEMS array. (A) Sequential closed loop operation of all devices. **(B)** Individual frequency trace of each device under a myoglobin ion beam.

frequency is proportional to the adsorption rate incident on each device, assuming uniform deposition. The different slopes for $\Delta f/f$ observed for each device is therefore due to variable ion flux over the array's entire field.

To map the ion flux of GroEL in three dimensions, we used the 3D XYZ positioner's (Attocube) capability for 0.1 nm positioning precision over 5 mm travel range. The array has a pitch of 20 µm in X and 60 µm in Y, so to achieve uniformly spaced data collection, the array was first staggered twice in Y by 20 µm, forming a field of 60 µm × 300 µm with a data point every 20 µm in X and Y. Then, this was repeated to form a field of 600 × 600 µm perpendicular to the ion beam axis. This was done with 3 minutes spent per position, so each



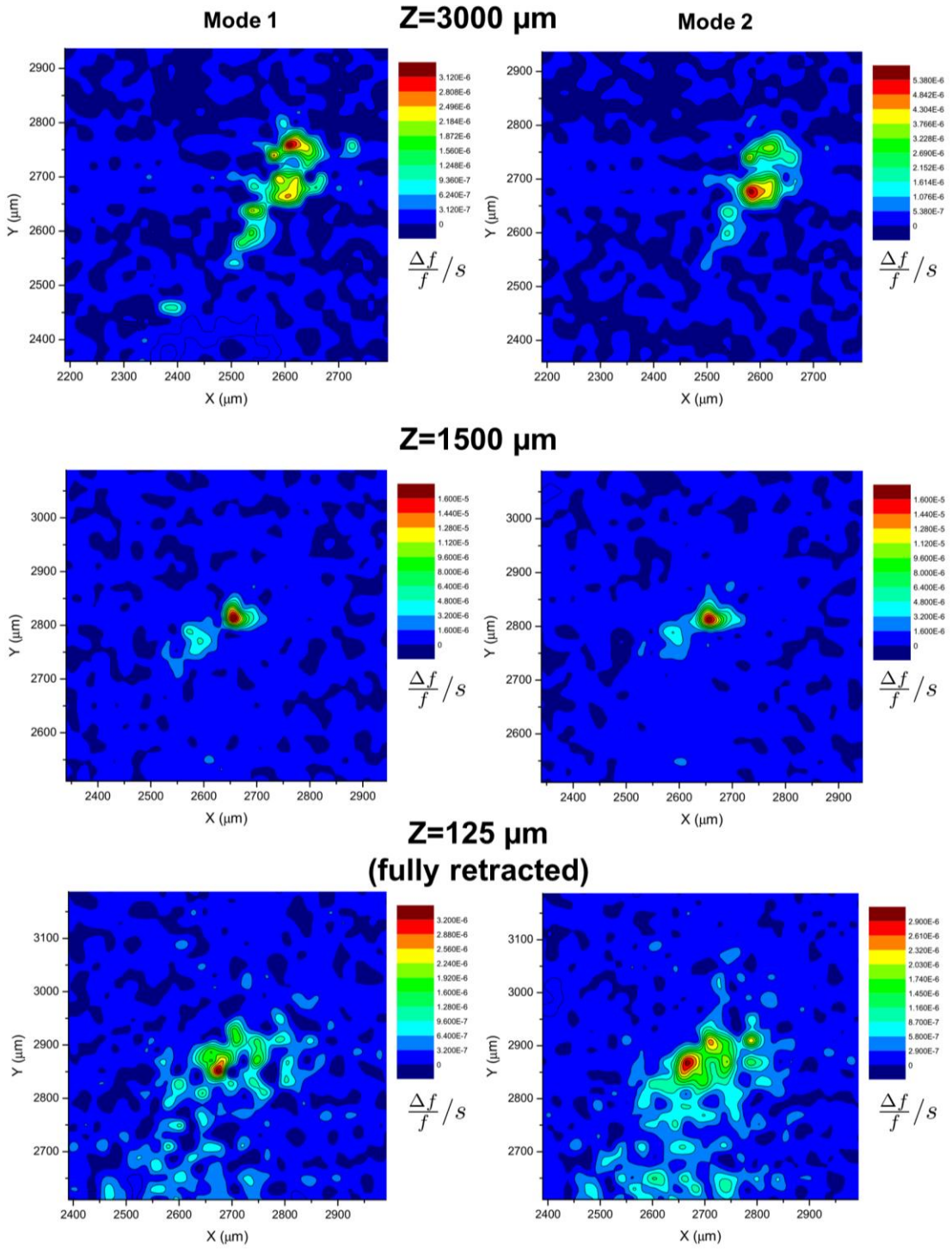


Figure 47. 3D map of GroEL flux rate on NEMS. A NEMS array was mounted on a 3D positioner to measure the ion flux rate, proportional to $(\Delta f/f)/s$, every 20 nm at several positions along the ion beam axis. The maximum flux rate at Z=1500 µm corresponds to 25 molecules per second.

field took 1 hour to capture. Finally, it was repeated for several Z values, from fully extended (closer to the lens, Z=5 mm) to fully retracted (further from the lens, Z=0). Positioning was performed in an automated manner using a custom Python script.

The initial measurement used a lens diaphragm voltage set to 0V. Results are shown using contour plots in Figure 47 with raw data given in terms of $(\Delta f/f)/s$. This can be converted to molecules/s using the mass responsivity for mode 1 at 12 Hz/ag or the responsivity for mode 2 at 32 Hz/ag, the approximate respective resonance frequencies of 25 MHz and 67 Hz, and a GroEL mass of 1.3 ag. The ions could be focused to a minimum beam size of 40 microns, with a flux rate of 25 molecules/s. Measurements were repeated with the positioner fully extended, and varying the diaphragm voltage. The ion beam size at different positions and diaphragm voltages are summarized in Table 7. Such data, once collected, is useful for finding the ions again after installing a new device, as well as setting the position of the device to achieve the desired flux rate. The experiment was repeated using myoglobin and the flux rate plotted with the positioner fully extended is shown in Figure 48. Myoglobin

Z (μm) keeping lens diaphragm at 0V	Ion beam size FWHM (μm)	Lens diaphragm (V) keeping Z=5 mm	Ion beam size FWHM (μm)
125	127.4	40	189.9
1500	40.5	20	146.4
3000	97.7	0	225.6
4200	212.1	–20	143.4
4625	245.2	–60	135.5
4995	225.6		

Table 7. Lateral extent of ion beam (in μm) while varying positioner Z or diaphragm voltage. Varying the positioner along the beam axis shows that the ion beam is maximally focused to a spot size of 40 μm at Z=1500.

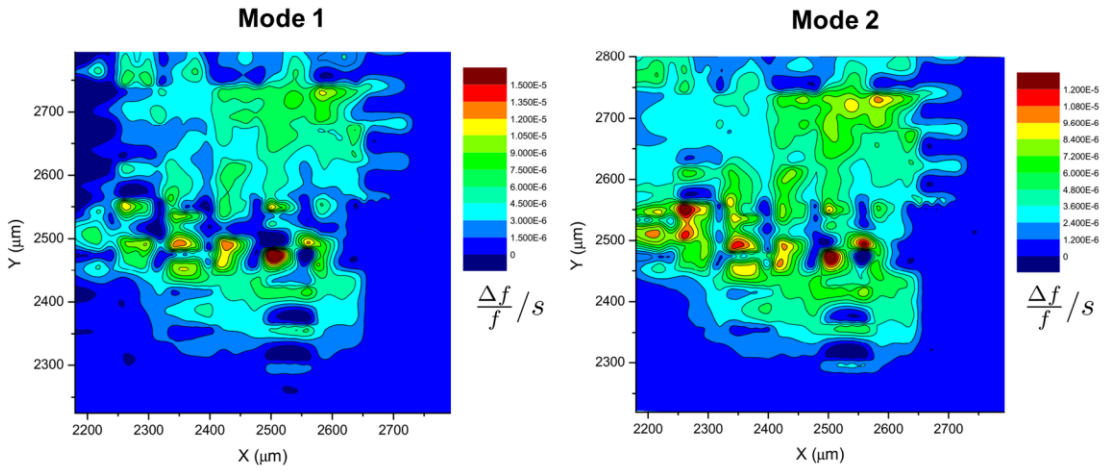


Figure 48. Flux rate of myoglobin on NEMS. Myoglobin was sent to a NEMS array similarly positioned to monitor flux rate every 20 nm. The maximum flux rate corresponds to 880 molecules per second.

could be detected over a much larger 5 mm spot; therefore, it could be used for preliminary device positioning.

3.3 High event rate NEMS-MS of GroEL with two modes

The ion flux data was used to position the smallest device in the array (7 μm long) with the best mass resolution at a location achieving approximately 1 molecule per second adsorption rate. This flux rate was chosen to demonstrate the ability to collect a large amount of data in a short period of time, while still having a low probability of multiple molecules landing within the measurement window of 100 ms. This measurement window is set according to the PLL time that minimizes the Allan deviation, and hence mass resolution, as shown in Figure 36b. Data was collected over 15 minutes, as shown in Figure 49. Jumps are automatically detected and evaluated to build a mass spectrum using a formalism based on transforming the joint probability distribution from the frequency domain of the first two

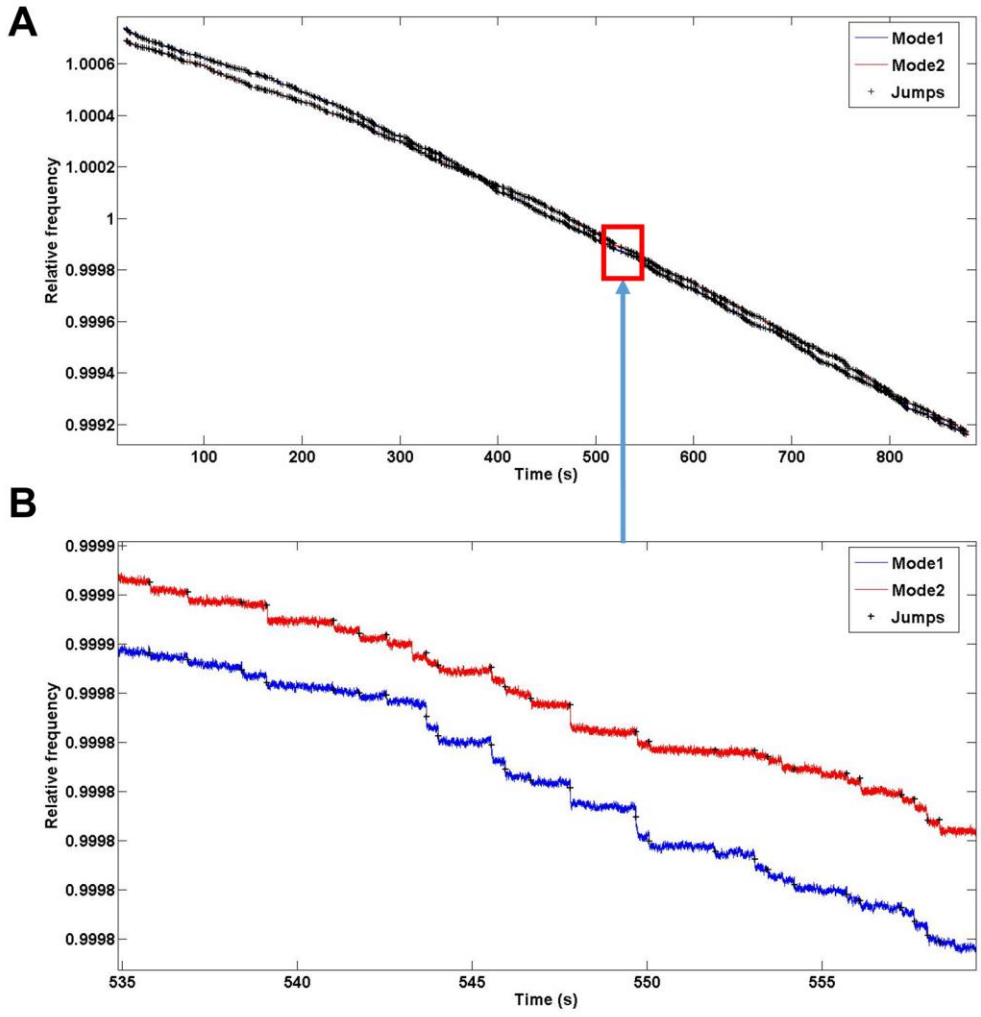


Figure 49. Two mode detection of GroEL molecules in real time. The highest mass-resolution device in the 20-device array was positioned to achieve approximately 1 molecule per second adsorption rate. The frequency of the first two resonance modes was tracked in real time using a PLL and frequency jumps exceeding a 5σ threshold were detected.

modes to the mass and position domains [1, 2, 6]. Using the first two modes, 656 frequency jumps were acquired.

In brief, frequency fluctuations over a given measurement window, for a device array with each NEMS element operated with the first two modes simultaneously excited, have been observed to form a 2D normal distribution. Adsorption events can be identified as fluctuations that exceed a specified threshold, such as 5σ from background noise [1, 6]. For

a given detected event with relative frequency shifts $\mu_1 = df_1/f_1$ and $\mu_2 = df_2/f_2$ for the first two modes respectively, the probability density function for a given mass and position that satisfies these two frequency shifts is given by [6]:

$$\mathrm{PDF}(\delta m, a) = |J| \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left(-\frac{\Gamma}{2(1-\rho^2)}\right);$$

$$J = \frac{\delta m}{\alpha_1\alpha_2}\left(\phi_1^2(a)\left.\frac{d\phi_2^2(x)}{dx}\right|_{x=a} - \left.\frac{d\phi_1^2(x)}{dx}\right|_{x=a}\phi_2^2(a)\right);$$

$$\Gamma = \frac{(h_1-\mu_1)^2}{\sigma_1^2} - \frac{2\rho(h_1-\mu_1)(h_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(h_2-\mu_2)^2}{\sigma_2^2};$$

$$h_1 = -\frac{\delta m}{2\alpha_1}\phi_1^2(a); \quad h_2 = -\frac{\delta m}{2\alpha_2}\phi_2^2(a), \tag{38}$$

where J is the Jacobian of the coordinate system transformation. The mass and position spectra are obtained by projecting this two-dimensional PDF into either mass or position:

$$\mathrm{PDF}(\delta m) = \int_{a=0}^{a=0.5} \mathrm{PDF}(\delta m, a)\, da; \tag{39}$$

$$\mathrm{PDF}(\delta a) = \int_{\delta m=0}^{\delta m=\infty} \mathrm{PDF}(\delta m, a)\, d(\delta m). \tag{40}$$

The frequency jumps were converted into a mass and position spectra according to this method, as shown in Figure 50. The mass spectrum features a peak mass of 747.9 kDa (compared with 801 kDa expected) and FWHM of 445 kDa. The device had Allan deviations of σ_1 = 2.23E–7 and σ_2 = 1.84E–7 for the first two modes and noise correlation coefficient $\rho = 0.28$, with an expected mass resolution of 118 kDa. This mass resolution is notably higher than the published value of 50 kDa for standalone CAL3 devices [7]; however, the measurements in this section used one of the 20 devices in an array, and the performance of

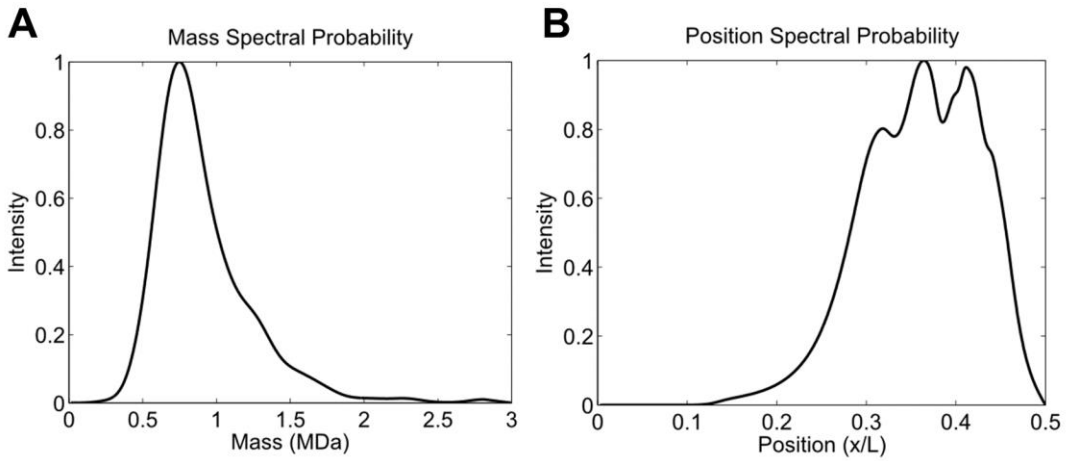


Figure 50. Mass and Position spectra of GroEL measured with two modes. Frequency jumps detected from GroEL adsorption events are used to form mass and position spectra using Equations 38-40. The average mass spectrum peak for 656 single-molecule adsorption events was 747.9 kDa with a FWHM of 445 kDa. The device had a mass resolution of 118 kDa. The position spectrum was not highly peaked, but not uniform either, due to low mass responsivity near the device clamp points (see text).

each such device in these first NEMS arrays is known to be worse than the performance of standalone devices [1]. The position spectrum was neither highly peaked nor uniform. Although uniform deposition over the beam is expected, due to the low device responsivity near the beam's clamp points, adsorption events do not uniformly exceed the 5σ threshold needed for event detection.

The effect of varying the jump detection threshold is shown in Figure 51. Because the GroEL molecule is close to the mass resolution of the device, the choice of jump detection threshold affects the mass spectrum peak. Having too low of a threshold results in a false peak at <500 kDa which can be ascribed to false positive events [1], while having too high of a threshold creates false negative events and artificially shifts the mass spectrum to the right. The choice of 5σ eliminates the false peak at <500 kDa without shifting the centroid of the mass spectrum to the right.

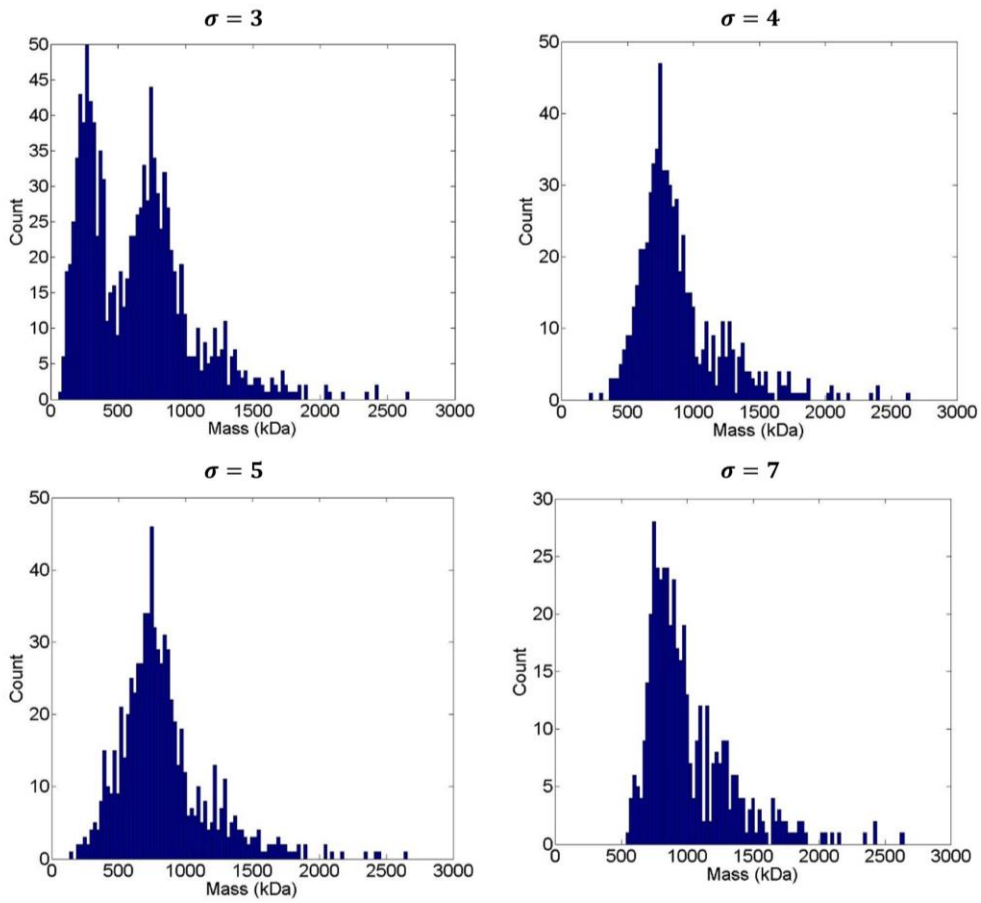


Figure 51. Mass spectrum for different choices of jump detection threshold. Histograms are shown for calculated mass of GroEL for different multiples of σ describing the 2D normal distribution for background device noise fluctuations.

The FWHM of the mass spectrum is significantly larger than expected. For a device with mass resolution of 118 kDa, a normal probability distribution would produce a FWHM of $2\sqrt{2\ln 2}\,\sigma \approx 2.355\sigma = 278$ kDa. Although the mass resolution is produced from a decidedly non-normal distribution via the nonlinear joint probability distribution transformation, the large difference is still notable. A simulation was performed to closely mimic the experiment to ensure that the calculations are internally correct and to investigate any sources of discrepancy. Results are shown compared to the experimental data in Figure 52. The FWHM of the simulation was 299 kDa, close to expected.

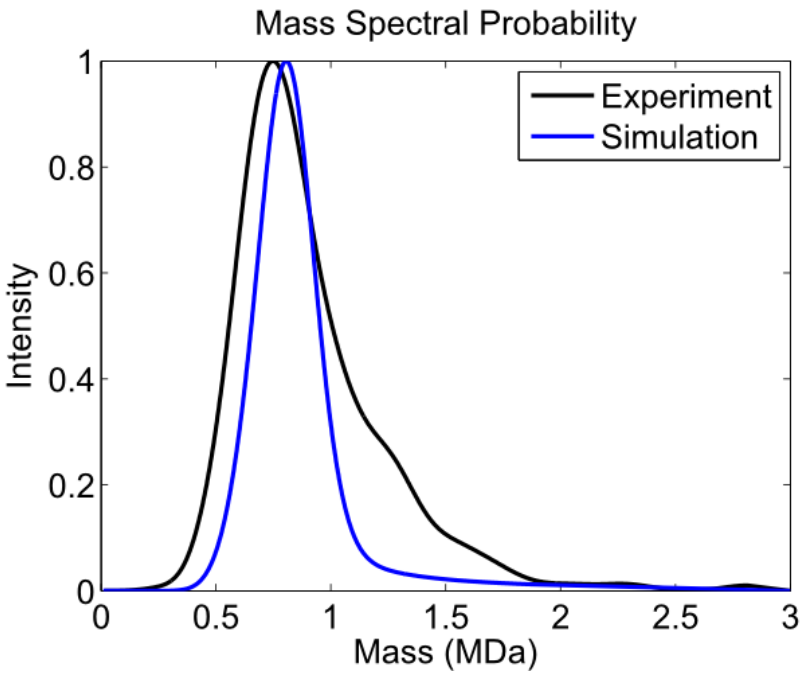


Figure 52. Simulation vs experiment for GroEL mass spectrum. A simulation was performed consisting of an 801 kDa point mass deposited uniformly 1000 times over a doubly-clamped beam. Frequency shifts were calculated, then noise was added to match the experimental device. A mass spectrum was then calculated using these frequency shifts. The simulated spectrum has a peak at 802 kDa and FWHM of 299 kDa.

Due to the nonlinear transformation, the peak value of the mass probability density for each individual particle does not equal the mean value. This difference will also vary depending on where the particle lands. To illustrate this, an additional simulation showing the peak value of the mass probability density versus the peak value of the position probability density is shown in Figure 53. The simulation was conducted both with noise added in for the device used in this Section and with no noise, showing that the effect is always present. It was found that the ensemble mass spectrum calculated using events over the entire device will produce a mass spectrum with an overall peak different from the molecule's mass, unless data inclusion is restricted to calculated positions $0.39 < x < 0.41$. Additionally, the ensemble mass spectrum will have a width larger than that expected from the device's mass resolution unless data is restricted to this extremely narrow region.

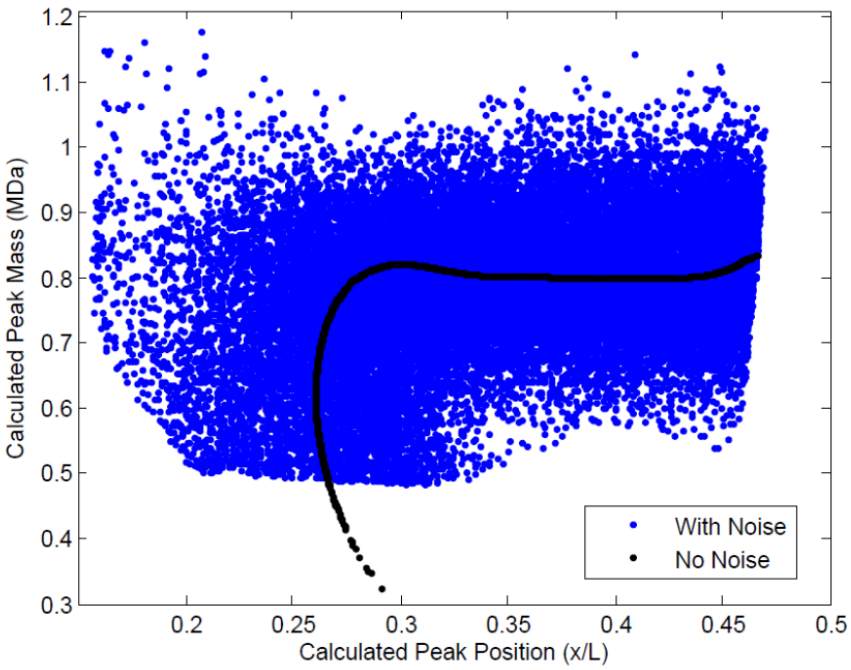


Figure 53. Simulation of calculated peak mass versus calculated peak position. Calculated mass vs position is shown for the same simulation performed, both with and without device noise for the device used to collect experimental data. The peak mass of each event differs from the molecule's 801 kDa due to the nonlinear nature of the probability distribution transformation. Combining many events collected over the device will therefore produce a mass spectrum with an overall peak and width different from that collected from a single event. This would occur even with no device noise.

The mass spectrum for the experimental GroEL data was recalculated using various choices of position restriction as shown in Figure 54. Restricting the data to $0.39 < x < 0.41$ improved the spectrum, not only eliminating the small unresolved peaks on the right shoulder of the spectrum, but also allowing for a second peak to be resolved near twice the mass of GroEL. The dominant peak of the spectrum for events with $0.39 < x < 0.41$ was 846 kDa, and the FWHM was 426 kDa. The second peak was at 1.756 kDa. This demonstrates that, since we can deduce where analytes land, we can select those events associated with the highest accuracy *a posteriori* to improve the final result. This position-restricted data was compared directly to the simulation in Figure 55. The peak mass was close to expected, but the width of the spectrum was still significantly larger than expected.

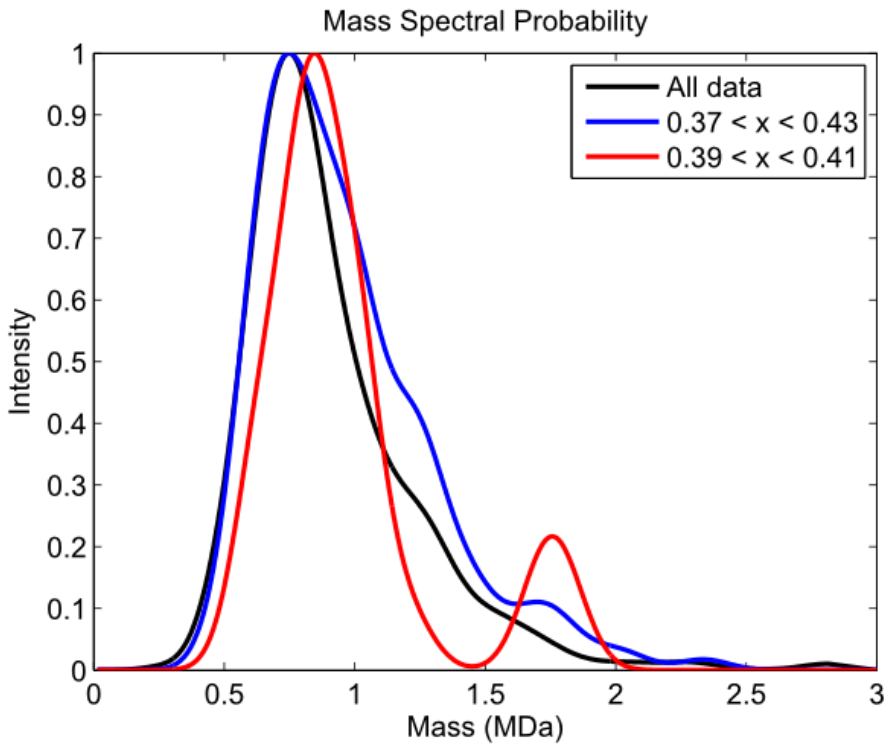


Figure 54. GroEL mass spectrum from restricted regions of doubly-clamped beam. The position spectrum for each individual event was calculated using Equation 40. The peak value of this spectrum was used to determine whether or not to include the event in constructing the overall mass spectrum. Significantly restricting the calculated position of the data improves the mass spectrum and allows events near double the GroEL mass (at 1.6 MDa) to be resolved.

The difference in calculated mass compared with expected has been previously observed in calibrated experiments with metallic nanoclusters [8]; this has been ascribed to the fact that the device's piezoresistive gauges change the effective device length (and hence, mass). This could be addressed with a calibration procedure. The increased width of the spectrum compared with the simulation may be due to two or more GroEL molecules landing within the measurement time of 100 ms; rather than both landing within the restricted position $0.39 < x < 0.41$, giving rise to the secondary peak in the spectrum, the molecules could land at two completely different positions, causing the spectrum to broaden.

A small note on the MATLAB calculations used to compute mass spectra: MATLAB natively relies on vectorized data for efficient and optimized computation. For the 2D

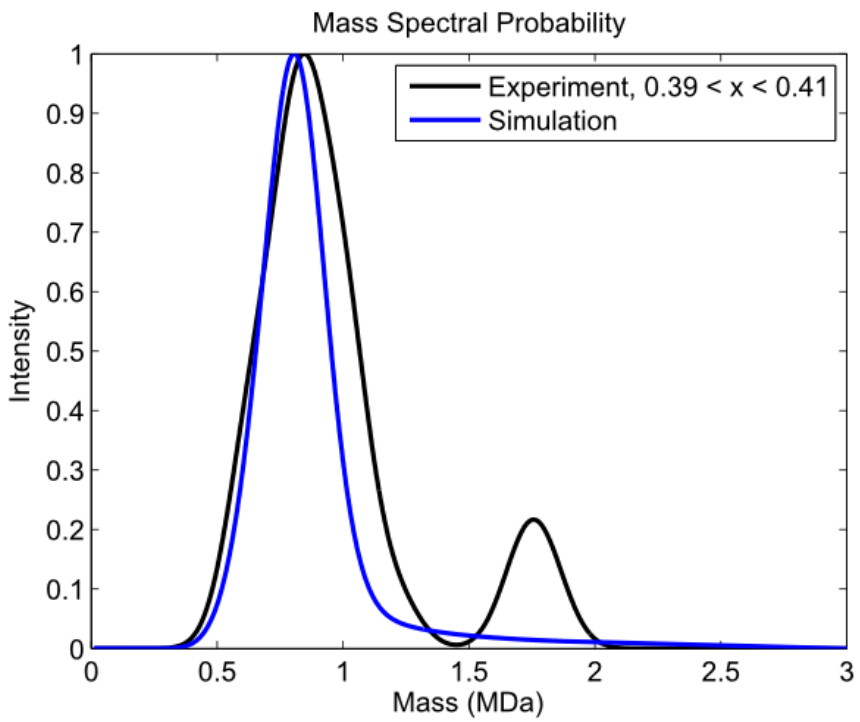


Figure 55. Simulation vs experiment for restricted region of doubly-clamped beam. Experimental data was much closer to the simulation for the highly restricted region $0.39 < x < 0.41$. The dominant mass of the experimental mass spectrum different from simulation by 4.9% and the width of the mass spectrum was larger than experiment by 42%. The experimental plot is constructed using 40 events, 6% of the total detected.

calculations involving mass and position data points, it was important to consistently use columns to store mass and rows for position. Calculations can then be performed in aggregate using matrices rather than looping over each dimension. Rewriting the code like this improved computation speed by a factor of 10, allowing for the large number of simulations in this section to be easily computed.

