Data-Driven Robust Acoustic Noise Filtering for Atomic Force Microscope Image

Jiarong Chen and Qingze Zou

***Abstract*—This article proposes a data-driven acoustic vibration filtering technique to eliminate acoustic-caused distortions in atomic force microscope (AFM) images. AFM measurement is sensitive to external disturbances including acoustic noises, as disturbance 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. The acoustic dynamics involved is complicated, broadband, and not decaying with frequency increase. More challenge arises in practice as the location of the acoustic noise source tends to be unknown and arbitrary, resulting in low signal to noise ratio (SNR) in the acoustic signal measurement, and large error in the acoustic dynamics quantified. In this work, we propose a Wiener-filter-based robust filtering technique to improve both the SNR of the acoustic signal measured and reduce the error in the acoustic dynamics obtained. Then, a coherence minimization approach is proposed to further enhance the accuracy of the filter without modeling. Experimental implementation is presented and discussed to illustrate the proposed technique.**

***Index Terms*—Acoustic noise filtering, atomic force microscope (AFM) imaging, coherence minimization, data-driven, Wiener filter.**

I. Introduction

IN THIS article, a data-driven dynamics-based postfiltering technique is proposed to eliminate acoustic vibration-caused distortions in atomic force microscope (AFM) images. As one of the most important tools/instruments for nanoscience and nanotechnology [1], AFM depends on the measurement and manipulation of the interaction between a nano-size cantilever probe and the sample—-based on the mechanical "touch to see" principle [2]. Thus, external disturbances including acoustic vibration can disturb the probe–sample interaction and distort the AFM measurement results [3]. Although passive noise apparatuses have been employed to combat acoustic noise, such a method—the *de facto* industry standard—faces limitations in performance, usability, and cost. Whereas active acoustic cancellation, however, is challenging, particularly when the acoustic noise is from an unknown and arbitrary location. Thus, this work is motivated to tackle these challenges by developing a robust filtering technique.

The effect of acoustic noise on AFM operation arises as maintaining the probe–sample interaction closely around the set point value is crucial in all AFM applications, ranging from imaging [4], nanomechanical measurement [5], to probe-based nanofabrication [6]. Extraneous probe vibration can be induced both externally by disturbances including acoustic noise and seismic vibrations and internally due to the excitement of the dynamics and hysteresis adverse effects of the nanopositioning system (from the piezo actuator to the cantilever probe) [7]. Although the internal adverse effects have been compensated for through both hardware improvements (by increasing the bandwidth of the piezo actuator and/or cantilever [8], [9]) and software (algorithmic) enhancement (by developing more advanced control techniques to better account for the dynamics and hysteresis effects [7], [10], [11]), compensation for the external adverse effects has been largely limited to hardware means through passive vibration/noise isolation apparatus [12], [13], [14], [15]. Although vibration and acoustic noise effects can be mitigated, these passive apparatus are costly, bulky, and not compatible in applications such as biomedical-related research, where AFM needs to be integrated with other instruments such as optical microscope [16]. Moreover, residual image distortion still persists, and the image quality obtained cannot meet stringent requirements in applications such as cleanroom nanometrology in semi-conductor industry [1]. These acoustic-related issues have limited the application and impact of AFM in these and other science and engineering fields. However, unlike the development of algorithms—control techniques—to account for the internal disturbances in high-speed AFM operations [7], [17], few work has been reported on active control of acoustic noise for AFM. Recently, an inversion-based feedforward controller has been proposed for online active noise control during AFM imaging [18]. Its performance, however, can be hindered by the hardware constraints (e.g., online computation power and data acquisition speed), the bandwidth limitation, and the unknown and arbitrary location of the noise source. Thus, advanced algorithms need to be developed to combat acoustic noise effect on AFM image.

Received 6 August 2024; revised 2 February 2025; accepted 26 February 2025. Recommended by Technical Editor S. Zhuang and Senior Editor K. Oldham. This work was supported by NSF under Grant CMMI-1851907, Grant IIBR-1952823, Grant PFI-2234449, and Grant NSF-PHY-2412551. *(Corresponding author: Qingze Zou.)*

The authors are with the Department of Mechanical and Aerospace Engineering, Rutgers, The State University of New Jersey, Piscataway, NJ 08854 USA (e-mail: jiarong.chen@rutgers.edu; qzzou@soe.rutgers.edu).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TMECH.2025.3550320.

Digital Object Identifier 10.1109/TMECH.2025.3550320

Challenges exist in eliminating acoustic-caused distortions in AFM images. As the noise-caused distortion is coupled with the sample topography in the images obtained, conventional filtering techniques based on frequency-separation (e.g., low-pass, band-pass, or notch filter) are ineffective. The acoustic-noise dynamics, i.e., the dynamics from the acoustic noise to the vibration response of the cantilever probe, is highly oscillatory (i.e., contains multiple poles and zeros) and broadband, and does not decay as frequency increases, making it difficult to employ model-based approaches along the framework of Kalman-filtering [19]. This difficulty becomes even more so when the location of the acoustic source is arbitrary and unknown—as usually the case in practice. The unknown and arbitrary acoustic source location can result in low signal to noise ratio (SNR) in the acoustic signal measured, particularly when the sensor (microphone) is distant away from the acoustic source, it is practically unfeasible to relocate the sensor with respect to the acoustic source. Moreover, the unknown acoustic source also results in errors in the acoustic dynamics obtained as the acoustic dynamics depends on the path of the noise propagation and varies significantly as the distance and direction of the noise source changes. Thus, for the filter to be effective, the complexity of the acoustic dynamics and the unknown acoustic source location effect must be addressed.

These challenges have been tackled recently through the development of data-driven finite-impulse-response (FIR)-based filtering techniques [20], [21]. The basic idea is to build the filter based on the acoustic dynamics involved. To robustly capture the complicated acoustic dynamics, the FIR-representation instead of a parameterized-model of the acoustic dynamics is utilized [21]. By combining the FIR-based filter with an additional nano-vibration sensor [21], the acoustic-caused AFM image distortion can be reduced. The method, however, requires special hardware modification (to install the specially designed sensor), and does not account for the adverse effect of the acoustic source location being arbitrary and unknown. This issue might be mitigated by measuring the acoustic dynamics responses for acoustic source at different locations and using the responses to construct a dictionary (of acoustic dynamics) to cover the entire operation area of interests [20]. Then, the acoustic dynamics for a noise source at an unknown location can be estimated via interpolation [20]. The filter is further enhanced through a frequency-dependent modulator based on coherence minimization. The efficacy of the technique has been demonstrated and validated through AFM imaging experiments [20]. The performance of this method [20], however, depends on the density and accuracy of the acoustic-dynamics (dictionary), and constructing the dictionary can be time consuming and prone to measurement uncertainties and acoustic dynamics variation—not favored in practice, particularly when the acoustic source location is unknown and arbitrary, and the SNR of the acoustic noise is low. Therefore, efforts are needed to further enhance the efficiency and robustness of the data-driven, dynamics-based filtering of AFM images.

The main contribution of this article is the development of a data-driven robust filtering technique to eliminate acoustic-vibration caused distortions in AFM images. It is shown that by only using the data measured during the imaging process itself–without any a priori acoustic dynamics characterization, an optimal filter can be constructed to minimize the the error in estimating the acoustic signal, and an optimal filter can be designed to minimize the image filtering error, respectively. We show that these two optimal filters are readily given by the Wiener filter [22] of the acoustic noise involved. Moreover, the proposed filters are further optimized through a coherence minimization based frequency modulator technique [20], such that the uncertainties in the Wiener filter parameters caused by limited measurements and other disturbances in practice are accounted for. Compared to the previous work [20], the proposed approach eliminates the need for constructing an acoustic dynamics dictionary while preserving the performance regardless the location of acoustic source. The proposed filtering scheme is implemented in an AFM imaging experiment, and the experimental results show that the image distortion can be significantly reduced [23].

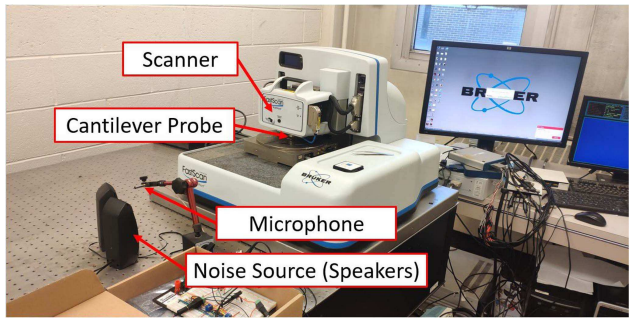


Fig. 1. Experimental setup for studying acoustic noise effect on AFM operation, where it is assumed that the location of the noise source (speakers) is unknown while the sensor (microphone) is placed at fixed and known location.

II. Data-Driven Acoustic Filtering of AFM Image

A. Acoustic-Caused AFM Image Distortions: Problem Formulation

It is important in AFM applications to eliminate external disturbances such as acoustic noise. The basic principle of AFM measurement is to manipulate the interaction between a nanometer-size cantilever-tip and the sample surface and regulate the probe–sample interaction with nanoscale force and displacement precision [4]. External disturbances like acoustic noise can induce extraneous perturbation to the probe–sample interaction, and thereby, loss of precision and quality in the AFM measurements (see Fig. 1). More specifically, during the tapping-mode (TM) imaging of AFM [2], the cantilever tip is excited around the resonant frequency of the cantilever and taps on the sample surface while scanning across the sampling surface. During the scanning process, the tapping amplitude is maintained around a prechosen constant level via feedback control [24] (see Fig. 2). When the tip–sample interaction (i.e., the tip tapping) is well maintained, the sample topography image can be obtained from the (vertical) displacement of the cantilever. However, the mechanical structure of the AFM can

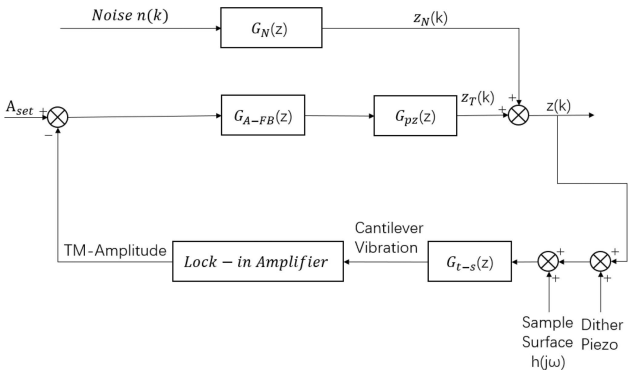


Fig. 2. Schematic block diagram of the AFM imaging process with acoustic noise.

be excited by the environmental disturbance like acoustic noise, resulting in unwanted cantilever vibration and thereby, image distortion (see Fig. 2).

