



A review of demodulation techniques for multifrequency atomic force microscopy

David M. Harcombe*, Michael G. Ruppert and Andrew J. Fleming

Review

Open Access

Address:

School of Electrical Engineering and Computing, The University of Newcastle, Callaghan, NSW, 2308, Australia

Email:

David M. Harcombe* - David.Harcombe@uon.edu.au;
Michael G. Ruppert - Michael.Ruppert@newcastle.edu.au;
Andrew J. Fleming - Andrew.Fleming@newcastle.edu.au

* Corresponding author

Keywords:

atomic force microscopy (AFM); multifrequency; demodulation; Kalman filter; Lyapunov filter; digital signal processing; field-programmable gate array (FPGA)

Beilstein J. Nanotechnol. **2020**, *11*, 76–91.
doi:10.3762/bjnano.11.8

Received: 04 September 2019

Accepted: 11 December 2019

Published: 07 January 2020

This article is part of the thematic issue "Advanced atomic force microscopy II".

Guest Editor: T. Glatzel

© 2020 Harcombe et al.; licensee Beilstein-Institut.
License and terms: see end of document.

Abstract

This article compares the performance of traditional and recently proposed demodulators for multifrequency atomic force microscopy. The compared methods include the lock-in amplifier, coherent demodulator, Kalman filter, Lyapunov filter, and direct-design demodulator. Each method is implemented on a field-programmable gate array (FPGA) with a sampling rate of 1.5 MHz. The metrics for comparison include the sensitivity to other frequency components and the magnitude of demodulation artifacts for a range of demodulator bandwidths. Performance differences are demonstrated through higher harmonic atomic force microscopy imaging.

Introduction

Atomic force microscopy (AFM) [1] has enabled innovation in nanoscale engineering since it was invented in 1986 by Binnig and co-workers. Atomic-scale topographical resolution is achieved by sensing the interaction between a sharp microcantilever probe and the sample [2]. Initial operation was in constant-force contact-mode, where a static deflection is maintained through a constant contact force [3].

In dynamic imaging modes [4], the cantilever is driven at, or near, a resonance frequency, which establishes the requirement for demodulation in AFM. In intermittent-contact constant-

amplitude AFM [5], a constant cantilever oscillation amplitude is maintained by feeding back the demodulated fundamental amplitude of the deflection signal. The imaging of delicate biological samples [6-8] is particularly suited to intermittent-contact AFM [9] when tip-sample contact is gentle.

Environmental damping has a large effect on the quality factor (Q) of the cantilever. Values can range from as low as $Q \approx 1$ in liquid [10], up to $Q \approx 10,000$ in ultra-high vacuum [11]. This affects the mechanical bandwidth of the cantilever according to the expression $f_{-3dB} = f_0/2Q$, where f_0 is the fundamental reso-

nance frequency. Assuming all other components in the z -axis feedback loop are also working at high speed [3], a low quality factor can demand a fast demodulator [12].

Multifrequency AFM (MF-AFM) is a major field within dynamic mode AFM. It involves studying multiple frequency components in the cantilever oscillation during tip–sample interactions [13]. Observing higher eigenmodes of the cantilever [14], higher harmonics of the fundamental resonance [15] and intermodulation products [16] have been shown to provide further nanomechanical sample information. These include properties such as sample elasticity, stiffness and adhesiveness [17], which are mapped simultaneously with the topography. Acquiring these observables requires the accurate demodulation of amplitude and phase of multiple frequency components.

Small interaction forces associated with higher-harmonic AFM have been imaged in free air [18] as well as liquid [19]. This has led to relatively large biological objects being imaged including viruses [20] and cells [21]. Multimodal AFM, where two or more resonance frequencies are driven, has theoretical foundations for determining secondary sample properties such as Young’s modulus [13,22]. Applications include the imaging of secondary properties of proteins [23] and polymers [24]. Intermodulation AFM actively drives the cantilever slightly below and above resonance with a two-tone drive. Compared to higher-harmonic AFM, this technique has more enhanced non-linear interactions [25]. Intermodulation products present in the cantilevers motion have been shown to be sensitive to material and chemical contrast [16,26], leading to enhanced nanomechanical insights [27]. Regardless of which MF-AFM technique is performed, the demodulator is an essential component for acquiring observables to characterize the sample.

Previously, the authors conducted an in-depth comparison of conventional and novel demodulation techniques for single-frequency amplitude-modulation atomic force microscopy [28]. It was found that conventional high-speed non-synchronous demodulators are incompatible with MF-AFM, due to the lack of robustness against unwanted frequency components [28]. These include the peak-hold [12], peak detector [29] and RMS-to-DC [30] conversion demodulators. In contrast, synchronous demodulators have been shown to provide accurate estimates in the presence of other frequency components [28]. As a result, MF-AFM experiments usually employ multiple lock-in amplifiers in parallel. However, this introduces an inherent bandwidth limitation as high-frequency mixing products must be low-pass filtered [28,31].

Motivated by improving high-speed MF-AFM demodulation capabilities, a multifrequency Kalman filter was developed

[32]. It outperformed a commercially available lock-in amplifier in terms of both tracking bandwidth and noise performance. However, a major disadvantage of the Kalman filter is its implementation complexity. This heavily limits the achievable sampling rate and ability to track a large number of signals. To alleviate this issue, the Lyapunov filter [33] was established, which is computationally more efficient than the Kalman filter while achieving similar performance [34]. This was extended to a multifrequency Lyapunov filter, which has seen success in higher-harmonic AFM for both amplitude and phase-contrast imaging [35,36]. A limitation, common to both the Kalman and the Lyapunov filter, is a fixed 1st-order response, which has motivated the development of techniques for the direct design of the demodulator frequency response [37,38].

This article aims to provide a rigorous experimental comparison of MF-AFM demodulation techniques. This includes the conventional lock-in amplifier and coherent demodulator, as well as the recently proposed Kalman filter, Lyapunov filter and direct-design method. For a fair comparison, each system is implemented on the same FPGA platform with a common sample rate. The sensitivity to unwanted frequency components for both low and high bandwidths is assessed along with implementation complexity. A final experimental comparison is conducted through higher-harmonic AFM imaging for both low and high tracking bandwidths.

Multifrequency AFM Modulation and Demodulation Fundamentals

Cantilever deflection signal model

A single component of the cantilever deflection signal is modeled as a sine wave with carrier frequency f_i , time-varying amplitude $A_i(t)$ and phase $\phi_i(t)$, that is of the form

$$y_i(t) = A_i(t) \sin(\omega_i t + \phi_i(t)). \quad (1)$$

For a better readability, explicit time-dependencies on the amplitude $A(t)$ and phase $\phi(t)$ are dropped from this point onward. By extension, a deflection signal consisting of multiple frequencies is given by

$$y(t) = \sum_{i=1}^n y_i(t) = \sum_{i=1}^n A_i \sin(\omega_i t + \phi_i), \quad (2)$$

where $i = 1, 2, \dots, n$ denotes the i -th modeled frequency. An alternative representation of a single signal component is of the linearly parameterized form

$$y_i(t) = x_{2i-1} \cos(\omega_i t) + x_{2i} \sin(\omega_i t), \quad (3)$$

where x_{2i-1} and x_{2i} represent quadrature and in-phase components respectively. This is convenient for MF-AFM, as the time-varying amplitude and phase of each frequency can be recovered by the output equations

$$\begin{aligned} A_i &= \sqrt{x_{2i-1}^2 + x_{2i}^2}, \\ \phi_i &= \arctan\left(\frac{x_{2i-1}}{x_{2i}}\right). \end{aligned} \quad (4)$$

Modulation

An amplitude-modulated signal (double-sideband full carrier) is obtained by mixing a modulating signal $y_m(t)$ at a frequency $\omega_m = 2\pi f_m$ with a carrier signal $y_i(t)$. The modulating signal oscillates at a frequency that is significantly slower than the carrier frequency ω_i . Figure 1a illustrates a cantilever driven at multiple frequencies being amplitude-modulated by a sample topography. In MF-AFM, the cantilever deflection signal contains frequency components originating from the fundamental resonance mode, as well as from higher eigenmodes and/or harmonics. If for simplicity we assume unity amplitudes, then amplitude-modulation of a distinct frequency component at ω_i is described by

$$\begin{aligned} y_m(t) \times y_i(t) &= [1 + M \sin(\omega_m t)] \times \sin(\omega_i t + \phi) \\ &= \sin(\omega_i t) + \frac{M}{2} \cos[(\omega_i - \omega_m)t + \phi] \\ &\quad - \frac{M}{2} \cos[(\omega_i + \omega_m)t + \phi]. \end{aligned} \quad (5)$$

Here, M is the modulation index, which for AM signals is the ratio of the peak value of the modulated signal relative to the carrier. Equation 5 shows that the modulation process creates distinct frequencies components at f_i and $f_i \pm f_m$. The latter components are termed the upper and lower sidebands and are centered symmetrically around the carrier frequency as illustrated in Figure 1b. As the modulating frequency increases, the sidebands move away from the carrier up until the limit where the left sideband is at DC and the right sideband is at $2f_i$. The scenario where $f_m > f_i$ is not of practical interest, as the amplitude changes would need to be faster than the cantilever oscillation frequency.

Demodulation

Demodulation is the process of estimating the modulating signal (sample) associated with a carrier frequency. Demodulators can be classified as either synchronous or non-synchronous. Non-synchronous methods do not require a reference oscillator.

