Ultramicroscopy 242 (2022) 113614

Contents lists available at ScienceDirect

Ultramicroscopy

journal homepage: www.elsevier.com/locate/ultramic

Data-driven dynamics-based optimal filtering of acoustic noise at arbitrary location in atomic force microscope imaging

Jiarong Chen, Qingze Zou *

Department of Mechanical and Aerospace Engineering, Rutgers University, Piscataway, NJ 08854, USA

ARTICLE INFO

Keywords:
Acoustic noise filtering
Atomic force microscopy
Coherence minimization
Topography imaging

ABSTRACT

This paper presents a post-filtering approach to eliminate distortions in atomic force microscope (AFM) images caused by acoustic noise from an unknown location. AFM operations are sensitive to external disturbances including acoustic noise, as disturbances to the probe-sample interaction directly results in distortions in the sample images obtained. Although conventional passive noise cancellation has been employed, limitation exists and residual noise still persists. Advanced online control techniques face difficulty in capturing the complex noise dynamic and limited system bandwidth imposed by robustness requirement. In this work, we propose a dynamics-based optimal filtering technique to remove the acoustic-caused distortions in AFM images. A dictionary-approach is integrated with time-delay measurement to localize the noise source and estimate the corresponding acoustic dynamics. Then a noise-to-image coherence minimization approach is proposed to minimize the acoustic-caused image distortion via a gradient-based optimization to seek an optimal modulator to the acoustic dynamics. Finally, the filter is obtained as the finite-impulse response of the optimized acoustic dynamics. Experimental implementation is presented and discussed to illustrate the proposed technique.

1. Introduction

In this paper, a data-driven dynamics-based post-filtering approach is proposed to eliminate acoustic-caused distortions in atomic force microscope (AFM) images. As an essential tool for nanoscale interrogation and measurement [1–3], AFM is sensitive to external and internal disturbances including acoustic noise, as disturbances to the tip-sample interaction directly result in distortions in AFM images [4]. Although the external-disturbance-caused adverse effects can be alleviated via conventional passive methods [5–8] or active control technique [9], residual noise distortion still persists, resulting in image distortion and loss of image quality. Thus, this work aims to develop a post-filtering approach to eliminate the acoustic noise distortions in AFM images.

Maintaining the tip-sample interaction closely around the desired value is important in AFM applications, including imaging [1–3], nanomechanical measurement [10–12], and tip-based nanofabrication [13–15]. Extraneous tip vibration relative to the sample can be induced by disturbances including acoustic noise and seismic vibrations, or by internal excitement of the dynamics and hysteresis adverse effects of the nanopositioning systems (from the piezo actuator to the cantilever tip). In contrast to the efforts to tackle internal adverse effects through hardware improvement to increase the bandwidth of the nanopositioning system [16,17], and/or through software/algorithm innovations of advanced control techniques [18,19]. Little has been done to compensation for the external disturbances [20] beyond the use of conventional passive vibration/noise isolation apparatus [5–8] such as vibration isolation table [21] and acoustic enclosure [22]. These passive noise/vibration isolation apparatus prevents the integration of other measurement instruments such as inverted optical microscope to the AFM platform [23], and residual noise distortion still exists [9]. Recently, inversion-based feedforward controllers [9,20] are introduced to compensate for the undesired piezo displacement. The performance of such an online active noise control can be hurdled by hardware constraints (e.g., online computation power and data acquisition speed) and limited system bandwidth imposed by the robustness requirements. Thus, post-filtering techniques shall be developed to improve the AFM image quality affected by the acoustic noise.

Challenges exist in filtering the acoustic-noise-caused distortions in AFM images. Conventional frequency-domain filtering approach (e.g., low-passor band-pass filtering) is ineffective as the acoustic-noise effect is largely random, broadband, and does not decay with frequency increase. As a result, sample topography can be significantly distorted or

* Corresponding author.
E-mail address: qzzou@soe.rutgers.edu (Q. Zou).

https://doi.org/10.1016/j.ultramic.2022.113614

Available online 16 September 2022

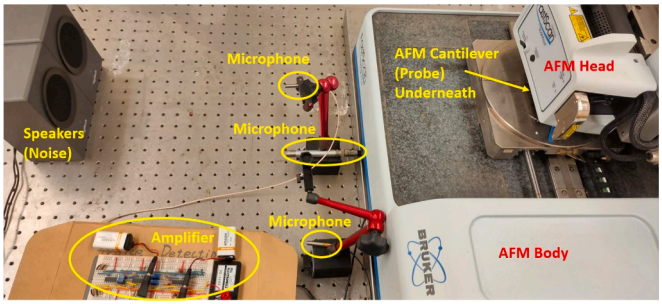

Fig. 1. An experimental setup for studying acoustic noise effect on AFM operation, where it is assumed that the location of the noise source (speakers) is unknown while the sensors (microphones) are placed at fixed and known locations.

removed along with the noise when these filters are used. While these limitations of conventional filtering might be alleviated via the filtered-x least-mean-square algorithms [24] and its modifications [25,26] to account for the coupling between the noise and the sample topography, such techniques rely on acquiring accurate error of the residual noise effect, i.e., an acoustic-caused image is needed, whereas it is mixed and coupled with the "true" sample topography that is unknown in general. Alternatively, the acoustic-caused image distortion might be decoupled from the sample topography image by exploring a model-based approach along the Kalman-filtering framework [27]. However, it is challenging to capture the complicated acoustic noise dynamics of broadband nature in practices accurately and robustly. This difficulty becomes even more serious when the location of noise source is unknown a priori—due to the unknown noise source, the un-collated sensing results in large uncertainties in the acoustic dynamics model (used in the Kalman-filter design) and the noise signal acquired, as the acoustic noise dynamics heavily depends on the noise propagation path, and thereby, varies significantly when the relative position between the noise source and the sensor changes. Thus, for the post-filtering technique to be effective, the complexity and the pronounced location-dependent variation of the acoustic noise dynamics must be accounted for.

The main contribution of this paper is the development of a data-driven dynamic-based technique to eliminate the acoustic noise-caused distortion in AFM imaging. Specifically, the filter is designed directly as a finite-impulse-response (FIR) representation of the underline primary acoustic dynamics (PAD) response. First, to account for the unknown noise location effect on the PAD, the unknown noise source is localized based on the time-delay measurement method [28]. Then, a dictionary-based approach is proposed to estimate the corresponding PAD based on the estimated noise location. Secondly, an acoustic-dynamics modulator is introduced to the estimated PAD to minimize the acoustic-caused image distortion via coherence minimization. Particularly, the modulator is sought via the gradient-based iterative algorithm to minimize the coherence between the acoustic noise and the residual acoustic-caused image distortion iteratively. Although the notion of noise-and-signal coherence has been exploited previously [29,30] in speech signal filtering and enhancement, the proposed approach is fundamentally different in exploring a gradient-based approach to directly minimize the noise-and-signal coherence to optimize the filter (i.e., the acoustic dynamics model). The proposed filtering scheme is implemented on AFM image examples, and the experimental results show that the image distortion can be significantly reduced.

2. Acoustic-caused AFM image distortion: challenges, & problem formulation

2.1. Acoustic-Noise effect on atomic force microscope operation: Imaging example

Eliminating external disturbances including acoustic noise is critical to AFM applications. The basic principle of AFM is to utilize a nanometer-size cantilever-tip attached to the AFM head (see Fig. 1) to interact with a sample [1–3], and manipulate the probe-sample interaction with nanoscale precision—-to interrogate topography and material properties of the sample, or to modify the sample and/or generate patterns on the sample [31]. Thus, external disturbances like acoustic noise can result in extraneous perturbation to the probe-sample interaction, and thereby, loss of precision and quality in AFM applications.

More specifically, acoustic noise effect reduces the AFM imaging quality. For example, in tapping mode (TM) imaging, the cantilever tip is excited (usually using a small piezo stack actuator, called a dither piezo) and taps on the sample surface, when the tip is scanning across the surface, and the tapping amplitude is maintained around a pre-chosen constant level under feedback control. [31]. Then, a sample topography image is obtained from the (vertical) displacement of the cantilever [2]. In the presence of environmental acoustic noise, however, the mechanical structure of the AFM (on which the cantilever is mounted, see Fig. 1) can be excited and vibrate, resulting in unwanted cantilever vibration with respect to the sample and thereby, image distortion. As an example, an image of a silicon wafer obtained under environmental acoustic noise (introduced by using the speaker in Fig. 1) is shown pronounced image distortion with respect to one obtained without the induced noise. (compare Fig. 2(a) and (b))

It is challenging to eliminate such acoustic-noise-caused distortions in AFM image. Conventional filtering techniques based on frequency-domain separation (e.g., low-pass, band-pass, or notch filter) are ineffective as the noise-caused distortion is largely coupled with the sample topography in the image obtained—These techniques will filter out and smear the sample topography features. Further challenges arise as the acoustic-noise dynamics is largely random, broadband, and do not decay with the frequency increases (see Fig. 3 for an example). These characteristics of the acoustic noise effect also present challenges to the model-based approaches along the Kalman-filtering framework [27]—-it is challenging to capture such a complicated acoustic noise dynamics (See Fig. 3) via a parameterized model accurately and robustly. This difficulty becomes even more so when the location of the noise source is arbitrary and unknown a priori—The arbitrarily noise location results in un-collocated sensing (see Fig. 1) and thereby, uncertainties in the acoustic dynamics and the noise signal measured—the acoustic noise dynamics depends on the noise propagation path, and varies significantly as the distance and direction of the noise source and the sensor changes. Thus, the complexity and the pronounced location-dependent uncertainty of the acoustic noise dynamics must be accounted for.

2.2. Filtering of acoustic-noise in AFM imaging: Problem formulation

We consider the scenario where the noise comes from an *arbitrary and unknown* but fixed location behind the sensors (whose locations are fixed and known)—as illustrated in Fig. 1 and appears in many AFM applications. Without loss of generality, we assume that

Assumption 1. The acoustic noise n[k] is a zero-mean, band-limited white noise, and the variation of the primary acoustic noise dynamics (PAD) is quasi static.