In this work, we aim to develop a postimaging filtering technique to eliminate the acoustic noise effect without localizing the acoustic source. Without loss of generality, we assume that

Assumption 1: The noise source is at a fixed but arbitrary and unknown location.

Assumption 2: The acoustic noise n[k] is a zero-mean, band-limited wide-sense stationary (WSS) random process [22], and the variation of the primary acoustic noise dynamics (PAD) is quasi static.

The PAD in Assumption 2 is the dynamics from the noise signal (as the input) to the AFM image signal (as the output response) [$G_N(z)$ in Fig. 2]. Assumption 2 is reasonable as the variation of the PAD is mainly caused by the change of the noise source location, i.e., the noise propagation route and the AFM configuration (e.g., mounting of the cantilever), both remain unchanged during an imaging process but otherwise can vary significantly in day-to-day operations.

The unknown acoustic source location (Assumption 1) results in source-sensor noncollocation, which, in turn, imposes two challenges to postimaging filtering of the acoustic noise: One, large uncertainty/variation in both the PAD and the acoustic noise measured (as both the PAD and the acoustic noise measurement depend on the media path of the acoustic propagation), and second, low SNR when the sensor is distant away from the acoustic source. In contract to our previous work that tackled this noncollocation problem through a dictionary-based approach [20], we propose to address these two challenges without constructing the dictionary, and without even acquiring additional acoustic signals a priori—We aim to only use the signals measured during the imaging process itself.

Considering that there exists measurement disturbances in the measured acoustic signal, e.g., other environmental and electrical thermal noise, the total measured acoustic signal $n_m[k]$ can be represented as

$$n_m[k] = n^*[k] + n_d[k] \tag{1}$$

where, respectively, $n^*[k]$ is the "true" acoustic noise responsible for the image distortion, and $n_d[k]$ is the measurement disturbance of the acoustic signal.

Thus, in the presence of acoustic noise, the measured AFM image signal, $z_m[k]$, becomes

$$z_m[k] = z_s[k] + z_n[k] + z_d[k], \text{ for } k = 0, \ldots, N_I - 1, \tag{2}$$

where $z_s[k]$ and $z_n[k]$ are the z-axis piezo displacement corresponding to the sample topography and that due to the acoustic noise, $z_d[k]$ is the measurement disturbance of the image signal, and N_I is the total number of sampling data acquired in the given imaging process, respectively. As $z_m[k]$ is used to plot the sample topography image, in the following, $z_m[k]$, $z_s[k]$, and $z_n[k]$ are called the measured image signal, the true sample image signal, and the image disturbance signal, respectively.

Assumption 3: The acoustic measurement disturbance $n_d[k]$ is a zero-mean WSS random process, mutually uncorrelated to the acoustic noise $n^*[k]$, the true image signal $z_s[k]$, and the measurement disturbance $z_d[k]$, respectively, i.e.,

$$\mathbf{E}\{n_d[k]z_s[k-j]\} = 0 \tag{3}$$

$$\mathbf{E}\{n^*[k]z_s[k-j]\} = 0, \text{and} \tag{4}$$

$$\mathbf{E}\{n^*[k]z_d[k-j]\} = 0 \tag{5}$$

for any given j ($j = 0, 1, 2, \ldots, N_i - 1$). Thus, the filtered image signal, $z_f[\cdot]$, can be obtained by estimating the image noise signal, $\hat{z}_n[k]$, and removing it accordingly, i.e.,

$$z_f[k] = z_m[k] - \hat{z}_n[k], \text{ for } k = 0, \ldots, N_I - 1 \tag{6}$$

and by Assumption 3, the filtering quality can be quantified by the residual image error $e_r[k]$

$$e_r[k] = z_n[k] - \hat{z}_n[k] = z_f[k] - z_s[k] - z_d[k]. \tag{7}$$

Next, we formally state the acoustic-vibration filtering problem.

Data-driven Robust Filtering (DDRF) of Acoustic-caused AFM Image Distortion: Let Assumptions 1–3 hold, the DDRF problem is to design an optimal acoustic-noise filter and an optimal acoustic-image filter directly from the data measured during the imaging process, i.e., without an a priori acoustic characterization process nor finding a parameterized model, such that

$\mathbb{O}$-1 By using the optimal acoustic-noise filter, the estimation of the acoustic noise is unbiased

$$\mathbf{E}(\hat{n}[k]) = \mathbf{E}(n^*[k]) = 0 \tag{8}$$

and the variance of the estimation is minimized, i.e., for any given k ($k = 1, 2, \ldots, N_i$)

$$\min_{\hat{n}[k]} J_n = \mathbf{E}\{e_n[k]\}^2 = \mathbf{E}\{n^*[k] - \hat{n}[k]\}^2. \tag{9}$$

$\mathbb{O}$-2 By using the optimal acoustic-image filter, the filtering of the image signal is unbiased

$$\mathbf{E}(e[k]) = \mathbf{E}\{z_s[k] - z_f[k]\} = 0. \tag{10}$$

and the variance of the filtering is minimized, i.e., for any given k ($k = 1, 2, \ldots, N_i$)

$$\min\ J_z = \mathbf{E}\{e_r[k]\}^2. \tag{11}$$

We proceed by achieving these two objectives in order.

B. [𝕆-1] Optimal Acoustic-Noise Filter Design: A Wiener Filter Approach

We propose to optimize the estimation of the acoustic noise signal $\hat{n}[\cdot]$ in (9) based on the Wiener filtering theory [22]. The key idea is to exploit the high sensitivity of the AFM measurement: the acoustic-caused disturbances in the AFM image signals (i.e., the z-axis displacement and the cantilever deflection signal as the responses to the acoustic signal)—can be utilized to enhance the SNR of the acoustic noise measurement, provided that the corresponding PAD can be accurately captured. This is based on that the "true" noise signal $n^*[k]$ is related to the image noise signal, $z_n[k]$, by

$$n^*[k] = (\mathbf{g_t} * \mathbf{z_n})[k] \tag{12}$$

where "*" denotes the discrete convolution operation, $\mathbf{g_t}$ is the impulse response of the inverse dynamics of the "exact" PAD from the acoustic signal $n^*[k]$ to the image noise signal $\mathbf{z_n} \in \Re^{N_I}$ (for j= $0, 1, \ldots, N_I - 1$), respectively,

$$\mathbf{g_t} = [g_z[0]\, g_z[1] \cdots g_z[N_I - 1]\}]^T \tag{13}$$

$$\mathbf{z_n} = [z_n[0]\, z_n[1] \cdots z_n[N_I - 1]]^T \tag{14}$$

However, as the image noise signal $z_n[k]$ is buried in the measured total image signal $z_m[k]$, thereby, unknown in general, we propose to estimate the acoustic noise signal $\hat{n}[k]$ by using the measured z-axis displacement, i.e., the measured image signal $z_m[k]$, via

$$\hat{n}[k] = [\hat{\mathbf{g}_z} * (\mathbf{z_m} - \bar{\mathbf{z_m}})][k] \triangleq [\hat{\mathbf{g}_z} * \check{\mathbf{z}}_m][k] \tag{15}$$

where $\hat{\mathbf{g}_z} \in \Re^{N_F}$ is the acoustic filter designed based on an estimation of the impulse response of the inverse dynamics $\mathbf{g_z}$, from the acoustic signal $n^*[k]$ to the measured total image signal $z_m[k]$, $\mathbf{z_m} \in \Re^{N_F}$ (for k= $0, 1, \ldots, N_I - 1$) is the vector form of the measured image signal, $\bar{\mathbf{z}_m}$ is the corresponding expectation value and $\check{\mathbf{z}}_m][k]$ is called the unbiased measured image signal, respectively,

$$\hat{\mathbf{g}_z} = [\hat{g}_z[0]\, \hat{g}_z[1] \cdots \hat{g}_z[N_I - 1]\}]^T \tag{16}$$

$$\mathbf{z_m} = [z_m[0]\, z_m[1] \cdots z_m[N_I - 1]]^T \tag{17}$$

$$\begin{aligned} \bar{\mathbf{z}_m} &= \mathbf{E}(\mathbf{z_m}) = \mathbf{E}(\mathbf{z_n} + \mathbf{z_s} + \mathbf{z_d}) \quad \text{(by (2))} \\ &= [z_s[0]\, z_s[1] \cdots z_s[N_I - 1]]^T = \mathbf{z_s} \end{aligned} \tag{18}$$

where the last equation follows by the expectation of $\mathbf{z_n}$ and $\mathbf{z_n}$ are zero by Assumptions 2 and 3. We seek the optimal filter $\hat{\mathbf{g}_z^*}$ to minimize the power of the correlation between the estimated acoustic measurement disturbance $\hat{n}_d[k]$

$$\hat{n}_d[k] = n_m[k] - \hat{n}[k] \tag{19}$$

and the total measured image signal $z_m[k]$

$$\min_{\hat{\mathbf{g}_z}} \hat{J}_n[j] \triangleq r_{nz}^2[j] = [\mathbf{E}\{\hat{n}_d[k] z_m[k-j]\}]^2 . \tag{20}$$

We show below that such a filter also minimizes the cost function J_n in (9) as well. As all the signals in the above cost function $\hat{J}_n$ are available, this leads to a computational scheme to obtain the optimal acoustic noise filter $\hat{\mathbf{g}_z}$.