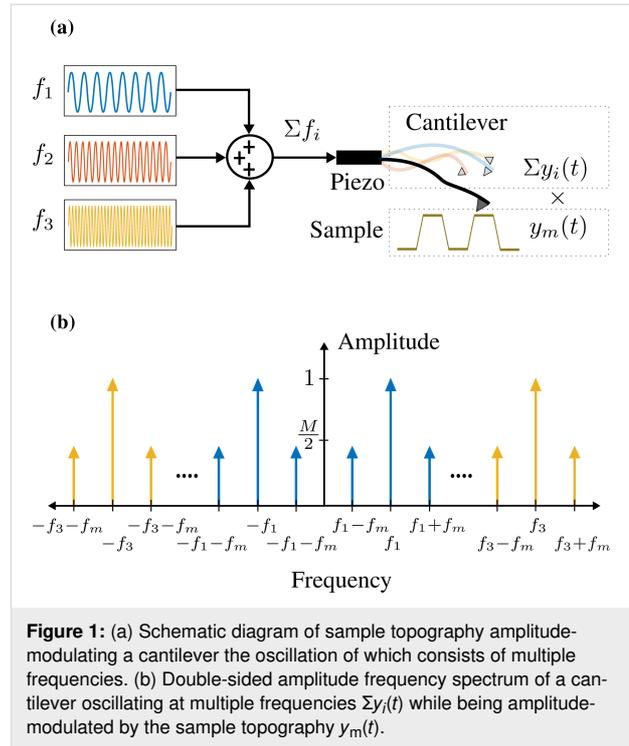


Figure 1: (a) Schematic diagram of sample topography amplitude-modulating a cantilever the oscillation of which consists of multiple frequencies. (b) Double-sided amplitude frequency spectrum of a cantilever oscillating at multiple frequencies $\Sigma y_i(t)$ while being amplitude-modulated by the sample topography $y_m(t)$.

However, these methods are incompatible with MF-AFM, due to their inability to reject unwanted frequency components [28]. For this reason, these techniques are not discussed in this article. Synchronous demodulation techniques employ a reference oscillator and can be categorized as either open-loop or closed-loop, depending on whether they use feedback to estimate parameters. Open-loop demodulators include the lock-in amplifier and coherent demodulator, while closed-loop methods include the Kalman filter, Lyapunov filter, and direct-design demodulator.

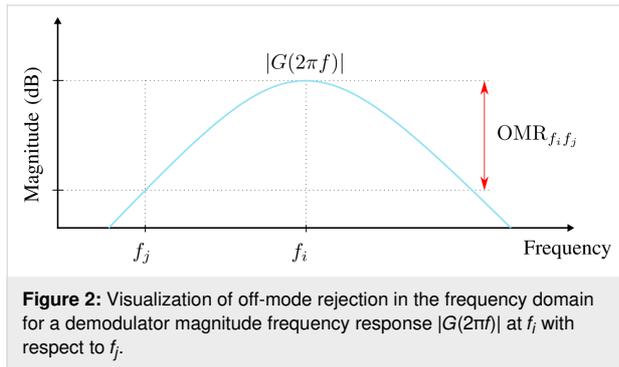
Performance metrics

In a previous work [28], the performance of single-frequency AFM demodulators was assessed by measuring the magnitude of demodulation artifacts and the sensitivity to measurement noise. However, multifrequency AFM applications require an additional metric due to the large number of potentially closely spaced frequencies. For example, higher-harmonic imaging with single-frequency excitation results in small harmonic amplitudes that must be estimated in the presence of both noise and much larger fundamental and/or harmonic components [19,36]. The performance of the demodulator in this regard can be quantified by a metric herein referred to as the off-mode rejection (OMR).

OMR is defined as the gain ratio between a modeled carrier frequency f_i and another frequency f_j as visualized in Figure 2. It can be evaluated by

$$\text{OMR}_{f_i f_j} = \frac{|G(2\pi f_i)|}{|G(2\pi f_j)|}, \quad (6)$$

where $G(2\pi f)$ is the demodulator frequency response. Additionally, implementation complexity is qualitatively discussed. It is assessed according to the maximum achievable sampling rate, timing requirements and computational scalability when modeling additional channels.



Review of Multifrequency Demodulation Methods

Lock-in amplifier

The multifrequency lock-in amplifier (LIA) [28,39-41] operates by multiplying an input signal described by Equation 2 with parallel in-phase and quadrature sinusoids tuned to frequencies the amplitude and phase of which are of interest. For simplicity, consider an ideal input signal consisting of a single sinusoid with a frequency ω_i , applied to a lock-in amplifier tuned to ω_i . During the mixing process, the following intermediate signals are generated

$$\begin{aligned} y_I(t) &= y(t) \times \sin(\omega_i t) \\ &= \frac{1}{2} A_i [\cos(\phi_i) - \cos(2\omega_i t + \phi_i)] \end{aligned} \quad (7)$$

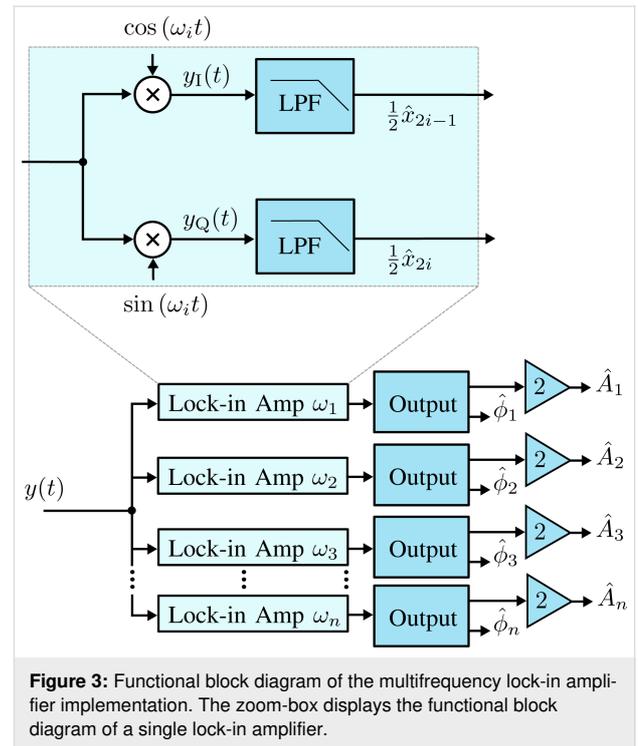
and

$$\begin{aligned} y_Q(t) &= y(t) \times \cos(\omega_i t) \\ &= \frac{1}{2} A_i [\sin(\phi_i) + \sin(2\omega_i t + \phi_i)]. \end{aligned} \quad (8)$$

From Equation 7 and Equation 8 it can be seen that in addition to the desired amplitude and phase, mixing products are generated at twice the carrier frequency $2\omega_i$. If the input contains more than one sinusoid and/or a noise process, further unde-

sired frequency components are present in the intermediate signals. These high-frequency mixing components and noise terms are removed by employing a low-pass filter, the cut-off frequency of which is determined by making a trade-off between tracking bandwidth and $2\omega_i$ ripple suppression [28]. Additionally, lock-in amplifiers should always be AC-coupled as any residual DC offset in the input signal (Equation 2) will generate a mixing component at ω_i .

The functional block diagram of the multifrequency lock-in amplifier is shown in Figure 3. Here, it can be seen that multiple frequencies are tracked by running several lock-in amplifiers in parallel, with each oscillator tuned to a specific frequency ω_i . The required components for digital implementation of each lock-in amplifier are a direct digital synthesizer (DDS) to generate the sine and cosine mixing signals, two multipliers, two low-pass filters and an output block. The output block, which calculates amplitude and phase, is described by Equation 4, meaning the square-root and arctan functions are required. Typically the phase is calculated by using either a polynomial approximation [42] or the CORDIC algorithm [43].



Coherent demodulator

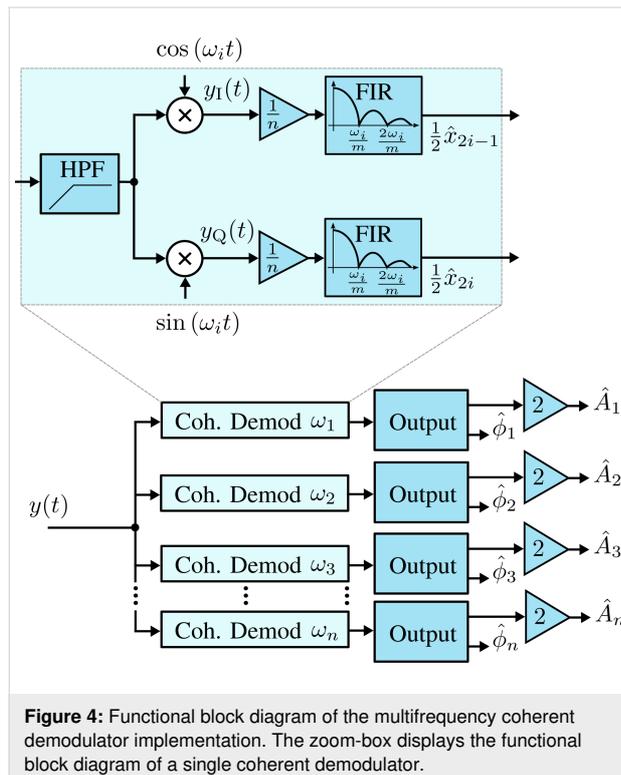
The multifrequency coherent demodulator is a digital demodulation method based on mixing and precise integration over a fixed time window [28,44-47]. Conceptually, it is a digital lock-in amplifier that utilizes mixing with in-phase and quadrature sinusoids

$$\begin{aligned} y_1(t) &= y(t) \times \sin(\omega_i t), \\ y_Q(t) &= y(t) \times \cos(\omega_i t), \end{aligned} \quad (9)$$

and implements low-pass filtering of mixing products and any other additional unwanted frequency components through precise fixed-length numerical integration [45]. If the input signal is a pure sinusoid (Equation 1) and the integration period T is chosen to be an integer multiple of the drive signal period, $T = mT_i$, the integrals over $y_1(t)$ and $y_Q(t)$ evaluate exactly to the in-phase and quadrature states

$$\begin{aligned} \frac{1}{2} x_{2i-1} &= \frac{1}{mT_i} \int_0^{mT_i} y_1(t) dt, \\ \frac{1}{2} x_{2i} &= \frac{1}{mT_i} \int_0^{mT_i} y_Q(t) dt. \end{aligned} \quad (10)$$

The functional block diagram of the multifrequency coherent demodulator is shown in Figure 4. It requires the same components as the lock-in amplifier, although the method in which the low-pass filter is implemented is different. Advanced implementation details can be found in the literature [28,45]. Practitioners should pay strong attention to timing considerations, otherwise the desired low-pass filtering effect will not occur.