The PAD is the dynamics from the noise signal (as the input) to the AFM image signal (as the output response), under the source-sensor collocated condition, i.e., the sensor (microphone) is placed at the noise source location, see$G_N(z)$ in Fig. 3. Moreover, the PAD can be

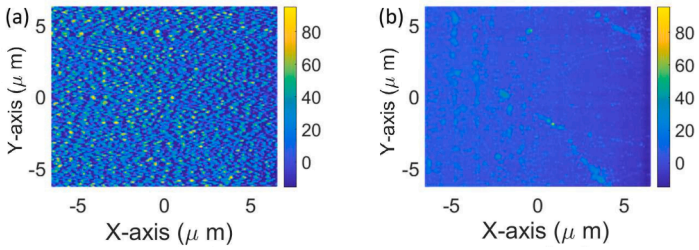

Fig. 2. Comparison of an AFM image of a silicon wafer sample obtained (a) with acoustic noise induced via the speakers in Fig. 1 to (b) that without the induced acoustic noise.

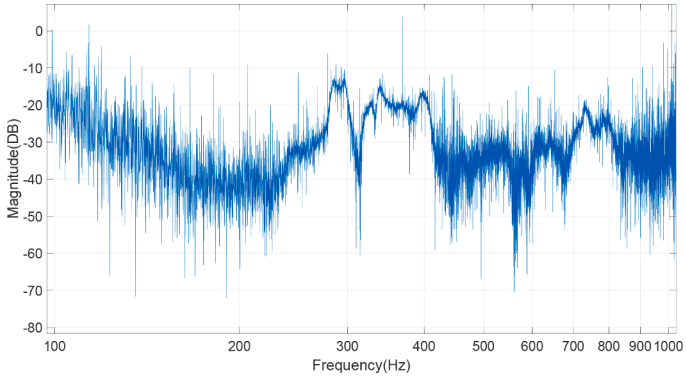

Fig. 3. An exemplary frequency response (magnitude part) of a primary acoustic dynamics (PAD) for the AFM setup in Fig. 1.

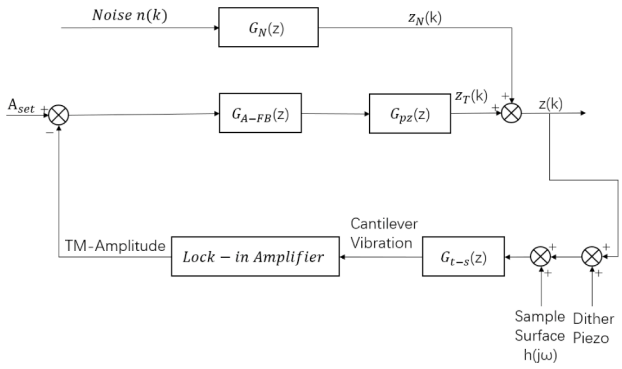

Fig. 4. Schematic block diagram of the AFM imaging process with acoustic noise.

modeled as a linear time-invariant dynamics(LTI) and the variation of the PAD is mainly caused by the change of the noise source location, i.e., the noise propagation route, and thereby, is largely unchanged during a given imaging process but can vary in day to day operation.

As depicted in the block diagram in Fig. 4, in the presence of acoustic noise, the total AFM image signal z[k] becomes:

$$z[k] = z_T[k] + z_n[k], \text{ for } k = 1, \ldots, N_I, \tag{1}$$

where$z_T[k]$ and$z_n[k]$ are the$z$-axis piezo displacement caused by the sample topography and that due to the acoustic noise, respectively, andN_I is the total number of sampling data acquired in the given imaging process. As$z[k]$ is used to plot the sample topography image, we call, in the following,$z[k]$,$z_T[k]$, and$z_n[k]$ the raw image signal, the true image signal, and the image noise signal, respectively. The image noise$z_n[k]$ is related to the acoustic noise$n[k]$ via

$$z_n[k] = n[k] * g_N[k], \tag{2}$$

where$g_N[k]$ is the impulse response of the PAD (from the noise source to the AFM cantilever), and '$*$' denotes the convolution operation, respectively.

Thus, a filtered image signal,$z_F[\cdot]$, can be obtained from the estimated image noise signal$\widehat{z}_n[k]$ via

$$z_F[k] = z[k] - \widehat{z}_n[k], \text{ for } k = 1, \ldots, N_I \tag{3}$$

$$\text{where } \widehat{z}_n[k] = \widehat{n}[k] * \widehat{g}_N[k] \tag{4}$$

with$\widehat{n}[k]$ and$\widehat{g}_N[k]$ the estimated acoustic noise and the estimated PAD (presented in the impulse response form), respectively. Thus, the filtering quality can be quantified by the residual image error$e_r[k]$

$$e_r[k] = z_F[k] - z_T[k]. \tag{5}$$

By Assumption 1,$\widehat{z}_n[k]$ is also a wide stationary random process process [32], so are$z_F[k]$ and$e_r[k]$. Hence, by Eqs. (3), (4), the expectation of the residual image error,$\mathbf{E}(e_r[\cdot]) = 0$.

Uncertainties exist in the estimated noise source$\widehat{n}[k]$ and the PAD,$\widehat{g}_N[k]$. Specifically, the noise and the sensor being un-collocated (due to the noise source location being arbitrary) implies that the measured noise$n_m[k]$ can be largely different from the "true" noise, and the PAD from the measured noise (acquired at the sensor location) to the image response$g_{N,m}[k]$, is also different from the "true" PAD$g_N[k]$. This difference increases pronouncely as the distance between the sensor and the noise source location or the direction difference (i.e., the noise-to-AFM direction vs. the sensor-to-AFM direction) increases. Thus, central to the acoustic-noise filtering of AFM image is to estimate$\widehat{n}[k]$ and$\widehat{g}_N[k]$ accurately from the measured noise and image response,$n_m[k]$ and$z_n[k]$, such that$\widehat{n}[k] \to n[k]$, and$\widehat{g}_N[k] \to g[k]$.

Data-driven Dynamics-based Optimal Filtering (DD-DOF) of Acoustic-Noise in AFM Image Let Assumption 1 hold, and the noise source location be fixed, arbitrary, and unknown a priori, and let three sensors located at three fixed locations be used to acquire the noise signal along with the AFM image data during the noise-effected imaging process, then the problem of DD-DOF in AFM imaging is to:

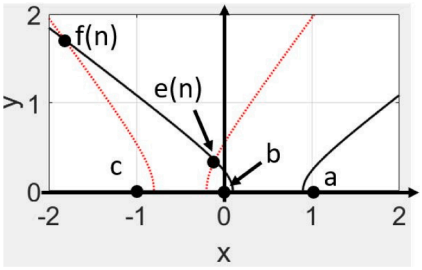

Fig. 5. Schematic representation of the acoustic source localization using time delay measurement, where *a*, *b*, and *c*, denote the locations of three microphones, the black-line and red-line are the hyperbola curves on which the possible noise source locations are. and *e* and *f* denote the location of the noise source *n*, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$\mathbb{O}-1$
Directly estimate both the noise signal and the corresponding PAD from the measured data,

$\mathbb{O}-2$
Construct the filter$\widehat{g}_N[k]$ by directly using the measured data, and minimize the following noise caused image distortion:

$$\min J_r = \sum_{k=1}^{N_I} \mathbf{E}\{e_r[k]\}^2 = \sum_{k=1}^{N_I} \mathbf{E}\{z_F[k] - z_T[k]\}^2 \tag{6}$$

In contrast to parameterized-model-based methods [27,33] where a model of the PAD needs to be identified a priori, the proposed data-driven approach aims to avoid not only the modeling process and the related issues, but also the performance-robustness trade-off—the filter can be automatically updated along with the measured data, and thereby, maintain the performance in the presence of dynamics and signal variations.

3. Data-driven acoustic filter optimization

We proceed by achieving the two objectives in order.

3.1. $\mathbb{O}-1$ *data-driven estimation of noise dynamics with unknown source location*

First, the noise source is localized through a time-delay measurement, and then used to estimate the PAD and the noise via a data-driven, dictionary-based approach.

3.1.1. *Acoustic source localization*

In the planar noise source scenario, localizing the noise source is equivalent to determining the coordinate of an unknown location*n* (noise source) with respect to other three known locations*a*,*b*, and*c* (three sensors). The possible locations of the noise source must fall onto a hyperbola (the black lines in Fig. 5) with its foci at*a* and*b*, respectively, (as the distance differences between the unknown point*n* and any two fixed points, e.g., between*n* and*a* vs between*n* and*b*,$\Delta d_{a-b,n} = |d_{a,n} - d_{b,n}|$, is fixed). By choosing the frame with point*b* as its origin and point*a* at coordinate (1, 0) (see Fig. 5), this hyperbola is given by

$$\frac{(x-0.5)^2}{\left(\Delta d_{a-b,n}/2\right)^2} - \frac{y^2}{0.5^2 - \left(\Delta d_{a-b,n}/2\right)^2} = 1. \tag{7}$$

Similarly, the possible location of noise source*n* also falls onto a hyperbola of foci at locations*b* and*c* (the red line in Fig. 5), respectively,

$$\frac{(x+0.5)^2}{\left(\Delta d_{b-c,n}/2\right)^2} - \frac{y^2}{0.5^2 - \left(\Delta d_{b-c,n}/2\right)^2} = 1, \tag{8}$$

where to simplify the expression, the third microphone location is at c (-1, 0), and$\Delta d_{b-c,n} = |d_{b,n} - d_{c,n}|$. Thus, the above Eqs. (7) and (8) implies that the noise source must be located at one of the intersections of these two hyperbolas (point*e* or point*f* in Fig. 5): The actual location is determined by which microphone picks up the noise signal first, e.g., in Fig. 5, the noise is at location*e*(n) when the microphone*b* received the noise signal ahead of the microphone*c*.

