nature photonics

Article

https://doi.org/10.1038/s41566-025-01657-6

Model-free estimation of the Cramér–Rao bound for deep learning microscopy in complex media

Received: 31 May 2024

Accepted: 11 March 2025

Published online: 28 May 2025

Check for updates

Ilya Starshynov[1,4], Maximilian Weimar[2,4], Lukas M. Rachbauer[2], Günther Hackl[2], Daniele Faccio[1], Stefan Rotter[2] & Dorian Bouchet[3]

Artificial neural networks have become important tools to harness the complexity of disordered or random photonic systems. Recent applications include the recovery of information from light that has been scrambled during propagation through a complex scattering medium, especially in the challenging case in which the deterministic input–output transmission matrix cannot be measured. This naturally raises the question of what the limit is that information theory imposes on this recovery process, and whether neural networks can actually reach this limit. To answer these questions, we introduce a model-free approach to calculate the Cramér–Rao bound, which sets the ultimate precision limit at which artificial neural networks can operate. As an example, we apply this approach in a proof-of-principle experiment using laser light propagating through a disordered medium, evidencing that a convolutional network approaches the ultimate precision limit in the challenging task of localizing a reflective target hidden behind a dynamically fluctuating scattering medium. The model-free method introduced here is generally applicable to benchmark the performance of any deep learning microscope, to drive algorithmic developments and to push the precision of metrology and imaging techniques to their ultimate limit.

Complexity and related chaotic processes form the foundation of many physical phenomena. At the same time, complexity has hindered the capability to predict the evolution of physical systems in various research fields ranging from biophotonics[1,2] to quantum optics[3]. However, over the past few years, many of these fields have seen remarkable progress as a result of data-driven models and machine learning approaches, which are capable of harnessing the complexity of physical dynamics with surprising effectiveness[4]. In essence, these data-driven models process physical information from the training data and build a representative statistical model from these data. The question of how exactly this information is maintained and distilled by an artificial neural network (ANN) then naturally arises, especially for applications in photonics that require the reconstruction of images[5] or the precise estimation of physical observables[6].

Light propagation through a complex medium is a typical example of a research area that has seen major advances owing to data-driven models[7]. Although the propagation of light is ruled by simple laws in homogeneous media, retrieving information through complex scattering media is a critical challenge, as light typically undergoes a number of unknown scattering and absorption events during propagation[8]. Several imaging techniques that were originally developed to address the challenges associated with complex light scattering are based on

[1]School of Physics and Astronomy, University of Glasgow, Glasgow, UK. [2]Institute for Theoretical Physics, Vienna University of Technology (TU Wien), Vienna, Austria. [3]Université Grenoble Alpes, CNRS, LIPhy, Grenoble, France. [4]These authors contributed equally: Ilya Starshynov, Maximilian Weimar. e-mail: dorian.bouchet@univ-grenoble-alpes.fr

https://doi.org/10.1038/s41566-025-01657-6

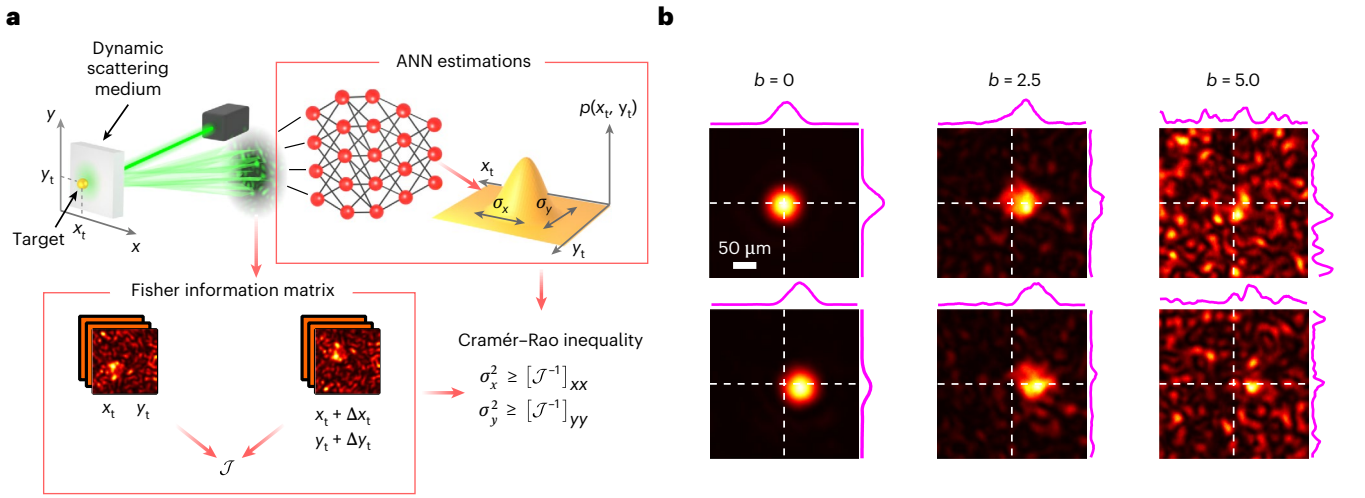

Fig. 1 | Principle of the experiment. a, A reflective target is placed behind a dynamic scattering sample. Our goal is to infer its position from the reflected coherent light using different ANNs, and compare the precision of the ANNs with the ultimate limit calculated using Fisher information theory. **b**, Examples of pairs of measured images with two adjacent target positions (see the top and bottom rows) at different scattering strengths (see the three columns labelled by the optical thickness $b = L/\ell$). In the case of a strongly scattering sample ($b = 5$), the precise target position cannot be easily estimated from such images due to random light scattering occurring within the dynamic medium.

the insight that the measured output random patterns (referred to as speckle patterns) are the deterministic results of millions of scattering events[9]. In particular, a popular approach is to experimentally measure the deterministic relation between object and image planes, which is conveniently done in a scattering matrix formalism[10–12]; the knowledge of this matrix can then be used to effectively transform the scattering medium into a simple optical element[13,14]. Nevertheless, in many cases of interest, including those involving dynamical scattering media that change in time, this deterministic relation cannot be measured without the use of a guidestar[15], which limits the practical applicability of the approach.

To reconstruct images through such media considering the occurrence of random scattering events, a number of techniques have been developed such as multiphoton microscopy[16,17], optical coherence tomography[18,19] and correlation imaging[20,21], but these are limited to specific application scenarios. In recent years, the emergence of ANNs has been a game changer in this field; indeed, deep neural networks were shown to be able to learn to image through a known or fixed scattering medium, or even reconstruct images hidden by unseen scattering media, with remarkable fidelity[22–38].