Lemma 1: Let Assumptions 1–3 be hold, and the optimal noise filter $\hat{g}_z[k]$ be given by (15) to (20), then

1) The estimation of the acoustic noise is unbiased, i.e., (8) holds, and
2) The variance of the noise estimation error J_n in (9) is minimized if and only if the square of the cross-correlation $r_{nz}[j]$ in (20) reaches its minimal value at zero, i.e., if and only if $\hat{J}_n[j] = 0$ for any given $j = 1, 2, \ldots, N_I - 1$.

Proof: The unbiased estimation (21) comes directly from Assumption 2, as by (15)–(16)

$$\begin{aligned} \mathbf{E}\{\hat{n}[k]\} &= \mathbf{E}\{[\hat{\mathbf{g}_z} * \check{\mathbf{z}}]\,[k]\} \\ &= [\hat{\mathbf{g}_z} * \{\mathbf{E}(\check{\mathbf{z}})\}][k] \\ &= [\hat{\mathbf{g}_z} * [\mathbf{E}(\mathbf{z_n}) + \mathbf{E}(\mathbf{z_d})][k] \\ &= [\hat{\mathbf{g}_z} * (\mathbf{0} + \mathbf{0})][k] = 0 \end{aligned} \tag{21}$$

Next, we proceed by substituting (1), (2), (12), (15) into (9) to rewrite the cost function J_n as

$$\begin{aligned} J_n &= \mathbf{E}\{n^*[k] - \hat{n}[k]\}^2 \\ &= \mathbf{E}\{(\mathbf{g_t} * \mathbf{z_n})[k] - [\hat{\mathbf{g}_z} * (\check{\mathbf{z}})][k]\}^2 \\ &= \mathbf{E}\{(\mathbf{g_t} * \mathbf{z_n})[k] - [\hat{\mathbf{g}_z} * (\mathbf{z_n} + \mathbf{z_d})][k]\}^2 \\ &= \mathbf{E}\{(\mathbf{g_t} - \hat{\mathbf{g}_z}) * \mathbf{z_n}]^2[k] - 2[(\mathbf{g_t} - \hat{\mathbf{g}_z}) * \mathbf{z_n}[k]] \\ &\quad [\hat{\mathbf{g}_z} * \mathbf{z_d}][k] + [\hat{\mathbf{g}_z} * \mathbf{z_d}]^2[k]\}. \end{aligned} \tag{22}$$

As the image noise signal $z_n[k]$ is driven by the acoustic noise $n^*[k]$ whereas the image measurement disturbance $z_d[k]$ is not, by Assumption 3

$$\mathbf{E}(z_d[k] z_n[k-j]) = 0, \quad \text{for } j = 1, 2, \ldots, N_I - 1 \tag{23}$$

Hence in (22)

$$\begin{aligned} &\mathbf{E}\{[(\mathbf{g_t} - \hat{\mathbf{g}_z}) * \mathbf{z_n}][k][\hat{\mathbf{g}_z} * \mathbf{z_d}][k]\} \\ &= \mathbf{E}\left\{\left[\sum_{i=0}^{N_i} \{g_t[i] - \hat{g}_z[i]\}\, z_n[k-j]\right]\left[\sum_{i=0}^{N_i} \hat{g}_z[i] z_d[k-i]\right]\right\} \\ &= 0 \end{aligned} \tag{24}$$

Substituting (24) into (22) yields

$$J_n = \mathbf{E}\{(\mathbf{g_t} - \hat{\mathbf{g}_z}) * \mathbf{z_n}]^2[k]\} + \mathbf{E}\{[\hat{\mathbf{g}_z} * \mathbf{z_d}]^2[k]\} \tag{25}$$

And thus

$$\begin{aligned} \frac{\partial J_n}{\partial \hat{g}_z[j]} &= 2\mathbf{E}\left\{\sum_{i=0}^{N_I}\{\{\hat{g}_z[i] - g_t[i]\} z_n[k-i]\} z_n[k-j]\right\} \\ &\quad + 2\mathbf{E}\left\{\sum_{i=0}^{N_I}\{\hat{g}_z[i] z_d[k-i]\} z_d[k-j]\right\} \\ &= 2\sum_{i=0}^{N_I}\{\hat{g}_z[i] - g_t[i]\} R_{nn}[i-j]\} \end{aligned}$$

$$+ 2\sum_{i=0}^{N_I}\{\hat{g}_z[i]R_{dd}[i-j]\} \tag{26}$$

where $R_{nn}[j]$ and $R_{dd}[j]$ are the auto-correlation of $z_n[k]$ and $z_d[k]$, respectively. As $z_n[k]$ and $z_d[k]$ are both real signals, their auto-correlations are symmetric, i.e., $R_{nn}[i-j] = R_{nn}[j-i]$ and $R_{dd}[i-j] = R_{dd}[j-i]$ [2]. Therefore,

$$\begin{aligned}\frac{\partial J_n}{\partial \hat{g}_z[j]} &= 2\sum_{i=0}^{N_I}\{\hat{g}_z[i] - g_t[i]\}R_{nn}[j-i] \\ &\quad + 2\sum_{i=0}^{N_I}\hat{g}_z[i]R_{dd}[j-i]\} \\ &= 2[(\hat{\mathbf{g}_z} - \mathbf{g_t}) * \mathbf{R_{nn}}][j] + 2[\hat{\mathbf{g}_z} * \mathbf{R_{dd}}][j]. \end{aligned}\tag{27}$$

Next we show that the above partial derivative (27) equals to the cost function in (20). Thus, the minimal of the cost function J_n is attained if and only if the cost function $\hat{J}_n$ reaches its minimal at zero. By (1) and (12), the cross-correlation in (20), $r_{nz}[j]$, can be represented as

$$\begin{aligned}\hat{J}_n &= [\mathbf{E}\{(n_m[k] - \hat{n}[k])(\check{z}_m[k-j])\}]^2 \\ &= [\mathbf{E}\{n_m[k]\check{z}_m[k-j] - \hat{n}[k](\check{z}_m[k-j])\}]^2 \\ &= [\mathbf{E}\{n_m[k]\check{z}_m[k-j]\} - \mathbf{E}\{\hat{n}[k]z_n[k-j]\} \\ &\quad - \mathbf{E}\{\hat{n}[k]z_d[k-j]\}]^2 \end{aligned}\tag{28}$$

By (15) and the definition of convolution

$$\begin{aligned}\mathbf{E}\{\hat{n}[k]z_n[k-j]\} &= \mathbf{E}\{(\hat{\mathbf{g}_z} * \check{\mathbf{z}}_\mathbf{m})[k]z_n[k-j]\} \\ &= \mathbf{E}\left[\left(\sum_{i=0}^{N_i}\hat{g}_z[i]\check{z}_m[k-i]\right)z_n[k-j]\right] \\ &= \mathbf{E}\left\{\sum_{i=0}^{N_i}(\hat{g}_z[i]z_n[k-i]z_n[k-j])\right\} \\ &\quad + \mathbf{E}\left\{\sum_{i=0}^{N_i}(\hat{g}_z[i]z_d[k-i]z_n[k-j])\right\} \end{aligned}\tag{29}$$

By Assumption 3

$$\begin{aligned}&\mathbf{E}\left\{\sum_{i=0}^{N_i}\hat{g}_z[i]z_d[k-i]z_n[k-j]\right\} \\ &= \sum_{i=0}^{N_i} g_z[i]\mathbf{E}(z_d[k-i]z_n[k-j]) = 0. \end{aligned}\tag{30}$$

Substituting (30) back into (29) yields

$$\begin{aligned}\mathbf{E}\{\hat{n}[k]z_n[k-j]\} &= \left[\sum_{i=0}^{N_i}\mathbf{E}\left(\hat{g}_z[i]z_n[k-i]z_n[k-j]\right)\right] \\ &= \left[\sum_{i=0}^{N_i}\hat{g}_z[i]\mathbf{E}(z_n[k-i]z_n[k-j])\right] \\ &\triangleq \left[\sum_{i=0}^{N_I}\hat{g}_z[i]R_{nn}[i-j]\right]. \end{aligned}\tag{31}$$

Similarly,

$$\begin{aligned}\mathbf{E}\{\hat{n}[k]z_d[k-j]\} &= \left[\sum_{i=0}^{N_i}\hat{g}_z[i]\mathbf{E}(z_d[k-i]z_d[k-j])\right] \\ &\triangleq \left[\sum_{i=0}^{N_i}\hat{g}_z[i]R_{dd}[i-j]\right]. \end{aligned}\tag{32}$$

And

$$\begin{aligned}&\mathbf{E}\{n_m[k]\check{z}_m[k-j]\} \\ &= \mathbf{E}\{n^*[k]z_n[k-j]\} + \mathbf{E}\{n_d[k]z_n[k-j]\} \\ &\quad + \mathbf{E}\{n^*[k]z_d[k-j]\} + \mathbf{E}\{n_d[k]z_d[k-j]\}. \end{aligned}\tag{33}$$

By Assumption 3,

$$\mathbf{E}\{n_d[k]z_n[k-j]\} = \mathbf{E}\{n^*[k]z_d[k-j]\} = 0. \tag{34}$$

Substituting (34) into (33) yields

$$\begin{aligned}\mathbf{E}\{n_m[k]\check{z}_m[k-j]\} &= \mathbf{E}\{n^*[k]z_n[k-j]\} \\ &= \sum_{i=0}^{N_I} g_t[i]R_{nn}[i-j]. \end{aligned}\tag{35}$$