3.4 Actuation and detection of more than two modes

With the development of inertial imaging theory, additional information about the shape of the molecule can be ascertained by actuating and detecting more than two modes for a doubly clamped beam. The remainder of this Chapter discusses efforts to perform this measurement experimentally on GroEL molecules. In the original configuration of the

experiment, each mode required two signal generators and one lock-in amplifier (LIA) (Figure 12); measuring many modes with this approach would start to become unwieldly. A custom instrument called a "D-Box," previously developed by our group, was used which

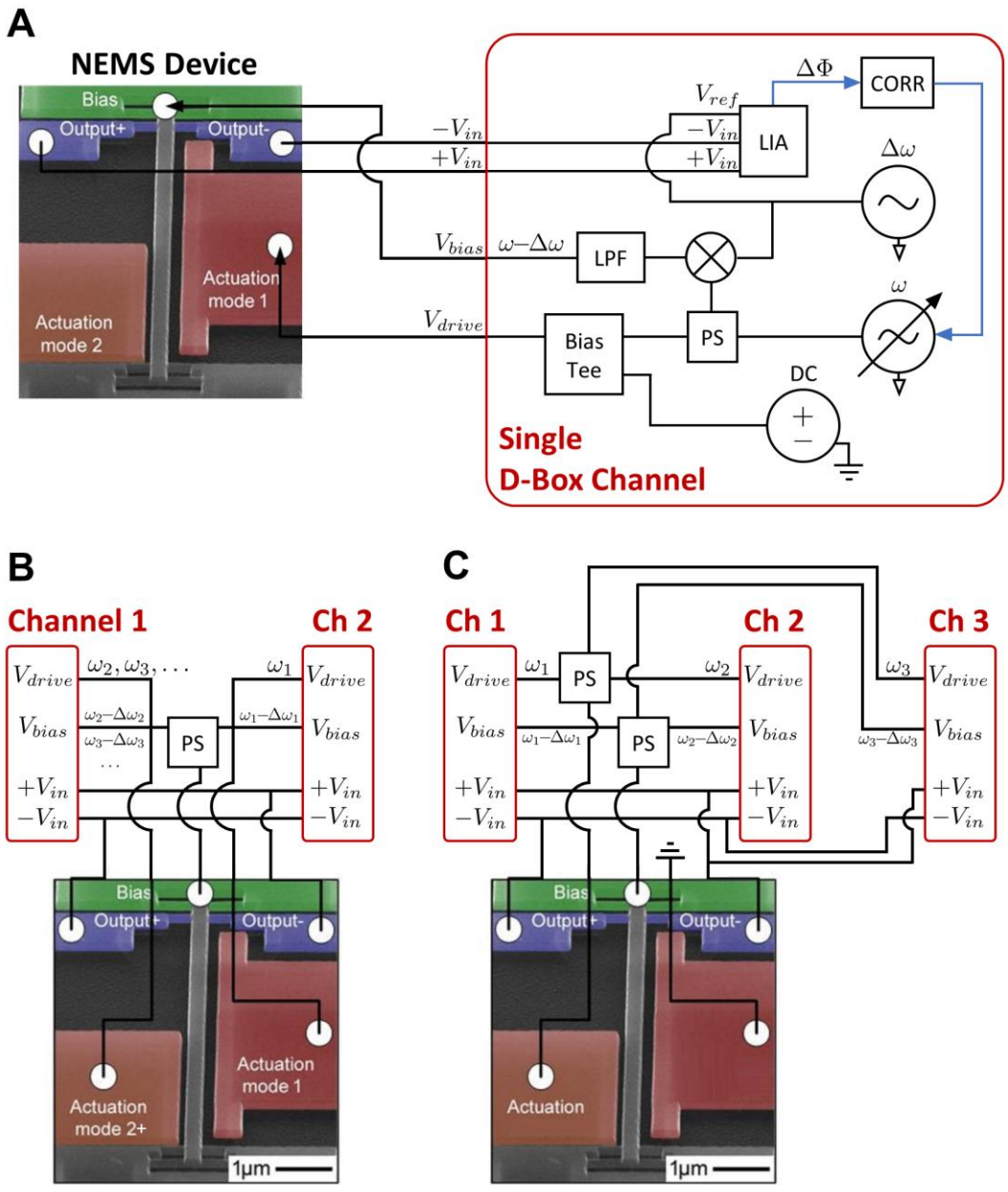


Figure 56. Measurement setups for more than two modes. (A) The D-Box contains 6 channels, each capable of actuating a single mode with the down-mixing scheme. Actuating more than two modes can be achieved by **(B)** sequential actuation or **(C)** simultaneous actuation. Details of operation are in the text.

includes six channels, with each channel capable of measuring a single mode with the down-mixing scheme (Figure 56a). More than two channels can be operated on the device by using power combiners for the drive and bias signals in the MHz range, and simple tees to split the <100 kHz down-mixed signal at the device output to the LIA inputs for each channel. Different values of $\Delta\omega$ for each mode are used to avoid cross-talk between modes for each channel.

The simplest setup would involve switching between modes (assuming device drift is minimal during the time it takes to switch between all modes). In principle, this would require only one D-Box channel and set of cables. In practice, two channels were used, with the first mode using one channel and connected to the larger actuation gate, and the second and higher modes using the other channel and connected to the smaller actuation gate (Figure 56b). This was done because actuation for the first mode was found to be significantly more efficient with the larger gate, and actuation for the higher modes was found to be significantly more efficient with the smaller gate.

Alternatively, all modes could be actuated simultaneously. A schematic showing how this was done for three modes is shown in Figure 56c. More modes could be achieved using power combiners with more inputs as well as additional cables, tees, and D-Box channels.

Modes activated	Allan deviation (Mode 1)
1	3E–7
1, 2	3.5E–7
1, 2, 3	1.2E–6
1, 2, 3 4	1.5E–6

Table 8. Allan deviation of first mode while combining actuation with higher modes. Room temperature measurement of a single CAL3 device with operation of 1, 2, 3, and 4 simultaneous modes. Each mode was actuated according to their optimal performance. The Allan deviation for the first mode is reported with a measurement time of 100 ms.

First attempts for operation of more than two modes focused on the simultaneous actuation approach due to the familiarity with simultaneously actuating two modes. An initial test was performed by operating a single CAL3 device at room temperature with increasing numbers of modes. As shown in Table 8, the Allan deviation of the first mode increased with each additional new mode, with significant performance degradation occurring in the jump from two to three modes. Similar experiments were performed on piezoelectric AlN membranes, and the decrease in performance for multimode operation was not observed [9]. So, with the piezoresistive devices in particular, the issue seems to be the large amount of heat dissipated during readout of their higher modes. Additionally, operation of additional higher modes to their onset of nonlinearity was found to significantly alter the resonance frequencies of lower modes; this inter-modal dependence of resonance frequency on amplitude indicates the presence of a nonlinearity which could alter device mode shape. Since inertial imaging requires extracting small differences in fractional frequency-shift data, this could alter the measurement.

Initial attempts at mode switching were also discouraging. A single CAL3 device was operated at room temperature, switching between modes 2 and 4. Each time switching to a new mode, a significant change in resonance frequency was observed, with the previous resonance frequency recovering after a long time interval of about 5 seconds regardless of PLL time (Figurea). Such a long time constant was not likely to be related to anyting electrical or mechanical. One possibility was a thermal issue. It was found that by carefully tuning the bias voltage V_{bias} of mode 4, the resonance frequency of mode 2 could remain the same frequency when switching back to it (Figureb). Selecting $V_{bias} = 1.87\text{V}$ for mode 4 evidently allowed for the resonance frequency of mode 2 to remain unchanged; higher V_{bias}

would cause the resonance frequency to increase, and lower V_{bias} would cause the resonance frequency to decrease. These measurements are consistent with the temperature of the device changing with V_{bias}, and the need to tune V_{bias} for each mode to achieve constant RMS temperature. (Note that the same value of V_{bias} cannot simply be used for all modes due to different signal attenuation at different frequencies). These issues are currently being

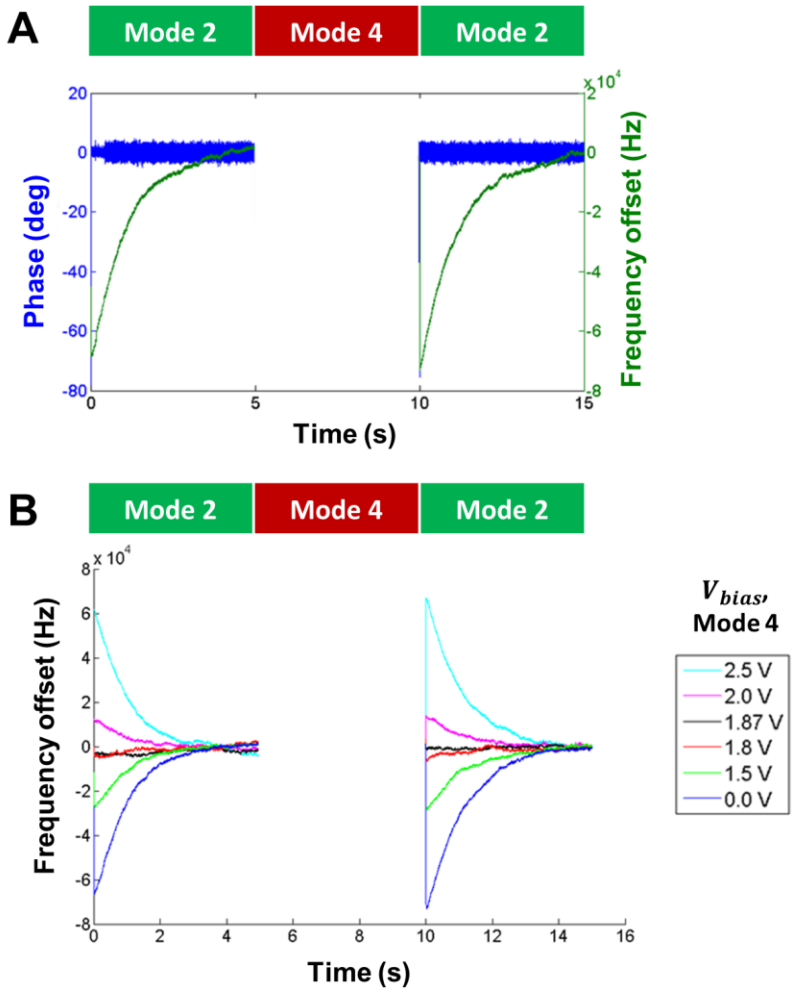


Figure 57. Mode switching with matched RMS heating. Data is shown for the tracked phase and resonance frequency for mode 2, while switching between modes 2 and 4, spending 5 seconds on each mode and operating with a 100 ms PLL time. **(A)** With each mode naively operated according to their individually optimized settings, phase recovers to zero within the correct PLL time, but the tracked frequency takes significant time to settle. This occurs each time switching to a new mode. **(B)** This could be avoided for mode 2 by tuning V_{bias} for mode 4 to 1.87V.

addressed in other work ongoing in our group by abandoning piezoresistive readout in pursuit of much lower power NEMS interrogation methods.

The Allan deviation for modes 2 and 4, operated individually at room temperature, were 6.6E–7 and 9.0E–7, respectively. Switching between the modes and stitching the data together for each mode using the bias tuning method, Allan deviation measurements were 7.0E–7 and 1.0E–6, respectively, showing that higher modes could be operated without significantly reducing performance and making inertial imaging measurements feasible.

The device was lowered to cryogenic temperatures and GroEL molecules were sent to the device while interrogating its lowest 4 modes. Unfortunately, large spikes in resonance frequency were often observed switching to a new mode, even after molecules were no longer sent to the device. This behavior only occurred at cryogenic temperatures, but did not have a long time constant like before. The origin of this phenomenon has not been resolved, and consequently, the data could not be used to perform inertial imaging. It is possible that the issue could be related to cabling imperfections, or issues related to molecular charge affecting the resonance frequency when switching actuation voltages.

Simultaneous actuation was tried again, except with lower levels of actuation for all three modes (V_{bias} at about 25% of that required for the onset of nonlinearity). Typically, each mode is driven up to the point of nonlinearity; this provides maximal signal-to-noise ratio. Lower actuation levels yield higher Allan deviations, and at very low drive levels, the device performance becomes limited by Johnson noise rather than its intrinsic 1/f mechanical domain noise. However, as discussed in Chapter 2, the performance of the piezoresistive CAL3 devices is instead limited by device heating as a consequence of the piezoresistive readout. By operating the devices at a lower temperature via a lower actuation voltage, it was

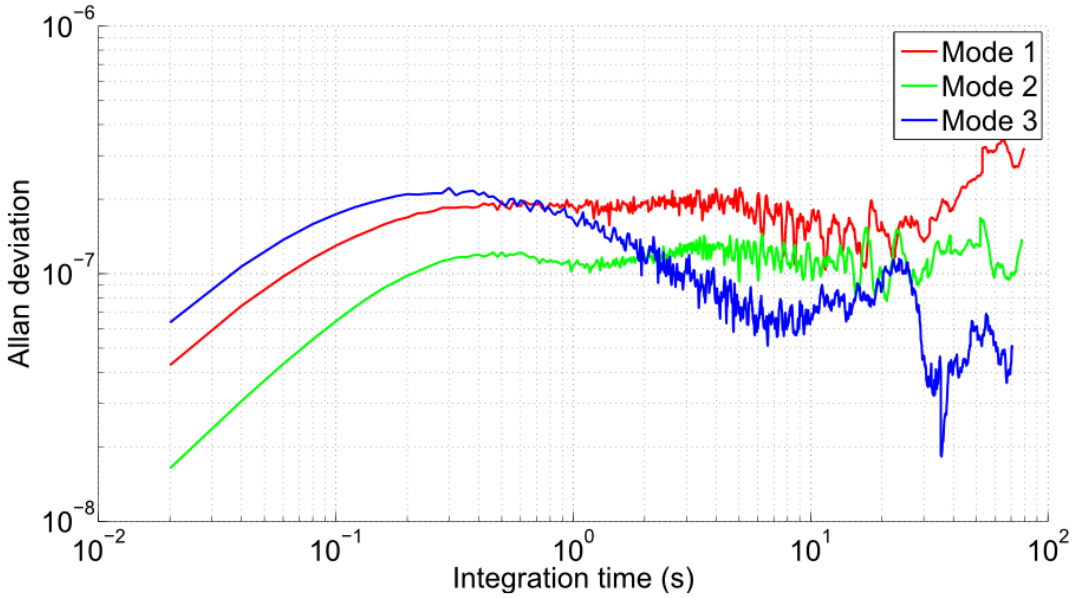


Figure 58. Allan deviation for CAL3 device actuated with 3 modes simultaneously. Devices were operated at 80K with liquid nitrogen and actuated at 25% of that required for onset of nonlinearity. Allan deviation was measured using a PLL time of 1 second.

found that the devices had lower 1/f noise floor which could be accessed by using longer integration times. The devices were cooled to 80K using liquid nitrogen (the actual device temperatures under bias conditions were not measured, but assumed to be higher than 80K). Operating at reduced actuation levels, Allan deviations of 1.91E–7, 1.08E–7, and 1.67E–7 were achieved for the first three modes, respectively, for a PLL time of 1 second (Figure 58). It was also observed that additional operation of the third mode did not significantly change the resonance frequencies of the first two modes, thus limiting the introduction of nonlinearities that could alter the measurement, as previously discussed.

3.5 Rudimentary adsorbate imaging using three modes

As discussed in Section 3.3, adsorption events are detected when the fractional frequency shifts exceed some threshold compared with the background noise level. This procedure can be generalized to three or more modes as follows. The raw frequency data are

averaged over the integration time τ_{avg} (3 seconds in these experiments), then the fractional frequency differences are calculated:

$$\bar{y}_n(i) = \frac{\bar{f}_n(i+1) - \bar{f}_n(i)}{\bar{f}_n}, \tag{41}$$

where $\bar{f}_n$ is the averaged frequency and $\bar{y}_n$ is the relative (averaged) frequency change for mode n. Using this data, the mean vector $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Sigma}$ can be calculated. In three dimensions:

$$\boldsymbol{\Sigma}(\bar{y}_1, \bar{y}_2, \bar{y}_3) = \begin{bmatrix} \sigma_1^2 & \rho_{12}\sigma_1\sigma_2 & \rho_{13}\sigma_1\sigma_3 \\ \rho_{12}\sigma_1\sigma_2 & \sigma_2^2 & \rho_{23}\sigma_2\sigma_3 \\ \rho_{13}\sigma_1\sigma_3 & \rho_{23}\sigma_2\sigma_3 & \sigma_3^2 \end{bmatrix}, \tag{42}$$

where σ_1, σ_2, σ_3 can be identified as the Allan deviations, and ρ_{12}, ρ_{13}, and ρ_{23} describe frequency noise correlation between modes. For the device used to collect data in this section, ρ_{12} = 0.675, ρ_{13} = 0.223, and ρ_{23} = 0.122. Assuming a multivariate normal distribution, contours of constant density are given by vectors $\boldsymbol{x}$ satisfying

$$(\boldsymbol{x} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{x} - \boldsymbol{\mu}) = c^2, \tag{43}$$

which are ellipsoids centered on $\boldsymbol{x} = \boldsymbol{\mu}$ with axes in the directions of eigenvectors of $\boldsymbol{\Sigma}$ and axis lengths $2c\sqrt{\lambda_i}$, where λ_i are the eigenvalues of $\boldsymbol{\Sigma}$. To define confidence regions, c^2 can be set to the quantiles of the χ^2 distribution; for 95% confidence values, one can use the inverse χ^2 for a p-value of 0.05 with 3 degrees of freedom. This works out to $c = 2.796$. A scatterplot for the averaged fractional frequency shifts of the first three modes are shown for the device along with the 95% confidence ellipsoid in Figure 59. Matlab code used to plot the confidence contour was adapted from [10].

Given the need for a slow integration time of 3 s for these experiments, GroEL molecules were sent to the device at a reduced rate of about once every 60 s. Over an 8-hour

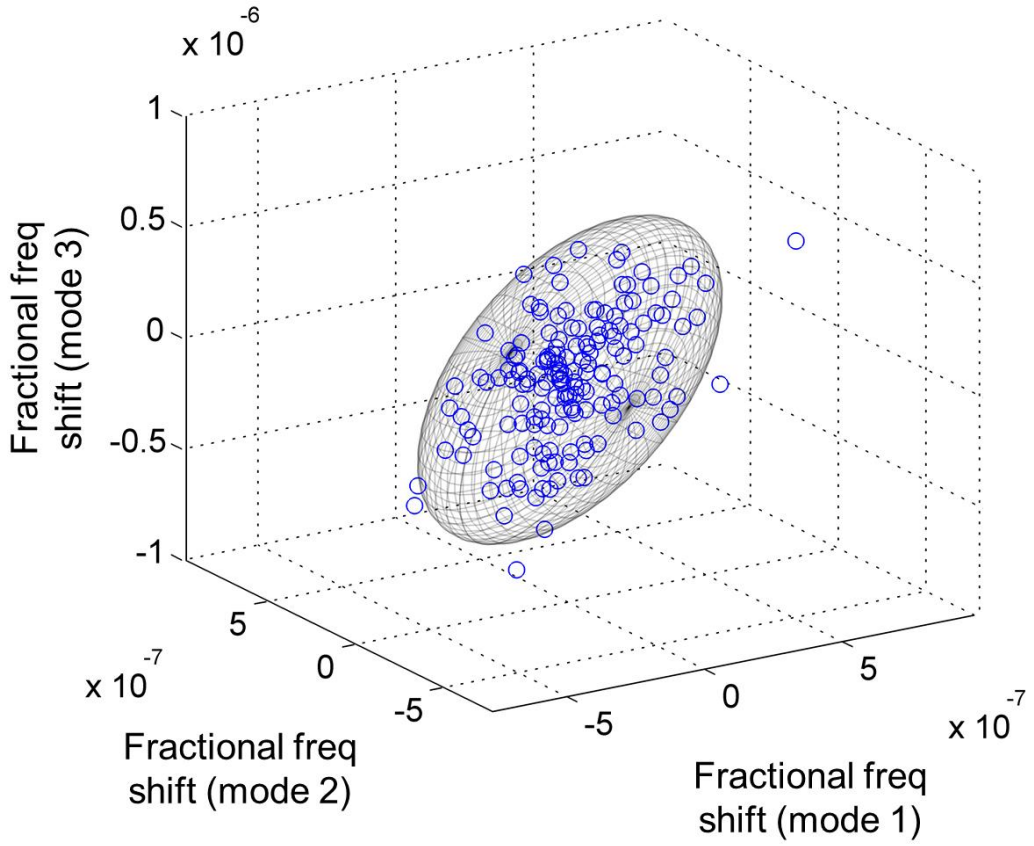


Figure 59. Noise characterization for three-mode acquisition. Relative frequency change is plotted for the first three modes actuated simultaneously with no adsorption events. An ellipsoid is plotted denoting the 95% confidence contour.

run, around 100 events were observed. A plot demonstrating adsorption events are shown in Figure 60. Actuation was performed using the 2f technique described in Section 2.7, with a constant amplitude provided throughout the entire experiment. This ensures that the frequency shifts are due to mass adsorption alone, and not conflated with charging effects that could potentially modify the frequency indirectly, as described in Section 2.7. Over the entire experiment, the device experienced a slow downward frequency drift, with each mode changing by $\Delta f/f \approx$ –3E–4. This is due to uniform adsorption of vapor and is minute enough not to affect mode shape. Over the course of the measurement window, frequency drift was not prominent as shown in Figure 60a. Additionally, the slope of the log-log plot

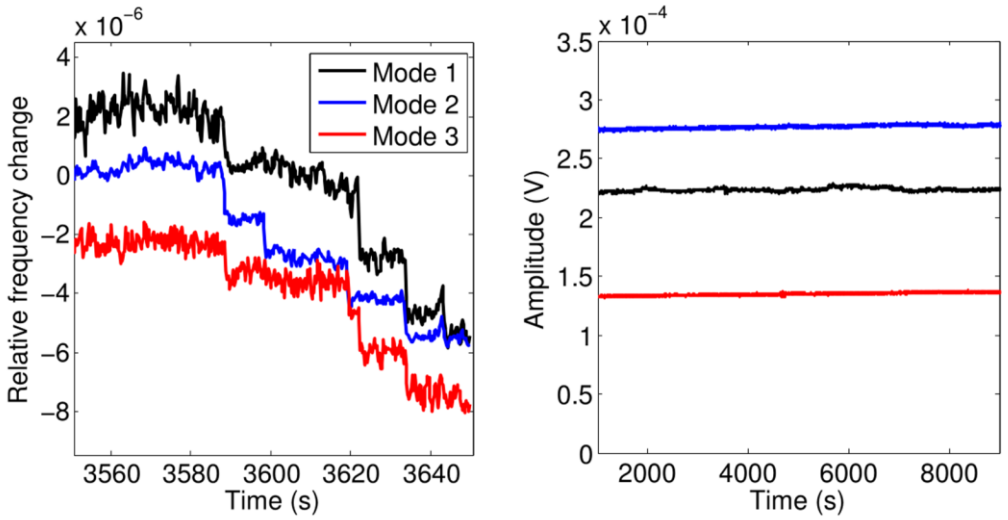


Figure 60. GroEL adsorbing to CAL3 device actuated with three modes. A long integration time of 3 seconds was used for low-noise operation. Representative frequency shifts are shown during adsorption events (left). The amplitude of all three modes was constant during the entire experiment (right).

of Allan deviation versus measurement time at the chosen PLL time as shown in Figure 58 does not indicate drift interferes with the measurement (compare with Figure 13).

Events were selected that exceeded the background noise with a confidence value of 99%. Detected frequency jumps were then used to calculate the mass, position, and variance using the truncated series approach and associated fitting algorithms (Chapter 4). Since the number of events collected was small, this was performed manually.

A sample of the mass moment calculations for the three-mode data are shown in Figure 61. Mass moments were calculated using the truncated series expansion approach introduced in Chapter 4 (of Part II, Theory) in this thesis. Some of the events (shown in yellow) could be immediately excluded due to the large uncertainty in adsorbate mass. A few of the events (shown in blue) demonstrated good precision in uncertainty for the standard deviation, but the calculated values (~700 nm) far exceeds the expected (1D) contact area for

Mass (kDa)	Pos	Std (nm)	Mass %Uncert	Std %Uncert
740	0.302	0	25.0	90.3
776	0.310	327	12.4	171.0
994	0.344	276	8.7	107.7
598	0.307	0	16.0	4753.6
477	0.297	0	109.5	70.4
860	0.297	396	13.5	138.5
1070	0.353	161	7.9	285.9
883	0.301	455	13.6	87.0
1432	0.343	670	12.4	12.6
537	0.293	0	46.0	106.8
860	0.260	298	30.4	647.6
246	0.377	1835	59.9	51.5
1486	0.382	736	9.3	13.6
821	0.315	0	11.8	244.5
2028	0.401	0	4.9	33.7

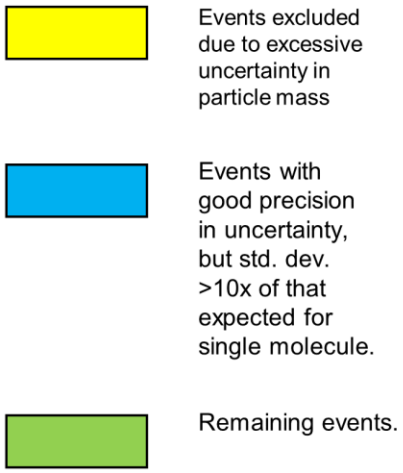


Figure 61. Representative sample of mass moment calculations using three-mode data. Relative frequency shifts were individually input into the truncated series method for calculating mass moments (Chapter 4). The method provides calculated moments as well as uncertainties.

GroEL (~15 nm). These events were interpreted to be multiple GroEL molecules landing within the measurement timeframe, but at significantly different locations on the beam. The remaining events (in green) had good uncertainty for mass, and the measured mass was typically within the range of a single GroEL molecule. These calculations were interpreted to demonstrate that the three-mode setup could distinguish between a single GroEL molecule landing, versus two molecules landing at different positions, but could not measure the spatial extent of a single molecule. The goal of inertial imaging of single molecules was not reached, but the data still demonstrates in some sense a rudimentary form of molecular imaging (adsorbate arriving in a single vs. two separate components). The failure to extract the extents of single-molecule events could be due to the low signal-to-noise present in the experiment, along with the possibility that single molecules have a significant rotational inertia

contribution to their kinetic energy due to their geometry relative to the in-plane beam motion (as discussed in Section 1.7).

To calculate a mass spectrum, the joint probability distribution formulation introduced in Section 3.3 needs to be extended to three dimensions. Specifically, the probability density function for a given *mass*, *position*, and *standard deviation* to satisfy data with relative frequency shifts $\mu_1 = df_1/f_1$, $\mu_2 = df_2/f_2$, and $\mu_3 = df_3/f_3$ is given by:

$$\mathrm{PDF}(\delta m, a, \langle x^2 \rangle) = |\boldsymbol{J}| \frac{1}{\sqrt{(2\pi)^3 |\boldsymbol{\Sigma}|}} \exp\left(-\frac{1}{2} (\boldsymbol{h} - \boldsymbol{\mu})^{\mathrm{T}} \boldsymbol{\Sigma}^{-1} (\boldsymbol{h} - \boldsymbol{\mu}) \right);$$

$$\boldsymbol{J} = \frac{(\delta m)^2}{\alpha_1 \alpha_2 \alpha_3} \det \begin{bmatrix} \phi_1^2(a) + \frac{1}{2}\frac{d^2\phi_1^2(a)}{dx^2}\langle x^2 \rangle & \phi_2^2(a) + \frac{1}{2}\frac{d^2\phi_2^2(a)}{dx^2}\langle x^2 \rangle & \phi_3^2(a) + \frac{1}{2}\frac{d^2\phi_3^2(a)}{dx^2}\langle x^2 \rangle \\ \frac{d\phi_1^2(a)}{dx} + \frac{1}{2}\frac{d^3\phi_1^2(a)}{dx^3}\langle x^2 \rangle & \frac{d\phi_2^2(a)}{dx} + \frac{1}{2}\frac{d^3\phi_2^2(a)}{dx^3}\langle x^2 \rangle & \frac{d\phi_3^2(a)}{dx} + \frac{1}{2}\frac{d^3\phi_3^2(a)}{dx^3}\langle x^2 \rangle \\ \frac{1}{2}\frac{d^2\phi_1^2(a)}{dx^2} & \frac{1}{2}\frac{d^2\phi_2^2(a)}{dx^2} & \frac{1}{2}\frac{d^2\phi_3^2(a)}{dx^2} \end{bmatrix};$$

$$h_{i=1,2,3} = -\frac{\delta m}{2\alpha_i} \left(\phi_i^2(a) + \frac{1}{2}\frac{d^2\phi_i^2(x)}{dx^2}\bigg|_{x=a} \langle x^2 \rangle \right), \tag{44}$$

where $\boldsymbol{J}$ is the Jacobian of the coordinate system transformation, $\boldsymbol{\Sigma}$ is given by Equation 42, and $h_{i=1,2,3}$ are the modeled frequency shifts induced by the random variables δm, a, and $\langle x^2 \rangle$ using the truncated series expansion approach developed in Chapter 4. The mass spectrum can be calculated by integrating over position and variance:

$$\mathrm{PDF}(\delta m) = \int_{\langle x^2 \rangle = 0}^{\langle x^2 \rangle = 1} \int_{a=0}^{a=0.5} \mathrm{PDF}(\delta m, a, \langle x^2 \rangle)\, da\, d\langle x^2 \rangle. \tag{45}$$

To compute the mass spectrum, the three-dimensional probability density function was calculated for each single-molecule adsorption event, determined as illustrated in Figure 61, and then projected onto the mass domain according to Equation 45. The results are shown in Figure 62. The primary peak at 795 kDa is extremely accurate. Similarly to the spectrum calculated with two modes and severe position restriction (Figure 54), a secondary peak can

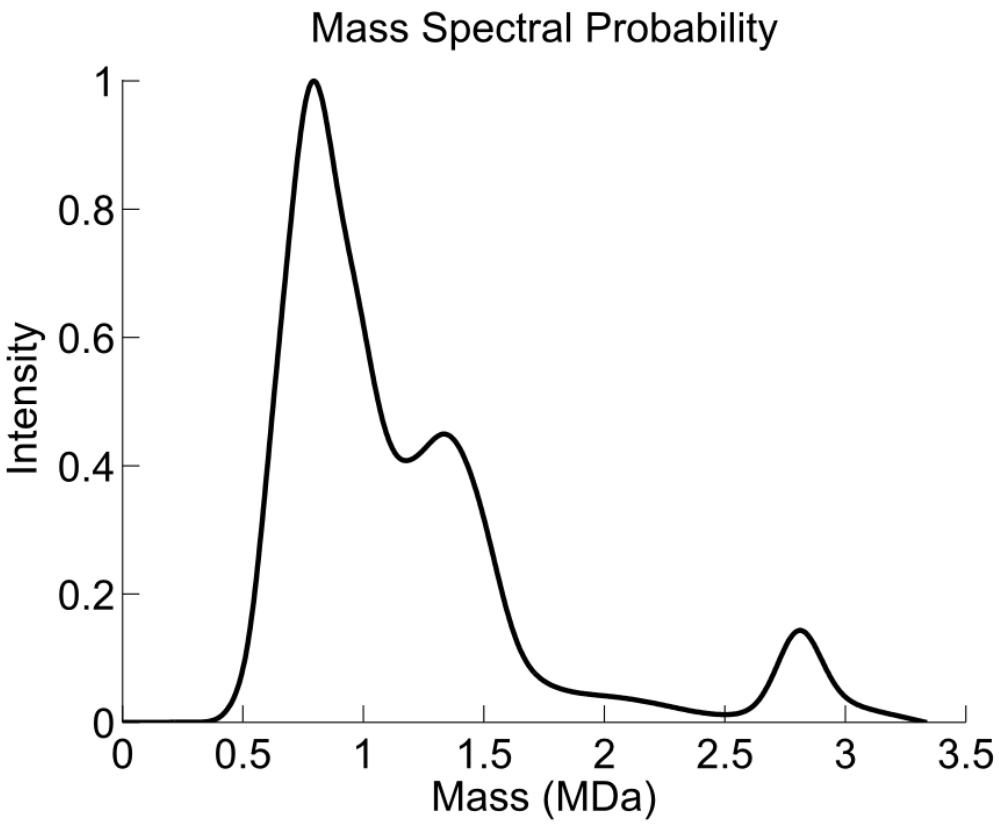


Figure 62. Mass spectrum for GroEL events calculated from three-mode data. The dominant peak is at 795 kDa and has a FWHM of 432 kDa. A secondary peak occurs near twice the mass of GroEL. A third peak is observed at 2.812 kDa.

be resolved near twice the mass of GroEL, indicating two molecules landing together in the same location.