For Equation 10 to hold, the integration period must be an integer multiple of the sampling period, $nT_s = mT_i$, where n is the number of samples in the integration. Since an arbitrary sample-to-carrier frequency ratio F_s/f_i is rarely an integer, this condition is hard to meet. Therefore, a practical solution is to find the smallest n such that $nT_s \leq mT_i \leq (n+1)T_s$ and perform a partial integration over the last sampling interval [45]. Such precise control over the integration period is achievable in digital systems, although the implementation of this method is still challenging.

The discrete-time integration in Equation 10 yields a very useful finite impulse response (FIR) filter, the frequency response of which is a sinc(\cdot) function with zeros occurring at integer multiples of the oscillation frequency [28]. Unlike the lock-in amplifier, this allows the coherent demodulator to achieve low-noise output estimates at high tracking bandwidths since it strongly rejects $2f_i$ mixing products [28]. In addition, this zeroing characteristic can provide strong attenuation of unwanted harmonics and intermodulation products. This has led to the multifrequency coherent demodulator being successfully applied to intermodulation AFM [26,27].

Kalman filter

The Kalman filter [48] has seen practical application in many fields including inertial navigation [49], robotics [50], and economics [51]. The Kalman filter uses a recursive algorithm to minimize the error between modeled and measured information to estimate an unknown process variable. Specifically, if the modeling and measurement noise processes have a Gaussian distribution, the Kalman filter produces an optimal estimate of a variable in the least-squares sense by minimizing the variance [52]. Fundamental to its operating principle, the Kalman filter utilizes a linear model of system dynamics and feedback of the state variables to update the Kalman gains, which controls the tracking bandwidth.

When the time-varying system is discretized for $t = kT_s$, where T_s is the sampling period, the process model of the Kalman filter is established as

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{w}_k, \\ y_k &= \mathbf{C}_k \mathbf{x}_k + v_k, \\ E[\mathbf{w}_k \mathbf{w}_k^T] &= \mathbf{Q}_k, \\ E[\mathbf{v}_k \mathbf{v}_k^T] &= R_k, \\ E[\mathbf{w}_k \mathbf{v}_k^T] &= 0, \end{aligned} \quad (11)$$

where w_k and v_k are the process noise and measurement noise with respective covariance matrices \mathbf{Q}_k and R_k . The state and output matrix are described by

$$\begin{aligned} \mathbf{A}_k &= I_{2n+1}, \\ \mathbf{C}_k &= \begin{bmatrix} \cos \theta_{1,k} & \sin \theta_{1,k} & \dots & \cos \theta_{n,k} & \sin \theta_{n,k} \end{bmatrix}, \end{aligned} \quad (12)$$

where I_{2n+1} is the identity matrix of dimension $2n + 1$, n is the number of modeled frequencies, $\theta_{i,k} = \omega_i k T_s$. In this representation, quadrature $x_{2i-1,k}$ and in-phase $x_{2i,k}$ states are assumed to be random variables describing the states of Equation 3.

The parameters \mathbf{Q}_k and R_k dictate the amount of uncertainty in the model and the measurement noise, respectively. To simplify tuning of the filter during operation it is recommended to fix R_k such that it reflects the standard deviation σ of the Gaussian noise in the input signal from the sensor $y(t)$ ($R = \sigma^2$). This leaves \mathbf{Q}_k as the only tuning variable that directly influences the Kalman gains and sets the tracking bandwidth.

The functional block diagram of the Kalman filter implementation is shown in Figure 5, it follows the standard recursive algorithm equations [53,54]. The prediction step is computed as

$$\begin{aligned} \hat{\mathbf{x}}_{k|k-1} &= \mathbf{A}_k \hat{\mathbf{x}}_{k-1|k-1}, \\ \mathbf{P}_{k|k-1} &= \mathbf{A}_k \mathbf{P}_{k-1|k-1} \mathbf{A}_k^T + \mathbf{Q}_k, \end{aligned} \quad (13)$$

where $\mathbf{P}_{k|k-1}$ denotes the predicted covariance matrix. This is followed by the Kalman gain and state measurement updates

$$\begin{aligned} \mathbf{K}_k &= \mathbf{P}_{k|k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + R_k)^{-1} \\ \hat{\mathbf{x}}_{k|k} &= \hat{\mathbf{x}}_{k|k-1} + \mathbf{K}_k (y_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}). \end{aligned} \quad (14)$$

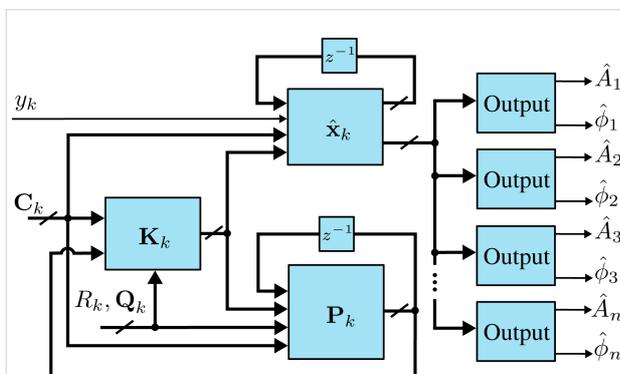


Figure 5: Functional block diagram of the Kalman filter implementation.

Lastly, the covariance matrix is updated according to

$$\mathbf{P}_{k|k} = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T, \quad (15)$$

which is in Joseph form, i.e., it is naturally symmetric and positive definite. These properties can be exploited in the implementation to reduce memory and computation requirements. In addition, it is the most numerically stable form of the covariance matrix and remains convergent and non-deterministic for any selection of \mathbf{Q}_k and R_k [53].

Specific amplitude and phase of a modeled frequency ω_i are shown to be recovered by employing the output equations in Equation 4. Although this method is simple to tune in real-time, a disadvantage is the fixed 1st-order response. Also, the Kalman filter equations have a complexity of $\mathcal{O}(n^3)$ for n modeled frequencies resulting in significant computational requirements beyond three modeled frequencies. This system representation has seen success in tracking power system voltage phasors [55] and more recently high-speed AFM [31,32].

Lyapunov filter

The Lyapunov filter [33,35,36] also uses a model-based feedback approach to obtain amplitude and phase of signals at desired frequencies. Under certain conditions, the Lyapunov filter has been shown to be equivalent to the Kalman filter [33]. However, the Lyapunov filter uses a tunable scalar gain γ instead of updating covariance matrix and Kalman gain equations. This gives the Lyapunov filter a computational complexity of $\mathcal{O}(n)$ as additional frequencies are modeled, a significant improvement over the Kalman filter.

A key property of the Lyapunov filter is exponential convergence of the estimated states [56], with the tunable loop gain γ governing the speed of convergence. The multifrequency Lyapunov filter is implemented as parallel linear observers tuned to a particular frequency ω_i , as depicted in Figure 6. An error signal is generated by feeding back an estimate of the input signal as per Equation 2, obtained from the parameterized states of each individual filter. Regulation of this error through feedback leads to the much desired suppression of the high-frequency mixing components.

