

# Machine Learning Based Localization and Classification with Atomic Magnetometers

Cameron Deans,<sup>1</sup> Lewis D. Griffin,<sup>2</sup> Luca Marmugi,<sup>1,\*</sup> and Ferruccio Renzoni<sup>1</sup>

<sup>1</sup>*Department of Physics and Astronomy, University College London, Gower Street, London WC1E 6BT, United Kingdom*

<sup>2</sup>*Department of Computer Science, University College London, Gower Street, London WC1E 6EA, United Kingdom*



(Received 14 September 2017; revised manuscript received 24 November 2017; published 18 January 2018)

We demonstrate identification of position, material, orientation, and shape of objects imaged by a  $^{85}\text{Rb}$  atomic magnetometer performing electromagnetic induction imaging supported by machine learning. Machine learning maximizes the information extracted from the images created by the magnetometer, demonstrating the use of hidden data. Localization 2.6 times better than the spatial resolution of the imaging system and successful classification up to 97% are obtained. This circumvents the need of solving the inverse problem and demonstrates the extension of machine learning to diffusive systems, such as low-frequency electrodynamics in media. Automated collection of task-relevant information from quantum-based electromagnetic imaging will have a relevant impact from biomedicine to security.

DOI: [10.1103/PhysRevLett.120.033204](https://doi.org/10.1103/PhysRevLett.120.033204)

Electromagnetic induction imaging (EMI), or magnetic induction tomography (MIT), with atomic magnetometers (AMs) was recently demonstrated for mapping the electric conductivity of objects and imaging of metallic samples [1–4].

EMI and its classical counterpart with conventional magnetic field sensors [5] rely on the detection of the ac magnetic field generated by eddy currents excited in media. This poses severe problems for image reconstruction, particularly in cluttered contexts or with low-conductivity specimens. The inherently diffusive and nonlinear nature of low-frequency electrodynamics in media makes conventional ray-optics analysis impossible. Consequently, back-projection approaches [6] are of limited use. Furthermore, the solution of the inverse problem for low-frequency electromagnetics is ill posed, undetermined, and computationally challenging [7]. Ultimately, these limitations reduce the attainable information from EMI and its spatial resolution.

In this Letter, we propose and demonstrate machine learning (ML) [8] as a method for enhancing the EMI capabilities and circumventing the problem of image reconstruction and interpretation. ML has thus far been applied in a wealth of fields [9–17]. ML-aided security screening in the  $X$  band [18,19] and biomedical imaging have been widely demonstrated [20,21], as well as image reconstruction through scattering media in the optical band [22,23]. All these applications to well-established imaging technologies are underpinned by linear systems, with raylike propagation.

Here we present proof-of-concept demonstrations of EMI by an AM with metallic and nonmetallic samples. Their localization and material, orientation, and shape classification from low-resolution images is aided by ML. AM-EMI supported by ML maximizes the information obtained from the images and provides relevant data for specific tasks, without requiring the inverse problem. This improves or

enables identification of critical features, such as structural defects in nondestructive evaluation [1,3], concealed threats in screening applications [4], or conductivity anomalies in biological tissues [24]. New perspectives open up for high-performance imaging based on EMI, in particular, with AMs, with a relevant impact on science and society.

In our setup, imaging relies on the position-resolved detection of the secondary magnetic field produced by eddy currents induced in the sample of interest by an ac magnetic field ( $B_{\text{rf}}$ ), oscillating at  $\nu_{\text{rf}} = 200$  kHz with amplitude  $1.9 \times 10^{-8}$  T. This primary field is produced by a pair of 180 mm diameter Helmholtz coils. The imaged object's response is detected by a  $^{85}\text{Rb}$  radio-frequency optically pumped AM, coherently driven by the same field ( $B_{\text{rf}}$ ), and described in [1,4]. We recall that a  $\sigma^+$ -polarized beam resonant with the  $F = 3 \rightarrow F' = 4$  transition of the  $^{85}\text{Rb}$   $D_2$  line aligns the atomic spins via optical pumping.  $B_{\text{rf}}$  excites a time-varying transverse component of the atomic polarization, which is perturbed by the secondary field produced by the objects of interest. The effects of this are imprinted in the polarization plane rotation of a  $\pi$ -polarized laser probe (Faraday rotation), detuned by +420 MHz. Further details can be found in the Supplemental Material [25].

