

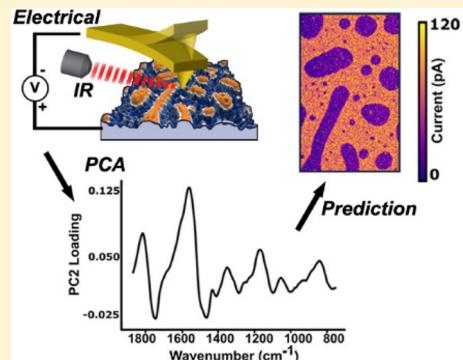
Identifying Nanoscale Structure–Function Relationships Using Multimodal Atomic Force Microscopy, Dimensionality Reduction, and Regression Techniques

Jessica Kong,^{1D} Rajiv Giridharagopal,^{1D} Jeffrey S. Harrison, and David S. Ginger^{*1D}

Department of Chemistry, University of Washington, Seattle, Washington 98195-1700, United States

Supporting Information

ABSTRACT: Correlating nanoscale chemical specificity with operational physics is a long-standing goal of functional scanning probe microscopy (SPM). We employ a data analytic approach combining multiple microscopy modes using compositional information in infrared vibrational excitation maps acquired via photoinduced force microscopy (PiFM) with electrical information from conductive atomic force microscopy. We study a model polymer blend comprising insulating poly(methyl methacrylate) (PMMA) and semiconducting poly(3-hexylthiophene) (P3HT). We show that PiFM spectra are different from FTIR spectra but can still be used to identify local composition. We use principal component analysis to extract statistically significant principal components and principal component regression to predict local current and identify local polymer composition. In doing so, we observe evidence of semiconducting P3HT within PMMA aggregates. These methods are generalizable to correlated SPM data and provide a meaningful technique for extracting complex compositional information that is impossible to measure from any one technique.



Chemical structure and composition are inextricably linked to the functional properties that give rise to materials with promising applications in systems from solar cells to batteries.^{1,2} Quantitative structure–function models are key to efficiently developing paradigms that enable materials discovery at a rate that outpaces the classical synthesis–characterization–theory approach.³ Images from multimodal atomic force microscopy (AFM) techniques which provide spatially resolved, quantitative structural and functional information make for particularly well-suited data sets for this purpose.⁴ Statistical learning tools can be leveraged to apply detailed quantitative analysis on structural and functional maps to develop insight for optimized materials and devices.

AFM-based techniques have become ubiquitous in functional nanoscale characterization.^{5–9} Imaging chemical composition on this scale, however, has been challenging; historically, such data were acquired using techniques including near-field scanning optical microscopy (NSOM) and tip-enhanced Raman spectroscopy (TERS), both of which can image chemical structure far beyond diffraction-limited optics.^{10–12} NSOM and TERS, however, impose a limitation of requiring careful optical alignment to “directly” collect optical information.^{11,13} To circumvent this issue, scanning probe methods like photoinduced force microscopy (PiFM)^{14,15} and photothermal infrared microscopy (PTIR, also called AFM-IR)^{16,17} have been developed to mechanically probe vibrational modes. These techniques use an infrared (IR) laser that spans the vibrational excitation range for many common organic materials. Instead of measuring the excitation via a change in absorption as in

NSOM, they measure the response via a mechanical response of the tip. Of these, PiFM has the distinction of being a nominally “non-contact” technique as opposed to a direct thermal expansion measurement in PTIR,¹⁸ although the ultimate result in either paradigm are IR spectra with nanoscale resolution measured via cantilever dynamics.

In brief, PiFM generates spatial maps with chemical contrast by measuring the time-integrated photoinduced gradient force between a tip and sample due to the interaction between their polarizabilities at wavenumbers corresponding to vibrational modes of different chemical species.¹⁴ PiFM can also generate hyperspectral images containing a photoinduced force infrared spectrum at each pixel, enabling visualization of absorption maps at all wavelengths. The current understanding of time-integrated photoinduced force^{19–23} and theory of PiFM can be found elsewhere.^{14,24}

One of the principle advantages of chemical imaging via PiFM is that correlated functional and chemical maps can be utilized to advance our understanding of the relationships between chemical structure and composition and electrical properties on the nanoscale. While hybrid systems that can obtain registered images of materials like AFM/mass spectrometry²⁵ are useful, they are more costly to implement and therefore motivate data-driven approaches to extract such information. Gleaning this level of insight requires combining

Received: April 2, 2018

Accepted: May 31, 2018

Published: May 31, 2018

domain knowledge and data science tools that in the field of scanning probe microscopy (SPM) have largely been applied to multidimensional SPM data sets.⁴ For example, principal component analysis (PCA) has been used as a tool to analyze band excitation SPM data²⁶ and extract statistically significant contributions to signals in general mode and general dynamic AFM.^{27–29} To gain physical insights from multidimensional first-order reversal curve current–voltage (FORC-IV) SPM data, techniques like *k*-means clustering and Bayesian linear unmixing have been applied.^{30,31} With the continued development of multidimensional SPM techniques, statistical learning tools will have an increasingly important role in processing, analyzing, and extracting information from progressively larger data sets.

A natural extension of the permeation of these data science techniques into the SPM community is to apply these tools to multimodal images with the goal of understanding the relationship between chemical structure and functional properties. As a simple demonstration of the utility of this approach, we apply statistical learning techniques to registered hyperspectral PiFM and conductive AFM (cAFM) images of a model polymer blend system comprising semiconducting poly(3-hexylthiophene) (P3HT) and insulating poly(methyl methacrylate) (PMMA). These polymers' distinct chemical fingerprints and the strong contrast between their electronic properties make this blend an ideal system for benchmarking analysis techniques to be used in increasingly complex systems with optoelectronic and chemical heterogeneity such as emerging mixed cation hybrid perovskites.^{32,33}

In this work, we employ affine transformations as an image registration technique, PCA as a dimensionality reduction tool, and principal component regression (PCR) as a statistical analysis method. This sequence of techniques is a representative analysis flow in deriving correlations between chemical and functional image data. (See Figure S1 for graphical overview.) We apply this approach to our model polymer blend to demonstrate the robustness of the registration and interpretability of the data resulting from PCA and PCR. Despite the linear nature and relative simplicity of PCR, it provides a sensible prediction of current with the larger deviations providing useful insight into the tomographic structure of the film. Importantly, this method is generalizable to correlated data sets from other chemical mapping tools.

