Stochastic noise in atomic force microscopy

Aleksander Labuda,[1] Martin Lysy,[2] William Paul,[1] Yoichi Miyahara,[1] Peter Grütter,[1] Roland Bennewitz,[3] and Mark Sutton[1]
[1]*Department of Physics, McGill University, Montreal, Quebec, Canada H3A 2T8*
[2]*Department of Statistics, Harvard University, Cambridge, Massachusetts 02138-2901, USA*
[3]*INM—Leibniz Institute for New Materials, 66123 Saarbrücken, Germany*
(Received 2 May 2012; published 5 September 2012)

Having reached the quantum and thermodynamic limits of detection, atomic force microscopy (AFM) experiments are routinely being performed at the fundamental limit of signal to noise. A critical understanding of the statistical properties of noise leads to more accurate interpretation of data, optimization of experimental protocols, advancements in instrumentation, and new measurement techniques. Furthermore, accurate simulation of cantilever dynamics requires knowledge of stochastic behavior of the system, as stochastic noise may exceed the deterministic signals of interest, and even dominate the outcome of an experiment. In this article, the power spectral density (PSD), used to quantify stationary stochastic processes, is introduced in the context of a thorough noise analysis of the light source used to detect cantilever deflections. The statistical properties of PSDs are then outlined for various stationary, nonstationary, and deterministic noise sources in the context of AFM experiments. Following these developments, a method for integrating PSDs to provide an accurate standard deviation of linear measurements is described. Lastly, a method for simulating stochastic Gaussian noise from any arbitrary power spectral density is presented. The result demonstrates that mechanical vibrations of the AFM can cause a logarithmic velocity dependence of friction and induce multiple slip events in the atomic stick-slip process, as well as predicts an artifactual temperature dependence of friction measured by AFM.

DOI: 10.1103/PhysRevE.86.031104 PACS number(s): 05.40.Ca, 81.40.Pq

I. INTRODUCTION

In the original implementation of the atomic force microscope in 1986 [1], nanoscale forces acting on a sharp tip were inferred by sensing the "static" deflection of the cantilever to which the tip was tethered. Soon after, "dynamic" methods [2,3] were implemented where the cantilever was oscillated near or at its resonance frequency; this circumvented many of the detection noise issues that occur at low frequencies in the static case by moving the relevant noise bandwidth into the kilohertz range. Today, both of these atomic force microscopy (AFM) methods serve the nanoscience community and have led experiments up to the boundaries imposed by thermodynamic and quantum limits [4] of physics.

At room temperature, stochastic thermal noise dominates the low signal-to-noise regime of nanoscale experiments. If thermal noise is well understood, it can even be exploited to extract information about the tip-sample interaction [5,6]. Importantly, the signal of interest in the AFM experiment itself might be stochastic in nature, and understanding its statistical properties can lead to new measurement techniques [7,8]. In fact, measuring the variance—rather than the mean—of a physical parameter can yield higher signal-to-noise ratio in nanoscale experiments where fluctuations dominate the signal of interest [9].

Inevitably, instrumental sources of noise are also present in AFM. It is therefore imperative to have a good understanding of all the sources of stochastic noise in order to properly interpret the results of AFM experiments in the presence of both signal and noise. Furthermore, understanding the sources of noise naturally leads to improvement of future instrumental design, and serves in the optimization of experimental protocols.

Given the growing complexity of AFM techniques, numerical simulations of AFM experiments help to understand the effects that instrumental parameters [10,11] and complex cantilever dynamics [12] have on the acquired signals. So far, virtual AFM simulators, such as VEDA [13], are optimized for deterministic calculations of cantilever dynamics. The next natural step is the inclusion of stochastic cantilever dynamics, as well as stochastic vibrations inherent to the instrument and colored detection noise, in order to more accurately reproduce true AFM experiments.

Therefore the goal of the statistical noise analysis presented in this paper is to provide a framework for the critical understanding of variability in data acquired by AFM, to provide AFM design guidelines for the minimization of noise in different types of experiments, to help optimize available parameters when constructing experimental protocols, and to establish statistical foundations for the simulation of stochastic noise in AFM.

The next section provides an overview of stochastic noise sources in AFM. Afterwards, a case study friction experiment is presented that will be referred to throughout the discussion of the four core sections. The first section describes the detection noise sources in AFM and how they scale with key experimental parameters. The second outlines the statistical properties of different types of noise sources in relation to their power spectral densities (PSDs). This leads to a method for integrating the PSDs to obtain quantitatively accurate estimates of the variance of linear measurements, presented in the third section. For nonlinear measurements, the fourth core section demonstrates the utility of a method for stochastic simulation of noise in AFM.

II. OVERVIEW OF STOCHASTIC NOISE IN AFM

In a recent article [14], the noise sources of AFM were divided into three categories: detection noise, force noise, and displacement noise.

Optical shot noise sets the fundamental limit of detection of the optical beam deflection method [15], and its reduction has played a primary role in enabling atomic resolution imaging with dynamic AFM in liquids [16], for example. Detection noise can have a measurable impact on the tip-sample physics and the outcome of an experiment by inducing unwanted vibrations when using feedback to track the tip-sample separation. Conversely to dynamic AFM, the detection noise in static AFM is far from white and can have an elaborate spectral distribution and be characterized by the enigmatic $1/f$ noise [17,18], power-line noise, various mechanical resonances, etc. Noise analysis in this situation becomes nontrivial because of the complicated spectral distribution of densities, which are poorly described by parametric physical models and must generally be estimated nonparametrically.

Force noise is mainly caused by thermal fluctuations which set a fundamental lower bound to the force fluctuations that the tip imparts on a sample during a measurement. Force noise can have a sizable impact on the outcome of an AFM experiment; for example, thermal fluctuations of the cantilever directly affect the maximal force observed in stick-slip measurements [19]. Thermal noise has a theoretical lower limit set by thermoelastic damping in vacuum environments [20–22]. In ambient conditions, it correlates with viscous damping [23–25] and can completely dominate detection noise in highly viscous environments [26]. Recently, the low-frequency fluctuations of cantilevers have been under scrutiny: structural damping causes a $1/f$ power spectral density in the cantilever at low frequencies where viscous forces are negligible [27]; $1/f$ noise can also be induced by the coupling of light power fluctuations to cantilever bending by thermally induced stress [14].

On the other hand, the sources of displacement noise in AFM are usually not of fundamental interest because they relate to instrument design and AFM engineering [28]. Nevertheless, their existence should not be disregarded: vibrations between the tip and the sample can have profound impact on the results of an AFM experiment, even if they lie outside of the measurement bandwidth. For example, deliberate tip-sample vibrations can enable dynamic superlubricity [29,30] and eliminate wear [31] in friction experiments. By the same token, undesirable vibrations of the tip-sample junction can affect the results of an AFM experiment well beyond simple deterioration of image quality.

III. CASE STUDY

This section presents a static AFM experiment which will serve as an example to guide the noise analysis of the following sections. A friction measurement was chosen as the benchmark because it represents a simple example of a linear measurement in static AFM.

Friction force microscopy [32] in liquid environments has recently reached atomic-scale resolution [33,34] with clear measurements of stick-slip tip motion acquired during the electrochemically activated transition between a copper chloride monolayer and Au(111), as shown in the image of Fig. 1. The lateral force curves in Fig. 1(a) demonstrate that roughly 2 eV of energy is elastically stored and released for every stick-slip cycle. However, roughly 50 meV is dissipated

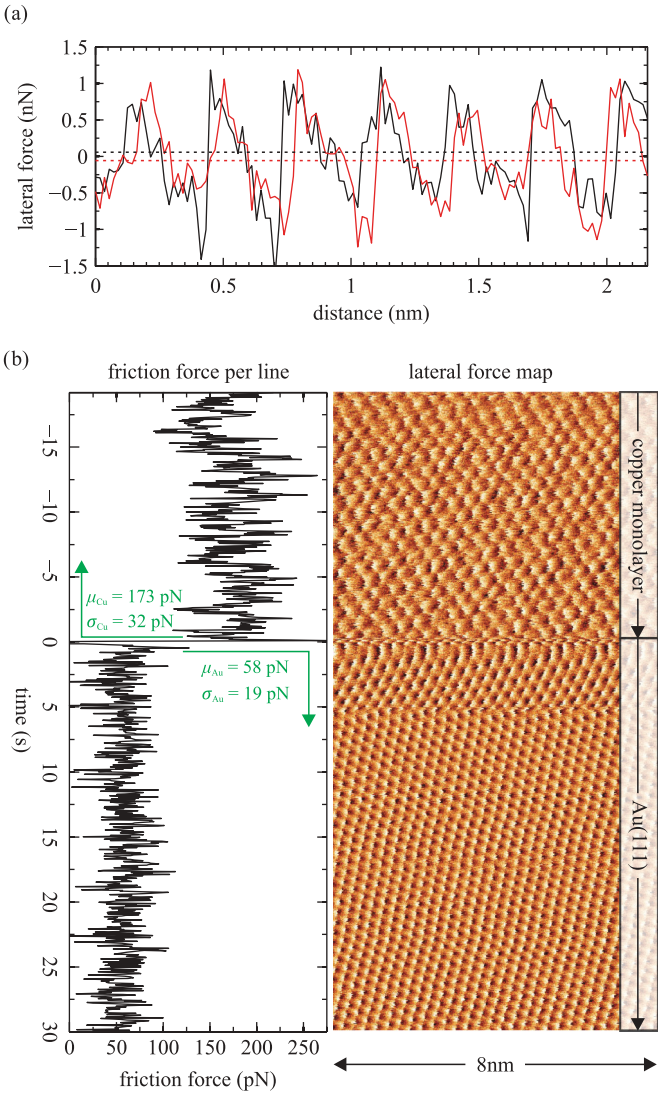

FIG. 1. (Color online) (a) A backward and forward lateral force scan is performed on Au(111), with a scan rate of 25 lines/s which corresponds to ~400 nm/s. The tip scans over the surface with stick-slip motion. Average lateral forces are shown in dotted lines. The area enclosed in the hysteresis loop is the energy lost due to friction. (b) A lateral force map is shown (right) with the calculated friction force per line (left) as a function of time; the copper chloride layer is electrochemically desorbed at $t = 0$ s, resulting in a change in observed lattice and a reduction in both friction average and standard deviation. (The slight kink at $t = 5$ s is caused by the change in image scan direction from downwards to upwards—two consecutive images were merged for this figure.)

during every stick-slip cycle, as determined by measuring the area enclosed by the hysteresis loop. The corresponding average friction force per line is plotted in Fig. 1(b), and demonstrates a clear reduction in friction on Au(111) vs the copper monolayer [33]. In the following sections, attention is focused on the variability in friction which gives rise to the following questions:

(i) Does the variability originate due to detection noise, or is it a true measurable variability in friction?

(ii) What are the origins of detection noise, and can it be reduced?

(iii) Is the true friction variability enhanced by vibrations of the instrument?

IV. NOISE IN OPTICAL BEAM DEFLECTION: THE POWER SPECTRAL DENSITY

This section introduces the "power spectral density" (PSD) in the context of analyzing noise of the optical beam deflection (OBD) method [15]. The PSD, sometimes referred to as the power spectrum, represents a frequency decomposition of the power of a signal. It is particularly suitable for quantifying stationary stochastic processes. The PSD will be central to the remainder of the document, and will first be presented in this section from an experimental point of view.

The OBD method is the most widely used detection method in AFM. Our AFM [33,35] employs a 1-mW fiber-coupled superluminescent diode (SLD). Like any light source, SLDs are afflicted by both angular fluctuations and power fluctuations of their emitted light beam. Understanding how both angular fluctuations and power fluctuations affect AFM measurements permits the optimization of experiments by the reduction of detection limits, and enables accurate interpretation of noise observed in AFM experiments.

In the most common application of the OBD method, the reflected light beam shines onto a split-diode photodetector. The difference in power between both sides of the photodetector is proportional to the angular deflection of the cantilever—to first order. As long as the light beam is perfectly centered on the photodetector, the detection of its position is practically impervious to power fluctuations. The differential nature of the measurement results in a high common-mode rejection ratio: the difference signal remains at zero even if the total optical power fluctuates.

Whereas this is a good approximation for dynamic AFM, static AFM is often performed under varying loads that cause a large nonzero difference signal during the experiment. Consequently, the detection power spectral density $n_\Delta^2(f)$ can vary throughout a static AFM experiment because it is a function of the time-averaged normalized difference signal Δ_0. The noise is minimized when $\Delta_0 = 0$, and increases as $|\Delta_0| > 0$. The generalized form $n_\Delta(f|\Delta_0)$ is measurable and predictable, as is demonstrated in the following section. Note that dynamic AFM can also be prone to the same problems if the light beam is not perfectly centered on the photodetector.

