nature portfolio

https://doi.org/10.1038/s41567-020-01137-4

Supplementary information

Maximum information states for coherent scattering measurements

In the format provided by the authors and unedited

Maximum information states for coherent scattering measurements

Supplementary Information

Dorian Bouchet[1], Stefan Rotter[2], and Allard P. Mosk[1]

[1] *Nanophotonics, Debye Institute for Nanomaterials Science, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, the Netherlands*

[2] *Institute for Theoretical Physics, Vienna University of Technology (TU Wien), A-1040 Vienna, Austria*

Supplementary Section S1 – Fisher information for the optimal detection scheme

In this section, we calculate the Fisher information $\mathcal{J}(\theta)$ for the optimal homodyne detection scheme, which we show to be equal to the quantum Fisher information $\mathcal{I}(\theta)$ for coherent states and statistically-independent noise fluctuations. We then define the Fisher information operator from the scattering matrix of the system, and we provide the expression of the minimum variance unbiased estimator. We demonstrate that maximum information states can also be iteratively identified by performing time-reversal of a small perturbation using phase conjugation. Finally, we show that the Fisher information operator can be used in the context of multiple parameter estimations, in order to maximize the trace of the Fisher information matrix.

S1.1 – Fisher information

Let us describe the measured data by a N-dimensional random variable X and a joint probability density function $p(X;\theta)$ parameterized by an arbitrary parameter θ. In general, the Fisher information is expressed by[1]

$$\mathcal{J}(\theta) = \mathbb{E}\left([\partial_\theta \ln p(X;\theta)]^2\right) \ , \tag{S1}$$

where $\mathbb{E}$ denotes the expectation operator acting over noise fluctuations. In the shot-noise limit, $p(X;\theta)$ can be obtained using two equivalent strategies[2]:

- Using a fully quantized picture, we can consider that $p(X;\theta)$ is determined by the quantum nature of the field. We define $|\alpha_k\rangle$ as being the single-mode coherent (Glauber) state for the k-th outgoing spatial mode. Superposing this state with a

reference coherent state $|r_k\rangle$, the occupation of the state follows a Poisson distribution of expectation value $|\alpha_k + r_k|^2$, where α_k and r_k are the eigenvalues of the annihilation operator $\hat{a}$ for the states $|\alpha_k\rangle$ and $|r_k\rangle$, respectively.

- Using a semi-classical picture, we can consider that $p(X;\theta)$ is determined by the quantum nature of the photodetection process. We define E_k^{out} as being the complex value of the classical field for the k-th outgoing spatial mode. Superposing this field with a reference field E_k^{ref}, the signal measured by the photodetector follows a Poisson distribution of expectation value $|E_k^{\text{out}} + E_k^{\text{ref}}|^2$.

Since the annihilation operator $\hat{a}$ is also the complex amplitude operator for coherent states[2], we can write $E_k^{\text{out}} = \langle\alpha_k|\hat{a}|\alpha_k\rangle = \alpha_k$ and $E_k^{\text{ref}} = \langle r_k|\hat{a}|r_k\rangle = r_k$, evidencing the equivalence between the two approaches.

Adopting the semi-classical notation and considering that noise fluctuations are statistically independent for any two different outgoing modes, the joint probability density function $p(X;\theta)$ is thus expressed by

$$p(X;\theta) = \prod_{k=1}^{N} e^{-|E_k^{\text{out}}+E_k^{\text{ref}}|^2} \frac{|E_k^{\text{out}} + E_k^{\text{ref}}|^{2X_k}}{X_k!} \,. \tag{S2}$$

Injecting this expression in Eq. (S1), we obtain

$$\mathcal{J}(\theta) = \sum_{k=1}^{N} \frac{\left[\partial_\theta \left(|E_k^{\text{out}} + E_k^{\text{ref}}|^2\right)\right]^2}{|E_k^{\text{out}} + E_k^{\text{ref}}|^2} \,. \tag{S3}$$

For a high-intensity reference field which does not depend on θ (i.e., for $|E_k^{\text{ref}}|^2 \gg |E_k^{\text{out}}|^2$ and $\partial_\theta E_k^{\text{ref}} = 0$), this expression simplifies to

$$\mathcal{J}(\theta) = \sum_{k=1}^{N} \frac{\left(\partial_\theta \left[E_k^{\text{out}} \left(E_k^{\text{ref}}\right)^* + E_k^{\text{ref}} \left(E_k^{\text{out}}\right)^*\right]\right)^2}{|E_k^{\text{ref}}|^2} \,. \tag{S4}$$

We now introduce $E_k^{\text{out}} = Q_k^{\text{out}} + iP_k^{\text{out}}$ and $E_k^{\text{ref}} = |E_k^{\text{ref}}|e^{i\phi_k}$. Equation (S4) becomes

$$\mathcal{J}(\theta) = 4\sum_{k=1}^{N} \left[(\partial_\theta Q_k^{\text{out}})\cos\phi_k + (\partial_\theta P_k^{\text{out}})\sin\phi_k\right]^2 \,. \tag{S5}$$