Hence, the noise source location can be determined once the sensor-source distance differences$\Delta d_{a-b,n}$ and$\Delta d_{b-c,n}$ in Eqs. (7) and (8), respectively, are quantified, or equivalently, the time difference between the noise signals captured by the three microphones at*a*,*b*, and*c*,τ_{ab}, andτ_{bc}, respectively, are quantified. This time difference is measured/quantified through the following cross-correlation between these three acoustic signals, e.g., the time differenceτ_{ab} is quantified through cross-correlation$\mathbf{R}_{ab}[j]$ [34]

$$\mathbf{R}_{ab}[j] = \sum_{k=0}^{N_I-1} n_a[k] n_b[(k+j) \bmod N_I] \quad j = 0, 1, 2, \ldots, N_I - 1 \tag{9}$$

where$\bmod$ denotes the modulo operation. Then the time delayτ_{ab} (for quantifying the distance difference,$\Delta d_{a-b,n}$) is obtained as the maximum value of the cross-correlation

$$\tau_{ab} = \frac{1}{f_s} \operatorname{argmax}_j (\mathbf{R}_{ab}[j]). \tag{10}$$

wheref_s is the sampling frequency. The distance difference$\Delta d_{b-c,n}$ is quantified in the same manner. Finally, the coordinates of the noise location, $\vec{\ell}_n = (x_n, y_n)$, is obtained by solving the 2nd-order algebraic equations Eqs. (7) and (8) (numerically).

To improve the localization accuracy, the above process can be repeated by considering combinatorial pairings of the three microphones (total of 3), and using all the pairs of distance difference obtained to estimate the noise source location via averaging. Further improvement can be sought by using more microphones and more advanced methods beyond averaging.

3.1.2. *Dictionary-based primary acoustic noise dynamic estimation*

The above identified noise location is then used to estimate the PAD$\widehat{g}_N[\cdot]$ —The PAD from the noise source in the given imaging process to the image signal (under the sensor-noise collocation condition (*SNC*)). We propose a data-driven dictionary-based approach. Specifically, a dictionary of the PAD frequency responses measured at selected locations,$\mathbb{D}_{G_n}$, is constructed,

$$\mathbb{D}_{G_n} = \left\{ G_n\left(\vec{\ell}_{p,q}, j\omega_k\right) \middle| \vec{\ell}_{p,q} \in \Re^2, \omega_k \in \mathbb{S}_\omega, \right.$$

$$\left. p = 1, \ldots, N_r, \text{ and } q = 1, \ldots, N_\theta \right\} \tag{11}$$

$$\text{where } \mathbb{S}_\omega = \left\{ \frac{k\omega_s}{N_I} \middle| \, k = 0, 1, \ldots, N_I - 1, \text{and } \omega_s = 2\pi f_s \right\} \tag{12}$$

is the set of the sampled discretized frequencies,$G_n(\vec{\ell}_{p,q}, j\omega_k)$ is the PAD from location$\vec{\ell}_{p,q}$ ($\vec{\ell}_{p,q} = (r_p, \theta_q)$: the polar coordinate of position) to the AFM image signal measured under the noise-sensor collocated condition, andN_θ andN_r are the total number of PADs measured in the polar direction and radical direction, respectively (See Fig. 5), i.e., for$\vec{\ell}_{p,q} = (r_p, \theta_q)$,

$$r_p = \frac{r_{\max}}{N_r - i}, \quad \theta_q = \frac{\theta_{\max}}{q} \tag{13}$$

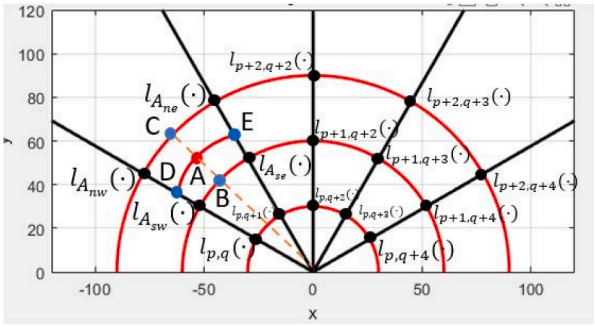

Fig. 6. Schematic illustration of the construction of the Dictionary of PADs, where the receiver (e.g., the AFM scanner in this work) is located at the origin, the black dots on the half-circle grids are the locations where the PADs in the dictionary have been measured, and blue dots at location point A, B, C, D and E denote the locations of PADs estimated by using the dictionary to estimate the PAD at location "A" (red-dot) where the unknown noise is,denote the interested noise locations, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

$$\text{for} \quad 1 \le p \le N_r \quad \text{and} \quad 1 \le q \le N_\theta \tag{14}$$

where$r_{\max}$ and$\theta_{\max}$ are the radius and angle (with respect to the AFM cantilever) corresponding to the maximum area of the noise effect considered, respectively.

Thus, the PAD at an unknown location can be estimated by using the dictionary$\mathbb{D}_{G_n}$ via, e.g., linear interpolation—More advanced estimation technique can be applied. Specifically, two constants,μ_r and$\mu_\theta \in (0,\ 1)$, are introduced to account for the dependence of the PADs on the distance (μ_r) and the direction (i.e., angle,μ_θ) with respect to the AFM cantilever, respectively. For example, to estimate the PAD at location$A$ in Fig. 6, first the smallest grid of the dictionary$\mathbb{D}_{G_n}$ within which location A resides can be determined by using the estimated location of A (identified in Sec. III.A), i.e.,A_{ne},A_{nw},A_{se}, andA_{sw} in Fig. 6 can be determined. Then the distance effect can be accounted for by using the Dictionary$\mathbb{D}_{G_n}$ to estimate the PAD at locations B and C (see Fig. 6),$\widehat{G}_{N,B}(j\omega_k)$ and$\widehat{G}_{N,C}(j\omega_k)$,

$$\widehat{G}_{N,B}(j\omega_k) = (1-\mu_\theta)G_n\left(\overrightarrow{\ell}_{A_{sw}}, j\omega_k\right) + \mu_\theta G_n\left(\overrightarrow{\ell}_{A_{se}}, j\omega_k\right) \tag{15}$$

$$\widehat{G}_{N,C}(j\omega_k) = (1-\mu_\theta)G_n\left(\overrightarrow{\ell}_{A_{nw}}, j\omega_k\right) + \mu_\theta G_n\left(\overrightarrow{\ell}_{A_{ne}}, j\omega_k\right) \tag{16}$$

for$\forall \omega_k \in \mathbb{S}_\omega$ where$G_n(\overrightarrow{\ell}_{A_{sw}}, j\omega_k)$,$G_n(\overrightarrow{\ell}_{A_{se}}, j\omega_k)$,$G_n(\overrightarrow{\ell}_{A_{nw}}, j\omega_k)$ and$G_n(\overrightarrow{\ell}_{A_{ne}}, j\omega_k) \in \mathbb{D}_{G_n}$ are the corresponding PADs at locationsA_{ne},A_{nw},A_{se}, andA_{sw}, respectively (see Fig. 6). Then the dynamic at location$A$ can be estimated via linear interpolation in the radical direction as

$$\widehat{G}_{N,A}(\omega_k) = (1-\mu_r)\widehat{G}_{N,B}(j\omega_k) + \mu_r \widehat{G}_{N,C}(j\omega_k) \tag{17}$$

Alternatively,$\widehat{G}_{N,A}(\omega_k)$ can also be estimated by accounting for the distance effect first—by estimating the PAD at location D and E (See Fig. 6),$\widehat{G}_{N,D}(j\omega_k)$ and$\widehat{G}_{N,E}(j\omega_k)$, respectively, then secondly, the direction effect by using the obtained$\widehat{G}_{N,D}(j\omega_k)$ and$\widehat{G}_{N,E}(j\omega_k)$ via linear interpolation in the polar direction. The accuracy of$\widehat{G}_{N,A}(\omega_k)$ can be further improved via averaging.

The above estimated PAD—in the non-parametric finite impulse response form—is used directly to quantify the noise-caused image distortion (i.e., the image noise signal$\widehat{z}_n[k]$ in Eq. (4)). Compared to parameterized-model based methods (e.g., the transfer function model), this data-driven approach is preferred as PAD is complicated with multiple poles and zeros, with no pronounced delaying as frequency increases (see Fig. 6 as an example)—parameterized-modeling approaches will not only result in a high-order model, and thereby, the complexity in computation and potential robustness issue with respect to parameter variations, but also can induce in extraneous modeling error.

3.1.3. *Noise signals estimation using dictionary*

Similarly, the effect of unknown noise location on the noise measurement can also be accounted for via the data-driven, dictionary-based approach. A dictionary of the noise propagation dynamics (NPD),$\mathbb{D}_{G_{np}}$, is constructed a priori,

$$\mathbb{D}_{G_{np}} = \left\{ G_{np}\left(\overrightarrow{\ell}_{p,q}, j\omega_k\right) \middle| \overrightarrow{\ell}_{p,q} \in \Re^2, \omega_k \in \mathbb{S}_\omega, i = 1,2,\ldots N_r, j = 1,2,\ldots N_\theta \right\} \tag{18}$$

where$G_{np}(\overrightarrow{\ell}_{p,q}, j\omega_k)$ is the NPD measured with the noise source (speaker) at location$\ell_{p,q} = (r_i, \theta_j)$ and the sensor (microphone) located at the given fixed location, e.g., with the noise at location$e(n)$ and the sensor at the origin in Fig. 5. The location of the microphone is chosen the same as that during the targeted imaging process.

Thus, the NPD at any unknown location A,$\widehat{G}_A(j\omega_k)$, can be obtained as the average of the direction-first and the distance-first estimations,$\widehat{G}_{A,\rho}(j\omega_k)$,$\widehat{G}_{A,d}(j\omega_k)$, respectively,

$$\widehat{G}_A(j\omega_k) = \frac{1}{2}\widehat{G}_{A,\rho}(j\omega_k) + \frac{1}{2}\widehat{G}_{A,d}(j\omega_k) \tag{19}$$

where as in the PAD estimation above, the direction-first estimation$\widehat{G}_{A,\rho}(j\omega_k)$ is obtained via

$$\widehat{G}_{A,\rho}(j\omega_k) = (1-\mu_r)\widehat{G}_B\left(\overrightarrow{\ell}_{p,q}, j\omega_k\right) + \mu_r \widehat{G}_C\left(\overrightarrow{\ell}_{p,q}, j\omega_k\right) \tag{20}$$

with$\widehat{G}_B(j\omega_k)$,$\widehat{G}_C(j\omega_k)$ the NPDs estimated at location-B and location-C, respectively, by using the dictionary$\mathbb{D}_{G_{np}}$—similar to Eqs. (16)–(18), and the distance-first estimation$\widehat{G}_{A,d}(j\omega_k)$ is obtained via

$$\widehat{G}_{A,d}(j\omega_k) = (1-\mu_\rho)\widehat{G}_D\left(\overrightarrow{\ell}_{p,q}, j\omega_k\right) + \mu_\rho \widehat{G}_E\left(\overrightarrow{\ell}_{p,q}, j\omega_k\right) \tag{21}$$

with$\widehat{G}_D(j\omega_k)$,$\widehat{G}_E(j\omega_k)$ the estimated NPDs at location-D and location-E, respectively, using the dictionary$\mathbb{D}_{G_{np}}$, similar to Eqs. (21) and (22).