The update law for the Lyapunov filter [33,36] for multiple frequencies is written as

$$\begin{aligned} \dot{\hat{\mathbf{x}}} &= \mathbf{\Gamma} \mathbf{C}^T [y(t) - \hat{y}(t)], \\ \hat{y}(t) &= \mathbf{C} \hat{\mathbf{x}}, \end{aligned} \quad (16)$$

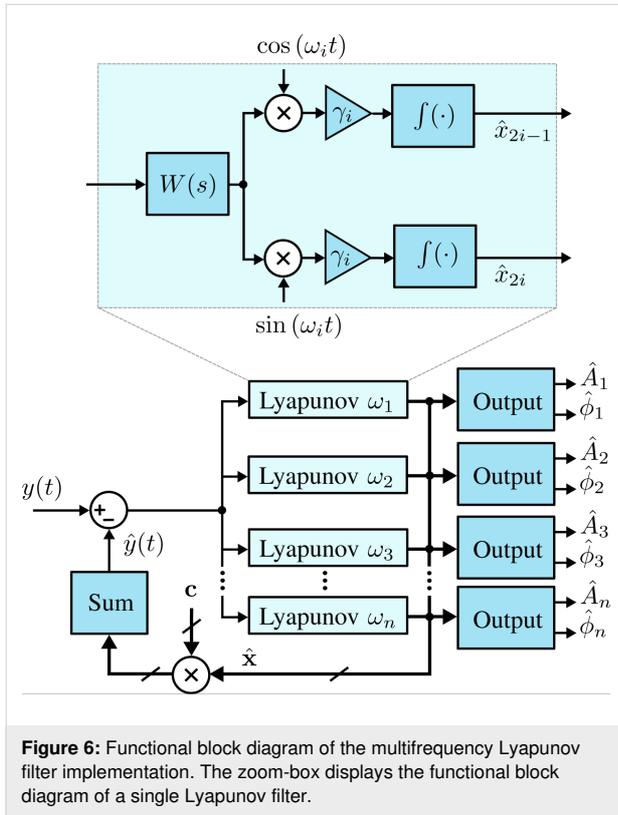


Figure 6: Functional block diagram of the multifrequency Lyapunov filter implementation. The zoom-box displays the functional block diagram of a single Lyapunov filter.

where

$$\mathbf{\Gamma} = \begin{bmatrix} \gamma_1 & & & & & \\ & \gamma_1 & & & & 0 \\ & & \ddots & & & \\ & & & \gamma_n & & \\ & & & & & \gamma_n \end{bmatrix} \quad (17)$$

and

$$\mathbf{C} = [\cos(\omega_1 t) \sin(\omega_1 t) \dots \cos(\omega_n t) \sin(\omega_n t)]. \quad (18)$$

In this form, $\hat{y}(t)$ represents the estimated input signal and the amplitude A_i and phase ϕ_i estimates are found by applying Equation 4 to each quadrature and in-phase pair of $\hat{\mathbf{x}}$. A key property to ensure exponential convergence of $\hat{\mathbf{x}}$ to \mathbf{x} is to guarantee that \mathbf{C} is persistently excited [56]. Convergence is shown for the single-frequency filter in [33] and can easily be extended for the multifrequency case. Furthermore, exponential convergence of $\hat{\mathbf{x}}$ means that \hat{A} and $\hat{\phi}$ also converge. This system representation has been shown to perform similarly to the Kalman filter [28], which is advantageous given its imple-

mentation simplicity. Recently, it has been used for higher-harmonic AFM for both amplitude and phase-contrast imaging [35,36].

Direct-design method

The direct-design method [37] also utilizes model-based feedback to obtain the amplitude and phase of signals at desired frequencies. However, intrinsic to its design methodology is the ability to implement an arbitrary filter response with a specified filter order and linearity in the bandpass region. For example, a demodulator can be implemented the frequency response of which resembles a Butterworth or Chebyshev filter with a desired filter bandwidth and order. This alleviates the limited 1st-order response of the Kalman and Lyapunov filters, creating stronger rejection of unwanted frequency components. This occurs while maintaining benefits such as low noise and low computational complexity.

In order to obtain an arbitrary demodulator response, consider the functional block diagram in Figure 7, where the integrator of the Lyapunov filter is replaced by the transfer function $F(s)$. In this form, the direct-design demodulator follows a modulated–demodulated control loop [57] with a unity plant. This method differs from the Lyapunov filter as it does not set the pre-filter to $W(s) = 1$, instead it utilizes $W(s)$ as part of the design of a desired closed-loop response. In the original work

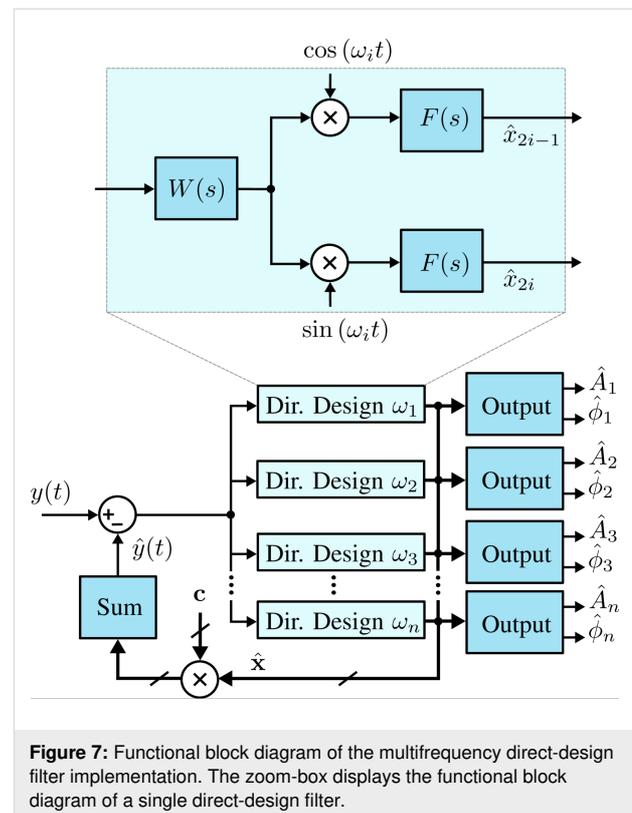


Figure 7: Functional block diagram of the multifrequency direct-design filter implementation. The zoom-box displays the functional block diagram of a single direct-design filter.

[37] on the direct-design method, a useful and relatively simple design methodology is detailed. Firstly, $F(s)$ is set as

$$F(s) = \frac{Y}{s}, \quad (19)$$

meaning the closed-loop equivalent transfer function is

$$\tilde{F}(s) = \frac{B(s)}{A(s)} = \frac{1}{2} [F(s + j\omega) + F(s - j\omega)]. \quad (20)$$

The pre-filter $W(s) = P(s)/L(s)$ is then found according to the pole-assignment equation

$$A_d(s) = L(s)A(s) + P(s)B(s) \quad (21)$$

to achieve a specified closed-loop response.

For example, to implement a 2nd-order bandpass Butterworth prototype, the closed-loop transfer function is

$$G_d(s) = \frac{B_d(s)}{A_d(s)} = \frac{b_2 s^2}{s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0}. \quad (22)$$

Here, b_2 and A_i are the filter coefficients the values of which are calculated based on the chosen filter order and bandwidth around the modeled frequency ω_i . As the desired closed-loop polynomial has five coefficients, the coefficients of the pre-filter $W(s)$ are of the form

$$W(s) = \frac{P(s)}{L(s)} = \frac{p_1 s + p_0}{l_2 s^2 + l_1 s + l_0}, \quad (23)$$

and are able to be obtained by solving Equation 21.

The existing literature on direct-design demodulation techniques [37,38] is concerned with single-frequency applications.

However, this article demonstrates the performance advantages that can also be achieved in multifrequency applications.

Summary

Table 1 compares the multifrequency demodulation techniques discussed in this section. Two distinct categories of synchronous demodulators can be seen; those that employ low-pass filtering of mixing products in open-loop configurations and those that use closed-loop model-based feedback to regulate the error. As shown in a previous work [28], the closed-loop methods are able to maintain very high tracking bandwidths, achieving single-cycle convergence ($f_{-3dB} \approx f_i$) with optimal noise performance.

Results and Discussion

Experimental setup

The multifrequency demodulation techniques detailed in the previous section were implemented on a Xilinx Kintex-7 KC705 evaluation board (model: XC7K325T) paired with a DC-coupled high-speed 4DSP input/output (I/O) card (model: FMC151). The FPGA clock is synchronized with the high-speed I/O card at 240 MHz. The I/O card has a two-channel 14-bit analog-to-digital converter (ADC) and a two-channel 16-bit digital-to-analog converter (DAC), which sample at 250 MHz and 800 MHz, respectively. All demodulation methods were run at a nominal sampling frequency of $F_s = 1.5$ MHz.

Implementation

Because of the high complexity, the sampling rate off the Kalman filter implementation was set to $F_s = 1.5$ MHz, which was the maximum achievable for three modeled frequencies employing floating point precision to ensure covariance matrix stability and a computationally optimized implementation [32].

The Lyapunov filter and direct-design method achieve sampling rates of $F_s = 7$ MHz for three modeled frequencies. This is due to the reduced complexity compared to the Kalman filter, floating point precision was also used to implement these methods.

Table 1: Summary of multifrequency demodulation estimation methods.

method	tuning	configuration	order	references
lock-in amplifier	low-pass filter	open-loop	specified n	[28,39-41]
coherent demodulator	number of hold cycles	open-loop	specified n	[28,44-47,58]
Kalman filter	\mathbf{Q}_k	closed-loop	1st	[28,31,32,48,53]
Lyapunov filter	$\mathbf{\Gamma}$	closed-loop	1st	[28,33,35,36,59]
direct-design method	desired poles	closed-loop	specified n	[37,38]

The open-loop methods include the lock-in amplifier and coherent demodulator, which are able to achieve $F_s = 120$ MHz for three modeled frequencies. In contrast to the closed-loop methods, the open-loop methods are compatible with pipelined fixed-point implementation, which results in significantly increased maximum sampling rates and reduced FPGA resource usage. A large number of modeled frequencies are possible.

In addition to processing requirements, the implementation complexity may also be increased by timing requirements. For example, the fixed-length numerical integration of the coherent demodulator results in $\text{sinc}(\cdot)$ frequency responses the zeros of which are related to F_s/f_i . This results in a limited number of possible high-bandwidth configurations. At low bandwidths, there is much more flexibility in achieving a desired bandwidth as the $(N + 1)$ -FIR filter is longer. Here, the group delay $(N/2)$ introduced should be considered with respect to the phase margin of the z -axis feedback loop.