For any integer m, the phase angles of the reference field which maximize Eq. (S5) are given by $\phi_k^{\max} = \arg(\partial_\theta E_k^{\text{out}}) + m\pi$, and those which minimize Eq. (S5) are given by $\phi_k^{\min} = \arg(\partial_\theta E_k^{\text{out}}) + (m + 1/2)\pi$. Choosing $\phi_k^{\min}$ as phase angles of the reference field yields $\mathcal{J}(\theta) = 0$. In contrast, choosing $\phi_k^{\max}$ as phase angles of the reference field yields

$$\mathcal{J}(\theta) = 4\sum_{k=1}^{N} |\partial_\theta E_k^{\text{out}}|^2 \,. \tag{S6}$$

This expression can be identified as the Fisher information for a random vector composed of N complex random variables whose real and imaginary parts are independent normally distributed random variables with variance $\sigma^2 = 1/4$.

S1.2 – Quantum Fisher information

The Fisher information $\mathcal{J}(\theta)$ sets a lower bound on the variance of unbiased estimators of θ for a definite measurement scheme (i.e. the homodyne scheme in our case). A more general lower bound exists, which applies to any quantum measurement described by a positive-operator-valued measure (POVM). This lower bound is given by the reciprocal of the quantum Fisher information $\mathcal{I}(\theta)$, which is defined by [3]

$$\mathcal{I}(\theta) = \mathrm{Tr}(\rho_{\mathrm{out}} L_{\mathrm{out}}^2) \ , \tag{S7}$$

where ρ_{out} is a density operator describing the quantum state of the system and L_{out} is the symmetrized logarithmic derivative of ρ_{out} with respect to θ defined as follows:

$$\rho_{\mathrm{out}} L_{\mathrm{out}} + L_{\mathrm{out}} \rho_{\mathrm{out}} = 2\, \partial_\theta \rho_{\mathrm{out}} \ . \tag{S8}$$

The Fisher information $\mathcal{J}(\theta)$ and the quantum Fisher information $\mathcal{I}(\theta)$ satisfy the inequality $\mathcal{I}(\theta) \geq \mathcal{J}(\theta)$, which is saturated when the POVM considered for the calculation of $\mathcal{J}(\theta)$ is optimal.

In our model, ρ_{out} describes an N-mode coherent state composed of simply-separable pure states, and thus Eq. (S7) simplifies to[4,5]

$$\mathcal{I}(\theta) = 4 \sum_{k=1}^{N} \left(\langle \partial_\theta \alpha_k | \partial_\theta \alpha_k \rangle - |\langle \partial_\theta \alpha_k | \alpha_k \rangle|^2 \right) \ , \tag{S9}$$

where $|\alpha_k\rangle$ is the single-mode coherent state associated with the k-th mode and $|\partial_\theta \alpha_k\rangle$ is its derivative with respect to θ. We can represent $|\alpha_k\rangle$ in the basis of Fock states $|n\rangle$ labeled by the occupation number n, which reads[2]

$$|\alpha_k\rangle = e^{-|\alpha_k|^2/2} \sum_{n=0}^{\infty} \frac{\alpha_k^n}{(n!)^{1/2}} |n\rangle \ , \tag{S10}$$

where α_k denotes the eigenvalue of the annihilation operator. Taking the derivative of this expression with respect to θ leads to

$$|\partial_\theta \alpha_k\rangle = -\,\mathrm{Re}\left(\alpha_k \partial_\theta \alpha_k^*\right) e^{-|\alpha_k|^2/2} \sum_{n=0}^{\infty} \frac{\alpha_k^n}{(n!)^{1/2}} |n\rangle + \partial_\theta \alpha_k e^{-|\alpha_k|^2/2} \sum_{n=1}^{\infty} \frac{n \alpha_k^{n-1}}{(n!)^{1/2}} |n\rangle \ . \tag{S11}$$

We can now use Eqs. (S10) and (S11) to calculate the two terms $\langle \partial_\theta \alpha_k | \partial_\theta \alpha_k \rangle$ and $|\langle \partial_\theta \alpha_k | \alpha_k \rangle|^2$ that appear in Eq. (S9). Let us first calculate $\langle \partial_\theta \alpha_k | \partial_\theta \alpha_k \rangle$ from Eq. (S11): using the orthonormality of Fock states, we obtain

$$\begin{aligned} \langle \partial_\theta \alpha_k | \partial_\theta \alpha_k \rangle = & \,\mathrm{Re}\left(\alpha_k \partial_\theta \alpha_k^*\right)^2 e^{-|\alpha_k|^2} \sum_{n=0}^{\infty} \frac{|\alpha_k|^{2n}}{n!} - \mathrm{Re}\left(\alpha_k \partial_\theta \alpha_k^*\right) \alpha_k \partial_\theta \alpha_k^* e^{-|\alpha_k|^2} \sum_{n=1}^{\infty} \frac{|\alpha_k|^{2(n-1)}}{(n-1)!} \\ & - \mathrm{Re}\left(\alpha_k \partial_\theta \alpha_k^*\right) \alpha_k^* \partial_\theta \alpha_k e^{-|\alpha_k|^2} \sum_{n=1}^{\infty} \frac{|\alpha_k|^{2(n-1)}}{(n-1)!} + |\partial_\theta \alpha_k|^2 e^{-|\alpha_k|^2} \sum_{n=1}^{\infty} \frac{n |\alpha_k|^{2(n-1)}}{(n-1)!}. \end{aligned} \tag{S12}$$