With the above estimated NPD, the noise at location$A$ during the targeted imaging process can be estimated from the measured noise$N_{A,m}(j\omega_k)$ as

$$\widehat{N}_{cl}(j\omega_k) = \frac{N_{A,m}(j\omega_k)}{\widehat{G}_A(j\omega_k)}, \quad \text{for } \omega_k \in \mathbb{S}_\omega \tag{22}$$

3.2. $\mathbb{O}-2$ *data-driven acoustic-dynamics-based filtering*

3.2.1. *Cross-coupling and imaging-dependent effect on PAD*

With the noise-sensor non-collocation issue addressed above, we next design the filter$\widehat{g}_N[\cdot]$ by assuming that the noise-sensor collocation condition holds. The issue is to account for the cross-axis coupling and the imaging process-dependent effect on the filter. Specifically, as the acoustic-noise-caused vibration of the AFM structure, in general, is three dimensional, the lateral vibration of the structure—due to the cross-axis coupling [35]—can also result in the tip disturbance in the vertical direction. This cross-coupling-caused tip disturbance becomes more pronounced and further amplified by the lateral scanning of the tip during the imaging process—-the cross-coupling effect depends on the lateral scanning speed and size. Such an imaging-dependent cross coupling dynamics effect is difficult to capture due to the mixture of the sample topography and the acoustic noise effects in the response signal.

Input : Measured total image signal, $z[k]$, raw noise signals by three microphones, $n_a[k]$, $n_b[k]$ and $n_c[k]$, and cantilever lateral torsional signal $f_{wor}[k]$

Output: Filtered image $z_F[k]$

Construct PAD and NPD dictionaries $\mathbb{D}_{G_n}$ and $\mathbb{D}_{G_{np}}$ in Eq. (12) and Eq. (19) offline a priori.

Localize the noise source location via Eq. (7) and Eq. (8).

Estimate the PAD using the PAD dictionary $\mathbb{D}_{G_n}$ in 1 using the estimated noise location and Eq. (18).

Estimate the noise signal using the NPD dictionary $\mathbb{D}_{G_{np}}$ via Eq. (23).

Use the estimated PAD and the estimated noise signal to obtain the initial filtered image $z_F[k]$ via Eq. (3).

Optimize the filter $\hat{g}_N[k]$ by using the adaptive coherence minimization:

while $\hat{C}_{nf,i}(j\omega_k) > Var[\hat{C}_{nf,i}(j\omega_k)]$ **do**

  Compute and update the coherence $\hat{C}_{nf,i}(j\omega_k)$ using Eq. (26-27), Eqs. (36-37) and Eqs. (39-44)

  Update the filter and the filtered image signal using Eqs. (36, 37).

end

Compute and update the coherence $\hat{C}_{nf,i}(j\omega_k)$

Obtain the filtered image.

Algorithm 1. Data-driven Optimal Filter of Acoustic Noise in AFM Image

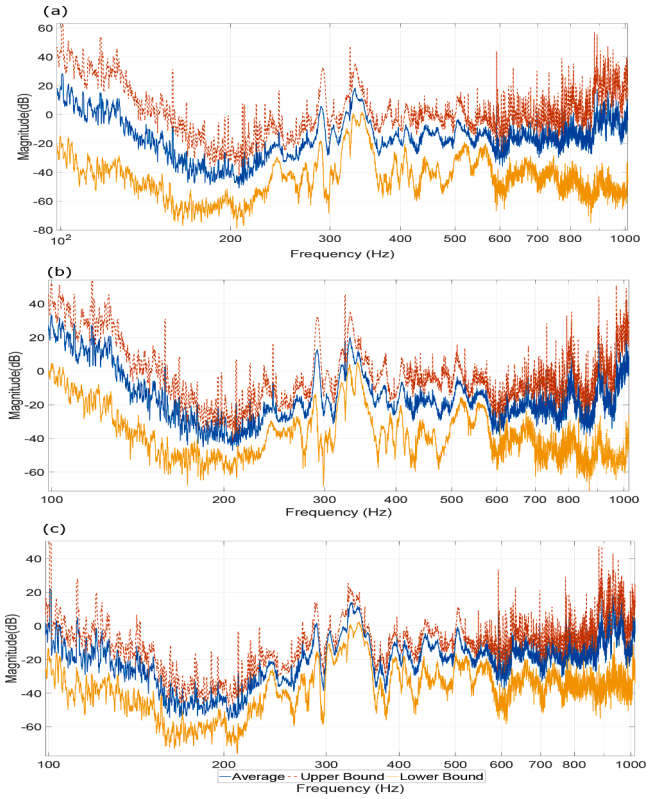

Fig. 7. Comparison between the averaged, the upper bound, and the lower bound of (a) all the PADs in dictionary $\mathbb{D}_{G_n}$. (b) those at the same distance but different angle, and (c) those at the same angle but different distance from the AFM cantilever, respectively.

Although the topography effect on the cross dynamics measurement might be avoided by imaging an atomically-flat sample (at the given imaging condition) a priori, this method is time consuming and not practically feasible. Therefore, we proposed the following data-driven approach to account for the cross-axis coupling effect first, then its imaging-dependence via an adaptive optimization scheme.

3.2.2. *Modulator-based filter optimization via coherence minimization*

Cross-axis coupling compensation First, as a base to account for the cross-axis coupling effect, a coupling-free PAD, $\widehat{G}_{N,nc}(j\omega_k)$, is obtained experimentally a priori by acquiring the image signal under a fixed-location acoustics excitation $Z_{WN}(j\omega_k)$ when the cantilever tip is engaged onto an atomically flat hard reference sample (e.g., a silicon wafer) without scanning, and a band-limited white noise acoustic noise $n_{WN}[k]$ is broadcasted to the environment (see the setup in Fig. 1),

$$\widehat{G}_{N,nc}(j\omega_k) = \frac{Z_{WN}(j\omega_k)}{N_{WN}(j\omega_k)}, \text{for} \quad \omega_k \in \mathbb{S}_\omega \tag{23}$$

Comparing the above coupling-free PAD to those captured during the imaging process under different acoustic noise conditions in experiments showed that the cross-coupling effect mainly appears as a phase distortion in the PAD obtained, i.e., the phase variations of the PAD measured during the imaging process is different from the noise-free one, but closely follows that of the friction force in the lateral direction. Thus, the phase of $\widehat{G}_{N,nc}(j\omega_k), \angle\phi_{N,nc}(j\omega_k)$, is replaced by the phase of the lateral torsional displacement of the cantilever tip measured during the imaging process, $\angle\phi_{N,f}(j\omega_k)$.

Modulator-based Filter Optimization To account for the imaging-dependent effect on the filter $\widehat{G}_N(j\omega_k)$ and further improve the filtering quality, we propose to seek a dynamic modulator $\alpha^*(j\omega_k) \in \mathbb{C}$ as in

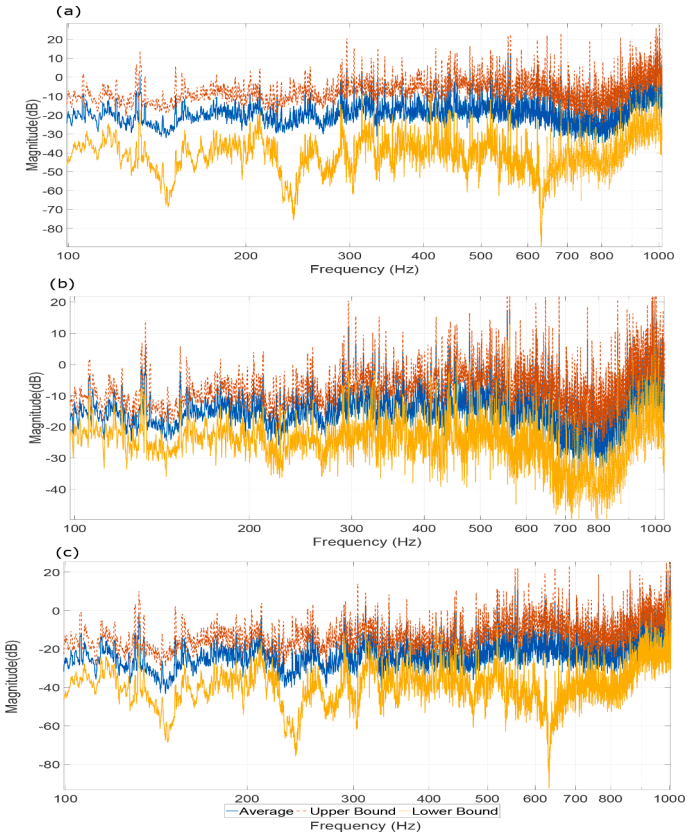

Fig. 8. Comparison between the averaged, the upper bound, and the lower bound of (a) all the NPDs in dictionary $\mathbb{D}_{G_n}$. (b) those at the same distance but different angle, and (c) those at the same angle but different distance from the AFM cantilever, respectively.

$$\widehat{G}_N^*(j\omega_k,\alpha_i)=\alpha^*(j\omega_k)\widehat{G}_N(j\omega_k),\quad \omega_k\in\mathbb{S}_\omega, \tag{24}$$

such that the cost function Eq. (6) is minimized. The optimal $\alpha^*(j\omega_k)$ is sought through the gradient descent search method, i.e., for any given $\omega_k\in\mathbb{S}_\omega$,

$$\widehat{G}_{N,i}(j\omega_k,\alpha_i)=\alpha_i(j\omega_k)\widehat{G}_N(j\omega_k),\quad \text{with} \tag{25}$$

$$\begin{aligned}\alpha_i(j\omega_k)&=\alpha_{i-1}(j\omega_k)+\lambda\frac{\partial\mathscr{J}_r}{\partial\widehat{G}_{N,i-1}(j\omega_k)}\\&=\alpha_{i-1}(j\omega_k)+\lambda\frac{\partial\mathscr{J}_r}{\partial\alpha_{i-1}(j\omega_k)}\frac{1}{\widehat{G}_N(j\omega_k)}\end{aligned} \tag{26}$$

for $i\geq 1$, where initially α_0 is chosen as one, $\lambda\in(0,\ 1)$ is a pre-chosen constant to ensure the descendence of the coherence, and $\mathscr{J}_r$ is the frequency-domain representation of the cost function J_r in Eq. (7), respectively, i.e.,

$$\mathscr{J}_r=\sum_{\omega_k=0}^{N_I}\left[Z_F(j\omega_k)-Z_T(j\omega_k)\right]^2, \tag{27}$$

as by the Parseval's Theorem [36], $\mathscr{J}_r=J_r$.

