Height Drift Correction in Non-Raster Atomic Force Microscopy

Travis R. Meyer,[1] Dominik Ziegler,[2] Christoph Brune,[3] Alex Chen,[4] Rodrigo Farnham,[5] Nen Huynh,[5] Jen-Mei Chang,[5] Andrea Bertozzi,[1,*] and Paul D. Ashby[2,†]

[1] *Department of Mathematics, University of California Los Angeles, Los Angeles, CA 90095 USA*
[2] *Molecular Foundry, Lawrence Berkeley National Laboratory, Berkeley, CA 94720 USA*
[3] *Institute for Computational and Applied Mathematics, University of Mnster, Germany*
[4] *Statistical and Applied Mathematical Sciences Institute, Research Triangle Park, NC 27709 USA*
[5] *Department of Mathematics and Statistics, California State University Long Beach, Long Beach, CA 90840 USA*

We propose a novel method to detect and correct drift in non-raster scanning probe microscopy. In conventional raster scanning drift is usually corrected by subtracting a fitted polynomial from each scan line, but sample tilt or large topographic features can result in severe artifacts. Our method uses self-intersecting scan paths to distinguish drift from topographic features. Observing the height differences when passing the same position at different times enables the reconstruction of a continuous function of drift. We show that a small number of self-intersections is adequate for automatic and reliable drift correction. Additionally, we introduce a fitness function which provides a quantitative measure of drift correctability for any arbitrary scan shape.

Keywords: Atomic force microscopy, self-intersecting scan, drift correction, non-raster scan

I. INTRODUCTION

Atomic Force Microscopy (AFM) measures the interaction force between a sharp tip and sample to acquire high resolution images by serially scanning a sample area while recording these minute interactions often at sub-nanometer resolution [1–4]. The fact that AFM achieves high resolution imaging over a large variety of sample types and environments makes it one of the most frequently used characterization tools in nanoscience. As AFM mechanically detects sub-nanometer size features, its accuracy is easily compromised by drift. Drift mainly originates from thermal fluctuations resulting in the slow expansion and contraction of instrument parts. Traditionally, in raster scanning height drift, or z-drift, is corrected using flattening. Flattening is performed by removing a low order polynomial fit from each scan line. This rudimentary technique often results in artifacts where sample features become partially removed or tilted. Extracting accurate topographic data, therefore, often requires the user to choose an appropriate combination of different flattening techniques.

Part of the rationale for a raster pattern is that the data samples align with a regular grid, making the data ideal for visualizing in a pixelated image. For each data point raster scanning requires the probe tip to be at a specific location at a given time. But non-linearities such as hysteresis and creep associated with the multidomain properties of high sensitivity piezoelectric materials make this a difficult engineering task [5–8]. We have shown that these problems can be overcome with Sensor Inpainting techniques [9] which use advanced inpainting algorithms [10–15] to generate accurate images based on position sensor data, see Appendix A. The technique, furthermore, frees scanning probe microscopy from the paradigm of raster scanning so that data recorded along any arbitrary path can be used to generate an image.

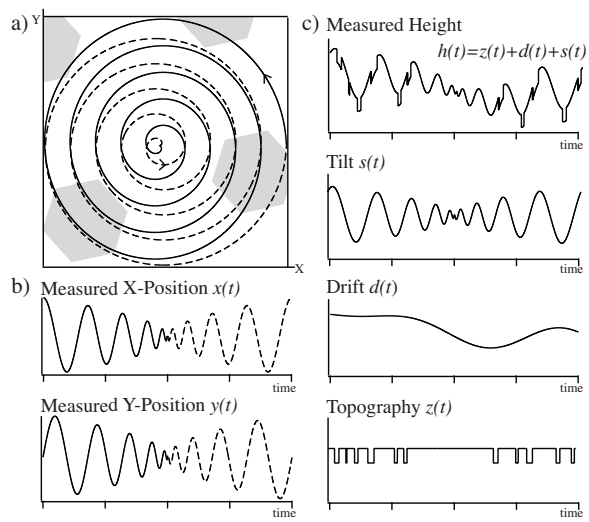

FIG. 1. **a)** Illustration of a five loop Archimedean spiral showing inward (solid line) and outward (dashed line) scan path. The gray hexagons illustrate recessed topographic features. **b)** measured X and Y positions versus time **c)** The measured height $h(t)$ along the scan pass comprises effects from the tilt $s(t)$, drift $d(t)$ as well as the real topography $z(t)$ which we aim to recover as accurately as possible to generate a clean topographic image.