This expression can be simplified by using the following properties of exponential series:

$$\sum_{n=0}^{\infty} \frac{|\alpha_k|^{2n}}{n!} = e^{|\alpha_k|^2} , \tag{S13}$$

$$\sum_{n=0}^{\infty} \frac{(n+1)|\alpha_k|^{2n}}{n!} = (1+|\alpha_k|^2)e^{|\alpha_k|^2} . \tag{S14}$$

Injecting Eqs. (S13) and (S14) into Eq. (S12) leads to

$$\langle \partial_\theta \alpha_k | \partial_\theta \alpha_k \rangle = |\partial_\theta \alpha_k|^2 + \mathrm{Im}(\alpha_k \partial_\theta \alpha_k^*)^2 . \tag{S15}$$

Let us now calculate $\langle \partial_\theta \alpha_k | \alpha_k \rangle$ from Eqs. (S10) and (S11): using again the orthonormality of Fock states, we obtain

$$\langle \partial_\theta \alpha_k | \alpha_k \rangle = -\,\mathrm{Re}\,(\alpha_k \partial_\theta \alpha_k^*)\, e^{-|\alpha_k|^2} \sum_{n=0}^{\infty} \frac{|\alpha_k|^{2n}}{n!} + \alpha_k \partial_\theta \alpha_k^* e^{-|\alpha_k|^2} \sum_{n=1}^{\infty} \frac{|\alpha_k|^{2(n-1)}}{(n-1)!} . \tag{S16}$$

Injecting Eq. (S13) into Eq. (S16) leads to

$$|\langle \partial_\theta \alpha_k | \alpha_k \rangle|^2 = \mathrm{Im}(\alpha_k \partial_\theta \alpha_k^*)^2 . \tag{S17}$$

Finally, injecting Eqs. (S15) and (S17) into Eq. (S9) results in

$$\mathcal{I}(\theta) = 4 \sum_{k=1}^{N} |\partial_\theta \alpha_k|^2 . \tag{S18}$$

Recalling that $\alpha_k = E_k^{\mathrm{out}}$, the Fisher information $\mathcal{J}(\theta)$ given by Eq. (S6) equals the quantum Fisher information $\mathcal{I}(\theta)$ given by Eq. (S18), thereby demonstrating that the homodyne detection scheme that we considered in Section S1.1 constitutes the optimal POVM for the estimation of θ.

S1.3 – Fisher information operator

In the formalism of the S-matrix, the outgoing field state is expressed by $|E^{\mathrm{out}}\rangle = S|E^{\mathrm{in}}\rangle$, where S denotes the scattering matrix of the system. (Note that the incident state $|E^{\mathrm{in}}\rangle$ and the outgoing state $|E^{\mathrm{out}}\rangle$ are defined here in the Hilbert space of all incident and outgoing spatial modes, whereas $|\alpha_k\rangle$ introduced in Section S1.1 was defined in the Hilbert space of all quantum states for the k-th outgoing spatial mode.) We can thus write E_k^{out} as a projection of the outgoing state on the state associated with the k-th mode, which in bra-ket notation reads $E_k^{\mathrm{out}} = \langle k|S|E^{\mathrm{in}}\rangle$, leading to

$$\mathcal{J}(\theta) = 4 \sum_{k=1}^{N} |\langle k|\partial_\theta S|E^{\mathrm{in}}\rangle|^2 . \tag{S19}$$

This expression can be expanded into

$$\mathcal{J}(\theta) = 4 \sum_{k=1}^{N} \langle E^{\mathrm{in}}|(\partial_\theta S)^\dagger|k\rangle\langle k|\partial_\theta S|E^{\mathrm{in}}\rangle . \tag{S20}$$

Using the completeness relation $\sum_k |k\rangle\langle k| = I_N$ where I_N is the N-dimensional identity matrix, we finally obtain

$$\mathcal{J}(\theta) = 4\ \langle E^{\text{in}}|(\partial_\theta S)^\dagger \partial_\theta S|E^{\text{in}}\rangle\ . \tag{S21}$$