The gradient in Eq. (28) is, in general, unknown—the true image signal $z_T[k]$ is unknown. Thus, next we convert the minimization of Eq. (29) to that of the coherence between the measured noise and the filtered image signal $Z_F[k]$ (both are known), as the true sample image, $z_T[k]$, is independent to the random noise $n[k]$, and therefore, the correlation between these two signal, $\mathbf{R}_{nt}[j]$, is zero

$$\mathbf{R}_{nt}[j]=\mathbf{E}\{n[k-j]z_T[j]\}=0,\quad j=0,1,2,...,N_I-1. \tag{28}$$

When the acoustic-noise caused distortion in the filtered image signal is completely removed, the filtered image sequence $z_F[k]$ shall also have a zero-correlation to the noise, i.e., when the noise is completely removed,

$$\mathbf{R}_{nf}[j]=\mathbf{E}\{n[k-j]z_F[j]\}\to 0\quad j=0,1,2,...,N_I-1. \tag{29}$$

Thus, minimizing J_r in $\mathbb{O}-2$ is equivalent to minimizing the following cost function J_c of the correlation between $z_F[k]$ and $n[k]$:

$$J_c=\sum_{j=1}^{N_I}\mathbf{R}_{nf}[j]^2=\sum_{j=1}^{N_I}\mathbf{E}\{n[k-j]z_F[j]\}^2 \tag{30}$$

By the Convolution Theorem of correlation [37], frequency domain representation of the above cost function is given by

$$\min\mathscr{J}_c=\sum_{k=1}^{N_F}\mathbf{E}[N^*(j\omega_k)Z_F(j\omega_k)]^2 \tag{31}$$

where N_F denotes the total number of effective frequency components in the noise source $n[k]$ and the filtered image sequence $z_F[k]$, i.e., ω_{N_F} is the lowest frequency at which the magnitude of $\widehat{N}(j\omega_k)$ becomes negligible, e.g., $|N(j\omega_k)|\leq\epsilon_n$ for all $\omega_k\geq N_F$ and ϵ_n is the chosen threshold value.

To facilitate the computation, the above correlation minimization is converted to minimizing the corresponding coherence between these two signals, $\mathbf{C}_{nf}(j\omega_k)$, by normalizing J_c with the (estimated) power spectral density of $Z_F(j\omega_k)$ and $N(j\omega_k)$

$$\min\mathscr{J}_c=\sum_{k=1}^{N_F}\frac{\mathbf{E}[N^*(j\omega_k)Z_F(j\omega_k)]^2}{P_N(j\omega_k)P_{Z_F}(j\omega_k)} \tag{32}$$

$$\triangleq\sum_{k=1}^{N_F}\mathbf{C}_{nf}(j\omega_k) \tag{33}$$

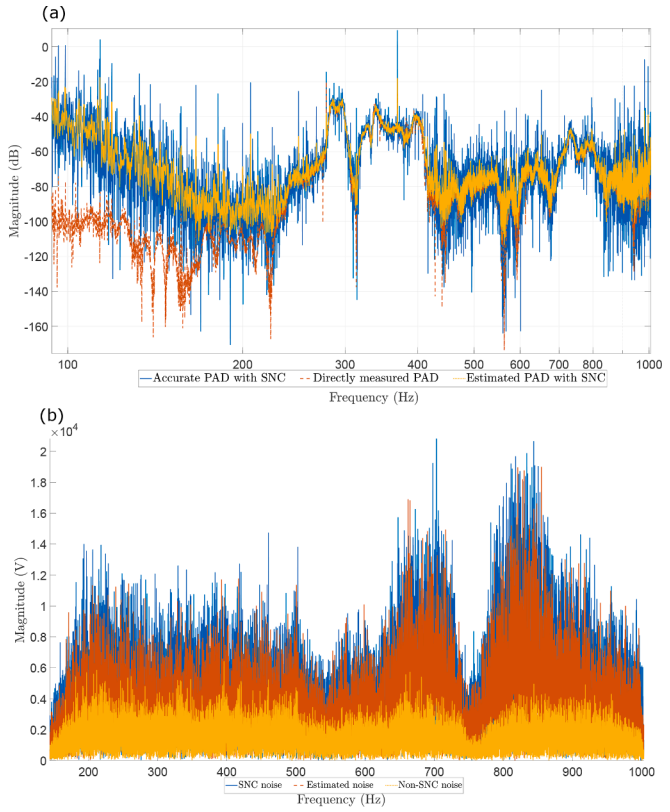

Fig. 9. (a) Comparison of the PAD estimated by using the proposed adaptive modulator optimization method to the accurately measured PAD (under the *SNC* condition) and the directly measured PAD without the *SNC* condition, respectively, and (b) comparison of the true noise measured under the *SNC* condition, noise estimated using the NPD, and the actual noise signal under the non-*SNC* condition, respectively.

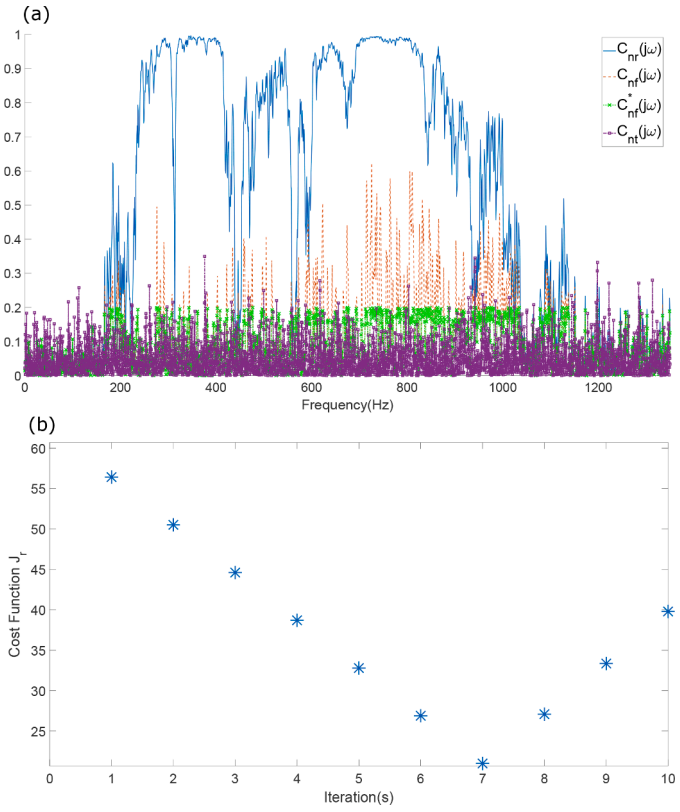

Fig. 10. (a) Comparison of the coherence between the acoustic noise $n[k]$ and the raw image signal $z[k]$, $C_{nr}(j\omega)$, to that between $n[k]$ and the directly filtered $z_F[k]$, $C_{nf}(j\omega)$, and that between $n[k]$ and optimal filtered $z_F^*[k]$, $C_{nf}^*(j\omega)$, and (b) change of the cost function J_r at the circled location in (a) during the adaptive modulator optimization process.

where the$P_N(j\omega_k)$ and$P_{Z_F}(j\omega_k)$ are power spectral densities of$N(j\omega_k)$ and$Z_F(j\omega_k)$, respectively. When the coherence$\mathbf{C}_{nf}(j\omega_k) \in (0,\ 1)$ is minimized, so is$\mathscr{J}_c$. Thus, the gradient-based iteration in Eq. (31) is converted to

$$\alpha_i(j\omega_k) = \alpha_{i-1}(j\omega_k) + \lambda \frac{\partial \mathbf{C}_{nf}(j\omega_k, \alpha_{i-1})}{\partial \widehat{G}_{N,i-1}(j\omega_k)} \tag{34}$$

$$= \alpha_{i-1}(j\omega_k) + \lambda \frac{\partial \mathbf{C}_{nf}(j\omega_k, \alpha_{i-1})}{\partial \alpha_{i-1}(j\omega_k)} \frac{1}{G_N(j\omega_k)} \tag{35}$$

where$\mathbf{C}_{nf}(j\omega_k, \alpha_i)$ is the coherence computed by modulator value obtained in the$i$th iteration,$\alpha_i(j\omega_k)$, i.e., in Eq. (33),

$$Z_F(j\omega_k) = Z_F(j\omega_k, \alpha_i) = Z(j\omega_k) - \widehat{N}(j\omega_k)\widehat{G}_{N,i}(j\omega_k)$$

$$= Z(j\omega_k) - \widehat{N}(j\omega_k)\alpha_i(j\omega_k)\widehat{G}_N(j\omega_k), \tag{36}$$

and the partial derivative$\partial \mathbf{C}_{nf}(j\omega_k)/\partial \alpha_i(j\omega_k)$ is computed numerically via the first-order Euler formula

$$\frac{\partial \mathbf{C}_{nf}(j\omega_k, \alpha_i)}{\partial \alpha(j\omega_k)} \approx \frac{\mathbf{C}_{nf}(j\omega_k, \alpha_i + \delta\alpha) - \mathbf{C}_{nf}(j\omega_k, \alpha_i)}{\delta\alpha} \tag{37}$$