The distinct chemical nature of P3HT and PMMA results in unique vibrational modes in their respective IR spectra. Figure 1a shows attenuated total reflectance Fourier transform IR (ATR-FTIR) spectra of PMMA, P3HT, and their blend in the range 1700–1760 cm^{−1}. (For full range spectra, see Figure S2.) PMMA exhibits a vibrational mode corresponding to a carbonyl (C=O) stretch at 1734 cm^{−1} that is also present in the blend but not in P3HT.³⁴ Figure 1b shows that within the range 780–840 cm^{−1} P3HT contains an aromatic C–H out-of-plane bending mode at 824 cm^{−1}.³⁵ As expected, this vibrational mode is also present in the blend but not in PMMA. Importantly, vibrational modes present in the bulk material are also present on the nanoscale, as shown in PiFM point spectra of films made of neat PMMA and P3HT in Figure 1c,d, indicating that they can be used as spectroscopic handles for nanoscale characterization.

According to the theoretical description of PiFM, the PiFM signal response originates from the real part of the polarizability,¹⁴ although photoinduced contributions can arise due to the imaginary part of the polarizability under certain imaging

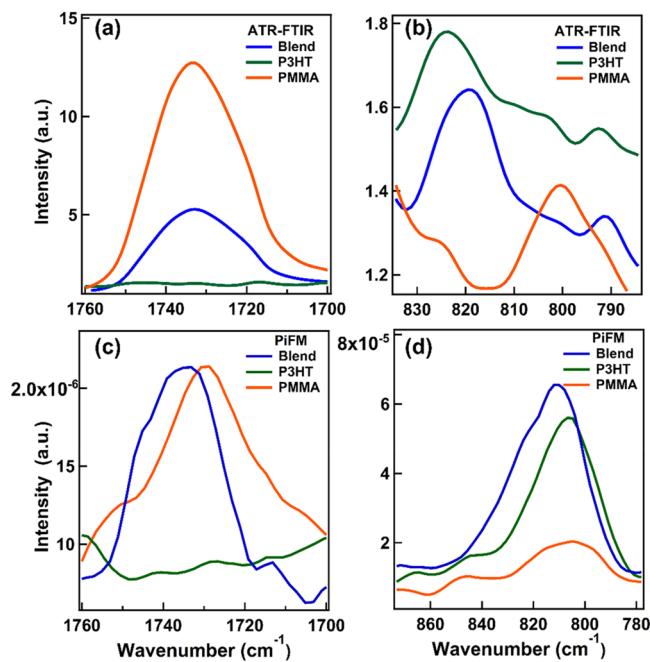


Figure 1. Comparison of attenuated total reflectance Fourier transform infrared (ATR-FTIR) (a,b) and photoinduced force microscopy (PiFM) point spectra (c,d) of neat polymer and blend films of poly(3-hexylthiophene) (P3HT) and poly(methyl methacrylate) (PMMA) showing that vibrational modes present in bulk ATR-FTIR spectra are also present in local PiFM spectra. (a,c) Spectral range of ~1700 to 1760 cm^{−1} corresponds to a C=O stretch in PMMA. (b,d) Spectral range of ~780 to 870 cm^{−1} corresponds to a C–H out-of-plane bending mode in P3HT.

conditions.²⁴ In contrast, the absorption in FTIR spectra is proportional to the imaginary part of the linear susceptibility.³⁶ Thus the signal in PiFM and FTIR spectra, at least in current understanding, probes different parts of a material's response to an electromagnetic field. Therefore, it is expected that there will be differences between local PiFM and macroscopic FTIR spectra.

For the purposes of this work, however, it is only important that the PiFM spectra of PMMA and P3HT both exhibit distinct spectral fingerprints, a condition we expect to hold generally for materials that have distinguishable FTIR spectra. For instance, Figure 1c shows the presence of a C=O stretch in neat PMMA and the polymer blend around 1728 and 1735 cm^{−1}, respectively; its absence in P3HT is in excellent agreement with the bulk ATR-FTIR spectra. However, the aromatic C–H bend of P3HT appears to shift from 824 cm^{−1} in the ATR-FTIR spectrum to 806 cm^{−1} in the PiFM point spectrum, as shown in Figure 1d. As P3HT exhibits no other vibrational modes in this region, it is reasonable to deduce that this peak corresponds to the aromatic C–H bend of this molecule.³⁵ While this shift places this vibrational mode in the vicinity of an unidentified vibrational mode of PMMA around 810 cm^{−1}, this peak is higher intensity in the P3HT blends and characteristic of P3HT.

We interpret the line-shape differences between our PiFM spectra compared with some of those previously reported in the literature (dispersive vs quasi-Lorentzian) based on the physical origin of the PiFM signal. Notably, Jahng et al. and Yang et al. have reported that there are both thermal and photoinduced force contributions in PiFM, with the former giving rise to dissipative, Lorentzian line shapes and the latter to

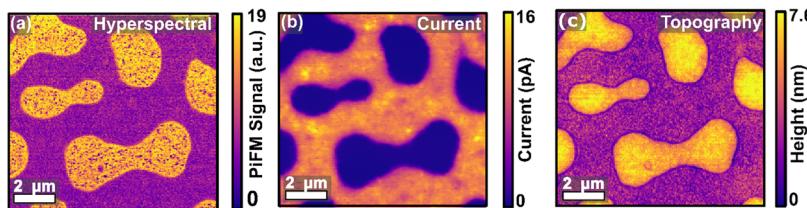


Figure 2. Registered multimodal atomic force microscopy data: (a) integrated hyperspectral PiFM signal over the range $760\text{--}1875\text{ cm}^{-1}$, (b) current under $+1\text{ V}$ sample bias, and (c) topography images of the same $10 \times 10\text{ }\mu\text{m}^2$ area of a P3HT/PMMA blend. In panel c, the bright aggregates are PMMA and the dark matrix is P3HT.