In this expression, we can identify the operator $F_\theta = (\partial_\theta S)^\dagger \partial_\theta S$, which we refer to as *Fisher information operator*. We obtain the quadratic form $\mathcal{J}(\theta) = 4\langle E^{\text{in}}|F_\theta|E^{\text{in}}\rangle$, which is the expression of the Fisher information given in the manuscript. Finally, for a unitary scattering matrix ($S^\dagger = S^{-1}$), the operator F_θ is expressed by

$$F_\theta = (-iS^{-1}\partial_\theta S)^2\ , \tag{S22}$$

where we used the identity $\partial_\theta S^{-1} = -S^{-1}(\partial_\theta S)S^{-1}$. Introducing the generalized Wigner-Smith operator[6], which is defined by $Q_\theta = -iS^{-1}\partial_\theta S$, we obtain the identity

$$F_\theta = Q_\theta^2\ , \tag{S23}$$

which implies that F_θ and Q_θ share the same eigenstates. Writing the eigenvalue equation $Q_\theta|\mathcal{E}_j^{\text{in}}\rangle = \Theta_j|\mathcal{E}_j^{\text{in}}\rangle$ where Θ_j is the j-th eigenvalue of Q_θ and $|\mathcal{E}_j^{\text{in}}\rangle$ is the associated eigenstate, the outgoing field state satisfies $|\partial_\theta \mathcal{E}_j^{\text{out}}\rangle = i\Theta_j|\mathcal{E}_j^{\text{out}}\rangle$. For a scattering matrix evaluated at θ_0 and a small parameter variation $\Delta\theta = \theta - \theta_0$, this results in $|\mathcal{E}_j^{\text{out}}(\theta)\rangle \simeq e^{i\Theta_j\Delta\theta}|\mathcal{E}_j^{\text{out}}(\theta_0)\rangle$. Thus, in the limit of a unitary scattering matrix, maximum information states are insensitive with respect to small variations in θ except for a global phase factor.

S1.4 – Minimum variance unbiased estimator

For small parameter variations around a given parameter value noted θ_0, the measured data can be described by the following linear model:

$$X_k = I_k^{\text{tot}} + (\partial_\theta I_k^{\text{tot}})(\theta - \theta_0) + W_k\ , \tag{S24}$$

where X_k represents the intensity data measured by the camera at the k-th sampling point, where I_k^{tot} and $\partial_\theta I_k^{\text{tot}}$ are evaluated at θ_0, and where W_k are N independent and normally distributed random variables with mean zero and variance I_k^{tot}. The normal distribution is indeed a good approximation of the Poisson distribution for large expectation values. In the case of linear models, general expressions exist for the minimum variance unbiased estimator, which depends on the noise statistics[1]. For the linear model expressed by Eq. (S24), the minimum variance unbiased estimator reads

$$\hat{\theta}(X) - \theta_0 = \frac{1}{\mathcal{J}(\theta_0)} \sum_{k=1}^{N} \frac{(\partial_\theta I_k^{\text{tot}})(X_k - I_k^{\text{tot}})}{I_k^{\text{tot}}}\ , \tag{S25}$$

where $\hat{\theta}(X)$ denotes the estimator of θ (i.e. the function devised to estimate θ from the measured data X). Writing $I_k^{\text{tot}} = |E_k^{\text{out}} + E_k^{\text{ref}}|^2$ and recalling that $|E_k^{\text{ref}}|^2 \gg |E_k^{\text{out}}|^2$ and $\partial_\theta E_k^{\text{ref}} = 0$, this expression simplifies to

$$\hat{\theta}(X) - \theta_0 = \frac{1}{\mathcal{J}(\theta_0)} \sum_{k=1}^{N} \frac{\left(\partial_\theta \left[E_k^{\text{out}} \left(E_k^{\text{ref}}\right)^* + E_k^{\text{ref}} \left(E_k^{\text{out}}\right)^*\right]\right)(X_k - |E_k^{\text{out}} + E_k^{\text{ref}}|^2)}{|E_k^{\text{ref}}|^2}\ . \tag{S26}$$

Introducing $E_k^{\text{ref}} = |E_k^{\text{ref}}|e^{i\phi_k}$ and using the phase angles $\phi_k^{\max} = \arg(\partial_\theta E_k^{\text{out}})$ which maximize the Fisher information leads to

$$\hat{\theta}(X) - \theta_0 = \frac{2}{\mathcal{J}(\theta_0)} \sum_{k=1}^{N} \frac{|\partial_\theta E_k^{\text{out}}|(X_k - |E_k^{\text{out}} + E_k^{\text{ref}}|^2)}{|E_k^{\text{ref}}|} \,. \tag{S27}$$

In this expression, $\mathcal{J}(\theta_0)$ is obtained from Eq. (S6), E_k^{out} is obtained from $E_k^{\text{out}} = \langle k|S|E^{\text{in}}\rangle$ and $\partial_\theta E_k^{\text{out}}$ is obtained from $\partial_\theta E_k^{\text{out}} = \langle k|\partial_\theta S|E^{\text{in}}\rangle$, with a scattering matrix S evaluated at θ_0.

