Lyapunov estimation for high-speed demodulation in multifrequency atomic force microscopy

David M. Harcombe*, Michael G. Ruppert¹, Michael R. P. Ragazzon² and Andrew J. Fleming¹

Abstract
An important issue in the emerging field of multifrequency atomic force microscopy (MF-AFM) is the accurate and fast demodulation of the cantilever-tip deflection signal. As this signal consists of multiple frequency components and noise processes, a lock-in amplifier is typically employed for its narrowband response. However, this demodulator suffers inherent bandwidth limitations as high-frequency mixing products must be filtered out and several must be operated in parallel. Many MF-AFM methods require amplitude and phase demodulation at multiple frequencies of interest, enabling both z-axis feedback and phase contrast imaging to be achieved. This article proposes a model-based multifrequency Lyapunov filter implemented on a field-programmable gate array (FPGA) for high-speed MF-AFM demodulation. System descriptions and simulations are verified by experimental results demonstrating high tracking bandwidths, strong off-mode rejection and minor sensitivity to cross-coupling effects. Additionally, a five-frequency system operating at 3.5 MHz is implemented for higher harmonic amplitude and phase imaging up to 1 MHz.

Introduction
Atomic force microscopy (AFM) [1] has been integral in the field of nanoscale engineering since its invention in 1986 by Binnig et al. By sensing microcantilever tip–sample interactions [2], atomic scale resolution imaging is achieved, which far exceeds the optical diffraction limit. An image generated by constant-force topography AFM depends entirely on its feed-
back control loop. The composition of a sample is visualized in three-dimensions by plotting the control signal against the lateral scan trajectories of the nanopositioner.

In static-mode AFM (contact mode), the control loop attempts to maintain a constant contact force [3]. Whereas in dynamic modes, for example intermittent-contact constant-amplitude AFM [4], the control loop acts to maintain a constant cantilever oscillation amplitude. This is achieved by feeding back the demodulated fundamental frequency present in the deflection signal. In intermittent-contact mode AFM [5], the tapping amplitude is chosen such that only gentle tip–sample interactions occur. This is particularly suitable for studying biological samples, allowing for biophysical processes to be studied [6-8].

Multifrequency AFM (MF-AFM) methods allow for the study of tip–sample interactions occurring at multiple frequencies [9]. This extends imaging information beyond the topography to a range of nanomechanical properties including sample stiffness, elasticity and adhesiveness [10]. The acquisition of these observables requires tracking the amplitude and phase of additional frequencies of interest. These include higher harmonics of the fundamental frequency [11], higher flexural eigenmodes [12] and intermodulation products [13]. Higher harmonic methods have demonstrated the ability to image relatively large biological objects, such as cells [14,15], while bimodal AFM has successfully imaged properties of protein complexes [16].

Intermodulation AFM is a novel extension to the bimodal method that focuses on the mixing products of a slightly below and above resonance bimodal drive. It has been shown to achieve increased image contrast [17] and lead to further insights into nanomechanical properties [18]. Regardless of which particular MF-AFM method is employed, they each require the demodulation of amplitude and phase to form observables for the characterization of nanomechanical properties.

Due to the large bandwidth requirements of tracking high frequencies in MF-AFM, every component of the z-axis feedback loop detailed in Figure 1 needs to be optimized for speed. This includes the lateral and vertical nanopositioner for each axis (x, y and z), cantilever, vertical feedback controller and demodulator. In this article, the demodulator component is improved with respect to its key performance metrics: tracking bandwidth, sensitivity to other frequency components, and implementation complexity. Tracking bandwidth is defined as the point in a frequency response where the output of the demodulator drops by $-3$ dB with respect to the input modulating signal. Off-mode rejection (OMR) is a term that describes the attenuation of unwanted frequencies present in the input signal, which lie outside the modeled carrier of interest.

It has been shown that conventional high-speed demodulation techniques are incompatible with MF-AFM, due to the lack of sensitivity to multiple frequency components [19]. These include the peak detector [20], peak-hold [21] and RMS-to-DC [22] conversion methods. A typical MF-AFM demodulator employs multiple lock-in amplifiers (LIA) in parallel, as each provides an accurate estimation of a particular frequency component. However, low-pass filters are employed to diminish mixing products, which severely limits the demodulator bandwidth [23].