A. Common-mode and differential-mode noise

Consider some irradiance profile $I(\theta')$ of a light beam shining on a split photodetector, where θ' is the angular coordinate in the far field. The integrated power incident on each section of the photodetector, P_A and P_B, is used to generate the normalized difference signal

$$\Delta(t) = \frac{P_B(t) - P_A(t)}{P_0}, \tag{1}$$

which provides a measure of cantilever deflection, and the normalized sum signal

$$\Sigma(t) = \frac{P_A(t) + P_B(t)}{P_0}, \tag{2}$$

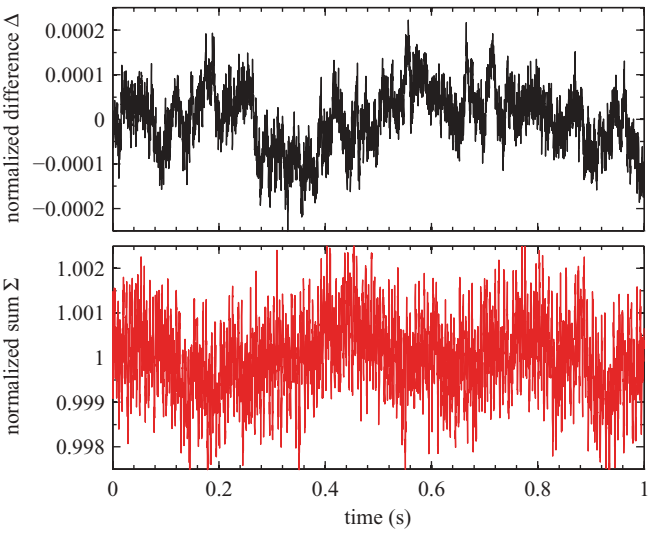

FIG. 2. (Color online) The normalized difference (Δ) and normalized sum (Σ) of our SLD light beam centered on the photodetector. Note that both vertical axes differ by a factor of 10 × in range. The Σ fluctuations are roughly 10 × larger and strongly afflicted by line noise (60 Hz and harmonics). Both signals are uncorrelated. The details of the optical beam deflection system and electronics used for this measurement are outlined elsewhere [33,35].

where P_0 is the total optical power time averaged across the entire experiment.

Figure 2 shows data acquired during 1 s of both signals in the situation where the light beam is centered on the photodetector: $\Delta_0 = 0$, where Δ_0, defined earlier, is the time-averaged value of $\Delta(t)$. In this situation, any instantaneous fluctuation $\delta I(\theta')$ of the irradiance profile can always be mathematically decomposed into two orthogonal components:

$$\delta I(\theta') = \delta I_{\mathrm{CM}} + \delta I_{\mathrm{DM}}. \tag{3}$$

The common-mode irradiance fluctuation δI_{CM} refers to the symmetric component of δI for which the deviations from the time-averaged irradiance $I_0(\theta')$ cause changes in the sum signal $\delta\Sigma$, but no change in the difference signal $\delta\Delta = 0$. Likewise, the differential-mode irradiance fluctuation δI_{DM} is antisymmetric, causes no net change in total optical power ($\delta\Sigma = 0$), and results only in a deviation $\delta\Delta$ which appears as a change in the measured angle. Figure 3 illustrates this decomposition.

Certain noise sources may cause deviations of only δI_{DM} or δI_{CM} or both. For example, shot noise causes the same amount of power fluctuation in both δI_{DM} and δI_{CM} because it is spatially uncorrelated on all length scales. On the other hand, angular movement in the light beam causes only δI_{DM} with no associated changes in total power because the fluctuation is perfectly spatially correlated (loss of light on one detector implies gain of light on the other detector). In the special case of $\Delta_0 = 0$ discussed so far, both of these fluctuations in the irradiance profile can be considered orthogonal, and therefore all fluctuations in $\delta\Sigma$ and $\delta\Delta$ are effectively temporally uncorrelated—assuming that $\delta\Sigma \ll 1$ and $\delta\Delta \ll 1$ such that cross-terms $\delta\Sigma\delta\Delta \approx 0$. Note that both time series in Fig. 2 show no (temporal) correlation.

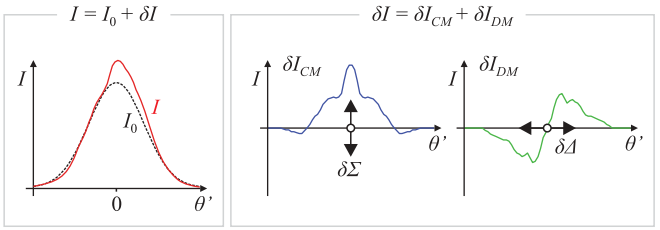

FIG. 3. (Color online) At any point in time, the irradiance profile of the light beam I has a deviation δI from its time-averaged profile I_0. This irradiance is defined with respect to the angular coordinate in the far field of the light beam. This deviation can be mathematically decomposed into two components: δI_{CM} and δI_{DM}. The common-mode fluctuation δI_{CM} is even about the center and causes changes in the sum signal $\delta\Sigma$. The differential-mode fluctuation δI_{DM} is odd about the center and causes changes in the difference signal $\delta\Delta$. Both $\delta\Sigma$ and $\delta\Delta$ are independent when the light beam is centered on the photodetector. Note that these plots are not to scale and that usually $\delta\Sigma \gg \delta\Delta$.

The time dependence of these fluctuations is now investigated. When $\Delta_0 = 0$, the common-mode noise and differential-mode noise can be estimated by measuring $\Sigma(t)$ and $\Delta(t)$, shown in Fig. 2, and calculating their respective average power spectral densities $\langle n_{\mathrm{CM}}^2 \rangle$ and $\langle n_{\mathrm{DM}}^2 \rangle$, as shown in Fig. 4(a). The next subsection describes how to measure a power spectral density.

These two measurements, $\langle n_{\mathrm{CM}}^2 \rangle$ and $\langle n_{\mathrm{DM}}^2 \rangle$, can be used to determine the noise affecting $\Delta(t)$ throughout an AFM experiment at any Δ_0. For $\Delta_0 \neq 0$, common-mode fluctuations couple into the difference signal Δ and cause additional variations $\delta\Delta$. The total noise for any Δ_0 can be calculated by

$$n_\Delta^2(f|\Delta_0) = n_{\mathrm{DM}}^2(f) + \Delta_0^2 n_{\mathrm{CM}}^2(f). \tag{4}$$

This equation holds when the fraction of the light on either photodetector ($\frac{P_A}{P_A+P_B}$) and the total light ($P_A + P_B$) are independent. While this equation is not expected to be strictly true, it can be a good approximation in practice where $n_{\mathrm{DM}}^2 \ll n_{\mathrm{CM}}^2 \ll 1$ as verified in Fig. 4, which compares the predictions of $n_\Delta^2(f|\Delta_0 = 0.057)$ by Eq. 4 with the corresponding empirical measurement.

Note that the electronic noise in the detection system is negligible compared to the optical noise in Fig. 4, which is necessary for Eq. (4) to hold. Appropriate circuit design achieves this condition [33,36].

B. Measuring a power spectral density

A power spectral density is estimated by the squared magnitude of the Fourier transform of a time series. In practice, any measurement of a spectral density is subject to bias, aliasing, spectral leakage, and drift; it must be carefully performed to obtain an accurate estimation of the true spectral density of noise. These problems are described in detail in the Appendix.

Each measurement in Fig. 4(a) was performed by acquiring a time series of the signal in question for 8 min at 200 kHz. A 70-kHz electronic antialiasing filter (eighth order) was used to prevent aliasing, as can be observed by the roll-off in Fig. 4(a). A linear fit to the data was subtracted to account for drift.

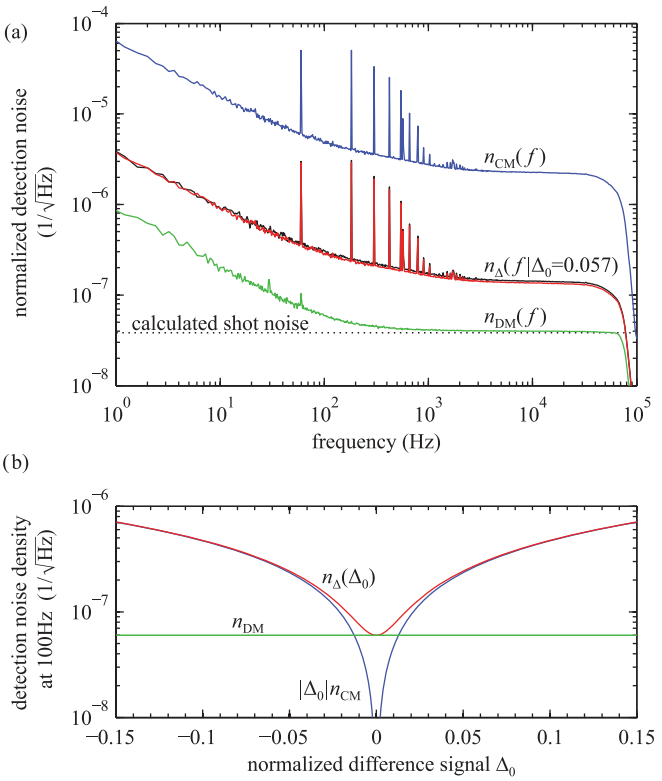

FIG. 4. (Color online) (a) The common-mode n_{CM} and differential-mode n_{DM} noise densities were measured by shining the light beam directly onto the photodetector, as described in Fig. 2. Note that the square root of the power spectral density is plotted here. Both noise densities were used to predict the detection noise density n_Δ for a normalized normal signal $\Delta_0 = 0.057$, using Eq. (4). The n_Δ was also measured after adjusting the photodetector position to achieve $\Delta_0 = 0.057$ experimentally; the measured noise density is overlaid and shows very good agreement with the calculated prediction. A $\Delta_0 = 0.057$ corresponds to a deflection of 150 nm and a force of 15 nN in our particular setup, and causes an increase in detection noise of $\sim 4\times$. The sharp noise peaks at 60 Hz and harmonics are caused by line noise in the SLD current source. The roll-off at 70 kHz was achieved by the use of an antialiasing filter. (b) The detection noise density n_Δ at 100 Hz is calculated from Eq. (4). At $\Delta_0 = 0$, the noise in the difference signal n_Δ is composed of differential-mode noise only. For $|\Delta_0| > 0.015$, common-mode noise dominates detection noise.

The entire time series was then multiplied by a Hann window to reduce spectral leakage and estimation bias of the $1/f$ noise. A Hann window was used as it offers a well-balanced trade-off between frequency resolution and dynamic range for our application. Then, the Fourier transform was performed and its squared magnitude resulted in a single PSD, which is renormalized by the power of the windowing function. Finally, the powers of adjacent frequencies of this PSD were averaged to reduce variance; this increases the signal-to-ratio at the expense of frequency resolution.

C. Discussion

On our AFM, common-mode noise dominates the deflection signal for $|\Delta_0| > 0.015$, which corresponds to deflections of >30 nm and forces >3 nN with a 225-μm-long cantilever

having a 0.1 N/m spring constant. This stresses the importance of centering the light beam on the photodetector before the experiment, such that $\Delta_0 = 0$ corresponds to 0 nN. Typically, experiments that are performed near 0 nN are most prone to detection noise and can benefit from the high common-mode rejection ratio at $\Delta_0 = 0$. Centering the light beam in the lateral direction is equally important for sensitive friction experiments.

The common-mode noise is fundamentally limited by source-induced noise, which exceeds shot noise by nearly ten times for our particular SLD, as calculated theoretically by Ref. [37]. In practice, the measured common-mode noise in Fig. 4 is actually dominated by electronic noise in the current driver which rolls off above 70 kHz, and has the characteristic 60-Hz harmonic noise peaks and the elusive $1/f$ noise at low frequencies. In either case, for drive-current or source-induced noise, experiments performed at $|\Delta_0| > 0.015$ do *not* benefit from an increase in optical power because they are *not* limited by optical shot noise.

Figure 4(b) demonstrates that differential-mode noise dominates the detection noise for $|\Delta_0| < 0.015$, and that shot noise only takes effect for frequencies above a few hundred hertz. Consequently, changing the optical power by either boosting the drive current or coating the cantilever reduces *only* shot noise, such that only fast experiments performed at $|\Delta_0| < 0.015$ benefit from an increase in optical power (an exact quantification of "fast" is the topic of Sec. VI). At low frequencies, the angular fluctuations of the light beam are measurable above shot noise. These fluctuations are spatially uncorrelated and intrinsic to the light beam as it exits the optical fiber; therefore they are independent of the AFM design and depend only on the specific light source.

These observations suggest that discussing the fundamental theoretical limits of OBD noise in static AFM experiments is not necessarily worthwhile, because they might be overshadowed by other noise sources in experimental settings. The differential-mode noise may be limited by $1/f$ noise, rather than shot noise, and the drive electronics may set the common-mode noise limit far above the shot noise or even source-induced noise. It is therefore advisable to empirically measure the common-mode and differential-mode noise for any particular setup and use the modeling summarized by Eq. (4) to make well-informed predictions about detection noise in AFM experiments.

So far, noise has been quantified by its power spectral density. At this stage, it is unclear how all the components of the PSD in Fig. 4 relate to the variability in the measured friction seen in Fig. 1. White noise, $1/f$ noise, and line noise all have unique statistical properties and dominate at different frequencies. The statistical interpretation of a PSD is the topic of the following section.

V. STATISTICS OF THE POWER SPECTRAL DENSITY

This section outlines the statistical properties of different noise sources in AFM, with the goal of performing meaningful noise analysis of linear measurements (Sec. VI) and efficient stochastic noise simulations (Sec. VII), using the power spectral density (PSD).