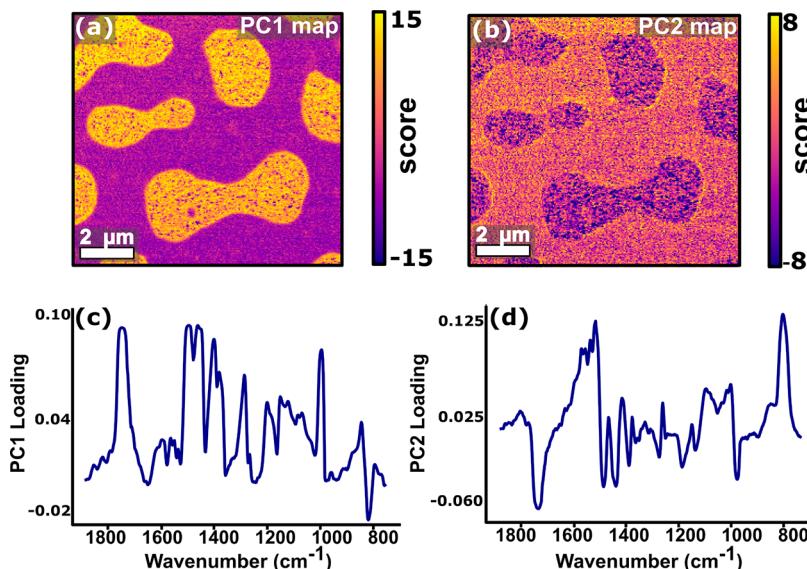


Figure 3. Principal component score images and principal components of the hyperspectral PiFM image as infrared (IR) spectra. Loading maps for the (a) first and (b) second eigenvectors showing that principal component analysis can be used to differentiate the PMMA from P3HT. IR spectra of the (c) first and (d) second eigenvectors that show that the principal components are not necessarily physically interpretable in that they maximize variance but may not correspond to physically observable spectra, a result that is well known for PCA.

dispersive lineshapes.^{24,37} However, if the sample's thermal relaxation rate is shorter than the interpulse time separation of the laser, then the sample does not have enough time to achieve significant thermal expansion and the photoinduced force gradient can be separated from the thermal contribution.²⁴ Because the thermal contribution is proportional to a wavelength-dependent thermal expansion coefficient, we expect there to be a wavenumber dependence to the thermal and photoinduced force contributions in most spectra. Therefore, a superposition of Lorentzian and dispersive line shapes can be anticipated under many common imaging conditions. Specifically, some studies have shown PiFM spectra that match well with those of FTIR in the $1300\text{--}1800\text{ cm}^{-1}$ window^{15,38} and therefore, at least in current understanding, those spectra would be interpreted as having a greater thermal contribution. In Figure 1a–d, the PiFM spectral peak at 823 cm^{-1} is not in agreement with those of FTIR, but appear more dispersive. Therefore, we interpret these peaks as having a greater photoinduced force gradient contribution.²⁴ The unidentified vibrational mode of PMMA around 800 cm^{-1} and carbonyl stretch around 1734 cm^{-1} are in good agreement with those of FTIR, so we interpret these peaks as having a greater thermal contribution.

To identify structure–function relationships, we take a hyperspectral PiFM image over the range $760\text{--}1875\text{ cm}^{-1}$ and current image at $+1\text{ V}$ sample bias with a gold tip. Prior to any data analysis, we register the hyperspectral PiFM and current

images using the affine transformation capabilities in Dipy.³⁹ In brief, translation, rigid, and affine transformations are sequentially optimized for the topography obtained in parallel with the cAFM scan. These transformations are obtained by automated maximization of the statistical dependence (as measured by mutual information) between the intensity distributions of the topography images obtained during hyperspectral PiFM and current imaging. Additional details regarding the image registration can be found in the SI and Jupyter notebook online in the associated repository.⁴⁰

We used separate instruments for PiFM and cAFM. In this case, image registration is important because differences due to variations in scan angle, piezo drift, and the different imaging modalities (intermittent contact for PiFM and contact mode in cAFM) can lead to distortions that can mask the sought-after structure–function relationships. We achieved successful colocalization across both microscopy methods with the following protocol: (1) we registered the images optically within several micrometers using visible alignment marks; (2) we imaged roughly the same region by locating the region of interest via large scans; and (3) we precisely aligned the images postcapture using affine transformations. (See Methods and SI section 3 for details.)

Figure 2 shows the registered integrated hyperspectral PiFM and current images along with the topography. The integrated hyperspectral PiFM image (Figure 2a) is obtained by integrating the area underneath the IR spectrum, spanning

760–1875 cm⁻¹ at every pixel. (See Figure S3 for associated full-range spectra at isolated points.) Contrast in this image is largely due to the presence of an intense C=O stretch in PMMA; areas rich in PMMA give a greater integrated signal due to this peak. Because P3HT is a hole conductor, the domains in the cAFM image (Figure 2b) correspond to hole extraction from the P3HT domains through the grounded tip.⁴¹ As expected, the current reflects the strong contrast in conductivity of this semiconductor–insulator blend, with the current over PMMA being significantly lower than that over P3HT. The topography of the blend (Figure 2c) shows the formation of PMMA aggregates (bright regions) within a P3HT matrix (dark background), a natural result of phase segregation within the blend.⁴²

We then utilize PCA to extract statistically important components from the hyperspectral image data cube of dimensionality 231 × 247 × 548. (See the SI for further information.) With PCA, the hyperspectral PiFM image is re-expressed as a linear combination of the principal components (PCs). Often, the variance of the original data captured by the first few principal components is large and quickly drops off for subsequent components, making PCA a useful data reduction technique by allowing one to work under the premise that most interesting information is contained in the first few, K , principal components and reduce the number of variables examined to K .⁴³

We apply PCA on the hyperspectral PiFM data set.⁴⁴ Figure 3 shows images of the first two scores along with their respective principal components. The score images (loading maps) (Figure 3a,b) are obtained by plotting the value of each principal component (Figure 3c,d) at each spatial pixel. The resulting score images (Figure 3a,b) can be used to differentiate between polymers, as seen in the loading maps, where the contrast in the image is the integrated intensity of the spectral loading at each pixel (Figure 3c,d).

Attempting to decipher a structure–function relationship from the principal component images of hyperspectral PiFM data alone, however, proves difficult. Within the framework of PCA, the loadings define the contributions of the original wavenumbers to the principal components. Loadings large in magnitude at particular wavenumbers signify greater contributions to the principal component. Thus, contrast seen in a score image can be largely attributed to wavenumbers with large loadings in the corresponding principal component. If one desires to interpret the principal components as spectra, then neither can be done so in the physical sense, as each has negative absorbance values. This outcome, namely, nonintuitive principal components without obvious physical analogues, is a well known result of PCA.^{30,31,45} We highlight it here merely as a reminder that while sometimes PCA of scanning-probe data may appear to decompose into desired physical components,^{26,39,46} one should not expect such results to always be the norm.