This enables the use of sinusoidal scan patterns that require less bandwidth and are better suited to drive high inertia nanopositioners [16–18]. FIG. 1 shows an Archimedean spiral as a typical non-raster scan path. For legibility we show a scan path with only five loops for the inward (solid line) and outward scans (dashed line). Raster scanning typically only uses either trace or retrace data for generating an image, but in Sensor In-

* E-Mail: bertozzi@math.ucla.edu
† E-Mail: pdashby@lbl.gov

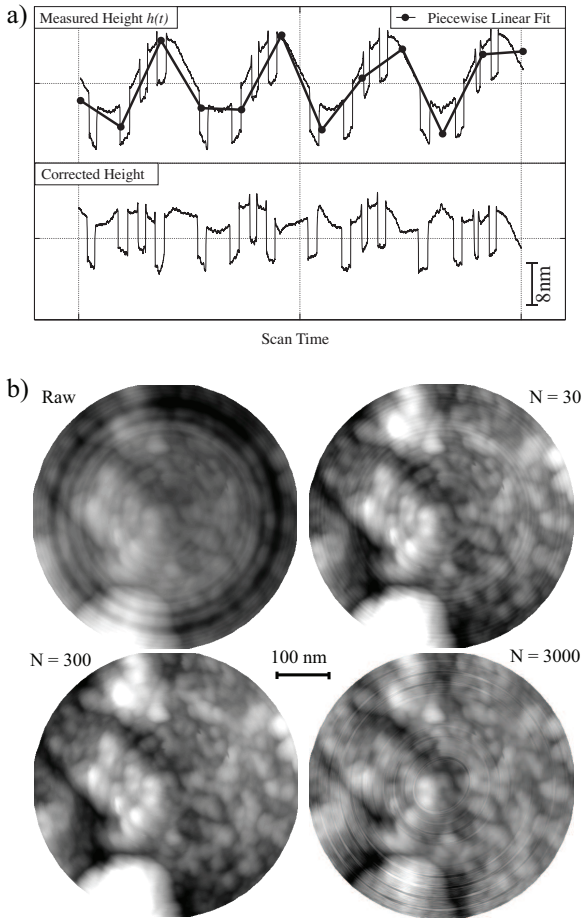

FIG. 2. Examples of piecewise linear fitting. The measured height signal along non-raster scan paths **a)** The recorded height $h(t)$ (top) is split into N segments onto which linear fitting is used to correct the drift component. A corrected height (bottom) results from subtracting the piecewise linear fit from the original data. Because of tilt flattening distorts many of the topographic features. **b)** 500 nm diameter Archimedean Spiral showing the grains of evaporated gold where drift data results in rings (raw). Piecewise linear fitting can flatten these artifacts (N=300). However, drift related artifacts remain present when too few segments are chosen (N=30). Height information is lost and topographic features are flattened when too many segments are used (N=3000).

painting 100% of the data can be displayed and at least a two-fold increase in temporal or spatial resolution is achieved. For high temporal resolution inward and outward scans can be used separately, and for higher lateral resolution a single image can be generated using the data collected on the inward and outward scan together. As shown in FIG. 1 c) the recorded height $h(t)$ typically contains contributions from the tilt of the sample $s(t)$ and drift of the instrument $d(t)$. In the presence of drift, the same point in space may have different measured heights because of temporal separation. It is therefore extremely important to have reliable methods to detect and correct drift in non-raster scan AFM. Using Sensor Inpainting the tip is not required to follow a linear motion. As a consequence, sample tilt does not necessarily result in a linear feature in the recorded height time trace. FIG. 2 shows our attempt to generalize the naïve fit-and-flatten approach to non-raster scan paths. FIG. 2 a) shows a time trace recorded using a non-raster scan pattern over a calibration grating with 8 nm deep features. A continuous piecewise linear fit is overlaid. The corrected height signal is shown in the bottom panel. It is clear that the linear fitting distorts many of the features. Without distinguishing drift from tilt any fitting algorithm will potentially distort the measured topography and result in loss of reliability. Furthermore, for a non-raster scan it is not obvious how to best choose the length of the linear segments for line flattening. In FIG. 2 b) we show an Archimedean spiral scan on evaporated gold; the individual grains are visible but rings in the raw image resulted from drift. Piecewise linear fitting can flatten these artifacts if the appropriate number of segments is chosen (N=300). The technique suffers from the same problems as flattening techniques in raster scan AFM where height information can easily be altered by the flattening and having too many segments results in a loss of topographic information (N=3000). The requirement to find the right flattening parameters makes the technique unreliable. In Section II we introduce drift detection using self-intersecting scan paths. Our new method requires no human interaction and is significantly more reliable since it detects and corrects drift independent of sample topography or tilt. In Section III we discuss the performance the self-intersection method, and finally in Section IV we present a summary and conclusion.