The statistical analysis in this section relates to stationary noise, defined as a stochastic process with an underlying probability distribution that does not vary in time. The covariance between any two outcomes of such a process depends only on the relative time separating their measurement, and is independent of absolute time. The mean must remain constant.

In contrast, nonstationary noise, such as vibrations caused by sporadic slamming of a door or intermittent discussions in the laboratory, does not constitute a fundamental detection limit in AFM and will be disregarded. Also, $1/f$ noise is nonstationary, but it can be approximated as stationary given certain conditions that will be discussed later. Linear drift is neither stochastic nor stationary on any time scale, and therefore cannot be described by a PSD; it should always be subtracted from the time signal before calculating a PSD (see Appendix for details).

Most fundamental noise sources in AFM can be approximated as stationary with high accuracy, which will be assumed henceforth.

A. Noise in the time domain

Let $\boldsymbol{x_t} = [\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_m]'$ be a $m \times 1$ vector of random variables, acquired on an equally spaced time-domain basis $t = [t_1, t_2 \ldots t_m]$. This is the form in which data are acquired by an analog-to-digital converter (ADC) card used in most modern data acquisition systems. The data vector $\boldsymbol{x_t}$ is always corrupted by stochastic noise, and can therefore be expressed as the sum of deterministic (d_t) and random ($\boldsymbol{\epsilon_t}$) components:

$$\boldsymbol{x_t} = d_t + \boldsymbol{\epsilon_t}, \tag{5}$$

where the "t" subscript emphasizes that these vectors are defined in the time domain. Note that vectors in bold represent vectors of random variables which have some associated probability density function, while d_t represents a nonprobabilistic data vector. Further in this section, $\boldsymbol{\epsilon_t}$ will be used to represent a single measurement drawn from the probability distribution of $\boldsymbol{\epsilon_t}$.

The distinction between d_t and $\boldsymbol{\epsilon_t}$ is a matter of perspective: for example, stochastic thermal fluctuations of the cantilever can be considered noise ($\boldsymbol{\epsilon_t}$) or signal (d_t) in different situations. If an experiment seeks to determine a hydration layer structure by its effect on the thermal motion of the cantilever [5], this thermal motion can be considered a desirable experiment signal corrupted by some stochastic detection noise. On the other hand, the measurement of forces exerted by a single titin molecule tethered to the cantilever tip [38] is clearly corrupted by undesirable stochastic thermal noise of that cantilever.

Assuming the noise $\boldsymbol{\epsilon_t}$ has zero mean, the measurement of interest d_t is equal to the expectation value of $\boldsymbol{x_t}$, which has some unwanted variability characterized by a temporal covariance matrix V_t:

$$V_t = \mathrm{Var}(\boldsymbol{x_t}) = \mathrm{E}(\boldsymbol{\epsilon_t} \cdot \boldsymbol{\epsilon_t}^{\dagger}) = \sigma^2 \begin{bmatrix} 1 & \rho_2 & \rho_3 & \ldots \\ \rho_2 & 1 & \rho_2 & \ldots \\ \rho_3 & \rho_2 & \ddots & \ldots \\ \vdots & \vdots & \vdots & 1 \end{bmatrix}, \tag{6}$$

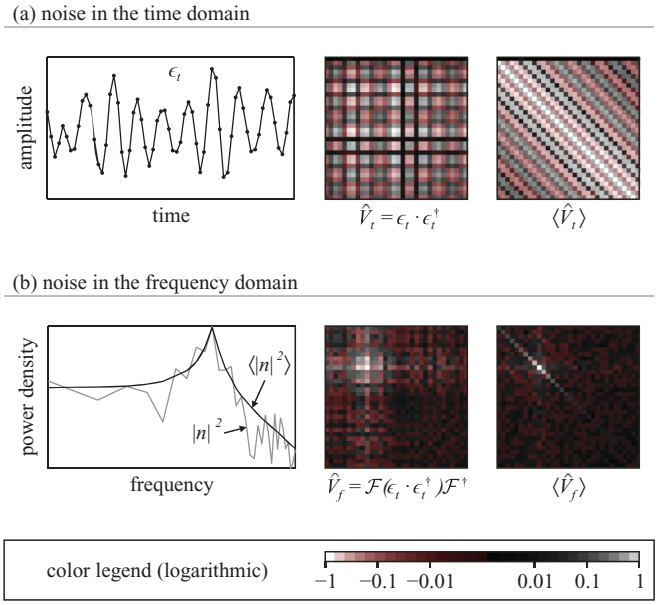

FIG. 5. (Color online) A thermally driven harmonic oscillator with $Q = 10$ is used as an example to illustrate the key differences between performing noise analysis in the time and frequency domains. (a) A single time series is shown, with a corresponding temporal power matrix $\hat{V}_t$, many of which can be averaged $\langle \hat{V}_t \rangle$ to estimate the temporal covariance matrix V_t. The latter has significant off-diagonal elements due to correlations between data points in the time domain. (b) A single power spectral density n^2 and the average $\langle n^2 \rangle$ of 1000 is shown, with the corresponding spectral power matrix $\hat{V}_f$ and its average $\langle \hat{V}_f \rangle$ which is asymptotically diagonal because the noise is stationary. For simplicity, only the real part of the first quadrant of the matrices is shown. $\langle \hat{V}_t \rangle$ and $\langle \hat{V}_f \rangle$ were averaged 1000 times.

where "†" represents the conjugate transpose, while Var and E are the variance and expectation value operators. Note that V_t is a Toeplitz matrix (identical elements across each diagonal) which is required for the time invariance property of stationary noise.

The matrix V_t is a theoretical quantity which can only be *estimated* experimentally. A single outcome $\boldsymbol{\epsilon_t}$ drawn from the probability distribution of $\boldsymbol{\epsilon_t}$ results in a temporal power matrix defined as $\hat{V}_t = \boldsymbol{\epsilon_t} \cdot \boldsymbol{\epsilon_t}^\dagger$. The average $\langle \hat{V}_t \rangle$ asymptotically converges to the temporal covariance matrix V_t. This is illustrated in Fig. 5(a) for a thermally driven harmonic oscillator.

In the special case of white noise, statistics are straightforward: all the off-diagonal elements of V_t are zero because every data point in $\boldsymbol{\epsilon_t}$ is uncorrelated. For nonwhite noise, the time-domain random variables are correlated, with associated off-diagonal covariance elements in V_t, as seen in Fig. 5(a). This makes the time domain a poor basis for performing noise analysis.

For example, a single friction measurement in Fig. 1 is a linear combination of the 1024 data points recorded in a lateral force loop: average forward lateral force minus average backward lateral force. In the case of white noise, the standard deviation of friction is simply the standard deviation of lateral force divided by $\sqrt{1024}$. However, in the presence of pure $1/f^2$ noise, for example, this approach is wrong and would underestimate the standard deviation of friction by nearly $8\times$. This error stems from the strong correlation of noise in the time domain; in short, covariances cannot be neglected if the noise is not white.

In the next section, variances will be decorrelated by transferring the noise analysis into the frequency domain.

B. Noise in the frequency domain

The signal $\boldsymbol{x_t}$ in the time domain can be mapped into the frequency domain by the linear transformation known as the discrete Fourier transform:

$$\boldsymbol{x_f} = \mathcal{F}\boldsymbol{x_t}, \tag{7}$$

where $\mathcal{F}$ is the complex-valued Fourier matrix. If the units of $\boldsymbol{x_t}$ are V, then the units of $\boldsymbol{x_f}$ are $V/\sqrt{Hz}$. The decomposition of the frequency-domain measurement $\boldsymbol{x_f}$ into signal and noise components remains valid due to linearity, and follows from Eq. (5) as

$$\boldsymbol{x_f} = \mathcal{F}d_t + \mathcal{F}\boldsymbol{\epsilon_t} = d_f + \boldsymbol{\epsilon_f}. \tag{8}$$

Although it might be more difficult to interpret the frequency-domain signal d_f, the interpretation of stationary noise $\boldsymbol{\epsilon_f}$ becomes much simpler because the corresponding spectral covariance matrix V_f is (asymptotically[1]) diagonalizable:

$$V_f = \mathrm{Var}(\boldsymbol{x_f}) = \mathrm{E}(\boldsymbol{\epsilon_f} \cdot \boldsymbol{\epsilon_f}^\dagger) = \begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_m^2 \end{bmatrix}. \tag{9}$$

In essence, the frequency domain represents the orthogonal basis of stationary noise [39], which accurately represents most noise sources in AFM. Because the random variables composing $\boldsymbol{x_f}$ are uncorrelated, their variances can be added without consideration of any off-diagonal terms. This makes the frequency domain an attractive basis for noise variance analysis. In analogy to the time domain, the spectral covariance matrix V_f is experimentally estimated by averaging the spectral power matrix $\hat{V}_f = \boldsymbol{\epsilon_f} \cdot \boldsymbol{\epsilon_f}^\dagger$. This is illustrated in Fig. 5(b), where $\langle \hat{V}_f \rangle$ is shown to converge to the diagonal V_f.

The power spectral density is the diagonal of the spectral covariance matrix V_f, which provides a complete statistical description of any stationary Gaussian noise. If the noise is stationary but non-Gaussian, further assumptions are necessary for a full characterization. However, meaningful noise variance analysis can still be performed on non-Gaussian stationary noise using the power spectral density, as will be shown in the next subsection.

[1]The term "asymptotically diagonalizable" refers to the fact that there is a true diagonal V_f defined in continuous time that can be approximated in discrete time to arbitrary accuracy as $m \to \infty$. In experimental AFM settings, where m is usually in the hundreds or even millions, this approximation is often overshadowed by the finite precision in estimating discrete time V_f.

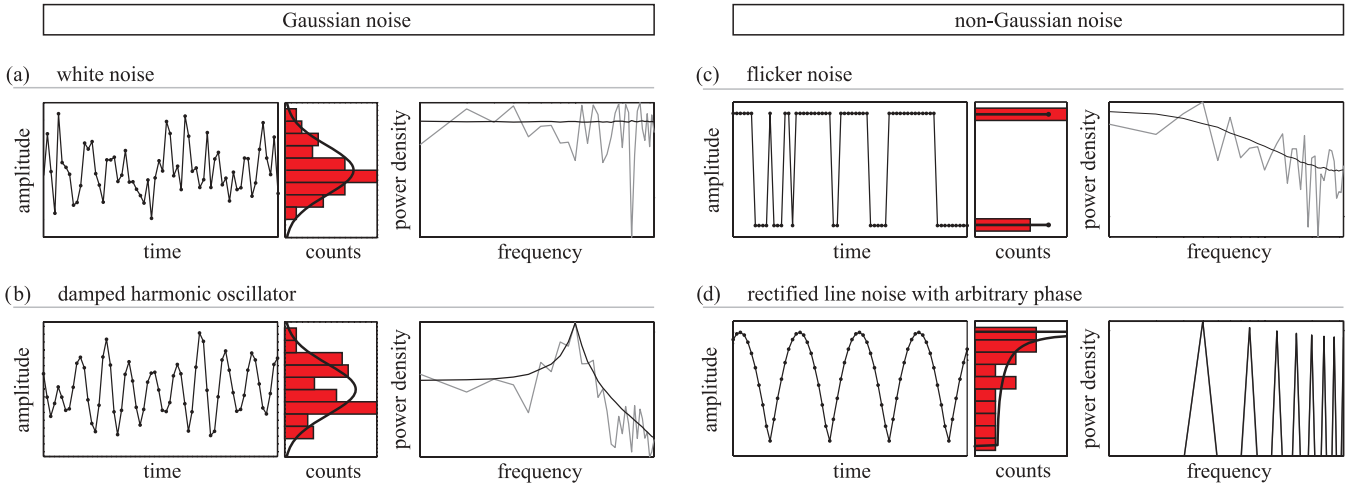

FIG. 6. (Color online) Four different stationary noise sources are analyzed. A single outcome is drawn, and histogrammed. Noise can have a Gaussian (a),(b) or a non-Gaussian (c),(d) probability density function, as overlaid on the histogram. Although flicker noise is non-Gaussian, it is stationary and therefore has a $\langle V_f \rangle$ which is asymptotically diagonalizable and a well-defined power spectral density. The rectified line noise is non-Gaussian and is only stochastic because it has a random phase. All harmonics are locked in phase, but since the noise is stationary, the analysis of variance using the power spectral density still holds.

Figures 6(a) and 6(b) shows examples of two stationary Gaussian noises: white noise and a thermally driven damped harmonic oscillator. The Gaussian property is illustrated by the histograms, which show convergence to a normal distribution in the time domain. Even though the harmonic oscillator has highly correlated data in the time domain, both spectral covariance matrices are diagonal, and therefore the power spectral densities describe these Gaussian noise sources completely.

Figure 6(c) depicts flicker noise [40], defined here as a two-state process with a switching time that is exponentially distributed. Although purely stochastic and stationary, this bimodal process is far from Gaussian, yet the spectral covariance matrix is diagonal because the process is stationary, and therefore an accurate analysis of variance can be performed in the frequency domain.