S1.5 – Finding maximum information states using phase conjugation

Time reversal using phase conjugation is a well-known technique that allows for instance to focus waves[7–9] or to identify open channels[10] in multiple scattering media. In some implementations, light waves are focused into a scattering medium by performing time-reversal of a perturbation using phase conjugation[11–13]. Here, we show that an iterative procedure based on such techniques actually converges toward maximum information states. Light waves are then focused onto those specific areas of the object that are most affected by the perturbation, in such a way that the Fisher information available to the observer is maximized. In order to demonstrate the link between maximum information states and time-reversed adapted perturbation, we must restrict the analysis to a $N \times N$ scattering matrix that satisfies the reciprocity relation $S^{\mathsf{T}} = S$ (or equivalently $S^\dagger = S^*$). The procedure then relies on an iterative approach used to calculate the incident state $|E_{(n)}^{\text{in}}\rangle$ at the n-th iteration from measurements based on an incident state $|E_{(n-1)}^{\text{in}}\rangle$. The medium is illuminated with the incident state $|E_{(n-1)}^{\text{in}}\rangle$ and, before a perturbation $\Delta\theta$ has occurred, the measured outgoing state is $|E_{(n-1)}^{\text{out}}(\theta)\rangle = S(\theta)|E_{(n-1)}^{\text{in}}\rangle$. For the same incident state, the outgoing state measured after the perturbation has occurred is $|E_{(n-1)}^{\text{out}}(\theta+\Delta\theta)\rangle = S(\theta+\Delta\theta)|E_{(n-1)}^{\text{in}}\rangle$. The new incident state $|E_{(n)}^{\text{in}}\rangle$ is then obtained by phase-conjugating the difference between these two outgoing states. In the limit of a small perturbation $\Delta\theta$, this leads to

$$|E_{(n)}^{\text{in}}\rangle = A_{(n)}^{-1/2}\, \partial_\theta S^* |E_{(n-1)}^{\text{in}}\rangle^* \,, \tag{S28}$$

where $A_{(n)}^{-1/2}$ is a (real-valued) normalization coefficient. Using the relation $S^* = S^\dagger$, we can write $\partial_\theta S^* = \partial_\theta S^\dagger$. Then, starting from a random incident state $|E_{(0)}^{\text{in}}\rangle$ and after n successive iterations, we obtain

$$|E_{(n)}^{\text{in}}\rangle = B_{(n)}^{-1/2} \left[(\partial_\theta S)^\dagger \partial_\theta S\right]^{n/2} |E_{(0)}^{\text{in}}\rangle \qquad \text{if } n \text{ is even}, \tag{S29}$$

$$|E_{(n)}^{\text{in}}\rangle = B_{(n)}^{-1/2} \left[(\partial_\theta S)^\dagger \partial_\theta S\right]^{n/2-1} \partial_\theta S^\dagger |E_{(0)}^{\text{in}}\rangle^* \qquad \text{if } n \text{ is odd}. \tag{S30}$$

Choosing the normalization condition $\langle E_{(n)}^{\text{in}}|E_{(n)}^{\text{in}}\rangle = 1$, the (real-valued) normalization coefficient $B_{(n)}$ is expressed by

$$B_{(n)} = \langle E_{(0)}^{\text{in}}| \left[(\partial_\theta S)^\dagger \partial_\theta S\right]^n |E_{(0)}^{\text{in}}\rangle \,. \tag{S31}$$

When the medium is illuminated with the incident state $|E^{\text{in}}_{(n)}\rangle$, the Fisher information associated with the resulting outgoing state reads

$$\mathcal{J}_{(n)}(\theta) = 4\langle E^{\text{in}}_{(n)}|(\partial_\theta S)^\dagger \partial_\theta S|E^{\text{in}}_{(n)}\rangle \,. \tag{S32}$$

Using either Eq. (S29) if n is even or Eq. (S30) if n is odd, Eq. (S32) becomes

$$\mathcal{J}_{(n)}(\theta) = 4\,\frac{\langle E^{\text{in}}_{(0)}|[(\partial_\theta S)^\dagger \partial_\theta S]^{n+1}|E^{\text{in}}_{(0)}\rangle}{\langle E^{\text{in}}_{(0)}|\,[(\partial_\theta S)^\dagger \partial_\theta S]^{n}\,|E^{\text{in}}_{(0)}\rangle} \,. \tag{S33}$$

Any incident state can be decomposed in the orthonormal basis formed by the eigenstates of F_θ. In this basis, the initial state $|E^{\text{in}}_{(0)}\rangle$ is expressed by

$$|E^{\text{in}}_{(0)}\rangle = \sum_{j=1}^{N} \gamma_j |\mathcal{E}^{\text{in}}_j\rangle \,, \tag{S34}$$

where $|\mathcal{E}^{\text{in}}_j\rangle$ is the j-th eigenstate of F_θ and $\gamma_j = \langle \mathcal{E}^{\text{in}}_j|E^{\text{in}}_{(0)}\rangle$. From Eq. (S34) and using the eigenvalue equation $F_\theta|\mathcal{E}^{\text{in}}_j\rangle = \Lambda_j|\mathcal{E}^{\text{in}}_j\rangle$, Eq. (S33) becomes

$$\mathcal{J}_{(n)}(\theta) = 4\,\frac{\sum_{j=1}^{N} |\gamma_j|^2 \Lambda_j^{n+1}}{\sum_{j=1}^{N} |\gamma_j|^2 \Lambda_j^{n}} \,. \tag{S35}$$