II. DRIFT CORRECTION USING SELF-INTERSECTING SCANS

A. Self-Intersecting Scan Patterns

Our method for measuring and correcting drift is based on measuring the height at a known position but different times. This requires the scan pattern to self-intersect, but as Sensor Inpainting [9] can generate an image from any arbitrary scan path, we are free to use any suitable path. FIG. 1 a) shows the most commonly used non-raster scan pattern: the double Archimedian spiral (DAS). Note that inverting the Y drive signal for the outward spiral (dashed line) maintains a counter clock wise motion and leads to a continuos path with no abrupt changes in scan direction and, more importantly, it introduces two self-intersections per loop. Fig. 3 a) shows such a spiral scan with 25 loops. The location of the self-intersections are indicated by red dots, all of which lie on a line. Measuring the height of a given position twice with only small temporal separation does not give a good measure of drift, and since thermal drift occurs on large time scales, scan paths with intersections of large temporal separation contain information better suited for

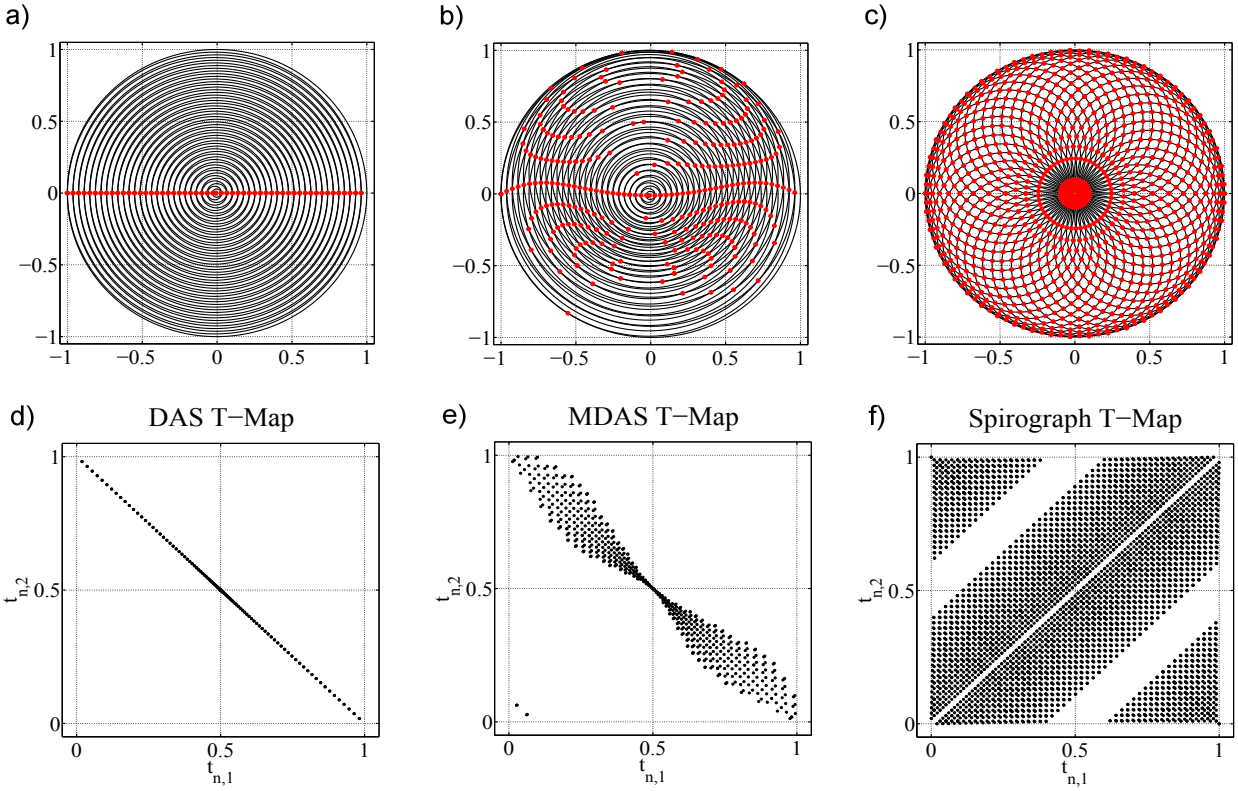

FIG. 3. Typical non-raster scan patterns where the red dots indicate the locations of self-intersections. **a)** A double Archimedean spiral (DAS) with 25 loops **b)** more intersections are achieved by adding a sine-wave modulation in one scan axis producing the modulated double Archimedean spiral (MDAS) **c)** a Spirograph scan pattern with a hundred loops. The temporal maps (T-maps) **d)**, **e)**, **f)** show the intersection time pairs for all scan paths. A distribution away from the diagonal from lower left to upper right is optimal for drift correction. A quantitative measure is provided by the fitness ζ. The DAS has a low fitness of $\zeta \approx 0.02$ indicating poor intersections, adding the modulation it is increased to $\zeta \approx 1.2$ for MDAS. The Spirograph has the highest fitness of $\zeta \approx 35$ but suffers from non-uniform data density, i.e. data collection is denser in the middle and outside.

the detection and correction of drift. In the modulated double Archimedean spiral (MDAS) shown in FIG. 3 b) the number of self-intersections and their temporal and spatial distribution is increased by perturbing one spatial coordinate with a sine wave of period equal to the scan time and amplitude of one tenth of the scan size. FIG. 3 c) shows a Spirograph, another type of sinusoidal scan pattern which generates many more self-intersections. In Section II D we will introduce a fitness function which provides a quantitative measure for drift correctability for each of these scan forms. However, its description requires the introduction of the least squares difference method first.