Off-mode rejection

Each multifrequency demodulator was assessed by applying a single-tone sine sweep of the carrier frequency ω_i on an input signal described by Equation 1. For each demodulator, an amplitude magnitude frequency response of all three channels was recorded as the input carrier frequency was swept from DC to 750 kHz with a constant amplitude A_i . The three channels were configured to model carrier frequencies of 50, 150, and 300 kHz. In this experiment, the noise floor is dictated by a residual DC offset, which is present due to the finite resolution of the DAC. However, as each demodulation technique was analyzed by using the same hardware, the relative OMR differences are a good indication of each methods performance.

The open-loop methods have the benefit of being easily configured to a desired filter order. For this experiment they are of 4th order, which for low-bandwidth settings in Figure 8 creates very steep roll-offs for the lock-in amplifier (Figure 8a,b) and coherent demodulator (Figure 8e,f). When compared to the

fixed 1st-order Kalman filter (Figure 8i,j) and Lyapunov filter (Figure 8m,n), the open-loop methods achieve stronger attenuation around the modeled carrier frequency.

The difference between the lock-in amplifier and coherent demodulator is the method used to employ the low-pass filter for suppressing mixing products. In this experiment, the lock-in amplifier utilizes a Butterworth filter, which generates a maximally flat frequency response around the modeled carrier frequency. Conversely, the coherent demodulator employs fixed-length numerical integration resulting in a $\text{sinc}(\cdot)$ envelope in its frequency response [28]. This leads to strong OMR at regular intervals at $\text{sinc}(\cdot)$ zero locations. However, there is less rejection in-between zeros compared to the Butterworth response.

In contrast to the open-loop methods, the Kalman and Lyapunov filters operate in a closed-loop configuration resulting in state cross-coupling during feedback. As seen in Figure 8 for the Kalman filter (Figure 8i–l) and the Lyapunov filter (Figure 8m–p), this leads to each channel zeroing frequency components corresponding to the other modeled channels. The direct-design method alleviates the fixed 1st-order frequency response of the Kalman and Lyapunov filters. In Figure 8, the direct-design method (Figure 8q–t) performance is shown when configured to a 2nd-order Butterworth filter. The higher filter order results in greater suppression of broadband noise and other frequency components around the modeled carrier frequency when compared to the other closed-loop methods.

Table 2 examines the channel-to-channel OMR performance of each multifrequency demodulator for the 300 kHz channel. It is clear that the open-loop demodulators have a significant performance decrease as the tracking bandwidth increases. The poor OMR is caused by insufficient roll-off of each frequency response with respect to the other modeled frequencies f_1 and f_2 . This occurs despite the coherent demodulator rejecting its own $2f_i$ mixing products. In contrast, the closed-loop Kalman filter, Lyapunov filter and direct-design methods benefit from cross-coupling zeros across all bandwidths allowing them to maintain

Table 2: Channel-to-channel off-mode rejection for the $f_3 = 300$ kHz channel.

method	Bandwidth $0.1\%f_i$		Bandwidth $1\%f_i$		Bandwidth $10\%f_i$		Bandwidth $50\%f_i$	
	OMR $_{f_3f_1}$	OMR $_{f_3f_2}$	OMR $_{f_3f_1}$	OMR $_{f_3f_2}$	OMR $_{f_3f_1}$	OMR $_{f_3f_2}$	OMR $_{f_3f_1}$	OMR $_{f_3f_2}$
	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)	(dB)
lock-in amplifier	-43.8	-44.0	-42.2	-42.8	-39.5	-29.0	-14.2	-3.1
coherent demodulator	-45.6	-46.3	-40.5	-40.0	-36.4	-27.8	-14.3	-4.6
Kalman filter	-44.0	-44.1	-53.0	-52.1	-47.8	-49.7	-42.6	-47.0
Lyapunov filter	-41.7	-40.5	-52.8	-53.1	-44.8	-46.2	-46.1	-47.4
direct-design method	-41.6	-46.7	-42.3	-43.5	-47.2	-46.0	-51.7	-52.3

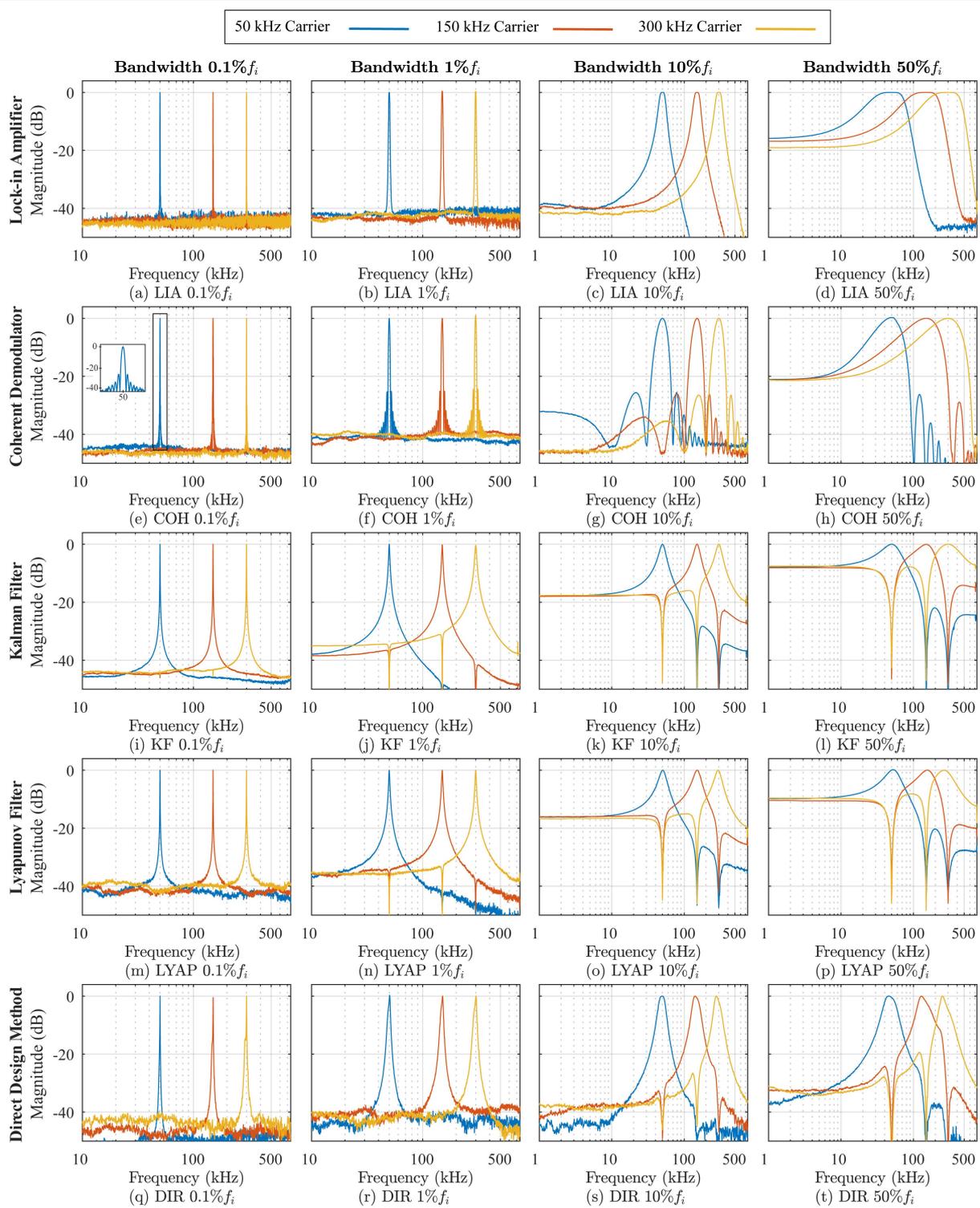


Figure 8: Experimental off-mode rejection results. Here each multifrequency demodulator is on a single row and the tracking bandwidths are adjusted per column with settings of 0.1%, 1%, 10% and 50%. For each system the three modeled carrier frequencies are $f_1 = 50$ kHz (blue), $f_2 = 150$ kHz (red) and $f_3 = 300$ kHz (yellow).

a strong OMR. The ability to precisely resolve the zeros is limited by the DAC resolution. However, the performance distinction between open-loop and closed-loop methods is clear.

Time-domain estimation analysis

Amplitude estimation performance of the three-channel multi-frequency demodulators was investigated when a three-tone

sinewave was applied as an input signal described by Equation 2. Here, $A_1 = 500$ mV, $A_2 = 100$ mV, $A_3 = 50$ mV, $f_1 = 50$ kHz, $f_2 = 150$ kHz, and $f_3 = 300$ kHz. Figure 9 shows the amplitude estimation error ($\hat{A}_i / A_i \times 100\%$) in the time-domain and amplitude estimate power spectral density (PSD) for both low ($1\%f_i$) and high ($50\%f_i$) tracking bandwidth settings.

When the three-tone sinewave is applied, the performance of each demodulator at low bandwidths is shown to be similar. Each channel is able to estimate the amplitude of its modeled frequency, while strongly attenuating the other frequency components present in the input signal.

At high bandwidths, closed-loop demodulators benefit from cross-coupling zeros at the modeled frequencies. Compared to the open-loop methods, this results in significantly less estimation error as seen in Figure 9. In Figure 9a, the lock-in amplifier 50 kHz estimate contains mixing products at $2f_i = 100$ kHz in the time-domain, shown as distinct peaks in the PSD. In contrast, the coherent demodulator in Figure 9b strongly attenuates the mixing products. The performance difference is due to the Butterworth filter not sufficiently attenuating the mixing products, while the sinc(-) envelope contains a zero at $2f_i$. However, both open-loop methods poorly estimate the 150 kHz and 300 kHz input signals at high tracking bandwidths due to weak OMR. The PSD reveals that the large estimation errors consist of intermodulation products, which arise from the input multiplying stage.