This expression can also be written in the following form:

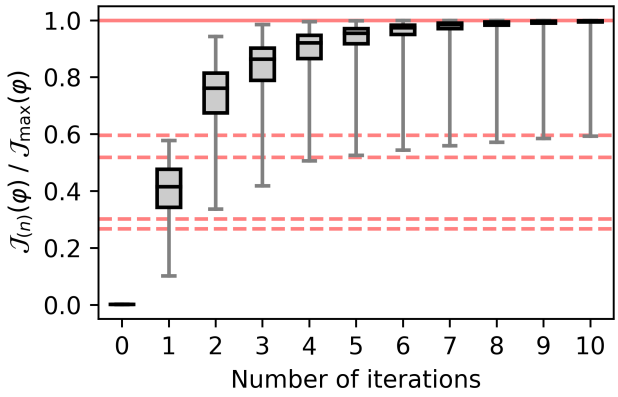
$$\mathcal{J}_{(n)}(\theta) = 4\Lambda_{\max} \times \frac{\sum_{j=1}^{N} |\gamma_j|^2 (\Lambda_j/\Lambda_{\max})^{n+1}}{\sum_{j=1}^{N} |\gamma_j|^2 (\Lambda_j/\Lambda_{\max})^{n}} \,, \tag{S36}$$

where $\Lambda_{\max}$ is the largest eigenvalue of F_θ. Assuming that $|\gamma_j|^2 \neq 0$ for the associated eigenstate, we end up with

$$\lim_{n\to+\infty} \mathcal{J}_{(n)} = 4\Lambda_{\max} \,. \tag{S37}$$

Thus, performing time-reversal of a perturbation using phase conjugation allows one to iteratively identify the maximum information state relative to the perturbation, provided that the scattering matrix describing the medium is a square matrix satisfying the reciprocity condition $S^{\mathsf{T}} = S$, and that the overlap between the initial state and the maximum information state is not equal to zero.

It clearly appears that the eigenstates of F_θ constitute a relevant basis to analyze this iterative procedure. Indeed, the number of iterations needed for the procedure to converge depends on both the values of γ_j and the eigenvalue spectrum of F_θ, according to Eq. (S35). In order to illustrate this feature, we consider the operator $f_\varphi = (\partial_\varphi r)^\dagger \partial_\varphi r$ measured by shifting the phase φ induced by the cross-shaped object. We numerically generate 10 000 random initial states and we use Eq. (S35) to calculate the ratio $\mathcal{J}_{(n)}(\varphi)/\mathcal{J}_{\max}(\varphi)$, where $\mathcal{J}_{\max}(\varphi)$ is the Fisher information associated with the maximum information state. After the first iteration, the median value of the ratio $\mathcal{J}_{(n)}(\varphi)/\mathcal{J}_{\max}(\varphi)$ is equal to 0.42 (Fig. S1), and we observe that 4 iterations are generally required to identify a state that reaches 90% of the optimal Fisher information. Moreover, some initial states show a small overlap with the maximum information state; for such states, the ratio $\mathcal{J}_{(n)}(\varphi)/\mathcal{J}_{\max}(\varphi)$ after 10 iterations is close to 0.6, corresponding to the second largest eigenvalue of f_φ.



Supplementary Figure S1 | Fisher information maximization using time-reversed adapted perturbation. Fisher information associated with light states identified using time-reversed adapted perturbation for 10 000 random initial states, normalized by the Fisher information of the maximum information state. A horizontal line goes through each box at the median value, edges of the boxes represent lower and upper quartiles, and whiskers represent minimum and maximum values. The box and whiskers for 0 iterations appear as a single horizontal line; indeed, random initial states generate a Fisher information that is much smaller than the maximum information state. The solid red line represents the Fisher information for the largest eigenvalue of f_φ, and the dashed red lines represent the Fisher information for the second, third, fourth and fifth largest eigenvalues.

S1.6 – Multiple parameter estimations

When several parameters $\theta = (\theta_1, \ldots, \theta_p)^\mathsf{T}$ need to be simultaneously estimated, the Cramér-Rao inequality reads[1]

$$\mathrm{Var}\left(\hat{\theta}_i\right) \geq \left[\mathcal{J}^{-1}(\theta)\right]_{ii} \ , \tag{S38}$$

where $\hat{\theta}_i$ is any unbiased estimator of θ_i, Var denotes the variance operator and $\mathcal{J}(\theta)$ is now a $p \times p$ Fisher information matrix defined as follows:

$$\left[\mathcal{J}(\theta)\right]_{ij} = \mathbb{E}\left(\left[\partial_{\theta_i} \ln p(X;\theta)\right]\left[\partial_{\theta_j} \ln p(X;\theta)\right]\right) \ . \tag{S39}$$

Diagonal elements of the Fisher information matrix can then be interpreted as the amount of information relevant to the estimation of each parameter taken individually, while off-diagonal elements describe the influence of possible correlations between estimated values of different parameters.