To benchmark the performance of these techniques, an insightful strategy consists of studying the fundamental limits imposed by physical laws on the quality of the reconstructed images. One such limit is the Abbe limit, which describes the ultimate resolution achievable with an imaging system. However, although this limit is well defined in homogeneous media, it cannot be used when imaging through scattering media. Indeed, multiple scattering effects alter light fields in a complex way, not only affecting the resolution but also impacting its contrast, possibly with the apparition of artefacts. Moreover, resolution is not a relevant metric to describe the performance of most computational imaging techniques, including those based on ANNs, since they often break the Abbe limit by including prior information about the object in the reconstruction procedure.

Instead of image resolution, we propose here to use a criterion that overcomes these shortcomings and sets a quantitative benchmark to assess the ability of ANNs to extract information from physical measurements. This criterion, which is commonly used in optical metrology[39,40] and which is applicable to any physical system, is implemented and demonstrated here for the specific case of light propagation in complex scattering media. The approach relies on the assumption that the required information is not a global image in itself, but specific features of interest that could potentially be extracted from this image, such as the size or the position of an object[41–44]. In this framework, the relevant physical limit is the Cramér–Rao bound, which sets the ultimate precision with which parameters of interest can be estimated[45]. However, calculating this bound is a challenging task in complex systems. Indeed, the expression of the Cramér–Rao bound is based on the probability density function (PDF) describing the data statistics, the analytical expression of which is typically unknown.

Here we overcome this hurdle by introducing an approach to calculate the Cramér–Rao bound solely from experimental data, even when a physical model describing the data is not available. We then use this benchmark to assess the performance of different types of ANNs that we have trained to estimate the position of an object through a dynamic scattering sample. For this task, our analysis demonstrates that a convolutional neural network (CNN) approaches the limit set by the Cramér–Rao bound. In our experiments, we consider the canonical case of a target object hidden behind a random scattering medium (Fig. 1a). We illuminate the medium with a laser, which gets transmitted to the hidden object as a random speckle pattern of light. The retro-reflected light passes back through the random medium before being collected by a camera (Methods provides a detailed description of the optical setup). Our aim is to precisely estimate the position $\theta = (x_t, y_t)$ of this object from a single-intensity frame collected by the camera. In practice, as an object, we use a reflecting target displayed by a digital micromirror device (DMD). This target is hidden behind a dynamic random scattering sample composed of a suspension of TiO_2 particles in glycerol, which is being pumped through a flow cuvette. The decorrelation time is such that two successive frames are uncorrelated (Methods). In this configuration, we can vary the concentration of TiO_2 particles to tune the scattering mean free path ℓ and control the optical thickness $b = L/\ell$ of the medium, where L is the optical path in the cuvette. To study different scattering regimes, we vary the optical thickness b from 1.7 to 5, which we also compared with the case in which light propagates in free space ($b = 0$). Although the target can be easily localized in free space, the occurrence of random scattering events results in the generation of speckle patterns that completely conceal the target in the case of large optical thicknesses (Fig. 1b and Supplementary Information, section 2).

https://doi.org/10.1038/s41566-025-01657-6

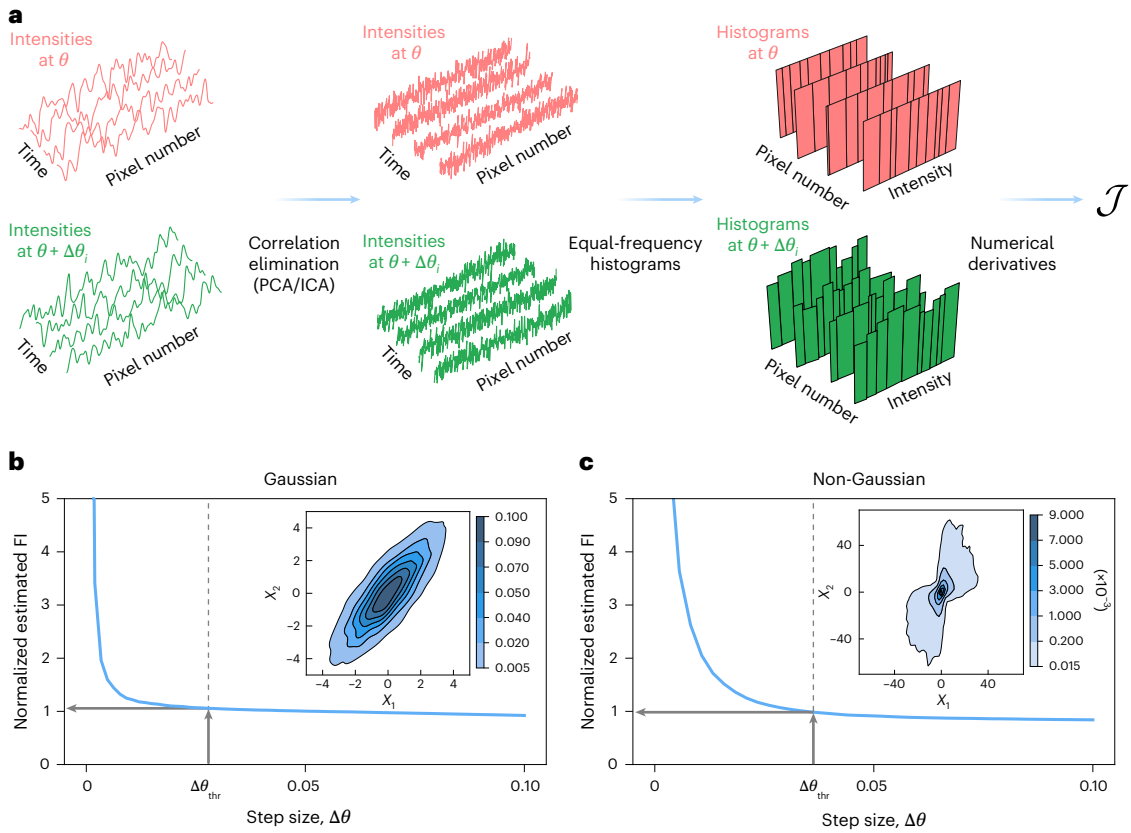

Fig. 2 | Demonstration of the Fisher information estimation procedure using numerically generated data. a, Representation of the procedure used to evaluate the Fisher information. Images (raw data) are measured as a function of time, that is, for different noise realizations (left). Statistical independence between different image pixels is restored using either PCA for Gaussian statistics or ICA for non-Gaussian statistics (middle). The underlying probability distributions are then estimated using equal-frequency binning (right). Replicating this analysis for different values of θ (represented in red and green) allows us to estimate the Fisher information with a finite-difference scheme. **b**, Numerical demonstration of the procedure for data with correlated Gaussian statistics (following the distribution shown in the inset). The blue curve depicts the estimated Fisher information (FI) divided by the true Fisher information, as a function of the step size of the finite-difference scheme. A clear plateau is observed when the estimated Fisher information reaches the true Fisher information, the optimal value of the step size $\Delta\theta_{thr}$ being identified using a stability criterion based on the second derivative of this function. **c**, Analogous to **b** but for data with correlated non-Gaussian statistics, demonstrating the efficiency of the method even for such complex statistics.