We note that numerous other techniques for spectral extraction such as independent component analysis (ICA), nonnegative matrix factorization (NMF), automatic target generation process (ATGP), pixel purity index (PPI), vertex component analysis (VCA), N-FINDR, and numerous others exist.^{47–53} Many methods available in software libraries rely on the presence of pure endmember spectra,^{47–49} whereas others require more overhead to implement. (See the SI for additional discussion regarding hyperspectral unmixing.) We plan to compare these approaches in more detail in the future, but here

we focus on what can be achieved with relatively simple and widely accessible PCA methods.

We now turn back to the PCA data. In principal component 1 (Figure 3c), peaks at 1739, 1453, and 1489 cm⁻¹, which are known vibrational modes of PMMA,³⁴ have the greatest loadings. In addition, there is a strong negative peak at 823 cm⁻¹, which is spectrally correlated with the P3HT aromatic C–H bending mode.³⁵ These peaks are therefore primarily responsible for the contrast observed in the first loading map (Figure 3a). Intuitively, pixels with high principal component 1 scores will have original spectra with strong absorption peaks at 1739, 1453, and 1489 cm⁻¹ corresponding to a C=O stretch, an unassigned vibrational mode, and O(CH₃) bend, respectively, in PMMA. In contrast, pixels with high principal component 1 scores will have original spectra with a weak absorption peak at 823 cm⁻¹. The peaks at 1739 and 823 cm⁻¹ are largely responsible for the contrast generated in the second loading map (Figure 3b), as indicated by principal component 2 (Figure 3d). We therefore broadly interpret principal component 2 as a lack of PMMA and presence of P3HT.^{34,35}

We extract physical insight by combining PCA data with cAFM information. We use PCR, which uses the principal components as the independent variables in a multiple linear regression model, to understand the relationship between local chemical composition (even without assigning composition fractions or endmember spectra) and electronic function.

To perform the analysis, we randomly selected 10% of the image pixels and fit a multiple linear regression model by regressing the electrical current on the first 10 principal components. With this model, we can reasonably predict the current of the remaining pixels from the hyperspectral information alone, as shown in Figure 4. (For quantitative measures of error and spatially resolved training and test data sets, see the SI.) Figure 4a shows that the predicted current is in good overall qualitative agreement with the measured current

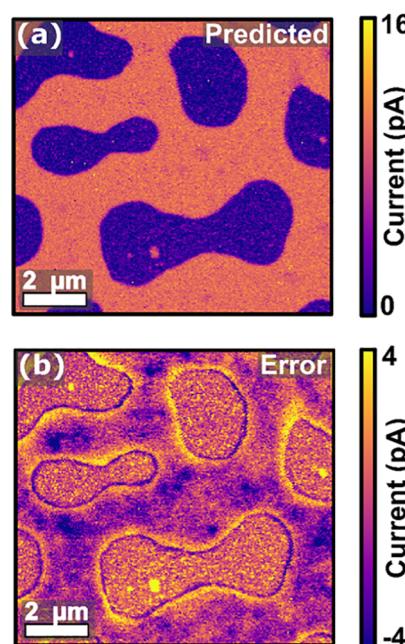


Figure 4. (a) Current as predicted with a principal component regression model obtained by regressing the electrical current onto the first 10 principal components. (b) Error image obtained by subtracting the real current, shown in Figure 1b, from the predicted current.

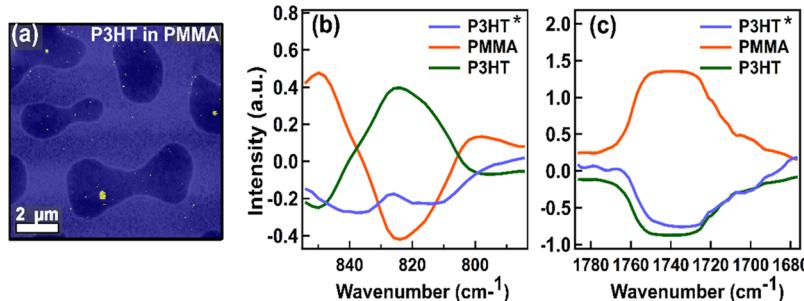


Figure 5. (a) Regions with predicted current that deviate more than three standard deviations from the real current (yellow pixels) overlaid on the hyperspectral PiFM image. Average spectrum, reconstructed from the first 10 principal components only, of pixels in panel a are in blue, PMMA aggregates in orange, and P3HT matrix in green for the range (b) 850–780 and (c) 1670–1790 cm^{-1} . The yellow pixels highlighted in panel a are likely to be P3HT regions within the PMMA aggregates that are electrically insulated.

(Figure 2b). Overall, the PMMA aggregates are predicted to be more insulating than the P3HT components. Importantly, this model also reasonably predicts the current images for different image sets of the same composition polymer blend. (See the SI for more information.)

Figure 4b shows the error image (measured current–predicted current) from the PCR. We speculate that regions where the PCR model underpredicts the current could be regions of better ordered P3HT with higher mobility.^{54–56} Interestingly, this image contains additional insight: regions that diverge most from the PCR model inform the identification of the polymer and appear to reveal tomographic structure indiscernible from PiFM or cAFM alone. Figure 5a shows the integrated hyperspectral PiFM image with the spatial distribution of the pixels (in yellow), where the PCR model overpredicts the real current by more than three standard deviations. (See the SI for statistical rationale.) Upon closer inspection, we see these pixels reside within regions that have negligible electrical current in cAFM, yet the PCR model predicts they are P3HT. The average spectrum, reconstructed from the first 10 principal components, for these pixels, denoted P3HT*, along with that of the PMMA aggregates and P3HT matrix, is shown in the spectral windows corresponding to the distinct vibrational modes of P3HT and PMMA in Figure 5b,c, respectively. (See Figure S8 for masks used to obtain average spectra.) The average spectrum for these pixels is similar to that of the P3HT matrix as it exhibits a negative peak at 1733 cm^{-1} (Figure 5c), distinct from the PMMA aggregates. In addition, it also has a peak at 823 cm^{-1} ; the weaker intensity is consistent with these regions being within PMMA.