For scattering measurements performed with the optimal homodyne detection scheme, the trace of the Fisher information matrix is expressed by

$$\mathrm{Tr}\left[\mathcal{J}(\theta)\right] = 4\langle E^{\mathrm{in}}| \sum_{i=1}^{p} F_{\theta_i} |E^{\mathrm{in}}\rangle \ , \tag{S40}$$

where $F_{\theta_i} = (\partial_{\theta_i} S)^\dagger \partial_{\theta_i} S$ is the Fisher information operator associated with the i-th parameter. Thus, the incident state that maximizes the trace of the Fisher information matrix

is found by calculating the eigenstates of the operator $\sum_i F_{\theta_i}$ and by choosing the one associated with the largest eigenvalue. Whenever the Fisher information matrix is diagonal, we can write $[\mathcal{J}^{-1}(\theta)]_{ii} = [\mathcal{J}(\theta)]_{ii}^{-1}$. In this case, maximizing the trace of the Fisher information matrix improves the measurement precision for all parameters simultaneously. Furthermore, any weighted sum of diagonal elements of the Fisher information matrix can be maximized using the same approach, a feature which could be useful to improve the precision for specific parameters of interest. In contrast, if the Fisher information matrix is not diagonal, we can only write $[\mathcal{J}^{-1}(\theta)]_{ii} \geq [\mathcal{J}(\theta)]_{ii}^{-1}$. It is then still needed to maximize the value of diagonal elements but it is also important to minimize the value of off-diagonal elements, thus requiring a different optimization strategy. If an iterative algorithm is used for this purpose, maximizing the trace of the Fisher information matrix can yield a good initial guess for the algorithm.

Supplementary Section S2 – Fisher information in the experiments

In this section, we give the expression of the Fisher information for our optical setup and we provide the expression of the minimum variance unbiased estimator used in the experiments.

S2.1 – Expression of the Fisher information in the experiments

In our proof-of-principle experiments, the reference beam is not optimally shaped but consists of a tilted plane wave. In this case, the Fisher information is obtained by averaging Eq. (S5) over the phase angle of the reference beam:

$$\mathcal{J}(\theta) = \frac{2}{\pi} \sum_{k=1}^{N} \int_0^{2\pi} \mathrm{d}\phi \left[(\partial_\theta Q_k^{\mathrm{out}}) \cos\phi + (\partial_\theta P_k^{\mathrm{out}}) \sin\phi \right]^2 . \tag{S41}$$

We obtain the following simplified expression:

$$\mathcal{J}(\theta) = 2 \sum_{k=1}^{N} |\partial_\theta E_k^{\mathrm{out}}|^2 . \tag{S42}$$

Note that this expression is identical to the expression of the optimal Fisher information given in Eq. (S6), except for a factor of two. Hence, the incoming state which maximizes Eq. (S42) also maximizes Eq. (S6).

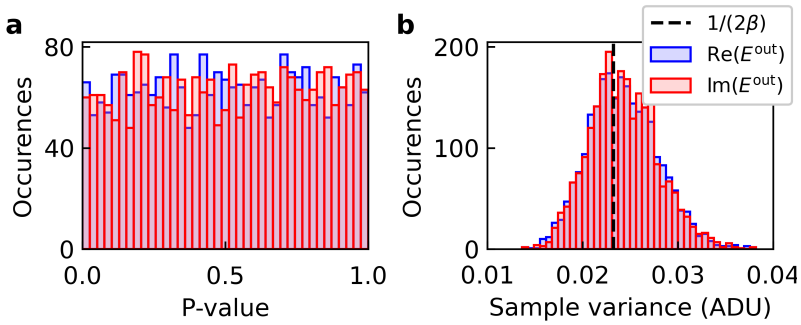
In the experiments, we must also take into account two additional considerations. First, the value of the $N = 2465$ sampling points are actually estimated from $N_{\mathrm{cam}} = 53\,044$ statistically independent camera pixels (oversampling is required in off-axis holography). The oversampling ratio $\beta = N_{\mathrm{cam}}/N$ must therefore be included as a prefactor in Eq. (S42). Moreover, while we measured the reflection matrix without any neutral density filter, we consistently use the strongest neutral density filter ND6 (fractional transmittance $\mathcal{T} = 0.83 \times 10^{-6}$) to compute all values of the Fisher information. Therefore, the expression of

the Fisher information relevant to our experiments is

$$\mathcal{J}(\theta) = \frac{\mathcal{T}}{\sigma^2}\langle E^{\text{in}}|(\partial_\theta r)^\dagger \partial_\theta r|E^{\text{in}}\rangle \ , \tag{S43}$$

where we noted $\sigma^2 = 1/(2\beta)$ in order to highlight that this expression can be identified as the Fisher information for a random vector composed of N complex random variables whose real and imaginary parts are independent normally distributed random variables with variance σ^2. The precision limit σ_{CRB}, which bounds the standard deviation of any estimator of θ in our experiments, is then expressed from Eq. (S43) by $\sigma_{\text{CRB}} = \mathcal{J}^{-1/2}$.