Motivated by improving MF-AFM demodulation capabilities, previous work by the authors includes a multifrequency Kalman filter [24]. It was shown to outperform a commercially available LIA in terms of both tracking bandwidth and noise performance. However, a major disadvantage of the Kalman filter is the large computational expense of each additional frequency modeled. This reduces its realizable performance through limitations of the sample rate. An estimator in the form of a Lyapunov filter [25] was demonstrated to perform similarly to the Kalman filter [26]. However, the Lyapunov filter complexity scales significantly better than the Kalman filter when multiple frequencies are modeled [27].

This article extends previous work by providing a thorough performance analysis of the multifrequency Lyapunov filter in terms of tracking bandwidth, off-mode rejection and cross-coupling effects. In addition, MF-AFM demodulation is demonstrated by performing higher harmonic imaging with amplitude and phase on both a stiff and compliant sample.

**Lyapunov Filter**

**System modeling**

A single-frequency cantilever deflection signal is modeled as a sine wave with carrier frequency $f_c$, time-varying amplitude $A(t)$ and phase $\phi(t)$ of the form.
\[ y(t) = A(t) \sin(\omega t + \phi(t)). \] (1)

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

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

where \( i = 1, 2, \ldots, n \) denotes the \( i \)-th modeled frequency. Through the double-angle trigonometric identity, this model is linearly parameterizable such that vector pairs within the state vector \( x = [x_1, x_2, \ldots, x_{2n-1}, x_{2n}]^T \) represent quadrature and in-phase components of each particular modeled frequency. That is, each individual sine wave is represented by

\[ y_i(t) = x_{2i-1} \cos(\omega_i t) + x_{2i} \sin(\omega_i t). \] (3)

Based on the parametrization of the signals in Equation 3, the time-varying amplitude and phase of a particular frequency is recovered by

\[ A_i = \sqrt{x_{2i-1}^2 + x_{2i}^2}, \]
\[ \phi_i = \arctan \left( \frac{x_{2i-1}}{x_{2i}} \right). \] (4)

**Filter description**

The Lyapunov filter [28] is implemented as a linear observer as shown in Figure 2. A key property of the filter is exponential convergence of the estimated states [29], with the tunable loop gain constant \( \gamma \) governing the speed of convergence. The filter is shown to have a negative feedback loop in which integral action regulates the error. By feeding back an estimate of the input signal obtained from the parameterized states in the form of Equation 3, an error signal is generated. Regulation of this error through feedback leads to the much desired suppression of the \( 2f_c \) mixing components.

The update law for the single-frequency Lyapunov filter [28] can be extended to a multifrequency form, resulting in

\[ \dot{x} = \Gamma c^T [y(t) - \tilde{y}(t)], \] (5)
\[ \tilde{y}(t) = c \hat{x}, \] (6)

where

\[ \Gamma = \text{diag}(\gamma_1, \gamma_1, \gamma_2, \gamma_2, \ldots, \gamma_n, \gamma_n) \] (7)

and

\[ c = [\cos(\omega_1 t) \sin(\omega_1 t), \ldots, \cos(\omega_n t) \sin(\omega_n t)]^T. \] (8)

In this form, \( \tilde{y}(t) \) represents the estimated output and the amplitude \( A_i \) and phase \( \phi_i \) are available by applying Equation 4 to each quadrature and in-phase pair of \( \hat{x} \). A key requirement to ensure exponential convergence of \( \hat{x} \) to \( x \), is to guarantee that \( c \) is persistently excited [29]. Convergence is shown for the single-frequency filter in [28], and can easily be extended for the multifrequency case. Furthermore, exponential convergence of \( \hat{x} \) means that \( \hat{A} \) and \( \hat{\phi} \) also converge.