AFM imaging

The lock-in amplifier and Lyapunov filter were compared through an MF-AFM imaging experiment where they estimate a signal in the presence of undesirable frequency components. These demodulators were chosen as they are the most simple methods to implement in their respective configurations. This experiment further investigated open-loop and closed-loop demodulator performance at low and high tracking bandwidths.

When compared to bimodal AFM, higher-harmonic AFM [15,60] has inherently greater demodulation challenges. Strong OMR is required as higher harmonics are separated by nf_0 , much closer than the approx. $6f_0$ second eigenmode spacing [61]. In addition, harmonic content from tip-sample interactions scales with approx. $1/n^2$ [13]. Therefore, the signals of interest are detected in the presence of a much larger fundamental resonance frequency, emphasizing the need for strong noise sensitivity from the demodulator.

Higher-harmonic AFM imaging was performed using an NT-MDT NTEGRA AFM on the second harmonic amplitude. The chosen cantilever (Budget Sensor TAP190G) has a funda-

mental resonance frequency of 156.75 kHz. The sample is a blend of polystyrene (PS) and polyolefin elastomer (LDPE) available from Bruker (PS-LDPE-12M). Due to the different elastic moduli of the PS and LPDE regions, the sample is used for evaluating imaging methods that are sensitive to elasticity.

Higher-harmonic amplitude images were obtained by the lock-in amplifier and Lyapunov filter on the second harmonic. Each demodulator was configured to track 313.50 kHz. In addition, the Lyapunov filter contained a channel modeling the fundamental resonance frequency. Although the cantilever is actively driven at its fundamental resonance frequency, during imaging its deflection signal contains additional frequency components. These include higher harmonics and intermodulation products excited by non-linear tip-sample forces during contact.

Second-harmonic amplitude images captured by both demodulators at low (1 kHz) and high (60 kHz) tracking bandwidths are shown in Figure 10. At low bandwidths, the lock-in amplifier (Figure 10e) and the Lyapunov filter (Figure 10c) perform comparably as demonstrated in a previous work [36]. However, at high bandwidths the lock-in amplifier image has large artifacts when compared to the Lyapunov filter. This is due to the different OMR achieved by each system with respect to the fundamental resonance frequency. Through channel cross-coupling, the Lyapunov filter is guaranteed to contain a zero at the desired location of 156.75 kHz.

In contrast, the lock-in amplifier insufficiently attenuates the fundamental resonance frequency. The PSD of the raw data from a single scan line (Figure 10g), taken from the image in Figure 10f, reveals that the estimate contains large intermodulation products. These signal components are aliased due to the low AFM sampling frequency ($F_s = 256$ Hz), resulting in the low-frequency artifacts seen in Figure 10f.

Conclusion

This article compares the performance of traditional and recently proposed demodulators for MF-AFM. These include conventional open-loop methods such as the lock-in amplifier and coherent demodulator, and closed-loop methods such as the Kalman filter, the Lyapunov filter and the direct-design method. The sensitivity of each demodulator to unwanted frequency components was assessed for low and high tracking bandwidths. Additionally, higher-harmonic AFM imaging was conducted for both low and high tracking bandwidths to further compare demodulator performance.

Open-loop demodulation schemes attenuate the high-frequency mixing component at $2f_i$ by employing a low-pass filter. The lock-in amplifier provides flexibility to implement a desired

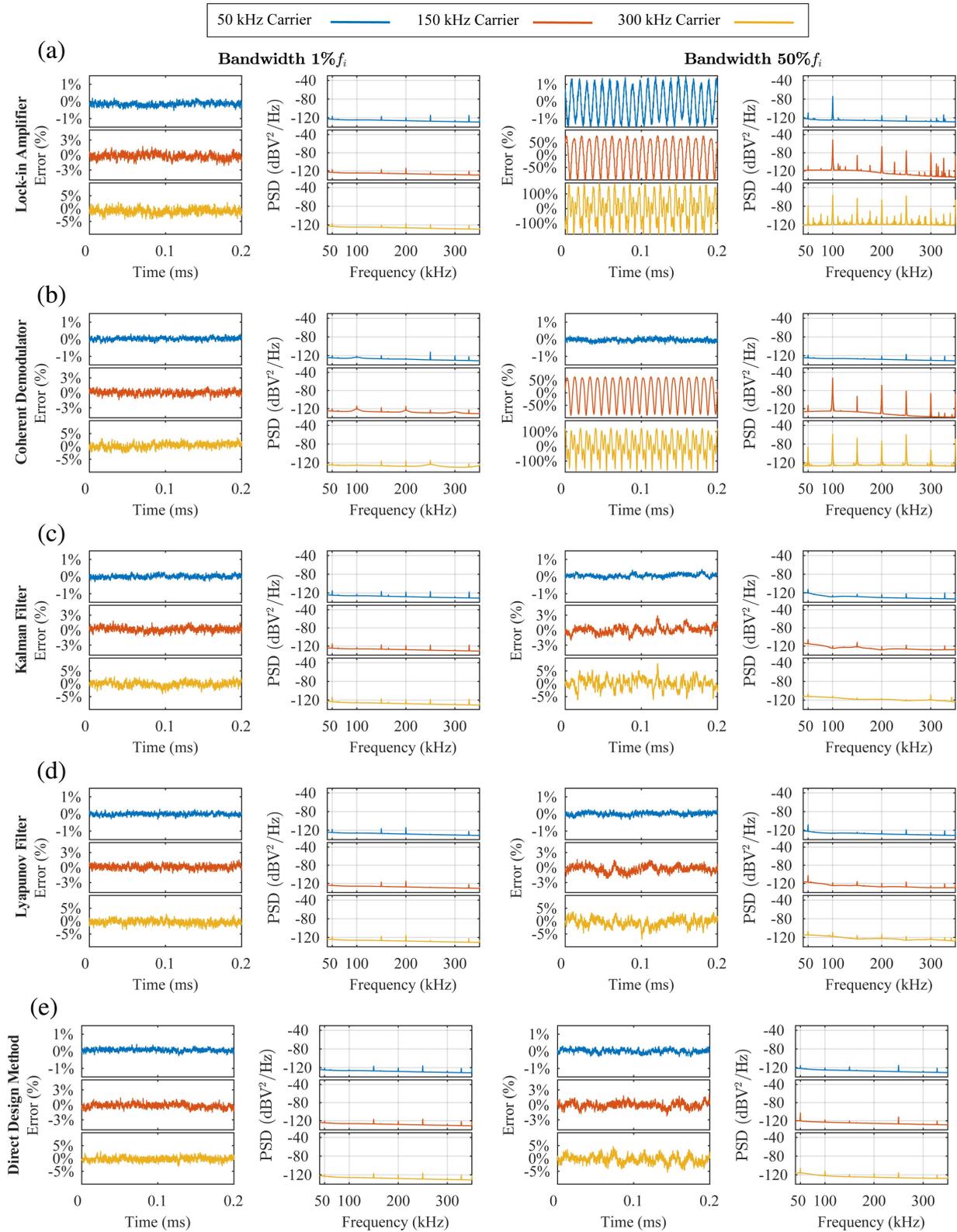


Figure 9: Experimental amplitude estimation error and power spectral density of amplitude estimation for the (a) lock-in amplifier, (b) coherent demodulator, (c) Kalman filter, (d) Lyapunov filter and (e) direct-design method for low (1% f_i) and high (50% f_i) tracking bandwidths. The input signal is described by Equation 2, where $A_1 = 500$ mV, $A_2 = 100$ mV, $A_3 = 50$ mV and $f_1 = 50$ kHz, $f_2 = 150$ kHz and $f_3 = 300$ kHz.