Figure 6(d) represents rectified line noise, modeled as the absolute value of a sine wave. Although stationary, its only stochastic component is the random phase. Such a noise source is very spectrally pure, and defined only by the fundamental frequency and harmonics. Despite the fact that the PSD has only deterministic harmonic powers, the spectral covariance matrix is still diagonal and the analysis of variance presented in the following section holds. Note that spectral leakage has been disregarded in the discussion so far because it can be made arbitrarily small by appropriate averaging methods (see Appendix).

C. Probability distribution of the power spectral density

This description begins with the special case of stationary Gaussian noise, and is later extended to stationary non-Gaussian noise. The limitations in statistically characterizing $1/f$ noise, which is nonstationary, are also discussed. Finally, the behavior of deterministic noise under stationary noise analysis is investigated.

1. Stationary Gaussian noise

Each random variable composing the frequency-domain noise vector $\boldsymbol{\epsilon}_f$ is a linear combination of all the random variables in the time-domain noise vector $\boldsymbol{\epsilon}_t$, as weighted by the complex-valued Fourier matrix $\mathcal{F}$. Since $\boldsymbol{\epsilon}_t$ is normally distributed for Gaussian noise, the real and imaginary random variables of $\boldsymbol{\epsilon}_f$ are also normally distributed, and any two frequencies are asymptotically independent for $m \to \infty$.

The power spectral density is the squared magnitude of $\boldsymbol{\epsilon}_f$,

$$\boldsymbol{n}^2 = \mathrm{diag}(\boldsymbol{\epsilon}_f \cdot \boldsymbol{\epsilon}_f^\dagger) \tag{10}$$

(for notational simplicity, the squaring in $\boldsymbol{n}^2$ implies squaring the magnitude $|\boldsymbol{n}|^2$). Note that $\boldsymbol{n}^2$ is a $m \times 1$ vector of random variables with an associated probability distribution. According to Eq. (10), every component of $\boldsymbol{n}^2$ is the sum of squares of the identically distributed real and imaginary parts of the corresponding component of $\boldsymbol{\epsilon}_f$. From basic probability theory, it is known that the sum of squares of two independent and identically distributed normal random variables is exponentially distributed. In other words, the power spectral density $\boldsymbol{n}^2$ of stationary noise is composed of independent exponentially distributed random variables.

Experimentally, the expectation value $E(\boldsymbol{n}^2)$ is estimated by the average power spectral density $\langle n^2 \rangle$, taken over many single observations n^2. Because of the large variance of the exponential distribution, it is important to perform many averages when measuring power spectral densities [41]. Note that the averaging should be performed in power $\langle |\boldsymbol{n}|^2 \rangle$, rather than in magnitude $\langle |\boldsymbol{n}| \rangle^2$, otherwise the PSD is underestimated by a factor of $4/\pi$ (see Appendix for details).

2. Stationary non-Gaussian noise

Stationary non-Gaussian noise sources may also comply with the statistical properties described so far—under certain conditions. Take flicker noise, as illustrated in Fig. 2(c); even though the time-domain probability density function is far

from Gaussian, the frequency-domain random variables are asymptotically normally distributed by virtue of the central limit theorem, and thus $\boldsymbol{n}^2$ is composed of exponentially distributed random variables. On the other hand, the rectified sine wave in Fig. 6(d) has only one random component—its phase—and therefore does not obey the criterion for the central limit theorem to take effect. In fact, any two $\boldsymbol{n}^2$ for this noise are identical; in other words, $\boldsymbol{n}^2$ is purely deterministic with no associated probability distribution (disregarding spectral leakage).

Although the power spectral density does not fully describe the statistical properties of stationary non-Gaussian noise, it is sufficient for the noise analysis of linear measurements presented in the next section.

3. 1/f noise

Measuring $1/f$ noise is a nontrivial task. For exponents $0 < c < 1$, noise with power spectra $\sim 1/f^c$ are stationary but with long-range memory. For $c \geqslant 1$, the noise is nonstationary, having variance which changes with time [42]. In certain cases, the $1/f$ power spectrum of the noise can be approximated on a limited bandwidth, as long as its PSD is measured across a much wider bandwidth. For nonstationary noise, however, the windowing function must be chosen according to c in order to obtain consistent estimates.

Meanwhile, any distributional assumption such as Gaussianity of $1/f$ noise should always be verified. A common deterministic signal leading to a $1/f$ noise spectrum is linear drift. Stochastic $1/f$ noise and deterministic linear drift are fundamentally different sources of noise. Using the PSD to represent linear drift is inconvenient. Estimation details in the presence of both stochastic $1/f$ noise and deterministic drift are discussed in the Appendix, where it is concluded that it is best to subtract linear drift from any signal before calculating its PSD to more accurately estimate the $1/f$ noise.

4. Deterministic periodic noise

In typical AFM experiments, the relative phase of the deterministic periodic noise is random, such that its effect on a measurement is not purely deterministic.

Take a square wave noise with constant amplitude and a random phase; the power spectral density $\boldsymbol{n}^2$ of this noise source has fixed harmonics with predetermined magnitudes (disregarding spectral leakage). Therefore, $\boldsymbol{n}^2$ is a vector of deterministic numbers, rather than random variables. The following simulation describes the effect of such deterministic noise on the probability distribution of linear measurements. For illustrative purposes, a friction experiment was simulated with three noise sources:

(1) Deterministic noise: a deterministic square wave with a period equal to the scan rate. The only random component is the relative phase of the square wave.

(2) Dephased noise: the noise from (1) was recycled but each frequency component was given a random phase (drawn from a uniform distribution).

(3) Stochastic Gaussian noise: noise was generated from the PSD from (1), whereby each complex-valued frequency amplitude was drawn from a complex normal distribution. This method is described in detail in Sec. VII.

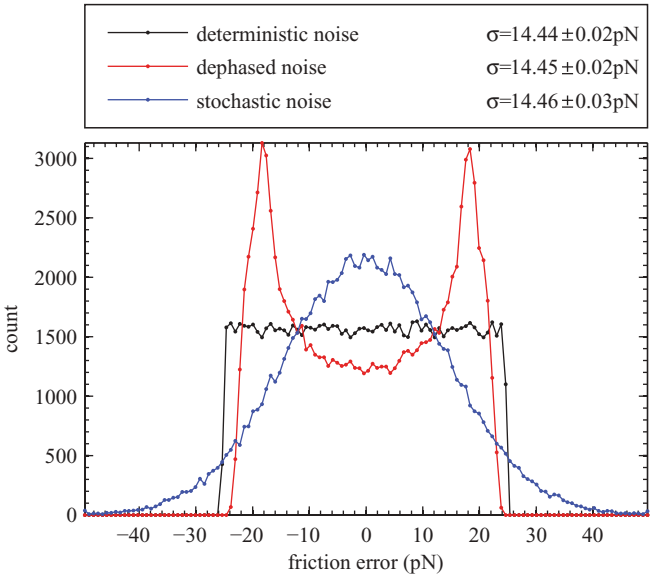

FIG. 7. (Color online) A friction experiment was simulated for a square wave noise, assuming (1) a single random phase (deterministic noise), (2) dephasing all the frequency components but keeping fixed magnitudes (dephased noise), and (3) drawing the complex-valued amplitudes from a normal distribution (stochastic noise). Note that all noise sources lead to an identical standard deviation in friction, even though their probability distributions are completely different. The square wave was given an amplitude of 50 pN, and its period was equal to the scan rate. Note that such a source of noise would be imperceptible in the lateral force channel in Fig. 1(a) due to its small amplitude, but could cause significant error in estimating the friction computed in Fig. 1(b).

The result is presented in Fig. 7. Because the simulation was performed with zero friction, any observed friction is an error caused by noise. Although each noise source results in completely different distributions in the observed friction error; their standard deviations are identical. This demonstrates that the analysis of variance using PSDs is often accurate for any type of stationary noise source (deterministic or stochastic); however, it cannot predict the probability distribution of a linear measurement unless the noise is purely stochastic and stationary, which are the assumptions required for the central limit theorem in Sec. V C 2 to hold.

D. Discussion

The frequency domain is best suited for noise analysis because its covariance matrix is diagonal. In other words, the frequency-domain data points are uncorrelated, while time-domain data points are correlated if the noise is not white. This significantly simplifies the analysis of variance presented in the next section, where variances can be added without consideration of covariances.

The diagonal of the covariance matrix, defined as the power spectral density (PSD), contains all the statistical information of stationary Gaussian noise, which describes fundamental noise sources in AFM such as optical shot noise and thermal noise. For stationary non-Gaussian noise, the PSD still provides an accurate measure of variance, but lacks information regarding the specific probability distribution of

a measurement performed in the presence of such noise. This statement applies to both stochastic noise, such as flicker noise, and deterministic noise, such as line noise.

The assumptions in describing thermal noise and optical shot noise as stationary are that the temperature is fixed, and that the average optical power is fixed, respectively. It is well known that there are fluctuations in both quantities, which set a lower limit to accuracy in describing these noise sources by their power spectral density. For example, $1/f$ noise in optical power causes the shot noise level to fluctuate during an experiment. More importantly, we are aware that the PSD of detection noise may vary during an experiment if the cantilever deflection changes (see Fig. 4). This can also violate the assumption of stationarity.

For stationary stochastic noise, the data points of a PSD are independent and exponentially distributed. The exponential distribution of PSD data points follows from the fact that the real and imaginary amplitude components at a given frequency are (asymptotically) independent and identically distributed Gaussian random variables.

VI. STANDARD DEVIATION OF LINEAR MEASUREMENTS

This section outlines a method for integrating the power spectral density that provides a quantitatively accurate standard deviation of any arbitrary linear measurement.

For an accurate noise analysis, it is best to characterize the noise of the system across a bandwidth much wider than the bandwidth used during the experiment. Ideally, the noise is measured from a frequency well below the scan rate (by a few orders of magnitude) and up to a frequency that exceeds the roll-off frequency of the detection system. Accurate PSD estimation is covered in the Appendix.

A. Linear measurements in the time domain

As described in the previous section, experimental data are typically acquired in the time domain, and recorded as a $m \times 1$ vector composed of a deterministic component (d_t) and stochastic component ($\boldsymbol{\epsilon_t}$), as in

$$\boldsymbol{x_t} = d_t + \boldsymbol{\epsilon_t}. \tag{11}$$

The deterministic component contains the information of interest to the experimenter. A single-valued linear measurement $\boldsymbol{\varphi}$ is defined as any linear combination of the m values of $\boldsymbol{x_t}$, summarized by the dot product

$$\boldsymbol{\varphi} = w_\varphi \cdot \boldsymbol{x_t}, \tag{12}$$

where the $1 \times m$ vector w_φ is the *measurement sampling function* which weighs the linear combination of $\boldsymbol{x_t}$ elements. Note that $\boldsymbol{\varphi}$ is a random variable as it is the weighted sum of many random variables.

For example, if $\boldsymbol{x_t}$ is a time series of the lateral force for a single scan loop during a friction experiment (see Fig. 1), the following measurement sampling function results in the average friction force:

$$w_\varphi = \frac{1}{m}[1, 1, \ldots -1, -1], \tag{13}$$

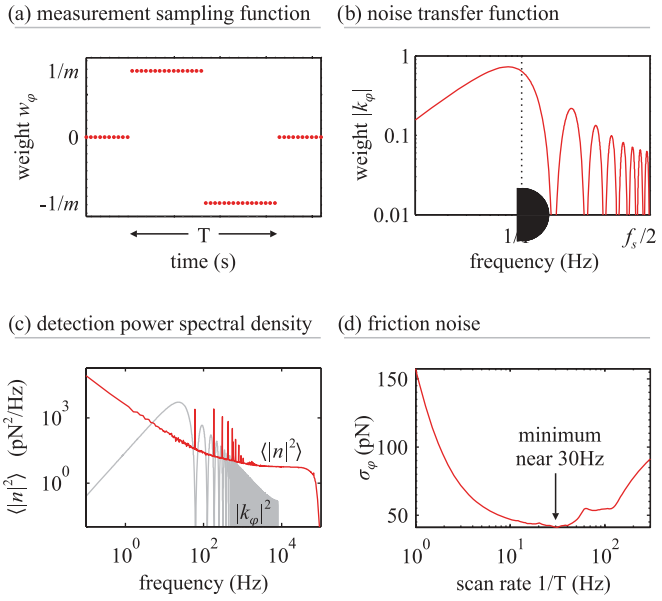

FIG. 8. (Color online) (a) The measurement sampling function for a friction measurement w_φ, and (b) the magnitude of its Fourier transform $|k_\varphi|$. Zero-padding of w_φ was used to increase the frequency resolution of $|k_\varphi|$, as explained in the Appendix. (c) The power spectral density of detection noise $\langle n^2 \rangle$, and the noise transfer function $|k_\varphi|^2$ (overlaid in light gray for a scan rate of 10 Hz), which are multiplied and integrated to obtain the standard deviation of a friction measurement σ_φ, at various scan rates as shown in (d). Detection noise in this friction experiment is minimized at a scan rate of ~30 Hz, where 60 Hz and harmonics are filtered out by the zeros of $|k_\varphi|$, as seen in (c). An ideal Nyquist filter at $f_s/2$ was assumed in this computation. In (a) and (b), only 32 data points were used to define w_φ (prior to zero-padding) for visual clarity. In (c) and (d), a 512-data-point w_φ was used to match our experiment.

with an equal number of positive and negative values as shown in Fig. 8(a). In fact, this is the measurement sampling function for estimating the hysteresis in a variety of AFM experiments, including force-distance spectroscopy, for example.