This interpretation, combined with the presence of a vibrational mode at 823 cm^{-1} , suggests that these regions are likely P3HT regions trapped within the PMMA aggregates. The advantage of the PCA/PCR approach, as shown in Figure 5, is clear when compared with the seemingly straightforward approach of using the raw hyperspectral PiFM signal at 823 cm^{-1} to correlate with the cAFM current image. The result (see Figure S9) is less clear than the PCA/PCR approach because of the noise at 823 cm^{-1} , while PCA allows more use of the acquired information and also denoises the data. The pixels shown in Figure 5a exhibit the spectral signatures of P3HT regions but lack an electrically connected pathway for hole transport through the entire film. Also notable is that these regions (Figure 5a) are not directly correlated with regions of lower integrated hyperspectral PiFM intensity (see Figure S10) in the same region, as might be expected if this information was

contained within the top few nanometers accessible to PiFM.⁵⁷ The error from PCR, combined with our *a priori* knowledge about the semiconducting and insulating nature of the two polymers, suggests an explanation for why P3HT aggregates within PMMA do not appear in the cAFM image: it is due to the lack of an electrically connected pathway to the extracting contact. Such inferences are impossible with either the cAFM or PiFM alone and reveal a level of tomography that is unique to a data-driven approach to scanning probe regression of electrical information on chemical nanostructure.

Herein we have acquired hyperspectral PiFM and cAFM images of the same area within a PMMA/P3HT blend film. We show that while the local PiFM spectra do not necessarily match the macroscopic FTIR, they are still useful in differentiating between the polymers. We described a representative workflow to register and analyze hyperspectral PiFM and cAFM images with PCA and PCR. Using these statistical tools, we demonstrate that the two techniques yield complementary information that when combined, provide additional insight into the tomography of P3HT/PMMA that neither provides in isolation. While the PCR model yields a reasonable prediction of the current image, we note that these results show that there could be additional value in moving beyond basic PCA approaches in the future. Furthermore, to improve predictions of the PCR model, it is likely that nonlinear techniques such as multiple adaptive regression splines (MARS),⁵⁸ which can flexibly model more sophisticated relationships by including nonlinearity and interaction terms, will be necessary. Incorporating prior information about systems can also improve the results of dimensionality reduction and regression; therefore, these techniques may be more appropriate when applied within a Bayesian framework.⁵⁹

EXPERIMENTAL METHODS

Materials. All materials were purchased from Sigma-Aldrich and used as received unless otherwise stated.

A 1:1 ratio, 1 wt % solution of P3HT and PMMA blend was prepared following methods in Kergoat et al.⁶⁰ PMMA was dissolved in dichlorobenzene by stirring overnight. Subsequently, P3HT was added to the PMMA solution and stirred until completely dissolved. The solution was then filtered with a 0.45 μm polytetrafluoroethylene (PTFE) filter.

Polymer Film Fabrication. Films were fabricated on ITO substrates (TFD, Inc.), which were sequentially cleaned by sonicating in dilute detergent, acetone, and propan-2-ol and subsequently treated with oxygen plasma for 10 min. A thin layer of NiO was deposited by spin-coating in air a precursor

consisting of 129 mg nickel(II) acetylacetone dissolved in 5 mL of anhydrous ethanol with the addition of 50 μ L of 38% HCl at 5000 rpm at 11 000 rpm/s for 30 s. The NiO layer was annealed at 320 °C for 45 min in air. Subsequently, the polymer precursor solution was spin-coated onto the NiO at 1500 rpm and 1500 rpm/s for 30 s in air. Films were then heated to 120 °C for 20 min.

Hyperspectral Infrared Imaging. Prior to imaging, a razor blade was used to scratch an optically visible "X" onto the sample; the area imaged was near this "X." A Molecular Vista VistaScope coupled to a LaserTune QCL was used for hyperspectral photoinduced force hyperspectral imaging. The PiFM was operated in sideband mode where the first and second mechanical resonance frequencies at 320 kHz and 1.9 MHz of a Si cantilever with Pt coating were used to detect the photoinduced force and surface topography, respectively. The spatial resolution of the hyperspectral image acquired was 256 \times 256; the spectroscopic resolution was 2 cm^{-1} . During imaging, the laser was swept from 760 to 1875 cm^{-1} , which gave 559 spectroscopic dimensions. The first wavenumber of the spectrum was removed because it contained a software artifact. The last eight wavenumbers in the range 760 to 775 cm^{-1} were omitted from the analysis as the laser profile was unreliable in this range and no relevant spectroscopic information is contained here. The sample was housed in a vacuum shroud backfilled with nitrogen. After imaging, the tip was put in contact with the sample to scratch a 17 μm box around the area imaged. The 17 μm box was optically visible.

Conductive Atomic Force Microscopy. Conductive AFM measurements were taken with an Asylum Research MFP-3D BIO atomic force microscope (Oxford Instruments) installed on an inverted Nikon Eclipse Ti microscope and Table Stable vibration isolation stage. A Budget Sensors ContE-GB tip with spring constant \sim 0.2 N/m was positioned near the center of the optically visible 17 μm box. Large topography scans were taken until the area imaged in the PiFM was captured in a scan, after which the scan size was reduced and scan orientation was adjusted to closely match that of the hyperspectral PiFM image. We note that we had to take into consideration the scan direction and orientation of the two instruments. A surface bias of +1 V was applied as the tip was raster scanned across the surface of the sample to simultaneously measure the topography and current between the tip and sample. The sample was imaged in the dark using a closed flow cell under flowing nitrogen, conditions that have been shown to permit stable imaging of P3HT by cAFM.⁶¹ The resolution of these images was 256 \times 256.

FTIR. FTIR spectra were collected on a Thermo Fischer Nicolette 8700 with a nitrogen-cooled MCT-A detector with the following settings: 256 scans at a resolution of 1.928 cm^{-1} with an optical velocity of 0.9494, and an aperture of 3. Films for FTIR were made on KBr substrates.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: [10.1021/acs.jpclett.8b01003](https://doi.org/10.1021/acs.jpclett.8b01003).