Samples are moved with respect to the AM by a computer-controlled  $XY$  stage in the  $(x, y)$  plane [Figs. 1(a)–1(b)], 30 mm above the AM and enclosed by the rf Helmholtz coils. Four images per scan are obtained by simultaneously measuring the in-phase ( $X$ ) and quadrature ( $Y$ ) components of the AM output, as well as its amplitude ( $R$ ) and phase-lag ( $\Phi$ ). Active environmental and stray fields control [4] ensure continuous and consistent operation for more than 240 consecutive hours in an unshielded environment, without requiring any human intervention. To maintain consistency over long measurement time scales, the AM sensitivity is

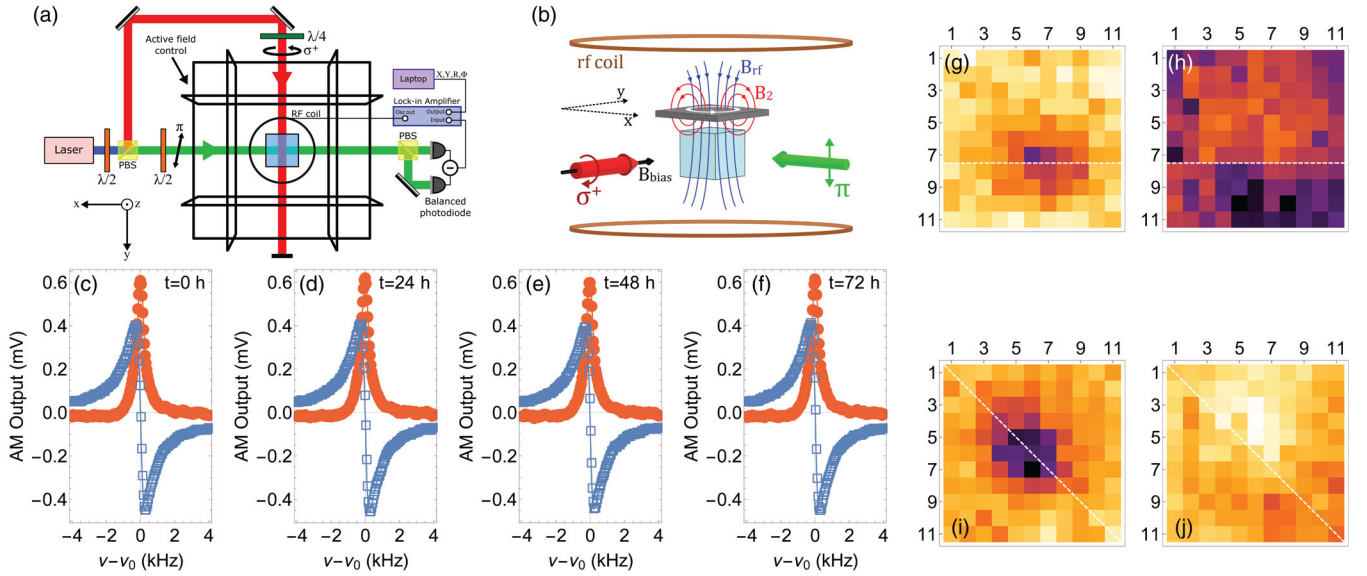


FIG. 1. (a) Sketch of the experimental apparatus; PBS indicates a polarizing beam splitter. (b) Details of the imaging technique. Samples are moved by an  $XY$  stage in the plane marked by dotted arrows.  $B_{rf}$  is the rf field driving the AM and inducing eddy currents (in white). The EMI secondary field is drawn in red. (c)–(f) Frequency response of the rf AM, in operational conditions, at  $\nu_0 = 200$  kHz: in-phase ( $X$ , filled circles) and quadrature ( $Y$ , empty squares) components. Traces were acquired in the same conditions every 24 h. (g)–(j) Examples of low-resolution EMI images used for this work: Al rectangle ( $50 \times 25 \times 3$  mm<sup>3</sup>,  $\sigma = 3.77 \times 10^7$  S/m), obtained with 8 mm scan step size at 200 kHz. (g) Amplitude ( $R$ ) image at  $0^\circ$  orientation. (h) Corresponding phase ( $\Phi$ ) image. (i)  $R$  image at  $+45^\circ$  orientation. (j) Corresponding  $\Phi$  image. Raw data displayed, with the same color scale for  $R$  and  $\Phi$ . A thin dashed line marks the orientation of the rectangle main axis.

purposely decreased to  $3.3 \times 10^{-11}$  T/ $\sqrt{\text{Hz}}$  [25], but it is stable throughout the measurement campaigns [Figs. 1(c)–1(f)].