It is inevitable that this measurement has some variability caused by true physical variability between repeated measurements d_t and/or due to noise $\boldsymbol{\epsilon_t}$. The topic of this section is to separate these two sources of variability by accurately quantifying the variability caused by $\boldsymbol{\epsilon_t}$.

For linear measurements, the deterministic and stochastic components are separable, which follows by combining Eqs. (11) and (12) into

$$\boldsymbol{\varphi} = w_\varphi \cdot d_t + w_\varphi \cdot \boldsymbol{\epsilon_t}. \tag{14}$$

The expectation value of the measurement is $\mathrm{E}(\boldsymbol{\varphi}) = w_\varphi \cdot d_t$. The variance of the measurement σ_φ^2 caused by stochastic noise $\boldsymbol{\epsilon_t}$ is a linear combination of the elements of the $m \times m$ temporal covariance matrix V_t.

$$\sigma_\varphi^2 = \mathrm{Var}(w_\varphi \boldsymbol{\epsilon_t}) = w_\varphi V_t w_\varphi'. \tag{15}$$

The following section uses the properties of stationary noise to simplify this matrix calculation by transferring this noise analysis into the frequency domain.

B. Linear measurements in the frequency domain

The variance analysis of the measurement $\boldsymbol{\varphi}$ can be transferred into the frequency domain by simply adding the identity $\mathcal{F}^{-1}\mathcal{F} = I$ to the previous equation:

$$\sigma_\varphi^2 = \mathrm{Var}(w_\varphi \mathcal{F}^{-1}\mathcal{F}\boldsymbol{\epsilon_t}) = k_\varphi V_f k_\varphi^\dagger, \tag{16}$$

where the *noise transfer function* $k_\varphi = w_\varphi \mathcal{F}^{-1}$ and the spectral covariance matrix $V_f = \mathrm{Var}(\mathcal{F}\boldsymbol{\epsilon_t})$ as defined in the previous section. For stationary noise, V_f is diagonal; therefore the variance of the measurement $\boldsymbol{\varphi}$ can be computed using vectors rather than matrices, and (16) simplifies to the dot product

$$\sigma_\varphi^2 = |k_\varphi|^2 \cdot \mathrm{diag}(V_f). \tag{17}$$

In other words, the variances in the frequency domain can be added without consideration of covariances because the random variables at all frequencies are independent (no off-diagonal terms). In any experiment setting, $\mathrm{diag}(V_f)$ is estimated by the time-averaged power spectral density $\langle n^2 \rangle$, such that

$$\sigma_\varphi^2 \approx |k_\varphi|^2 \cdot \langle n^2 \rangle. \tag{18}$$

This equation summarizes a simple method for numerically integrating a power spectral density to obtain the standard deviation of any linear measurement. In summary, the time-domain measurement sampling function w_φ is known to any experimenter, and its squared Fourier transform $|k_\varphi|^2$ provides the weighting factors for integrating the power spectral density $\langle n^2 \rangle$.

As an example, the noise transfer function $|k_\varphi|$ for a friction measurement is shown in Fig. 8(b). The detection noise of our AFM, in Fig. 8(c), was integrated with this noise transfer function at various scan rates. Some of the technical details relating to spectral leakage and aliasing when calculating the standard deviations in Fig. 8(d) are discussed in the Appendix.

Figure 8(d) demonstrates that the standard deviation of the friction measurement due to the detection noise is minimized at a scan rate of 30 Hz. This result might be surprising, as it is typically assumed that scanning at a slower rate results in lower detection noise. However, the opposite can be true in the dominance of $1/f$ noise.

With an accurately calculated value of the friction variability due to stochastic noise, any additional variability can be attributed to true variation in friction and interpreted accordingly.

C. Various linear measurements

To demonstrate the general utility of this approach, the measurement sampling functions w_φ of four examples of linear measurements along with the noise transfer functions $|k_\varphi|$ are shown in Fig. 9(a). The measurements are integrated in Fig. 9(b), where the detection noise was modeled as $1/f^2$ and white noise with a corner frequency near 100 Hz. Depending on the type of measurement, the noise may be minimized anywhere between 0 Hz and 1 kHz.

Most measurements have a local minimum in detection noise where the contribution of $1/f^2$ noise and white noise are equal; in the dominance of $1/f^2$ noise faster scanning is desirable, whereas slower scanning reduces the impact of white noise.

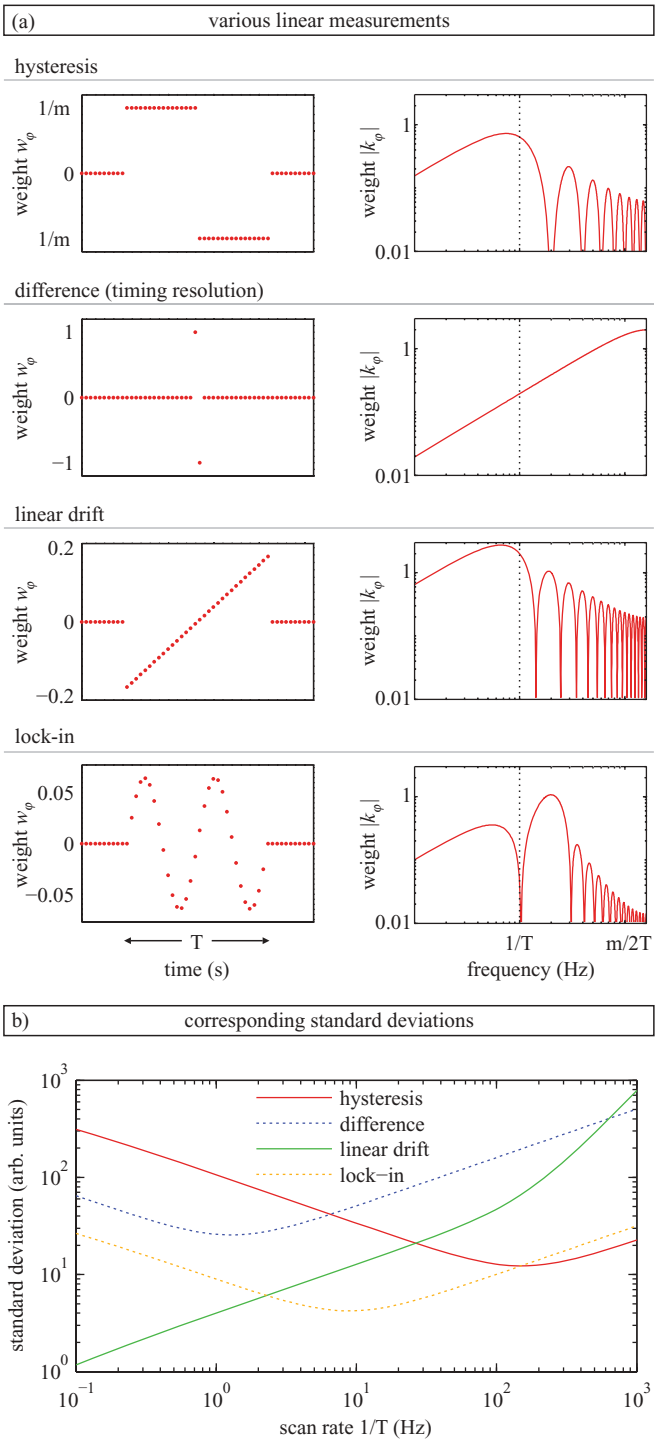

FIG. 9. (Color online) Four linear measurements are shown in the time and frequency domains. The number of sample $m = 32$ was set low for illustrative purposes only. (b) The standard deviation of detection noise was computed for $m = 256$, which is a typical number of data points per scan line in AFM, at various scan rates $1/T$, where T is the duration of a single scan line. The noise was modeled as $1/f^2$ and white noise with a corner frequency near 100 Hz. Depending on the measurement, the noise minimum can be anywhere between 0 Hz and 1 kHz.

In Fig. 9, linear drift is the only linear measurement with a standard deviation that decreases monotonically as the

scan rate decreases. This decrease occurs because drift is an intrinsically time-dependent quantity: the longer the acquisition, the larger the drift signal becomes. This effect outweighs $1/f^2$ noise in the case of such time-dependent measurements.

D. Discussion

The data in Fig. 1 were acquired with a difference signal $|\Delta_0| = 0.007$, which allowed us to reconstruct the PSD of detection noise by the weighted sum of common-mode and differential-mode noise of our SLD, using Eq. (4). The resulting PSD was integrated using Eq. (18) to provide an accurate measure of the standard deviation expected for detection noise in our friction experiment. The calculated detection noise of 15 pN is just below the measured variability of 19 pN in Fig. 1. This implies that the true friction variability on gold is barely resolvable above detection noise. On the other hand, 28 pN of the observed 32 pN of variability on copper is attributed to true changes in friction.

The true variability in friction on copper chloride is distinguishable from detection noise as it is not white. In fact, it is closer to brown noise, suggesting that there are slow processes that cause most of the friction variability. Experiments have shown $1/f$ variations in friction due to wear debris [43]. Although our experiment is well below the onset of wear, studying the differences in statistical distributions of friction on a copper chloride monolayer vs gold can help understand the effects of adsorbed monolayers on nanoscale friction [44]. This would provide additional information regarding changes of the tip-sample junction composition, such as rearrangement of the tip atoms or sporadic changes in the ionic composition near the surface that may affect the coadsorption of chloride with the copper monolayer, all of which can contribute to changes in friction. Furthermore, the distribution of friction can be related to the statistics of the stick-slip jump events to understand the effects of thermal fluctuations as well as the atomic structure and orientation of the surface on the outcome of a nanoscale friction experiment [45]. Alas, a reduction in detection noise is required to study the true fluctuations in friction on copper chloride vs gold and their statistical properties in our experiment.

For the data in Fig. 1, our scan rate was set to 25 Hz, which falls within the minimum region around 30 Hz as seen in Fig. 8(d). This region occurs because of the joint minimization of $1/f$ and white noise, as well as the rejection of 60 Hz and harmonics for a scan rate of exactly 30 Hz. Notice that the noise transfer function $|k_\varphi|$ of a friction experiment has zeros at even harmonics of the scan rate. At these frequencies, an integer number of periods fit into the friction measurement sampling function w_φ, and therefore cause zero error in friction regardless of their amplitude.

It would have been worthwhile to properly center the detection light beam ($|\Delta_0| = 0$) at the beginning of the experiment to remove common-mode noise which contributed $\sim$20% of additional detection noise. Any further reduction in detection noise would require design changes to our AFM. The use of softer cantilevers [33] would reduce the detection noise of lateral force, which to first order scales linearly with torsional stiffness. Secondly, the reduction of the light beam divergence would also lead to lower detection noise. The light beam divergence can be reduced by using a smaller collimated light beam diameter, a longer effective focal length for focusing the light beam, exploiting the stress-induced cantilever radius of curvature [46], or patterning the cantilever with a diffraction grating [14].

It is instructive to note that boosting the light power, by coating the cantilever or increasing the SLD drive current, would *not* reduce the detection noise because the friction detection noise is dominated by $1/f$ noise for scan rates below 100 Hz on our system. This occurs because the hysteresis noise transfer function, seen in Fig. 9(a), has a peak at low frequencies around the scan rate, whereas higher-frequency shot noise is mostly averaged out by the measurement sampling function. Increasing optical power only reduces detection shot noise. Actually, coating the cantilever would only *increase* the normal force noise at low frequencies due to coupling between optical power fluctuations and cantilever thermally induced bending [14].

VII. STOCHASTIC SIMULATION

Certain experiments require simulation, rather than calculation such as performed in the previous section, to quantify the effects of noise on the outcome of the experiment. Specifically, this is necessary for nonlinear situations, for which a linear measurement sampling function does not exist and a PSD cannot be simply integrated. In this section, a method for simulating stochastic Gaussian noise from any arbitrary spectral distribution is presented, and is used in a friction experiment simulation to quantify the effects of mechanical vibrations of the AFM on atomic-scale friction.

A. The inverse fast Fourier transform method for simulating Gaussian noise

The detailed description of the probability distribution of power spectral densities in Sec. V contains the ingredients for simulating stochastic Gaussian noise. The true power spectral density $\boldsymbol{n}^2$ may be approximated by a measured $\langle n^2 \rangle$ or calculated from a theoretical model. Given an equally spaced $m \times 1$ vector $\langle n^2 \rangle$, a frequency-domain noise vector ϵ_f can be simulated as follows.

The first half of ϵ_f must be drawn from a complex normal distribution; the real and imaginary elements of $\epsilon_{f_{1/2}}$ for each frequency bin are drawn from independent normal distributions with mean 0 and variance $\langle n^2 \rangle/2$, as in

$$\epsilon_{f_{1/2}} = \mathcal{N}\left(0, \frac{\langle n^2 \rangle}{2}\right) + i\mathcal{N}\left(0, \frac{\langle n^2 \rangle}{2}\right). \tag{19}$$

Then, the ϵ_f must be arranged to comply with the properties of a real signal in the time domain [39]: the negative frequencies must be complex conjugates of the positive frequencies, such that $\epsilon_f = [\epsilon_{f_{1/2}} \epsilon^*_{f_{1/2}}]$, and $\epsilon^*_{f_{1/2}}$ must be flipped. Finally, the resulting $2m \times 1$ vector ϵ_f can be converted to a time-domain noise vector by the inverse Fourier transform matrix:

$$\epsilon_t = \mathcal{F}^{-1} \epsilon_f, \tag{20}$$

which can be efficiently computed using the fast Fourier transform (FFT).