**Results and Discussion**

**Hardware**

The Lyapunov filter was implemented on a high-speed FPGA to achieve the necessary sample rate for accessing higher harmonics during imaging. A Xilinx Kintex-7 KC705 evaluation board (model: XC7K325T) paired with a DC-coupled high-speed 4DSP input/output (I/O) card (model: FMC151) was utilized. The FPGA clock is synchronized with the high-speed I/O card at 250 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.
Implementation

Figure 2 shows the block diagram of a single-frequency Lyapunov filter (SF-LYAP). Here, the digital components required for FPGA implementation can be seen: multipliers, adders, registers, sample rate control for feedback and a programmable direct digital synthesizer (DDS). The DDS generates the sine (in-phase) and cosine (quadrature) signals required to model carrier frequencies. It may be tuned through control of its frequency word which is calculated by

\[ \text{FW} = \frac{2^n f_{\text{out}}}{f_{\text{clk}}} \]  

where \( f_{\text{out}} \) is the desired output frequency, FW represents the binary word required to program the DDS, \( n \) is the length of FW and \( f_{\text{clk}} \) is the speed of the FPGA board.

A SF-LYAP was successfully implemented at a sampling rate of \( f_s = 5 \text{ MHz} \). As stability is of priority, the chosen data representation is floating point in the standard IEEE 754 format. The integration method used is backward Euler, as this ensures stability when \( \gamma \) is large [30]. The output equation (Equation 4) is realized with the Xilinx Coordinate Rotation Digital Computer (CORDIC), set to a 16-bit configuration such that the amplitude and phase are formatted for the I/O card. The carrier frequency \( f_c, \gamma \) and any necessary output gains for amplifying very small signals during imaging are tunable in real-time using the Xilinx Virtual Input Ouput (VIO) tool.

In Figure 3, the block diagram of the implemented multifrequency Lyapunov filter (MF-LYAP) is shown. Here, it can be seen that an MF-LYAP involves several SF-LYAPs running in parallel. Channel cross-coupling occurs in the combined output feedback as dictated by the output equation (Equation 6). The Lyapunov filters timing constraints for a five-frequency system result in a maximum sampling rate of \( f_s = 3.5 \text{ MHz} \). This is a large improvement over the multifrequency Kalman filter [24], which was 1.5 MHz for a three-frequency system. The Kalman filter equations [24] can be shown to have a complexity of \( O(n^3) \), while that of the Lyapunov filter is \( O(n) \) for \( n \) modeled frequencies. This stark difference in complexity arises from the computations required for the Kalman gain and covariance matrix update.

Experimental setup

A LIA (Zürich Instruments HF2LI) was used in-conjunction with a laboratory function generator (Agilent 33521A waveform generator) to experimentally verify the performance of the implemented Lyapunov filter. These investigations include a frequency response experiment to measure the tracking bandwidth and channel cross-coupling. Additionally, off-mode rejection of channels in both high-speed and slow configurations was explored through a carrier sweep.

Tracking bandwidth

The tracking bandwidth of the Lyapunov filter was characterized through frequency responses from both a simulated and experimentally implemented system. For each frequency response, the modulating signal \( A(t) \) in Equation 1 was swept from DC to 1.5 MHz while the carrier frequency was held constant. The tracking bandwidth experiment examines the relationship between the \(-3 \text{ dB} \) point and \( \gamma \) for a 1 MHz carrier frequency and \( \gamma \) values ranging between \( 5 \times 10^4 \) and \( 1 \times 10^7 \). Figure 4 shows the results of (a) simulated and (b) experimental tracking frequency responses, where it can be seen that the two systems match closely. The similarity was achieved by maintaining a consistent sample rate and integration method for both simulation and experimental implementation. In Figure 4c, the simulated and experimental \(-3 \text{ dB} \) points are shown as a function of the tunable loop constant \( \gamma \). For both systems, the tracking bandwidth approaches the carrier frequency \( f_c \).

Cross-coupling

The effect of channel cross-coupling on the tracking bandwidth was examined for both a simulated and experimentally implemented system. For simplicity, cross-coupling was demonstrated with a two-frequency MF-LYAP wherein the modeled carrier frequencies are 100 kHz and 500 kHz for channels 1 and 2, respectively. Each channel is considered for two fixed
Figure 4: (a) Simulated and (b) experimentally obtained frequency responses of the Lyapunov filter with a carrier frequency of 1 MHz and varying tracking bandwidths, as indicated by the increasing $\gamma$ values. Results in (c) show the $-3$ dB tracking bandwidth of each system as a function of $\gamma$.