Results

Estimation of the Cramér–Rao bound

Our first goal is to assess the ultimate localization precision achievable through dynamic scattering media. For this purpose, we use a general framework based on statistical estimation theory[45]. Owing to statistical fluctuations arising from random interactions between the complex medium and the probe field, the connection between measured data X and parameters of interest θ is intrinsically probabilistic. We can, thus, describe it using the PDF $p(X;\theta)$, which is parameterized by the vector θ containing all parameters of interest, such as the spatial coordinates x_t and y_t of the target. Note that this PDF is also known as the likelihood function. In this picture, the measured data are represented by the random vector X, composed of N random variables X_k representing the intensity values measured by each pixel of the camera. Then, the Cramér–Rao inequality sets an ultimate limit on the precision of the estimated values of θ. More precisely, the standard deviation $\sigma_i(\theta)$ of any unbiased estimator of the ith component of θ satisfies[45]

$$\sigma_i(\theta) \geq C_i(\theta) = \sqrt{\left[\mathcal{J}^{-1}(\theta)\right]_{ii}}, \tag{1}$$

where $\mathcal{J}(\theta)$ is the Fisher information matrix and $C_i(\theta)$ is the Cramér–Rao bound on the ith component of θ. The Fisher information matrix is defined by

$$\left[\mathcal{J}(\theta)\right]_{ij} = \left\langle \left(\frac{\partial \ln[p(X;\theta)]}{\partial \theta_i}\right)\left(\frac{\partial \ln[p(X;\theta)]}{\partial \theta_j}\right)\right\rangle, \tag{2}$$

where $\langle \ldots \rangle$ denotes the average over statistical fluctuations.

This formalism is very general, as it sets an ultimate limit on the precision of the estimated values of θ regardless of the physical origin of statistical fluctuations. Typically, the Cramér–Rao bound is calculated for shot-noise-limited measurements or for measurements corrupted by an additive Gaussian noise. Examples include single-molecule localization microscopy[46,47], non-line-of-sight configurations[48] and static scattering media described by known scattering matrices[49–51]. In such cases, measurements follow either Gaussian or Poissonian statistics, and calculating the Cramér–Rao bound is relatively straightforward as equation (2) then takes a simple analytical form. By contrast, for complex scattering media that change in time, statistical fluctuations are dominated by random scattering events and no simple analytical solutions are available for equation (2). Indeed, the noise statistics is unknown, and the random variables X_k (which represent the measured pixel values) are correlated with each other and do not follow a simple parametric model. We, thus, need to rely on non-parametric estimations of the distribution, which is challenging due to the high dimension of the random variable.

In this work, we show how to overcome this difficulty by evaluating the Cramér–Rao bound solely from experimental measurements,

https://doi.org/10.1038/s41566-025-01657-6

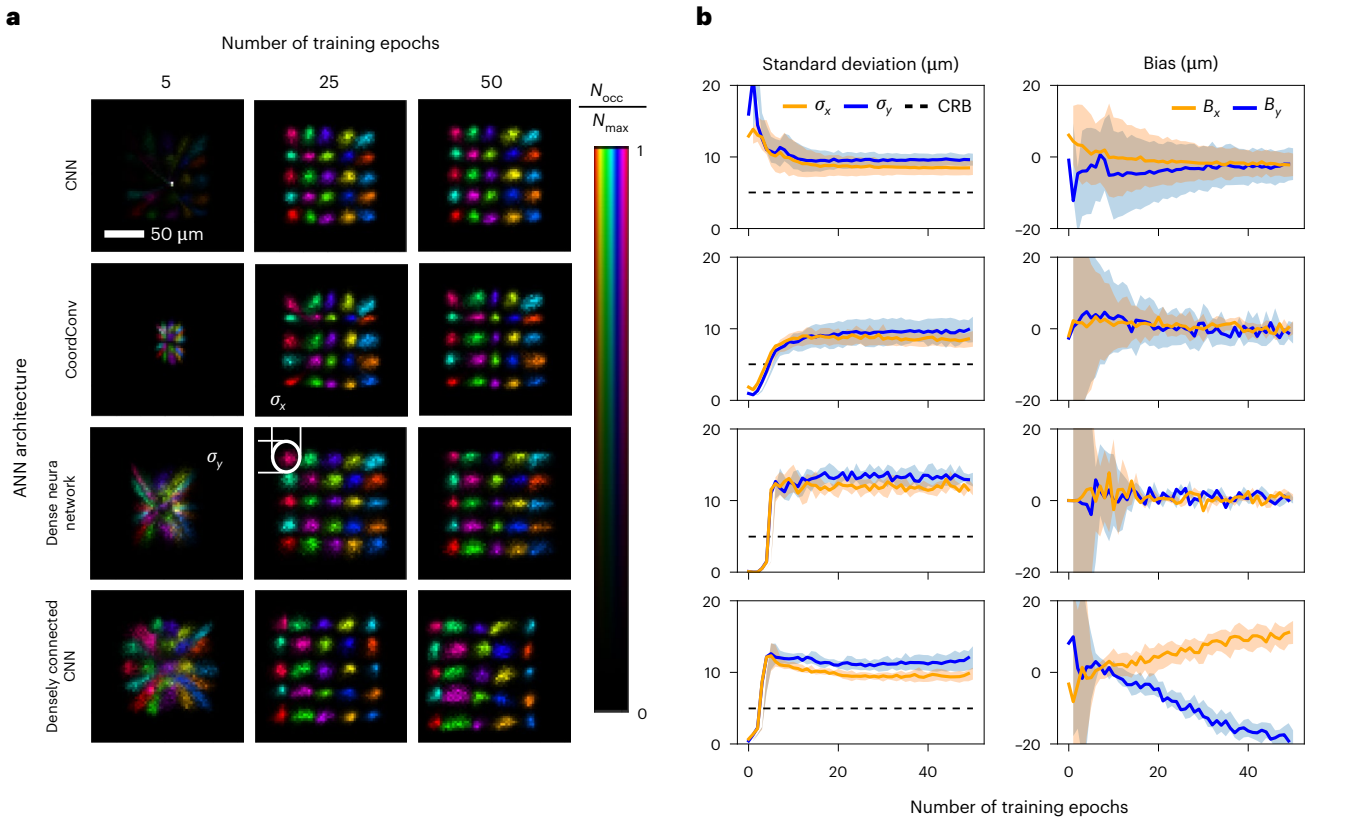

Fig. 3 | Comparison of different ANN architectures for the sample with optical thickness $b = 4.2$. The rows in **a** and **b** correspond to (from top to bottom) CNN, CoordConv, dense and densely connected CNN architectures. **a**, Two-dimensional histograms of the estimated target position. Different colours encode different ground-truth positions, and brightness encodes the probability. $N_{max} = 5{,}000$ testing examples are used for each target position, and N_{occ} denotes the number of occurence for each estimated target position. The spread of the histograms characterizes the ANN precision. The columns show the progression of the prediction performance throughout the training process (after 5, 25 and 50 epochs, respectively). **b**, Evolution of the standard deviations (left column) and biases (right column) during training. The thick lines show the values averaged over all target the positions, and the shaded areas represent the interquartile range (third minus first quartile) calculated over 25 positions. The black dashed lines in the left column represent the Cramér–Rao bound (CRB), calculated assuming that the estimators are unbiased.