B. Finding Intersections

The first step in using the differences method for drift compensation is to find the self-intersections of the curve that represents the scan path. For a well-defined curve in continuous space, intersections are equally well-defined. When the curve is represented by discrete samples, however, the problem becomes somewhat more complex and ill-defined. As the AFMs position sensor gives quantized information at a finite sampling rate, the intersection detection algorithm can easily generate false intersections, throwing off the differences fitting energy. To combat this issue, we convolve each of the position sensor signals with a Gaussian smoothing function with a standard deviation on the same order as the smallest resolved distance by the sensors. One solution is to compose the N discrete samples $\{\vec{x}_i\}_{i=1}^{N}$ of the curve into $N-1$ connected line segments joining adjacent samples described by

$$X_i = \{(1-t)\vec{x}_i + t\vec{x}_{i+1} \mid t \in (0,1]\}.$$

Then the line segments can be checked against each other for intersections in $O(N^2)$ time by checking all possible intersections pair-wise. For a more detailed description of the algorithm see Appendix B.

C. Least Squares Differences Algorithm

Error in the height measurement due to z-drift is locally approximately linear and can be modeled as a smooth function with small second derivative. This assumption together with the existence of path self-intersections motivate our approach. Let $\vec{x}(t)$ be vector containing $x(t)$ and $y(t)$ that describes the scan path i.e. position of the AFM probe on the sample. Using accurate sensors to measure these positions we can assume minimal position errors. As introduced in FIG. 1 the height signal measured by the AFM during a particular scan is denoted by $h(t)$, which may be decomposed into $h(t) = z(t) + d(t) + s(t)$ representing the effective topography $z(t)$, drift $d(t)$ and tilt $s(t)$ components. It is assumed that $|d''| \ll |\vec{x}''|$. For a scan path which intersects itself M times, define the times of self-intersection $t_{n,1}$, $t_{n,2}$ so that $\vec{x}(t_{n,1}) = \vec{x}(t_{n,2})$ where $n = 1, 2, ..., M$. At each of these M points of intersections, we observe that

$$h(t_{n,2}) - h(t_{n,1}) = z(t_{n,2}) - z(t_{n,1}) + d(t_{n,2}) - d(t_{n,1}) + s(t_{n,2}) - s(t_{n,1})$$

because the topography z and the tilt s are well-defined functions of position, the difference vanishes at points of self-intersection and thus,

$$h(t_{n,2}) - h(t_{n,1}) = d(t_{n,2}) - d(t_{n,1}) := d_n.$$

In order to exploit the existence of self-intersections, we propose the minimum of the energy functional $E(f)$ for the continuous drift $d(t)$.

$$E(f) = \sum_{n=1}^{M} (f(t_{n,2}) - f(t_{n,1}) - d_n)^2 + \lambda \int_0^1 |f''(t)|^2 dt \tag{1}$$

The minimization reduces to solving an Euler-Lagrange equation. We find a sufficiently differentiable function which minimizes $E(f)$ and thus provides the continuous drift function $d(t) \equiv f(t)$. For this particular problem, these equations result in the system of the form $L_\lambda[f] = \vec{h}$ where L_λ is a linear operator and $\vec{h}$ is a vector function of the data d_n. The derivations for L_λ and $\vec{h}$ using a finite basis expansion are given in Appendix C. While a finite differences approximation is the usual approach to solving this, restricting the problem to only the span of a small basis significantly reduces the computational complexity. Since the drift function is assumed to be smooth, a basis of smooth time-localized functions such as uniformly spaced splines or Gaussian curves, the latter of which which we use, is appealing. The solution to this problem, f, is thus a smoothed fit to the difference errors d_n. The first term in $E(f)$ enforces fidelity to the measured drift while the second term enforces smoothness of the final result and removes noise. Once f is calculated, the signal $h(t) - f(t)$ is the corrected height from the algorithm. Since the height differences at the self-intersection points are invariant under sample tilt, we may first subtract the z-drift fit then secondly subtract a plane fit from the data, henceforth known as tilt removal, without worrying about an interaction between the algorithms. Performing the tilt removal first may mistakenly interpret drift as tilt and subtract it off, permanently distorting the signal. Thus self-intersection drift correction should precede tilt removal.

D. Quality of Self-Intersections

The different scan paths shown in FIG. 3 a) b) and c) result in different distributions of the intersection times throughout the scan. This distribution contributes to the quality of information about drift that the intersections provide. FIG. 3 d) e) and f) show the temporal maps for the discussed DAS, MDAS, and Spirograph scan patterns. The temporal maps plot intersection time pairs $(t_{n,1}, t_{n,2})$. The diagonal which runs from the lower-left to upper-right represents intersections with $t_{n,1} = t_{n,2}$, while distance from this diagonal indicates temporal separation $|t_{n,1} - t_{n,2}|$. Since thermal drift occurs on large time scales, scan paths with intersections of large temporal separation contain the best information for thermal drift discovery and removal.