To improve the computation accuracy, the Welch's Method [38] is employed to compute the coherence,$\mathbf{C}_{nf}(j\omega_k, \alpha_i)$, as the average of the short-segment coherence$\widehat{\mathbf{C}}_{nf,r}(j\omega_k, \alpha_i)$, i.e., for any given$\alpha$,

$$\mathbf{C}_{nf}(j\omega_k, \alpha) \approx \widehat{\mathbf{C}}_{nf}(j\omega_k, \alpha) = \frac{\sum_{r=1}^{M_c} \mathbf{C}_{nf,r}(j\omega_k, \alpha)}{M_c} \tag{38}$$

with $\mathbf{C}_{nf,r}(j\omega_k, \alpha) = \dfrac{|N_r^*(j\omega_k) Z_{F,r}(j\omega_k, \alpha)|^2}{P_{N,r}(j\omega_k) P_{F,r}(j\omega_k, \alpha)}$ (39)

for $i \geq 1$, where$N_r(j\omega_k)$ and$Z_{F,r}(j\omega_k, \alpha)$ are the Fourier transform of the$r$th segment of the estimated noise signal$\hat{n}[k]$ and the filtered image signal$z_F^\alpha[k]$, respectively, obtained by partitioning the sequence of$n[k]$ and$z_F^\alpha[k]$ intoM_c segments, respectively,

$$\widehat{\mathbf{n}} = [\mathbf{n}_1\ \mathbf{n}_2\ \ldots\ \mathbf{n}_r\ \ldots\ \mathbf{n}_{M_c}] \quad \text{and} \tag{40}$$

$$\mathbf{z}_F^\alpha = \left[\mathbf{z}_{F,1}^\alpha\ \mathbf{z}_{F,2}^\alpha\ \ldots\ \mathbf{z}_{F,r}^\alpha\ \ldots\ \mathbf{z}_{F,M_c}^\alpha\right], \quad \text{where} \tag{41}$$

$$\mathbf{n}_r = [n((r-1)L+1),\ n((r-1)L+2,\ \ldots, n(rL)] \tag{42}$$

$$\mathbf{z}_{F,r}^\alpha = \left[z_{F,r}^\alpha[(r-1)L+1],\ z_{F,r}^\alpha[(r-1)L+2],\ \ldots, z_{F,r}^\alpha(rL)\right]. \tag{43}$$

where the length of each segment$L$ is chosen such that the sampled frequency resolution of each segment is high enough to capture the frequencies where the peak of coherence occurs, i.e.,$f_s/L \leq \epsilon_f$ forϵ_f is the chosen frequency resolution.

To ensure consistency in the above estimation, the minimization in Eq. (23) is modified as

$$\min \mathscr{J}_c = \sum_{k=1}^{N_F} \left\{ \widehat{\mathbf{C}}_{nf}(j\omega_k) - \mathbf{Var}\left[\widehat{\mathbf{C}}_{nf}(j\omega_k)\right] \right\} \tag{44}$$

where$\mathbf{Var}[\widehat{\mathbf{C}}_{nf}(j\omega_k)]$ is the variance of the estimated$\widehat{\mathbf{C}}_{nf}(j\omega_k)$,

$$\lim_{N_I \to \infty} \mathbf{Var}\left[\widehat{\mathbf{C}}_{nf,i}(j\omega_k)\right] = \mathbf{E}\left\{ \left[\widehat{\mathbf{C}}_{nf,i}(j\omega_k) - \mathbf{C}_{nf,i}(j\omega_k)\right]^2 \right\} \tag{45}$$

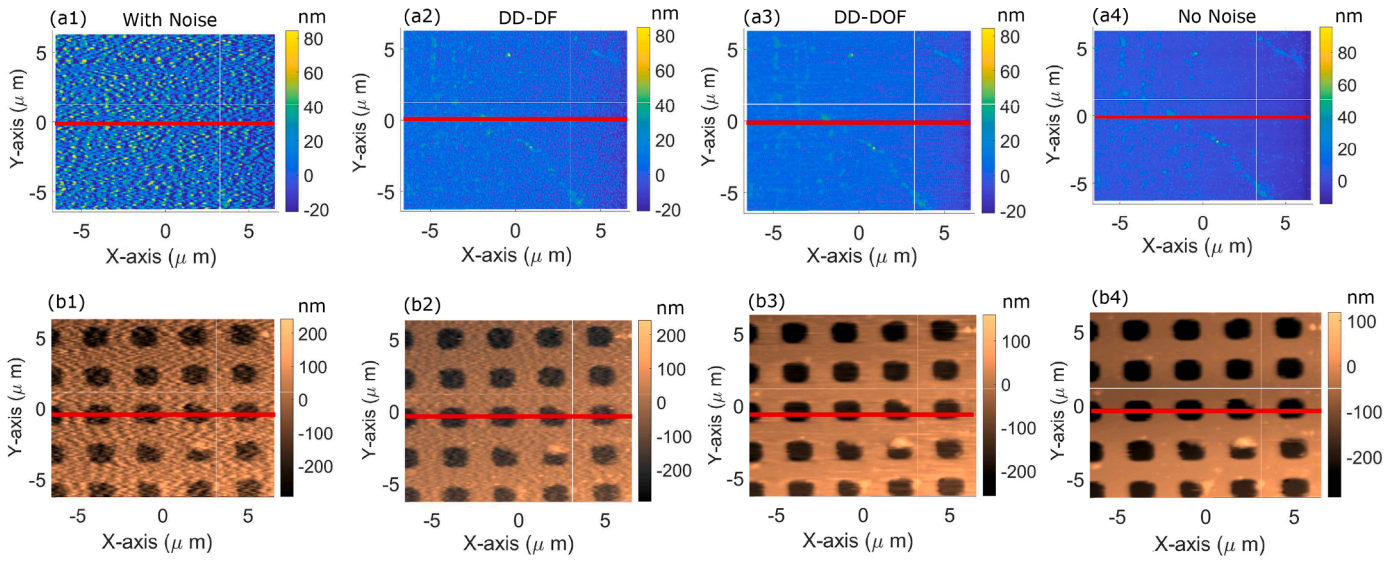

Fig. 11. Comparison of the topography images (first row) of the silicon sample and (second row) of the calibration sample obtained at 5 Hz (a1,b1) with the induced acoustic noise, (a2,b2) with the induced acoustic noise filtered by data-driven dynamics-based filtering (DD-DF), (a3,b3) with the induced acoustic noise and by using the Data-driven dynamics-based optimal filtering (DD-DOF), and (a4,b4) without the induced acoustic noise, respectively.

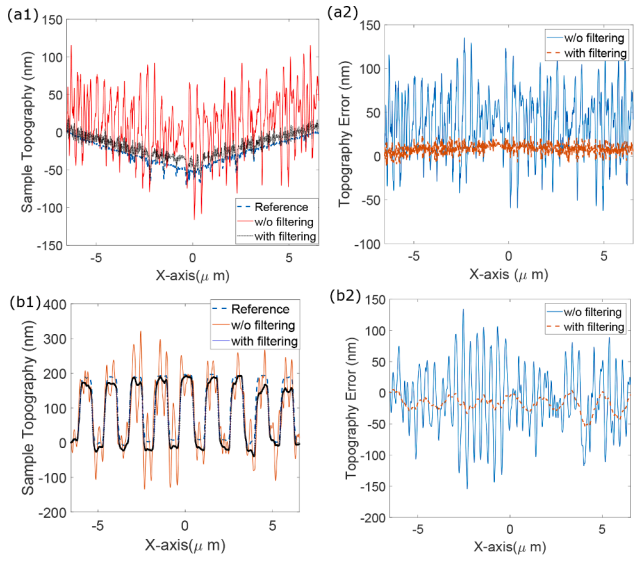

Fig. 12. Comparison of the filtered (using the DD-DOF method) cross-section topography profile (at the red-line in Fig. 11) to the un-filtered one and the reference (without induced acoustic noise) for (a1) the silicon sample, (b1) the calibration sample, and the comparison of the corresponding topography error with respect to the reference for (a1) the silicon sample and (b2) the calibration sample, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

The proposed filtering technique is summarized in Algorithm 1.

4. Experiment example

The proposed approach was demonstrated through an AFM imaging experiment. The objective was to show that by using the proposed approach, the image distortion caused by acoustic noise from an arbitrary location can be substantially reduced.

The AFM imaging experiment was performed on a commercial AFM system (Dimension FastScan, Bruker Nano Inc.), where the acoustic noise was induced by a speaker (Logitech S150, Logitech, Inc.) placed near the AFM scanner head, and measured via a precision array microphone (BK 4958, Bruel Kjaer Inc.), (see Fig. 1 for the experiment setup). The noise signal was first pre-filtered and amplified using a homemade Op-Amp circuit, then measured via a data acquisition system (a Labview-FPGA system NI RIO, USB-7856R, National Instrument Inc.). All the filtering algorithms were designed and implemented in MATLAB.

Implementation of the DD-DOF algorithm

A dictionary of PADs and a dictionary of NPDs, $\mathbb{D}_{G_n}$ and $\mathbb{D}_{G_{np}}$, respectively, were constructed a priori first. A total of 15 PADs, $G_n(\vec{\ell}_{i,j}, j\omega_k)$ were acquired in the dictionary $\mathbb{D}_{G_n}$, where each PAD $G_n(\vec{\ell}_{i,j}, j\omega_k)$ was measured by maintaining a stable probe-surface contact with a silicon sample at the chosen interaction force amplitude under a proportional-integral (PI) feedback control. Then a band-limited (20-1kHz) white noise acoustic sound with zero-mean and constant variance of 40 mV (generated via MATLAB) was broadcasted to the environment through the speaker for 10 s, and both the acoustic noise and the z-axis displacements signals were acquired simultaneously under the noise-sensor collocation condition. The 15 locations were chosen to cover an area in front of the AFM scanner head with the distance from the AFM head at $\vec{\ell}_{i,j} = (\rho_i, \theta_j)$ with $\rho_i = \{40cm, 80cm, 120cm\}$ *and* $\theta_j = \{30^\circ, 60^\circ, 90^\circ, 120^\circ, 150^\circ\}$. Then, the NPD dictionary $\mathbb{D}_{G_{np}}$ was constructed with the noise source placed at the same 15 locations but the microphone was fixed at the origin in Fig. 5 instead.