approximating the underlying PDF from a finite number of samples (Fig. 2a shows a graphical illustration of this approach). First, we address the problem of correlations by applying a transformation to the data that preserves the information content but removes the dependencies between the components of X. A common approach, based on a principal component analysis (PCA)[52], is to estimate the covariance matrix of the data and choose a transformation that diagonalizes this matrix. This approach, however, does not guarantee statistical independence between the components. To improve on this method, we use a so-called independent component analysis (ICA)[53], which allows to construct a linear transformation $Y = AX$ that minimizes the dependencies between components of the random variable, harnessing more degrees of freedom for choosing the transformation. From the transformed data, we then estimate the underlying density function $p(Y; \theta)$ using equal-frequency histograms, which allows for more stable derivative estimates compared with histograms based on uniform bin widths. This enables us to construct an estimator of the Fisher information by approximating the derivatives of $p(Y; \theta)$ with a finite-difference scheme. The value of the step size $\Delta\theta$ is chosen by evaluating our estimator as a function of $\Delta\theta$ and using a stability criterion based on the second derivative of this function (Methods and Supplementary Information, section 3.4).

To demonstrate the performance of our approach, we test it on numerically generated data with known Fisher information. We first generate 50-dimensional multivariate Gaussian data that we decorrelate using PCA (non-Gaussianity of the data is a requirement for the ICA algorithm to converge, but PCA does guarantee independence for Gaussian statistics). Figure 2b shows the estimated normalized Fisher information as a function of the step size of the finite-difference scheme; a clear plateau can be identified when the normalized Fisher information reaches unity, which means that the Fisher information is correctly estimated. We then generate correlated non-Gaussian data, which we process using ICA. Even for these complex statistics, we again observe a clear plateau when the normalized Fisher information reaches unity, illustrating the broad applicability of the method.

Comparison of different ANN architectures

We now want to assess whether the precision of ANNs can approach the Cramér–Rao bound. For this purpose, we investigate various ANN architectures to evaluate their performance in estimating the target position from the measured images. We first collect a number of speckle patterns with the target located in the centre of the field of view, that is, $\theta = (0, 0)$. We then randomly translate these patterns in the transverse plane to generate an augmented dataset (Methods), with which we train different ANNs. Finally, we test our networks on unseen data, for which the target displayed by the DMD has been physically translated. In general, networks can retrieve the position θ of the target with precision that depends on many factors, including the optical thickness of the scattering media, the network architecture and the number of training epochs. To quantitatively evaluate the precision of the ANN predictions, we compute the expected value of the target coordinates based on the distribution provided by the ANN (this is achieved using

https://doi.org/10.1038/s41566-025-01657-6

a softmax activation function, as described in the Methods). Subsequently, we construct individual histograms for each target position and use them to calculate the corresponding standard deviations σ_x and σ_y, which we use as a figure of merit for the ANN precision. In addition, we calculate the average deviations of the predicted coordinates from the true positions to characterize the ANN biases B_x and B_y.

The most straightforward ANN choice for image processing tasks is a CNN[54]. It has been shown, however, that ANNs with purely convolutional layers typically fail in the task of accurately tracking the coordinates of an object in an image[55]. This is due to the intrinsic translational invariance of the convolution operation, which leads to the loss of information about the feature position. To break the spatial invariance of the convolution operation, one possibility is to explicitly add coordinate meshes to the layers (CoordConv layers), which leads to faster convergence and lower bias compared with usual CNNs[55]. Although fully connected layers (dense neural networks) are usually harder to train, they can also be used in our case since the target of interest has a simple spatial structure. Finally, a CNN modification featuring structured skip connections was recently introduced[56]. Such connections form a structure akin to a dense layer, leading to the name densely connected CNNs, also known as 'DenseNets'.

We compared the performance of these four classes of ANN. Before training, we performed hyperparameter tuning for each architecture (Supplementary Information, section 4.2). The best-performing ANN of each class was then trained and tested to locate the hidden target. The comparison of their performance is shown in Fig. 3. In Fig. 3a, we present the estimated target positions for the four different architectures along the training process, for the scattering sample with optical thickness $b = 4.2$. On these two-dimensional histograms, distinct ground-truth target positions are represented by different colours, and the brightness of each colour indicates the corresponding probability predicted by the ANN for that particular position. We observe that the choice of the ANN architecture does not seem to have a large influence over the estimated target position. To quantitatively analyse the performance of the different ANN architectures, Fig. 3b shows the evolution of the standard deviations σ_x and σ_y (left column) and biases B_x and B_y (right column), calculated from the histograms in Fig. 3a. These values were obtained by averaging over statistical fluctuations of the ANN predictions.

As it turns out, the CNN approaches the (unbiased) Cramér–Rao bound, which here is equal to 5.0 µm for the scattering sample under consideration ($b = 4.2$). However, as evident from the large variability in the associated bias plot (Fig. 3b, first row, right column), the CNN develops a substantial bias. This issue can be resolved by including coordinate layers, which strongly mitigates the bias and leads to the same standard deviation as the usual CNN. In the case of the dense neural network, the observed bias is smaller, but the standard deviation is larger than that of CoordConv. The densely connected CNN architecture exhibits a very different behaviour: after only a few epochs, it gives low bias and approaches the Cramér–Rao bound. Nevertheless, this advantage comes with a trade-off: a large bias is observed later in the training process, requiring careful consideration to stop the training at an appropriate point. Note that in the case of the CoordConv, dense neural network and densely connected CNN architectures, the Cramér–Rao bound is overpassed in the first few epochs; this occurs because the initial guess of these architectures is zero for each coordinate, resulting in standard deviations equal to zero. Estimations are then highly biased in the early stage of the training, which makes the unbiased Cramér–Rao bound calculated using equation (1) not relevant to the estimations performed during these first epochs.