We propose the first non-trivial eigenvalue of L_λ, labeled henceforth as ζ, as a measure for the quality of the intersection information provided by the scan path with intersection times described by $t_{n,j}$. This quantity is related to both the likelihood of z-drift being detected and the strength of the minimum of the energy function $E(f)$. The corresponding eigenfunction represents a drift profile d which changes the energy the least, and is thus closest to being missed by the model. Scan paths which are better suited for the proposed algorithm will have large eigenvalues associated with L_λ, and therefore ζ provides a quantitative measure of scan path quality. The appropriate value of the penalty λ used for both the correction and fitness calculation remains an ongoing problem in the image processing field. Generally, well-structured curves with many self-intersections are effectively corrected with $\lambda = 0$; however, a low ζ_λ fitness results when significant noise is present or the intersection information is insufficient to describe the drift. Indeed, for particularly poor intersection data the fitting is singular. Penalization is suggested to mitigate these situations and guarantee non-singularity. However, we have generally observed that penalization is only necessary with scans of exceptionally poor ζ fitness. For the comparison of different scan paths we henceforth use $\lambda = 10^{-3}$ and the corresponding ζ as a description of the fitness. Compared to a simple DAS the modulation in the MDAS increases the fitness ζ from 0.02 to 1.2. The Spirograph pattern attains an even higher fitness value than the DAS and MDAS at $\zeta = 35$. However, non-uniform sampling over the surface area and

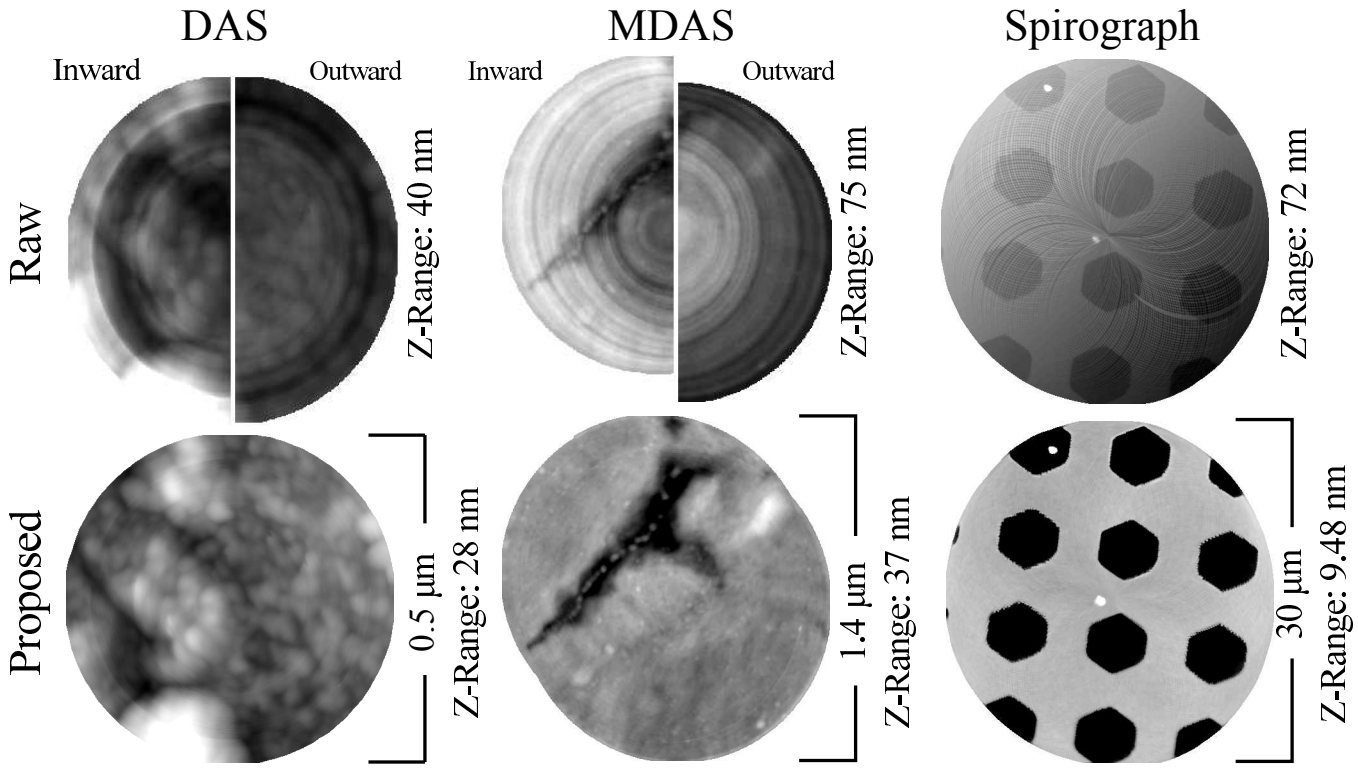

FIG. 4. **a)** Raw data of a DAS scan of a gold sample inward (left) and outward (right) scans are shown for comparison. **b)** Raw data of a larger MDAS scan with large drift in z **c)** Raw data of a Spirograph scan over an AFM calibration sample. **d)** Drift corrected topography using self-intersection method. **e)** The combination of drift correction and Sensor Inpainting perfectly reconstructs the real topography even in the presence of significant drift in X,Y and Z. **f)** Drift corrected Spirograph scan.

significantly more self-intersections than necessary to recover $d(t)$ makes this geometry inherently less attractive. Nevertheless, the Spirograph demonstrates the wide variety of possibilities and the generality of the proposed algorithm.