Each image is an  $11 \times 11$  ( $22 \times 22$ ) matrix, where each point corresponds to an 8 mm (4 mm) step of the  $XY$  stage. The center position of the sample is randomly distributed within the interval  $[0, 30] \times [0, 30]$  mm<sup>2</sup>, with a resolution of  $10^{-3}$  mm. The large rf coils and the long  $XY$  stage step severely reduced the images' resolution and contrast (see also [25]), beyond the point where the samples and their details can be directly identified from the images. As an example, we present images of an Al rectangle obtained at 200 kHz with 8 mm scan step size in Figs. 1(g)–1(j). We show that ML compensates for such degradation, shifting the burden of image recognition from the observer to the computer and from the imaging phase to the training process.

We have tested the following combined tasks: (i) classification of four materials and localization, (ii) classification of four orientations and localization, and (iii) classification of four shapes and localization. In each case, 255 sets of images of a same class (e.g., same material) were collected, with random position. Each set comprises four images, namely,  $X$ ,  $Y$ ,  $R$ , and  $\Phi$ . In total, 4080 data files per task were available for training. Sets of 160 *blind* data files were acquired to test more realistic conditions: the blind images were acquired in different measurement runs, without any direct correlation to the training data sets. Details of these images were hidden during ML analysis.

ML algorithms were tested, tuned, and validated using random splits of the images, grouped into disjoint testing and training subsets (cross validation) [31,32]. The latter had an arbitrarily chosen size of  $N$ . The optimum parameters were then applied to the corresponding blind data sets. Deviations from the ground truth (i.e., the set of nominally true values) were measured with the root-mean-square error (RMSE) of the distance between the predicted position and the ground truth positions for localization. Failure rate ( $\epsilon$ ) was introduced for classifications [25]. RMSE and  $\epsilon$  were then investigated as a function of the number of training images  $N$ . Results were compared to baseline and ceiling performance. Baseline performance is the error obtained with the best blind guess. For localization, this corresponds to  $\text{RMSE}_l = 12.38$  mm, for classifications to  $\epsilon_l = 75\%$ . The ceiling performance is the minimum error achievable given the characteristics of data and their variability. For localization, this corresponds to half the  $XY$  stage step ( $\text{RMSE}_h = 4$  mm, unless otherwise stated), for classifications, to  $\epsilon_h = 0\%$ . The corresponding 95% confidence intervals (CIs) were calculated via bootstrap resampling with replacement of the blind data, to avoid the assumption that errors were normally distributed [33].

Among the available data sets, we found the  $\{R, \Phi\}$  combination to be the most effective in the explored tasks [25]. It is therefore presented as the primary combination in the following analysis. Linear regressors (LRs), nearest neighbors, random forest, neural networks, and support vector machine (SVM) [34] algorithms were tested. Linear regression was found to uniformly perform better for the

localization tasks. Radial basis function (RBF) kernel SVM was the best performing for the classification tasks [25]. RBF SVM parameters, namely, margin softness and Gaussian width, were tuned by cross validation [35] within the training data.

For material classification and localization, four squares of equal size, made of copper, aluminium, Ti90/Al6/V4 alloy, and graphite, respectively, were tested. Their conductivities vary between  $5.98 \times 10^7$  and  $7.30 \times 10^4$  S/m, making the graphite the lowest conductivity and the first nonmetallic sample to be imaged with EMI performed by an AM [1,3,4,36].

Cross validation of localization with LR performed better than baseline performance beyond only  $N = 8$  training images: RMSE = 9.7 mm with CI = [7.2, 12.3] mm. The ceiling performance was achieved with  $N = 64$  training images. A final RMSE = 3.1 mm with CI = [3.0, 3.2] mm was obtained with  $N = 2048$ . We underline that this residual error is 2.6 times smaller than the  $XY$  step size: this constraint is surpassed by ML.