Achievable precision for different scattering strengths

On the basis of the combination of minimal uncertainty and small bias (using the lowest mean squared error as a criterion, as described in Supplementary Information, section 4.3), we select CoordConv as our best architecture and study how it performs at different scattering strengths. As an illustration, Fig. 4a shows the ANN predictions in the absence of a scattering sample as well as for two other optical thicknesses ($b = 2.5$ and $b = 5$). Although the bias seems unaffected by the presence of the scattering sample, the variance of the estimated positions greatly increases for thick scattering samples. To quantitatively describe this behaviour, Fig. 4b shows the standard deviations achieved with the CoordConv architecture as a function of the optical thickness, obtained by training and testing this architecture 25 times with different random initial weights. These results also include a bias correction step, which is necessary to compare the ANN predictions with the unbiased Cramér–Rao bound (Supplementary Information, section 4.4).

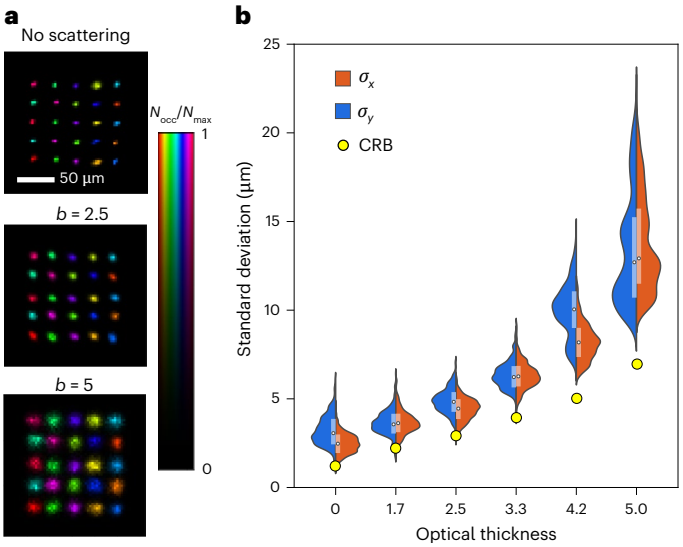

Fig. 4 | Dependence of ANN precision on the optical thickness of the scattering sample. a, Examples of reconstructed histograms of the target positions using CoordConv architecture in the absence of a scattering sample ($b = 0$), as well as for $b = 2.5$ and $b = 5$. **b**, Violin plots showing the CoordConv precision achieved for 25 random initializations of the network. The white dots show the median values of σ_x and σ_y, vertical bars indicate the corresponding first and third quartiles, and the coloured areas show the associated histograms. The yellow circles represent the Cramér–Rao bound (CRB).

As expected, the average standard deviation increases with the optical thickness, ranging from approximately 2.5 µm in the absence of a scattering sample up to 12.8 µm for a strongly scattering sample ($b = 5$). Remarkably, for small optical thicknesses, the standard deviation of the ANN comes very close to the Cramér–Rao bound (Fig. 4b, yellow circles), demonstrating that the CoordConv architecture can reach the ultimate precision limit. In some cases, it even seems that the Cramér–Rao bound can even be overpassed; this artefact can be explained by the approximations made during the calculation of Fisher information (Supplementary Information, section 3.5) and by the finite size of the test dataset that was used to calculate the standard deviation of the ANN (Supplementary Information, section 4.5). For larger optical thicknesses, we observe a large variability in the ANN precision, as captured by the broader histograms shown in Fig. 4b. This variability comes essentially from the dependence of ANN precision on the target position. We indeed observe that off-central positions, which are less connected to the output of the network, are harder to predict than the central ones (Supplementary Information, section 4.6). Moreover, it is also plausible that with a larger number of trainable weights, the ANN would perform better at high optical thicknesses (Supplementary Information, section 4.2).

In Fig. 4b, we can also observe that the dependence of the Cramér–Rao bound on the optical thickness is approximately linear

between $b = 0$ and $b = 5$. Although no analytical expression is currently available to describe this behaviour, we suppose that it arises here as the result of a competition between shot noise and mechanical vibrations of the sample (mostly relevant for small optical thicknesses), as well as exponential attenuation of ballistic photons (mostly relevant for large optical thicknesses).

Finally, we also performed cross-tests of the models, which we trained and tested using datasets associated with different optical thicknesses and different object sizes. We observe that models trained on the strongest scattering dataset retain their ability to predict the target's position in weakly scattering conditions, whereas the models trained on a weak scattering dataset become imprecise to estimate the target position in strongly scattering conditions (Supplementary Information, section 5.1). Moreover, the models seem to generalize relatively well for small variations in the object size and shape (Supplementary Information, section 5.2).

Discussion

We have shown that the Cramér–Rao inequality, which is the fundamental inequality that limits the precision of any estimator, also provides the ultimate benchmark for different ANN architectures trained to provide estimates on complex photonic systems. For the problem of estimating the position of a target behind a dynamic scattering sample, the performance of all architectures is similar, with a slight advantage for the CNN with CoordConv layers. We have then demonstrated that for different optical thicknesses, this architecture approaches the Cramér–Rao bound.

Calculating the Cramér–Rao bound entails accessing the Fisher information matrix describing the measured data. Although this matrix is typically derived from analytical models, here we presented a general model-free approach to approximate the Fisher information matrix of an unknown statistical distribution based on experimental data. Note that, in general, the Fisher information matrix depends on the true value of parameter θ, and therefore, the Cramér–Rao bound allows one to assess the local performance of ANNs around a given value in parameter space. Although here we studied a system with translational invariance for which the Fisher information does not depend on θ, our work could also serve as a building block to study more involved situations requiring the Fisher information to be calculated over the whole parameter space. Another perspective of our work is to analyse the achievable precision when large images need to be reconstructed. Although a pixel-based imaging strategy seems impractical here due to the large number of parameters typically involved[57,58], finding a relevant sparse representation of these images appears to be a promising strategy[59,60]. Our analysis performed with two parameters (the transverse coordinates of the target), thus, paves the way towards the analysis of more complex imaging scenarios in which many parameters need to be estimated.

The formalism inherently includes all the physical effects that affect the precision with which parameter values can be estimated. In our experiments, the statistics of noise was of fundamentally different origin depending on the scattering strength. Indeed, in free space, the precision was essentially limited by shot noise as well as by mechanical vibrations of the experimental setup. By contrast, in the presence of complex scattering media, the precision was limited by the occurrence of random scattering events in the media. Yet, we have shown that the Cramér–Rao bound can be successfully calculated regardless of the physical origin of the noise, and without any analysis of the speckle correlations characterizing our scattering systems[61–63]. In fact, the key quantity is rather the likelihood function, which is assessed directly from the measured experimental data and without any assumption on the data statistics. Moreover, all prior information is naturally included in the parameterization of the system. As such, even far-field super-resolution effects are inherently described by the formalism.