The uncertainty for the dominant peak is 432 kDa, which is similar to that obtained with the two-mode measurement, here despite the lower Allan deviation. To see why this might be the case, contour plots were produced to show the probability distribution for each event in mass and variance space (integrating the 3D probability density function over the position dimension). Presumed single- and double-molecule landing events are shown with the experimental GroEL data in Figure 63 and compared with simulation data (using the

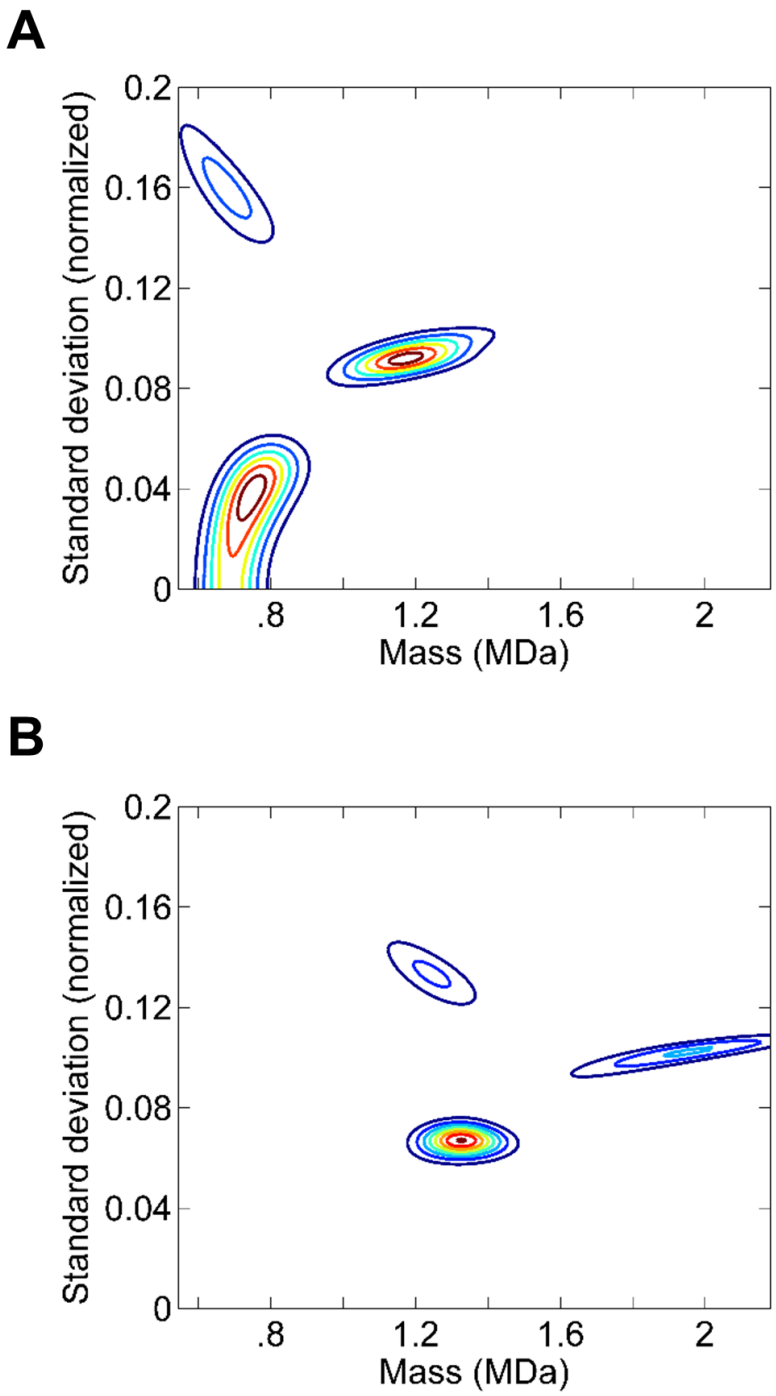


Figure 63. Probability density for mass and standard deviation for representative adsorption events (experimental). Density functions integrated over the position dimension are shown for **(A)** a suspected single-molecule event and **(B)** a suspected double-molecule event. The single-molecule event features an additional region in the distribution at higher mass and higher variance, which would broaden the final mass spectrum if not removed.

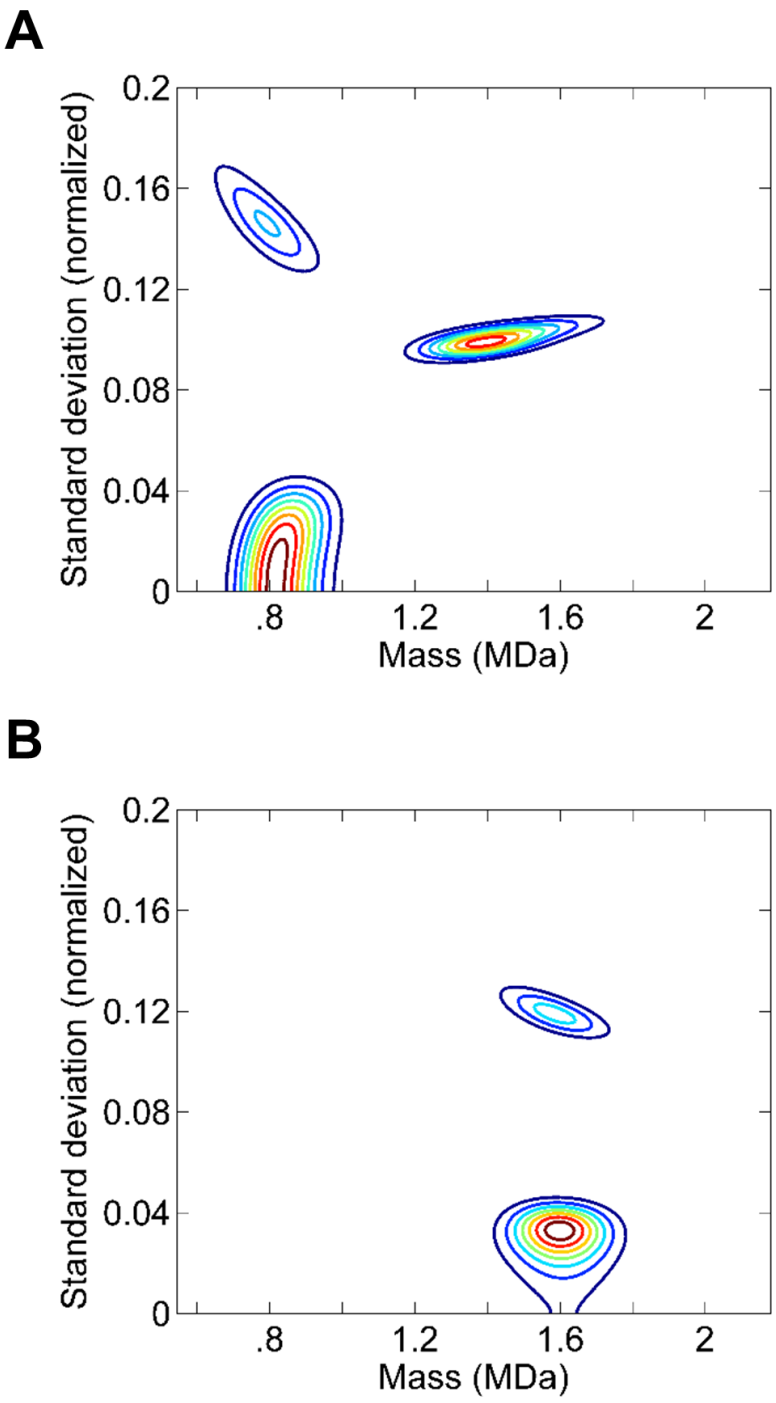


Figure 64. Probability density for mass and standard deviation for simulated adsorption events. Density functions integrated over the position dimension are shown for **(A)** a single-molecule event with mass of exactly 801 kDa and **(B)** two molecules of 801 kDa landing at different positions. The single-molecule event features an additional region in the distribution at higher mass and higher variance similar to that seen in the experimental data.

device's measured noise characteristics) in Figure 64. Both the experimental and simulation-based density contours show that the addition of variance as a third degree of freedom to describe a single-particle adsorption event results in a new region in the distribution allowing for the possibility of a higher mass, higher variance solution to the data. This possibility, while less probable than the correct mass, lower variance solution, gets projected onto the mass spectrum, leading to unnecessary broadening of the mass spectrum even with more information available. Therefore, once the particle has been identified as a single-molecule event (with the peak value of the distribution belonging to a low-variance region), other regions in the distribution should be removed prior to calculating the mass spectrum. A demonstration of this for the particle shown in Figure 63a is shown in Figure 65. Removing

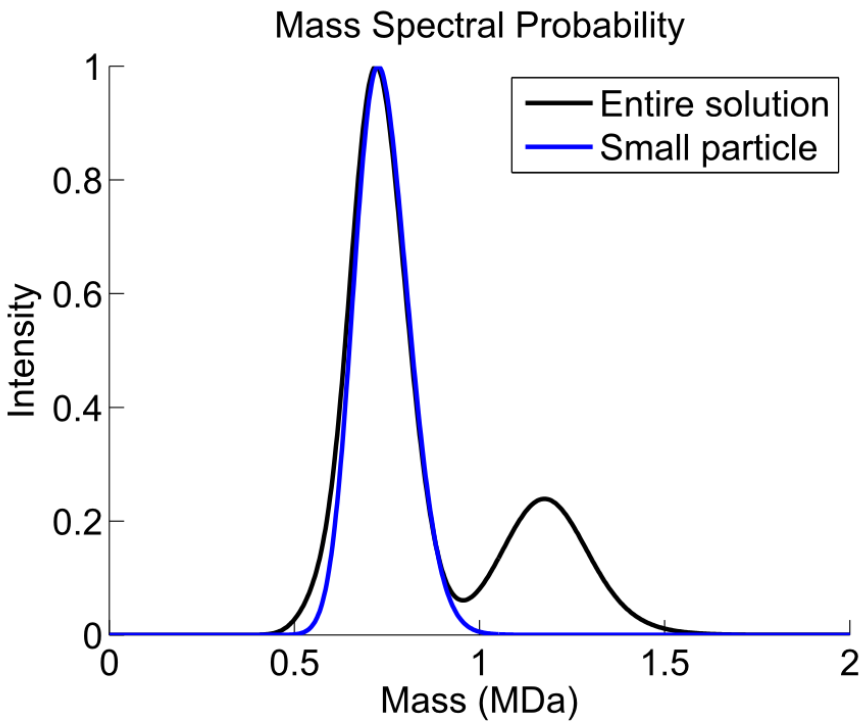


Figure 65. Mass spectrum for GroEL event after identifying as single particle. The 3D probability distribution in mass, position, and variance space for the GroEL event shown in Figure 63a has three regions; the region corresponding to the most probable solution (in the lower left) corresponds physically to a small particle solution and was integrated over other variables to produce a mass spectrum featuring less uncertainty.

data for the probability distribution for normalized standard deviations exceeding 0.075 results in a simplified mass spectrum for the particle.

This process was repeated for all GroEL data identified as single molecule events through the method illustrated in Figure 61. Each individual event was found to have multiple regions to the probability distribution in mass-variance space similar to the representative event shown in Figure 63a, with at least one region extending to zero variance, indicating a small particle solution. The small particle region of the probability distribution was isolated for each event, and integrated over other variables to produce a more representative mass spectrum for GroEL single particle events as shown in Figure 66. This mass spectrum

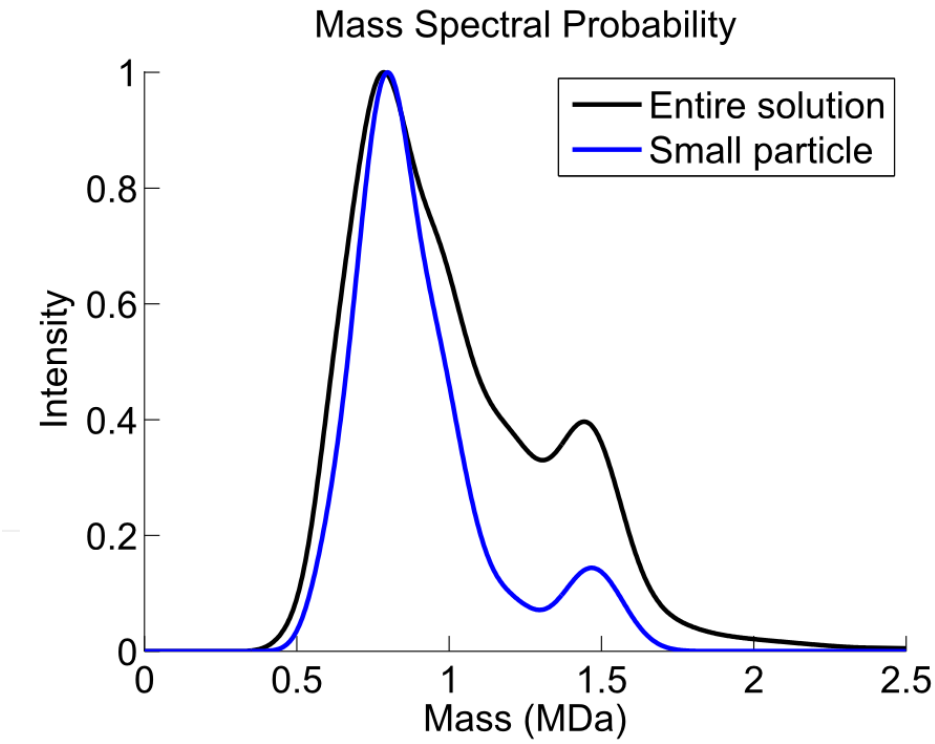


Figure 66. Mass spectrum for GroEL single particle events. The 3D probability distribution in mass, position, and variance space for the GroEL data identified as single particle events were projected onto mass ("Entire solution"). Regions of their distributions corresponding to a small particle solution (those that include zero variance) were isolated and integrated over other variables to produce a more accurate mass spectrum ("Small particle").

features a peak of 797 kDa with a FWHM of 319 kDa, as well as a secondary peak at 1.469 kDa. These peaks are close to the mass and twice the known mass of GroEL (801 kDa).

Finally, the process was repeated for events identified as two molecules landing at different locations on the beam. As illustrated in Figure 63b, there are multiple regions in the multidimensional probability distribution. The region with the highest probability density was isolated and integrated over other variables to produce mass spectra. These mass spectra were combined with those obtained with the single particle events to obtain a final mass spectra as shown in Figure 67. Selecting an appropriate region of the probability distribution for each event led to an improved mass spectrum, with the main peak of 805 kDa with a

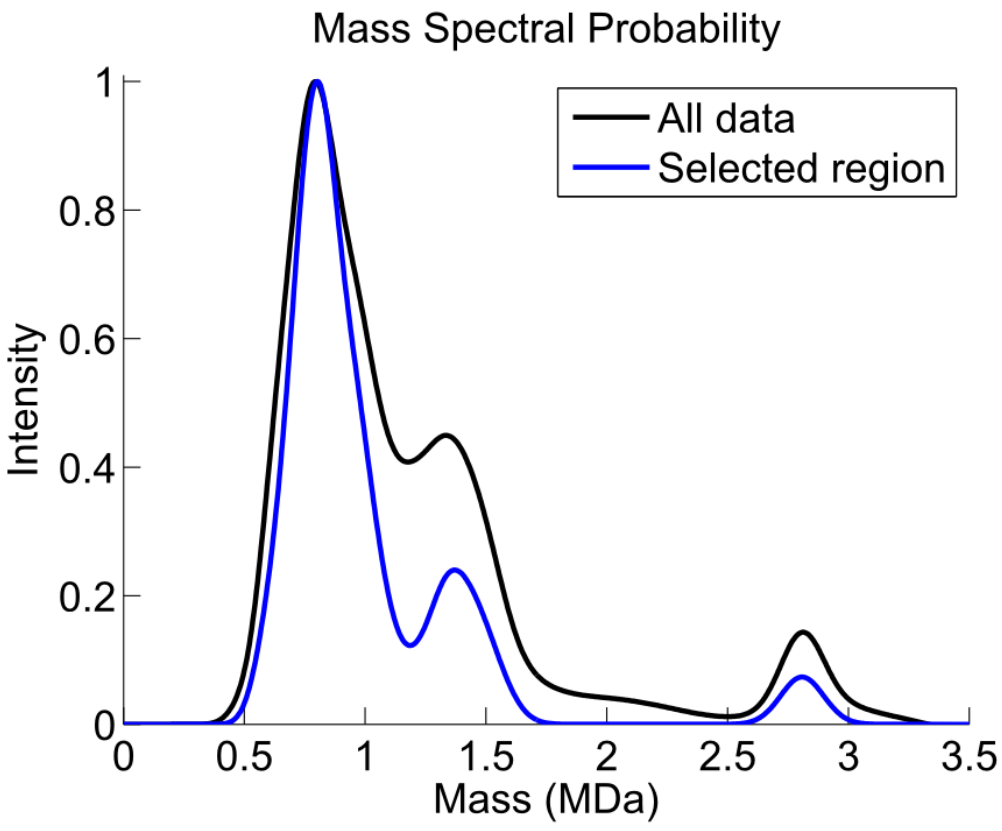


Figure 67. Mass spectrum for all GroEL events calculated from three-mode data. The spectrum is shown integrating the 3D probability density over position and variance, using either the entire distribution or the appropriate selected region as discussed in the text. The effective mass resolution for the final mass spectrum is 131 kDa.

FWHM of 310 kDa near the exact mass of GroEL and nearly resolved from the second and third peaks. The data features an effective mass resolution of σ = FWHM / 2.355 = 131 kDa. It bears repeating, however, that single-molecule events could have significant rotational inertia contribution based on their geometry relative to the in-plane device motion. This effect could possibly be mitigated using devices with transverse motion, or adding additional corrects to the model to account for this effect.

3.6 Summary

In this chapter, GroEL molecules were sent to NEMS devices operated with two or three modes using the hybrid Orbitrap-NEMS instrument. An array of 20 NEMS devices was initially used to characterize the ion beam in three dimensions, allowing for precise positioning of the devices to allow for high (or low) ion flux. High flux measurements were performed at ~1 molecule per second with two modes, along for a large amount of GroEL adsorption events to be collected. The mass spectrum featured a large uncertainty in mass, but enough events were collected that all but 6% of the data landing in a small region of the device could be discarded. The remaining data provided more information and yielded a higher-quality spectrum revealing two peaks, one centered at GroEL's expected mass, and one near twice the mass, indicating some events were associated with two molecules landing simultaneously.

GroEL molecules were also collected with three modes operated simultaneously. Good noise performance could be obtained by operating the devices at lower bias voltages (and hence, lower temperature) with long integration times. In this regime, only ~100 events could be collected over 8 hours, but the results readily demonstrate a proof-of-concept that single- vs double-molecule landing events can be distinguished with the second mass

moment (variance) calculated using inertial imaging theory. While this falls short of our starting goal of imaging the spatial mass distribution of individual proteins, the ability to distinguish between a single point adsorbate and one with an extended mass distribution can be considered a form of rudimentary imaging. Carefully analyzing the full three-dimensional probability distribution in mass, position, and variance space allows for the construction of an improved mass spectrum compared with that collected using only two modes.

3.7 Bibliography

1. Sage, E., *Nouveau concept de spectromètre de masse à base de réseaux de nanostructures résonantes (Doctoral dissertation)*. 2013, Grenoble University.
2. Hanay, M.S., et al., *Single-protein nanomechanical mass spectrometry in real time.* Nature Nanotechnology, 2012. **7**(9): p. 602-608.
3. Hanay, M.S., et al., *Inertial imaging with nanomechanical systems.* Nature nanotechnology, 2015. **10**(4): p. 339-344.
4. Sader, J.E., et al., *Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation.* Nano Letters, 2018. **18**(3): p. 1608-1614.
5. Sage, E., et al. *Frequency-addressed NEMS arrays for mass and gas sensing applications*. in *Transducers & Eurosensors XXVII: The 17th International Conference on Solid-State Sensors, Actuators and Microsystems*. 2013. IEEE.
6. Hanay, M.S., *Towards single-molecule nanomechanical mass spectrometry (Doctoral Dissertation)*. 2011, California Institute of Technology.
7. Mile, E., et al., *In-plane nanoelectromechanical resonators based on silicon nanowire piezoresistive detection.* Nanotechnology, 2010. **21**(16): p. 165504.
8. Sage, E., et al., *Neutral nanomechanical mass spectrometry and direct comparisons to time-of-flight mass spectrometry.* Nature Communications, 2015. **6**(1): p. 1-5.
9. Hung, P.S.-H., *Advanced applications of nanoelectromechanical systems (Doctoral dissertation)*. 2016, California Institute of Technology.
10. "How to draw a covariance error ellipse?" Comment by Meysam. Available from: https://www.visiondummy.com/2014/04/draw-error-ellipse-representing-covariance-matrix/#comment-190.

PART II: Theory
Chapter 4

BEYOND THE POINT-MASS APPROXIMATION FOR NEMS-MS

The mass measurement of single molecules, in real time, is performed routinely using resonant nanomechanical devices. This approach models the molecules as point particles. A recent development now allows the spatial extent, and indeed image, of the adsorbate to be characterized using multi-mode measurements [1]. This "inertial imaging" capability is achieved through virtual re-engineering of the resonator's vibrating modes, by linear superposition of their measured frequency shifts. Here, we present a complementary and simplified methodology for the analysis of these inertial imaging measurements that exhibits similar performance, while streamlining implementation. This development, together with the software that we provide, enables the broad implementation of inertial imaging which opens the door to a range of novel characterization studies of nanoscale adsorbates.

This chapter was completed in collaboration with Profs. John Sader, Selim Hanay, and Michael Roukes, and published in [2].

4.1 Introduction

Nanomechanical resonators can be used as fast and sensitive mass balances due to their small mass, high vibrational frequencies, and low intrinsic energy dissipation [3-10]. The strong dependence of mass responsivity on device size has driven the development of a new type of mass spectrometer based on inertial mass sensing using nanoelectromechanical systems (NEMS) [11, 12]. This technology enables very precise measurements particularly

for high molecular weight biomaterials. Measurements to date include the weighing of individual proteins, metal nanoparticles, large biomolecules and demonstration of mass sensing with near atomic-scale-mass resolution [13-16].

The above-mentioned measurements use the property that the resonant frequency of a nanomechanical device is directly related to its inertial mass. Increasing the device's inertial mass, via mass adsorption on its surface, reduces the resonant frequencies of its vibrating modes in a deterministic fashion. Central to all such measurements of adsorbed mass is a theoretical model enabling conversion of the observed frequency shifts to an added (adsorbed) mass. The most common approach is to approximate the adsorbate by a point mass. This leads to the well-known result for a one-dimensional elastic beam:

$$\frac{\delta f_n}{f_n} \approx -\frac{m_{\text{added}}}{2M_{\text{device}}}\phi_n^2(x), \tag{46}$$

where δf_n is the change in resonant frequency of the device upon mass adsorption, f_n is the resonant frequency of the device without the adsorbate, m_{added} is the adsorbed mass at position x, M_{device} is the device mass, and $\phi_n(x)$ is the scaled displacement mode shape of mode n; this formula is valid for $m_{\text{added}} \ll M_{\text{device}}$. Equation 46 shows that the frequency shift depends not only on the adsorbed mass but also on its position, because $\phi_n(x)$ varies spatially in general – without knowledge of this position, the adsorbed mass cannot be determined. However, use of two (or more) modes can disentangle this ambiguity, leading to simultaneous determination of the adsorbate's position and mass [11, 12]; at least three modes are required for a cantilever beam. Despite the demonstrated success of this so-called "two-mode theory," data pertaining to the adsorbate's shape is absent in all such point particle analyses.

Recently, a methodology was proposed that discards the point particle approximation and allows the spatial distribution of the adsorbate's mass density to be measured, i.e., an "inertial image" of the adsorbate is obtained [1]. This inertial imaging methodology uses the property that within all the measured fractional frequency shifts, $\delta f_n/f_n$, lie untapped information on the adsorbate's shape. By linearly superposing the measured frequency shifts, the square of the device's vibrating modes are virtually re-engineered to yield flat (constant), linear, parabolic, cubic, etc. spatial dependencies. This virtual re-engineering enables all moments of the analyte's mass distribution to be measured simply by adding constant multiples of the observed fractional frequency shifts of several device modes. That is, the mass, position, variance, skewness etc. of the adsorbate's mass density distribution are determined simultaneously. This "linear superposition" approach was proposed and demonstrated in Ref. [1] using both synthetic numerical data and experimental measurements of liquid droplets deposited on a microcantilever in air (Figure 68).

Critically, the spatial resolution of this inertial imaging technique is limited by uncertainty (noise) in the measured frequency shifts, not the spatial wavelength of the vibrating modes, i.e., diffraction-like phenomena play no role. Theoretical projections show that, with frequency noise levels in current NEMS devices, simultaneous mass measurement and imaging with molecular-scale resolution is possible. The methodology presented in Ref. [1] involves a numerical algorithm to perform the required virtual mode re-engineering and iteratively determine the measurement zone (i.e., the spatial region of the device) over which

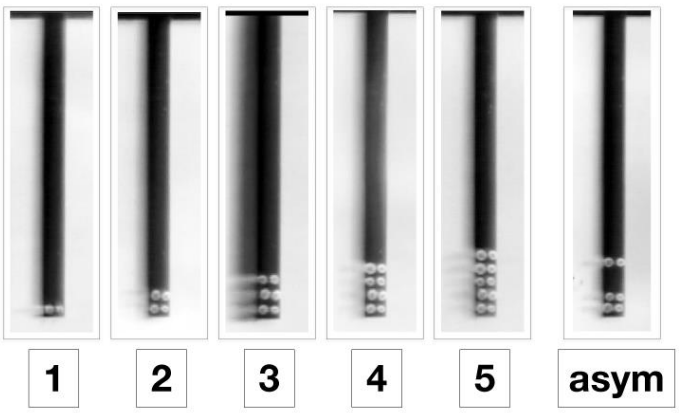


Figure 68. Images of microcantilevers with known mass distribution. Optical image for droplets deposited on a microcantilever using AFM dip-pen lithography to generate samples with known mass, mean position, extent, and skewness. After each droplet deposition, the frequency shift in the cantilever modes were measured. The analysis of mechanical frequency shifts yielded the spatial properties of deposited droplets in Ref. [1]. A simplified theoretical approach is presented here to obtain the spatial properties of the same analytes. Cantilever dimensions are 397 μm long, 29 μm wide, and 2 μm thick. Figure is taken from Ref. [1] and replicated here.

the moments are to be determined. Since this approach requires some computational analysis, it is desirable to simplify its underlying theoretical framework to enable its general use.

Here, we develop such a complementary theoretical framework for measuring the mass distribution of an analyte adsorbed to a resonating nanomechanical device. Rather than relying on linear mode superposition to virtually re-engineer the mode shapes, and thus determine the required moments of the mass distribution [1], we directly determine the moments of the mass distribution using a nonlinear analytical formula; requirement to determine a measurement zone is also eliminated. Formulas for both one- and two-dimensional devices are presented. These formulas are simple and can be employed using standard packages like Mathematica and Matlab; Wolfram CDF Player apps are available upon request, screen shots of which are provided in the Supporting Information for [2]. The formulas are derived using a local (asymptotic) expansion of the mass distribution, whereas the linear superposition approach of Ref. [1] uses a global approach of linearly superposing

the mode shapes over a finite region – these frameworks present complementary approaches for interpreting multi-mode NEMS mass adsorption data.

We examine the performance of the new nonlinear formulas using both numerical (synthetic) data and the above-mentioned published experimental measurements [1], and contrast them to the properties of the linear superposition theory. Strikingly, both theories exhibit similar properties and dependencies on frequency noise, i.e., their spatial resolutions are identical. In fact, these complementary theories present an inverse set of relations in the limit of small particle size, as we shall discuss. We show that the two-mode theory outlined above, which implements a point-mass approximation, is simply a subset of this general nonlinear theoretical formula. The one- and two-dimensional formulas are applicable to beam and plate resonators, respectively, under arbitrary boundary conditions – they can therefore be used to analyze a multitude of measurement configurations. The number of modes required to determine the moments is also explored.

4.2 Theory

We first consider a one-dimensional elastic beam resonator. The leading-order expression for the fractional frequency shift of this resonator with an adsorbate mass, m_{added}, that is far smaller than that of the resonator, M_{device}, is:

$$\frac{\delta f_n}{f_n} = -\frac{1}{2M_{\text{device}}} \int_0^L \mu_{1\text{D}}(x)\phi_n^2(x)\, dx, \tag{47}$$

where $\mu_{1\text{D}}(x)$ is the linear mass density of the adsorbate, L is the device length and the scaled displacement mode shapes, ϕ_n, form an orthonormal basis set, i.e., $\int_0^L \phi_m \phi_n\, dx = L\, \delta_{mn}$, where δ_{mn} is the Kronecker delta function.

We expand $\phi_n^2(x)$ in its Taylor series about the center-of-mass (position) of the adsorbate, $\bar{x}$, and substitute the result into Equation 46, giving:

$$\frac{\delta f_n}{f_n} = -\frac{1}{2M_{\text{device}}}\left[\phi_n^2(\bar{x})\int_0^L \mu_{1\text{D}}(x)dx + \left.\frac{\partial \phi_n^2}{\partial x}\right|_{x=\bar{x}}\int_0^L \mu_{1\text{D}}(x)(x-\bar{x})dx \right.$$

$$\left. + \left.\frac{\partial^2 \phi_n^2}{\partial x^2}\right|_{x=\bar{x}}\int_0^L \mu_{1\text{D}}(x)\frac{(x-\bar{x})^2}{2}dx + \cdots\right]. \tag{48}$$

Knowledge of the position, $\bar{x}$, is not required yet – it is determined from measurements of the fractional frequency shifts of multiple modes using the resulting formula; see below. Importantly, the second integral in Equation 48 vanishes because it is the difference of the first moment with itself. This gives the required result

$$\frac{\delta f_n}{f_n} = -\frac{m_{\text{added}}}{2M_{\text{device}}}\left\{\phi_n^2(\bar{x}) + \sum_{p=2}^{p=\infty}\frac{1}{p!}\left.\frac{\partial^p \phi_n^2}{\partial x^p}\right|_{x=\bar{x}}\langle x^p\rangle\right\}, \tag{49}$$

where m_{added} is the adsorbate's mass and the central moments are defined

$$\langle x^p\rangle \equiv \frac{1}{m_{\text{added}}}\int_0^L \mu_{1\text{D}}(x)(x-\bar{x})^p dx\ . \tag{50}$$

Equation 49 enables the central moments, $\langle x^p\rangle$, of the adsorbate's mass distribution to be determined from the measured fractional frequency shifts, $\delta f_n/f_n$, of multiple modes.

The corresponding result for a two-dimensional structure, e.g., an elastic membrane, is obtained in an analogous fashion, yielding the required formula:

$$\frac{\delta f_n}{f_n} = -\frac{m_{\text{added}}}{2M_{\text{device}}}\left\{\Phi_n^2(\bar{x},\bar{y}) + \sum_{p=2}^{p=\infty}\sum_{j=0}^{p}\frac{\langle x^{p-j}y^j\rangle}{(p-j)!\,j!}\left(\left.\frac{\partial^p \Phi_n^2}{\partial x^{p-j}\partial y^j}\right|_{(x,y)=(\bar{x},\bar{y})}\right)\right\}, \tag{51}$$

where (x,y) is the Cartesian coordinate system in the plane of the structure, $(\bar{x},\bar{y})$ is the adsorbate's center-of-mass, and the two-dimensional moments are

$$\langle x^p y^q \rangle \equiv \frac{1}{m_{\text{added}}} \int_S \mu_{2\text{D}}(x,y)(x-\bar{x})^p (y-\bar{y})^q \, dS, \tag{52}$$

whereas $\mu_{2\text{D}}\ (x,y)$ is the areal mass density distribution of the adsorbate and the region of integration is the structure's surface area, *S*. The scaled displacement mode shapes $\Phi_n(x,y)$ in Equation 51 form the orthonormal basis set, $\int_S \Phi_m \Phi_n \, dS = \ S\ \delta_{mn}$. Equation 51 allows determination of the central moments, $\langle x^p y^q \rangle$, of a two-dimensional adsorbed mass distribution from the measured fractional frequency shifts of multiple modes of a two-dimensional mechanical resonator.

As mentioned above, Equations 49 and 51 apply to all one-dimensional and two-dimensional resonators, regardless of their boundary conditions, such as cantilevered/simply supported beams, square plates etc. The mode shapes intrinsically embody the structure's boundary conditions.

We now discuss some features and practicalities of the above nonlinear formulas, Equations 49 and 51, for inertial imaging using nanomechanical devices.

Number of Modes: Equations 49 and 51 can be implemented in practice by truncating their infinite series and solving for the unknown moments using a numerical root finding procedure. Numerical solutions of Equations 49 and 51 are easily found using built-in functions of standard packages such as Mathematica and Matlab, e.g., using Newton's method. Importantly, the number of required moments of the mass distribution specifies the number modes that need to be measured. For example, to measure the variance of a (1D) mass attached to a beam resonator requires at least three modes, because its specification involves the three lowest moments of the distribution. We discuss the convergence of the moments with increasing mode number below.

We focus on the one-dimensional resonator formula, Equation 49, because (i) these structures are widely used [3-10], and (ii) the properties of this formula are identical to Equation 51 for two-dimensional structures.

Two-mode (point-mass) theory: Two-mode theory, Equation 46 for a one-dimensional elastic beam, is clearly a subset of the general formula, Equation 49 – the latter includes all moments of the adsorbate's mass distribution. Two-mode theory is obtained by truncating this general formula to include only the zeroth and first-order moments of the mass distribution; in-line with the methodology described above. Use of Equation 49, rather than Equation 46, enables higher-order moments to be evaluated using the same methodology as two-mode theory, through systematic inclusion of higher-order terms in the expansion.

Central Moments: Another feature of the nonlinear formulas, for one-dimensional and two-dimensional structures, is that the moments are specified about the mean position of the adsorbate's mass distribution. This contrasts to the linear superposition theory reported in Ref. [1] that evaluates the moments about a (user-specified and fixed) reference point on the device, e.g., the clamped position of a cantilever beam. The required central moments, about the center-of-mass of the adsorbate, are subsequently determined by combining these fixed-reference-point moments. The present formulation eliminates this requirement because the central moments are evaluated directly; see Equations 49 and 51. As we shall see, this does not affect the final result (and its uncertainty), but direct determination of the central moments does simplify analysis.

Convergence: While Equations 49 and 51 use an (asymptotic) Taylor expansion of the mode shape squared, they are expected to converge with increasing mode number, regardless of the spatial extent of the mass distribution. This is because the mode shapes can

be expressed in terms of elementary functions such as trigonometric and hyperbolic functions. That is, exponential functions whose Taylor expansions are themselves uniformly convergent for all values of their arguments, i.e., they are entire functions. This convergence can be proved for one-dimensional devices by performing a ratio test for the series in Equation 49 (see Supporting Information of Ref. [2]).

Equations 49 and 51 and linear superposition theory are both derived from Equation 47, albeit using different assumptions. The linear superposition theory takes a global approach of virtually re-engineering the modes over a finite spatial domain of the device to obtain the required weights for each moment. In contrast, Equations 49 and 51 locally expand the device modes about the adsorbate's center-of-mass using Taylor expansions. In the limit of small adsorbate size, both formulations are expected to yield identical results – this is observed in the numerical results reported below.

Moreover, Equations 49 and 51 express the fractional frequency shifts in terms of the adsorbate's moments, whereas linear superposition theory presents the inverse relation: moments as a function of the fractional frequency shifts. These complementary theories therefore present a set of inverse functions. Indeed, a Taylor expansion can be used to formulate linear superposition theory in the limit of small adsorbate size; see Supplementary Section 5 of Ref. [1].