Material classification with RBF SVM produced an error  $\epsilon = 2\%$  with  $N = 2048$  training images and of only 12% with  $N = 8$  [25]. This result is aided by the different EMI signals of the four materials: the amplitude and the phase lag of the secondary field are proportional to the specimen's conductivity [5]. The difference in the EMI images demonstrates the capability of our system of discriminating different materials. However, it also makes material identification—even in nonoptimum conditions—potentially achievable by a human operator (see also [25]).

We therefore focus on more challenging tasks in the following. In these cases, a blind data set is also collected at a later stage and tested with the ML algorithms tuned during training.

Task (ii) comprises varying position and orientation. In this case, Al rectangles with aspect ratio 2:1 were aligned in four different orientations, namely  $\{-45^\circ, 0^\circ, +45^\circ, +90^\circ\}$  with respect the system's quantization axis ( $\hat{y}$ ). Two rectangles were tested, sample A ( $50 \times 25 \times 3$  mm<sup>3</sup>) and sample B ( $40 \times 20 \times 3$  mm<sup>3</sup>). For A, the  $XY$  stage step size was maintained at 8 mm. For B, two settings at 8 and 4 mm were used. Figure 2 shows the results in terms of classification errors  $\epsilon$ .

Cross validation was successfully applied in each case. A consistent improvement in performance is observed for increases in training samples and reduction in step size. Sample A error rapidly converges toward  $\epsilon \leq 20\%$  and reaches  $\epsilon = 4\%$  with  $N = 512$ . This performance is obtained also with a  $N = 40$  blind data set, where we have obtained a residual  $\epsilon = 20\%$ , with CI = [8%, 32%]. We note that such a task, even in the most favorable case of sample A, would be virtually impossible for a human observer [see Figs. 1(g)–1(j)].

A similar behavior for cross validation is observed also in the worst case of sample B at 8 mm step size:  $\epsilon < \epsilon_l = 75\%$  with  $N = 8$  training images. However, overall weaker

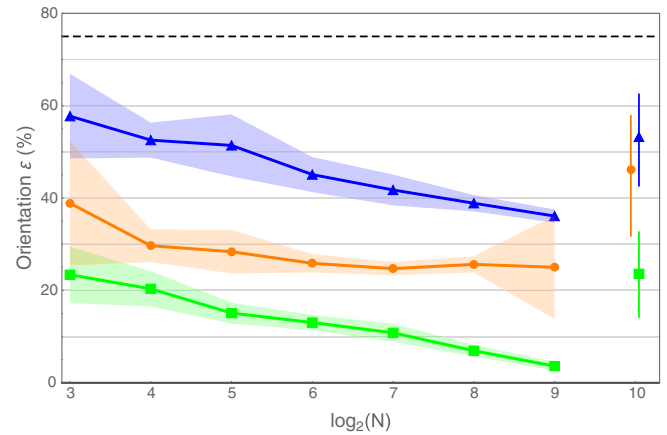


FIG. 2. Orientation  $\epsilon$  versus  $N$ : (green squares) sample A, 8 mm step; (blue upward triangles) sample B, 8 mm step; (orange disks) sample B, 4 mm step. Isolated markers are the blind data set results. The shaded area marks the 95% CI of the mean cross-validated performance. The error bars for the blinded results indicate the 95% CIs of system performance given the limited amount of blinded data. The dashed line marks the baseline performance ( $\epsilon_l = 75\%$ ). The thick horizontal line marks the ceiling performance ( $\epsilon_h = 0\%$ ).

performance of sample B is observed. This is related to the size of the test object: sample B is 1.56 times smaller than sample A. Consequently, the level of EMI signals obtained is lower, thus reducing the images' contrast. Statistical fluctuations explain the larger CI of the blind B data set with 4 mm steps: this set has only  $N = 20$  acquisition. Such small random variations in blind data sets were identified by the ML algorithms only. Nevertheless, classification via ML is not hampered: the blind results for both samples are well below the random choice level. Furthermore, results are improved by reducing the  $XY$  step size. The scan density impact is highlighted in Fig. 2: the 2 times shorter step produces a reduction between 50% and 25% of  $\epsilon$ .