Figure S1. Overview of work flow. Figure S2. ATR-FTIR and PiFM point spectra of neat PMMA, neat P3HT, and blend films along with the laser profile of the QCL in the region 800 to 1800 cm^{-1} . Figure S3. Optical microscopy image of the 17 μm box used for locating the region of

interest. Unregistered topography scans obtained during hyperspectral PiFM imaging and conductive AFM imaging. Figure S4. Various points in the region imaged and their associated photoinduced force microscopy point spectra in the range 775–1875 cm^{-1} . Figure S5. Extracted endmember components from NMF, ICA, VCA, ATGP, NFINDR, and PPI. Figure S6. Randomly selected pixels overlaid on the integrated hyperspectral PiFM image used to generate the training and test data sets. ([PDF](#))

AUTHOR INFORMATION

Corresponding Author

*E-mail: ginger@chem.washington.edu.

ORCID

Jessica Kong: [0000-0003-3009-6288](#)

Rajiv Giridharagopal: [0000-0001-6076-852X](#)

David S. Ginger: [0000-0002-9759-5447](#)

Notes

The authors declare no competing financial interest.

Images and associated workflow of analysis are available at https://github.com/kongjy/MultimodalAFMonPMMAP3HT_processing.

ACKNOWLEDGMENTS

This paper is primarily based on work supported by the Department of Energy BES under award number DESC0013957. J.K. is supported in part by the Clean Energy Institute and National Science Foundation Research Traineeship under award NSF DGE-1633216. We thank Ariel Rokem (The University of Washington eScience Institute) for help with image registration. We thank Daniel J. Graham (University of Washington) for helpful discussions regarding PCA. We also thank Derek B. Nowak (Molecular Vista) and William A. Morrison (Molecular Vista) for help with PiFM. Part of this work was conducted with instrumentation supported by the University of Washington Student Technology Fee at the Molecular Analysis Facility, a National Nanotechnology Coordinated Infrastructure site at the University of Washington, which is a user facility supported in part by the National Science Foundation (Grant ECC-1542101), the University of Washington, the Molecular Engineering & Sciences Institute, the Clean Energy Institute, and the National Institutes of Health.

REFERENCES

- (1) Balke, N.; Bonnell, D.; Ginger, D. S.; Kemerink, M. Scanning Probes for New Energy Materials: Probing Local Structure and Function. *MRS Bull.* **2012**, *37*, 633–637.
- (2) Tennyson, E. M.; Gong, C.; Leite, M. S. Imaging Energy Harvesting and Storage Systems at the Nanoscale. *ACS Energy Lett.* **2017**, *2*, 2761–2777.
- (3) Kalinin, S. V.; Sumpter, B. G.; Archibald, R. K. Big-Deep-Smart Data in Imaging for Guiding Materials Design. *Nat. Mater.* **2015**, *14*, 973–980.
- (4) Kalinin, S. V.; Strelcov, E.; Belianinov, A.; Somnath, S.; Vasudevan, R. K.; Lingerfelt, E. J.; Archibald, R. K.; Chen, C.; Proksch, R.; Laanait, N.; et al. Big, Deep, and Smart Data in Scanning Probe Microscopy. *ACS Nano* **2016**, *10*, 9068–9086.
- (5) Giridharagopal, R.; Rayermann, G. E.; Shao, G.; Moore, D. T.; Reid, O. G.; Tillack, A. F.; Masiello, D. J.; Ginger, D. S. Submicrosecond Time Resolution Atomic Force Microscopy for Probing Nanoscale Dynamics. *Nano Lett.* **2012**, *12*, 893–898.