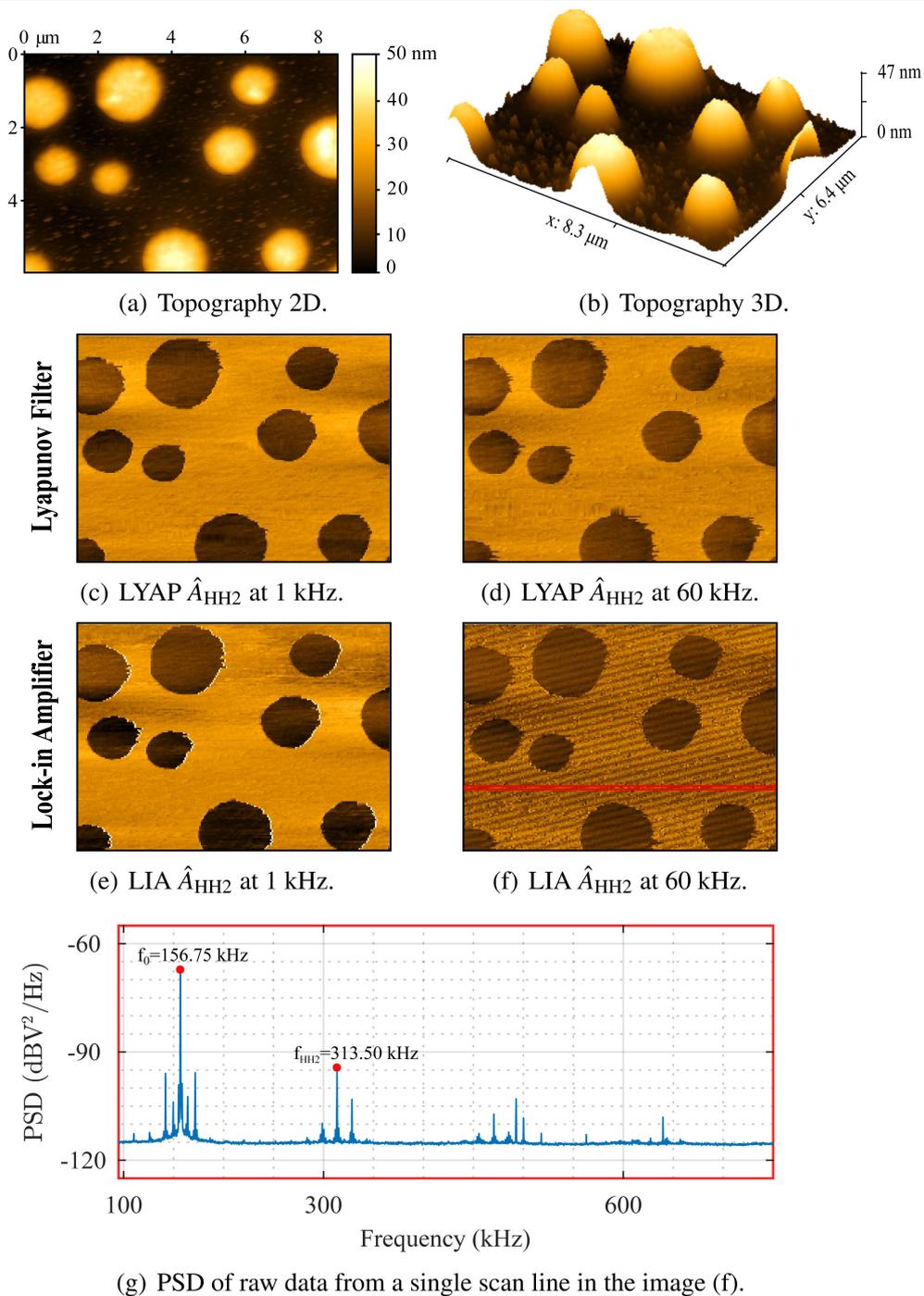


Figure 10: Higher-harmonic amplitude AFM imaging performed with the fundamental mode of a TAP190G cantilever on a PS/LPDE polymer blend. Images shown are the (a,b) topography in nanometers at 3 kHz, with parallel second-harmonic amplitude estimates from the (c,d) lock-in amplifier and (e,f) Lyapunov filter at tracking bandwidths of 1 kHz and 60 kHz. The PSD (g) is shown for the raw data of a single scan line from the image in (f).

filter response and order. Conversely, the coherent demodulator contains a sinc(\cdot) envelope as it performs numerical integration over a fixed-length time window. Both demodulators excel at low bandwidths due to steep roll-offs, while having poor OMR at high tracking bandwidths. Although the lock-in amplifier implementation is simpler, the coherent demodulator

sinc(\cdot) lobes are advantageous for higher harmonic and inter-modulation AFM.

The closed-loop Kalman filter, Lyapunov filter and direct-design method employ internal feedback of the estimated states to reject the mixing products. This allows them to maximize the

Table 3: Summary of MF-AFM applications and required demodulator properties. Note: f_0 is the fundamental resonance frequency and f_i is the demodulated frequency.

MF-AFM application	application characteristics		demodulator properties	
	frequency spacing	number of channels	tracking bandwidth	off-mode rejection
intermodulation	very small ($f_i/f_0 \ll 1$)	40+	very low ($\ll 1\%f_0$)	very important
higher harmonic	medium ($f_i/f_0 \approx 1$)	10+	low ($1\%f_0$)	important
higher mode	large ($f_i/f_0 \gg 1$)	1–5	user choice ($\geq 1\%f_0$)	less important

tracking bandwidth without introducing additional noise in the amplitude estimate [28]. An added benefit of this approach is cross-coupling zeros occurring at modeled frequencies, which was demonstrated to reduce estimation artifacts. The direct-design method alleviates the limited 1st-order response of the Kalman and Lyapunov filters. When configured to a 2nd-order Butterworth response, it achieved an increased roll-off which increases broadband noise suppression while still maintaining strong OMR performance.

Table 3 is provided as a reference of MF-AFM application characteristics and required demodulator properties. A recommendation for which demodulator is most suited to three major MF-AFM applications is given as follows:

Intermodulation AFM: This MF-AFM application tracks a large number of closely spaced intermodulation products [16]. As each signal of interest has a frequency separation of the order of 100 Hz, a very low bandwidth and a very strong OMR are essential. The requirement to track up to 40+ signals is most suited to a computationally inexpensive open-loop method. The coherent demodulator is recommended for intermodulation AFM, since the sinc(\cdot) response of each channel can be configured to zero other intermodulation products [58].

Higher-harmonic AFM: This MF-AFM application tracks integer multiples of the cantilever fundamental resonance frequency, resulting in frequency spacing of the order of the fundamental resonance frequency, which ranges between 100 and 300 kHz. Since each harmonic is in the presence of a much larger fundamental resonance frequency, a low tracking bandwidth and a strong OMR is required. The open-loop lock-in amplifier and coherent demodulator are recommended at low tracking bandwidths. While the closed-loop Lyapunov filter and direct-design method are recommended if a higher tracking bandwidth is desired. This is because the closed-loop methods have the added benefit of zeroing the large fundamental resonance frequency and other harmonics.

Higher-mode AFM: This MF-AFM application tracks the fundamental resonance frequency and higher resonance modes. Frequency content of interest is typically separated by 500 kHz or more, depending on the cantilever geometry. This provides flexibility to the user to operate at either a low or high tracking bandwidth. At low tracking bandwidths, the lock-in amplifier is recommended, as it is of lower complexity than the coherent demodulator, which offers little benefit for widely spaced signals. At high tracking bandwidths, a closed-loop method is recommended as they achieve single-cycle convergence ($f_{-3dB} \approx f_i$) with optimal noise performance. Also, each channel has the added benefit of zeroing other resonant modes. The Lyapunov filter and direct-design method are preferred over the Kalman filter, as they are significantly easier to implement.

ORCID® iDs

David M. Harcombe - <https://orcid.org/0000-0002-9866-5806>

Michael G. Ruppert - <https://orcid.org/0000-0003-2286-4929>

Andrew J. Fleming - <https://orcid.org/0000-0001-7132-8458>

References

- Binnig, G.; Quate, C. F.; Gerber, C. *Phys. Rev. Lett.* **1986**, *56*, 930–933. doi:10.1103/physrevlett.56.930
- Rabe, U.; Janser, K.; Arnold, W. *Rev. Sci. Instrum.* **1996**, *67*, 3281–3293. doi:10.1063/1.1147409
- Abramovitch, D. Y.; Andersson, S. B.; Pao, L. Y.; Schitter, G. A tutorial on the mechanisms, dynamics, and control of atomic force microscopes. In *2007 American Control Conference*, 2007; pp 3488–3502.
- García, R.; Perez, R. *Surf. Sci. Rep.* **2002**, *47*, 197–301. doi:10.1016/s0167-5729(02)00077-8
- García, R. *Amplitude Modulation Atomic Force Microscopy*; Wiley-VCH Verlag GmbH: Weinheim, Germany, 2010. doi:10.1002/9783527632183
- Möller, C.; Allen, M.; Elings, V.; Engel, A.; Müller, D. *J. Biophys. J.* **1999**, *77*, 1150–1158. doi:10.1016/s0006-3495(99)76966-3
- Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. *Nature* **2010**, *468*, 72–76. doi:10.1038/nature09450
- Chiaruttini, N.; Redondo-Morata, L.; Colom, A.; Humbert, F.; Lenz, M.; Scheuring, S.; Roux, A. *Cell* **2015**, *163*, 866–879. doi:10.1016/j.cell.2015.10.017