The caveat in using this method is that the $\mathcal{F}^{-1}$ is a circulant matrix, making the first and last data points of ϵ_t contiguous, and therefore strongly correlated. This assumption has no

consequence for generating noise with no periodicity or long-range correlation, such as white noise, for example. The circulant property is also relatively inconsequential for an oscillatory signal which decorrelates within the time window of the generated data set. In any case, the safest approach is to generate $2\times$ more data than requested and to discard half of the data. This results in a $m \times 1$ noise vector ϵ_t. For $1/f$ noise, which is divergent, it may be necessary to discard more than half the data to achieve a desired level of statistical accuracy, depending on the exponent of the $1/f$ noise. This method was used to generate the Gaussian noises in Figs. 5 and 6 from their PSDs, for example.

B. Simulating displacement noise

The lateral mechanical vibrations of our AFM were measured in air by bringing a cantilever (PPP-NCHAuD, Nanosensors) into contact with a flat sapphire sample. A stiff lateral contact was ensured by applying a large normal load of several hundred micronewtons [47]. As long as the lateral contact stiffness exceeds the lateral cantilever stiffness, the power spectral density of the lateral signal is an accurate measure of the lateral displacement noise. Because our light beam is circularly symmetric, the lateral and normal sensitivity (in nm/V) is related by a simple geometric factor [33] (calibrated before making contact). Lastly, detection noise was subtracted from the measurement to isolate the lateral displacement noise, shown in Fig. 10(a). The lateral displacement noise integrates to 51 pm. Although this approach is highly prone to calibration errors (up to a factor of 2), this does not qualitatively affect the conclusions drawn in the Discussion subsection.

In an ideal friction experiment, the cantilever base is scanned back and forth smoothly across the sample with some velocity v. In reality, displacement noise causes the scanning to be erratic, as shown in Fig. 10(b). In this case, displacement noise ϵ_{disp} in the time domain was simulated using the inverse fast Fourier transform (iFFT) method from the spectral density defined in Fig. 10(a), and then added to the smooth triangular wave.

Although it is strictly incorrect to simulate 180 Hz line noise using the iFFT method (because it is a spectrally pure noise source with fixed amplitude and random phase), this noise source carries a negligible amount of power in the measured PSD. Note that the noise peak near 540 Hz is not spectrally pure. Using a "direct" method of averaging which suppresses spectral leakage (see Appendix), the noise peak near 540 Hz is unequivocally shown to span a wide bandwidth: 534–542 Hz. The breadth of this peak reflects the fact that this mechanical vibration of the sample is a stochastic process with a short coherence time relative to the time scale of the experiment. This justifies the iFFT method's assumption of uncorrelated frequencies for simulating this noise source.

Both the ideal and erratic scanning profiles in Fig. 10(b) were used as inputs to the following atomic stick-slip simulation.

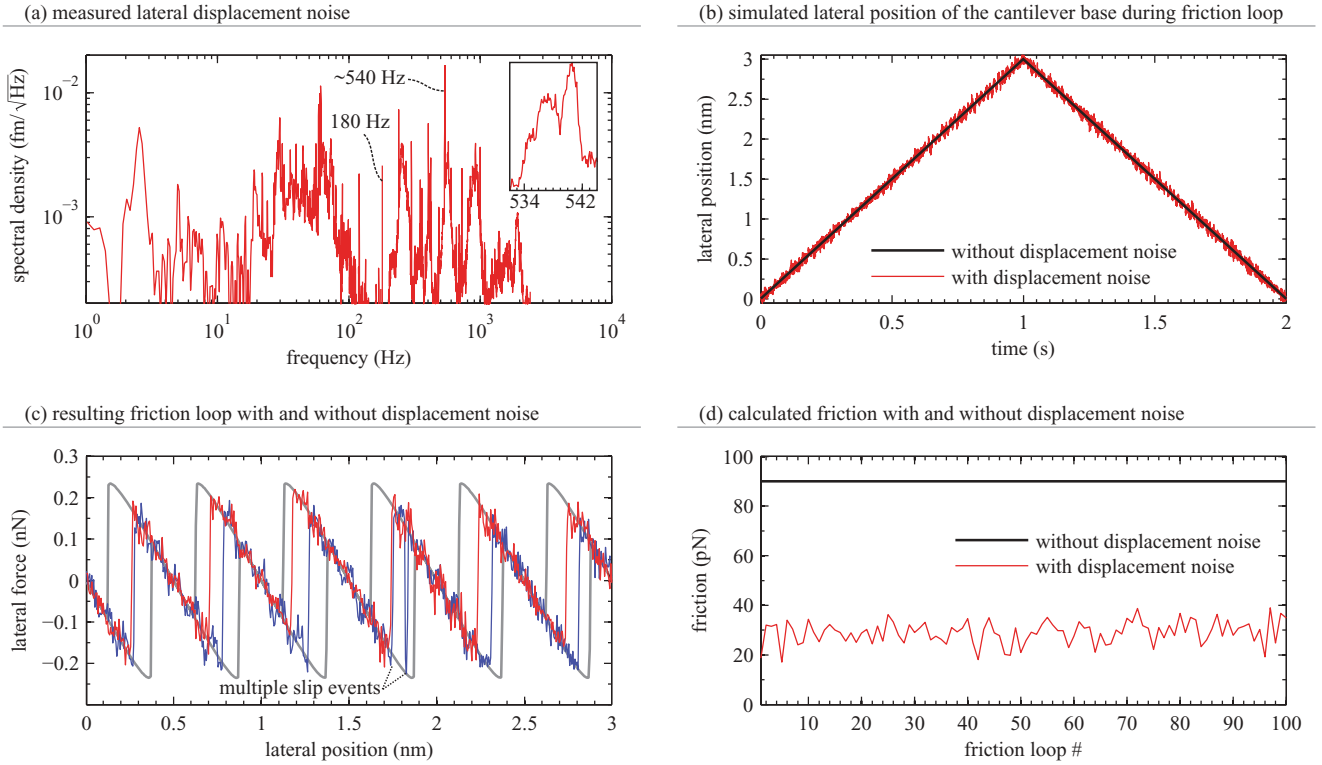

FIG. 10. (Color online) (a) The lateral displacement noise was measured on our system for 400 s. The direct method of averaging was used to prevent spectral leakage; a 51-point average led to a frequency resolution of 0.1275 Hz. The detection noise was subtracted. The inset shows the prominent source of noise near 540 Hz, which is stochastic in nature. (b) The iFFT method was used to simulate displacement noise in the time domain, which affects the lateral scanning of the tip in a friction experiment. (c) A friction loop was simulated at zero temperature without and with displacement noise; the results are shown as thick and thin lines, respectively. (d) The displacement noise lowers the friction by $\sim 3\times$ and causes signal variability.

C. Simulation of atomic stick-slip

This simulation is based on the one-spring Tomlinson model [48], which assumes that the potential $U_a(x)$ in the lateral direction x is sinusoidal with a period defined by the atomic lattice of the substrate. A time-varying harmonic potential of the contact is defined by the effective lateral contact stiffness k, as in

$$U_c(x,t) = \tfrac{1}{2}k(x - vt + \epsilon_{\text{disp}})^2, \tag{21}$$

which includes the scanning velocity v and displacement noise ϵ_{disp} discussed earlier. At every time step in the simulation, the new tip position is reassigned as the new local minimum in the combined potential $U = U_a + U_c$. This implicitly assumes that the system temperature is zero. It also conveniently makes the following simulation purely deterministic, such that the effects of stochastic displacement noise, simulated in the previous section, can be isolated unambiguously. To this end, detection noise was also omitted from the simulation.

The parameters used for the simulation were taken from a previous simulation of atomic stick-slip performed by Socoliuc *et al.* [49]. A tip was scanned at $v = 3$ nm/s across NaCl with a lattice constant $a = 0.5$ nm, and an effective lateral contact stiffness $k = 1$ N/m. The corrugation energy E_0 of the sinusoidal potential was set to 0.237 eV, resulting in $\eta = 3$, where $\eta = 2\pi^2 E_0/ka^2$. The η parameter defines the stick-slip ($\eta > 1$) and the continuous sliding ($\eta < 1$) regimes of friction.

By assuming no displacement noise, the simulation results in a predictable stick-slip pattern represented by thick gray lines in Fig. 10(c). Including displacement noise in the simulation strongly affects the outcome of the experiment, as shown by the overlaid thin lines in Fig. 10(c). The timing of the slip events has significant variability. Furthermore, the measured friction force is much lower relative to the ideal case, as summarized by the 100 simulations in Fig. 10(d). The friction loops of all 100 simulations are available elsewhere (see Supplemental Material [50]).

D. Discussion

The peak lateral force required to initiate a slip event seems to have dropped significantly due to displacement noise, as seen by comparing both friction loops in Fig. 10(c). However, this is simply an illusion caused by averaging the data to an effective 512-Hz sampling rate. On the time scale of each simulation step (5 kHz), the lateral force that leads to a slip event is exactly 0.239 nN by virtue of the assumed Tomlinson model. The displacement noise at high frequencies causes this lateral force threshold to be reached prematurely, resulting in early slip events and a threefold reduction in friction. This illustrates the nonlinear nature of this simulation: displacement noise can cause an early slip event, which is usually irreversible because the cantilever subsequently remains in its new local minimum of the combined potential U. Displacement noise can also induce back-and-forth slip events, as pointed out in Fig. 10(c), where a large stochastic mechanical disturbance causes the tip to jump back and forth between two lattice positions. Similar events have been observed experimentally [51] and postulated theoretically [52,53]. Lastly, the stochastic nature of the displacement noise also causes significant variability in friction, as seen in Fig. 10(d). Such variability would not be observed if dynamic superlubricity were induced by a deterministic sine wave [31], for example.

The mechanism that causes the suppression of friction described so far is akin to a thermal activation process described by Arrhenius transition state theory. Thermal activation of slip events has been theoretically proposed by Gnecco *et al.* [48] and Sang *et al.* [54], and is still under study today [55]. Such models predict a decrease in friction with increasing temperature as well as a logarithmic dependence of friction with scanning velocity.

Similarly, stochastic displacement noise can assume a logarithmic velocity dependence because the mechanical activation process reported here is analogous to thermal activation. For Gaussian displacement noise, the amplitude of the tip-sample vibrations follows a normal distribution, such that the instantaneous potential energy (amplitude squared) of the system follows an exponential distribution. Secondly, the attempt frequency (the rate parameter in the Arrhenius equation) relates directly to the coherence time of the mechanical noise.

The logarithmic velocity dependence of displacement noise is shown in Fig. 11. The slight deviations from ideal logarithmic behavior are due to the fact that the attempt frequency is not single-valued in this model; rather, it is defined by the displacement noise spectrum in Fig. 10(a). However, most of the power of the displacement noise falls in the 534–542 Hz frequency range, described earlier, which dominates the logarithmic dependence observed in Fig. 11 for velocities up to 400 nm/s. At that velocity, the stick-slip frequency (800 Hz) exceeds the mechanical vibration frequency (~540 Hz), and the idea of an attempt frequency breaks down. Although the displacement noise still causes variability in the friction data at high sliding velocities, the mechanical vibrations are just as likely to increase as to decrease the friction, such that the average friction with and without displacement noise both coincide.

Recently, the effects of athermal instrumental noise on friction have been modeled parametrically by Dong *et al.* [56] as a white force noise amplified by the cantilever lateral resonance. This instrumental *force* noise was assumed temperature independent and predicts a friction plateau at cryogenic temperatures where thermal activation becomes

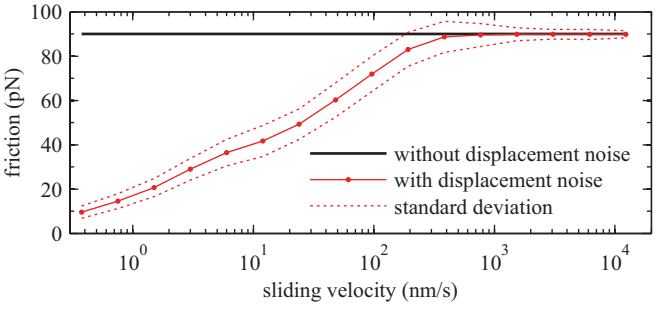

FIG. 11. (Color online) The friction simulation was performed at varying velocities to demonstrate the logarithmic velocity dependence of friction in the presence of displacement noise. A plateau occurs once the stick-slip frequency exceeds the displacement noise frequency (dominated by noise near 540 Hz). The signal variability decreases once this plateau is reached. The error on the mean roughly corresponds to the diameter of the circles.

negligible. In other words, the instrument was assumed to have a finite effective temperature caused by instrumental noise.