The method itself is very general and can be used to assess whether any ANN is optimal with respect to the task it was designed for. We expect our approach to be especially useful to drive algorithmic developments in the field of computational imaging[5]. Major potential applications include imaging through multimode fibres[64] and through complex scattering tissues[8], for instance, for in vivo neuronal imaging[65,66].

Online content

Any methods, additional references, Nature Portfolio reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41566-025-01657-6.

References

1. Moen, E. et al. Deep learning for cellular image analysis. *Nat. Methods* **16**, 1233–1246 (2019).
2. Tian, L. et al. Deep learning in biomedical optics. *Lasers Surg. Med.* **53**, 748–775 (2021).
3. Gebhart, V. et al. Learning quantum systems. *Nat. Rev. Phys.* **5**, 141–156 (2023).
4. Tang, Y., Kurths, J., Lin, W., Ott, E. & Kocarev, L. Introduction to focus issue: when machine learning meets complex systems: networks, chaos, and nonlinear dynamics. *Chaos* **30**, 063151 (2020).
5. Barbastathis, G., Ozcan, A. & Situ, G. On the use of deep learning for computational imaging. *Optica* **6**, 921–943 (2019).
6. Zuo, C. et al. Deep learning in optical metrology: a review. *Light Sci. Appl.* **11**, 39 (2022).
7. Gigan, S. et al. Roadmap on wavefront shaping and deep imaging in complex media. *J. Phys. Photon.* **4**, 042501 (2022).
8. Yoon, S. et al. Deep optical imaging within complex scattering media. *Nat. Rev. Phys.* **2**, 141–158 (2020).
9. Mosk, A. P., Lagendijk, A., Lerosey, G. & Fink, M. Controlling waves in space and time for imaging and focusing in complex media. *Nat. Photon.* **6**, 283–292 (2012).
10. Popoff, S. M. et al. Measuring the transmission matrix in optics: an approach to the study and control of light propagation in disordered media. *Phys. Rev. Lett.* **104**, 100601 (2010).
11. Conkey, D. B., Caravaca-Aguirre, A. M. & Piestun, R. High-speed scattering medium characterization with application to focusing light through turbid media. *Opt. Express* **20**, 1733–1740 (2012).
12. Yu, H. et al. Measuring large optical transmission matrices of disordered media. *Phys. Rev. Lett.* **111**, 153902 (2013).
13. Popoff, S. M., Lerosey, G., Fink, M., Boccara, A. C. & Gigan, S. Image transmission through an opaque material. *Nat. Commun.* **1**, 81 (2010).
14. Choi, Y. et al. Overcoming the diffraction limit using multiple light scattering in a highly disordered medium. *Phys. Rev. Lett.* **107**, 023902 (2011).
15. Horstmeyer, R., Ruan, H. & Yang, C. Guidestar-assisted wavefront-shaping methods for focusing light into biological tissue. *Nat. Photon.* **9**, 563–571 (2015).
16. Denk, W., Strickler, J. H. & Webb, W. W. Two-photon laser scanning fluorescence microscopy. *Science* **248**, 73–76 (1990).
17. Horton, N. G. et al. In vivo three-photon microscopy of subcortical structures within an intact mouse brain. *Nat. Photon.* **7**, 205–209 (2013).
18. Huang, D. et al. Optical coherence tomography. *Science* **254**, 1178–1181 (1991).
19. Kashani, A. H. et al. Optical coherence tomography angiography: a comprehensive review of current methods and clinical applications. *Progr. Retin. Eye Res.* **60**, 66–100 (2017).