- (6) Shao, G.; Glaz, M. S.; Ma, F.; Ju, H.; Ginger, D. S. Intensity-Modulated Scanning Kelvin Probe Microscopy for Probing Recombination in Organic Photovoltaics. *ACS Nano* **2014**, *8*, 10799–10807.
- (7) Salvador, M.; Vorpahl, S. M.; Xin, H.; Williamson, W.; Shao, G.; Karatay, D. U.; Hillhouse, H. W.; Ginger, D. S. Nanoscale Surface Potential Variation Correlates with Local S/Se Ratio in Solution-Processed CZTSSe Solar Cells. *Nano Lett.* **2014**, *14*, 6926–6930.
- (8) Jesse, S.; Rodriguez, B. J.; Choudhury, S.; Baddorf, A. P.; Vrejoiu, I.; Hesse, D.; Alexe, M.; Eliseev, E. A.; Morozovska, A. N.; Zhang, J.; et al. Direct Imaging of the Spatial and Energy Distribution of Nucleation Centres in Ferroelectric Materials. *Nat. Mater.* **2008**, *7*, 209–215.
- (9) Coffey, D. C.; Reid, O. G.; Rodovsky, D. B.; Bartholomew, G. P.; Ginger, D. S. Mapping Local Photocurrents in Polymer/Fullerene Solar Cells with Photoconductive Atomic Force Microscopy. *Nano Lett.* **2007**, *7*, 738–744.
- (10) Stöckle, R. M.; Suh, Y. D.; Deckert, V.; Zenobi, R. Nanoscale Chemical Analysis by Tip-Enhanced Raman Spectroscopy. *Chem. Phys. Lett.* **2000**, *318*, 131–136.
- (11) Dunn, R. C. Near-Field Scanning Optical Microscopy. *Chem. Rev.* **1999**, *99*, 2891–2928.
- (12) Bao, W.; Melli, M.; Caselli, N.; Riboli, F.; Wiersma, D. S.; Staffaroni, M.; Choo, H.; Ogletree, D. F.; Aloni, S.; Bokor, J.; et al. Mapping Local Charge Recombination Heterogeneity by Multidimensional Nanospectroscopic Imaging. *Science* **2012**, *338*, 1317–1321.
- (13) Schmid, T.; Opilik, L.; Blum, C.; Zenobi, R. Nanoscale Chemical Imaging Using Tip-Enhanced Raman Spectroscopy: A Critical Review. *Angew. Chem., Int. Ed.* **2013**, *52*, 5940–5954.
- (14) Jahng, J.; Fishman, D. A.; Park, S.; Nowak, D. B.; Morrison, W. A.; Wickramasinghe, H. K.; Potma, E. O. Linear and Nonlinear Optical Spectroscopy at the Nanoscale with Photoinduced Force Microscopy. *Acc. Chem. Res.* **2015**, *48*, 2671–2679.
- (15) Nowak, D.; Morrison, W.; Wickramasinghe, H. K.; Jahng, J.; Potma, E.; Wan, L.; Ruiz, R.; Albrecht, T. R.; Schmidt, K.; Frommer, J.; et al. Nanoscale Chemical Imaging by Photoinduced Force Microscopy. *Sci. Adv.* **2016**, *2*, e1501571.
- (16) Dazzi, A.; Prazeres, R.; Glotin, F.; Ortega, J. M. Local Infrared Microspectroscopy with Subwavelength Spatial Resolution with an Atomic Force Microscope Tip Used as a Photothermal Sensor. *Opt. Lett.* **2005**, *30*, 2388–2390.
- (17) Dazzi, A.; Prater, C. B.; Hu, Q.; Chase, D. B.; Rabolt, J. F.; Marcott, C. AFM-IR: Combining Atomic Force Microscopy and Infrared Spectroscopy for Nanoscale Chemical Characterization. *Appl. Spectrosc.* **2012**, *66*, 1365–1384.
- (18) Centrone, A. Infrared Imaging and Spectroscopy Beyond the Diffraction Limit. *Annu. Rev. Anal. Chem.* **2015**, *8*, 101–126.
- (19) Chaumet, P. C.; Nieto-Vesperinas, M. Time-Averaged Total Force on a Dipolar Sphere in an Electromagnetic Field. *Opt. Lett.* **2000**, *25*, 1065–1067.
- (20) Dholakia, K.; Zemánek, P. Colloquium: Gripped by Light: Optical Binding. *Rev. Mod. Phys.* **2010**, *82*, 1767–1791.
- (21) Arias-González, J. R.; Nieto-Vesperinas, M. Optical Forces on Small Particles: Attractive and Repulsive Nature and Plasmon-Resonance Conditions. *J. Opt. Soc. Am. A* **2003**, *20*, 1201–1209.
- (22) Jahng, J.; Brocious, J.; Fishman, D. A.; Huang, F.; Li, X.; Tammar, V. A.; Wickramasinghe, H. K.; Potma, E. O. Gradient and Scattering Forces in Photoinduced Force Microscopy. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2014**, *90*, 155417.
- (23) Ladani, F. T.; Potma, E. O. Dyadic Green's Function Formalism for Photoinduced Forces in Tip-Sample Nanojunctions. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95*, 205440.
- (24) Jahng, J.; Park, S.; Morrison, W. A.; Kwon, H.; Nowak, D.; Potma, E. O.; Lee, E. S. Nanoscale Spectroscopic Studies of Two Different Physical Origins of the Tip-Enhanced Force: Dipole and Thermal. 2017, *arXiv:1711.02479v2 [physics.optics]*. arXiv.org e-Print archive. <https://arxiv.org/abs/1711.02479>.
- (25) Ovchinnikova, O. S.; Tai, T.; Bocharova, V.; Okatan, M. B.; Belianinov, A.; Kertesz, V.; Jesse, S.; Van Berk, G. J. Co-Registered Topographical, Band Excitation Nanomechanical, and Mass Spectral Imaging Using a Combined Atomic Force Microscopy/Mass Spectrometry Platform. *ACS Nano* **2015**, *9*, 4260–4269.
- (26) Jesse, S.; Kalinin, S. V. Principal Component and Spatial Correlation Analysis of Spectroscopic-Imaging Data in Scanning Probe Microscopy. *Nanotechnology* **2009**, *20*, 085714.
- (27) Belianinov, A.; Kalinin, S. V.; Jesse, S. Complete Information Acquisition in Dynamic Force Microscopy. *Nat. Commun.* **2015**, *6*, 6550.
- (28) Somnath, S.; Collins, L.; Matheson, M. A.; Sukumar, S. R.; Kalinin, S. V.; Jesse, S. Imaging via Complete Cantilever Dynamic Detection: General Dynamic Mode Imaging and Spectroscopy in Scanning Probe Microscopy. *Nanotechnology* **2016**, *27*, 414003.
- (29) Collins, L.; Belianinov, A.; Somnath, S.; Balke, N.; Kalinin, S. V.; Jesse, S. Full Data Acquisition in Kelvin Probe Force Microscopy: Mapping Dynamic Electric Phenomena in Real Space. *Sci. Rep.* **2016**, *6*, 30557.
- (30) Strelcov, E.; Belianinov, A.; Hsieh, Y.-H.; Jesse, S.; Baddorf, A. P.; Chu, Y.-H.; Kalinin, S. V. Deep Data Analysis of Conductive Phenomena on Complex Oxide Interfaces: Physics from Data Mining. *ACS Nano* **2014**, *8*, 6449–6457.
- (31) Strelcov, E.; Belianinov, A.; Hsieh, Y.-H.; Chu, Y.-H.; Kalinin, S. V. Constraining Data Mining with Physical Models: Voltage- and Oxygen Pressure-Dependent Transport in Multiferroic Nanostructures. *Nano Lett.* **2015**, *15*, 6650–6657.
- (32) Moerman, D.; Eperon, G. E.; Precht, J. T.; Ginger, D. S. Correlating Photoluminescence Heterogeneity with Local Electronic Properties in Methylammonium Lead Tribromide Perovskite Thin Films. *Chem. Mater.* **2017**, *29*, 5484–5492.
- (33) de Quilettes, D. W.; Vorpahl, S. M.; Stranks, S. D.; Nagaoka, H.; Eperon, G. E.; Ziffer, M. E.; Snaith, H. J.; Ginger, D. S. Impact of Microstructure on Local Carrier Lifetime in Perovskite Solar Cells. *Science* **2015**, *348*, 683–686.
- (34) Dirlíkov, S.; Koenig, J. L. Infrared Spectra of Poly(Methyl Methacrylate) Labeled with Oxygen-18. *Appl. Spectrosc.* **1979**, *33*, 551–555.
- (35) Hotta, S.; Rughooputh, S. D. D. V.; Heeger, A. J.; Wudl, F. Spectroscopic Studies of Soluble Poly(3-Alkylthienylenes). *Macromolecules* **1987**, *20*, 212–215.
- (36) Mukamel, S. Quantum Electrodynamics. In *Principles of Nonlinear Optical Spectroscopy*; Oxford University Press: New York, 2006; pp 94–96.
- (37) Yang, H. U.; Raschke, M. B. Resonant Optical Gradient Force Interaction for Nano-Imaging and -Spectroscopy. *New J. Phys.* **2016**, *18*, 053042.
- (38) Gu, K. L.; Zhou, Y.; Morrison, W. A.; Park, K.; Park, S.; Bao, Z. Nanoscale Domain Imaging of All-Polymer Organic Solar Cells by Photo-Induced Force Microscopy. *ACS Nano* **2018**, *12*, 1473–1481.
- (39) Garyfallidis, E.; Brett, M.; Amirbekian, B.; Rokem, A.; Van Der Walt, S.; Descoteaux, M.; Nimmo-Smith, I. Dipy, a Library for the Analysis of Diffusion MRI Data. *Front. Neuroinformatics* **2014**, *8*, 1–17.
- (40) Kong, J. si_notebook. https://github.com/kongjy/MultimodalAFMonPMMAP3HT_processing/blob/master/jupyter_notebooks/si_notebook.ipynb.
- (41) Pingree, L. S. C.; Reid, O. G.; Ginger, D. S. Imaging the Evolution of Nanoscale Photocurrent Collection and Transport Networks during Annealing of Polythiophene/Fullerene Solar Cells. *Nano Lett.* **2009**, *9*, 2946–2952.
- (42) Kergoat, L.; Battaglini, N.; Miozzo, L.; Piro, B.; Pham, M.-C.; Yassar, A.; Horowitz, G. Use of Poly(3-Hexylthiophene)/Poly(Methyl Methacrylate) (P3HT/PMMA) Blends to Improve the Performance of Water-Gated Organic Field-Effect Transistors. *Org. Electron.* **2011**, *12*, 1253–1257.
- (43) Shlens, J. A Tutorial on Principal Component Analysis. 2014, *arXiv:1404.1100v1 [cs.LG]*. arXiv.org e-Print archive. <https://arxiv.org/abs/1404.1100v1>.
- (44) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-Learn: Machine Learning in Python. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