4.3 Results and Discussion

We now assess the utility of Equation 49 for recovering the moments of an adsorbate's mass distribution. This is first performed by generating (synthetic) numerical frequency shift data from Equation 47, for a specified mass distribution, and then recovering the moments of this distribution using Equation 49. Second, we reanalyze the experimentally

measured frequency shift data reported in Ref. [1] using Equation 49, and compare its results to those of the linear superposition method [1]. The one-dimensional resonator considered here is a cantilevered elastic beam, and we use its lowest-order flexural modes of vibration in the analysis.

Synthetic data. The adsorbate's mass distribution chosen for this analysis is illustrated in Figure 69. It consists of two individual rectangular distributions of different magnitude, leading to an overall skewed distribution. The width of each rectangular distribution is the same, while the overall width, ϵ, of the adsorbed mass is varied. This enables assessment of the robustness of Equation 49 in recovering the required central moments as a function of the width of the adsorbate. The cantilever length, L, scales all distances.

We first examine the performance of Equation 49 using data for the lowest four flexural modes of the cantilever beam. This approach is identical to the measurement analysis

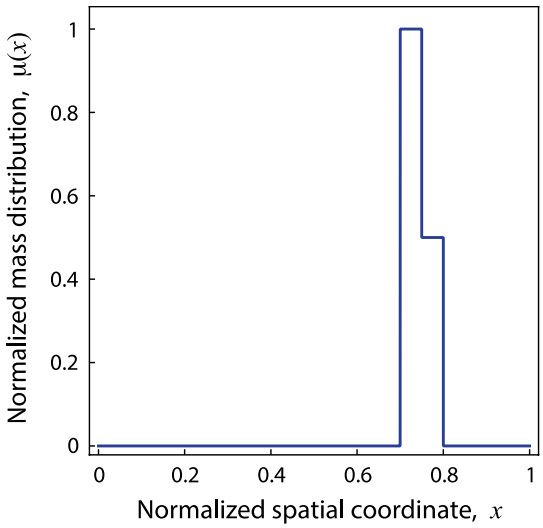


Figure 69. Simulated 1D mass density distribution of a sample adsorbate on an elastic beam. The adsorbate's spatial extent (width) is ϵ, and the adsorbate is formed from two rectangular density distributions of equal width that differ in magnitude by a factor of two. The lower boundary of the adsorbate is positioned at $x = 0.7$ and its overall width is $\epsilon = 0.1$. The spatial coordinate is scaled by the device length, L, whereas the mass distribution is scaled by its maximum value, μ_0.

reported in Ref. [1], and allows the robustness of using a fixed number of modes to be assessed; the exact moments of the adsorbate's mass distribution are known in the present case.

Table 9 gives a comparison of (i) the recovered mass (zeroth moment), center-of-mass (position), variance, and skewness of the sample mass distribution using Equation 49, and (ii) the exact moments, as a function of its overall width ϵ which is varied from $\epsilon = 0.01$ to 0.3. The largest value of ϵ is an adsorbate whose spatial extent covers 30% of the device. The results in Table 9 show that mass, position, and variance are all recovered accurately, with the error in Equation 49 growing as ϵ increases. This rise in error is expected because Equation 49 is based on a Taylor expansion about the adsorbate's center-of-mass. It is striking that, despite this asymptotic approach, Equation 49 is robust in recovering these low-

Width, ϵ	Mass	Center-of-Mass	Variance	Skewness
0.01	7.50×10^{-3} (0.00 %)	0.704 (0.00 %)	7.64×10^{-6} (-0.03 %)	0.440 (0.41 %)
0.03	2.25×10^{-2} (0.00 %)	0.712 (0.00 %)	6.86×10^{-5} (-0.23 %)	0.429 (-2.20 %)
0.07	5.25×10^{-2} (0.02 %)	0.729 (-0.01 %)	3.70×10^{-4} (-1.10 %)	0.335 (-24.0 %)
0.1	7.51×10^{-2} (0.11 %)	0.741 (-0.04 %)	7.50×10^{-4} (-1.80 %)	0.178 (-59.0 %)
0.3	0.217 (-3.60 %)	0.837 (1.40 %)	5.67×10^{-3} (-18.0 %)	0.706 (61.0 %)

Table 9. Recovery of skewed distribution with small extents. Accuracy of Equation 49 in recovery of the moments of a specified mass distribution (illustrated in Figure 69). Recovered moments are reported; percentage error relative to the exact result is in parentheses. Frequency shift data for the lowest four flexural modes of a cantilever beam is generated using Equation 47. Recovered mass (second column) is scaled by $\mu_0 L$; see Figure 69.

order moments, even for the largest value of $\epsilon = 0.3$. Skewness is also captured accurately for the smallest, ϵ, but fails for larger spatial extents. This decrease in accuracy with increasing adsorbate size is enhanced for the higher-order moments, as expected due to the asymptotic nature of Equation 49. Nonetheless, the reduction in accuracy can be overcome by using more beam modes, which is now explored.

Results for the moments as a function of the number of modes are given in Figure 70 and Figure 71. Figure 70 shows the first two moments of the mass distributions recovered using Equation 49 and those obtained directly from the distribution (the exact values); Figure 71 gives corresponding results for the variance and skewness. Critically, determination of the p^{th} moment using Equation 49 requires use of at least p mechanical modes. As such, results for the higher-order moments do not exist when a smaller number of modes are used; see Figure 71. While a minimum of three modes are required (in general) to unequivocally determine the position of an adsorbate using a cantilever, results for two modes are also shown for completeness; any ambiguity in the solution is removed here because the exact solution is known.

Recovered moments for the narrowest mass distribution ($\epsilon = 0.01$) are given in the first row of Figure 70 and Figure 71. Note that as the number of modes is increased, the moments determined using Equation 49 systematically approach their exact values. Indeed, accurate results for all the presented moments are found regardless of the number of modes used. This is because the spatial extent of the adsorbate is only 1% of the total length of the device, ensuring that the Taylor expansion in Equation 47 converges rapidly. Equation 49 is thus ideally suited to measuring the moments of adsorbates whose sizes are much smaller than that of the device – this is particularly advantageous when exploring the ultimate limits

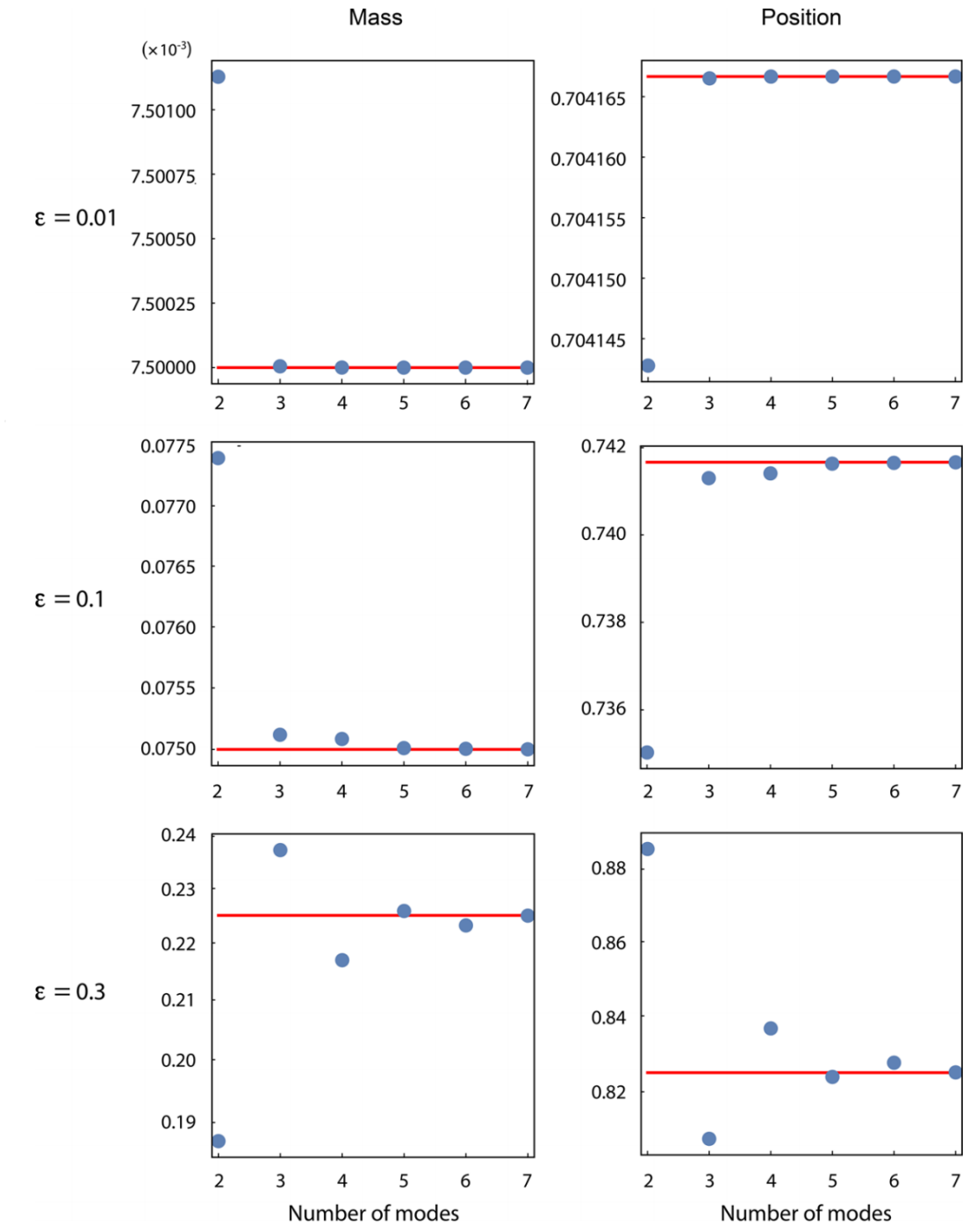


Figure 70. Recovery of mass and position of simulated adsorbate. Adsorbed mass and position of sample adsorbate in Figure 69 is recovered using Equation 49 as a function of number of modes and spatial extent of adsorbate, $\epsilon = 0.01, 0.1, 0.3$. Exact solutions are shown as red horizontal lines. Mass distribution is given in Figure 69.

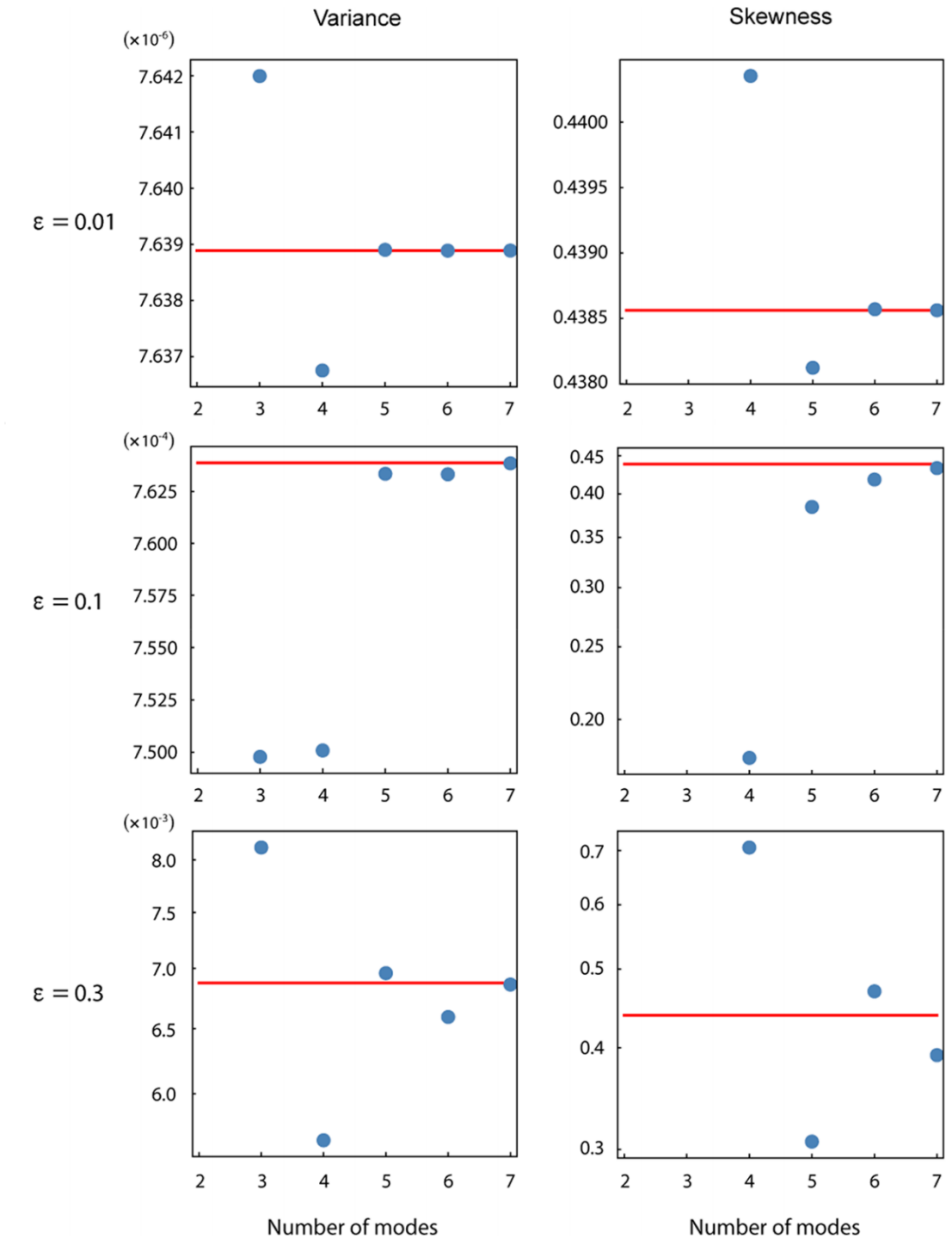


Figure 71. Recovery of variance and skewness of simulated adsorbate. As for Figure 70, but for the variance and skewness of the mass distribution.

of the sensitivity of the technique [1]. We remind the reader that inertial imaging is independent of any wavelength-dependent phenomena – this feature is illustrated here with

the adsorbate size being one to two orders-of-magnitude smaller than spatial wavelengths of the beam modes.

The second and third rows of Figure 70 and Figure 71 show that rates of convergence to the exact solutions decrease with increasing ϵ, i.e., more modes are required. For the largest spatial extent, $\epsilon = 0.3$, the moments recovered using Equation 49 oscillate about their exact values and even using 7 modes does not achieve convergence in highest-order moment, i.e., the skewness. While increasing the number of modes will in principle lead to convergence (see above), a large number of modes may be required and truncation error may provide a limitation in practice (see discussion above). This limitation is in addition to obvious practical issues involved in measuring many modes of a real device. This shows that the number of modes used in inertial imaging should be systematically increased until convergence is achieved in the required moments. If convergence does not occur using the available modes, the recovered moments may not be accurate. This is expected to be an issue only for adsorbates whose spatial extent is a large fraction of the device size. In Supporting Information of Ref. [2], we present an approach to estimate the number of modes needed as a function of the geometric moment to be calculated and the spatial extent of the analyte.

Experimental data. We now reanalyze the experimental data shown in Figure 68 and reported in Ref. [1] for liquid droplet arrays deposited on a microcantilever using dip pen lithography. Supplementary Tables 1 and 2 of Ref. [1] report numerical data for the measured resonant frequencies of the lowest four flexural modes of the cantilever, and their uncertainties (based on a 95% confidence interval). Six cantilever/droplet configurations are reported: the unloaded cantilever and 5 droplet arrays. In Ref. [1], linear superposition theory (discussed above) was used to recover the mass, position, variance, and skewness of the

droplets arrays, with the measured frequency shift noise used to determine uncertainties in the measured moments. This published (linear superposition theory) data is now compared to the same moments recovered with the nonlinear formula, Equation 4-4.

Table 10 presents numerical results for the measured variances of the droplet arrays using Equation 49, and those obtained using the linear superposition theory. We refrain from presenting results for the mass and position since they exhibit similar agreement to the data in Table 10. This comparison shows that numerical values for the recovered variances are nearly identical using these two complementary approaches. Even more striking is the uncertainties in the variances due to the measured frequency noise of each mode: the two theoretical frameworks give virtually identical results. While there are some minor differences especially for the droplet distributions of greater spatial extent, this comparison

Droplet array	Linear superposition	Equation 49	
		4 modes	3 modes
2 rows	5.35×10^{-4} (±38%)	5.32×10^{-4} (±34%)	5.33×10^{-4} (±31%)
3 rows	1.60×10^{-3} (±6.9%)	1.56×10^{-3} (±6.5%)	1.48×10^{-3} (±4.8%)
4 rows	2.92×10^{-3} (±1.4%)	2.79×10^{-3} (±1.2%)	2.81×10^{-3} (±0.74%)
5 rows	4.33×10^{-3} (±0.39%)	3.87×10^{-3} (±0.36%)	4.12×10^{-3} (±0.34%)
Asymmetric rows	5.21×10^{-3} (±1.5%)	4.67×10^{-3} (±1.3%)	4.23×10^{-3} (±1.1%)

Table 10. Comparison to linear superposition theory. Measured normalized variance of the droplet arrays of Ref. [1]; distance is again scaled by the cantilever length, *L*. Results obtained using linear superposition theory of Ref. [1] and Equation 49. Listed uncertainties in the parentheses specify a 95% confidence interval and are determined by the uncertainty in the measured resonant frequencies. The first four modes of the cantilever are used in the linear superposition theory.

shows that Equation 49 can be used with confidence to recover the spatial extent (variance) of an adsorbate. This finding is also consistent with above observations regarding the applicability of Equation 49 to adsorbates of finite extent – while Equation 49 is derived using an asymptotic Taylor expansion, it is expected to hold for adsorbates of arbitrary size provided a sufficient number of modes are used to verify convergence.

Experimental skewness data. Since more than 4 modes were not measured in Ref. [1], an assessment of the convergence of the results for skewness with mode number cannot be made. However, Figure 71 clearly shows that increasing the spatial extent of the adsorbate reduces the rate of convergence with increasing mode number. Thus, results for the droplet array consisting of 2 rows are expected to be more accurate than other arrays. We find nearly identical results using linear superposition and Equation 49, with their results differing by less than 1% (data not shown). This agreement decreases with increasing spatial extent of the adsorbate. For the asymmetric droplet array, linear superposition theory gives a value of -0.453 [1] whereas the optically measured value is -0.537; Equation 49 predicts a value of -0.695, which is consistent with these results. We refrain from analyzing this data further because its convergence, as a function of mode number, cannot be assessed.

The only inertial imaging measurements to date were reported in Ref. [1]. Importantly, these were performed passively by measuring the thermal noise spectra of the cantilevers over a finite bandwidth (to 1 MHz) – hence only 4 modes were collected. As discussed, analysis of these measurements (using the linear superposition theory of Ref. [1]) required significant computation because a measurement zone needed to be computed. The present theory eliminates this requirement, simplifying analysis and ensuring unequivocal measurement interpretation – uncertainty due to the measurement zone is eliminated. Use of

this theory in future measurements, e.g., using active drive and a larger bandwidth to overcome the above-stated limitations, will enable the full potential of inertial imaging to be experimentally realised.

4.4 Conclusion

In this chapter, we have presented simple nonlinear formulas for measuring the spatial distribution of a mass adsorbed to a nanomechanical resonator. Formulas for both one- and two-dimensional devices have been reported, which facilitates inertial imaging measurements using a range of current NEMS devices. The linear superposition theory of Ref. [15] and the nonlinear formulas in Equations 49 and 51 formally represent a set of inverse relations in the limit of small adsorbate size. They produce identical numerical results in this limit. The presented nonlinear formulas can be implemented trivially using standard packages such as Mathematica and Matlab, and simply involve the use of a root-finding procedure, e.g., Newton's method. These formulas are expected to be advantageous in the application of inertial imaging in practice. Wolfram CDF Player™ apps that implement the nonlinear formulas for cantilever and doubly-clamped beams are available from the authors; details are provided in the Supporting Information [16].

4.4 Bibliography

1. Hanay, M.S., et al., *Inertial imaging with nanomechanical systems.* Nature nanotechnology, 2015. **10**(4): p. 339-344.
2. Sader, J.E., et al., *Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation.* Nano Letters, 2018. **18**(3): p. 1608-1614.
3. Ekinci, K.L., X.M.H. Huang, and M.L. Roukes, *Ultrasensitive nanoelectromechanical mass detection.* Applied Physics Letters, 2004. **84**(22): p. 4469-4471.
4. Ilic, B., et al., *Attogram detection using nanoelectromechanical oscillators.* Journal of Applied Physics, 2004. **95**(7): p. 3694-3703.

5. Yang, Y.T., et al., *Zeptogram-scale nanomechanical mass sensing.* Nano Letters, 2006. **6**(4): p. 583-586.
6. Li, M., H.X. Tang, and M.L. Roukes, *Ultra-sensitive NEMS-based cantilevers for sensing, scanned probe and very high-frequency applications.* Nature Nanotechnology, 2007. **2**(2): p. 114-120.
7. Gil-Santos, E., et al., *Nanomechanical mass sensing and stiffness spectrometry based on two-dimensional vibrations of resonant nanowires.* Nature Nanotechnology, 2010. **5**(9): p. 641-5.
8. Burg, T.P., et al., *Weighing of biomolecules, single cells and single nanoparticles in fluid.* Nature, 2007. **446**(7139): p. 1066-1069.
9. Schmid, S., et al., *Real-time single airborne nanoparticle detection with nanomechanical resonant filter-fiber.* Scientific Reports, 2013. **3**: p. 1288.
10. Naik, A.K., et al., *Towards single-molecule nanomechanical mass spectrometry.* Nature Nanotechnology, 2009. **4**(7): p. 445-450.
11. Hanay, M.S., et al., *Single-protein nanomechanical mass spectrometry in real time.* Nature Nanotechnology, 2012. **7**(9): p. 602-608.
12. Dohn, S., et al., *Position and mass determination of multiple particles using cantilever based mass sensors.* Applied Physics Letters, 2010. **97**(4): p. 044103.
13. Jensen, K., K. Kim, and A. Zettl, *An atomic-resolution nanomechanical mass sensor.* Nature Nanotechnology, 2008. **3**(9): p. 533-7.
14. Chiu, H.Y., et al., *Atomic-scale mass sensing using carbon nanotube resonators.* Nano Letters, 2008. **8**(12): p. 4342-6.
15. Lassagne, B., et al., *Ultrasensitive mass sensing with a nanotube electromechanical resonator.* Nano Letters, 2008. **8**(11): p. 3735-8.
16. Chaste, J., et al., *A nanomechanical mass sensor with yoctogram resolution.* Nature Nanotechnology, 2012. **7**(5): p. 300-303.

Chapter 5

RECONSTRUCTION OF AN ADSORBATE'S IMAGE

Inertial imaging with nanomechanical resonators uses the measured frequency shifts of multiple modes induced by the physisorption of an analyte to calculate its mass *moments* [1, 2]. These moments include the analyte's total mass, its center-of-mass, as well as higher moments such as the variance (average size), skewness, etc. Subsequently, the adsorbate's mass density distribution, or *image*, is deduced from these fitted moments.

In this chapter, we examine a series of techniques that provide *direct* access to the adsorbate's image. Obtaining this image might be the primary motivation of an experiment; for example, it may be useful to ascertain structural features of individual molecules that are not ordinarily accessible in proteomics or native mass spectrometry experiments. In this case, intermediate calculations involving a finite number of fitted moments as well as their associated errors could only negatively impact the desired result.

At first glance, the problem of deducing the adsorbate's image appears intractable. Intuitively, the image might be arbitrarily complicated, while we are limited to a finite amount of data points. However, we find that we can use a technique called regularization that uses additional information known about the adsorbate and the experiment used to deliver it to the device in order to reduce the complexity of the problem and find a reasonable solution. Strikingly, we find that the mass moments of the image recovered directly from the data with this approach are more accurate than if the mass moments were fit to the data alone via previously introduced methods [1, 2].

5.1 Introduction

Thus far, two methods have been developed to calculate the mass moments of an analyte that has physisorbed to a nanomechanical resonator. One approach to these calculations involves creating linear superpositions of the squared mode shapes to approximate constant, linear, parabolic, etc., functions, which allows the mass moments to be calculated as linear combinations of the measured frequency shifts [1]. A complementary approach, introduced in Chapter 4 and [2], instead represents the squared mode shapes as a truncated series expansion, allowing the mass moments to be calculated from the fractional frequency shifts using a nonlinear fit. Both techniques deduce the adsorbate's image from these fitted moments.

Obtaining the adsorbate's image is the primary goal of this chapter; mass moments can then be calculated from this image as a secondary objective for quantitative purposes (for example, calculating the total mass of the adsorbate and comparing it to the value obtained from a traditional mass spectrometer). Reconstructing the unknown image from the data can be categorized as a type of inverse problem; a schematic is shown in Figure 72. We will find that this inverse problem is ill-posed: if a finite amount of frequency-shift data is available, there will always be an infinite number of valid mass density distributions that could have produced that data.

Inverse problems that are ill-posed arise in many fields, especially in situations where one wishes to determine the unknown input or internal structure of a physical system giving rise to a measured output signal. Some prominent examples include computed tomography,

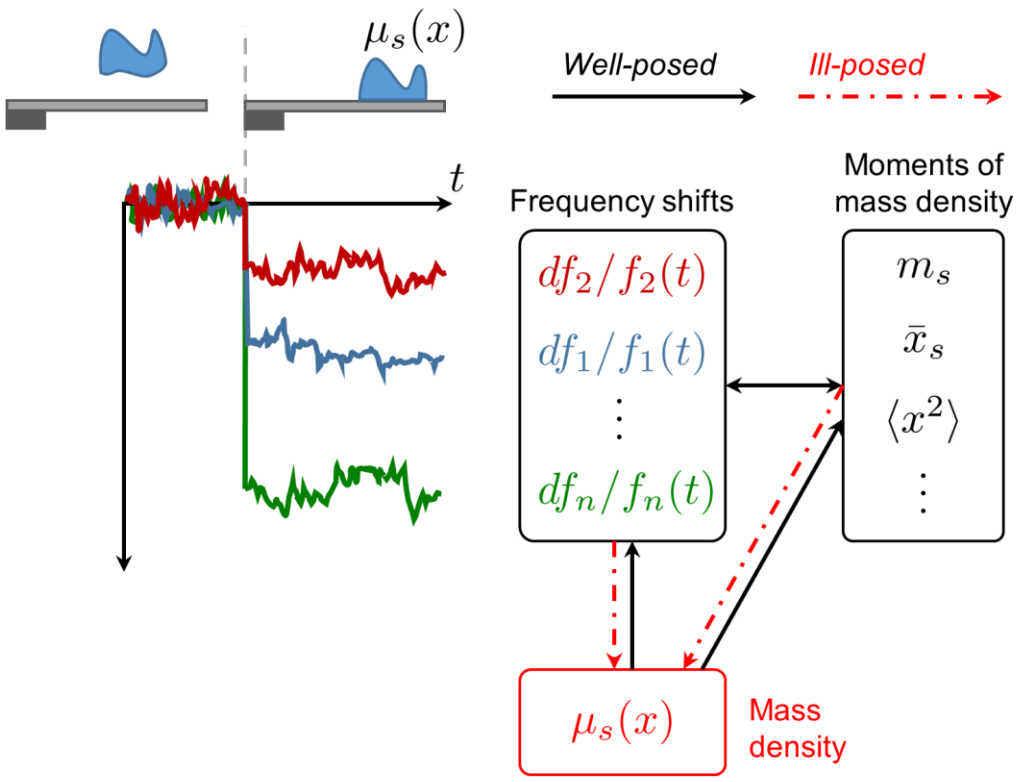


Figure 72. Reconstructing shape information about an adsorbate. A particle adhering to a nano- or micromechanical device induces a series of frequency shifts for each resonant mode of the device. These frequency shifts can be used to deduce information about the shape of the adsorbate in terms of the mass moments $m_s, \bar{x}_s, \langle x^2 \rangle$, ... in a well-posed manner. Recovering the underlying distribution $\mu_s(x)$ from these calculated moments, or directly from the frequency shifts, is an ill-posed problem.

where the density inside a body is reconstructed from the recorded intensity of X-ray beams at different angles, solving the diffusion or heat equation backwards in time to determine the sources, as well as a variety of inverse problems in acoustics, optics, electrostatics, geophysics, digital image recovery, and others [4-8].

For any ill-posed problem, including the mass density reconstruction problem considered in this chapter, additional information is needed beyond the data to restrict the solution space to a single valid solution, or at least a limited number of possibilities. Certain assumptions can be included to constrain or reduce the complexity of the problem if additional information is known about the adsorbate and the nature of how it is delivered to

the device. For example, if it is known that the experiment delivers one, or at most two, distinct particles to the device within a single measurement interval, with each particle having a limited spatial extent, then we find that we can distinguish between these possibilities as well as recover the first few mass moments with only a small number of measured modes.

If frequency-shift information is measured from a larger number of modes, we find that more details about the adsorbate's image can be reliably reconstructed. This requires the use of additional criteria such as minimizing the complexity, or maximizing the smoothness or disorder of the adsorbate, which can be achieved via the process of regularization. The physical motivation for such criteria are that, while proteins come in a large variety of shapes such as globular, fibrous, and disordered, they generally fold into the lowest energy conformation [9]. Additionally, we would like to avoid nonphysical solutions (such as singularities) when a simpler shape fits the data just as well, so we would prefer to bias the solution toward a realistic physical conformation using a mathematical framework that achieves that outcome.

We consider several variants of regularization: zeroth-order or L_2 regularization, first-order or Tikhonov regularization, and maximum entropy regularization. The first two correspond to minimizing the total mass and overall first derivative of the image. Each variant relies on the choice of a regularization parameter λ that dictates a tradeoff between fitting the solution to the data and optimizing the smoothness of the solution. Zeroth- and first-order regularization are found to be extremely sensitive to λ, the ideal choice of which cannot be known *a priori*. In contrast, maximum entropy is found to be robust to the choice of λ, and it is demonstrated for a set of simulated density distributions for which excellent

agreement is obtained. This method is then used experimentally to recover the image of droplets placed on a microcantilever with high fidelity.

5.2 Theory Overview

From Chapter 4, the fractional frequency shift of the *n*th flexural mode induced by a thin adsorbate is given by

$$\frac{\delta f_n}{f_n} = -\frac{1}{2M_{\text{device}}} \int_0^L \mu_s(x)\phi_n^2(x)\, dx. \tag{53}$$

where $\phi_n(x)$ are the mode shapes normalized such that $\int_0^L \phi_m(x)\phi_n(x)\, dx = \delta_{mn}$. In this chapter, we will normalize the frequency-shift data as well as the mass density distribution for brevity:

$$c_n = \int_0^1 \mu(x)\phi_n^2(x)\, dx. \tag{54}$$

The problem of finding $\mu(x)$ for the case when $\phi_n^2(x)$ and $c_n, 1 \ldots N$ are known is termed the generalized moment problem [10]. It is a generalized form of the classical moment problem, in which the data is obtained from a series of monomial functions, i.e.

$$c_n = \int_0^1 \mu(x)x^n\, dx, \tag{55}$$

and since the integral is performed over a finite region, the problem is more specifically classified as the Hausdorff moment problem [11]. Both the classical moment problem and the generalized moment problem are severely ill-posed. More formally, the solution is not guaranteed to exist; if the solution exists, it is not unique unless an infinite amount of data is

known ($n = 1 \ldots \infty$); and finally, even when the solution is uniquely specified, it can change arbitrarily with small perturbations in the data [12].

The second two points are problematic given that experimental data is always finite and embedded within the inevitable noise associated with the measurements. Additionally, both the generalized moment problem and the Hausdorff moment problem are equally ill-posed. In other words, in seeking a means to reconstruct the unknown mass density $\mu(x)$, it is not helpful to first construct the mass moments, which could be reframed as transforming the generalized moment problem (Equation 54) into the Hausdorff moment problem (Equation 55). In fact, information may be lost in doing so due to errors associated with truncation or approximation.

In this chapter, several methods will be proposed to regularize the ill-posed problem; that is, to make the problem well-posed so that the solution is guaranteed to be unique and vary continuously with the data. In practice, this involves explicitly adding a constraint or smoothness condition to the problem *a priori*, i.e., beyond the data itself. Although this may initially seem less preferable to solutions that involve no *a priori* knowledge, it is important to note that several implicit assumptions have already been made in the derivation of Equation 1, namely, knowledge of the device mode shapes $\phi_n(x)$ as well as the physical characteristics of the adsorbate (thin and flexible). Hence, such additional constraints may naturally arise from detailed physical considerations the solution must always satisfy. A violation of these assumptions will lead to incorrect results; for example, if the adsorbate is not thin relative to the device motion, this will add a significant rotational inertia component that can contribute significantly to the observed frequency shift. Additional assumptions that have previously been applied include the point-particle approximation as well as the

assumption that only one such point particle lands within the timeframe of the measurement [13, 14]. Therefore, the use of *a priori* assumptions has typically been considered a reasonable aspect of the interpretation of the data.

As there are an infinite number of valid solutions to the original ill-posed problem, there exist an infinite variety of methods to make the problem well-posed. This chapter will consider a few canonical methods, including zeroth-order regularization [15], first-order Tikhonov regularization [15], and the method of maximum entropy [16]. Additional prominent methods discussed in the literature include the Backus-Gilbert method [8, 15], the kernel density function method [17], as well as the method of adaptive splines [18]. These methods may be worth exploring in future work, but they will not be explored here for this specific application where the moment data is produced from the frequency shifts arising from the physical dynamics of an Euler-Bernoulli beam.