III. RESULTS AND DISCUSSION

We modified a commercial AFM (MFP-3D, Asylum Research) to steer the tip along any desired scan path on the sample while recording the height using tapping mode. To visualize the data we used Sensor Inpainting [9] using a heat equation algorithm [12]. (A more detailed description on this image reconstruction technique is given in Appendix A.) The raw height data is shown in the top row of FIG. 4 and the bottom row shows the corrected height using the self-intersection drift correction method. The first dataset FIG. 4 a) and d) is the same dataset as used for piecewise linear fitting (see FIG. 2) showing a 500 nm diameter scan over an annealed gold sample. To better illustrate the fluctuations in the measured height we generated an inpainted image using the raw data of the inward and outward scans and put two halves of the images next to each other (FIG. 4 a)). The result obtained using the self-intersection drift correction method (FIG. 4 d)) shows a clean image which preserves the height of topographic features. Compared to the piecewise linear fitting technique our drift correction algorithms does not need any input parameters and hence no human interaction is required. The double Archimedian scan has 1700 loops which result in 111k self-intersections with a fitness ζ of 0.8. The second example is taken on the same gold sample but this time we use a larger scan size of 1.4 μm with 27k intersections using the MDAS scan pattern and 471 loops (FIG. 4 b)). With the resulting fitness ζ of 126 our algorithm easily corrects for the large z drift present. Note that the large lateral drift in X and Y is accurately recorded by the sensors, and thanks to Sensor Inpainting such drift correctly results in a slightly elongated image which perfectly represents the scanned surface (FIG. 4 e)). The third experiment was taken from a silicon calibration sample with 8 nm deep hexagonal features. The raw data (FIG. 4 c)) is recorded over 30 μm and shows tilt of the sample. When using a Spirograph scan path, drift does not result in concentric rings but artifacts all over the image. As the scan pattern with 414 loops and 58k intersection results in a very high fitness of $\zeta = 447$ we achieve excellent correction of drift. In a final step a simple plane fit is subtracted to correct for tilt (FIG. 4 f)). The length of the recorded datasets for the DAS, MDAS, and Spiro-

graph were 1.5M, 164k, and 500k samples, respectively. Without any optimization for execution time the corresponding computation times were 5, 0.7, and 0.8 seconds using an Intel's Core 2 Duo P9500 processor and 4GB of system memory. Optimization of the code and use of distributed processing including multithreading, graphics processing units, or FPGAs could easily shorten the computation time in order to be used for a live display at video rates.

IV. CONCLUSIONS

Drift in the height signal during image acquisition is a common issue in all scanning probe techniques. In this work, we have shown that self-intersection scan paths can be used to effectively remove z-drift. Our method significantly outperforms a generalization of the typical flattening algorithms (piecewise linear fit). The proposed self-intersection algorithm is invariant under both tilt and sample geometry, and depends only on the scan path used to guarantee reliability. We proposed the fitness ζ which measures how likely the self-intersection method can discover the drift component. We demonstrate for varying scan patterns, scan size, and samples that an error-free topography is recovered when drift is corrected. As thermal drift occurs at low frequencies we find that the existence of a meager set of self-intersections with a ζ of 0.8 already is sufficient to correct the topographic data. Our technique is applicable to a large set of possible scan paths now being explored as fast-scanning alternatives to the raster pattern, and may be useful for real-time object tracking or cycloid scans along an object of interest [19, 20].

ACKNOWLEDGMENTS

This research is partially supported by NSF grants DMS-1045536 and CBET-0940417, UC Lab Fees Research grant 12-LR-236660, and the W. M. Keck Foundation. Work at Molecular Foundry was supported by the Office of Science, Office of Basic Energy Sciences, of the US Department of Energy under contract No. DE-AC02-05CH11231.

Appendix A: Image Reconstruction from Non-Raster Scan Datasets

With non-raster paths, visualizing the data generally requires it being put on a grid using inpainting techniques. We do this for the discrete innterpretation of $h(t)$ by first using bilinear interpolation weights for each sample to distribute it into the four nearest grid cells. Each grid cell of the boundary data $z(\vec{x})$ contains a weighted average height of nearby samples, and the fidelity parameter $\bar{\lambda}(\vec{x})$ is the sum of the weights for each grid cell.