Hence, submitting (31), (32) and (35) back into (28) yields

$$\begin{aligned}\hat{J}_n &= \left\{\sum_{i=0}^{N_i}(g_t[i] - \hat{g}_z[i])R_{nn}[i-j] - \sum_{i=0}^{N_i}\hat{g}_z[i]R_{dd}[i-j]\right\}^2 \\ &= \{[(\mathbf{g_t} - \hat{\mathbf{g}_z}) * \mathbf{R_{nn}}][j] - [\hat{\mathbf{g}_z} * \mathbf{R_{dd}}][j]\}^2. \end{aligned}\tag{36}$$

Thus, the proof is completed by combining (27) with (36). □

Next, we obtain the optimal filter $g_z[k]$ by minimizing the cost function $\hat{J}_n$ to zero. We rewrite $\hat{J}_n$ in the matrix form as

$$\begin{aligned}\hat{J}_n &= \mathbf{E}\{\hat{n}_d[k-j]z_m[j]\}^2 \\ &= \mathbf{E}\{(n_m[k-j] - (\mathbf{g_z} * \mathbf{z_m}[k-j])z_m[j]\}^2 \\ &= \{\mathbf{E}\{\mathbf{n_m}\mathbf{z_m^T}\} - \mathbf{E}\{\mathbf{z_m}\mathbf{z_m^T}\}\mathbf{g_z}\}^2 \\ &= \mathbf{p_n} - \mathbf{R_z}\mathbf{g_z} \end{aligned}\tag{37}$$

where $\mathbf{R}_z$ is the auto-correlation matrix of measured image signal $z_m[k]$

$$\begin{aligned}\mathbf{R_z} &= \mathbf{E}\{\mathbf{z_m}\mathbf{z_m^T}\} \\ &\triangleq \begin{bmatrix} R_z[0] & R_z[1] & \dots & R_z[N_I-1] \\ R_z[1] & R_z[0] & \dots & R_z[N_I-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_z[N_I-1] & R_z[N_I-2] & \dots & R_z[0] \end{bmatrix} \end{aligned}\tag{38}$$

and the vector $\mathbf{p_n}$ is the cross-correlation between the measured acoustic noise $n_m[k]$ and the measured image signal $z_m[k]$

$$\mathbf{p_n} = \{\mathbf{E}\{\mathbf{n_m z_m^T}\} \triangleq [p_n(0)\, p_n(-1)\, \cdots\, p_n(-N_I+1)]^T. \tag{39}$$

Thus, by setting $\hat{J}_n = 0$ in (37), the optimal filter $g_z^*[k]$ can be readily obtained from (37) as a Wiener filter [22]

$$\mathbf{g_z^*} = \mathbf{R_z^{-1}} \mathbf{p_n}. \tag{40}$$

As both the acoustic measurement disturbance $n_d[k]$ and the total measured noise $n_m[k]$ are WSS, the auto-correlation matrix $\mathbf{R}_z$ above is nonsingular for any nonzero measurement disturbance $n_d[k]$ [22]. Therefore, the estimated acoustic noise is obtained as

$$\hat{n}^*[k] = [\mathbf{g_z^*} * (\mathbf{z_m} - \bar{\mathbf{z_m}})][k] \triangleq [\mathbf{g_z^*} * \hat{\mathbf{z_m}}][k]. \tag{41}$$

Steepest Descent Method to Compute the Wiener Filter: Finding the Wiener filter can be computationally costly when the related acoustic dynamics is complicated and a high-order FIR truncation of the Wiener filter (i.e., the order number N_F is large) is needed. Thus, we utilize the steepest descent method [22] to seek the optimal $g_z^*[k]$ via the following recursive iteration:

$$\mathbf{g}_{z,i+1} = \mathbf{g}_{z,i} + \mu \hat{n}_d[i] \mathbf{E}\{\mathbf{z_{m,i}}\} \tag{42}$$

where

$$\mathbf{z_{m,i}} = [z_m[i]\, z_m[i+1]\, \cdots\, z_m[i+N_F]]^T \quad \text{for } i = 0, 1, \ldots, N_I - N_F. \tag{43}$$

$\mu \in (0, 1]$ is the step size, chosen to ensure the convergence of the recursive iteration in (42), and initially $\mathbf{g}_{z,0} = \mathbf{0}$. As the expectation of $\mathbf{E}\{\mathbf{z}_i\}$ is, in general, unknown, the following unbiased estimation is used [2]:

$$\mathbf{E}\{\mathbf{z}_i\} = \lim_{Na \to \infty} \frac{1}{N_a} \sum_{i=0}^{N_a - 1} \mathbf{z}_{k-i} \tag{44}$$

where N_a is the number of samples used in the estimation. Thus, the following estimation is employed by using the values of the two signals at current sampling instant

$$\mathbf{E}\{\mathbf{z_{m,i}}\} \approx \mathbf{z_{m,i}}. \tag{45}$$

In the following, the optimal estimated filter obtained via (41)–(45) is denoted as $\mathbf{g}_z^\dagger$.

Modulator-Based Filter Optimization: As approximation [see (45)] is inevitable in seeking the optimal acoustic-noise filter $\hat{g}_z^*[k]$ above, and extraneous error can also be introduced due to the truncation involved in the Wiener filter, we propose the following modulator-based method to minimize the approximation error and further enhance the optimal estimated disturbance noise filter $\hat{g}_z^\dagger[\cdot]$ in (40) in the frequency domain via

$$\hat{G}_z^\ddagger(e^{j\omega_k}) = \alpha_\ddagger(e^{j\omega_k}) \hat{G}_z^\dagger(e^{j\omega_k}) \tag{46}$$

where $\hat{G}_z^\dagger(e^{j\omega_k})$ denotes the z-transform of the acoustic-noise filter $\hat{g}_z^\dagger[\cdot]$, $\omega_k = \frac{k\omega_s}{N_I}$ for $k = 0, 1, \ldots, N_I - 1$ with $\omega_s = 2\pi f_s$ are the discrete sampled frequencies, and $\alpha_\ddagger(e^{j\omega_k})$ is the optimal modulator to be determined, respectively. Specifically, the optimal modulator $\alpha_\ddagger(e^{j\omega_k})$ is to minimize the filter error $\delta g_z[k]$

$$\delta g_z[k] = g_z^*[k] - \hat{g}_z^\ddagger[k] \tag{47}$$

where $\hat{g}_z^*[k]$ is the exact optimal acoustic-noise filter given by (40).

It is shown in Lemma 2 below that minimization of the error of the acoustic-noise filter $\delta g_z[k]$ can be converted to minimizing the coherence between the estimated measurement disturbance $\hat{n}_d^*[k]$ and the unbiased measured image signal $\breve{z}_m[k]$, i.e., find $\alpha_\ddagger(e^{j\omega_k})$ to minimize the coherence between $\hat{n}_d^*[k]$ and $\breve{z}_m[k]$

$$\mathbf{C}_{nz}(e^{j\omega_k}) = \frac{\mathbf{E}[\hat{N}_d^*(e^{j\omega_k}) \breve{Z}_m(e^{j\omega_k})]^2}{\hat{P}_{N_d}(e^{j\omega_k}) P_{\breve{Z}_m}(e^{j\omega_k})} \quad \text{where} \tag{48}$$

$$\hat{N}_d(e^{j\omega_k}) = N_m(e^{j\omega_k}) - \alpha_\ddagger(e^{j\omega_k}) G_z^*(e^{j\omega_k}) \breve{Z}_m(e^{j\omega_k}) \tag{49}$$

and $\hat{P}_{N_d}(e^{j\omega_k})$ and $P_{\breve{Z}_m}(e^{j\omega_k})$ are the power spectral densities of $\hat{N}_d(e^{j\omega_k})$ and $\breve{Z}_m(e^{j\omega_k})$, respectively.

$$\hat{P}_{N_d}(e^{j\omega_k}) = \|\hat{N}_d(e^{j\omega_k})\|^2 \quad P_{\breve{Z}_m}(e^{j\omega_k}) = \|\breve{Z}_m(e^{j\omega_k})\|^2. \tag{50}$$

Lemma 2: Let Assumptions 1–3 hold, then the error of the noise filter $\delta g_z[k]$ is minimized in two-norm sense if and only if at any given ω_k the coherence between the estimated measurement disturbance $\hat{n}_d^*[k]$ and the measured image signal $\breve{z}_m[k]$, $\mathbf{C}_{nz}(e^{j\omega_k})$, is minimized.