In contrast, our simulation uses an empirical measurement of the *displacement* noise of our instrument as input. The dynamics of the cantilever play no role in this simulated experiment because the lateral resonance frequency of the cantilever, say 150 kHz, greatly exceeds the mechanical vibrations of the AFM (<3 kHz). In that respect, the cantilever can be accurately described as an ideal force sensor that does not affect the outcome of the experiment. Nevertheless, the displacement noise can be attributed a cantilever-equivalent effective temperature. By the equipartition theorem, the cantilever in this simulation is expected to have 12 pm of vibrational noise at room temperature ($\sqrt{k_B T/k_c}$), while there is 51 pm of integrated displacement noise (across the 1 Hz to 10 kHz bandwidth). In other words, the cantilever temperature would have to be 5500 K for the integrated cantilever noise to exceed the integrated displacement noise.

We now note the problematic fact that the displacement noise of an AFM is expected to be temperature dependent. The mechanical response of an AFM has been shown to change drastically at cryogenic temperatures [57]. Due to the lowered damping of hardware components in the AFM, the external noise coupling drives the mechanical components with larger amplitudes and can promote the suppression of friction at cryogenic temperatures. This may result in a nonmonotonic temperature dependence of friction that would be characterized by a drop in friction at low temperatures.

We conclude that the reduction of friction by mechanical dithering of the sample, known as dynamic superlubricity [31], inadvertently occurs due to inevitable mechanical vibrations of an AFM. The degree of this friction suppression depends on a multitude of experimental parameters and the specific displacement noise of the AFM in question. Therefore, a general quantitative assessment of the problem cannot be made. However, the method proposed here can be used to estimate the expected impact of displacement noise on any particular system and experiment. Measuring the temperature dependence of the displacement noise is also recommended before performing an experiment for studying the temperature dependence of friction.

VIII. SUMMARY AND CONCLUSIONS

The power spectral density (PSD) of detection noise may change during an experiment if the light beam does not remain perfectly centered on the photodetector. This PSD is measurable and predictable after characterizing the common-mode and differential-mode noise of the light beam used to measure the cantilever deflections.

An intuitive method for integrating a PSD to provide accurate estimates of any linear measurement was presented, allowing predictive power over experiments, and informed analysis of measured data. We have shown that the variability in measured friction may be dominated by detection noise even in situations with clearly resolved stick-slip. Interestingly, increasing the light power (by a reflective coating, for example) does not reduce this friction variability, which is limited by $1/f$ angular noise of the light beam. In the dominance of $1/f$ noise, faster scanning reduces the impact of noise on the measurement. From a noise analysis perspective, the optimal scan rate for friction experiments is 30 Hz, on our particular AFM.

The power spectral density is a powerful tool for noise analysis of stationary noise, either stochastic or deterministic. Understanding the statistical properties of measured PSDs for different noise sources provides the necessary foundation for the simulation of AFM experiments affected by stochastic noise. Such simulations allow the assessment of the relative contribution of detection, force, and displacement noise sources in any particular AFM experiment.

In this work, stochastic simulation of measured displacement noise was used to quantify the effects of mechanical vibrations on the outcome of a friction experiment. Assuming a one-spring Tomlinson model with typical scanning conditions and zero temperature, our simulation predicts that the displacement noise can cause a logarithmic velocity dependence of friction, induce multiple back-and-forth slip events in the atomic stick-slip process, and cause significant variability in measured friction.

An attractive feature of the iFFT method of stochastic simulation is that it is not restricted to parametric models: any arbitrary numerically defined PSD can be used as input. Mechanical vibrations between the tip and the sample, for example, can only be empirically measured, and may be difficult to accurately or efficiently describe parametrically. We have attributed a stochastic sample vibration near 540 Hz as the largest contributor to athermal friction suppression and increased friction variability on our system.

Although we have demonstrated noise analysis and stochastic simulations in the context of friction experiments, these methods can be used for a wide range of AFM experiments, and extend to dynamic AFM measurements as well [58]. Stochastic simulations can be used for the optimization of experimental protocols by aiding in the selection of experimental parameters. Tandem simulations can also be used alongside actual experiments to diagnose the cause of variability, and to assess whether a measurement is limited by fundamental sources of noise, mechanical vibrations, or simply due to true variability of the physical system being measured.

ACKNOWLEDGMENTS

We acknowledge valuable discussions with Philip Egberts and Jason Cleveland, and funding from NSERC and FQRNT.

APPENDIX

This Appendix outlines certain technical details regarding the integration of PSDs, as well as their estimation using Fourier transform methods. Before proceeding, the reader should be familiar with the concepts of "spectral leakage" and "aliasing," described in the following paragraphs.

Spectral leakage occurs because a perfect sine wave (infinitely long in time) cannot be measured in practice: it is inevitable that its measurement is performed within a finite time window. Truncating a perfect sine wave introduces new frequency components into the measurement. In other words, the power of the single frequency sine wave "leaks" into

adjacent frequencies. This description generalizes to signals of any arbitrary shape.

Aliasing is due to the finite-time sampling of a signal. This effect causes familiar phenomena such as the appearance of car wheels rotating backwards in films, for example. Similarly, in signal processing, slowly sampling a fast signal "aliases" fast frequencies onto slower ones. For a sampling rate of 100 Hz, different sine waves at 10, 90, 110, 190, 210 Hz, etc., all appear identical and indistinguishable from a 10-Hz sine wave.

1. Numerical details for PSD integration

In order to numerically integrate an average power spectral density $\langle n^2 \rangle$ weighted by a noise transfer function $|k_\varphi|$, both vectors must share an identical frequency basis. Typically, the system noise is characterized before the experiment over a large frequency range such as 10^{-2}–10^{6} Hz, which is much wider than the experiment bandwidth of, say, 10^{0}–10^{3} Hz. However, noise at frequencies below and above the experiment bandwidth affect the measurement by the mechanisms of spectral leakage and aliasing, respectively.

We use zero-padding and zero-interleaving in the time domain as methods for extending the frequency range of $|k_\varphi|$ to match the frequency range of $\langle n^2 \rangle$ for an accurate integration of the entire frequency spectrum.

Before performing the Fourier transform on the measurement sampling function w_φ to obtain $|k_\varphi|$, zero-padding should be adjusted to match the frequency resolution of $|k_\varphi|$ to the frequency resolution of $\langle n^2 \rangle$, as illustrated in Fig. 12. Zero-padding takes into account the effects of spectral leakage on the measurement.

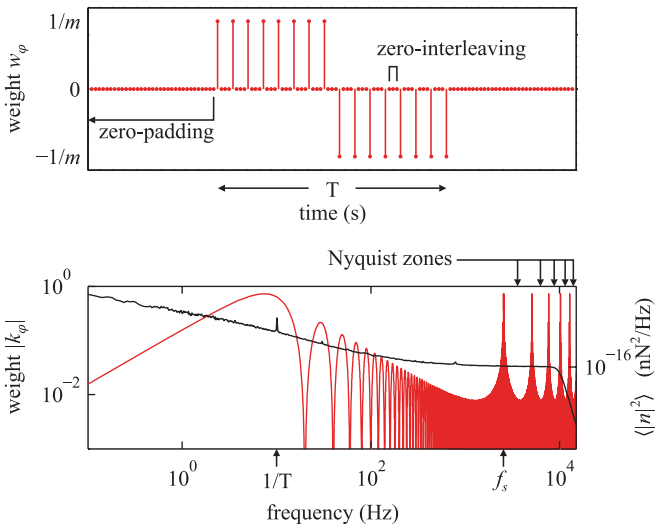

FIG. 12. (Color online) The sampling function w_φ of a hysteresis measurement is shown. The number of data points was selected as $m = 16$ for plotting purposes only. Zero-padding the signal increases the frequency resolution and zero-interleaving increases the maximum frequency. The creation of Nyquist zones above the sampling frequency simulates aliasing. Both are tuned such that the frequency basis of the noise transfer function $|k_\varphi|$ corresponds to the measured detection noise $\langle n^2 \rangle$. The $|k_\varphi|$ here is shown for $m = 256$ in the second plot, for a scan rate of 10 Hz.

Next, zero-interleaving is used to increase the maximum frequency of $|k_\varphi|$ beyond the roll-off frequency of $\langle n^2 \rangle$. Zero-interleaving takes into account the effects of aliasing during analog-to-digital conversion, by extending $|k_\varphi|$ to higher frequencies. Every zero that is interleaved generates a new Nyquist zone, as shown in Fig 8(b). Typical ADCs have a fixed integration time per sample, such that lowering the sampling rate from its maximum causes "dead time" in between samples leading to aliasing; this process is mimicked by zero-interleaving as illustrated by Fig. 12(a). Note that the peaks for each Nyquist zone indicate the unnecessary and undesirable aliasing of high-frequency noise into the measurement. Two methods can be used to reduce aliasing: (1) oversampling the signal by operating the ADC at its maximum sampling rate removes the Nyquist zones as all the high-frequency noise is measured and averaged by the measurement sampling function; (2) low-pass filtering the signal (prior to sampling) at a frequency below the Nyquist frequency ($f_s/2$), such that no noise remains in the Nyquist zones. Both of these methods prevent high-frequency noise from aliasing into the measurement bandwidth.

The Fourier transform of the zero-padded and zero-interleaved sampling function is shown in Fig. 12, overlaid on the detection noise. In practice, the matching of both frequency bases cannot be achieved perfectly, such that appropriate interpolation is used as a final step. Finally, the variance of the measurement can be computed by the dot product: $|k_\varphi|^2 \cdot \langle n^2 \rangle$.

2. Accurate PSD estimation

The covariance matrix V_f is experimentally estimated by the average power spectral density vector $\langle n^2 \rangle$. The accuracy of this estimation is mainly limited by spectral leakage, drift, and statistical biases, which are the topics of this section.

a. Spectral leakage and windowing. Omitting windowing during the Fourier transform (i.e., using a rectangular window) causes biases when estimating $1/f$ noise due to spectral leakage. For example, a $2\times$ bias in the estimation of $1/f^2$ noise occurs because the rectangular window falls off with $1/f^2$ itself; convoluting a $1/f^2$ noise with a $1/f^2$ window results in $2/f^2$. Furthermore, spectral leakage correlates adjacent frequencies of a spectral density, thereby violating the assumption of a diagonal spectral covariance matrix presented in Eq. (9). This section presents a method for reducing the effects of spectral leakage compared to traditional methods used for calculating $\langle n^2 \rangle$ from a time series.

Bartlett's method [59] for power spectral density estimation amounts to dividing a time series into smaller segments, calculating their respective spectral densities, and averaging them. This reduces the variance of the $\langle n^2 \rangle$ estimation, at the expense of reducing the frequency resolution. Spectral leakage in this situation becomes much more problematic, because the duration of each time series is much shorter. Welch's method [60] offsets this problem to a certain extent, by windowing each time series. This reduces the effects of the time-window edge discontinuities of each time series. However, significant spectral leakage still occurs regardless of the windowing function that is used.

Arguably, the main advantage of these two methods is the reduction in necessary computing power, which was

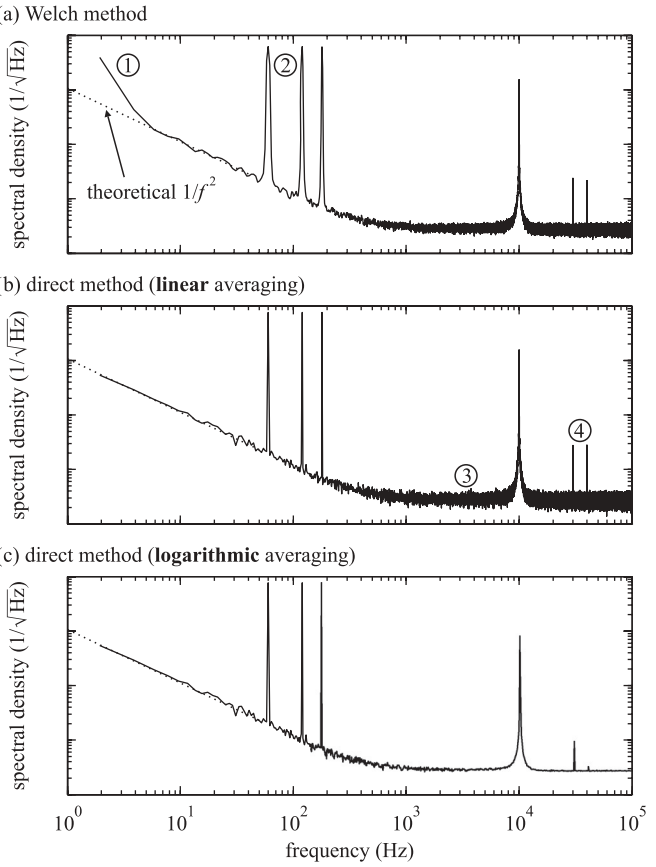

FIG. 13. The same simulated noise data were used for all three power spectral density estimations. The line noise peaks (60, 120, 180 Hz) were modeled as perfect sine waves with an arbitrary phase. (a) The Welch method causes estimation bias at low frequencies (1) and spectral leakage (2). (b) The direct method solves these problems, but the increase in variance at high frequencies (3) remains illusory. Also, the bandwidth of noise peaks can become thinner than the graphical linewidth of 0.176 mm (4), thereby distorting the perception of the power of each noise source. (c) These problems are solved using logarithmic averaging. All methods used a Hann window when computing the Fourier transform.

scarce in the 1960s when these methods were developed. The availability of modern computing led us to adopt a more simple and direct method of PSD estimation for which spectral leakage can mostly be eliminated: a single PSD of the entire time series is calculated (after appropriate windowing) followed by averaging groups of adjacent frequency bins to reduce the variance [61]. The same trade-off between variance and resolution remains, but spectral leakage can be tuned arbitrarily small and the estimation bias at low frequencies is prevented. This is illustrated in Figs. 13(a) and 13(b), where Welch's method is contrasted with the "direct method."