- (45) Belianinov, A.; He, Q.; Kravchenko, M.; Jesse, S.; Borisevich, A.; Kalinin, S. V. Identification of Phases, Symmetries and Defects through Local Crystallography. *Nat. Commun.* **2015**, *6*, 7801.
- (46) Collins, L.; Ahmadi, M.; Wu, T.; Hu, B.; Kalinin, S. V.; Jesse, S. Breaking the Time Barrier in Kelvin Probe Force Microscopy: Fast Free Force Reconstruction Using the G-Mode Platform. *ACS Nano* **2017**, *11*, 8717–8729.
- (47) Nascimento, J. M. P.; Dias, J. M. B. Vertex Component Analysis: A Fast Algorithm to Unmix Hyperspectral Data. *Ieee Trans. Geosci. Remote Sens.* **2005**, *43*, 898–910.
- (48) Winter, M. E. *N-FINDR: An Algorithm for Fast Autonomous Spectral End-Member Determination in Hyperspectral Data*; International Society for Optics and Photonics, 1999; Vol. 3753, pp 266–276.
- (49) Chang, C.-I.; Plaza, A. A Fast Iterative Algorithm for Implementation of Pixel Purity Index. *IEEE Geosci. Remote Sens. Lett.* **2006**, *3*, 63–67.
- (50) Ren, H.; Chang, C.-i. Automatic Spectral Target Recognition in Hyperspectral Imagery. *IEEE Trans. Aerosp. Electron. Syst.* **2003**, *39*, 1232–1249.
- (51) Erichson, N. B.; Mendible, A.; Wihlborn, S.; Kutz, J. N. Randomized Nonnegative Matrix Factorization. *Pattern Recognit. Lett.* **2018**, *104*, 1–7.
- (52) Hyvärinen, A.; Oja, E. Independent Component Analysis: Algorithms and Applications. *Neural Netw. Off. J. Int. Neural Netw. Soc.* **2000**, *13*, 411–430.
- (53) Paatero, P.; Tapper, U. Positive Matrix Factorization: A Non-Negative Factor Model with Optimal Utilization of Error Estimates of Data Values. *Environmetrics* **1994**, *5*, 111–126.
- (54) Wood, D.; Hancox, I.; Jones, T. S.; Wilson, N. R. Quantitative Nanoscale Mapping with Temperature Dependence of the Mechanical and Electrical Properties of Poly(3-Hexylthiophene) by Conductive Atomic Force Microscopy. *J. Phys. Chem. C* **2015**, *119*, 11459–11467.
- (55) Jiang, X.; Patil, R.; Harima, Y.; Ohshita, J.; Kunai, A. Influences of Self-Assembled Structure on Mobilities of Charge Carriers in π -Conjugated Polymers. *J. Phys. Chem. B* **2005**, *109*, 221–229.
- (56) Mauer, R.; Kastler, M.; Laquai, F. The Impact of Polymer Regioregularity on Charge Transport and Efficiency of P3HT:PCBM Photovoltaic Devices. *Adv. Funct. Mater.* **2010**, *20*, 2085–2092.
- (57) Murdick, R. A.; Morrison, W.; Nowak, D.; Albrecht, T. R.; Jahng, J.; Park, S. Photoinduced Force Microscopy: A Technique for Hyperspectral Nanochemical Mapping. *Jpn. J. Appl. Phys.* **2017**, *56*, 08LA04.
- (58) Friedman, J. H. Multivariate Adaptive Regression Splines. *Ann. Stat.* **1991**, *19*, 1–67.
- (59) Christian, P.; Robert, A. Defense of the Bayesian Choice. In *The Bayesian Choice*; Springer Texts in Statistics; Springer: New York, 2007; pp 507–530.
- (60) Kergoat, L.; Battaglini, N.; Miozzo, L.; Piro, B.; Pham, M.-C.; Yassar, A.; Horowitz, G. Use of Poly(3-Hexylthiophene)/Poly(Methyl Methacrylate) (P3HT/PMMA) Blends to Improve the Performance of Water-Gated Organic Field-Effect Transistors. *Org. Electron.* **2011**, *12*, 1253–1257.
- (61) Reid, O. G.; Munechika, K.; Ginger, D. S. Space Charge Limited Current Measurements on Conjugated Polymer Films Using Conductive Atomic Force Microscopy. *Nano Lett.* **2008**, *8*, 1602–1609.