Proof: By the Convolution theorem [25], the cost function in (20) can be represented in frequency domain as

$$\min \mathcal{J}_n = \{\mathbf{E}[\hat{N}_d^*(e^{j\omega_k}) \breve{Z}_m(e^{j\omega_k})]\}^2 \tag{51}$$

and the coherence between $\hat{n}_e[k]$ and $z_m[k]$ at any given sampled frequency ω_k is given as

$$\begin{aligned} \mathbf{C}_{nz}(e^{j\omega_k}) &= \frac{\{\mathbf{E}[\hat{N}_d^*(e^{j\omega_k}) \breve{Z}_m(e^{j\omega_k})]\}^2}{\hat{P}_{N_d}(e^{j\omega_k}) P_{\breve{Z}_m}(e^{j\omega_k})} \\ &= \frac{1}{P_{N_d}(e^{j\omega_k}) P_Z(e^{j\omega_k})} \\ &\quad \{[G_t(e^{j\omega_k}) - G_z(e^{j\omega_k}) + \delta G_z(e^{j\omega_k})] R_{zn}(e^{j\omega_k}) \\ &\quad - [G_z(e^{j\omega_k}) - \delta G_z(e^{j\omega_k})] R_{dd}(e^{j\omega_k})\}^2 \end{aligned} \tag{52}$$

where $P_{N_d}(e^{j\omega_k})$ is the power spectral density of the estimated measurement disturbance $\hat{n}_d[k]$ [see (1)]. The above (52) and (36) implies that the error in the estimated acoustic noise filter $\hat{g}_z^\ddagger[k]$ compared with the true disturbance dynamic $G_z(e^{j\omega_k})$ is minimized, if and only if the coherence $\mathbf{C}_{nz}(e^{j\omega_k})$ is minimized. Thus, by minimizing the coherence $\mathbf{C}_{nz}(e^{j\omega_k})$ at every sampled frequency ω_k, the optimal estimated acoustic noise filter $\hat{g}_z^\ddagger[k]$ converges to the exact acoustic noise filter $g_z^*[k]$. This completes the proof. □

The optimal $\alpha^\ddagger(e^{j\omega_k})$ that minimizes the cost function J_n in (9) can be obtained through a gradient-based iterative descend method [20] by

$$\alpha_i^\ddagger(j\omega_k) = \alpha_{i-1}^\ddagger(j\omega_k) + \lambda \frac{\partial \mathbf{C}_{nz}(j\omega_k, \alpha_{i-1}^\ddagger)}{\partial \hat{G}_{z,i-1}^\ddagger(j\omega_k)}$$

$$= \alpha_{i-1}^{\ddagger}(j\omega_k) + \lambda \frac{\partial \mathbf{C}_{nz}(j\omega_k, \alpha_{i-1}^{\ddagger})}{\partial \alpha_{i-1}^{\ddagger}(j\omega_k)} \frac{1}{\hat{G}_z^{\dagger}(j\omega_k)} \tag{53}$$

where

$$\frac{\partial \mathbf{C}_{nz}(j\omega_k, \alpha_i^{\ddagger})}{\partial \alpha^{\ddagger}(j\omega_k)} \approx \frac{\mathbf{C}_{nf}(j\omega_k, \alpha_i^{\ddagger} + \delta\alpha^{\ddagger}) - \mathbf{C}_{nf}(j\omega_k, \alpha_i^{\ddagger})}{\delta\alpha^{\ddagger}}. \tag{54}$$

Readers are referred to [20] for more details.

C. [𝕆-2] Optimal Acoustic-Image Filter Design

The optimal acoustic-image filter $\hat{g}_N[k]$ to minimize the error between the true image and the filtered image (objective 𝕆-2) can also be designed through the above Wiener-filter-based modulator optimization approach. Specifically, a Wiener Filter $\hat{g}_n[k]$ is designed to optimize the estimated true image $z_f[k]$ by minimizing the power of correlation between the estimated acoustic noise $\hat{n}[k]$ and the measured image signal $z_f[k]$—It can be shown similarly that minimizing the cost function J_z in (9) is equivalent to minimizing the following correlation $\hat{J}_z$:

$$\min \hat{J}_z = \mathbf{E}\{\hat{n}[k-j] z_f[j]\}^2. \tag{55}$$

Thus, the same minimization process above [in (37)–(40)] can be applied to obtain the optimal acoustic filter $g_n^*[k]$ as a Wiener Filter

$$\hat{\mathbf{g}}_n^* = \mathbf{R}_n^{-1} \mathbf{p_{nz}} \tag{56}$$

where $\mathbf{R_n}$ is the auto-correlation matrix of the estimated acoustic noise obtained from Section-II B, $\hat{\mathbf{n}}$,

$$\mathbf{R_n} = \mathbf{E}\{\hat{\mathbf{n}}\hat{\mathbf{n}}^T\} = \begin{bmatrix} R_n[0] & R_n[1] & \dots & R_n[N_I-1] \\ R_n[1] & R_n[0] & \dots & R_n[N_I-2] \\ \vdots & \vdots & \ddots & \vdots \\ R_n[N_I-1] & R_n[N_I-2] & \dots & R_n[0] \end{bmatrix} \tag{57}$$

and $\mathbf{p_{nz}} = \mathbf{E}\{\hat{\mathbf{n}}\mathbf{z_m^T}\}$ is the cross-correlation between the estimated noise $\hat{n}^*[k]$ and the measured sample image signal $z_m[k]$, respectively. The filtered image is then obtained as

$$z_f[k] = z[k] - (\hat{\mathbf{g}}_n^* * \hat{\mathbf{n}}^*)[k] \tag{58}$$

with $\hat{n}^*[k]$ given in (41).

Similarly, the steepest descend method is used to obtain the acoustic image filter by replacing the $\hat{n}_d[i]$ and $\mathbf{z_{m,i}}$ with $z_m[i] - z_f[i]$ and $\mathbf{n_i^*}$, respectively, where

$$\mathbf{n_i^*} = [n^*[i]\, n^*[i+1] \cdots n^*[i+N_F]]^T \quad \text{for } i = 0, 1, \dots, N_I - N_F. \tag{59}$$

We denote the optimal estimated filter obtained via (41)–(44) as $\hat{\mathbf{g}}_n^{\dagger}(\cdot)$. Then, $\hat{\mathbf{g}}_n^{\dagger}(\cdot)$ is further enhanced by a frequency-dependent modulator $\beta_{\ddagger}$ via

$$\hat{G}_n^{\ddagger}(e^{j\omega_k}) = \beta_{\ddagger}(e^{j\omega_k}) G_n^{\dagger}(e^{j\omega_k}) \tag{60}$$

where the optimal modulator $\beta_{\ddagger}$ is to minimize the filter error $\delta g_n[k]$

$$\delta g_n[k] = g_n^*[k] - \hat{g}_n^{\ddagger}[k] \tag{61}$$

and obtained by minimizing the following coherence between $n^*[k]$ and the filtered image $z_F[k]$:

$$\min \mathbf{C}_{nf}(e^{j\omega_k}) = \frac{\mathbf{E}[N^*(e^{j\omega_k}) Z_F(e^{j\omega_k})]^2}{P_N(e^{j\omega_k}) P_{Z_F}(e^{j\omega_k})}. \tag{62}$$

Lemma 3: Let Assumptions 1–3 hold, then the error of the image filter $\delta g_n[k]$ is minimized in two-norm sense if and only if at any given ω_k, the coherence between the estimated acoustic noise $\hat{n}^*[k]$ and the filtered image signal $z_f[k]$, $\mathbf{C}_{nf}(e^{j\omega_k})$, is minimized.

The proof is similar to that of Lemma 2 and is omitted.

The gradient-based iterative descend method is also employed to obtain the optimal $\beta^{\ddagger}(e^{j\omega_k})$ to minimize the cost function (62).

III. Experiment Example

The proposed approach was demonstrated through an AFM imaging experiment. The objective was to show that by using the proposed approach, the acoustic-caused image distortion from an arbitrary and unknown acoustic source 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.) as shown in Fig. 1. The ground vibration induced by the speaker (when it was broadcasting sound to the environment) was damped by the optical table (on which the AFM system was placed, see Fig. 1), and its effect on the cantilever vibration became part of the PAD of the AFM system, thereby, was accounted for through the proposed approach. The noise sensor signal was first pre-filtered and amplified using a homemade Op-Amp circuit, and then measured via a data acquisition system (NI RIO, USB-7856R, National Instrument Inc.). All the filtering algorithms were designed in MATLAB (Mathworks Inc.) and then implemented using NI-USB-7856R FPGA board.

Experimental Implementation: First, the acoustic-noise-effected AFM images were acquired with the noise source placed at three different "unknown" locations and the sensor (microphone) placed at a fixed location, respectively, as shown in Fig. 3. A silicon sample and a calibration sample (STR-1800R) were imaged at a scan rate of 5 Hz under the TM when a band-limited (20–1 kHz) white noise of zero-mean and constant variance of 100 dB was broadcasted to the room through the speaker. To mimic the scenario of low SNR acoustic noise measurement, a different band-limited zero-mean white noise (20–1 kHz) acoustic signal was introduced as the measurement disturbance through an earphone sticked tightly to the microphone. The amplitude of this disturbance noise was adjusted such that it can be clearly captured by the microphone but didn't induce measurable AFM probe vibration. Both the noise signals and

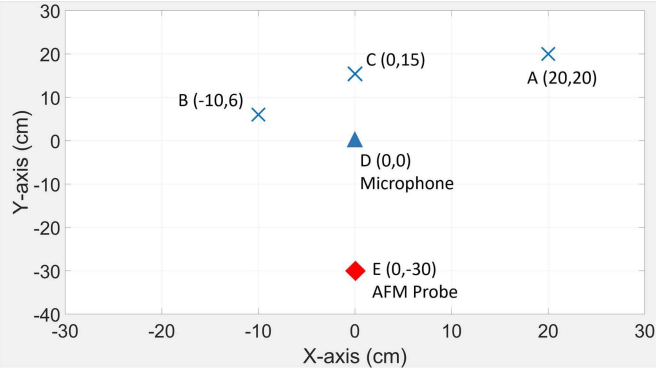


Fig. 3. Location of the acoustic noise source (speaker) at three "unknown" locations, A, B, and C, with respect to that of the sensor (microphone), D, and the AFM probe, E, respectively.