9. Zhong, Q.; Inniss, D.; Kjoller, K.; Elings, V. B. *Surf. Sci.* **1993**, *290*, L688–L692. doi:10.1016/0039-6028(93)90582-5
10. Tamayo, J.; Humphris, A. D. L.; Miles, M. J. *Appl. Phys. Lett.* **2000**, *77*, 582–584. doi:10.1063/1.127051
11. Lübke, J.; Tröger, L.; Torbrügge, S.; Bechstein, R.; Richter, C.; Kühnle, A.; Reichling, M. *Meas. Sci. Technol.* **2010**, *21*, 125501. doi:10.1088/0957-0233/21/12/125501
12. Ando, T.; Kodera, N.; Takai, E.; Maruyama, D.; Saito, K.; Toda, A. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 12468–12472. doi:10.1073/pnas.211400898
13. García, R.; Herruzo, E. T. *Nat. Nanotechnol.* **2012**, *7*, 217–226. doi:10.1038/nnano.2012.38
14. Martínez, N. F.; Lozano, J. R.; Herruzo, E. T.; Garcia, F.; Richter, C.; Sulzbach, T.; Garcia, R. *Nanotechnology* **2008**, *19*, 384011. doi:10.1088/0957-4484/19/38/384011
15. Stark, R. W.; Heckl, W. M. *Rev. Sci. Instrum.* **2003**, *74*, 5111–5114. doi:10.1063/1.1626008
16. Platz, D.; Tholén, E. A.; Pesen, D.; Haviland, D. B. *Appl. Phys. Lett.* **2008**, *92*, 153106. doi:10.1063/1.2909569
17. García, R.; Proksch, R. *Eur. Polym. J.* **2013**, *49*, 1897–1906. doi:10.1016/j.eurpolymj.2013.03.037
18. Hillenbrand, R.; Stark, M.; Guckenberger, R. *Appl. Phys. Lett.* **2000**, *76*, 3478–3480. doi:10.1063/1.126683
19. Raman, A.; Trigueros, S.; Cartagena, A.; Stevenson, A. P. Z.; Susilo, M.; Nauman, E.; Contera, S. A. *Nat. Nanotechnol.* **2011**, *6*, 809–814. doi:10.1038/nnano.2011.186
20. Preiner, J.; Tang, J.; Pastushenko, V.; Hinterdorfer, P. *Phys. Rev. Lett.* **2007**, *99*, 046102. doi:10.1103/physrevlett.99.046102
21. Cartagena-Rivera, A. X.; Wang, W.-H.; Geahlen, R. L.; Raman, A. *Sci. Rep.* **2015**, *5*, 11692. doi:10.1038/srep11692
22. Forchheimer, D.; Borysov, S. S.; Platz, D.; Haviland, D. B. *Nanotechnology* **2014**, *25*, 485708. doi:10.1088/0957-4484/25/48/485708
23. Herruzo, E. T.; Perrino, A. P.; Garcia, R. *Nat. Commun.* **2014**, *5*, 3126. doi:10.1038/ncomms4126
24. Benaglia, S.; Amo, C. A.; Garcia, R. *Nanoscale* **2019**, *11*, 15289–15297. doi:10.1039/c9nr04396a
25. Platz, D.; Tholén, E. A.; Hutter, C.; von Bieren, A. C.; Haviland, D. B. *Ultramicroscopy* **2010**, *110*, 573–577. doi:10.1016/j.ultramic.2010.02.012
26. Forchheimer, D.; Forchheimer, R.; Haviland, D. B. *Nat. Commun.* **2015**, *6*, 6270. doi:10.1038/ncomms7270
27. Thorén, P.-A.; de Wijn, A. S.; Borgani, R.; Forchheimer, D.; Haviland, D. B. *Nat. Commun.* **2016**, *7*, 13836. doi:10.1038/ncomms13836
28. Ruppert, M. G.; Harcombe, D. M.; Ragazzon, M. R. P.; Moheimani, S. O. R.; Fleming, A. J. *Beilstein J. Nanotechnol.* **2017**, *8*, 1407–1426. doi:10.3762/bjnano.8.142
29. Ando, T. *Nanotechnology* **2012**, *23*, 062001. doi:10.1088/0957-4484/23/6/062001
30. Kitchin, C.; Counts, L. *RMS to DC Conversion Application Guide*, 1st ed.; Analog Devices, 1986.
31. Ruppert, M. G.; Karvinen, K. S.; Wiggins, S. L.; Reza Moheimani, S. O. *IEEE Trans. Control Syst. Technol.* **2016**, *24*, 276–284. doi:10.1109/tcst.2015.2435654
32. Ruppert, M. G.; Harcombe, D. M.; Moheimani, S. O. R. *IEEE/ASME Trans. Mechatronics* **2016**, *21*, 2705–2715. doi:10.1109/tmech.2016.2574640
33. Ragazzon, M. R. P.; Ruppert, M. G.; Harcombe, D. M.; Fleming, A. J.; Gravdahl, J. T. *IEEE Trans. Control Syst. Technol.* **2018**, *26*, 765–772. doi:10.1109/tcst.2017.2692721
34. Ruppert, M. G.; Harcombe, D. M.; Ragazzon, M. R. P.; Moheimani, S. O. R.; Fleming, A. J. Frequency Domain Analysis of Robust Demodulators for High-Speed Atomic Force Microscopy. In *2017 American Control Conference (ACC)*, 2017.
35. Harcombe, D. M.; Ruppert, M. G.; Ragazzon, M. R.; Fleming, A. J. Higher-harmonic AFM imaging with a high-bandwidth multifrequency Lyapunov filter. In *2017 IEEE International Conference on Advanced Intelligent Mechatronics (AIM)*, 2017; pp 725–730.
36. Harcombe, D. M.; Ruppert, M. G.; Ragazzon, M. R. P.; Fleming, A. J. *Beilstein J. Nanotechnol.* **2018**, *9*, 490–498. doi:10.3762/bjnano.9.47
37. Ruppert, M. G.; Harcombe, D. M.; Moore, S. I.; Fleming, A. J. Direct Design of Closed-loop Demodulators for Amplitude Modulation Atomic Force Microscopy. In *2018 Annual American Control Conference (ACC)*, 2018; pp 4336–4341.
38. Moore, S. I.; Ruppert, M. G.; Harcombe, D. M.; Fleming, A. J.; Yong, Y. K. *IEEE/ASME Trans. Mechatronics* **2019**, *24*, 1861–1870. doi:10.1109/tmech.2019.2928592
39. Cosens, C. R. *Proc. Phys. Soc., London* **1934**, *46*, 818–823. doi:10.1088/0959-5309/46/6/310
40. Michels, W. C.; Curtis, N. L. *Rev. Sci. Instrum.* **1941**, *12*, 444–447. doi:10.1063/1.1769919
41. Morris, E. D., Jr.; Johnston, H. S. *Rev. Sci. Instrum.* **1968**, *39*, 620–621. doi:10.1063/1.1683462
42. Lyons, R. G. *Understanding Digital Signal Processing*, 1st ed.; Addison-Wesley Longman Publishing Co., Inc.: Boston, MA, U.S.A., 1996.
43. Volder, J. E. *IRE Trans. Electron. Comput.* **1959**, *EC-8*, 330–334. doi:10.1109/tec.1959.5222693
44. Kokavecz, J.; Tóth, Z.; Horváth, Z. L.; Heszler, P.; Mechler, Á. *Nanotechnology* **2006**, *17*, S173–S177. doi:10.1088/0957-4484/17/7/s12
45. Abramovitch, D. Y. Coherent Demodulation with Reduced Latency Adapted for use in Scanning Probe Microscopes. U.S. Patent US20100128342A1, May 27, 2010.
46. Abramovitch, D. Y. Low latency demodulation for Atomic Force Microscopes, Part I efficient real-time integration. In *Proceedings of the 2011 American Control Conference*, 2011; pp 2252–2257.
47. Abramovitch, D. Y. Low Latency Demodulation for Atomic Force Microscopes, Part II: Efficient Calculation of Magnitude and Phase. In *IFAC Proceedings Volumes*, 2011; pp 12721–12726.
48. Kalman, R. E. *J. Basic Eng.* **1960**, *82*, 35–45. doi:10.1115/1.3662552
49. Grewal, M. S.; Henderson, V. D.; Miyasako, R. S. *IEEE Trans. Autom. Control* **1991**, *36*, 3–13. doi:10.1109/9.62283
50. Chen, S. Y. *IEEE Trans. Ind. Electron.* **2012**, *59*, 4409–4420. doi:10.1109/tie.2011.2162714
51. Harvey, A. C. *Forecasting, Structural Time Series Models and the Kalman Filter*; Cambridge University Press: Cambridge, United Kingdom, 1990.
52. Anderson, B.; Moore, J. *Optimal Filtering*; Prentice-Hall, 1979.
53. Brown, R. G.; Hwang, P. Y. *Introduction to Random Signals and Applied Kalman Filtering*, 3rd ed.; John Wiley & Sons, 1985.
54. Simon, D. *Optimal State Estimation*; John Wiley & Sons, Inc.: Hoboken, NJ, U.S.A., 2006. doi:10.1002/0470045345
55. Girgis, A.; Daniel Hwang, T. *IEEE Trans. Power Appar. Syst.* **1984**, *PAS-103*, 2943–2951. doi:10.1109/tpas.1984.318297
56. Ioannou, P. A.; Sun, J. *Robust Adaptive Controls*; Docer Publications, Inc.: Mineola, NY, U.S.A., 2012.

57. Karvinen, K. S.; Moheimani, S. O. R. *Mechatronics* **2014**, *24*, 661–671. doi:10.1016/j.mechatronics.2013.11.011
58. Tholén, E. A.; Platz, D.; Forchheimer, D.; Schuler, V.; Tholén, M. O.; Hutter, C.; Haviland, D. B. *Rev. Sci. Instrum.* **2011**, *82*, 026109. doi:10.1063/1.3541791
59. Ragazzon, M. R. P.; Gravdahl, J. T.; Fleming, A. J. On Amplitude Estimation for High-Speed Atomic Force Microscopy. In *2016 Annual American Control Conference (ACC)*, 2016; pp 2635–2642.
60. Stark, M.; Stark, R. W.; Heckl, W. M.; Guckenberger, R. *Appl. Phys. Lett.* **2000**, *77*, 3293–3295. doi:10.1063/1.1325404
61. García, R.; Magerle, R.; Perez, R. *Nat. Mater.* **2007**, *6*, 405–411. doi:10.1038/nmat1925

License and Terms

This is an Open Access article under the terms of the Creative Commons Attribution License (<https://creativecommons.org/licenses/by/4.0>). Please note that the reuse, redistribution and reproduction in particular requires that the authors and source are credited.

The license is subject to the *Beilstein Journal of Nanotechnology* terms and conditions: (<https://www.beilstein-journals.org/bjnano>)

The definitive version of this article is the electronic one which can be found at:
[doi:10.3762/bjnano.11.8](https://doi.org/10.3762/bjnano.11.8)