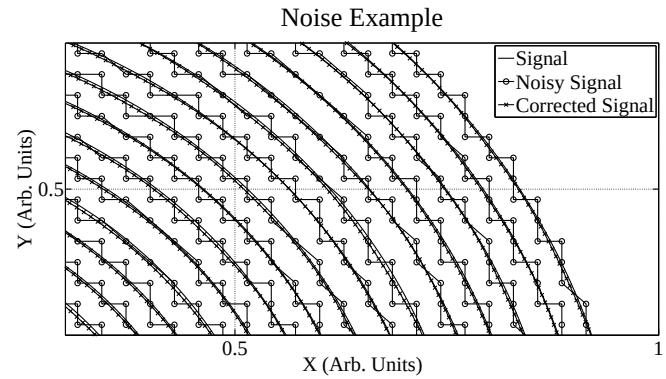

FIG. 5. The position sensor noise results from the signal being compressed into a low-bandwith signal and causes significant problems when detecting self-intersections. In this synthetic case, a scan path is measured with error and corrected with the convolution approach we descibe.

Therefore areas with more data will have higher fidelity, and cells with no data will reasonably have zero fidelity. $\bar{\lambda}(\vec{x})$ is then scaled by a user-specified amount indicating the strength of denoising. The grid points which get no data are filled using image inpainting. We herein use heat equation inpainting, which roughly corresponds to diffusing the information out to the unknown areas following the physical laws of heat flow. Briefly, this is performed as follows: the domain is separated into $D = \{\mathbf{x}(t) \big| t \in [0, T]\}$ where data is known and $\Omega \backslash D$ where data is missing, with Ω the entire sample area. The topography map $\hat{z}(\vec{x})$ is completed by solving the system of equations

$$\begin{aligned} 0 &= \Delta \hat{z} \ \text{ on } \Omega \backslash D \\ \bar{\lambda}(\hat{z} - z) &= \Delta \hat{z} \ \text{ on } D, \end{aligned}$$

which may be solved very quickly with multiscale methods [21], or by directly solving the finite difference system. $\bar{\lambda}(\vec{x})$, denoted with a bar to differentiate from λ of the algorithm, varies with space and determines the point-wise fidelity to $z(\vec{x}) = z(\vec{x}(t))$; a low value indicates strong denoising.

Appendix B: Finding Intersections on Quantized Datapoints

The data provided by the AFM has multiple sources of error, including a noise source due to the truncation of the position signal for finite bit representation (bitnoise). Let l_x and l_y be the smallest measurable distance represented by least significant bit for X and Y direction. The noisy signal received is described in terms of the path $< x(t), y(t) >$ by the function $< l_x \lfloor x(t)/l_x \rfloor, l_y \lfloor y(t)/l_y \rfloor >$ where $\lfloor \cdot \rfloor$ denotes the floor function. This is problematic for the intersection detection algorithm because the noise can generate many false intersections, throwing off the differences fitting energy. To combat this issue, we

convolve each of the position sensor signals with a Gaussian smoothing function with a standard deviation on the same order as l_x and l_y. An example of a synthetic signal being corrected is shown in FIG. 5. The focus of this work is not to remove this noise, however some processing has shown necessary for successful performance of the self-intersection method when considerable noise is present.

The recursive algorithm for finding intersections is as follows:

```
Function quad_tree_recur( list_of_segments,
           bounding_box, depth )
  If count(list_of_segments) < MIN_SIZE or
     depth > MAX_RECURSION then
    For all segments A,B in list_of_segments
      If A intersects B in bounding_box then
        Add intersection to global list.
      EndIf
    EndFor
  Else
    Subdivide bounding_box into equal area
       rectangles B1, B2, B3, B4.
    For n = 1, 2, 3, and 4
      Let sub_list =
    those of list_of_segments
            which intersect with Bn.
      Call quad_tree_recur( sub_list, Bn,
       depth + 1 );
    EndFor
  EndIf
EndFunction
```

This is called with an initial list containing all segments, a bounding box which contains all segments, and a recursion depth of zero. When the algorithm is implemented efficiently, it may process a million line segments in well under a second using Intel's Core 2 Duo P9500 processor and 4GB of system memory. The algorithm is highly parallelizable through the assignment of independent threads to different branches of the recursion. The ideal choice of the two constants in the algorithm depend on the system and language of implementation, though reasonable default values are 100 for MIN_SIZE and 8 for MAX_RECURSION. The running time can generally be reduced to $O(N \log N)$ by pre-processing the data with a quad-tree structure [22].

Appendix C: Derivation of Euler-Lagrange Equations

Our intention is to solve for the minimizing function f in equation (1) with representation restricted to the span of a basis of functions $\{\phi_i\}$ for $i = 1, 2, ..., N$. Thus, expanding f over this basis with coefficients c_i, notice that

$$f(t_{n,2}) - f(t_{n,1}) = \sum_{i=1}^{N} c_i \left[\phi_i(t_{n,2}) - \phi_i(t_{n,1})\right]$$

and proceed much the same way as the classical least-squares approximation. Recall that $d_j = d(t_{j,2}) - d(t_{j,1})$ is the error in height taken at the jth difference of the function being fit, $d(t)$.