5.3 Discrete Density Reconstruction

As previously noted, reconstructing the entire mass density function $\mu(x)$ from a limited number of data points is an impossible task. We can at least reduce the dimensionality of the problem from an infinite number of degrees of freedom to a limited number by replacing the integration with a finite sum:

$$c_n \approx \sum_i \mu(x_i)\phi_n^2(x_i)\,\Delta x_i. \tag{56}$$

Assuming $\mu(x)$ and $\phi_n^2(x)$ do not vary too much between the points and the points are sufficiently dense, the sum in Equation 56 accurately approximates the integral in Equation 54. Rewriting $m_i \equiv \mu(x_i)\Delta x_i$, the underlying density distribution leading to the sum in Equation 56 can be represented as:

$$\mu(x) = \sum_i m_i \delta(x - x_i), \tag{57}$$

where $\delta(x)$ is the Dirac delta function. This representation for the mass density distribution is equivalent to one used to calculate the mass and position of multiple distinct point particles on a microcantilever using multiple modes [19]; in this section, Equation 57 will be used not only to determine the minimum number of modes needed to distinguish between one or two separate particles in a similar fashion, but also in a novel way: to recover mass moments of a single, isolated distribution of limited spatial extents.

To adapt Equation 57 for a small number of modes, it can be adjusted so that there are an equal number of unknown parameters as data points N:

$$\mu(x) = \sum_{i=1}^{N/2} m_i \delta(x - x_i). \tag{58a}$$

for N even, and

$$\mu(x) = m_1[\delta(x - x_0) + \delta(x - x_1)] + \sum_{i=2}^{(N-1)/2} m_i \delta(x - x_i). \tag{58b}$$

for N odd. Substituting into Equation 54, we obtain (for N even):

$$c_n = \sum_{i=1}^{N/2} m_i \phi_n^2(x_i), \tag{59}$$

and a similar formula for N odd; in particular, for $N = 3$:

$$c_n = m_1[\phi_n^2(x_1) + \phi_n^2(x_2)]. \tag{60}$$

In this way, the underdetermined inverse problem has been replaced with a series of nonlinear equations with the same number of equations as unknown parameters. In the minimal case $N = 2$, the formula reduces to the point-mass approximation formula [2]:

$$c_n = -2\frac{\delta f_n}{f_n} = \frac{m_{\text{added}}}{M_{\text{device}}}\phi_n^2(x), \tag{61}$$

which, in the case of a doubly-clamped beam, allows for the recovery of the mass and position of the adsorbate using only two modes [13]. Additionally, we can see that, if the fit for unknown parameters is successful, with three frequency-shift measurements, the underlying mass distribution $\mu(x)$ contains enough information to specify the mass, position, and spatial variance (but not asymmetry) of the adsorbate's mass density. With four

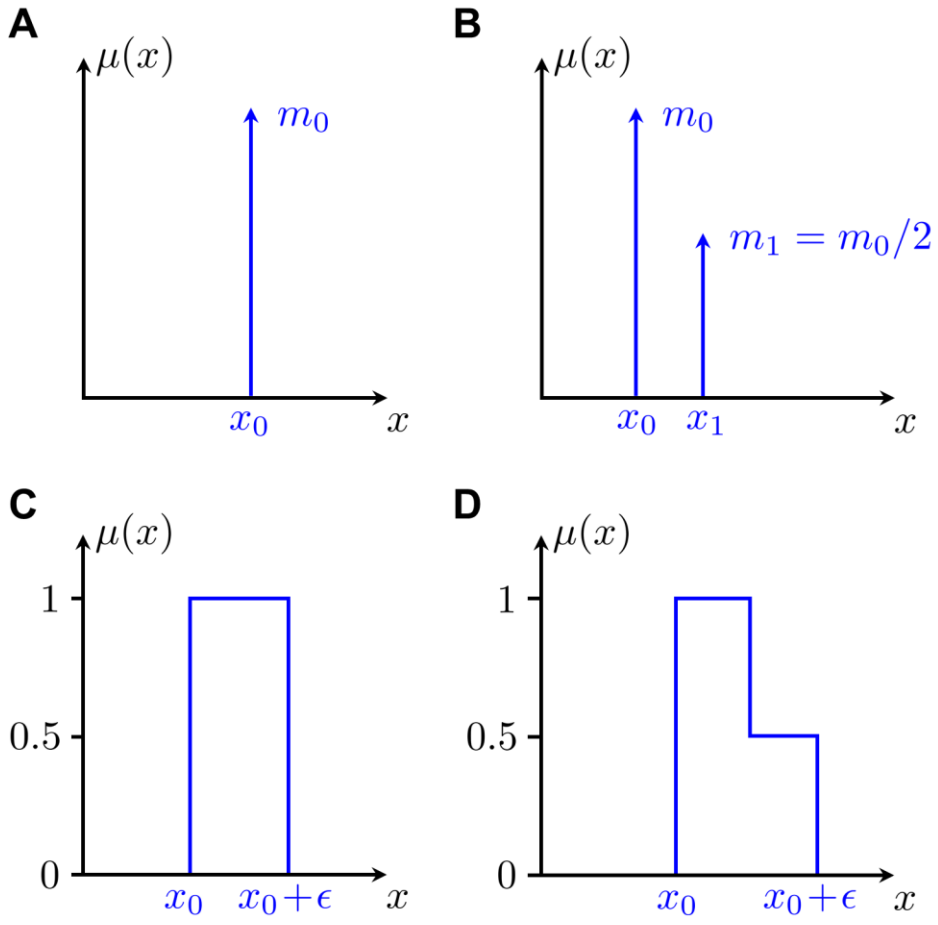


Figure 73. Test distributions used to simulate frequency-shift data. Synthetic frequency-shift data was calculated using these test distributions, which represent **(A)** a single point particle $\mu(x) = m_0\delta(x - x_0)$**, (B)** two point particles, **(C)** a uniform distribution beginning at x_0 and having width ϵ, and **(D)** a piecewise uniform distribution having total extents ϵ.

frequency shifts, $\mu(x)$ is also able to specify the adsorbate's asymmetry, since it consists of two point masses of different weights.

This approach was taken using a set of frequency-shift data simulated using Equation 53, based on a series of test distributions. These test distributions, shown in Figure 73, include a single point particle, two point particles, a uniform distribution, and a skewed distribution. This represents a range of simplified distributions that approximate what might be encountered in a measurement. In particular, while the experiment may be designed to deposit one molecule at a time to the sensor, a few molecules may arrive within the measurement window, and it would be beneficial to be able to distinguish between these situations. If it can be readily established that only a single analyte is detected, this improves the ability to distinguish between more complex distributions (e.g. type C vs type D in Figure 73). We do not include the possibility of three or more separate particles landing at arbitrary locations; more modes would be needed to distinguish between possible scenarios.

Doubly-clamped beam (three frequency shifts). In brief, it was found that for a doubly-clamped beam, three frequency shifts are sufficient to reliably distinguish between a single point particle, two point particles, and a uniform or skewed distribution with limited spatial extents ($\epsilon \leq 0.01$) using Equation 60. The fitting was performed using Mathematica's `FindRoot` routine with the initial positions for the fitting chosen arbitrarily, and then adjusted to a new arbitrary position only if the fitting routine produced a clear error. In the case of a single point particle, the recovered mass and position are accurate to machine precision. For the case of two point particles, the fitting produces two distinct positions, thus correctly classifying the distribution as type B. However, the positions obtained are sometimes wildly inaccurate depending on the initial position choice used for the fitting.

For a uniform distribution with limited spatial extents, the recovered mass, position, and variance are accurate to within 0.02%. For the case of a skewed distribution with limited spatial extents, the recovered mass, position, and variance are also highly accurate, and the accuracy improves as the extents approached zero. A summary of the results of the fitting for the skewed distribution are shown in Table 11.

It should be mentioned that no method could possibly distinguish between a single point particle and two particles landing on the mirror positions of a doubly-clamped beam $0.5 - \Delta x$ and $0.5 + \Delta x$ due to the symmetry of the device and its responsivity.

Cantilever (three frequency shifts). For a cantilever, the mass and position of a single point particle can reliably be recovered if $0.16 < x_0 < 0.58$ or $0.79 < x_0 < 1$. For $0.58 < x_0 < 0.79$, `FindRoot` mistakenly indicates a two-particle solution for some initial choices of the two positions. The frequency shifts generated with the two-particle solution are indistinguishable from those generated with the single-particle test distribution for the first

Width, ϵ	Mass	Center-of-Mass	Variance
0.0001	7.50×10^{-5} (0.00 %)	0.300 (0.00 %)	7.64×10^{-10} (0.01%)
0.001	7.50×10^{-4} (0.00 %)	0.300 (0.00 %)	7.64×10^{-8} (0.06 %)
0.01	7.50×10^{-3} (0.00 %)	0.304 (0.00 %)	7.68×10^{-6} (0.57 %)
0.1	7.50×10^{-2} (0.066 %)	0.341 (-0.069 %)	7.52×10^{-4} (1.60 %)

Table 11. Recovery of skewed distribution with small extents (three modes, doubly-clamped beam). Skew test distribution (type D shown in Figure 2) was used to generate frequency shifts for three modes for a doubly-clamped beam. For all distributions, $x_0 = 0.3$, but the spatial extent ϵ was varied. Recovered moments are reported; percentage error is in parentheses.

three modes, so, from the results alone, it was not possible to tell that the identification of the class of distribution was performed in error. Therefore, for a cantilever, three modes are not sufficient to infer information about the distribution if it is not known *a priori* whether the distribution is localized. If it is known *a priori* that the distribution was localized (e.g. only a single adsorbate with limited spatial extents within the measurement window), the mass and position of the distribution can be fit using three modes with Equation 61 first, then `FindRoot` can be constrained to search for solutions within a small region surrounding that initial fit position using Equation 60. When done iteratively in this way, similar results were obtained for a skewed distribution with limited extents ($\epsilon \leq 0.01$) as shown in Table for the doubly-clamped beam.

Doubly-clamped beam (four frequency shifts). Equation 59 was used with $N = 4$ to attempt to distinguish between all four classes of test distributions shown in Figure 73 as well as reliably recover information such as total mass, position, and so on. It was found that this can always be done in a robust manner; however, in the case where the test distribution is two distinct particles at arbitrary locations, it is necessary to attempt many different guesses for the initial positions to obtain any fit at all; for most choices of the initial positions, no fit could be made. When a fit can be made, the position and mass of both particles can be recovered to machine precision. The weight of the second particle can also be varied from $m_0/10 < m_1 < 10m_0$ with the same result. For the case of a skewed distribution, the recovered distribution resembles the test distribution, as shown in Figure 74. The first four mass moments of the distribution can be recovered with excellent agreement as well (Table 12).

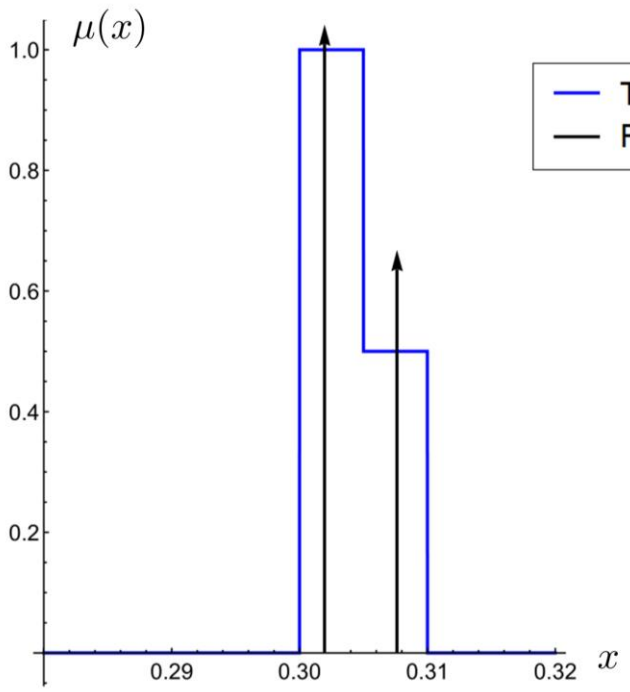


Figure 74. Discrete density reconstruction using four modes. Synthetic frequency-shift data was calculated using a skewed test distribution for four modes for a doubly-clamped beam, and the image was recovered by fitting to a discretized distribution, which consists of two delta functions of varied weight and location. The height of the arrows depicted for the delta functions are proportional to their weights. Excellent agreement to the test distribution was obtained.

Width, ϵ	Mass	Center-of-Mass	Variance	Skewness
0.01	7.50×10^{-3} (0.00 %)	0.304 (0.00 %)	7.64×10^{-6} (-0.03%)	0.448 (2.09 %)
0.03	2.25×10^{-2} (0.00 %)	0.313 (0.00 %)	6.86×10^{-5} (-0.25 %)	0.458 (4.33 %)
0.07	5.25×10^{-2} (-0.01 %)	0.329 (0.00 %)	3.70×10^{-4} (-1.23 %)	0.442 (0.75 %)
0.1	7.50×10^{-2} (-0.02 %)	0.342 (0.00 %)	7.44×10^{-4} (-2.62 %)	0.397 (-9.43 %)

Table 12. Recovery of mass moments for skewed distribution (four modes, doubly-clamped beam). Test distribution type D shown in Figure 2 (a skewed distribution) was used to generate frequency shifts for four modes for a doubly-clamped beam. For all distributions, $x_0 = 0.3$, but the spatial extent ϵ was varied. Recovered moments are reported; percentage error is in parentheses.

Cantilever (four frequency shifts). In contrast with the doubly-clamped beam, recovery of the four unknown parameters m_1, x_1, m_2, x_2 using Equation 59 is not guaranteed. Reliable fitting can no longer be achieved for any test distribution using `FindRoot`, but instead the fitting must be reformulated as a form of (non-negative) least squares optimization to achieve any degree of success:

$$\arg\min_{m_i, x_i} \sum_{n=1}^{4} \left(c_n - m_1 \phi_n^2(x_1) - m_2 \phi_n^2(x_2)\right)^2 \quad \text{subject to } m_i \geq 0. \tag{62}$$

After extensive empirical testing, the results obtained are similar to the case of the doubly-clamped beam with three modes: the fitting can reliably distinguish between a localized distribution (for example, a point particle or a distribution with arbitrary complexity, but with total width of $\epsilon < .01$) and two point particles at arbitrary locations. In the case of two point particles, fitting produces masses and locations substantially different from each other, but the fitted positions were often completely inaccurate. In the case of a single point particle

Width, ϵ	Mass	Center-of-Mass	Variance	Skewness
0.01	7.50×10^{-3} (0.00 %)	0.704 (0.00 %)	7.64×10^{-6} (-0.01%)	0.439 (0.17 %)
0.03	2.25×10^{-2} (0.00 %)	0.713 (0.00 %)	6.87×10^{-5} (-0.10 %)	0.434 (-1.05 %)
0.07	5.25×10^{-2} (0.01 %)	0.729 (0.00 %)	3.72×10^{-4} (-0.49 %)	0.305 (-11.39 %)
0.1	7.50×10^{-2} (0.06 %)	0.742 (-0.02 %)	7.57×10^{-4} (-0.87 %)	0.397 (-30.50 %)

Table 13. Recovery of mass moments for skewed distribution (four modes, cantilever). Test distribution type D shown in Figure 2 was used to generate frequency shifts for four modes for a cantilever. For all distributions, $x_0 = 0.7$, but the spatial extent ϵ was varied. Recovered moments are reported; percentage error is in parentheses.

test distribution or localized distribution, the fitting frequently returned $m_1 = 0$ for one of the masses, m_2 close to the correct value, and the x_2 close to the center of mass. The distribution can be further interrogated via `FindRoot` after localizing in the first pass least squares fitting by restricting the fitting to a small region around the first fitted position. Doing so in this iterative manner on the skewed test distribution produces results in excellent agreement up to skewness; results are summarized in Table 13.

More than four frequency shifts. A least-squares fit was performed for $N = 5$ for the unknown parameters m_1, m_2, x_0, x_1, x_2 for the skewed test distribution with width $\epsilon = 0.01$ for both a doubly-clamped beam and a cantilever. An exact fit using `FindRoot` was found to be impossible. In both cases, the recovered mass, center-of-mass, and variance are highly accurate, but both the skewness and kurtosis of the recovered distribution are very inaccurate. For the doubly-clamped beam, for example, with the distribution beginning at $x_0 = 0.3$, an error of 0.00%, 0.00%, –0.43% is obtained for the mass, center-of-mass, and variance, respectively, but a value of –3.46 was obtained for the skewness (compared with 0.439, expected), and 16.08 for the kurtosis (2.12, expected). The recovered distribution for the doubly-clamped beam, along with the initial distribution used to simulate the frequency-shift data, is shown in Figure 75.

It is clear from Figure 75 that the recovered distribution using least squares fitting applied to the discretized problem is not able to accurately reconstruct a complicated test distribution using five modes. This could be due to numerical accuracy or stability issues associated with fitting to data with five modes, or due to the fact that the least squares fitting found a local minimum among multiple distributions that are numerically indistinguishable from the correct one. A similar recovered image was obtained for the cantilever.

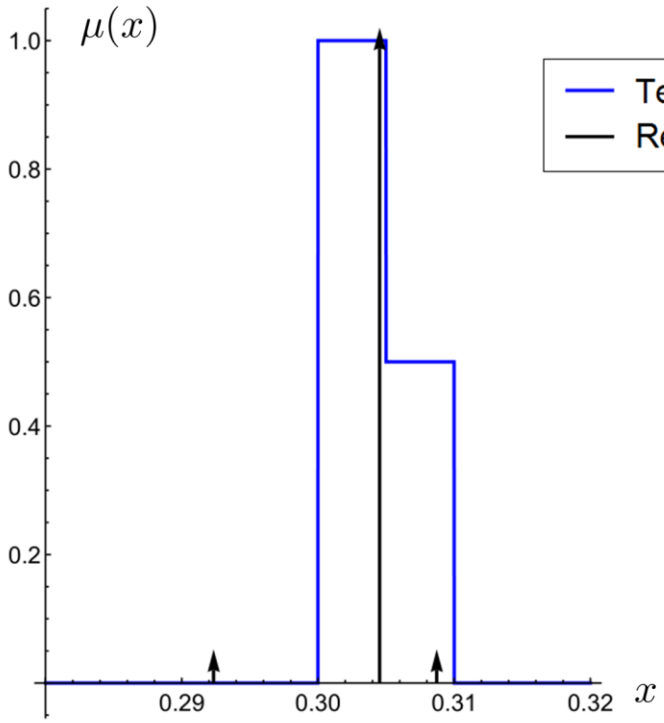


Figure 75. Discrete density reconstruction using five modes. Synthetic frequency-shift data was calculated using a skewed test distribution for five modes for a doubly-clamped beam, and the image was recovered by fitting to a discretized distribution, which consists of a series of delta functions of varied weight and location. The height of the arrows depicted for the recovered distribution are proportional to the weights associated with the corresponding delta functions.

5.4 Discrete Density Reconstruction with Regularization

As noted in Section 5.3, fitting more than four unknown parameters with more than four frequency shifts is generally found to be unreliable for a localized distribution. Although the discretized formula in Equation 59 converges to the integral in Equation 54 as $N \to \infty$, in practice, the fit cannot be done due to the complexity of fitting so many unknown parameters with a set of highly nonlinear equations. It is evident that there are too many valid solutions to a given set of frequency-shift data within the numerical accuracy available for calculation. This problem is exacerbated when dealing with noisy data – as the noise level

increases, so does the number of solutions that can reproduce the data with equal validity within the noise. Thus, it is necessary to impose some scheme outside the data itself to select one solution out of many. A standard approach to do so is known as regularization, and the application of this method to recovering an adsorbate's image will be discussed in this section.

Regularization involves a trade-off between optimizing the agreement between the solution and the data (denoted $\boldsymbol{\mathcal{A}}$), and optimizing the stability or smoothness of the solution (denoted $\boldsymbol{\mathcal{B}}$): since two optimization problems cannot be solved simultaneously, the approach optimizes the combined problem

$$\boldsymbol{\mathcal{A}} + \lambda \boldsymbol{\mathcal{B}}, \tag{63}$$

for some choice of $0 < \lambda < \infty$ that dictates the relative preference of $\boldsymbol{\mathcal{A}}$ or $\boldsymbol{\mathcal{B}}$; as $\lambda \to 0$, the problem approaches optimizing $\boldsymbol{\mathcal{B}}$ with $\boldsymbol{\mathcal{A}}$ as a constraint, and as $\lambda \to \infty$, the problem approaches optimizing $\boldsymbol{\mathcal{A}}$ with $\boldsymbol{\mathcal{B}}$ as a constraint [15]. Even if $\boldsymbol{\mathcal{A}}$ is degenerate in the sense of having multiple valid solutions, if $\boldsymbol{\mathcal{B}}$ is non-degenerate, the overall expression in Equation 63 will be too, thus there will be a unique solution to the overall problem [15].

As demonstrated in Section 5.3, fitting the solution to the data involving a large number of modes requires a least-squares optimization rather than an exact nonlinear fit, so a least-squares approach will be used for $\boldsymbol{\mathcal{A}}$ going forward. In choosing least-squares for $\boldsymbol{\mathcal{A}}$, a larger number of unknowns compared with data points can be used since the degeneracy is broken by the overall problem $\boldsymbol{\mathcal{A}} + \lambda \boldsymbol{\mathcal{B}}$. Therefore, for ease of implementation, rather than the discretization in Equation 58a and 58b using variable positions for each delta function, the mass density function will be evaluated over a large number of uniformly spaced points. Doing so allows us to rewrite the discretized frequency shift formula in Equation 56 as:

$$c_i = \sum_{k=1}^{M} \mu(x_k)\phi_i^2(x_k)\Delta x_k \equiv \sum_{k=1}^{M} \Phi_{ik}\mu_k, \qquad n = 1 \ldots N \tag{64}$$

where $\Phi_{ik} \equiv \phi_i^2(x_k)\Delta x_k$, $M \geq N$, and $x_k = x_0 + (k-1)\Delta x$ are preselected coordinates with x_0 and Δx chosen such that the distribution encompasses a region of limited spatial extents surrounding the estimated center-of-mass, as discussed in Section 5.3. (Recall that it is always possible to distinguish a localized distribution from two separate particles at arbitrary locations by using at least 3 modes for a doubly-clamped beam or 4 modes for a cantilever. If the distribution is localized, it is possible to obtain an accurate estimate of the center-of-mass.) Using this approach, we can formulate $\boldsymbol{\mathcal{A}}$ as follows:

$$\arg\min_{\mu_k} \sum_{i=1}^{N} \left(c_i - \sum_{k=1}^{M} \Phi_{ik}\mu_k, \right)^2 \text{ subject to } \mu_k \geq 0. \tag{65}$$

In the case of noisy data, a χ^2 estimate can be used [15]:

$$\arg\min_{\mu_k} \chi^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \left[c_i - \sum_{k=1}^{M} \Phi_{ik}\mu_k \right] S_{ij}^{-1} \left[c_j - \sum_{k=1}^{M} \Phi_{jk}\mu_k \right] \text{ subject to } \mu_k \geq 0, \tag{66}$$

where S_{ij} is the covariance matrix. In the case of non-correlated noise, off-diagonal entries are not used and the sum simplifies:

$$\arg\min_{\mu_k} \chi^2 = \sum_{i=1}^{N} \left[\frac{c_i - \sum_{k=1}^{M} \Phi_{ik}\mu_k}{\sigma_i} \right]^2 \text{ subject to } \mu_k \geq 0. \tag{67}$$

If $M \gg N$, the values of μ_k can be chosen to make Equation 66 go to zero. A more realistic outcome is $\chi^2 = N$, so the value of λ in the overall problem $\boldsymbol{\mathcal{A}} + \lambda\boldsymbol{\mathcal{B}}$ is often adjusted until this is achieved for the final solution [15].

Zeroth-order regularization. For the stability problem $\boldsymbol{\mathcal{B}}$, the simplest choice is to use the smallest $\mu(x)$ in the sense of:

$$\arg\min_{\mu_k} \sum_{k=1}^{M} \mu_k^2 \quad \text{subject to } \mu_k \geq 0. \tag{68}$$

The regularization problem using the above for $\boldsymbol{\mathcal{B}}$ is known as *zeroth-order regularization* [15] or L_2 regularization (in the sense of giving preference to solutions with smaller L_2 norm) [20]. Applied to our problem, this amounts to finding the density function with minimum total mass that satisfies the frequency-shift constraints to desired precision. It should be pointed out that there is a small distinction between Equation 68 and the typical regularization problem in that here the mass density function is required to be positive. This positivity constraint means that a closed-form solution is not available; however, the problem can still be solved numerically with relative ease due to its stability.

As a demonstration, the following zeroth-order regularization problem was solved in Mathematica for the case of the skewed distribution (type D in Figure 73) on a doubly-clamped beam starting at $x_0 = 0.3$ and having width $\epsilon = 0.01$:

$$\arg\min_{\mu_k} \sum_{i=1}^{N} \left(c_i - \sum_{k=1}^{M} \Phi_{ik}\mu_k, \right)^2 + \lambda \sum_{k=1}^{M} \mu_k^2 \text{ subject to } \mu_k \geq 0. \tag{69}$$

Regularization was then repeated in a more realistic way: the center-of-mass of the distribution was first estimated by fitting to a discrete distribution without regularization (Section 5.3). Then, $M = 21$ coordinates were chosen centered around this value and having small extents ($\epsilon = 0.02$). The choice of extents is fairly well-matched to the exact value of $\epsilon = 0.01$, but ϵ can easily be adjusted to "zoom in" on the distribution in an iterative manner. The regularization parameter λ was then adjusted, and the recovered distributions are shown in Figure 76 along with the recovered moments in Table 14. Evidently, too large of a choice of λ favors the smoothness criteria too strongly, creating a smooth function that does not

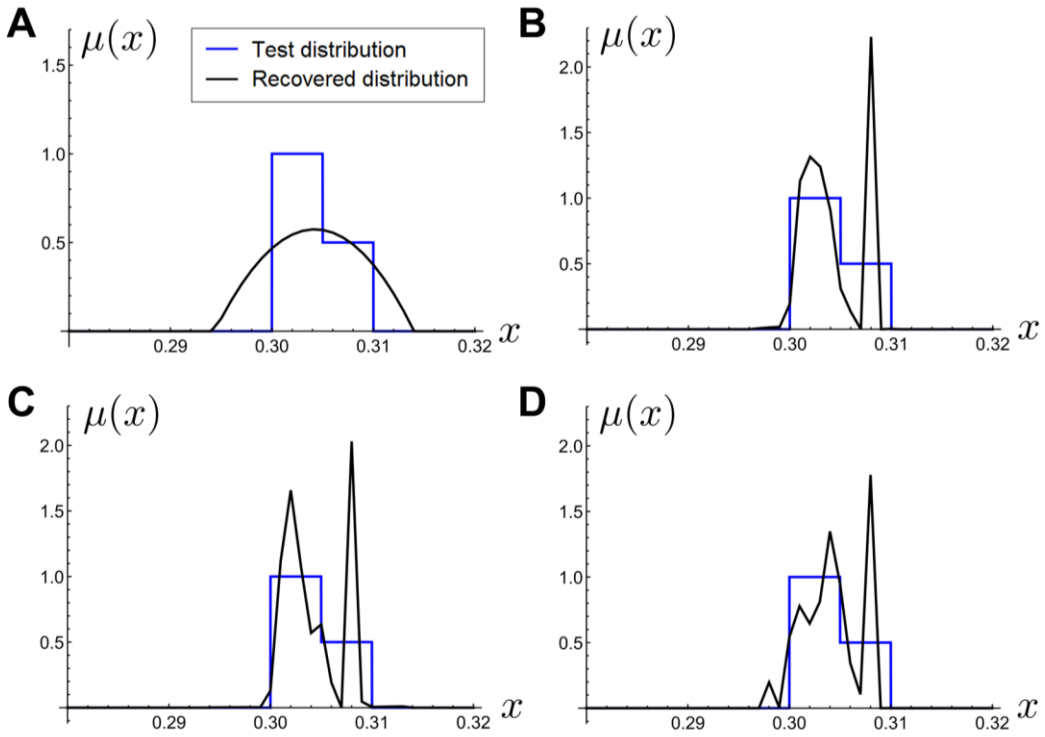


Figure 76. Zeroth-order regularization. Regularization was performed on the test distribution using $M = 21$ coordinates centered about the estimated center-of-mass obtained from the discretized problem without regularization. The recovered distributions are shown for various choices of the regularization parameter λ: **(A)** $\lambda = 10^{-3}$, **(B)** $\lambda = 10^{-8}$, **(C)** $\lambda = 10^{-10}$, **(D)** $\lambda = 10^{-18}$.

match the original data well. Decreasing λ to 10^{-8} or less was found to generate distributions with frequency shifts indistinguishable to the simulation data, but still led to a variety of outcomes. For $\lambda = 10^{-8}$ and $\lambda = 10^{-10}$, the recovered distributions visually match the original distribution quite well, and the recovered moments were accurate. For $\lambda = 10^{-18}$, the recovered distribution was irregular and also had moments that poorly matched the original distribution.

Evidently, employing $\mu(x)$ as a smoothness parameter enables recovery of a complicated distribution, but is not an ideal choice in that the outcome is highly dependent on the choice of regularization parameter λ. Additionally, amongst the variety of possible solutions, there is no clear reason for preferring the solution with the least mass—other than

Table 14. Recovery of mass moments using zeroth-order regularization. A skewed distribution with $x_0 = 0.3$ and $\epsilon = 0.1$ was used to generate frequency shifts for five modes for a doubly clamped beam. The distributions were recovered with different choices of regularization parameter. Recovered moments are reported; percentage error is in parentheses.

Regularization Parameter, λ	Mass	Center-of-Mass	Variance	Skewness
10^{-3}	7.51×10^{-2} (0.11 %)	0.304 (0.01 %)	1.92×10^{-5} (151%)	-0.01 (-102 %)
10^{-8}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	7.68×10^{-6} (0.50 %)	0.335 (-23.7 %)
10^{-10}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	7.64×10^{-6} (0.01 %)	0.415 (-5.40 %)
10^{-18}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	7.72×10^{-6} (1.03 %)	-0.09 (-120 %)

its simplicity in the formulation of the problem.

Tikhonov regularization. A generalization of zeroth-order regularization is known as linear or Tikhonov regularization, in which the stability problem $\boldsymbol{\mathcal{B}}$ uses a matrix $\boldsymbol{\Gamma}$ [15]:

$$\arg\min_{\mu_k} |\boldsymbol{c} - \boldsymbol{\Phi} \cdot \boldsymbol{\mu}|^2 + \lambda \ |\boldsymbol{\Gamma} \cdot \boldsymbol{\mu}|^2. \tag{70}$$

This generality allows for more sophisticated smoothness criteria, such as minimizing the first derivative (or higher) of the unknown function. In minimizing the first derivative, the regularization problem becomes:

$$\arg\min_{\mu_k} \sum_{i=1}^{N} \left(c_i - \sum_{k=1}^{M} \Phi_{ik} \mu_k, \right)^2 + \lambda \sum_{k=1}^{M-1} (\mu_k - \mu_{k+1})^2 \ \text{subject to} \ \mu_k \geq 0. \tag{71}$$

Similarly to the previous case, the non-negativity constraint $\mu_k \geq 0$ makes this problem nonlinear (in that sense, Equation 71 is only inspired by Tikhonov regularization). As a result, closed-form solutions are not available in contrast with most of the literature on the

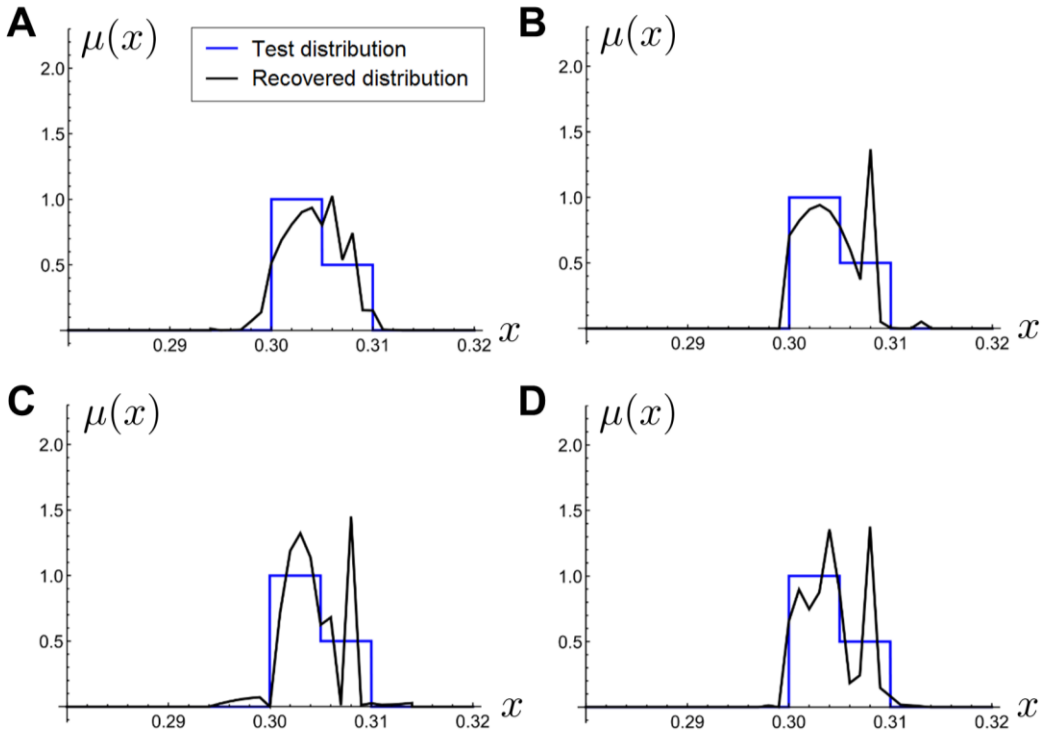


Figure 77. Tikhonov regularization. Regularization was performed similar to the zeroth-order approach, but using the first derivative of the function as the smoothness criteria. The recovered distributions are shown for various choices of the regularization parameter λ: **(A)** $\lambda = 10^{-3}$, **(B)** $\lambda = 10^{-8}$, **(C)** $\lambda = 10^{-10}$, **(D)** $\lambda = 10^{-18}$.

subject, but as before, it can be solved numerically in a stable manner. Reconstructed density functions from the same test distribution using Equation 71 as the optimization problem are shown in Figure 77. The results are marginally smoother than before, but exhibit a similarly undesirable sensitivity upon the regularization parameter λ (with recovered moments similarly becoming less accurate with $\lambda < 10^{-10}$).