Next, the acoustic-noise-effected AFM images were acquired. Two different types of samples, a silicon wafer and a calibration sample (STR3-1800P), respectively, were imaged at the scan rate of 5 Hz under the tapping mode, when a band-limited (20-1kHz) white noise acoustic sound of zero-mean and constant variance of 40 mV was broadcasted to the environment through the speaker at an "unknown" position (at $\vec{\ell}_n = (71.4cm, 135^\circ)$ from the AFM cantilever, i.e., the speaker location and the associated noise dynamics knowledge were not used in the following filtering). During the imaging process, the noise was acquired using the three microphones placed at locations at $(0, 0^\circ)$, $(-20cm, 0^\circ)$, and $(20cm, 0^\circ)$, respectively. Both the z-axis piezo displacements and the friction force in the lateral direction were also captured simultaneously. The time delays between the acoustic signals and the three measured noise signals were used to estimate the noise location as described in Sec.III (B) above. Then the estimated noise location was

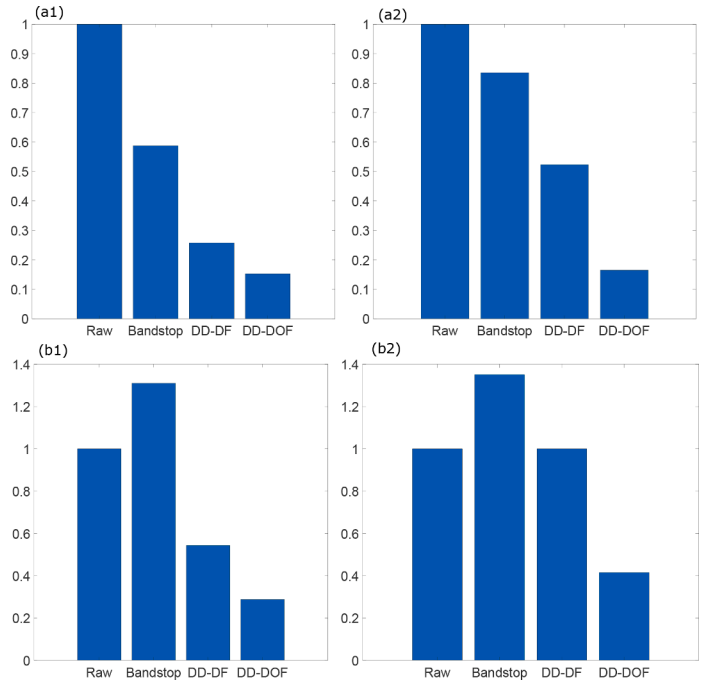

Fig. 13. Comparison of the normalized image error (with respect to the raw image error) for (a) the silicon sample and (b) the calibration sample in (1) 2-norm and (2) ∞-norm, respectively.

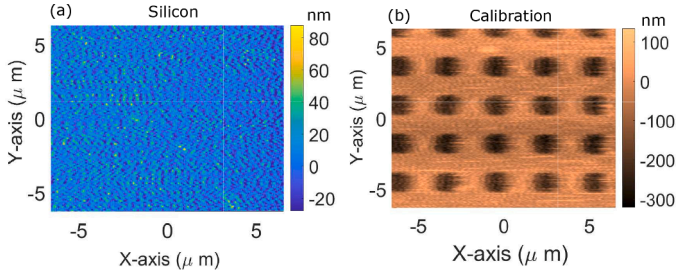

Fig. 14. The topography images of the silicon sample (a) and the calibration sample (b) obtained at 5 Hz filtered by conventional bandstop filtering technique, respectively.

used along with the dictionary$\mathbb{D}_{G_n}$ to estimate the PAD at the noise location. The estimated PAD was first corrected by replacing its phase with the phase of the lateral friction force signal measured, then used as the initial value of the FIR filter (Eq. (26)) in the proposed gradient-based optimization of the dynamic modulator$\alpha^*(j\omega_k)$, by following Algorithm 1, whereα_0 andλ are chosen as 1 and 0.1, respectively. The length of the segments in coherence calculation$L$ was chosen at 41875 sampling points. (Total: 8 segments) Finally, the acoustic-caused z-axis displacement was estimated via Eq. (3), and the filtered sample image was obtained via Eq. (4). For comparison, the images were also filtered by using the directly measured noise and directly measured FIR filter—i.e., both the noise and the PAD-based FIR filter were obtained under the *SNC* condition. Images of the same sample area without the induced acoustic noise were acquired as a reference.

Experimental results and discussion

First, the PADs and the NADs in Dictionary$\mathbb{D}_{G_n}$ and Dictionary$\mathbb{D}_{G_{np}}$ are presented in Figs. 7 and 8, in terms of the averaged, the upper bound and the lower bound between the PADs and the NADs, respectively. The location of the "unknown" noise source was estimated at$(70.9cm, 126\circ)$, compared to the true location at$(71.4, 135\circ)$. The corresponding PAD estimated by using the proposed adaptive modulator optimization technique is compared to accurately measured one (measured under the *SNC* condition) in Fig. 9, along with the directly measured one (with the fixed-location microphone at$(0,\ 0\circ)$). The coherence between the induced acoustic noise signal$n[k]$ and the filtered image signal$z_F[k]$ before and after the modulator optimization process are compared to the coherence between$n[k]$ and the raw image signal$z[k]$ in Fig. 10 (a), and the change of the cost functionJ_r during the iterative optimization process is shown in Fig. 10(b), respectively. The images of the silicon wafer sample and the calibration sample obtained with or without applying the proposed approach are compared in Fig. 11(a1, b1) and (a3, b3), along with the noise-free reference in Fig. 11(a), respectively. The cross section sample profile of these cases are compared in Fig. 12(a1, a2) and the corresponding differences (topography signal under noise with or without filtering vs. the noise-free one) are compared in Fig. (b1, b2) for the two samples, respectively. The filtering quality of the proposed technique was also qualified by the topography difference with respect to the noise-free image error in 2-norm and∞-norm over the entire image, normalized using the values for the unfiltered noise images, as compared in Fig. 13 along with the values for the bandstop filtering and directly filtering results, respectively. Finally, the filtered images obtained by using the proposed approach are compared to the bandstop results in Fig. 14 for the two samples, respectively.

The experimental results clearly demonstrated the location-dependence of the PADs and NADs. As shown in Fig. 7, the PADs varied over 180% in 2-norm and 560% in infinity-norm, respectively, and the 2-norm variation was over 200% and 700% with respect to the distance and the angle/direction changes, respectively. Similar variations were also observed in the NADs too (see Fig. 8). Due to such a large location-dependent variations of the PAD, the directly measured PAD with the unknown-location noise was largely different from the "true" one (measured under the *SNC* condition). As shown in Fig. 9, difference as large as 150% in 2-norm occurred in the lower frequency range from 100 Hz to 220 Hz. Thus, it was evident that the effect of unknown noise location needed to be accounted for in the filtering process.

The effectiveness of the proposed DD-DOF in accurately estimating the PAD of the unknown-location noise can be clearly seen from the

experimental results. First, by using the proposed localization method, the location of the unknown noise was accurately estimated. With the error around 16% and 8% in the distance and the angle estimation, respectively, the residual location-caused variation in the PAD was negligible. As such, the corresponding PADs of the noise source was accurately estimated. Shown in Fig. 9, the large error in the directly measured PAD was substantially reduced, with the 2-norm error of the PAD reduced by over 82% in the frequency range 100 Hz to 220 Hz, and over 33% in the entire measured frequency range. Such an improvement in PAD quantification was directly gained from the proposed gradient-based adaptive modulator optimization. As shown in Fig. 10, the noise-to-image coherence, i.e., the coherence between the acoustic noise$n[k]$ and the raw image signal$z[k]$,C_{nr}, was significantly larger than that between n[k] and the image signal without the induced acoustic noise,C_{nt}, indicating that the noise-to-image coherence was a good indicator of the noise effect on the image result. By using the PAD-based filter, the noise-to-image coherence was significantly reduced by more than 55% in 2-norm for frequenciesω in [20, 1000] Hz, indicating the removal of the acoustic noise from the image signal. However, it was still larger than the coherence without the induced acoustic noise,C_{nt}. Through the proposed gradient-based adaptive optimization process, the noise-to-image coherence was further reduced by 33%, as shown in Fig. 10(a). As a result, the cost function of the acoustic-noise-caused image errorJ_r was consistently reduced by 80% (see Fig. 10(b)). Thus, the experimental results demonstrated the effectiveness of the proposed DD-DOF technique in accurately capturing the PAD and the acoustic noise.

Finally, the AFM image filtering results demonstrated that with the proposed DD-DOF technique, the acoustic-noise-caused image distortions were largely removed. The acoustic-noise-caused image distortion was pronounced in the images obtained (Fig. 11(a1,b1)), expressed as large ripple-like artifacts across the entire image—the acoustic-caused artificial topography fluctuation was over 59% and 66% for the silicon sample and the calibration sample, respectively. On the contrarily, by using the proposed filtering technique, such noise-caused artifacts were substantially reduced by 75% and 46% for the two samples (see Fig. 11 (a2), (b2)), respectively, as also can be seen from the cross-section profile comparison in Fig. 12. Such a significant noise reduction was further enhanced by using the proposed adaptive coherence-based optimization method. As shown in Fig. 11(c1, c2), the DD-DOF images compared well to the noise-free ones, with the noise-caused topography variation further reduced—to a total reduction of 85% and 72% for the two samples, respectively (see Fig. 13). This noise reduction, however, was difficult (if not impossible) to be attained by using the conventional bandstop filtering technique. As shown in Fig. 14, when using the bandstop filter, the acoustic-caused image distortion was, although reduced, still persisted and pronounced, much larger than those using the proposed technique. The acoustic-caused distortion was still at 58% and 131% for the silicon and the calibration sample, respectively, two times and three times larger than that via the proposed technique, respectively. Thus, the experimental results clearly demonstrated that the proposed DD-DOF technique can effectively reduced acoustic-noise-caused distortion in AFM imaging.