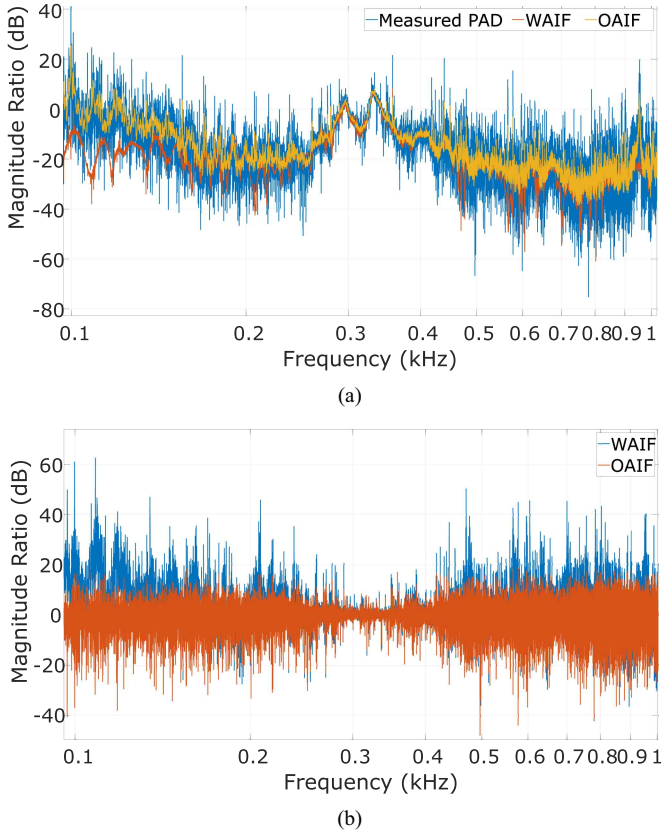


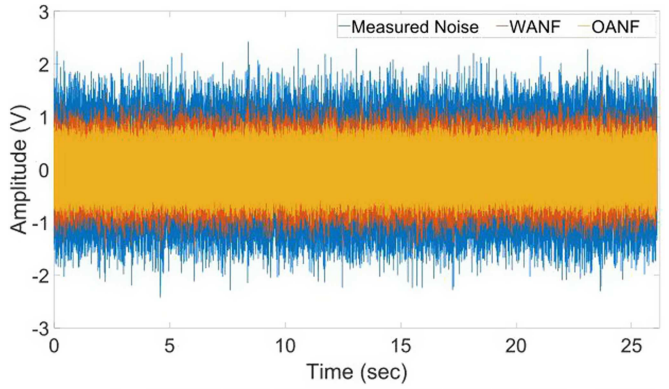
Fig. 4. (a) Comparison of the frequency response (magnitude part) of the measured PAD $G_n(e^{j\omega_k})$ and those estimated by using the Wiener-FIR acoustic-image filter (WAIF) and the optimal acoustic-image filter (OAIF) method, and (b) the comparison of the error (magnitude part) of the estimated dynamics using the WAIF, and that using the OAIF with respect to the "true" measured PAD, respectively.

the z-axis piezo displacements were acquired simultaneously during the imaging processes. This imaging and measurement process was repeated for the other two unknown noise locations "B" and "C", respectively.

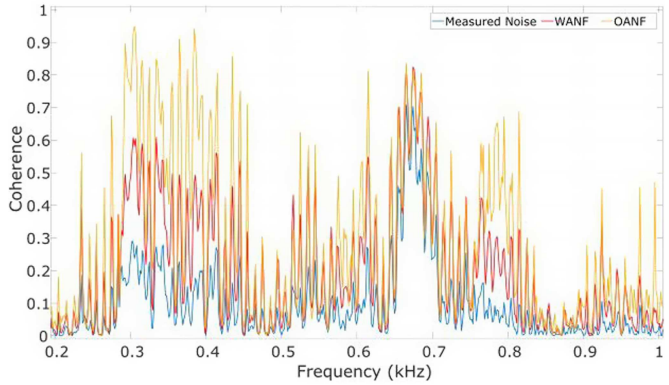
To filter the acoustic noise acquired, the optimal acoustic-noise filter, $\mathbf{g}_{\mathbf{z}}^*$, and the optimal acoustic-image filter, $\mathbf{g}_{\mathbf{n}}^*$, were constructed using the PAD from the measured noise signal $n_m[k]$ and total image signal $z_m[k]$. First, the filter coefficients of the Wiener filter were obtained using (42)–(45), where the Wiener filter was truncated at the order of 4096 and step size μ was chosen at 0.1, called the Wiener-FIR acoustic-noise and Wiener-FIR acoustic-image filter below, respectively. Next, both filters were further optimized by using the modulator-based coherence minimization method (described in Section 2). Specifically, the optimal modulator $\alpha_\dagger(e^{j\omega_k})$ and $\beta_\dagger(e^{j\omega_k})$ were obtained iteratively described in [20], where the step size was chosen at 0.05. Finally, the filtered images were obtained by using, first, the optimized acoustic-noise filter $\mathbf{g}_{\mathbf{z}}^*$ to "clean up" the measured acoustic noise $n_m[k]$, and secondly the optimized acoustic-noise filter $\mathbf{g}_{\mathbf{z}}^*$ along with the filtered acoustic signal $\hat{n}^*[k]$ to filter the measured raw image $z_m[k]$.

The same procedure was repeated for filtering the images captured at the other two unknown locations "B" and "C", respectively. For comparison, the images were also filtered directly by using the Wiener-FIR acoustic-noise filter followed by the Wiener-FIR acoustic-image filter, and were filtered by using a bandstop filter with the cutoff frequency in the range between 250 and 400 Hz (where the PAD was the largest in magnitude) was chosen, respectively [see Fig. 4(a)]. Finally, as a benchmark filtering result—to evaluate the proposed approach in improving the SNR ($\mathbb{O}$-1) and accounting for the unknown sensor-noise un-collocation effect ($\mathbb{O}$-2), the images were also processed by using the acoustic noise measured under the sensor-noise collocation and the filter constructed by using the PAD obtained under this collocation condition and further improved via the modulator-optimization method for the silicon and the calibration sample, respectively.

Experimental Results and Discussion: The experimental results are shown in Figs. 4–11. To evaluate the efficacy of the proposed Wiener-FIR acoustic-noise filter (WANF) and the optimal acoustic-noise filter (OANF) in improving the SNR of the measured acoustic signals from an unknown location, the PAD obtained by using the filtered acoustic signal via the proposed WANF and proposed OANF technique, are compared to that quantified by using the measured signal directly (without filtering) in Fig. 4(a). The difference of these two PADs with respect to the experimentally measured "true" PAD (obtained by measuring the acoustic noise under the sensor-speaker collocation and without induced acoustic measurement noise), i.e., the error of these two PADs, are compared in Fig. 4(b). Also, the measured raw noise signal is compared to the filtered noise signal by using the WANF and the OANF in Fig. 5 for location "A" (as an example), respectively. The corresponding raw acoustic-effected images of the silicon sample and the calibration sample are compared to those filtered by using the bandstop filter, Wiener-FIR acoustic-image filter (WAIF) and the optimal acoustic-image filter (OAIF) in Figs. 6 and 7, respectively. Then, the filtering results of the proposed OANF and the OAIF for the images captured at the three unknown speaker locations are shown Figs. 6 and 9, with comparison to those obtained in location "D" setup (under the sensor-noise collocation condition). Finally, the image quality was evaluated by relative 2-norm error $E_2(\%)$ for



(a)



(b)

Fig. 5. (a) Comparison of the measured acoustic noise signal of low SNR compared to the filtered noise by using the Wiener-FIR acoustic-noise filter (WANF) and the optimal acoustic-noise filter (OANF), and (b) comparison of the coherence between the z-axis displacement to the measured noise, to the noise filtered by Wiener-FIR acoustic-noise filter (WANF) and the optimal acoustic-noise filter (OANF), respectively.

the raw image, WANF followed by the WAIF, and the OANF followed by the OAIF in Fig. 10, respectively, and the image error by using the proposed OANF and OAIF techniques for the results obtained in the three unknown location set-ups ("A", "B," and "C") are also compared to the benchmark reference (i.e., obtained in the collocation set-up, location "D" set-up) in Fig. 11, respectively. The relative 2-norm error $E_2(\%)$ was computed by using the image obtained in the quiet condition, and across all scanned lines of the entire image.

The experimental results showed that by using the proposed approach, the SNR of the acoustic signal measured in AFM imaging was substantially improved. First, the WANF captured the measured PAD well, with its dynamics (represented by its poles and zeros) overlapped with those of the measured PAD closely [see Fig. 4(a)]. However, significant deviation from the measured PAD existed in the low frequency range between 100 and 200 Hz. Such an error on the PAD was significantly reduced by using the proposed OANF as shown in Fig. 4(b). Next, as shown in Fig. 5, the SNR of the measured acoustic noise signal was significantly improved by using the proposed WANF as more than 50% of the irrelevant noise was removed from the signal. This substantial improvement in SNR was further enhanced by using the proposed OANF with another 10% SNR improvement. Moreover, the coherence between the noise and image was improved by more than 20% by the proposed WANF and 40% by the proposed OANF. Such an improvement of the SNR of the acoustic measurement directly contributed to the filtering improvement of image distortions obtained later.

The experimental results also demonstrated that by using the proposed technique, the image distortion caused by acoustic disturbance from an unknown location was significantly reduced. As shown in Figs. 6(c) and 7(c), by using the WAIF, the acoustic-caused image distortion was largely removed, whereas when using the bandstop filter, although the acoustic-caused ripple artifacts across the topography image was also largely removed and the image quality had improved considerably on silicon sample, the edge of the pitches were severely smeared in the filtered image of the calibration sample [see Fig. 7(b)]. However, by using the proposed OAIF, the image quality was further improved, with the image quality visually improved further [see Figs. 6(d) and 7(d)]. This improvement can be seen more clearly from the relative image error comparison—the relative 2-norm error was reduced by 78% on the silicon sample and 68% on the calibration sample by using the WAIF, and then further reduced by another 10–12% by using the proposed OAIF (see Fig. 10). Finally, the experimental results also demonstrated the proposed OAIF is robust against the variation of the noise source location. As shown in Fig. 11, the 2-norm image error at three different locations are consistent between 11% and 19% and compared well to the 12% to 13% image error under the condition that the noise source and microphone collocated on both samples. Overall, the improvement attained was similar to that reported recently [20], without constructing dictionaries of acoustic dynamics via a pirori measurement and characterization—as needed previously [20]—to account for unknown acoustic source location. Instead, such an adverse effect was effectively accounted for through the proposed DDRF technique. Such an enhancement substantially reduced the cost and efforts needed in practical implementations, making the technique potentially an effective tool for practical AFM imaging applications. The proposed technique can be easily extended to other areas that are sensitive to acoustic and vibration effects (e.g., precision robotics, Scanning Tunnel Electron Microscope and E-beam Lithography, etc.). This will be explored in our future work soon.