Maximum entropy regularization (discrete case). As demonstrated for the case of Tikhonov regularization above with the non-negative mass constraint, computational solutions for regularization are not necessarily limited by linearity if the overall problem is numerically stable. Therefore, it seems reasonable to allow for a generally nonlinear smoothness criteria for $\boldsymbol{\mathcal{B}}$. One popular approach involves maximizing the *entropy* associated with the distribution (equivalently, minimizing the negative entropy) [15]:

$$\arg\min_{\mu_k} \sum_{i=1}^{N} \left(c_i - \sum_{k=1}^{M} \Phi_{ik}\mu_k, \right)^2 + \lambda \sum_{k=1}^{M-1} (\mu_k \log \mu_k)^2. \quad (72)$$

It should be noted that, while using a functional form with parallels to thermodynamics and information theory, other nonlinear functional forms for such as $-\log \mu$ or $-\sqrt{u}$ are known to produce similar results [15]. Thus, while it could be justified on philosophical or axiomatic grounds [21], the practical utility of the functional can be equally considered and should include the following physical constraints: the solution should be constant (i.e. unbiased) in the absence of data, sharp features are permitted among a smooth background, positivity must be ensured, and missing data such as $c_i, i > N$ should be smoothed over. Maximum entropy regularization features these constraints [15].

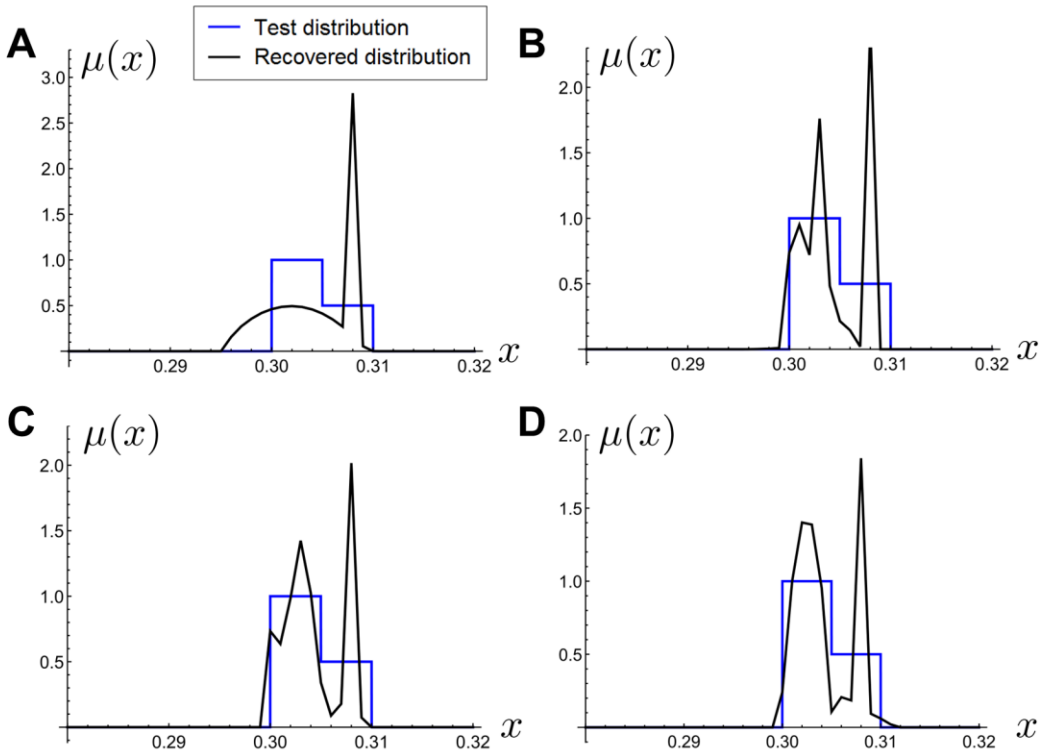


Figure 78. Discretized maximum entropy regularization. Regularization was performed using an entropy functional as the smoothness criteria. The recovered distributions are shown for various choices of the regularization parameter λ: **(A)** $\lambda = 10^{-18}$, **(B)** $\lambda = 10^{-21}$, **(C)** $\lambda = 10^{-24}$, **(D)** $\lambda = 10^{-30}$.

Table 15. Recovery of mass moments using discretized maximum entropy regularization. A skewed distribution with $x_0 = 0.3$ and $\epsilon = 0.1$ was used to generate frequency shifts for five modes for a doubly-clamped beam. The distributions were recovered with different choices of the regularization parameter. Recovered moments are reported; percentage error is in parentheses.

Regularization Parameter, λ	Mass	Center-of-Mass	Variance	Skewness
10^{-18}	7.51×10^{-2} (-0.07 %)	0.304 (0.00 %)	1.49×10^{-5} (94.5%)	-0.49 (-212 %)
10^{-21}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	8.73×10^{-6} (14.3 %)	0.216 (-50.7 %)
10^{-24}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	7.68×10^{-6} (0.59 %)	0.271 (-38.2 %)
10^{-30}	7.50×10^{-2} (0.00 %)	0.304 (0.00 %)	7.64×10^{-6} (-0.03 %)	0.458 (4.49 %)

When applied to the skewed test distribution, the discretized maximum entropy regularization recovers a distribution that resembles the original for any $\lambda < 10^{-20}$ as shown in Figure 78. As λ is reduced to zero, corresponding to maximizing the entropy with the frequency shifts as constraints, the recovered mass moments become increasingly accurate as shown in Table 15. This contrasts with the zeroth- and first-order regularization procedures, in which λ must be within a certain range of values for a successful outcome.

5.5 Maximum Entropy Regularization: Continuous Solution

In the previous section, three functionals were compared for regularizing the discretized inertial imaging problem: these minimize the total mass and the overall first derivative, and maximize the entropy, respectively. Among the three methods, the maximum entropy approach was found to produce the most consistent results for a broad choice of values for the regularization parameter λ, and the results improve as $\lambda \to 0$, thereby

eliminating the ambiguities associated with the choice of λ for the other approaches considered.

In this section, maximum entropy regularization will be applied to the original moment problem in Equation 54 without discretization. In addition to inheriting the practical benefits of the maximum entropy method in the previous section—an unbiased solution in the absence of data, sharp features are permitted, missing data (i.e., those that would be obtained from higher modes) is appropriately smoothed over, and a positive mass density is guaranteed—the continuous solution features a closed form solution in terms of the basis functions [16, 22-26]; in this case, these are the squared mode shapes $\phi_n^2(x)$.

Maximum entropy regularization seeks to maximize the entropy functional:

$$\max S \equiv -\int_0^1 \mu(x) \ln \mu(x)\, dx, \tag{73}$$

given the fractional frequency-shift data as constraints. Shannon suggested this integral as a continuous version of his information entropy [27], but it has been noted that it is not a proper generalization, due to not being dimensionally correct, among other objections that have been raised [28]. Unlike entropy, relative entropy or cross-entropy does generalize correctly for continuous probability densities [21, 28]. Under that approach, the goal is to minimize the relative entropy with respect to a bias, or prior distribution estimate $\Pi(x)$:

$$\min S_{\text{rel}} \equiv \int_0^1 \mu(x) \ln \frac{\mu(x)}{\Pi(x)}\, dx, \tag{74}$$

which becomes equivalent to maximizing entropy when choosing a uniform bias distribution.

Beyond the practical benefits enumerated above, justification for maximizing entropy, or minimizing cross-entropy, is frequently given in terms of intuitive arguments

about entropy and cross-entropy as information measures [29-31]. The resulting distribution $\mu(x)$ is the one that agrees with what is known, but minimally commits to missing information. Alternatively, $\mu(x)$ minimizes the amount of information necessary to change from a prior $\Pi(x)$ into the posterior $\mu(x)$. There remains some controversy to this approach; many find the argument plausible but lacking a sufficient level of proof for choosing to maximize entropy and not some other function [32, 33]. The most rigorous argument has been given by Shore and Johnson [21]. They argue that minimizing cross-entropy is the unique method for performing inductive inference when provided with constraint information in the form of expected values (i.e., moments) that is self-consistent. Self-consistency in this context is defined as the ability to yield a unique solution, regardless of the approach. In other words, the same result should be obtained regardless of the choice of coordinate system or whether one uses all experimental data simultaneously, or updates the solution $\mu(x)$ one mode at a time. For example, one can choose to solve for $\mu(x)$ given the frequency shift for the first mode, then use that $\mu(x)$ as the prior $\Pi(x)$ and solve for $\mu(x)$ given the frequency-shift data for the second mode, and so on.

The inclusion of the bias function $\Pi(x)$ is useful as it allows us to impose *a priori* knowledge about the distribution. For example, if the distribution can be determined to be a single, isolated particle of limited spatial extent, $\Pi(x)$ can be set to zero everywhere except in a small region surrounding the particle's estimated center-of-mass. In terms of the regularization procedure established in the previous section, we hope to solve

$$\arg\min_{\mu(x)} \sum_{n=1}^{N} \left(c_n - \int_0^1 \mu(x)\phi_n^2(x)\, dx, \right)^2 + \lambda \int_0^1 \mu(x) \ln \frac{\mu(x)}{\Pi(x)} dx \qquad (75)$$

for infinitesimal λ. This can be equivalently formulated by associating a Lagrange multiplier with each frequency-shift constraint [16]:

$$\arg\min_{\mu(x)} \sum_{n=1}^{N} \lambda_n \left(c_n - \int_0^1 \mu(x)\phi_n^2(x)\, dx, \right)^2 + \int_0^1 \mu(x) \ln \frac{\mu(x)}{\Pi(x)} dx. \tag{76}$$

The solution is obtained by performing functional variation of Equation 75 with respect to $\mu(x)$:

$$\mu(x) = \Pi(x) \exp\left(- \sum_{n=1}^{N} \lambda_n \Phi_n^2(x) \right), \tag{77}$$

where the λ_n's can be solved numerically to match the frequency-shift constraints:

$$c_n = \int_0^1 \Pi(x) \exp\left(- \sum_{n=1}^{N} \lambda_n \Phi_n^2(x) \right) \phi_n^2(x)\, dx\ . \tag{78}$$

If a solution to Equation 78 exists, it is unique and varies continuously with the data [16]. Thus, it represents a well-posed solution to our present problem of obtaining the spatial mass density distribution given the frequency-shift data arising from physisorption to an Euler-Bernoulli beam.

Application of maximum entropy to a skewed test distribution with no bias (corresponding to $\Pi(x) = 1$) is shown in Figure 79 for the case of four modes of a cantilever. Frequency shifts generated from the solution match those of the test distribution to machine precision. While some general resemblance to the test distribution was obtained, the solution also shows mass density near the base of the cantilever ($x < 0.1$). Evidently, the algorithm is impartial toward adding mass at this location, at which the device provides low responsivity. This erroneous result can be avoided only by including additional assumptions, which can be encoded precisely with the choice of the bias function $\Pi(x)$. These assumptions

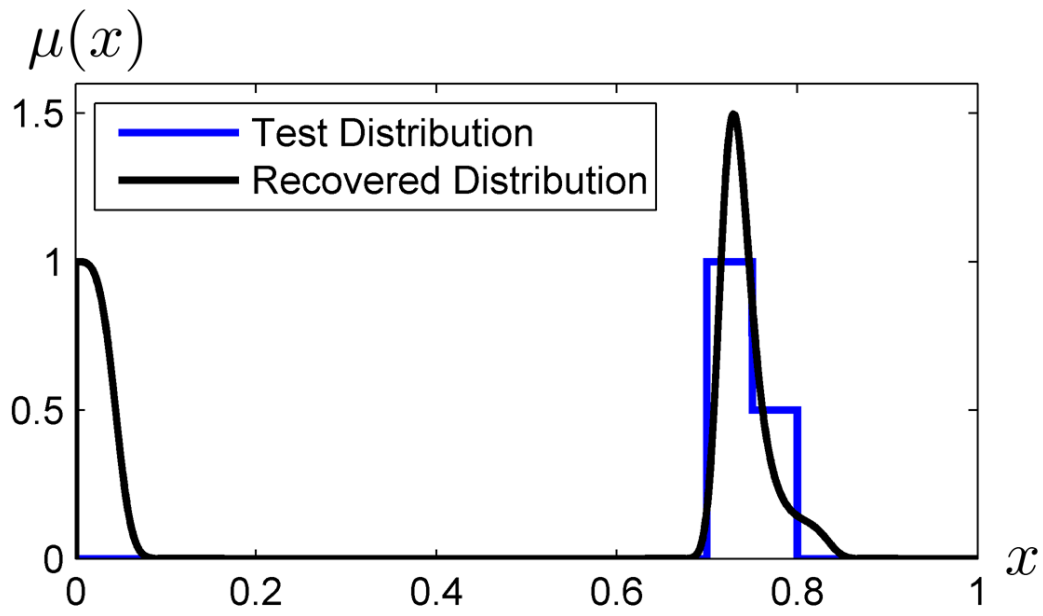


Figure 79. Maximum entropy solution with no bias (four modes, cantilever). The entropy of the unknown test distribution was maximized, given that it would also satisfy the frequency shift formula in Equation 1. No extra assumptions were used to localize the solution, and the solution was numerically indistinguishable from the frequency shifts simulated from the test distribution.

should be based on the physics of the problem, and could include that the adsorbate is expected to be a single, localized distribution, as opposed to two distinct distributions. Recall from Section 5.3 that this is, in general, not possible unless we severely constrain the overall solution space—for example, limiting the possibilities to a single point mass, a single localized distribution, or two point masses. With this strong constraint, it is possible to fit two delta functions to the data as shown in Figure 74. These can be used as the left and right endpoints of the bias function:

$$\Pi(x) = \begin{cases} 1, & x_0 - \delta/2 < x < x_1 + \delta/2 \\ 0, & \text{otherwise} \end{cases} \tag{79}$$

where δ can be increased from 0 until one of the endpoints reaches the boundary of the device or, alternatively, until the overall extents reach a maximum width considered reasonable given *a priori* knowledge about the adsorbate. Since the bias function constructed in

Equation 79 dictates whether or not the solution exists within a certain range, it will be referred to as the *measurement zone*.

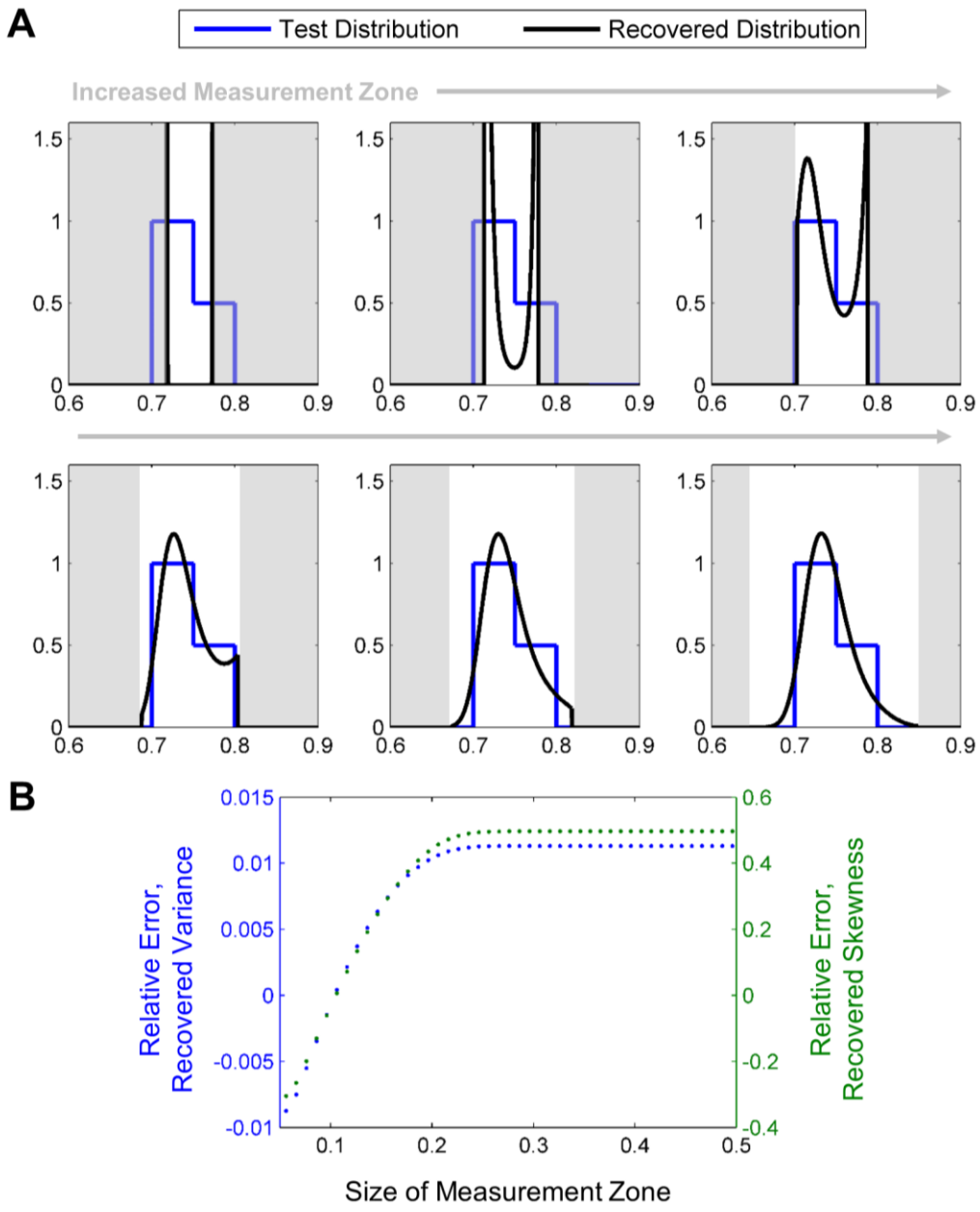


Figure 80. Maximum entropy solution with restricted measurement zone (4 modes). **(A)** A bias function was used to restrict the region where the solution was permitted to exist. As the measurement zone increases, the solution changes from two singular distributions, to resembling the true solution, to resembling a skew-normal distribution. *X-axis*: normalized position, *Y-axis*: normalized mass density. **(B)** Relative error of variance and skewness of recovered distribution versus width of measurement zone. Units of the measurement zone are normalized to the length of the beam.

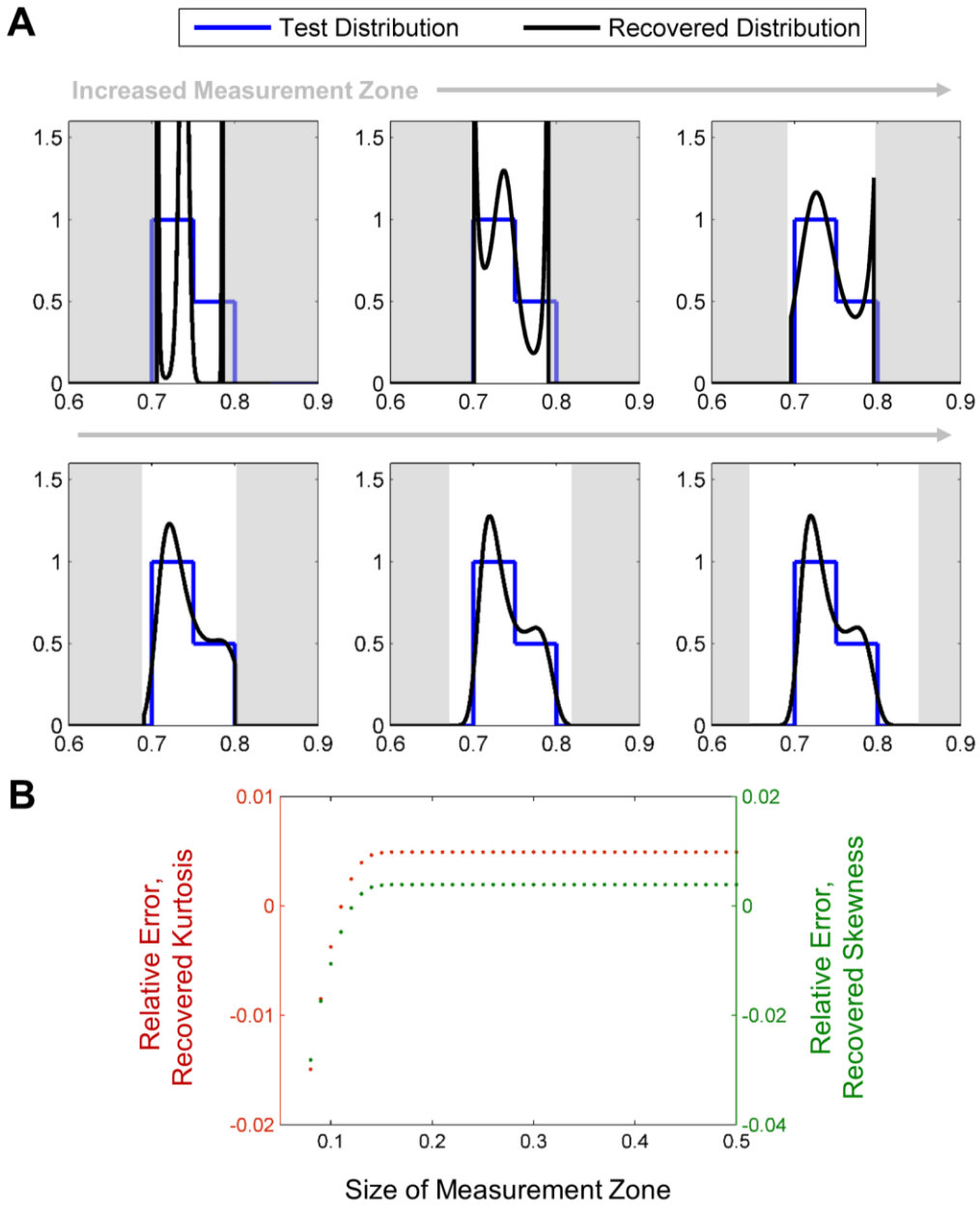


Figure 81. Maximum entropy solution with restricted measurement zone (5 modes). **(A)** As the measurement zone increased, the solution changed from several singular distributions to resembling the true solution. *X-axis*: normalized position, *Y-axis*: normalized mass density. **(B)** Relative error of skewness and kurtosis of distribution versus choice of measurement zone.

A demonstration of the dependence of the solution on the choice of measurement zone is shown in Figure 80. It is observed that, when the left and right endpoints of the measurement zone are chosen to equal the locations of the two delta functions fitted with

Equation 59, the maximum entropy solution similarly features two distinct, singular distributions. The code does not converge if the measurement zone is smaller than this initial choice, but it always successfully converges if the measurement zone is larger. When the measurement zone is increased symmetrically until $x_1 = 1$ (the end of the cantilever), the solution resembles a skew normal distribution. The mass, position, and variance of the recovered solution have an error of order 1%, regardless of the choice of measurement zone. The skewness of the recovered solution has an error between –30% to +50%. Thus, without imposing additional assumptions about the nature of the unknown distribution, four modes are sufficient to reliably recover the mass, position, and variance, but not the skewness.

This approach was repeated with the same test distribution using five modes, as shown in Figure 81. In this case, the first five fitted moments (up to kurtosis) are all found to be highly accurate regardless of the choice of measurement zone. Additionally, the fitted moments and the appearance of the recovered distribution are stable over a wide range of measurement zones once this zone exceeded the test distribution's extents.

Figure 80 and Figure 81 demonstrate that with four or five modes, a conservative measurement zone of up to 50% of the length of the beam can be used to bias the maximum entropy solution in order to recover the unknown mass distribution. However, the physical adsorbate likely has sharply defined spatial extents, rather than the long tails obtained in the computed distributions using the largest measurement zones. Similarly, the physical adsorbate will not have singularities associated with the smallest measurement zones. These basic properties can serve to justify the choice of a measurement zone, ranging from the smallest possible zone allowing for code convergence and larger ones that provide the most

stable results. Imposing this stronger form of bias will improve the accuracy of the fitted moments when only four modes are available.

When choosing a conservative measurement zone that is 50% of the length of the beam, simple canonical distributions such as normal, skew, and bimodal distributions can be accurately recovered using five modes, as shown in Figure 82. More complicated distributions, such as a trimodal distribution, require more modes. Figure 83 shows the recovery of a trimodal distribution with increasing number of modes: with only three modes available, the recovered distribution incorrectly resembles a normal distribution and the recovered moments are accurate only up to the variance; with five modes, the recovered

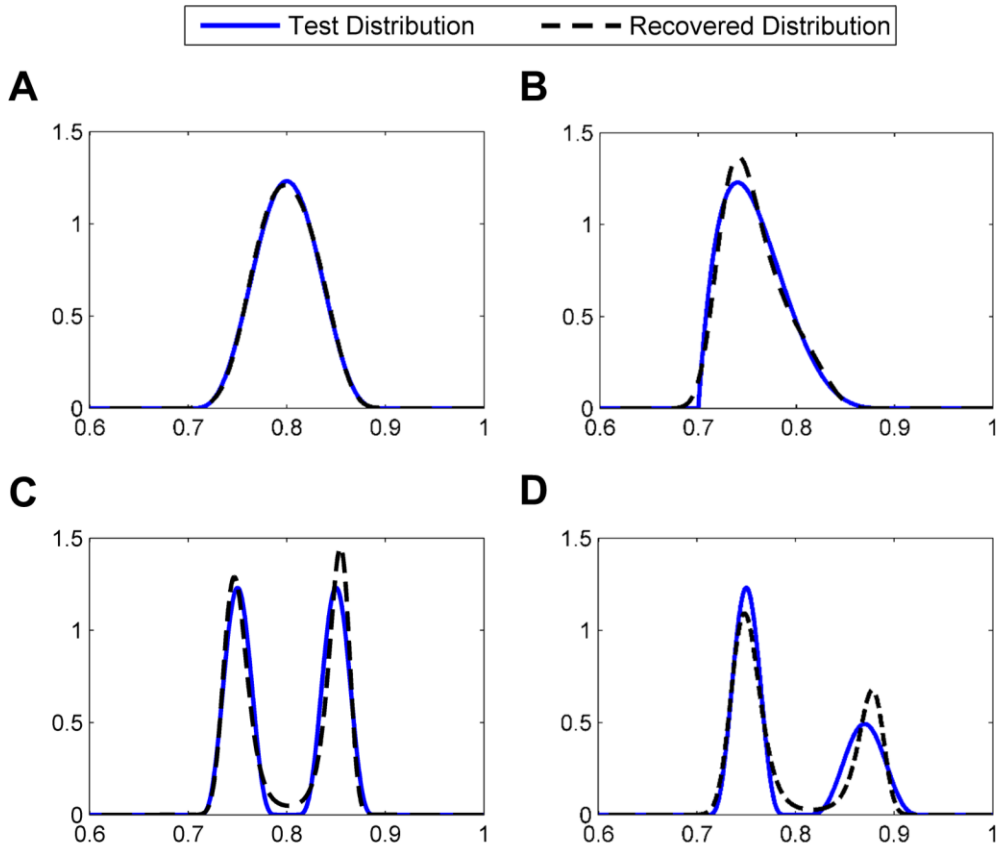


Figure 82. Maximum entropy solution for canonical test distributions. A series of canonical test distributions could be recovered with excellent agreement using the maximum entropy solution with five modes for a cantilever: **(A)** normal, **(B)** skew, **(C)** bimodal, and **(D)** skew bimodal. *X-axis*: normalized position, *Y-axis*: mass density.

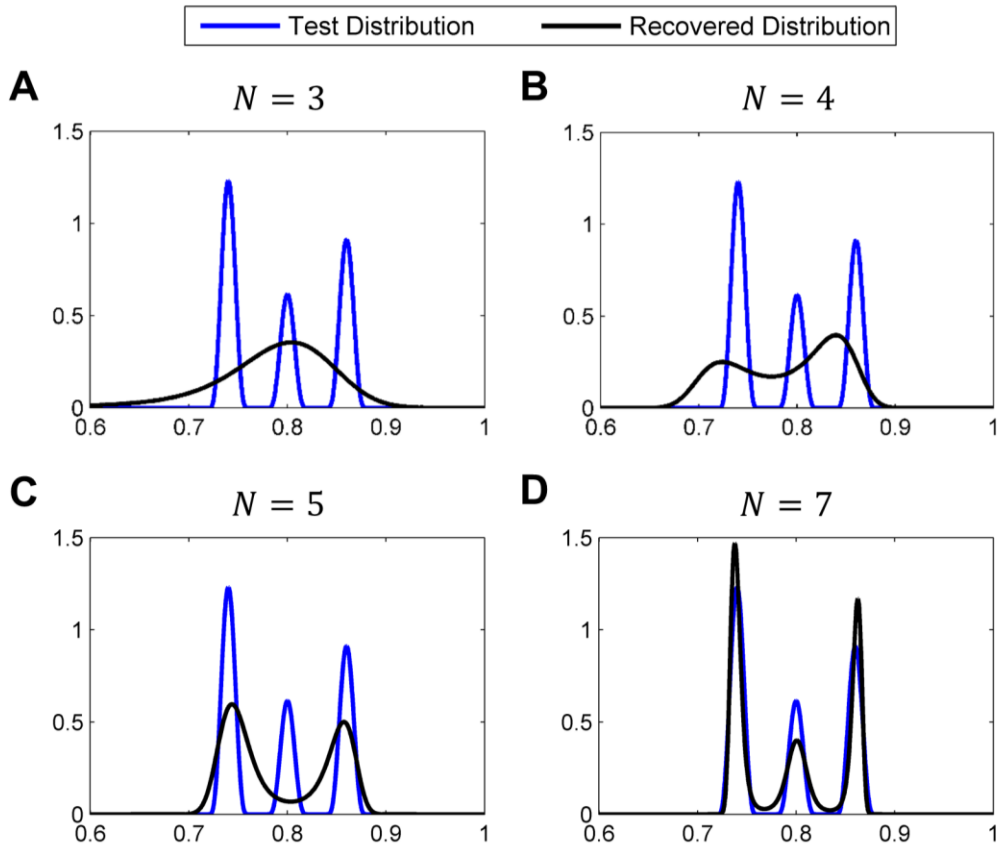


Figure 83. Maximum entropy solution for a complicated test distribution. A trimodal test distribution was used to simulate a series of droplets. Maximum entropy was used to recover the distribution with an increasing number of modes. *X-axis*: normalized position, *Y-axis*: normalized mass density.

distribution incorrectly resembles a bimodal distribution and the recovered moments are accurate only up to the kurtosis; seven modes are necessary to fully recover the distribution. This demonstrates that determining whether a molecule has two or three distinguishable structural domains requires seven modes.

5.6 Maximum Entropy Regularization for a Small Adsorbate

The maximum entropy solution, given in Equation 77, is constructed using the squared mode shapes and is therefore capable of reconstructing a wide variety of shapes, provided these squared mode shapes vary substantially over the measurement zone of

interest. Initial simulations failed to achieve numerical convergence for test distributions with widths less than 2% of the length of the beam. It was found that smaller adsorbates required increasingly larger values of λ_n to converge to the solution, and the requirement to exponentiate these large values during the computation led to numerical stability issues.

To see this, assume that the mass density distribution represented by Equation 77 has a global maximum at x_0 so that $\mu'(x_0) = 0$. If we write the maximum entropy solution as follows:

$$\mu(x) = \Pi(x)\exp\big(-f(x)\big), \qquad f(x) = \sum_{n=1}^{N} \lambda_n \Phi_n^2(x), \tag{80}$$

the function $f(x)$ can be expanded about x_0:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + O((x - x_0)^3) \tag{81}$$

If the bias function $\Pi(x)$ represents a uniform function over a chosen measurement zone, the location of the global maximum x_0 will also describe the maximum for $f(x)$, so $f'(x_0) = 0$. In this case, $\mu(x)$ can be represented approximately as a Gaussian distribution:

$$\mu(x) \approx \exp\big(-f(x_0)\big)\exp\left(-(x - x_0)^2 \sum_{n=1}^{N} \lambda_n \frac{\partial^2 \Phi_n^2(x)}{\partial x^2}\bigg|_{x_0}\right) \tag{82}$$

which has variance

$$\sigma^2 = \frac{1}{2\sum_{n=1}^{N} \lambda_n \frac{\partial^2 \Phi_n^2(x)}{\partial x^2}\Big|_{x_0}}\,. \tag{83}$$

Since the mode shapes have been normalized, we therefore obtain

$$\lambda_n \sim \frac{1}{\sigma^2}\,. \tag{84}$$

Thus, as $\sigma \to 0, \lambda_n \to \infty$.