Interestingly, the different orientations and the different  $\epsilon$  do not affect the localization prediction (as shown in Fig. 3). Even when the blind classification is challenged, as in the case of sample B, the localization RMSE becomes smaller than the ceiling performance with only  $N = 256$  (8 mm data set). The trend is confirmed by the analysis of the 4 mm step size data set with sample B (RMSE<sub>l</sub> = 2 mm, in this case). A RMSE = 4 mm is obtained with  $N = 512$ , and—as a further confirmation of statistical fluctuations in blind images—the blind data set performed better than the cross-validation sets (RMSE = 2.6 mm, CI = [1.9, 3.2] mm).

For the last task, four shapes of Al, 3 mm thick, were used: a 50 mm square, a 50 mm disk, a  $50 \times 25$  mm<sup>2</sup> rectangle, and an isosceles triangle, base 50 mm and height 50 mm.

Localization performs very well, aligned with previous results, confirming its substantial independence from other degrees of freedom. We obtained RMSE = 4.8 mm, CI = [4.6, 5.0] mm with  $N = 512$ . The  $N = 40$  blind data set gave a consistent RMSE = 4.5 mm, CI = [3.6, 5.4] mm [25].

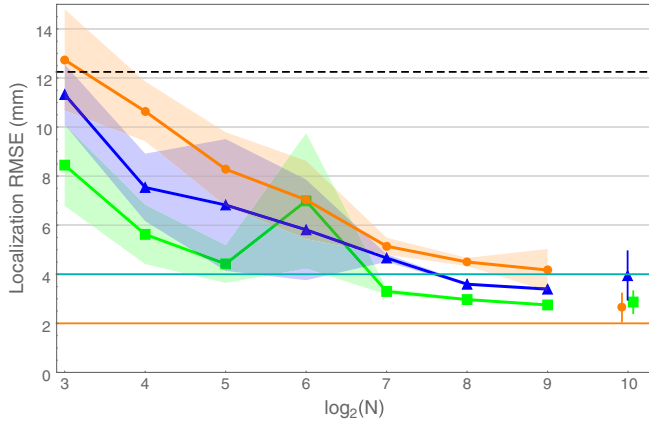


FIG. 3. Localization RMSE versus  $N$ : (green squares) sample A, 8 mm step; (blue upward triangles) sample B, 8 mm step; (orange disks) sample B, 4 mm step. Isolated markers are the blind data set results. The shaded area marks the 95% CI of the mean cross-validated performance. The error bars for the blinded results indicate the 95% CIs of system performance given the limited amount of blinded data. The dashed line marks the baseline performance ( $\text{RMSE}_l = 12.38$  mm). The thick horizontal lines mark the ceiling performance  $\text{RMSE}_h = 4$  and 2 mm, respectively.

Shape classification results are shown in Fig. 4. Classification outperforms chance with only  $N = 8$  images. The best performance,  $\varepsilon = 3\%$ ,  $\text{CI} = [2\%, 4\%]$  is obtained with  $N = 512$ . Consistent performance is observed with the  $N = 40$  blind data set:  $\varepsilon = 22\%$ ,  $\text{CI} = [9\%, 35\%]$ .

Virtually no incorrect attributions were observed between the square and the disk, whereas around 6%, symmetric confusion was observed between the rectangle and the triangle (Table I). This is attributed to the equal surface ( $1250 \text{ mm}^2$ ) of the two samples. Given that the rf

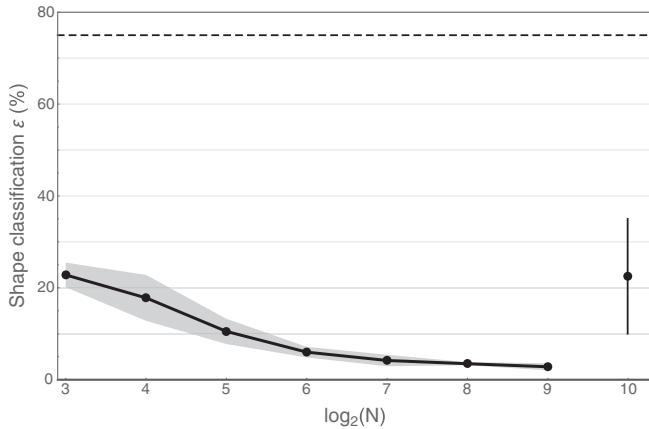


FIG. 4. Shape classification  $\varepsilon$  versus  $N$ . The isolated mark is the blind data set result. The shaded area marks the 95% confidence interval of the mean cross-validated performance. The error bars for the blinded results indicate the 95% CIs of system performance given the limited amount of blinded data. The dashed line marks the baseline performance ( $\varepsilon_l = 75\%$ ). The thick horizontal line marks the ceiling performance ( $\varepsilon_h = 0\%$ ).