Figure 5: Experimental frequency responses from a single-frequency Lyapunov filter with five carrier frequencies tuned to a $-3$ dB tracking bandwidth of $f_c$.

Off-mode rejection

The off-mode rejection of the multifrequency Lyapunov filter was analyzed by performing a single-tone sine sweep on the input signal and recording the demodulated amplitude magnitude of each channel. For each frequency response, the carrier frequency $f_c$ in Equation 1 was swept from DC to 1.25 MHz with a constant amplitude $A$. This experiment used a five-frequency MF-LYAP with channels set to carriers of 100 kHz, tracking bandwidth settings, low (1 kHz) and high (50% of $f_c$), while the other channel is increased in speed.

Figure 6a–c shows that the tracking bandwidth of a channel will increase from its original setting as the other channel is tuned faster. Conversely, Figure 6d shows channel 2 slowing down as channel 1 is increased in speed. This is explained by the fact that channel 2 is set to a tracking bandwidth of 250 kHz (50% $f_c$), which is higher than the maximum obtainable speed of channel 1. Throughout this investigation, the simulation and experimental results agree. The results show that cross-coupling effects are more pronounced in low-speed channels. They are, however, negligible if tracking bandwidths of channels remain below 10% of $f_c$.
200 kHz, 500 kHz, 700 kHz and 1 MHz for both a simulated and experimentally implemented system.

Figure 7a–e shows off-mode rejection for a fast (10% of $f_c$) tracking bandwidth setting. For each channel, a full recovery (0 dB) of the signal can be seen to occur at its modeled carrier frequency, as expected. There is strong off-mode rejection occurring at the other modeled carrier frequencies, due to output feedback cross-coupling sharing state information between channels. Figure 7f–j demonstrates off-mode rejection for a slow (1 kHz) tracking bandwidth setting. Here, the narrowband response is a direct result of the reduced $\gamma_i$ values of each channel. This causes a less distinct, but still visible, modeled off-mode rejection at the other carrier frequencies. It can be seen that the slower system achieves greater off-mode rejection outside of the modeled frequencies than the fast system. Again, a similar performance between the simulated and experimental results can be observed. The less distinct off-mode rejection in the experimental results compared to the simulations is due to a finite DC offset from the DAC. This precludes the direct measurement of signals smaller than this value.

AFM Imaging

Imaging Setup

The Lyapunov filter as a multifrequency AFM demodulator was validated through a series of imaging experiments where it is compared side-by-side to a lock-in amplifier. To ensure a fair comparison, the demodulators were tuned to the same tracking bandwidth in both experiments. This is required as the noise performance has been shown to be a function of the tracking bandwidth [19]. The lock-in amplifier is the state-of-the-art multifrequency method due to its strong off-mode rejection, however it can not achieve the same speed as the Lyapunov filter due to post-mixing filtering [19]. As the high-speed superiority of the Lyapunov filter is well established, it is compared to the lock-in amplifier in a low-speed environment.

Using an NT-MDT NTEGRA AFM, amplitude and phase higher harmonic imaging was performed with a NT-MDT NSG01 and Bruker DMASP cantilever. These cantilevers were found to have fundamental resonance frequencies of 168.8 kHz and 46.1 kHz, respectively. The samples used are a z-calibration grating (NT-MDT TGZ3) with periodic height features of approx. 500 nm and 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 LDPE regions, the sample is widely used for qualitative imaging the material contrast.

Imaging a TGZ3 Calibration Grating

Higher-harmonic amplitude images with the first, second, third, sixth and seventh harmonics were obtained by the MF-LYAP and multiple parallel LIAs. The frequency response of the NSG01 cantilever in free air and the power spectrum of its deflection signal during contact are shown in Figure 8. Here, the fundamental and second resonance frequencies can be seen in the cantilever frequency response. The deflection signal spectrum shows additional higher harmonics and minor intermodulation products are present. These are due to non-linear atomic forces exciting the cantilever during contact.
Amplitude imaging results are shown in Figure 9. As the sixth and seventh harmonics are closely spaced to the second resonance frequency of the cantilever, they provide an increased signal-to-noise ratio. The MF-LYAP can be seen to perform comparably to the LIA when tuned to similar measurement bandwidth settings. When imaging with higher harmonics, the off-mode rejection of each channel was tuned to suppress the large fundamental frequency.