Let $\vec{f}$ and $\vec{d}$ denote the length-M column vectors with, respectively, components $f(t_{j,2}) - f(t_{j,1})$ and d_j for $j = 1, 2, ..., M$. Let $\vec{c}$ be the length-N column vector formed by the coefficients c_i where $i = 1, 2, ..., N$. Denote by $\mathbf{A}$ the $M \times N$ matrix containing the basis differences at the crossing points with entries $\mathbf{A}_{ij} = \phi_j(t_{i,2}) - \phi_j(t_{i,1})$. The error on the differences is thus $||\vec{d} - \vec{f}||^2 = ||\vec{d} - \mathbf{A}\vec{c}||^2$. Define the $N \times N$ matrix $\mathbf{M}$ with entries

$$\mathbf{M}_{ij} = \int_0^1 \phi_i''(t)\phi_j''(t)\, dt.$$

By algebraic manipulation it may be shown that

$$\int_0^1 |f''(t)|^2 dt = \vec{c}^T \mathbf{M} \vec{c}.$$

Using these results, the functional in (1) may be now restated in terms of a minimization over $\vec{c}$

$$\min_{\vec{c}} \|\vec{d} - \mathbf{A}\vec{c}\|^2 + \lambda \vec{c}^T \mathbf{M} \vec{c}.$$

Differentiation with respect to $\vec{c}$ leads to the necessary optimality condition

$$L_\lambda \vec{c} = \left(\mathbf{A}^T\mathbf{A} + \lambda\mathbf{M}\right)\vec{c} = \mathbf{A}^T\vec{d} = \vec{h}.$$

The matrix L_λ is invertible and positive definite if $\lambda > 0$, in which case the solution is additionally guaranteed to be unique.

For the calculation of the ζ_λ fitness, we use the orthonormal Fourier basis and 500 basis elements. This is because the Fourier basis is simple to calculate, orthogonal, and because ζ_λ is not improved by using a more exotic basis choice.

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

[2] L. Gross, F. Mohn, N. Moll, P. Liljeroth, and G. Meyer, Science **325**, 1110 (2009).

[3] L. Gross, F. Mohn, N. Moll, B. Schuler, A. Criado, E. Guitin, D. Pea, A. Gourdon, and G. Meyer, Science

337, 1326 (2012).

[4] L. Gross, F. Mohn, N. Moll, G. Meyer, R. Ebel, W. M. Abdel-Mageed, and M. Jaspars, Nature Chemistry **2**, 821 (2010).

[5] B. Mokaberi and A. Requicha, Transactions on Automation Science and Engineering **3**, 199 (2006).

[6] M. D'Acunto and O. Salvetti, Pattern Recognition and Image Analysis **21**, 9 (2011).

[7] D. Croft, G. Shed, and S. Devasia, Journal of Dynamic Systems, Measurement, and Control **123**, 35 (2001).

[8] S. Devasia, E. Eleftheriou, and S. Moheimani, IEEE Transactions on Control Systems Technology **15**, 802 (2007).

[9] D. Ziegler, T. R. Meyer, R. Farnham, C. Brune, A. Bertozzi, and P. D. Ashby, Submitted to Nanotechnology (2013).

[10] M. Bertalmio, L. Vese, G. Sapiro, and S. Osher, in *Proceedings of Computer Vision and Pattern Recognition Conference*, Vol. 2 (IEEE Computer Society, 2003) pp. 707–712.

[11] G. Aubert and P. Kornprobst, *Mathematical Problems in Image Processing: Partial Differential Equations and the Calculus of Variations*, 2nd ed., edited by S. Antman, J. Marsden, and L. Sirovich, Applied Mathematical Sciences, Vol. 147 (Springer Science + Business Media, LLC, 2006).

[12] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, in *Proceedings of the 27th annual conference on Computer Graphics and Interactive Techniques*, SIGGRAPH '00 (ACM Press/Addison-Wesley Publishing Co., New York, NY, USA, 2000) pp. 417–424.

[13] V. Caselles, J. Morel, and C. Sbert, Transactions on Image Processing **7**, 376 (1998).

[14] T. Goldstein and S. Osher, SIAM Journal on Imaging Sciences **2**, 323 (2009).

[15] A. Chen, T. Wittman, A. Tartakovsky, and A. Bertozzi, Applied Mathematics Research eXpress **2011**, 182 (2011).

[16] I. A. Mahmood and S. O. R. Moheimani, Nanotechnology **20**, 365503 (2009).

[17] I. A. Mahmood, S. O. R. Moheimani, and B. Bhikkaji, Transactions on Nanotechnology **10**, 203 (2011).

[18] Y. K. Yong, S. O. R. Moheimani, and I. R. Petersen, Nanotechnology **21**, 365503 (2010).

[19] P. Chang and S. B. Andersson, in *Proceedings of the American Control Conference* (2008) pp. 3207–3212.

[20] P. Chang and S. Andersson, in *Proceedings of the 48th IEEE Conference on Decision and Control held jointly with the 28th Chinese Control Conference* (2009) pp. 8290–8295.

[21] W. Dahmen, B. Kleemann, S. Prssdorf, and R. Schneider, in *Boundary Element Topics*, edited by W. Wendland (Springer Berlin Heidelberg, 1997) pp. 189–219.

[22] R. A. Finkel and J. Bentley, Acta Informatica **4**, 1 (1974).