To address this issue, the modified Gram-Schmidt procedure was used to orthonormalize the squared mode shapes for a given choice of measurement zone. This procedure has been discussed as a means for preconditioning data for maximum entropy algorithms to make them numerically stable [3]. Intuitively, this process transforms the mode shapes, which would all identically look like constant functions over a small interval with $\sigma \to 0$ such that they vary appreciably from each other over a new interval with larger effective σ. The frequency-shift data is transformed as well such that the reformulated problem is completely equivalent and no data is lost. We find this transformation to be an essential step to find solutions that converge over a small measurement zone. In brief, the algorithm constructs a new set of functions $u_i(x)$ using a linear superposition of the squared mode shapes:

$$u_i(x) = \sum_{j=1}^{N} \mathrm{G}_{ij} \Phi_j^2(x) \tag{85}$$

that satisfy the orthonormality condition:

$$\int_{x_0}^{x_1} u_i(x) u_j(x)\, dx = \delta_{ij} \tag{86}$$

over the new measurement zone $x_0 \leq x \leq x_1$. Note that while the original mode shapes $\Phi_n(x)$ are orthonormalized over the entire beam $0 \leq x \leq 1$, the shared mode shapes $\Phi_n^2(x)$ are *not*, so in principle, this Gram-Schmidt procedure could be carried out over any interval to improve the numerical performance of the algorithm. Once the coefficients G_{ij} have been calculated, the frequency-shift data represented by the integral in Equation 54 can be transformed to the new basis:

$$\int_{x_0}^{x_1} \mu(x) u_i(x)\, dx = \sum_{j=1}^{N} \mathrm{G}_{ij} \int_{x_0}^{x_1} \mu(x)\, \Phi_j^2(x)\, dx$$

$$= \sum_{j=1}^{N} \mathrm{G}_{ij} \int_{0}^{1} \mu(x)\, \Phi_j^2(x)\, dx = \sum_{j=1}^{N} \mathrm{G}_{ij} c_j, \qquad (87)$$

where the second equality holds provided the measurement zone completely encompasses the adsorbate. A demonstration of the modified Gram-Schmidt procedure applied to the first five modes of a cantilever, over a measurement zone representing 0.2% of the extents of the

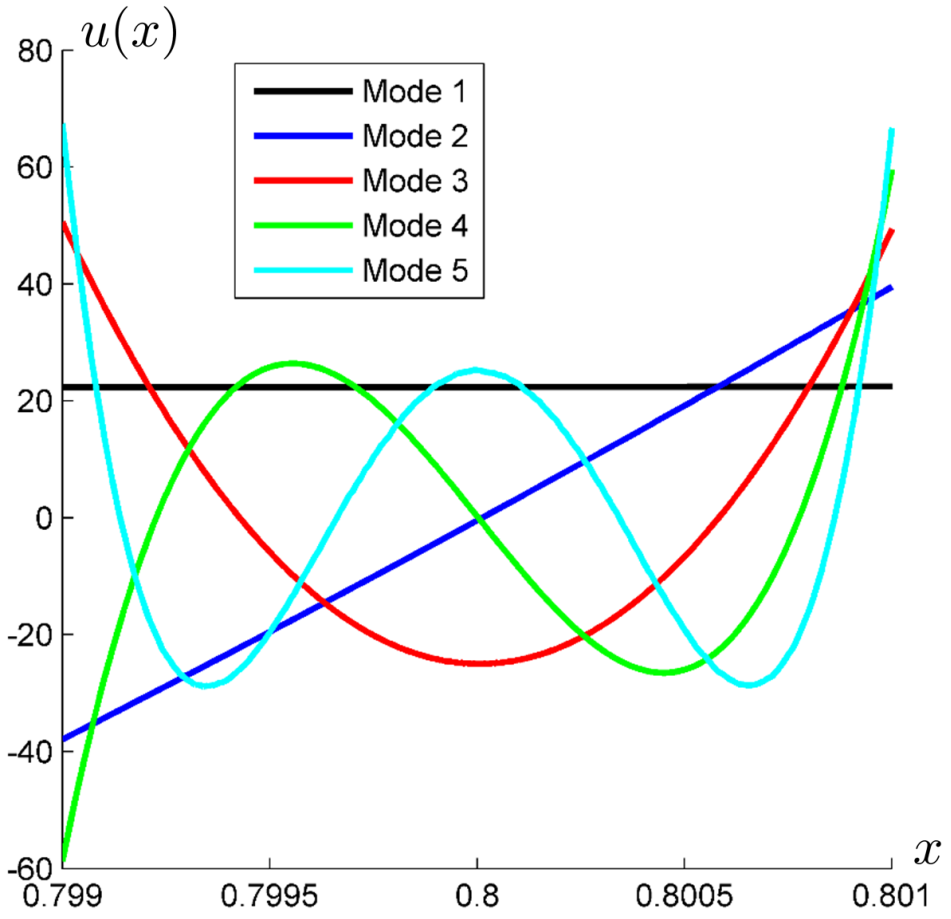


Figure 84. Orthonormalized modes over a small measurement zone. The modified Gram-Schmidt procedure was used to create linear superpositions of the squared mode shapes that are orthonormal over a small interval. These new mode shapes can be used to precondition the frequency-shift data to allow for numerical convergence of maximum entropy over a small interval.

beam, is shown in Figure 84. Over a small measurement zone, they resemble Legendre polynomials.

Figure 85 shows the test distribution with extents $\epsilon = 0.001$ reconstructed using maximum entropy using four or five modes after orthogonalizing the squared mode shapes over the measurement zone. This measurement zone is chosen conservatively in the following manner: the data is initially fitted to a two point-mass solution with positions x_0 and x_1 using techniques discussed in Section 5.3. The left and right endpoints of the

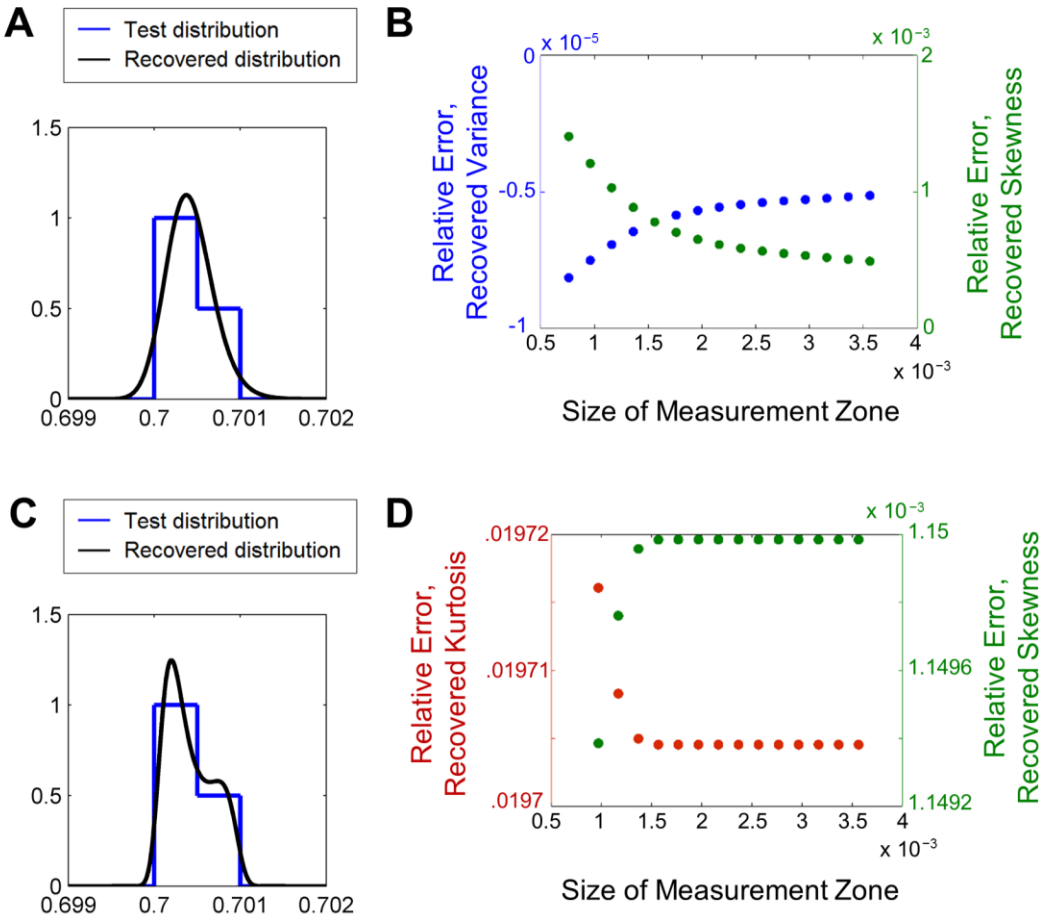


Figure 85. Maximum entropy solution over a small measurement zone. After orthogonalizing the squared mode shapes over a small measurement zone, maximum entropy was used to recover a test distribution using **(A)** 4 modes and **(C)** 5 modes with a conservative choice of measurement zone. The relative error for higher mass moments versus measurement zone (normalized to length of the beam) are shown for **(B)** 4 modes and **(D)** 5 modes.

measurement zone are chosen to be $x_0 - \delta/2$ and $x_1 + \delta/2$, with $\delta = 5(x_1 - x_0)$. Mass moments up to skewness could be accurately using four modes and in a manner that was relatively insensitive to this choice of measurement zone. Similarly, all moments up to kurtosis could be similarly recovered using five modes without a strong dependence on measurement zone. This is shown in Figure 85, where the measurement zone was varied from $\delta = 0$ to $\delta = 5(x_1 - x_0)$.

In Table 16, I show the accuracy of recovered moments from the maximum entropy solution for the same test distribution over a wide range of widths using four modes. In each case, the measurement zone is chosen conservatively, as discussed above. For distributions with widths of $\epsilon < 0.03$, moments up to the fourth (skewness) could be recovered with a high degree of accuracy. Above $\epsilon = 0.03$, moments up to third (variance) were still accurate, but the skewness became inaccurate. This is despite the fact that the maximum entropy

Width, ϵ	Mass	Center-of-Mass	Variance	Skewness
0.001	7.50×10^{-4} (0.00 %)	0.700 (0.00 %)	7.64×10^{-8} (0.00%)	0.439 (0.05 %)
0.003	2.25×10^{-2} (0.00 %)	0.701 (0.00 %)	6.88×10^{-7} (0.00 %)	0.439 (0.06 %)
0.01	7.50×10^{-3} (0.00 %)	0.704 (0.00 %)	7.64×10^{-6} (0.02%)	0.438 (0.16 %)
0.03	2.25×10^{-2} (0.00 %)	0.713 (0.00 %)	6.89×10^{-5} (0.18 %)	0.452 (3.14 %)
0.1	7.49×10^{-2} (-0.08 %)	0.742 (0.03 %)	7.73×10^{-4} (1.13 %)	0.657 (49.70 %)

Table 16. Recovery of mass moments for skewed distribution (cantilever). Test distribution type D shown in Figure 2 was used to generate frequency shifts for four modes for a cantilever. For all distributions, $x_0 = 0.7$, but the spatial extent ϵ was varied. Recovered moments for the maximum entropy solution using a conservative measurement zone are reported; percentage error is in parentheses.

solution generally resembles the test distribution. As discussed previously, the error for the recovered skewness for this larger width distribution can be improved by using a more aggressive (smaller) measurement zone, which eliminates the long tails characteristic of the continuous maximum entropy distribution—instead favoring a more localized distribution. However, because the choice of measurement zone is somewhat arbitrary or would require specific justification using *a priori* knowledge about the adsorbate, the most conservative zone is selected for consistency and comparison purposes.

5.7 Maximum Entropy Regularization in the Asymptotic Limit

The Gram-Schmidt procedure discussed in the previous section can be used with the method of maximum entropy to reconstruct mass density distributions over extremely small measurement zones, $\epsilon \ll 1$, relative to the beam length. In this section, we will investigate the maximum entropy solution as $\epsilon \to 0$ in the asymptotic limit.

The approximation for $\mu(x)$ for small adsorbates given in Equation 82 approaches a delta function as the standard deviation $\sigma \to 0$. This can be shown in another way using Laplace's method, which is used to estimate integrals of the form

$$\int_a^b h(x)e^{Mg(x)}dx, \tag{88}$$

where M is large and $h(x)$ is positive. In this case, the integral can be approximated as [34]:

$$\int_a^b h(x)e^{Mg(x)}dx \approx \sqrt{\frac{2\pi}{M|g''(x_0)|}}h(x_0)\, e^{Mg(x_0)}\,. \tag{89}$$

As noted in Section 5.6, $\lambda_n \to \infty$ as $\sigma \to 0$, so this approximation applies when calculating integrals involving the maximum entropy solution, such as the frequency shifts generated by

the solution via Equation 54, or mass moments of the solution. The frequency shifts can be expressed as:

$$c_n = \int_{x_0}^{x_1} \exp\left(-\sum_{n=1}^{N} \lambda_n \Phi_n^2(x)\right) \Phi_n^2(x) dx$$

$$\approx \sqrt{\frac{2\pi}{\left|\sum_{n=1}^{N} \lambda_n \frac{\partial^2 \Phi_n^2}{\partial x^2}(x_s)\right|}} \exp\left(-\sum_{n=1}^{N} \lambda_n \Phi_n^2(x_s)\right) \Phi_n^2(x_s)$$

$$= m_s \Phi_n^2(x_s) \tag{90}$$

where x_0 and x_1 define the limits of the measurement zone, and m_s and x_s are the adsorbate's mass and position. Therefore, the maximum entropy solution reduces to a point mass in the asymptotic limit of a small adsorbate.

The approximation formula in Equation 89 is accurate to order $O\left(M^{-1/2}\right)$. Additional terms can be calculated using standard techniques for asymptotic expansion of integrals [34]. This was done using integrals involving the frequency shifts as in Equation 90, as well as integrals involving moments of the maximum entropy solution, *i.e.*

$$\int_{x_0}^{x_1} g(x) \exp\left(-\sum_{n=1}^{N} \lambda_n \Phi_n^2(x)\right) dx, \tag{91}$$

where $g(x) = 1, x, x^2$, etc, in order to simplify the final result. With this approach, the frequency shifts can be approximated as:

$$c_n \approx m_s \left[\Phi_n^2(x_s) + (\bar{x} - x_s) \left.\frac{\partial \Phi_n^2}{\partial x}\right|_{x_s} + \frac{\langle x^2 \rangle}{2} \left.\frac{\partial^2 \Phi_n^2}{\partial x^2}\right|_{x_s}\right]. \tag{92}$$

We note the striking similarity to Equation 49 using $p = 3$ (4 modes). Minor differences arise due to the requirement of Laplace's method to expand about the location of the mass density distribution's maximum value, denoted x_s, compared to its center-of-mass, denoted

$\bar{x}$. Some advantages may arise in using Equation 92 to fit mass moments using frequency-shift data when four modes are present: namely, the formula only uses up to the second derivative of the squared mode shapes, yet is able to access asymmetry via the metric $\bar{x} - x_s$ (which is proportional to the Pearson mode skewness) rather than the skewness $\langle x^3 \rangle$ which is a much smaller quantity.

5.8 Experimental Validation

In [1], droplets were deposited onto microcantilevers using dip-pen lithography, and the frequency of four modes of the cantilevers in air was measured before and after droplet deposition. These samples were also imaged optically using a microscope to obtain the known center-of-mass, variance, and skewness for comparison. Techniques in [1] and [2] have been used to recover the mass moments from the generated frequency-shift data. In this section, techniques from Section 5.3 and Section 5.5 will be applied to directly recover the density distribution of the published droplet data, then these recovered distributions will be compared to the optical images. Additionally, the mass moments of the recovered mass density distribution will be compared to the known mass moments.

Discrete Density Reconstruction. As discussed in Section 5.3, one possible generalization to the point-mass formula is given by Equation 59, which gives in the case of four modes ($N = 4$):

$$c_n = m_1 \phi_n^2(x_1) + m_2 \phi_n^2(x_2), \tag{93}$$

where m_1 and m_2 are relative to the mass of the beam. Equation 93 was used to fit two point masses of different weights to the experimental frequency-shift data collected with 4 modes and originally published in [1]. Reconstructed density distributions using this approach are

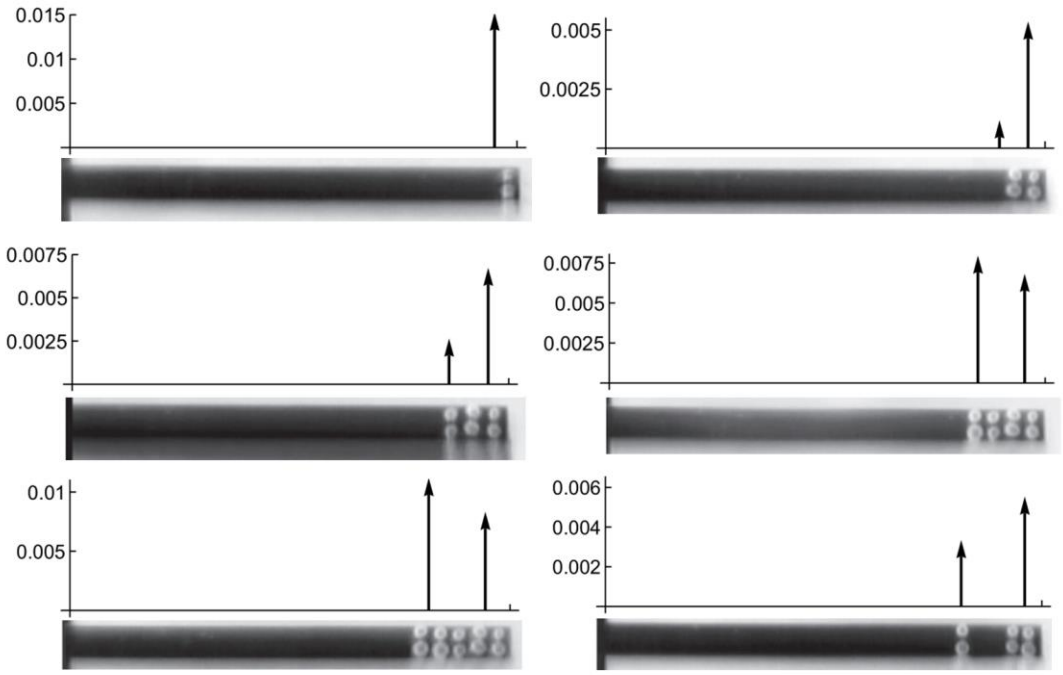


Figure 86. Discrete density reconstruction of experimental droplet data. Optical images for droplets placed on microcantilever using dip-pen lithography are taken from [1] and replicated here. Discretization is used to reconstruct the mass density distributions using the frequency-shift data for four modes, and these distributions are overlain above of the optical images such that the X-axis is precisely aligned with the cantilever extents. Recovered distribution represented as delta functions with height on Y-axis representing mass relative to the beam. Cantilever dimensions are 397 μm long, 29 μm wide and 2 μm thick.

shown in Figure 86. The general location and extents of the distribution can be inferred from these recovered distributions; however, they do not give much of an impression as to the overall shape. The reconstructed distribution for the asymmetric droplets is somewhat distinguishable in its asymmetry from the case of four or five rows of droplets, but the fidelity of reproduction is weaker in the case of two or three rows of droplets, where the distribution incorrectly gives an impression of asymmetry. In the case of a single droplet, Equation 93 can be used to recognize that a single localized droplet was present, although the location of the droplet was somewhat inaccurate, and the width of the droplet was inaccessible.

Table 17 shows the variance of the recovered distributions and compares with the two other techniques, linear superposition and the mode shape expansion approach. For the case of four rows, five rows, and asymmetric rows, the discretized density distribution recovered from the data had a variance almost exactly between the two other approaches. It also produced results consistent with the other approaches in the case of two-row and three-row droplet arrays. The skewness of the distribution recovered using Equation 93 for the asymmetric row data is −0.509; this is also between the value provided from linear superposition theory (−0.453) [1] and the value provided from the truncated series approach (−0.695) [2], and provides the most accurate result compared with the known optical value (−0.537) [2].

Maximum Entropy Regularization. As discussed in Section 5.5, this method is most successful when careful consideration of the bias function $\Pi(x)$ is used to localize the solution. Discrete density reconstruction as shown in Figure 86 can be used first, indicating

Droplet array	Linear superposition	Mode shape Taylor expansion	Equation 93
2 rows	5.35 E−4 (±38%)	5.32 E−4 (±34%)	6.30 E−4
3 rows	1.60 E−3 (±6.9%)	1.56 E−3 (±6.5%)	1.62 E−3
4 rows	2.92 E−3 (±1.4%)	2.79 E−3 (±1.2%)	2.86 E−3
5 rows	4.33 E−3 (±0.39%)	3.87 E−3 (±0.36%)	4.04 E−3
Asymmetric rows	5.21 E−3 (±1.5%)	4.67 E−3 (±1.3%)	4.97 E−3

Table 17. Recovery of variance using different approaches. Measured normalized variance of the droplet arrays of [1, 2]. Results obtained using linear superposition theory of [1], mode shape expansion theory of [2], and discretization approach using Equation 93. Listed uncertainties in the parentheses specify a 95% confidence interval and are determined by the uncertainty in the measured resonant frequencies. The first four modes of the cantilever are used for all theories.

that the distribution always ends near the tip of the cantilever, but varies somewhat in the starting point. For simplicity, we can choose for our bias function:

$$\Pi(x) = \begin{cases} 1, & x_0 < x < 1 \\ 0, & \text{otherwise} \end{cases}, \tag{94}$$

such that we only vary the beginning of the measurement zone $0 < x_0 < 1$ over the length of the device. Frequency-shift data was reconstructed from the maximum entropy solution and compared with the experimental frequency-shift data as shown in Figure 87. For $x_0 < 0.79$, frequency-shift data reconstructed from the maximum entropy solution is nearly

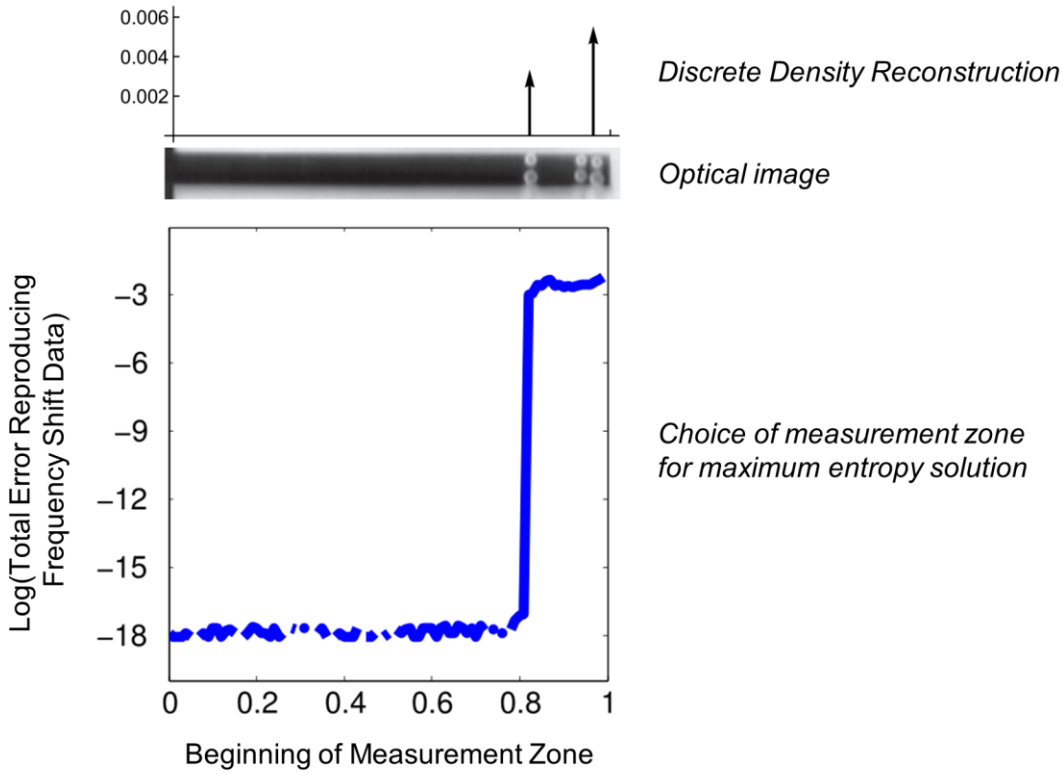


Figure 87. Choice of measurement zone for experimental droplet data. Discrete density reconstruction (top) previously indicated the distribution was near the tip of the cantilever. The measurement zone for the maximum entropy method was selected to begin at a variable position x_0 along the length of the cantilever and end at the tip of the cantilever. After the maximum entropy solution was found using a given measurement zone, the frequency-shift data was reconstructed and the total error calculated with respect to the experimental data. This total error is shown with an X-axis scaled to align with the optical image and has a sharp transition at $x_0 = 0.79$.

indistinguishable from the experimental data. At $x_0 = 0.79$, a sharp transition occurs and the solution no longer accurately reproduces the original data. This contrasts with the situation using simulated data in Section 5.5, which failed to converge at all when the measurement zone chosen for the algorithm became smaller than the original distribution. The differing behavior between simulation and experiment for measurement zones that are too small may be due, in part, to device noise. Fortunately, the device noise in these experiments was small enough for the transition between a successful image reconstruction and an unsuccessful one to be clear. This transition occurs precisely where the droplets physically begin and is evident without any *a priori* knowledge of the adsorbate's image. The measurement zone is therefore chosen to begin immediately prior to this transition point for each set of droplet data.

This procedure was used for the asymmetric droplet data to select the bias function $\Pi(x)$ given in Equation 42 for three and four modes. The resulting maximum entropy solution is shown along with this choice of bias function in Figure 88. Of note, three modes are evidently sufficient to indicate a bimodal distribution, rather than a normal distribution, which would normally be constructed given knowledge of three moments (mass, center-of-mass, variance). The recovered images have sharp spikes near the edges of the measurement zone, and these are an artifact arising from the choice of a rectangular bias function that does not match the physical reality of the adsorbate. This can be avoided by choosing a smoother bias function; depending on the *a priori* knowledge of the adsorbate, its shape can be chosen more precisely. Figure 88 shows that with a parabolic bias function, the reconstructed image is smoother and matches the data with a high degree of fidelity. Also of note is that asymmetry is evident in the reconstructed image using only 3 modes. This is not normally

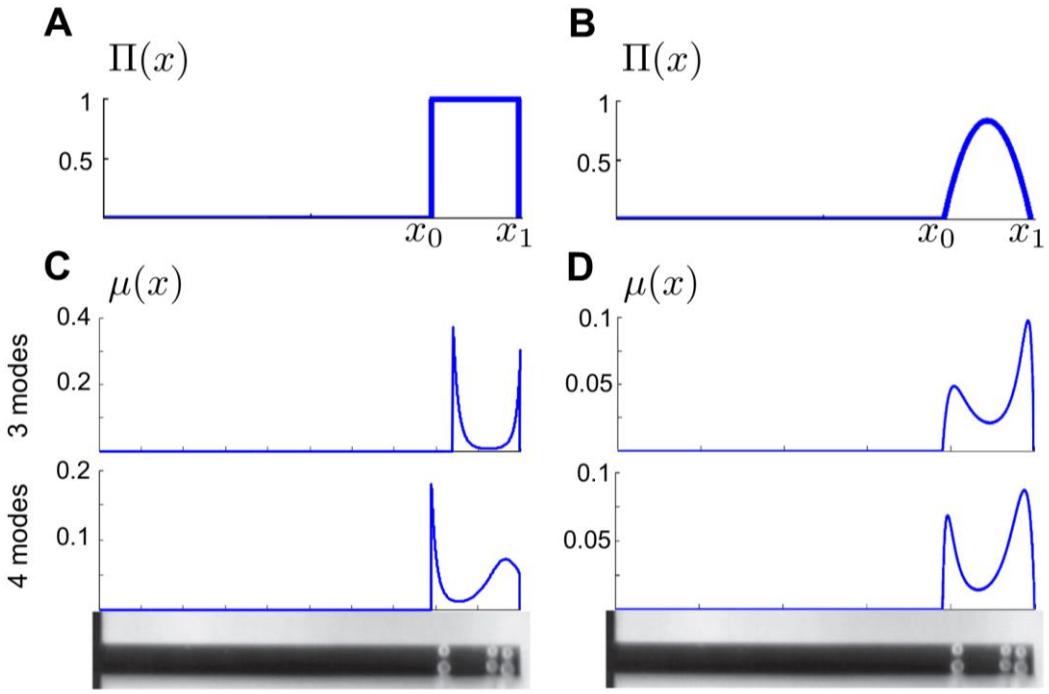


Figure 88. Maximum entropy reconstruction of asymmetric droplet data. A **(A)** rectangular or **(B)** parabolic bias function was chosen for the maximum entropy algorithm, and the reconstructed image is shown in **(C)** and **(D)** for these bias functions, respectively. In each case, $x_1 = 1$ and x_0 was chosen to minimize the measurement zone while still accurately reproducing the experimental data

expected, but, in this case, arises due to the fortuitous positioning of the adsorbate at the edge of the cantilever. One end of the measurement zone can't go past the tip of the cantilever, but the other end can be chosen in a way that closely matches the physical start of the distribution. Thus, the center of the measurement zone does not necessarily align with the center-of-mass of the adsorbate.

This procedure was repeated for all droplet data published in [1] using the parabolic bias function, and the results are shown in Figure 89. Compared with the discrete density reconstruction method shown in Figure 86, the maximum entropy method appears to give a more informative picture of unknown distribution.

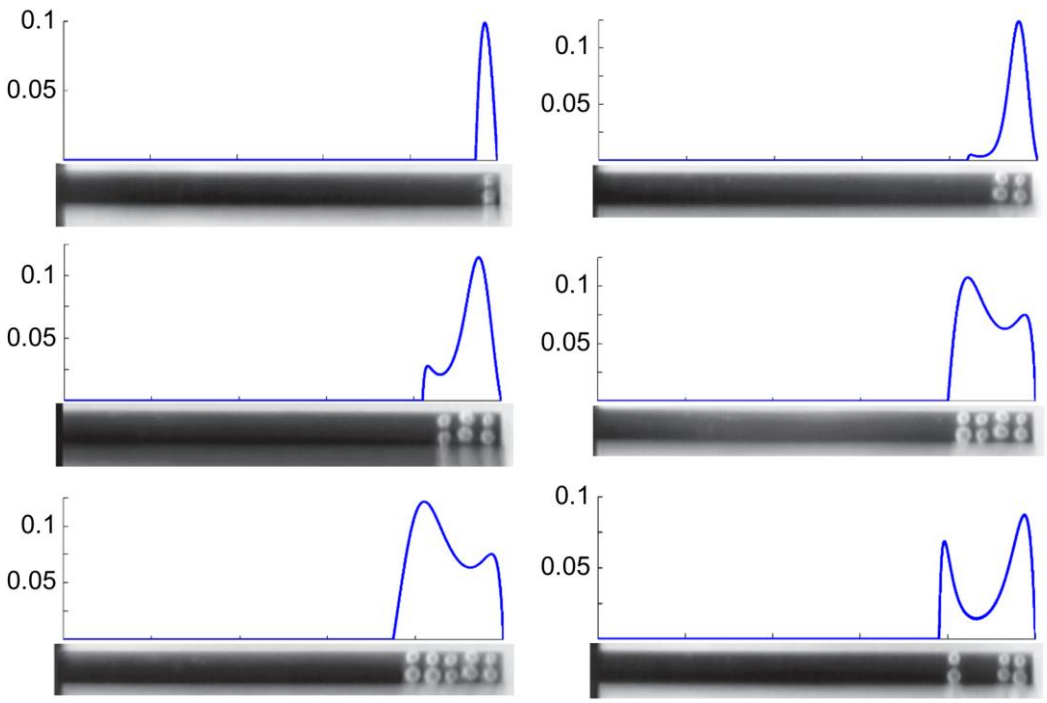


Figure 89. Maximum entropy reconstruction of experimental droplet data. A smooth (parabolic) bias function was used with 4 modes for all droplet data published in [1], and maximum entropy regularization was used to recover the image. The recovered image is shown above the optical data with the X-axis aligned precisely with the cantilever dimensions.

5.9 Two-Dimensional Imaging

As has been noted in the literature, the method of maximum entropy can be readily extended to higher dimensions [3]. Here, the only difficulty lies in the increased computational difficulty. The problem can be formulated as follows:

$$\min S \equiv \int_{\Omega} \mu(\vec{x}) \ln \frac{\mu(\vec{x})}{\Pi(\vec{x})} d\vec{x}, \tag{95}$$

subject to

$$c_n \equiv \int_{\Omega} \mu(\vec{x}) \Phi_n^2(\vec{x}) d\vec{x}, \qquad n = 1 \ldots N \tag{96}$$

which has the solution:

$$\mu(\vec{\boldsymbol{x}}) = \Pi(\vec{\boldsymbol{x}}) \exp\left(-\sum_{n=1}^{N} \lambda_n \Phi_n^2(\vec{\boldsymbol{x}}) \right), \tag{97}$$

Algorithms and code to calculate the λ_n in Equation 45 given a set of moments expressed in the form:

$$c_{ij} \equiv \int_{\Omega} \mu(\vec{\boldsymbol{x}}) x^i x^j \, dx \, dy, \tag{98}$$

are provided in [3] and could be adapted to use known mode shapes $\Phi_n^2(\vec{\boldsymbol{x}})$ from a two-dimensional membrane device. Alternatively, the moments in the form of Equation 46 can be calculated using the truncated expansion approach (which has been shown to be feasible

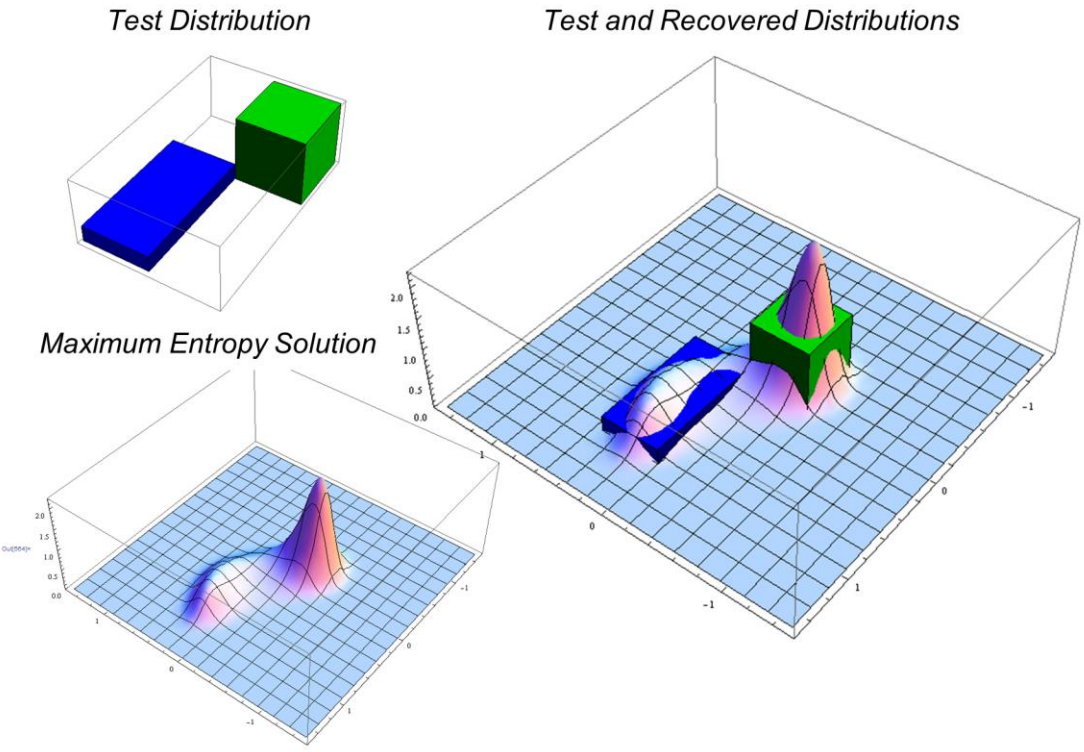


Figure 90. Two-dimensional maximum entropy imaging. A two-dimensional test distribution was used to calculate the first 15 standard moments. The distribution was recovered using these moments using software provided in [2, 3].

in two dimensions [2]), and then the image can be reconstructed from these standard moments. Figure 90 demonstrates use of the software to reconstruct the two-dimensional density distribution of an adsorbate assuming that these standard moments are accurately known.