5. Conclusion

A data-driven dynamic-based filtering approach was developed to eliminate AFM image distortion caused by acoustic noise from an unknown location. The location of the noise source was estimated based on the time-delay measurement method, and a dictionary of PADs and a dictionary of NADs at selected locations were constructed a priori, and used along with the noise source location identified to estimate the PAD and the noise signal. An acoustic-dynamics modulator was introduced to minimize the acoustic-caused image distortion by minimizing the coherence between the acoustic noise and the residual acoustic-caused image distortion through gradient-based optimization. The efficacy of the proposed approach was demonstrated by filtering experimentally measured AFM images of a silicon sample and a calibration sample. The filtering results showed that by using the proposed technique, the image distortion was substantially reduced.

Declaration of Competing Interest

None.

Acknowledgments

The financial support of NSF grants CMMI-1663055, CMMI-1851907 and IIBR-1952823 is gratefully acknowledged.

References

[1] Y. Yan, Y. Wu, Q. Zou, C. Su, An integrated approach to piezoactuator positioning in high-speed atomic force microscope imaging, Rev. Sci. Instrum. 79 (2008) 073704, https://doi.org/10.1063/1.2956980.
[2] J. Ren, Q. Zou, Adaptive-scanning, near-minimum-deformation atomic force microscope imaging of soft sample in liquid: live mammalian cell example, Ultramicroscopy 186 (2018) 150–157, https://doi.org/10.1016/j.ultramic.2017.12.020.ISSN 0304-3991
[3] T. Uchihashi, H. Watanabe, S. Fukuda, M. Shibata, T. Ando, Functional extension of high-speed AFM for wider biological applications, Ultramicroscopy 160 (2016) 182–196, https://doi.org/10.1016/j.ultramic.2015.10.017.
[4] F. Gołek, P. Mazur, Z. Ryszka, S. Zuber, AFM image artifacts, Appl. Surf. Sci. 304 (2014) 11–19, https://doi.org/10.1016/j.apsusc.2014.01.149.
[5] W. Zhang, G. Meng, Z. Peng, Nonlinear dynamic analysis of atomic force microscopy under bounded noise parametric excitation, IEEE/ASME Trans. Mechatron. 16 (6) (2011) 1063–1072, https://doi.org/10.1109/TMECH.2010.2073715.
[6] S. Ito, D. Neyer, S. Pirker, J. Steininger, G. Schitter, Atomic force microscopy using voice coil actuators for vibration isolation. Proceedings of the IEEE International Conference on Advanced Intelligent Mechatronics (AIM), IEEE, 2015, pp. 470–475.
[7] F. Benmouna, D. Johannsmann, Hydrodynamic interaction of AFM cantilevers with solid walls: an investigation based on AFM noise analysis, Eur. Phys. J. E 9 (2002) 435–441, https://doi.org/10.1140/epje/i2002-10096-x.
[8] D. Tranchida, S. Piccarolo, R.A.C. Deblieck, Some experimental issues of AFM tip blind estimation: the effect of noise and resolution, Meas. Sci. Technol. 17(10) 2630–2636.
[9] S. Yi, T. Li, Q. Zou, Active control of acoustics-caused nano-vibration in atomic force microscope imaging, Ultramicroscopy 195 (2018) 101–110, https://doi.org/10.1016/j.ultramic.2018.07.006.
[10] Z. Xu, D. Tramp, Q. Zou, et al., Nanoscale broadband viscoelastic spectroscopy of soft materials using iterative control, Exp. Mech. 52 (2012) 757–769, https://doi.org/10.1007/s11340-011-9547-3.
[11] J. Ren, Q. Zou, A control-based approach to accurate nanoindentation quantification in broadband nanomechanical measurement using scanning probe microscope, IEEE Trans. Nanotechnol. 13 (1) (2014) 46–54, https://doi.org/10.1109/TNANO.2013.2287505.
[12] J. Ren, A. Mousavi, X. Li, Q. Zou, N. Erina, C. Su, Enhanced measurement of broadband nanomechanical property of polymers using atomic force microscope, Appl. Phys. Lett. 102 (2013) 183116, https://doi.org/10.1063/1.4804381.
[13] H. Xie, S. Régnier, High-efficiency automated nanomanipulation with parallel imaging/manipulation force microscopy, IEEE Trans. Nanotechnol. 11 (1) (2012) 21–33, https://doi.org/10.1109/TNANO.2010.2041359.
[14] Y. Yan, H. Wang, Q. Zou, A decoupled inversion-based iterative control approach to multi-axis precision positioning: 3D nanopositioning example, Automatica 48 (1) (2012) 167–176, https://doi.org/10.1016/j.automatica.2011.09.032.
[15] Z. Wang, J. Tan, Q. Zou, W. Jiang, Mechanical-plowing-based high-speed patterning on hard material via advanced-control and ultrasonic probe vibration, Rev. Sci. Instrum. 84 (2013) 113704, https://doi.org/10.1063/1.4832046.
[16] G. Schitter, K.J. Astrom, B.E. DeMartini, P.J. Thurner, K.L. Turner, P.K. Hansma, Design and modeling of a high-speed AFM-scanner, IEEE Trans. Control Syst. Technol. 15 (2007) 906–915, https://doi.org/10.1109/TCST.2007.902953.
[17] T. Ando, T. Uchihashi, N. Kodera, et al., High-speed AFM and nano-visualization of biomolecular processes, Pflug. Arch. Eur. J. Phys. 456 (2008) 211–225, https://doi.org/10.1007/s00424-007-0406-0.
[18] J. Ren, Q. Zou, High-speed adaptive contact-mode atomic force microscopy imaging with near-minimum-force, Rev. Sci. Instrum. 85 (2014) 073706, https://doi.org/10.1063/1.4884343.
[19] G.M. Clayton, S. Tien, K.K. Leang, Q. Zou, S. Devasia, A review of feedforward control approaches in nanopositioning for high-speed SPM, ASME J. Dyn. Syst. Meas. Control 131 (6) (2009) 061101, https://doi.org/10.1115/1.4000158.
[20] S. Ito, G. Schitter, Atomic force microscopy capable of vibration isolation with low-stiffness z-axis actuation, Ultramicroscopy 186 (2018) 9–17, https://doi.org/10.1016/j.ultramic.2017.12.007.
[21] M. Vladimir, S. Vladimir, Optical tables vibration isolation during precision measurements, Procedia Eng. 111 (2015) 561–568, https://doi.org/10.1016/j.proeng.2015.07.043.

[22] C.H. Hansen, Solutions to example problems in engineering noise control, 2003.
[23] Y. Suzuki, N. Sakai, A. Yoshida, et al., High-speed atomic force microscopy combined with inverted optical microscopy for studying cellular events, Sci. Rep. 3 (2013) 2131, https://doi.org/10.1038/srep02131.
[24] S.M. Kuo, D.R. Morgan, Active noise control: a tutorial review, Proc. IEEE 87 (6) (1999) 943–971, https://doi.org/10.1109/5.763310.
[25] D.P. Das, G. Panda, S.M. Kuo, New block filtered-x LMS algorithms for active noise control systems, IET Signal Process. 1 (2) (2007) 73–81, https://doi.org/10.1049/iet-spr:20060220.
[26] C.Y. Chang, K.K. Shyu, Active noise cancellation with a fuzzy adaptive filtered-x algorithm, IEE Proc. Circuits Devices Syst. 150 (5) (2003) 416–422, https://doi.org/10.1049/ip-cds:20030406.
[27] M.G. Ruppert, K.S. Karvinen, S.L. Wiggins, S.O.R. Moheimani, A Kalman filter for amplitude estimation in high-speed dynamic mode atomic force microscopy, IEEE Trans. Control Syst. Technol. 24 (1) (2016) 276–284, https://doi.org/10.1109/TCST.2015.2435654.
[28] A.K. Tellakula, Acoustic source localization using time delay estimation, 2007.
[29] N. Yousefian, P.C. Loizou, A dual-microphone speech enhancement algorithm based on the coherence function, IEEE Trans. Audio Speech Lang. Process. 20 (2) (2012) 599–609, https://doi.org/10.1109/TASL.2011.2162406.
[30] H.W. Löllmann, A. Brendel, W. Kellermann, Generalized coherence-based signal enhancement. Proceedings of the ICASSP 2020 - 2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), IEEE, 2020, pp. 201–205.
[31] N. Gadegaard, Atomic force microscopy in biology: technology and techniques, Biotech. Histochem. 81 (2–3) (2006) 87–97, https://doi.org/10.1080/10520290600783143.
[32] M. Gruber, Statistical digital signal processing and modeling, Technometrics 39 (3) (1997) 335–336.
[33] M.S. Rana, H.R. Pota, I.R. Petersen, High-speed AFM image scanning using observer-based MPC-notch control, IEEE Trans. Nanotechnol. 12 (2) (2013) 246–254, https://doi.org/10.1109/TNANO.2013.2243752.
[34] A.K. Tellakula, Acoustic source localization using time delay estimation, 2007.
[35] Y. Wu, J. Shi, C. Su, Q. Zou, A control approach to cross-coupling compensation of piezotube scanners in tapping-mode atomic force microscope imaging, Rev. Sci. Instrum. 80 (2009) 043709, https://doi.org/10.1063/1.3124183.
[36] L.J.W.S. Rayleigh, On the character of the complete radiation at a given temperature, Philos. Mag. 27 (1889) 460–469, https://doi.org/10.1080/14786448908628386.
[37] C. McGillem, G. Cooper, Continuous and Discrete Signal and System Analysis (2 ed.), Holt, Rinehart and Winston, 1984, p. 118.(3-102)
[38] P. Welch, The use of fast fourier transform for the estimation of power spectra: a method based on time averaging over short, modified periodograms, IEEE Trans. Audio Electroacoust. 15 (2) (1967) 70–73, https://doi.org/10.1109/TAU.1967.1161901.