20. Bertolotti, J. et al. Non-invasive imaging through opaque scattering layers. *Nature* **491**, 232–234 (2012).
21. Katz, O., Heidmann, P., Fink, M. & Gigan, S. Noninvasive single-shot imaging through scattering layers and around corners via speckle correlations. *Nat. Photon.* **8**, 784–790 (2014).
22. Ando, T., Horisaki, R. & Tanida, J. Speckle-learning based object recognition through scattering media. *Opt. Express* **23**, 33902–33910 (2015).
23. Horisaki, R., Takagi, R. & Tanida, J. Learning-based imaging through scattering media. *Opt. Express* **24**, 13738–13743 (2016).
24. Li, S., Deng, M., Lee, J., Sinha, A. & Barbastathis, G. Imaging through glass diffusers using densely connected convolutional networks. *Optica* **5**, 803–813 (2018).
25. Li, Y., Xue, Y. & Tian, L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media. *Optica* **5**, 1181–1190 (2018).
26. Sun, Y., Xia, Z. & Kamilov, U. S. Efficient and accurate inversion of multiple scattering with deep learning. *Opt. Express* **26**, 14678–14688 (2018).
27. Turpin, A., Vishniakou, I. & Seelig, J. D. Light scattering control in transmission and reflection with neural networks. *Opt. Express* **26**, 30911–30929 (2018).
28. Rahmani, B., Loterie, D., Konstantinou, G., Psaltis, D. & Moser, C. Multimode optical fiber transmission with a deep learning network. *Light Sci. Appl.* **7**, 69 (2018).
29. Lyu, M., Wang, H., Li, G., Zheng, S. & Situ, G. Learning-based lensless imaging through optically thick scattering media. *Adv. Photon.* **1**, 036002 (2019).
30. Caramazza, P., Moran, O., Murray-Smith, R. & Faccio, D. Transmission of natural scene images through a multimode fibre. *Nat. Commun.* **10**, 2029 (2019).
31. Zhu, C. et al. Image reconstruction through a multimode fiber with a simple neural network architecture. *Sci. Rep.* **11**, 896 (2021).
32. Resisi, S., Popoff, S. M. & Bromberg, Y. Image transmission through a dynamically perturbed multimode fiber by deep learning. *Laser Photon. Rev.* **15**, 2000553 (2021).
33. Li, Y., Cheng, S., Xue, Y. & Tian, L. Displacement-agnostic coherent imaging through scatter with an interpretable deep neural network. *Opt. Express* **29**, 2244–2257 (2021).
34. Starshynov, I., Turpin, A., Binner, P. & Faccio, D. Statistical dependencies beyond linear correlations in light scattered by disordered media. *Phys. Rev. Res.* **4**, L022033 (2022).
35. Rahmani, B. et al. Learning to image and compute with multimode optical fibers. *Nanophotonics* **11**, 1071–1082 (2022).
36. Bai, B. et al. All-optical image classification through unknown random diffusers using a single-pixel diffractive network. *Light Sci. Appl.* **12**, 69 (2023).
37. Abdulaziz, A., Mekhail, S. P., Altmann, Y., Padgett, M. J. & McLaughlin, S. Robust real-time imaging through flexible multimode fibers. *Sci. Rep.* **13**, 11371 (2023).
38. Huang, Z., Gu, Z., Shi, M., Gao, Y. & Liu, X. OP-FCNN: an optronic fully convolutional neural network for imaging through scattering media. *Opt. Express* **32**, 444–456 (2024).
39. Giovannetti, V., Lloyd, S. & Maccone, L. Advances in quantum metrology. *Nat. Photon.* **5**, 222 (2011).
40. Polino, E., Valeri, M., Spagnolo, N. & Sciarrino, F. Photonic quantum metrology. *AVS Quantum Sci.* **2**, 024703 (2020).
41. Cohen, S. D., Cavalcante, H. L., DdeS. & Gauthier, D. J. Subwavelength position sensing using nonlinear feedback and wave chaos. *Phys. Rev. Lett.* **107**, 254103 (2011).
42. del Hougne, P., Imani, M. F., Fink, M., Smith, D. R. & Lerosey, G. Precise localization of multiple noncooperative objects in a disordered cavity by wave front shaping. *Phys. Rev. Lett.* **121**, 063901 (2018).
43. del Hougne, P. Robust position sensing with wave fingerprints in dynamic complex propagation environments. *Phys. Rev. Res.* **2**, 043224 (2020).
44. Jauregui-Sánchez, Y., Penketh, H. & Bertolotti, J. Tracking moving objects through scattering media via speckle correlations. *Nat. Commun.* **13**, 5779 (2022).
45. Van Trees, H. L., Bell, K. L. & Tian, Z. *Detection Estimation and Modulation Theory, Part I* (John Wiley & Sons, 2013).
46. Ober, R. J., Ram, S. & Ward, E. S. Localization accuracy in single-molecule microscopy. *Biophys. J.* **86**, 1185 (2004).
47. Deschout, H. et al. Precisely and accurately localizing single emitters in fluorescence microscopy. *Nat. Methods* **11**, 253–266 (2014).
48. Song, X., Xu, J., Liu, F., Han, T. X. & Eldar, Y. C. Intelligent reflecting surface enabled sensing: Cramér-Rao lower bound optimization. In 2022 IEEE Globecom Workshops *(GC Wkshps)* 413–418 (IEEE, 2022).
49. Bouchet, D., Carminati, R. & Mosk, A. P. Influence of the local scattering environment on the localization precision of single particles. *Phys. Rev. Lett.* **124**, 133903 (2020).
50. Bouchet, D., Rotter, S. & Mosk, A. P. Maximum information states for coherent scattering measurements. *Nat. Phys.* **17**, 564–568 (2021).
51. Horodynski, M., Bouchet, D., Kühmayer, M. & Rotter, S. Invariance property of the Fisher information in scattering media. *Phys. Rev. Lett.* **127**, 233201 (2021).
52. Jolliffe, I. T. & Cadima, J. Principal component analysis: a review and recent developments. *Phil. Trans. R. Soc. A* **374**, 20150202 (2016).
53. Hyvärinen, A. & Oja, E. Independent component analysis: algorithms and applications. *Neural Netw.* **13**, 411–430 (2000).
54. Goodfellow, I., Bengio, Y. & Courville, A. *Deep Learning* (MIT Press, 2016).
55. Liu, R. et al. An intriguing failing of convolutional neural networks and the CoordConv solution. In *Proc. 32nd International Conference on Neural Information Processing Systems NIPS'18* 9628–9639 (Curran Associates, 2018).
56. Huang, G., Liu, Z., Maaten, L. V. D. & Weinberger, K. Q. Densely connected convolutional networks. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* 2261–2269 (IEEE Computer Society, 2017).
57. Barrett, H. H., Denny, J. L., Wagner, R. F. & Myers, K. J. Objective assessment of image quality. II. Fisher information, Fourier crosstalk, and figures of merit for task performance. *J. Opt. Soc. Am. A* **12**, 834–852 (1995).
58. Bouchet, D., Dong, J., Maestre, D. & Juffmann, T. Fundamental bounds on the precision of classical phase microscopes. *Phys. Rev. Appl.* **15**, 024047 (2021).
59. Szameit, A. et al. Sparsity-based single-shot subwavelength coherent diffractive imaging. *Nat. Mater.* **11**, 455–459 (2012).
60. Bouchet, D., Seifert, J. & Mosk, A. P. Optimizing illumination for precise multi-parameter estimations in coherent diffractive imaging. *Opt. Lett.* **46**, 254–257 (2021).
61. Feng, S., Kane, C., Lee, P. A. & Stone, A. D. Correlations and fluctuations of coherent wave transmission through disordered media. *Phys. Rev. Lett.* **61**, 834 (1988).
62. Akkermans, E. & Montambaux, G. *Mesoscopic Physics of Electrons and Photons* (Cambridge Univ. Press, 2007).
63. Starshynov, I. et al. Non-Gaussian correlations between reflected and transmitted intensity patterns emerging from opaque disordered media. *Phys. Rev.* **8**, 021041 (2018).
64. Cao, H., Čižmár, T., Turtaev, S., Tyc, T. & Rotter, S. Controlling light propagation in multimode fibers for imaging, spectroscopy, and beyond. *Adv. Opt. Photon.* **15**, 524–612 (2023).

65. Stibůrek, M. et al. 110 µm thin endo-microscope for deep-brain in vivo observations of neuronal connectivity, activity and blood flow dynamics. *Nat. Commun.* **14**, 1897 (2023).
66. Sarafraz, H. et al. Speckle-enabled in vivo demixing of neural activity in the mouse brain. *Biomed. Opt. Express* **15**, 3586–3608 (2024).

Methods

Experimental setup

The setup used in our experiments is shown in Supplementary Information, section 1. A laser beam from a diode (Thorlabs DJ532-40) is expanded to a diameter of 2 mm before entering the optical system. Lens L2 forms the image of the front surface of the scattering sample on a scientific complementary metal–oxide–semiconductor camera (Andor Zyla 5.5). Lens L1 is placed in such a way that the input beam after L2 is collimated. Using the lens L3 and the microscope objective MO (Olympus Plan N 10X), we image the DMD (Vialux V-7001; pixel size, 13.7 µm) onto the back surface of the scattering sample, as well as forming the illumination area of 8-mm diameter on the DMD. The resulting magnification of the optical system formed by L2, MO and L3 is 2.9. The DMD is oriented such that the surface of the 'on' pixels is perpendicular to the optical axis of the imaging system, that is, these pixels back-reflect the incident light towards the sample. By contrast, the surface of the 'off' pixels is tilted such that light reflected by these pixels escapes the optical system, because it goes outside the aperture of lens L3. Note that in this configuration, the DMD is tilted by a few degrees in the y direction, which leads to a slight defocus for different y positions.