Although the proposed technique is for offline postimaging process, it can be readily expanded to more general AFM imaging applications. For real-time applications, the technique can be implemented by inserting a short "pause" period without interfering the imaging process. For example, the proposed algorithm can be used to process the data at the end of the scanning of each line (right before the start of the scanning on the next line). Additional sampling periods can be added if needed to process the algorithm and the algorithm can be transformed to the frequency domain and computed across multiple sampling periods via the time-distributive FFT algorithm [26]. Also, although the proposed technique is for arbitrary and unknown but *fixed* acoustic noise source, it can be easily extended to address the change of noise source location being known a priori (e.g., preprogrammed) and slow, or in a discrete manner

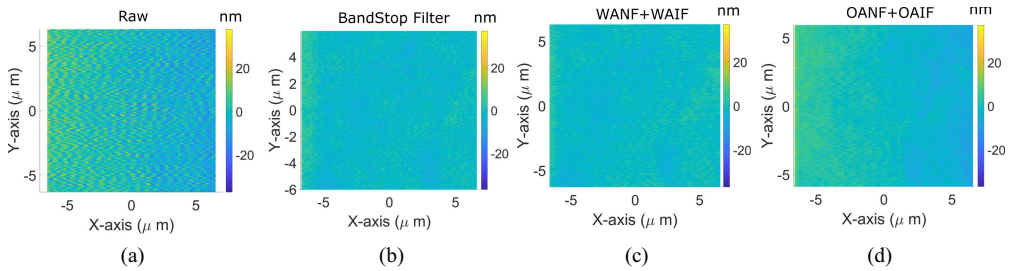


Fig. 6. Comparison of the original raw image of a silicon sample obtained at scan rate of 5 Hz (a) under the induced acoustic noise for location "A" set up in Fig. 3, to those filtered by (b) the bandstop filter, (c) the Wiener-FIR acoustic-noise combined with Wiener-FIR acoustic-image filters and (d) the optimal acoustic-noise combined with optimal acoustic-image filters, respectively.

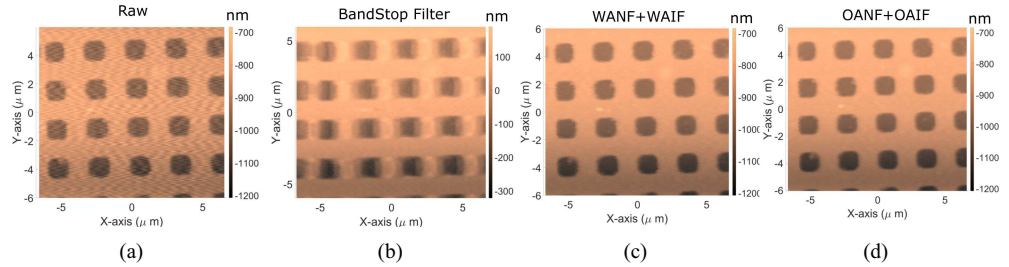


Fig. 7. Comparison of the original raw image of a calibration sample obtained at scan rate of 5 Hz (a) under the induced acoustic noise for location "A" set up in Fig. 3, to those filtered by (b) the bandstop filter, (c) the Wiener-FIR acoustic-noise combined with Wiener-FIR acoustic-image filters and (d) the optimal acoustic-noise combined with optimal acoustic-image filters, respectively.

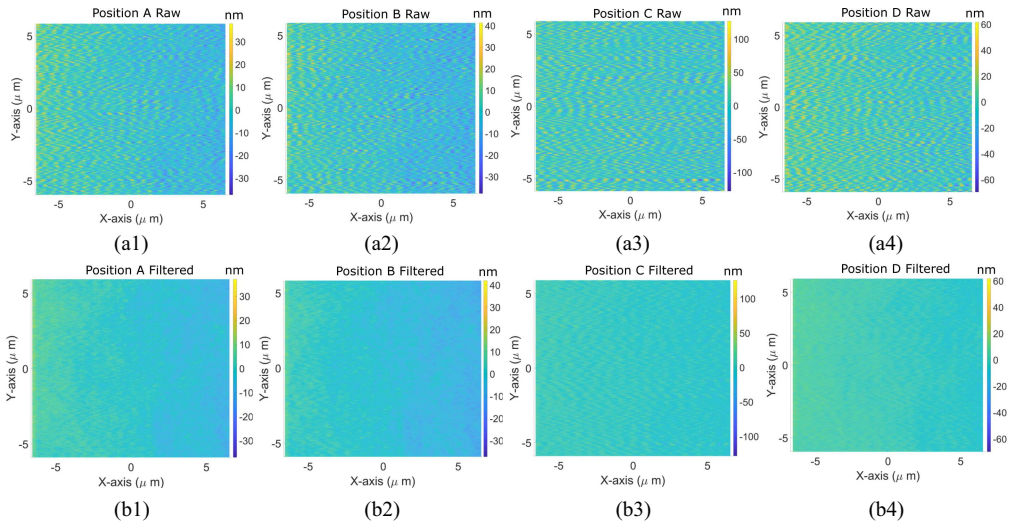


Fig. 8. Comparison of (first row) the raw image of the silicon sample obtained at scan rate of 5 Hz, under the induced acoustic noise (a1 to a4) in the location "A" to "D" setup in Fig. 3, respectively, to (second row, b1–b4) the correspond ones filtered by using the DDRF technique, respectively, where (a4) and (b4) (location "D" set-up) are the results for the sensor-noise collocated case.

(i.e., the noise source transits to different locations swiftly and stays at those locations for a known period of time). In these scenarios, the proposed technique can be applied to construct and apply the filters in a section-by-section manner. More generally, the proposed technique can be extended to the case where the change of the acoustic noise source arbitrarily and unknown (not common in AFM applications). This will be pursued as future work. Moreover, the proposed technique is not AFM manufacturer dependent and can be equally applied to other types of AFM systems under other operation conditions.

IV. Conclusion

A data-driven robust-optimal filtering technique was developed to eliminate AFM image distortion caused by acoustic

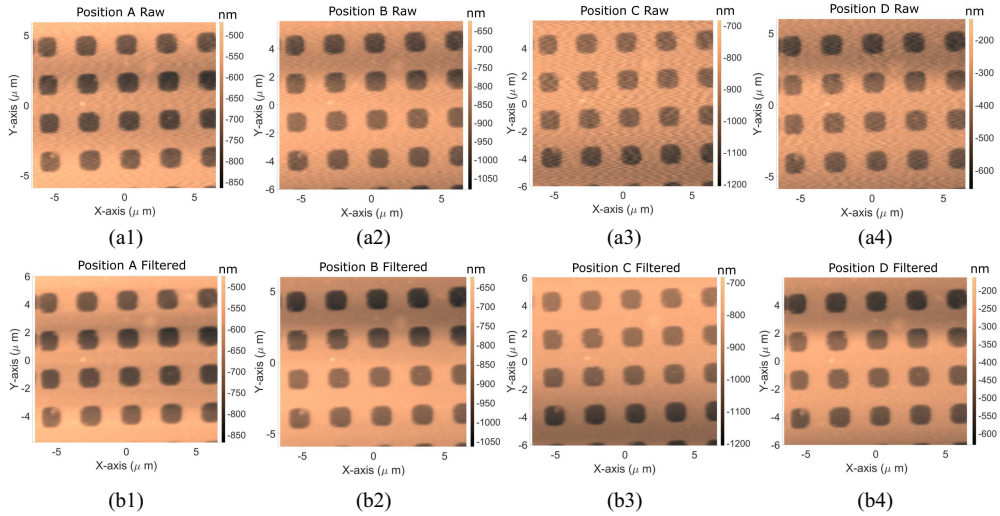


Fig. 9. Comparison of (first row) the raw image of the calibration sample obtained at scan rate of 5 Hz, under the induced acoustic noise (a1 to a4) in the location "A" to "D" setup in Fig. 3, respectively, to (second row, b1–b4) the correspond ones filtered by using the DRF technique, respectively, where (a4) and (b4) (location "D" set-up) are the results for the sensor-noise collocated case.

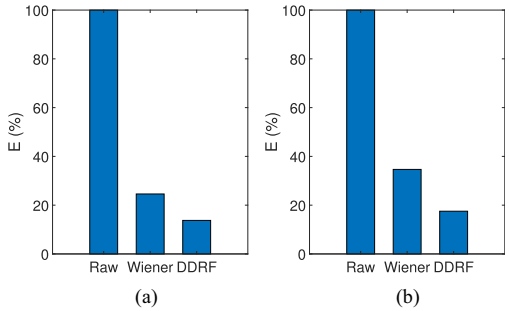


Fig. 10. Comparison of the normalized image error of the calibration sample for the Wiener-FIR filters and DDRF technique in the relative 2-norm, for location "A" set-up.