Figure 13(c) uses the same data as Fig. 13(b) which were subsequently averaged using logarithmically spaced frequency bins, making it more suitable for plotting on a logarithmic scale. Basically, for a linearly spaced PSD in the frequency range $[a,b]$, the N-sized logarithmically spaced PSD is generated by grouping all data points that fall within the frequency range $[ae^{(i-1)\ln(b/a)/N}, ae^{i\ln(b/a)/N}]$, where i is the bin index from 1 to N.

This presents a quantitatively more accurate visual representation of the power spectral density for applications where a logarithmic frequency axis is more appropriate, such as static AFM. In reality, the variance of shot noise *decreases* towards the right of a logarithmic graph where there are many more data points which estimate shot noise for a given fractional bandwidth $\Delta\omega/\omega$. (Note that in certain applications, such as dynamic AFM, the figure of merit is the linear bandwidth $\Delta\omega$ in which case logarithmic averaging is not appropriate.)

Also, certain noise peaks appear significant in Fig. 13(b), but should be disregarded because their bandwidth is very small and therefore the corresponding power is insignificant; for example, of the two peaks labeled (4) that appear identical in Fig. 13(b), the one to the right has a bandwidth that is 30× smaller, which is apparent by inspection of the graph plotted in Fig. 13(c). We stress that logarithmic averaging in Fig. 13(c) is used only for visual representation of PSDs, whereas the direct method is used in noise analysis as it contains more information.

b. 1/f noise vs drift. As a disclaimer, it should be noted that the distinction between drift and $1/f$ noise is the topic of debate, and the discussion presented herein is based on the assumption that drift is caused by a deterministic process, while $1/f$ noise is inherently stochastic. As described in Sec. V A, the distinction between stochastic and deterministic noise can be a matter of perspective.

Although $1/f$ noise and drift can be difficult to distinguish experimentally [62], they require different treatment. For example, the average PSDs of both linear drift and Brownian motion converge to a $1/f^2$ power spectrum. Brownian motion is a stochastic process, with a zero mean, and the PSD correctly estimates its variance (assuming spectral leakage is appropriately handled). In contrast, drift can be thought of as a deterministic process on the short time scale of a single experiment. Drift has a predictable effect on the mean (rather than variance) of a signal, and therefore is not accurately described by a PSD.

Properly quantifying the variability of drift itself would require a PSD measurement taken across many days, and a strong component is expected at a frequency of day^{-1} because temperature cycles between day and night are prominent. In fact, the $1/f^2$ power spectrum of linear drift at shorter experimental time scales is simply a result of spectral leakage: the long time scale drift oscillations are truncated by a finite time window, and the edge effects of this truncation result in spectral leakage. The $1/f^2$ power spectrum of linear drift relates only to the rectangular windowing function, and can be changed to $1/f^8$ simply by choosing a Hann windowing function, for example. Furthermore, a PSD of any finite signal is equivalent to the PSD of an infinite periodic signal; the PSD of linear drift is equivalent to the PSD of an infinite sawtooth wave. The $1/f^2$ spectrum quantifies the harmonics of that sawtooth wave, which is clearly an incorrect representation of linear drift. In summary, the observed spectrum of drift is simply an artifact of the finite Fourier transform method, and its shape provides no information about the actual drift.

For these reasons, drift must be subtracted from any signal before taking its PSD. It is inevitable that any estimate of drift (to be subtracted) will be corrupted by $1/f$ noise, and conversely the measurement of $1/f$ noise will be altered

by the drift subtraction. For short experiments, a simple linear drift can be a very good approximation, while longer experiments may require the subtraction of nonlinear drift that may be modeled as polynomial drift, or logarithmic creep, for example. All the PSDs reported in this paper were calculated after subtraction of a linear drift, judging that the remaining noise can accurately be described as "variance" on the time scales considered.

c. Amplitude averaging. The amplitude spectral density $|\boldsymbol{n}|$ rather than the power spectral density $|\boldsymbol{n}|^2$ is often used in noise analysis. As derived in Sec. IV, each random variable in $|\boldsymbol{n}|^2$ is exponentially distributed, implying that each random variable in $|\boldsymbol{n}|$ follows a Rayleigh distribution. Squaring and averaging are noncommutative operations: $\langle|n|^2\rangle \neq \langle|n|\rangle^2$ (the average of the squares is not the square of the average). However, the two relate by a calculable statistical bias, and the average power spectral density of stochastic noise can be determined from the average amplitude spectral density by

$$\langle|n|^2\rangle = \frac{4}{\pi}\langle|n|\rangle^2. \tag{A1}$$

[1] G. Binnig, C. F. Quate, and C. Gerber, Phys. Rev. Lett. **56**, 930 (1986).

[2] Y. Martin, C. C. Williams, and H. K. Wickramasinghe, J. Appl. Phys. **61**, 4723 (1987).

[3] T. R. Albrecht, P. Grutter, D. Horne, and D. Rugar, J. Appl. Phys. **69**, 668 (1991).

[4] S. D. Bennett, L. Cockins, Y. Miyahara, P. Grütter, and A. A. Clerk, Phys. Rev. Lett. **104**, 017203 (2010).

[5] J. P. Cleveland, T. E. Schaffer, and P. K. Hansma, Phys. Rev. B **52**, R8692 (1995).

[6] A. Roters, M. Gelbert, M. Schimmel, J. Rühe, and D. Johannsmann, Phys. Rev. E **56**, 3256 (1997).

[7] S. Kawai, F. F. Canova, T. Glatzel, A. S. Foster, and E. Meyer, Phys. Rev. B **84**, 115415 (2011).

[8] A. Gannepalli, A. Sebastian, J. Cleveland, and M. Salapaka, Appl. Phys. Lett. **87**, 111901 (2005).

[9] C. L. Degen, M. Poggio, H. J. Mamin, and D. Rugar, Phys. Rev. Lett. **99**, 250601 (2007).

[10] L. Nony, A. Baratoff, D. Schär, O. Pfeiffer, A. Wetzel, and E. Meyer, Phys. Rev. B **74**, 235439 (2006).

[11] M. Gauthier, R. Pérez, T. Arai, M. Tomitori, and M. Tsukada, Phys. Rev. Lett. **89**, 146104 (2002).

[12] J. Melcher, C. Carrasco, X. Xu, J. L. Carrascosa, J. Gómez-Herrero, P. José de Pablo, and A. Raman, Proc. Natl. Acad. Sci. USA **106**, 13655 (2009).

[13] D. Kiracofe, J. Melcher, and A. Raman, Rev. Sci. Instrum. **83**, 013702 (2012).

[14] A. Labuda, J. R. Bates, and P. H. Grütter, Nanotechnology **23**, 025503 (2012).

[15] G. Meyer and N. M. Amer, Appl. Phys. Lett. **53**, 1045 (1988).

[16] T. Fukuma, T. Ichii, K. Kobayashi, H. Yamada, and K. Matsushige, Appl. Phys. Lett. **86**, 034101 (2005).

[17] F. N. Hooge, T. G. M. Kleinpenning, and L. K. J. Vandamme, Rep. Prog. Phys. **44**, 479 (1981).

[18] P. Bak, C. Tang, and K. Wiesenfeld, Phys. Rev. Lett. **59**, 381 (1987).

[19] Y. Sang, M. Dubé, and M. Grant, Phys. Rev. E **77**, 036123 (2008).

[20] C. Zener, Phys. Rev. **52**, 230 (1937).

[21] V. T. Srikar and S. D. Senturia, J. Microelectromech. Syst. **11**, 499 (2002).

[22] S. Prabhakar and S. Vengallatore, J. Micromech. Microeng. **17**, 532 (2007).

[23] J. E. Sader, J. Appl. Phys. **84**, 64 (1998).

[24] S. Rast, C. Wattinger, U. Gysin, and E. Meyer, Nanotechnology **11**, 169 (2000).

[25] M. T. Clark, J. E. Sader, J. P. Cleveland, and M. R. Paul, Phys. Rev. E **81**, 046306 (2010).

[26] A. Labuda and P. Grütter, Langmuir **28**, 5319 (2012).

[27] P. Paolino and L. Bellon, Nanotechnology **20**, 405705 (2009).

[28] J. B. Thompson, B. Drake, J. H. Kindt, J. Hoskins, and P. K. Hansma, Nanotechnology **12**, 394 (2001).

[29] R. Guerra, A. Vanossi, and M. Urbakh, Phys. Rev. E **78**, 036110 (2008).

[30] A. Socoliuc, E. Gnecco, S. Maier, O. Pfeiffer, A. Baratoff, R. Bennewitz, and E. Meyer, Science (NY) **313**, 207 (2006).

[31] M. Lantz, D. Wiesmann, and B. Gotsmann, Nat. Nanotechnol. **4**, 586 (2009).

[32] C. M. Mate, G. M. McClelland, R. Erlandsson, and S. Chiang, Phys. Rev. Lett. **59**, 1942 (1987).

[33] A. Labuda, W. Paul, B. Pietrobon, R. B. Lennox, P. H. Grütter, and R. Bennewitz, Rev. Sci. Instrum. **81**, 083701 (2010).

[34] A. Labuda, F. Hausen, N. N. Gosvami, P. H. Grütter, R. B. Lennox, and R. Bennewitz, Langmuir **27**, 2561 (2011).

[35] A. Labuda, K. Kobayashi, Y. Miyahara, and P. Grütter, Rev. Sci. Instrum. **83**, 053703 (2012).

[36] T. Fukuma, M. Kimura, K. Kobayashi, K. Matsushige, and H. Yamada, Rev. Sci. Instrum. **76**, 053704 (2005).

[37] A. Garcıia-Valenzuela and J. Villatoro, J. Appl. Phys. **84**, 58 (1998).

[38] M. B. Viani, T. E. Schäffer, A. Chand, M. Rief, H. E. Gaub, and P. K. Hansma, J. Appl. Phys. **86**, 2258 (1999).

[39] P. Brockwell and R. Davis, *Time Series: Theory and Methods, 2nd ed.* (Springer, New York, 1991).

[40] E. Milotti, Phys. Rev. E **51**, 3087 (1995).

[41] S. F. Nørrelykke and H. Flyvbjerg, Rev. Sci. Instrum. **81**, 075103 (2010).

[42] J. Brophy, Phys. Rev. **166**, 827 (1968).

[43] I. Vragovic, J. M. Molina, R. Prieto, M. Duarte, J. Narciso, and E. Louis, Phys. Rev. E **80**, 066123 (2009).

[44] S. Cheng, B. Luan, and M. O. Robbins, Phys. Rev. E **81**, 016102 (2010).

[45] A. Schirmeisen and H. Hölscher, Phys. Rev. B **72**, 045431 (2005).

[46] A. Labuda and P. H. Grütter, Rev. Sci. Instrum. **82**, 013704 (2011).

[47] R. W. Carpick, D. F. Ogletree, and M. Salmeron, Appl. Phys. Lett. **70**, 1996 (1997).

[48] E. Gnecco, R. Bennewitz, T. Gyalog, C. Loppacher, M. Bammerlin, E. Meyer, and H. J. Guntherodt, Phys. Rev. Lett. **84**, 1172 (2000).
[49] A. Socoliuc, R. Bennewitz, E. Gnecco, and E. Meyer, Phys. Rev. Lett. **92**, 134301 (2004).
[50] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.86.031104 for 100 simulated friction loops.
[51] S. Maier, Y. Sang, T. Filleter, M. Grant, R. Bennewitz, E. Gnecco, and E. Meyer, Phys. Rev. B **72**, 245418 (2005).
[52] S. Y. Krylov and J. W. M. Frenken, J. Phys. Condens. Matter **20**, 354003 (2008).
[53] Y. Dong, A. Vadakkepatt, and A. Martini, Tribol. Lett. **44**, 367 (2011).
[54] Y. Sang, M. Dubé, and M. Grant, Phys. Rev. Lett. **87**, 174301 (2001).
[55] M. H. Müser, Phys. Rev. B **84**, 125419 (2011).
[56] Y. Dong, H. Gao, and A. Martini, Europhys. Lett. **98**, 16002 (2012).
[57] A. Labuda, Y. Miyahara, L. Cockins, and P. H. Grütter, Phys. Rev. B **84**, 125433 (2011).
[58] A. Labuda, M. Lysy, and P. Grütter, Appl. Phys. Lett. (in press).
[59] M. S. Bartlett, Nature **161**, 686 (1948).
[60] P. Welch, IEEE Trans. Audio Electroacoust. **15**, 70 (1967).
[61] P. J. Daniell, J. R. Stat. Soc., Ser. B **8**, 88 (1946).
[62] T. Brettschneider, G. Volpe, L. Helden, J. Wehr, and C. Bechinger, Phys. Rev. E **83**, 041113 (2011).