TABLE I. Shape classification confusion matrix based on  $N = 256$ . Rows are predictions.

	Square (%)	Disk (%)	Rectangle (%)	Triangle (%)
Square	99.5	0.5	0.0	0.0
Disk	0.2	99.8	0.0	0.0
Rectangle	0.0	0.2	93.6	6.2
Triangle	0.0	0.0	6.7	93.3

Helmholtz coils' diameter is about three times larger than the samples' cross section, eddy currents are excited within their entire surface. With the four samples having the same thickness, the surface area becomes relevant for discrimination.

In light of this, we note that shape classification is successful even with same areas. This clearly shows that the ML is capable of implicitly using and integrating all the information, including those hidden in the samples' area and in other parameters, when areas are not effective, to improve the classification.

In conclusion, we have demonstrated that EMI with AMs provides information for identifying position, material, orientation, and shape of test samples, thanks to ML support. ML was applied for the first time to diffusive, non-ray-optics images produced by a  $^{85}\text{Rb}$  rf AM operating in EMI modality. This demonstrates the suitability of ML for diffusive and complex physical systems. Localization better than the smallest pixel of the image was demonstrated. Classification by ML allowed the identification of low-conductivity materials such as graphite, imaged for the first time with an AM. ML has also revealed the use of hidden information, such as the deduced area of the samples, without any specific input. Based on the present results, no evidence preventing scaling to larger number of classes was found, provided that the chosen imaging resolution matches the scale of the relevant features.

Our findings demonstrate that the conventional approach for EMI/MIT image reconstruction (the electromagnetic inverse problem) can be efficiently circumvented by ML, with relevant impact on computational burden. This opens up new perspectives for EMI/MIT with atomic magnetometers.

This work was funded by the UK Quantum Technology Hub in Sensing and Metrology, Engineering and Physical Sciences Research Council (EPSRC) (EP/M013294/1). C.D. acknowledges support from the Engineering and Physical Sciences Research Council (EPSRC) (EP/L015242/1).

\*l.marmugi@ucl.ac.uk

- [1] C. Deans, L. Marmugi, S. Hussain, and F. Renzoni, *Appl. Phys. Lett.* **108**, 103503 (2016).
- [2] A. Wickenbrock, S. Jurgilas, A. Dow, L. Marmugi, and F. Renzoni, *Opt. Lett.* **39**, 6367 (2014).
- [3] A. Wickenbrock, N. Leefer, J.W. Blanchard, and D. Budker, *Appl. Phys. Lett.* **108**, 183507 (2016).