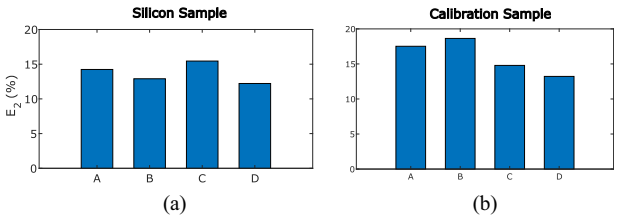


Fig. 11. Comparison of the normalized image error (with respect to the raw image error) for the calibration sample in 2-norm at four different locations, respectively.

noise. The Wiener filter in the FIR representation is explored to construct the filter and improve the SNR of the measured acoustic signal. It is shown that by introducing a modulator into the filters, the error in the estimated acoustic dynamics and the low SNR of the measured acoustic noise—both due to the acoustic noise location being arbitrary and unknown—can be eliminated by optimizing the modulator via a gradient-based coherence minimization approach. The efficacy of the proposed approach was demonstrated by filtering experimentally measured AFM images of a calibration sample. The filtering results showed that by using the proposed technique, the image distortion was substantially reduced. The future work includes the extension of the proposed technique to other acoustic noise and vibration cancellation areas including medical imaging.

REFERENCES

[1] S. Ducourtieux and B. Poyet, "Development of a metrological atomic force microscope with minimized Abbe error and differential interferometer-based real-time position control," *Meas. Sci. Technol.*, vol. 22, no. 9, 2011, Art. no. 94010.

[2] P. Eaton and P. West, *Atomic Force Microscopy*. London, U.K.: Oxford Univ. Press. 2010.

[3] F. Gołek, P. Mazur, Z. Ryszka, and S. Zuber, "AFM image artifacts," *Appl. Surf. Sci.*, vol. 304, pp. 11–19, Jun. 2014, doi: 10.1016/j.apsusc.2014.01.149.

[4] T. Uchihashi, H. Watanabe, S. Fukuda, M. Shibata, and T. Ando, "Functional extension of high-speed AFM for wider biological applications," *Ultramicroscopy*, vol. 160, 2016, pp. 182–196, doi: 10.1016/j.ultramic.2015.10.017.

[5] J. Ren and Q. Zou, "A control-based approach to accurate nanoindentation quantification in broadband nanomechanical measurement using scanning probe microscope," *IEEE Trans. Nanotechnol.*, vol. 13, no. 1, pp. 46–54, Jan. 2014, doi: 10.1109/TNANO.2013.2287505.

[6] Z. Wang, J. Tan, Q. Zou, and W. Jiang, "Mechanical-plowing-based high-speed patterning on hard material via advanced-control and ultrasonic probe vibration," *Rev. Sci. Instrum.*, vol. 84, 2013, Art. no. 113704, doi: 10.1063/1.4832046.

[7] G. M. Clayton, S. Tien, K. K. Leang, Q. Zou, and S. Devasia, "A review of feedforward control approaches in nanopositioning for high-speed SPM," *ASME J. Dyn. Sys., Meas., Control*, vol. 131, no. 6, Nov. 2009, Art. no. 061101. Accessed: Oct. 28, 2009, doi: 10.1115/1.4000158.

[8] G. Schitter, K. J. Astrom, B. E. DeMartini, P. J. Thurner, K. L. Turner, and P. K. Hansma, "Design and modeling of a high-speed AFM-Scanner," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 5, pp. 906–915, Sep. 2007, doi: 10.1109/TCST.2007.902953.

[9] T. Ando et al., "High-speed AFM and nano-visualization of biomolecular processes," *Pflugers Arch. - Eur. J. Physiol.*, vol. 456, pp. 211–225, 2008, doi: 10.1007/s00424-007-0406-0.

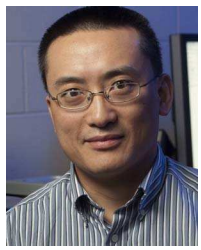
[10] Y. Wu, H. Chen, N. Sun, Z. Fan, and Y. Fang, "Neural network based adaptive control for a piezoelectric actuator with model uncertainty and unknown external disturbance," *Int. J. Robust Nonlinear Control*, vol. 33, no. 3, pp. 2251–2272, 2023.

[11] G. Liu, N. Sun, T. Yang, Z. Liu, and Y. Fang, "Equivalent-input-disturbance rejection-based adaptive motion control for pneumatic artificial muscle arms via hysteresis compensation models," *Control Eng. Pract.*, vol. 138, 2023, Art. no. 105609.
[12] W. Zhang, G. Meng, and Z. Peng, "Nonlinear dynamic analysis of atomic force microscopy under bounded noise parametric excitation," *IEEE/ASME Trans. Mechatron.*, vol. 16, no. 6, pp. 1063–1072, Dec. 2011, doi: 10.1109/TMECH.2010.2073715.
[13] S. Ito, D. Neyer, S. Pirker, J. Steininger, and G. Schitter, "Atomic force microscopy using voice coil actuators for vibration isolation," in *Proc. IEEE Int. Conf. Adv. Intell. Mechatron.*, 2015, pp. 470–475, doi: 10.1109/AIM.2015.7222578.
[14] F. Benmouna and D. Johannsmann, "Hydrodynamic interaction of AFM cantilevers with solid walls: An investigation based on AFM noise analysis," *Eur. Phys. J. E*, vol. 9, pp. 435–441, 2002, doi: 10.1140/epje/i2002-10096-x.
[15] D. Tranchida, S. Piccarolo, and R. A. C. Deblieck, "Some experimental issues of AFM tip blind estimation: The effect of noise and resolution," *Meas. Sci. Technol.*, vol. 17, no. 10, pp. 2630–2636, Aug. 2006.
[16] Y. Suzuki, N. Sakai, and A. Yoshida, "High-speed atomic force microscopy combined with inverted optical microscopy for studying cellular events," *Sci. Rep.*, vol. 3, 2013, Art. no. 2131, doi: 10.1038/srep02131.
[17] A. Sebastian, A. Gannepalli, and M. V. Salapaka, "A review of the systems approach to the analysis of dynamic-mode atomic force microscopy," *IEEE Trans. Control Syst. Technol.*, vol. 15, no. 5, pp. 952–959, Sep. 2007.
[18] S. Yi, T. Li, and Q. Zou, "Active control of acoustics-caused nano-vibration in atomic force microscope imaging," *Ultramicroscopy*, vol. 195, pp. 101–110, Dec. 2018, doi: 10.1016/j.ultramic.2018.07.006.
[19] M. G. Ruppert, K. S. Karvinen, S. L. Wiggins, and S. O. R. Moheimani, "A Kalman filter for amplitude estimation in high-speed dynamic mode atomic force microscopy," *IEEE Trans. Control Syst. Technol.*, vol. 24, no. 1, pp. 276–284, Jan. 2016, doi: 10.1109/TCST.2015.2435654.
[20] J. Chen and Q. Zou, "Data-driven dynamics-based optimal filtering of acoustic noise at arbitrary location in atomic force microscope imaging [J]," *Ultramicroscopy*, vol. 242, 2022, Art. no. 113614.
[21] S. Zhai et al., "AFM vibration noise reduction via squeeze-film-damping-effect sensing method," *IEEE Trans. Instrum. Meas.*, vol. 73, 2024, Art. no. 1007411, doi: 10.1109/TIM.2024.3440377.
[22] S. S. Haykin, *Adaptive Filter Theory*. India: Pearson Education, 2013.
[23] J. Chen and Q. Zou, "Data-driven robust acoustic noise filtering for atomic force microscope image," in *Proc. IEEE/ASME Int. Conf. Adv. Intell. Mechatron.*, Seattle, WA, USA, 2023, pp. 99–104, doi: 10.1109/AIM46323.2023.10196180.
[24] N. Gadegaard, "Atomic force microscopy in biology: Technology and techniques," *Biotechnic Histochemistry*, vol. 81, pp 87–97, 2006, doi: 10.1080/10520290600783143.
[25] C. D. McGillem and G. R. Cooper, *Continuous and Discrete Signal and System Analysis*, 2nd ed., R. Holt and Winston, Eds., Philadelphia, PA: Saunders College Pub, 1984.
[26] J. Liu, B. Yan, and Q. Zou, "Optimal time-distributed fast Fourier transform: Application to online iterative learning control—experimental high-speed nanopositioning example," *Mechatronics*, vol. 41, pp. 114–124, 2017, doi: 10.1016/j.mechatronics.2016.11.007.

Jiarong Chen received the B.S. degree in automatic control from the University of Electronic Science and Technology of China, Chengdu, China in 2016, and the M.S. degree in electrical and computer engineering from Rutgers, the State University of New Jersey, Piscataway, NJ, USA, in 2018.

He is currently working toward the Ph.D. degree in the mechanical and aerospace engineering with the Department of Rutgers, the State University of New Jersey, Piscataway, NJ, USA.

His research interests include control of high-speed scanning probe microscopy and nanoscale acoustic noise control.



Qingze Zou received the B.S. degree in automatic control from the University of Electronic Science and Technology of China, Chengdu, China, in 1994, the M.S. degree in mechanical engineering from Tsinghua University, Beijing, China, in 1997, the Ph.D. degree in mechanical engineering from the University of Washington, Seattle, WA, USA, in 2003.

He is currently a Professor of mechanical and aerospace engineering with Rutgers, the State University of New Jersey, Piscataway, NJ, USA. Previously he had taught at the Iowa State University, USA. His research interests include data-driven, trajectory-mapping-based precision tracking and motion control, control of high-speed scanning probe microscopy, nanoscale acoustic noise and vibration cancellation and control, stomata dynamics modeling and measurement, and robotic manufacturing.

He had been an Associate Editor of *ASME Journal of Dynamic Systems, Measurement and Control*, and IFAC-Control Engineering Practice, a Technical Editor of IEEE/ASME TRANSACTIONS ON MECHATRONICS, and the lead Guest Editor of the 4th and 5th focused section on the TMECH/AIM emerging topics. Currently he is an Associate Editor of *IFAC-Mechatronics*. He is a Fellow of ASME. Dr. Zou was a recipient of the NSF CAREER Award in 2009, and the O. Hugo Schuck Best Paper Award from the American Automatic Control Council in 2010.