Imaging a PS/LPDE calibration grating
Higher-harmonic phase images were obtained for the first five harmonics of a Bruker DMASP cantilever. The frequency response of the cantilever in free air and the power spectrum of its deflection signal during contact are shown in Figure 10. Here, the fundamental resonance frequency and higher eigenmodes can be seen in the cantilever frequency response. As before, the deflection signal contains additional higher harmonics and intermodulation products due to the non-linear atomic excitation. Note that the DMASP cantilever uses integrated piezoelectric actuation [31], which results in a clean frequency response when compared to the base excited NSG01 as seen in Figure 8.

The higher-harmonic phase imaging results are shown in Figure 11. For both demodulators, we see a strong material contrast between the PS and LPDE regions. This was expected from the rich frequency content present in the deflection signal, as visible in Figure 10. We note that the images show particularly strong contrast for the second harmonic, which is due to its proximity to the second mode of the cantilever. This fact is also visible in the increased noise floor around that frequency in the deflection signal. Similarly to the amplitude imaging, the large fundamental frequency contribution required tuning higher harmonics for increased off-mode rejection. For this reason, we tuned the first harmonic demodulators to 1 kHz bandwidth (LIA LPF 1 kHz, LYAP $\gamma = 20 \times 10^3$ and the higher harmonics to 200 Hz (LIA LPF 200 Hz, LYAP $\gamma = 2 \times 10^3$).

Conclusion
This article describes a multifrequency Lyapunov filter for high-speed demodulation in MF-AFM. The performance and flexibility of the proposed Lyapunov filter is demonstrated through simulations and experiments. The filter may reach tracking bandwidths up to the modeled carrier frequency, the equivalent of single-cycle tracking. Additionally, the off-mode
rejection of the system was found to be controlled by its bandwidth as dictated by the tunable loop constant $\gamma$. The relationship between $\gamma$ and the bandwidth was shown to be linear, up to the modeled carrier frequency. Channel cross-coupling, which occurs due to output feedback, was found to cause distinct rejection of other modeled frequencies during the off-mode rejection experiments. An investigation into this cross-coupling revealed it has negligible effect on the tracking bandwidth of the system.

The multifrequency Lyapunov filter as a flexible, high-speed demodulator was verified through higher harmonic MF-AFM imaging for both amplitude and phase. This demonstrates the filters ability to be used as a demodulator in various MF-AFM techniques involving higher harmonic, higher eigenmode or intermodulation frequency components. In the presented AFM images, the proposed filter performed comparably to a state-of-the-art lock-in amplifier setup. In comparison to the Kalman filter, the Lyapunov filter is similar in terms of speed, off-mode rejection and operation. However, it was found to be significantly easier to implement, which is a priority when considering an extension to multiple frequencies.

ORCID® iDs
David M. Harcombe - https://orcid.org/0000-0002-8866-5806
Michael G. Ruppert - https://orcid.org/0000-0003-2286-4929
Michael R. P. Ragazzon - https://orcid.org/0000-0003-3770-5154
Andrew J. Fleming - https://orcid.org/0000-0001-7132-8458

References
1. Binnig, G.; Quate, C. F.; Gerber, C. Phys. Rev. Lett. 1986, 56, 930–933. doi:10.1103/PhysRevLett.56.930
2. Rabe, U.; Janser, K.; Arnold, W. Rev. Sci. Instrum. 1996, 67, 3281–3293. doi:10.1063/1.1147409
3. Abramovitch, D. Y.; Andersson, S. B.; Pao, L. Y.; Schitter, G. A tutorial on the mechanisms, dynamics and control of atomic force microscopes. 2007 American Control Conference; 2007; pp 3488–3502.
4. García, R. Amplitude modulation atomic force microscopy; John Wiley & Sons: New York, NY, U.S.A., 2011.
5. Zhong, Q.; Inniss, D.; Kjoller, K.; Elings, V. B. Surf. Sci. 1993, 290, 688–692. doi:10.1016/0039-6028(93)90582-5
6. 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
7. Kodera, N.; Yamamoto, D.; Ishikawa, R.; Ando, T. Nature 2010, 468, 72. doi:10.1038/nature09450