- [4] C. Deans, L. Marmugi, and F. Renzoni, *Opt. Express* **25**, 17911 (2017).
- [5] H. Griffiths, *Meas. Sci. Technol.* **12**, 1126 (2001).
- [6] Z. Dek, J. M. Grimm, M. Treitl, L. L. Geyer, U. Linsenmaier, M. Krner, M. F. Reiser, and S. Wirth, *Radiology* **266**, 197 (2013).
- [7] R. Merwa, K. Hollaus, P. Brunner, and H. Scharfetter, *Physiol. Meas.* **26**, S241 (2005).
- [8] M. I. Jordan and T. M. Mitchell, *Science* **349**, 255 (2015).
- [9] J. C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, *Phys. Rev. Lett.* **108**, 253002 (2012).
- [10] Q. Wei, R. G. Melko, and J. Z. Y. Chen, *Phys. Rev. E* **95**, 032504 (2017).
- [11] E. D. Cubuk, S. S. Schoenholz, J. M. Rieser, B. D. Malone, J. Rottler, D. J. Durian, E. Kaxiras, and A. J. Liu, *Phys. Rev. Lett.* **114**, 108001 (2015).
- [12] S.-M. Zhou, F. Fernandez-Gutierrez, J. Kennedy, R. Cooksey, M. Atkinson, S. Denaxas, S. Siebert, W. G. Dixon, T. W. O'Neill, E. Choy, C. Sudlow, UK Biobank Follow-up and Outcomes Group, and S. Brophy, *PLoS One* **11**, 1 (2016).
- [13] M. W. Libbrecht and W. S. Noble, *Nat. Rev. Genet.* **16**, 321 (2015).
- [14] X.-D. Cai, D. Wu, Z.-E. Su, M.-C. Chen, X.-L. Wang, L. Li, N.-L. Liu, C.-Y. Lu, and J.-W. Pan, *Phys. Rev. Lett.* **114**, 110504 (2015).
- [15] Y. Zhang and E.-A. Kim, *Phys. Rev. Lett.* **118**, 216401 (2017).
- [16] H.-K. Lau, R. Pooser, G. Siopsis, and C. Weedbrook, *Phys. Rev. Lett.* **118**, 080501 (2017).
- [17] M. Bukov, A. G. Day, D. Sels, P. Weinberg, A. Polkovnikov, and P. Mehta, [arXiv:1705.00565](https://arxiv.org/abs/1705.00565).
- [18] T. W. Rogers, N. Jaccard, E. J. Morton, and L. D. Griffin, *J. X-Ray Sci. Technol.* **25**, 33 (2017).
- [19] N. Jaccard, T. W. Rogers, E. J. Morton, and L. D. Griffin, *J. X-Ray Sci. Technol.* **25**, 323 (2017).
- [20] A. Abós, H. C. Baggio, B. Segura, A. I. García-Díaz, Y. Compta, M. J. Martí, F. Valldeoriola, and C. Junqué, *Sci. Rep.* **7**, 45347 (2017).
- [21] U. S. Kamilov, I. N. Papadopoulos, M. H. Shoreh, A. Goy, C. Vonesch, M. Unser, and D. Psaltis, *Optica* **2**, 517 (2015).
- [22] G. Satat, M. Tancik, O. Gupta, B. Heshmat, and R. Raskar, *Opt. Express* **25**, 17466 (2017).
- [23] A. Sinha, J. Lee, S. Li, and G. Barbastathis, *Optica* **4**, 1117 (2017).
- [24] L. Marmugi and F. Renzoni, *Sci. Rep.* **6**, 23962 (2016).
- [25] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevLett.120.033204>, which also includes Refs. [26–30], for further details on experimental setup, imaging performance, data analysis, and material and shape data sets.
- [26] I. M. Savukov, S. J. Seltzer, M. V. Romalis, and K. L. Sauer, *Phys. Rev. Lett.* **95**, 063004 (2005).
- [27] W. Chalupczak, R. M. Godun, S. Pustelny, and W. Gawlik, *Appl. Phys. Lett.* **100**, 242401 (2012).
- [28] Goodfellow, Carbon foil, C000470 datasheet, <https://tinyurl.com/yajm38p8>.
- [29] R. C. Weast, D. Lide, M. Astle, and W. Beyer, *1990 CRC Handbook of Chemistry and Physics* (CRC Press, Boca Raton, FL, 1989).
- [30] Goodfellow, Titanium/aluminium/vanadium foil, Ti010500 datasheet, <https://tinyurl.com/y7btntgc>.
- [31] C. Bishop, *Pattern Recognition and Machine Learning*, 1st ed. (Springer-Verlag, New York, 2006).
- [32] K. Hansen, G. Montavon, F. Biegler, S. Fazli, M. Rupp, M. Scheffler, O. A. von Lilienfeld, A. Tkatchenko, and K.-R. Müller, *J. Chem. Theory Comput.* **9**, 3404 (2013).
- [33] B. Efron, *J. Am. Stat. Assoc.* **82**, 171 (1987).
- [34] C. Cortes and V. Vapnik, *Mach. Learn.* **20**, 273 (1995).
- [35] C. Schaffer, *Mach. Learn.* **13**, 135 (1993).
- [36] C. Deans, L. Marmugi, S. Hussain, and F. Renzoni, *Proc. SPIE Int. Soc. Opt. Eng.* **9900**, 99000F (2016).