5.10 Conclusion

In this chapter, we have investigated several methods to reconstruct the mass density distribution of an adsorbate directly from the frequency-shift data. A simple extension to the point-mass approximation, representing the adsorbate as a series of point masses, is found to be versatile in that it could be used to distinguish when multiple different particles are present, similar to [19], but can also be used within a single distribution of arbitrarily small spatial extents to infer information about the adsorbate's size and asymmetry. Applied to the latter situation, the mass moments of the recovered distribution are unexpectedly found to be more accurate than if the mass moments were fit to the data directly via the previously introduced linear superposition approach [1] or truncated expansion approach [2]. The improvement over prior theories becomes negligible as the size of the particle goes to zero.

Further aspects of the adsorbate's shape are able to be recovered with a large number of modes by making use of maximum entropy regularization. This method provides a solution with the most disorder relative to a bias function. When given a uniform bias function extending over the entire beam, the solution tended to occupy areas of low mass responsivity in the beam. However, if the experiment is carefully designed to deposit one, or at most two, adsorbates at a time, these aspects of the solution can be eliminated by choosing a bias function that limits the solution to a small region near the adsorbate's estimated location. This approach works for arbitrarily small adsorbates, and is also demonstrated on

experimental data of droplets placed on a microcantilever. The mass moments calculated from the maximum entropy solution that employs the experimentally measured frequency-shift data are also unexpectedly found to be more accurate than moments calculated from the data directly.

The techniques in this chapter solely require knowledge of the device mode shapes and do not explicitly rely on derivatives. This could provide an advantage in situations where the mode shapes are not known with complete accuracy, let alone their higher derivatives. Experimental attempts to measure the mode shapes have been reported; in some cases, they differ substantially from the expected theory [35]. If the experimental mode shapes are thought to be more trustworthy than the theory, they could be used as inputs for the maximum entropy method discussed here, provided they are linearly independent.

5.11 Bibliography

1. Hanay, M.S., et al., *Inertial imaging with nanomechanical systems.* Nature nanotechnology, 2015. **10**(4): p. 339-344.
2. Sader, J.E., et al., *Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation.* Nano Letters, 2018. **18**(3): p. 1608-1614.
3. Abramov, R.V., *The multidimensional maximum entropy moment problem: A review of numerical methods.* Communications in Mathematical Sciences, 2010. **8**(2): p. 377-392.
4. Smith, K.T. and F. Keinert, *Mathematical foundations of computed tomography.* Applied Optics, 1985. **24**(23): p. 3950-3957.
5. Engl, H.W., M. Hanke, and A. Neubauer, *Regularization of inverse problems*. Vol. 375. 1996: Springer Science & Business Media.
6. Weber, C.F., *Analysis and solution of the ill-posed inverse heat conduction problem.* International Journal of Heat and Mass Transfer, 1981. **24**(11): p. 1783-1792.
7. O'Sullivan, F., *A statistical perspective on ill-posed inverse problems.* Statistical Science, 1986. **1**(4): p. 502-518.
8. Ang, D.D., et al., *Moment theory and some inverse problems in potential theory and heat conduction*. 2004: Springer.
9. Alberts, B., et al., *The shape and structure of proteins*, in *Molecular Biology of the Cell. 4th edition*. 2002, Garland Science.
10. Georgiou, T.T., *Solution of the general moment problem via a one-parameter imbedding.* IEEE Transactions on Automatic Control, 2005. **50**(6): p. 811-826.

11. Talenti, G., *Recovering a function from a finite number of moments.* Inverse problems, 1987. **3**(3): p. 501.
12. Akhiezer, N.I.i., *The classical moment problem: and some related questions in analysis*. Vol. 5. 1965: Oliver & Boyd.
13. Hanay, M.S., et al., *Single-protein nanomechanical mass spectrometry in real time.* Nature Nanotechnology, 2012. **7**(9): p. 602-608.
14. Naik, A.K., et al., *Towards single-molecule nanomechanical mass spectrometry.* Nature Nanotechnology, 2009. **4**(7): p. 445-450.
15. Press, W.H., *Numerical recipes 3rd edition: The art of scientific computing*. 2007: Cambridge university press.
16. Mead, L.R. and N. Papanicolaou, *Maximum entropy in the problem of moments.* Journal of Mathematical Physics, 1984. **25**(8): p. 2404-2417.
17. Athanassoulis, G. and P. Gavriliadis, *The truncated Hausdorff moment problem solved by using kernel density functions.* Probabilistic Engineering Mechanics, 2002. **17**(3): p. 273-291.
18. De Souza, L., et al., *Reconstruction of a distribution from a finite number of moments with an adaptive spline-based algorithm.* Chemical Engineering Science, 2010. **65**(9): p. 2741-2750.
19. Dohn, S., et al., *Position and mass determination of multiple particles using cantilever based mass sensors.* Applied Physics Letters, 2010. **97**(4): p. 044103.
20. Ng, A.Y. *Feature selection, L1 vs. L2 regularization, and rotational invariance*. in *Proceedings of the twenty-first international conference on machine learning*. 2004. ACM.
21. Shore, J. and R. Johnson, *Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy.* IEEE Transactions on Information Theory, 1980. **26**(1): p. 26-37.
22. Tagliani, A., *Hausdorff moment problem and maximum entropy: a unified approach.* Applied Mathematics and Computation, 1999. **105**(2-3): p. 291-305.
23. Tagliani, A., *Numerical aspects of finite Hausdorff moment problem by maximum entropy approach.* Applied Mathematics and Computation, 2001. **118**(2-3): p. 133-149.
24. Bretthorst, G.L. *The maximum entropy method of moments and Bayesian probability theory*. in *AIP Conference Proceedings*. 2013. AIP.
25. Ormoneit, D. and H. White, *An efficient algorithm to compute maximum entropy densities.* Econometric reviews, 1999. **18**(2): p. 127-140.
26. Antolin, J., *Maximum entropy formalism and analytic extrapolation.* Journal of Mathematical Physics, 1990. **31**(4): p. 791-797.
27. Shannon, C.E., W. Weaver, and A.W. Burks, *The mathematical theory of communication.* 1951.
28. Jaynes, E.T., *Information theory and statistical mechanics*, in *Brandeis University Summer Institute Lectures in Theoretical Physics: 1962 Lectures, Volume 3 - Statistical Physics*, K.W. Ford, Editor. 1963, Benjamin: New York. p. 181-218.
29. Jaynes, E.T., *Information theory and statistical mechanics.* Physical Review, 1957. **106**(4): p. 620.
30. Jaynes, E.T., *Prior probabilities.* IEEE Transactions on systems science and cybernetics, 1968. **4**(3): p. 227-241.

31. Hobson, A., *A new theorem of information theory*. Journal of Statistical Physics, 1969. **1**(3): p. 383-391.
32. Friedman, K. and A. Shimony, *Jaynes's maximum entropy prescription and probability theory*. Journal of Statistical Physics, 1971. **3**(4): p. 381-384.
33. Hobson, A., *The interpretation of inductive probabilities*. Journal of Statistical Physics, 1972. **6**(2): p. 189-193.
34. Bleistein, N. and R.A. Handelsman, *Asymptotic expansions of integrals*. 1975: Courier Corporation.
35. Hung, P.S.-H., *Advanced applications of nanoelectromechanical systems (Doctoral dissertation)*. 2016, California Institute of Technology.

Chapter 6

CONCLUDING REMARKS

The developments presented in this thesis provide much needed progress towards the realistic use of NEMS devices to advance the field of mass spectrometry, however much remains to be done. Acquiring a modern, sophisticated MS instrument such as the Q Exactive was extremely fortuitous for our lab. Previous incarnations of MS used home-built setups which had no method to independently verify the identity of ions delivered to the NEMS and very low event rates (one every ~20 minutes). In contrast, we had regular communications with Dr. Alexander Makarov and his expert R&D team at Bremen, and I was able to visit the factory where the instruments were made to discuss finer points about instrument operation. Additionally, I was able to meet Professor Albert Heck's lab and learn the finer points about Native MS. With these resources, we could build an extraordinarily sophisticated front-end for delivering molecules to NEMS, as shown in Part I.

With this hybrid instrument, several important experiments could be performed. The hybrid instrument could be used to verify that GroEL ions were sent to NEMS devices, and single molecule adsorption events could be easily resolved. Mass spectra collected with two and three modes showed excellent agreement with the expected mass. It was shown that careful analysis is necessary to compile final mass spectra from an ensemble of measurements. Additionally, we observed that charge from molecules sent from the Q Exactive could confound frequency-based mass measurements in an indirect way, but this effect could be eliminated with the use of a 2f detection technique. Finally, the use of improved cryogenics indicated the mass resolution of LETI style devices are limited by self-

heating. Future device designs with improved mass resolution (now in progress with our lab) will need to take note of these observations to avoid pitfalls.

In Chapter 3, we showed that ion optics can be used to produce a beam with a spot size of 40 μm^2, allowing for a maximum detection rate of 25 molecules per second. Using a 20-device array consisting of 10 μm by 300 nm devices thus currently allows for the detection of 5% of molecules reaching the device chip. This suggests that NEMS measurements can achieve a good limit of detection (the number of molecules in the initial solution needed to produce a signal). The only other factors affecting the detection limit using NEMS are the electrospray ionization efficiency (typically 5-10%, with the theoretical limit up to 85% [1]) and the transfer efficiency of ions to the NEMS (estimated to be 13% in Section 2.3). Thus, the current detection limit achieved in this thesis is estimated to be 10^4-10^5 molecules. Work is underway to improve this number significantly by combining an array of NEMS devices with an electrostatic focusing chip as a second layer to ensure each molecule in the ion beam reaches a NEMS device. As discussed in Chapter 1, conventional MS instruments configured for proteomics rely on LC separation, and this front-end stage leads to significant sample dilution, reducing the overall limit of detection to 10^7-10^8 molecules, even though the final detectors can sense a small number of charges [2]. Avoiding this LC stage, sending molecules directly to NEMS, therefore improves on the status quo by three orders of magnitude with more room for improvement. In principle, this will enable proteomic experiments on extremely small sample sizes, such as that from a single cell.

The work in this thesis relied exclusively on NEMS devices fabricated at LETI. The large quantity of expendable, reliable devices was a necessity for the initial calibration stage, since over a hundred devices were likely destroyed during use in the course of this thesis. This would not have been possible with individually-constructed devices at Caltech's Kavli

Nanoscience Institute; however, our locally-sourced devices have featured significantly improved mass resolution. Specifically, in parallel to the work described here, efforts were undertaken by Dr. Warren Fon and Jarvis Li, a graduate student in our group, to obtain further improvements in mass resolution. Following the completion of this work, they have developed a device providing 5 kDa mass resolution. Additional efforts are now in progress to develop superconducting devices, to implement readout by cavity optomechanics with estimated potential for mass resolution at the single-Dalton level. With the extensive engineering and documentation presented here to obtain robust delivery to NEMS devices, along with continued collaboration with Dr. Makarov and Professor Heck, collection of MS data with such next generation devices should prove straightforward.

In Part II, we introduced two new theories for analyzing frequency-shift data. Chapter 4 develops a Taylor expansion approach that allows for direct extraction of mass moments of the adsorbed molecule from frequency-shift data. This theory is shown to extract the moments with increasing accuracy with smaller particles and was demonstrated using previously published measurements involving droplets deposited on AFM microcantilevers. This theory also greatly facilitated interpretation of the three-mode measurements of GroEL molecules as shown in Chapter 3. Limited signal to noise, along with a possible rotational inertia contribution of the molecules relative to the in-plane device motion, prevented extraction of size or shape information of single molecules. However, the new inertial imaging theory was shown to be very robust at distinguishing when two molecules landed at different locations within the measurement window.

Chapter 5 introduced a new method to directly extract the mass density distribution of adsorbates using regularization. Maximum entropy regularization was found to provide a robust, closed-form, unique solution and accurate image reconstruction given a series of

frequency-shift data for a large number modes. The technique was also shown to work with small adsorbates on the order of 0.1% of the beam length, thus showing molecular shape information could be reconstructed that was independent of modeshape wavelength. Additionally, the accuracy of mass moments of the reconstructed density distributions were found to be similar to that of the expansion theory from Chapter 4 as well as previously published theory [3], with some improvement noticed for larger molecules (>10% of the size of the beam). The accuracy approached 100% in the limit of small molecules.

The new theory and simulations shown in Chapter 5 were not usable with the GroEL data collected in Chapter 3 due to the poor signal to noise, but as the new instrumentation efforts discussed above come online, mass density reconstruction should become possible for single molecules. In this case, four modes could be used to determine whether each molecule has asymmetry. Five modes could be used to distinguish whether the molecule has one or two distinct molecular domains. Seven modes could be used to distinguish between one, two, or three domains. Such simulations provide reason for optimism for a transformative contribution of NEMS to the field of mass spectrometry.

Bibliography

1. Cole, R.B., *Electrospray and MALDI mass spectrometry: Fundamentals, instrumentation, practicalities, and biological applications*. 2011: John Wiley & Sons.
2. Rose, R.J., et al., *High-sensitivity Orbitrap mass analysis of intact macromolecular assemblies.* Nature Methods, 2012. **9**(11): p. 1084-1086.
3. Hanay, M.S., et al., *Inertial imaging with nanomechanical systems.* Nature nanotechnology, 2015. **10**(4): p. 339-344.

Appendix A

LINEAR QUADRUPOLE TRANSMISSION

The NEMS transfer quadrupoles are constructed using two pairs of circular rods, with the two pairs operated with a DC and RF voltage U and V added to or subtracted from an overall DC potential Φ_0 (Figure 91a). Transmission of ions through these quadrupoles will occur for a range of *m/z* values based on the values of U, V, and the RF frequency ω. The equation of motion for an ion entering a linear quadrupole is given by [1]:

$$\begin{matrix} \frac{d^2x}{d\tau^2} + (a_x + 2q_x \cos 2\tau)x = 0 \\ \frac{d^2y}{d\tau^2} + \left(a_y + 2q_y \cos 2\tau\right)y = 0 \end{matrix}, \qquad \text{(A-1)}$$

where the dimensionless parameters a and q are related to ion *m/z* value, quadrupole rod spacing from the central axis r_0, as well as U, V, and ω:

$$a_x = -a_y = \frac{4U}{r_0^2\omega^2(m/z)}, \qquad q_x = -q_y = \frac{2V}{r_0^2\omega^2(m/z)}, \qquad \tau = \frac{\omega t}{2}. \qquad \text{(A-2)}$$

Equations A-1 are Mathieu differential equations. The amplitude of oscillation is proportional to $|P(a,q)^{-1}|$, where $P(a,q)$ is given by

$$P(a,q) = -C(a,q,0)\dot{S}(a,q,0), \qquad \text{(A-3)}$$

and $C(a,q,\tau)$ and $S(a,q,\tau)$ refer to the even and odd Mathieu functions, respectively [2]. Ion trajectory stability is necessary in both dimensions and can be roughly associated with

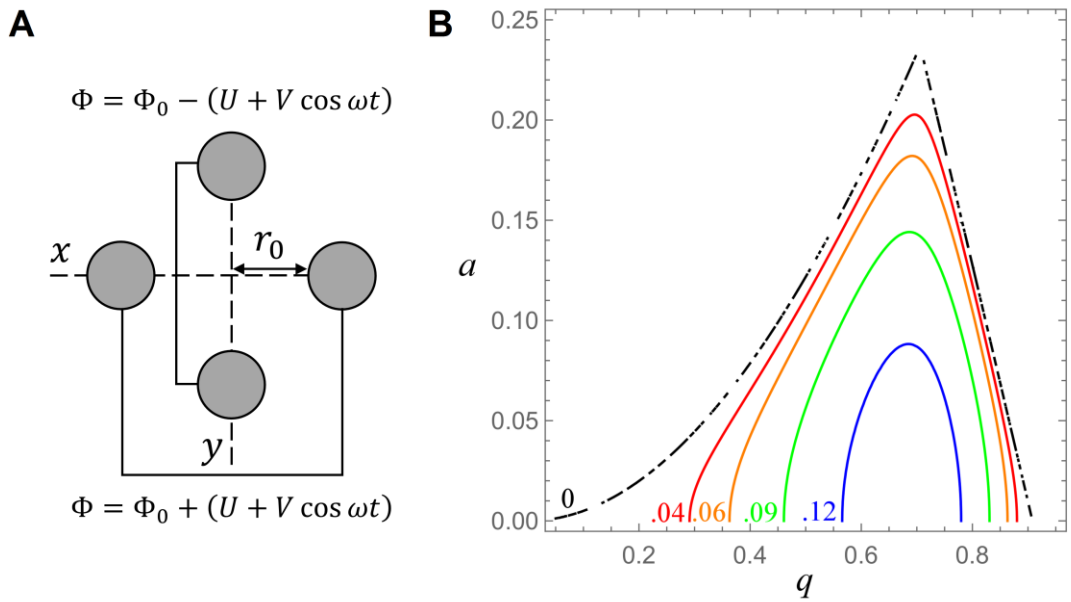


Figure 91. Linear quadrupole geometry and stability diagram. (A) A linear quadrupole can be constructed using two pairs of circular rods, energized with equal and opposite DC and RF voltages superimposed on an overall DC value. Adapted from [1]. **(B)** Contour plot of ion transmission. Ions will be transmitted if dimensionless values a and q lie inside the outer contour. The figure is symmetric with respect to both the a and q axes.

the quantity $P(a,q)P(-a,-q)$. Larger values of this product correspond with smaller ion amplitudes in both dimensions, and the stability boundary can be found by setting $P(a,q)P(-a,-q) = 0$. A contour plot of $P(a,q)P(-a,-q)$ is shown in Figure 91b. Other stability regions with larger values of a and q exist, but would require much higher voltages for the ions of interest.

From the figure, we see that stable ion trajectories can be obtained by setting the DC voltage U to zero ($a = 0$), provided $0 < q < 0.9$, and in particular, peak transmission occurs at roughly $q = 0.7$. The power supply is set to operate at 650 kHz by design and this frequency is not easily adjusted. The spacing between quadrupole rods and the central axis is given by $r_0 = 2.25$ mm. These values can be used to find the optimal RF voltage for an ion with a given *m/z*. Sample values for common ions are given in Table 18. The maximum

Table 18. Calculated RF voltage for optimal quadrupole transmission. RF voltage is calculated to achieve $q = 0.7$ given the *m/z* value for each ion using Equation A-2.

Ion	m/z	RF voltage for optimal transmission
Calmix component (caffeine)	195	60V
Myoglobin	2200	680V
GroEL	11500	3.5 kV

voltage for the power supply is 600 V, so that voltage must be used even if the optimal voltage is higher.

Bibliography

1. Gross, J.H., *Mass spectrometry: A textbook.* Springer, 2004. **16**: p. 29.
2. Baranov, V.I., *Analytical approach for description of ion motion in quadrupole mass spectrometer.* Journal of the American Society for Mass Spectrometry, 2003. **14**(8): p. 818-824.

Appendix B

MODE SHAPE CORRECTION FOR INERTIAL IMAGING

Recall from Chapter 1 that the mode shapes are assumed to remain unchanged with each adsorbed molecule. This appendix calculates the effect that a particle with mass density $\mu_{ad}(x)$ has on the NEMS resonator's mode shapes in order to ascertain under what conditions that assumption remains valid. The new mode shapes can be calculated alongside the new resonance frequency using perturbation theory. This approach was taken in [1] with a point particle; here, the derivation is extended to an arbitrary mass density.

Consider an Euler-Bernoulli beam (or cantilever) with mass density μ_0, Young's modulus E, and second moment of area I. The equation of motion for the nth flexural mode is given by

$$EI\frac{d^4\phi_{n,0}(x)}{dx^4} - \mu_0\omega_{n,0}^2\phi_{n,0}(x) = 0, \tag{B-1}$$

where $\phi_{n,0}(x)$ and $\omega_{n,0}$ refer to the unperturbed modeshape and resonant frequency of the nth mode. Upon adding a flat adsorbate with linear mass density $\epsilon\mu_{ad}(x)$, both the resonant frequency and the modeshape of the beam will change:

$$EI\frac{d^4\phi_n(x)}{dx^4} - (\mu_0 + \epsilon\mu_{ad})\omega_n^2\phi_n(x) = 0, \tag{B-2}$$

where the mass density profile of the adsorbate has been expressed as a product of a dimensionless parameter ϵ that we assume to be small, and a function $\mu_{ad}(x)$ describing the spatial variation of added mass. Written this way, we can express the new resonant frequency and modeshape as a perturbation series involving powers of ϵ:

$$\omega_n^2(x) = \omega_{n,0}^2 + \epsilon\omega_{n,1}^2 + \epsilon^2\omega_{n,2} + O(\epsilon^3), \tag{B-3}$$

$$\phi_n(x) = \phi_{n,0}(x) + \epsilon\sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0}(x) + \epsilon^2\sum_{m\neq n}^{\infty} c_{nm}^{(2)}\phi_{m,0}(x) + O(\epsilon^3). \tag{B-4}$$

Note that the correction to each modeshape only includes summations over other modeshapes ($m \neq n$). This is because the self-corrective term $m = n$ can be normalized out. To see this, note that we can instead write the full correction:

$$\phi_n(x) = \left(1 + \epsilon c_{nn} + \epsilon^2 c_{nn}^{(2)} + O(\epsilon^3)\right)\phi_{n,0}(x) + \epsilon\sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0}(x) + \epsilon^2\sum_{m\neq n}^{\infty} c_{nm}^{(2)}\phi_{m,0}(x) + O(\epsilon^3). \tag{B-5}$$

Inserting into equation B-2, we can divide through by $\left(1 + \epsilon c_{nn} + \epsilon^2 c_{nn}^{(2)} + O(\epsilon^3)\right)$ and absorb this factor into the other coefficients $c_{nm}, m \neq n$. Substituting both perturbation series B-3 and B-4 into B-2 allows us to group together like powers of ϵ. The first order equation yields

$$EI\sum_{m\neq n}^{\infty} c_{nm}\frac{d^4\phi_{m,0}}{dx^4} - \mu_{ad}\omega_{n,0}^2\phi_{n,0} - \mu_0\omega_{n,1}^2\phi_{n,0} - \mu_0\omega_{n,0}^2\sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0} = 0, \tag{B-6}$$

which can be simplified by using the unperturbed equation B-1 to express the fourth order derivative of the modeshapes in terms of the modeshapes themselves:

$$\mu_0\sum_{m\neq n}^{\infty} c_{nm}\left(\omega_{m,0}^2 - \omega_{n,0}^2\right)\phi_{m,0} - \left(\mu_{ad}\omega_{n,0}^2 - \mu_0\omega_{n,1}^2\right)\phi_{n,0} = 0. \tag{B-7}$$

To obtain the first order correction to the resonant frequency, we multiply this equation for the nth mode by $\phi_{n,0}(x)$ and integrate over the length of the beam. Due to the orthogonality of the modeshapes, terms involving $\int_0^L \phi_{m,0}(x)\phi_{n,0}(x)dx\,, m \neq n$ vanish and we obtain:

$$\omega_{n,1}^2 = -\frac{\omega_{n,0}^2}{M_0}\int_0^L \mu_{ad}(x)\phi_{n,0}^2(x)\,dx, \tag{B-8}$$

where $M_0 = \int_0^L \mu_0 dx$ is the total mass of the beam, and we have used the fact that the original modeshapes can be normalized such that $\int \phi_{n,0}^2(x)\,dx = 1$. Thus, to first order, we obtain the following frequency change induced by the analyte:

$$\frac{\omega_n^2}{\omega_{n,0}^2} - 1 = -\frac{1}{M_0}\int_0^L \epsilon\mu_{ad}(x)\phi_{n,0}^2(x)\,dx. \tag{B-9}$$

The coefficients describing the first order correction to the modeshape, c_{nm}, can be obtained by multiplying equation B-7 by $\phi_{m,0}(x)$ and integrating over the length of the beam:

$$c_{nm} = \frac{\omega_{n,0}^2}{\omega_{m,0}^2 - \omega_{n,0}^2}\frac{1}{M_0}\int_0^L \mu_{ad}(x)\phi_{m,0}(x)\phi_{n,0}(x)\,dx. \tag{B-10}$$

Note that the more the resonance frequency for mode m differs from mode n, the smaller the correction, and as $\omega_{m,0} \to \infty$, $c_{nm} \to 0$. Thus, even though the full correction to the modeshape given by equation B-4 involves an infinite sum over all modes, in practice the corrections for a given mode n due to modes $m \gg n$ have an increasingly negligible effect.

The next correction to the resonant frequency induced by analyte adsorption can be obtained by substituting equations B-3 and B-4 into equation B-2 and grouping together powers of ϵ^2:

$$EI\sum_{m\neq n}^{\infty} c_{nm}^{(2)}\frac{d^4\phi_{m,0}}{dx^4} - \mu_0\omega_{n,2}^2\phi_{n,0} - \mu_0\omega_{n,1}^2\sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0}$$

$$-\mu_0\omega_{n,0}^2\sum_{m\neq n}^{\infty} c_{nm}^{(2)}\phi_{m,0} - \mu_{ad}\omega_{n,1}^2\phi_{n,0} - \mu_{ad}\omega_{n,0}^2\sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0} = 0. \tag{B-11}$$

This equation can again be simplified by using the unperturbed equation B-1 to eliminate terms involving fourth order derivatives:

$$\mu_0 \sum_{m\neq n}^{\infty} c_{nm}^{(2)}\left(\omega_{m,0}^2 - \omega_{n,0}^2\right)\phi_{m,0} - \mu_0\omega_{n,2}^2\phi_{n,0} - \mu_0\omega_{n,1}^2 \sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0}$$

$$- \mu_{ad}\omega_{n,1}^2\phi_{n,0} - \mu_{ad}\omega_{n,0}^2 \sum_{m\neq n}^{\infty} c_{nm}\phi_{m,0} = 0. \quad \text{(B-12)}$$

Multiplying by $\phi_{n,0}(x)$ and integrating over the length of the beam:

$$\omega_{n,2}^2 = -\frac{\omega_{n,1}^2}{M_0}\int_0^L \mu_{ad}(x)\phi_{n,0}^2(x)\,dx$$

$$-\frac{\omega_{n,0}^2}{M_0^2}\sum_{m\neq n}^{\infty}\frac{\omega_{n,0}^2}{\omega_{m,0}^2 - \omega_{n,0}^2}\left(\int_0^L \mu_{ad}(x)\phi_{m,0}(x)\phi_{n,0}(x)\,dx\right)^2, \quad \text{(B-13)}$$

from which we obtain the frequency change induced by the analyte:

$$\frac{\omega_n^2 - \omega_{n,0}^2}{\omega_{n,0}^2} = -\frac{1}{M_0}\int_0^L \epsilon\mu_{ad}(x)\phi_{n,0}^2(x)dx + \left(\frac{1}{M_0}\right)^2\left(\int_0^L \epsilon\mu_{ad}(x)\phi_{n,0}^2(x)dx\right)^2$$

$$-\left(\frac{1}{M_0}\right)^2\sum_{m\neq n}^{\infty}\frac{\omega_{n,0}^2}{\omega_{m,0}^2 - \omega_{n,0}^2}\left(\int_0^L \epsilon\mu_{ad}(x)\phi_{m,0}(x)\phi_{n,0}(x)dx\right)^2. \quad \text{(B-14)}$$

For an analyte with significant spatial extents, we can estimate the relative contribution to the frequency change from the modeshape perturbation and the higher mass moments of the adsorbate. This can be done by performing a Taylor expansion of the squared modeshapes about the particle's center of mass, similar to [2]:

$$\int_0^L \epsilon\mu_{ad}(x)\phi_{n,0}^2(x)dx$$

$$= \left[\phi_{n,0}^2(\bar{x})\int_0^L \epsilon\mu_{ad}(x)dx + \left.\frac{\partial^2\phi_{n,0}^2}{\partial x^2}\right|_{\bar{x}}\int_0^L \epsilon\mu_{ad}(x)\frac{(x-\bar{x})^2}{2}dx + \cdots\right]. \quad \text{(B-15)}$$

Note that the integral involving the first moment vanishes above:

$$\int_0^L \epsilon\mu_{ad}(x)(x-\bar{x})dx = 0$$

because we have defined

$$\bar{x} = \frac{\int_0^L \epsilon\mu_{ad}(x)xdx}{\int_0^L \epsilon\mu_{ad}(x)dx}$$

to be the center of mass of the particle. The integral can be represented as follows:

$$\int_0^L \epsilon\mu_{ad}(x)\phi_{n,0}^2(x)dx = m_{ad}\left[\phi_{n,0}^2(\bar{x}) + \left.\frac{\partial^2\phi_{n,0}^2}{\partial x^2}\right|_{\bar{x}}\langle x^2\rangle + \cdots\right], \tag{B-16}$$

with each higher order moment having an increasingly negligible effect. Substituting the first two terms into equation B-14, we obtain

$$\frac{\omega_n^2-\omega_{n,0}^2}{\omega_{n,0}^2} = -\frac{m_{ad}}{M_0}\phi_{n,0}^2(\bar{x}) - \frac{1}{2}\frac{m_{ad}}{M_0}\left.\frac{\partial^2\phi_{n,0}^2}{\partial x^2}\right|_{\bar{x}}\langle x^2\rangle + \left(\frac{m_{ad}}{M_0}\right)^2\phi_{n,0}^4(\bar{x})$$

$$-\left(\frac{m_{ad}}{M_0}\right)^2\sum_{m\neq n}^{\infty}\frac{\omega_{n,0}^2}{\omega_{m,0}^2-\omega_{n,0}^2}\phi_{m,0}^2(\bar{x})\phi_{n,0}^2(\bar{x}) + \cdots \tag{B-17}$$

Comparing with the nonlinear fit formula proposed in [2] to estimate the variance of a distribution from a set of frequency shifts:

$$\frac{\omega_n^2-\omega_{n,0}^2}{\omega_{n,0}^2} = -\frac{m_{ad}}{M_0}\left[\phi_{n,0}^2(\bar{x}) + \frac{1}{2}\left.\frac{\partial^2\phi_{n,0}^2}{\partial x^2}\right|_{\bar{x}}\langle x^2\rangle + \cdots\right]. \tag{B-18}$$

We see that this fit is likely to be inaccurate without accounting for the higher-order corrections to the frequency shift formula obtained via perturbation analysis unless

$$\frac{m_{ad}}{M_0}\phi_{n,0}^4(\bar{x}) \ll \frac{1}{2}\left.\frac{\partial^2\phi_{n,0}^2}{\partial x^2}\right|_{\bar{x}}\langle x^2\rangle. \tag{B-19}$$

Since each derivative of the squared modeshape introduces a factor of L (the length of the beam), the nonlinear fit formula via B-18 used to determine the variance is likely to yield inaccurate results unless

$$\frac{m_{ad}}{\langle x^2 \rangle} \ll \frac{M_0}{2L^2}. \tag{B-20}$$

In other words, the linear mass density of the particle must be substantially less than half the beam's linear mass density. Accounting for height, we obtain (approximately):

$$\rho_{ad} h_{ad} \ll \rho_0 h. \tag{B-21}$$

A biomolecule has a density close to water (1 g/cm^3), which is of the same order of magnitude as silicon at 2.32 g/cm^3. We see that accurately measuring a molecule's variance (size) will be possible provided the adsorbate is not too tall compared with the beam. For the measurements used in this thesis, $h_{ad} \approx$ 15 nm and $h =$160 nm, so there is no issue. However, if future studies use ultra-thin devices or much larger molecules relative to the device, Equation B-17 would be needed to accurately calculate the adsorbate's mass and variance.

Bibliography

1. Yüksel, M. and M.S. Hanay, *Mode-shape perturbation induced by analyte adsorption in nanomechanical sensors.* arXiv preprint arXiv:1709.01496, 2017.
2. Sader, J.E., et al., *Mass spectrometry using nanomechanical systems: Beyond the point-mass approximation.* Nano Letters, 2018. **18**(3): p. 1608-1614.