As a scattering sample, we use a suspension of TiO_2 nanoparticles (Sigma-Aldrich) in glycerol, which is pumped through the flow cuvette F (Helma, 6.2 µl; optical path length, 100 µm) using the pressure chamber and compressor at a rate of around 7 ml h^{-1}. With this setup, the measured decorrelation time is around 30 ms (Supplementary Information, section 1). We chose the time interval between two successive frames (33.3 ms) so that the two successive frames are uncorrelated, and we chose the exposure time (200 µs) so that the speckle is stable during the measurement of one frame. We use a DMD area of 5 × 5 pixels (68.5 µm × 68.5 µm) as a model dynamic target and we record a sequence of reflected speckle patterns with different target positions, whereas the scattering liquid provides different realizations of the optical disorder for each speckle. With our optical setup, the DMD is directly imaged onto the sample plane, with a magnification factor of 0.18.

Data collection

For each optical thickness, a total of 5×10^5 images composed of 128 × 128 pixels (1 pixel corresponding to an area of 6.5 µm^2) were taken in the following sequence: B, S, B, D, B, S, B, D..., where B represents the background speckle measured with all the DMD pixels in the 'off' position (that is, none of the measured light comes from the DMD), S represents the static target dataset (with the target fixed at the centre of the field of view) and D represents the dynamic target dataset (with the target moving in a snake pattern across a 5 × 5 grid of positions separated by two DMD pixels both vertically and horizontally).

The collected background images were used to subtract the slowly changing speckle pattern occurring due to TiO_2 particles attaching to the walls of the cuvette. The time series of each pixel of the background dataset was filtered using a Savitzky–Golay filter (5,000-sample window or approximately 170 s) to form a dynamic background signal, which is then subtracted from the consecutive static and dynamic dataset images. Note that the subtraction of the slowly varying average background occurs on a timescale (170 s) that is much larger than the decorrelation time (30 ms), which means that background subtraction cannot allow us to isolate the ballistic light. We observed that without background subtraction, ANNs perform much worse: indeed, since we use a data augmentation procedure, ANNs incorrectly learn to rely on the numerically shifted background features during training.

The static dataset images were numerically shifted both horizontally and vertically by a random number of pixels (extracted from a uniform distribution) within the interval [−40, 40] pixels. Both static and dynamic datasets were then resized to obtain 32 × 32 images by pixel binning and were then normalized from 0 to 1 by dividing each image by its maximum intensity value. As a result of the procedure above, we obtained a training dataset of 1.25×10^5 images (32 × 32 pixels) and a testing dataset of the same size with 25 target positions. Note that the size of the testing dataset has been chosen to be sufficiently large so that we could assess the variance in ANN predictions (5,000 patterns for each of the 25 different positions of the target).

ANN structure and optimization

We have observed that for the task of predicting the target coordinates on a discrete grid, a classification approach works better than a regression approach. For each architecture, the target's vertical and horizontal coordinates were one-hot encoded into vectors of N_h length, which were then merged into a $2N_h$-length vector that served as the ground truth for the network. We used categorical cross-entropy loss and softmax activation for the final layer, enabling the network to produce the probability density of the target position. The final prediction for the target position was taken as the expected value of the horizontal and vertical position distributions computed from a single frame in the test dataset. For each architecture, we performed a hyperparameter grid search. The optimization space for each model was selected in such a way that the maximal number of parameters would be similar for different network architectures. The details about the search parameter space and the optimal hyperparameter values used in the training can be found in Supplementary Information, section 4.2.

The CoordConv architecture is found to be the optimal architecture for our task. Although the bias it develops is small, it is still non-negligible. We correct this bias by splitting a part of a training set, making predictions from it and comparing it with the ground truth. We fit a two-dimensional spline function to the set of points for which we can infer the bias in that way so that this function covers the whole field of view (Supplementary Information, section 4.4).

Numerical simulations

To generate the Gaussian dataset, we sample 125,000 data points from a multidimensional Gaussian distribution with independent and identically distributed entries, number of dimensions $N_{dim} = 50$, mean value $\mu = \theta$ and standard deviation $\sigma = 1$. To introduce correlations, we apply a matrix consisting of ones along the main diagonal and the first off-diagonals to every data point. For the image data, this would correspond to neighbouring pixels being strongly correlated but distant pixels being independent. We then estimate the Fisher information of this correlated dataset for the parameter θ, which is known analytically to be $\mathcal{J}(\theta) = N_{dim}$ (ref. 45). To generate the non-Gaussian dataset, we repeat the procedure above with the difference that we apply an additional nonlinear transformation $f(x) = x^3$ to each component of the correlated random variable. We choose this transformation as it preserves the Fisher information (this can be demonstrated, e.g., using singular value decomposition). Hence, the value of the Fisher information is still $\mathcal{J}(\theta) = N_{dim}$, and we can verify whether our method returns an estimated Fisher information that is close to the true Fisher information.

In Fig. 2b,c (insets), we show the Gaussian and non-Gaussian distributions, respectively, by showing point clouds of two generic components X_1 and X_2 of the 50-dimensional random variables X_k. Although the Gaussian distribution has an elliptical structure due to correlations, the non-Gaussian distribution clearly has a more complex shape. The blue curves in Fig. 2b,c correspond to the dependence of the estimator of the total Fisher information $\sum_k \hat{\mathcal{J}}^k$ on the step size $\Delta\theta$ of the finite-difference scheme for the Gaussian data and non-Gaussian data, respectively. These values are normalized by the true value of the Fisher information (known analytically), such that a value of 1 corresponds to a perfectly accurate estimate.

In these simulations, the optimal value of the step size $\Delta\theta_{thr}$ is selected by choosing the value at which the fluctuations of the second derivative of the curve become greater than those for large step sizes

by a factor of 10 (Supplementary Information, section 3.4). We observe that in both cases, the estimated Fisher information evaluated at this step size is close to the true Fisher information.

Data availability

The data generated in this study and scripts for data processing have been deposited in the University of Glasgow's repository for research data with the following URL: https://doi.org/10.5525/gla.researchdata.1926.

Acknowledgements

I.S. and D.F. acknowledge financial support from the UK Engineering and Physical Sciences Research Council (EPSRC grant nos. EP/T00097X/1 and EP/Y029097/1). D.F. acknowledges support from the UK Royal Academy of Engineering Chairs in Emerging Technologies Scheme. L.M.R. and S.R. were supported by the Austrian Science Fund (FWF) through project no. P32300 (WAVELAND).

Author contributions

D.B. and I.S. conceived the idea and initiated the project. I.S. performed the experiments and data analysis under the supervision of D.F. and D.B. M.W. developed the numerical approach to calculate the Cramér–Rao bound and helped with the data analysis under the supervision of L.M.R., S.R. and D.B. G.H. helped in developing numerical methods and with the data analysis. I.S., M.W. and D.B. wrote the manuscript, with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s41566-025-01657-6.

Correspondence and requests for materials should be addressed to Dorian Bouchet.

Peer review information *Nature Photonics* thanks Christophe Moser and the other, anonymous, reviewer(s) for their contribution to the peer review of this work.