We can verify that the measured field quadratures follow a Gaussian distribution when the neutral density filter ND6 is in the signal path. To this end, we illuminate the medium with the optimal incident state relative to the estimation of φ and we perform 100 successive measurements of the outgoing field. We then compute the p-value for each sampling point according to the Shapiro-Wilk test. The uniform distribution of p-value (Fig. S2a) confirms that the measured field quadratures follow a Gaussian distribution. We also construct the histogram of the sample variance (Fig. S2b), with a mean value of 0.0243 ADU. This value is in good agreement with the theoretical value given by $1/(2\beta) = 0.0232$ ADU. We consistently choose the mean value experimentally determined ($\sigma^2 = 0.0243$ ADU) to calculate the Fisher information in Eq. (S43).



Supplementary Figure S2 | Statistics of the field quadratures. a, P-value distribution for the 2465 sampling points. The p-value for each point is calculated using Shapiro-Wilk test statistics from 100 different realization of the random noise. **b**, Distribution of the sample variance for the 2465 sampling points.

S2.2 – Minimum variance unbiased estimator in the experiments

For small variations of θ around θ_0, measured data can be described by the following linear model:

$$Z_k = E_k^{\text{out}} + (\partial_\theta E_k^{\text{out}})(\theta - \theta_0) + W_k \ , \tag{S44}$$

where Z_k represents here the complex field retrieved from the measured data using off-axis holography and evaluated at the k-th sampling point, where E_k^{out} and $\partial_\theta E_k^{\text{out}}$ are evaluated at θ_0, and where W_k are N independent complex random variables whose real and imaginary

parts are independent normally distributed random variables with mean zero and variance σ^2. For this linear model, the minimum variance unbiased estimator reads[1]

$$\hat{\theta}(Z) - \theta_0 = \frac{\mathrm{Re}\left[\langle \partial_\theta E^{\mathrm{out}} | Z\rangle - \langle \partial_\theta E^{\mathrm{out}} | E^{\mathrm{out}}\rangle\right]}{\langle \partial_\theta E^{\mathrm{out}} | \partial_\theta E^{\mathrm{out}}\rangle} \,. \tag{S45}$$

All estimations presented in the manuscript are performed from data measured with the neutral density filter ND6 (characterized by a fractional transmittance $\mathcal{T}$) placed in the signal optical path. Moreover, due to measurement noise during reflection matrix measurements, the intensity and the Fisher information predicted from the knowledge of the reflection matrix differ from direct measurements by a factor η_{I} and η_{F}, respectively (see Supplementary Information S3). The estimator of θ associated with our experimental setup is thus obtained from Eq. (S45) by taking $|E^{\mathrm{out}}\rangle = (\eta_{\mathrm{I}}\mathcal{T})^{1/2}\, r|\tilde{E}^{\mathrm{in}}\rangle$ and $|\partial_\theta E^{\mathrm{out}}\rangle = (\eta_{\mathrm{F}}\mathcal{T})^{1/2}\, \partial_\theta r|\tilde{E}^{\mathrm{in}}\rangle$. We obtain the following expression:

$$\hat{\theta}(Z) - \theta_0 = \frac{\mathrm{Re}\left[(\mathcal{T}\eta_{\mathrm{F}})^{-1/2}\langle \tilde{E}^{\mathrm{in}}|(\partial_\theta r)^\dagger|Z\rangle - (\eta_{\mathrm{I}}/\eta_{\mathrm{F}})^{1/2}\langle \tilde{E}^{\mathrm{in}}|(\partial_\theta r)^\dagger r|\tilde{E}^{\mathrm{in}}\rangle\right]}{\langle \tilde{E}^{\mathrm{in}}|\,(\partial_\theta r)^\dagger\,\partial_\theta r|\tilde{E}^{\mathrm{in}}\rangle} \,. \tag{S46}$$

Evaluating this estimator using measured data Z directly yields the estimates shown in the manuscript. Importantly, all experimental parameters involved in this expression (i.e. $\mathcal{T}$, η_{I} and η_{F}) are characterized using independent measurements. Note that, whereas estimates shown in the manuscript are almost unbiased, small biases often appear when applying Eq. (S46) to experimental data. Such biases, which are usually smaller than $2\sigma_{\mathrm{CRB}}$, can be explained by the influence of measurement noise in the reflection matrices and by an imperfect correction of the global phase variations induced by thermal instabilities of the setup.

Supplementary Section S3 – Characterization of the outgoing field and its derivative

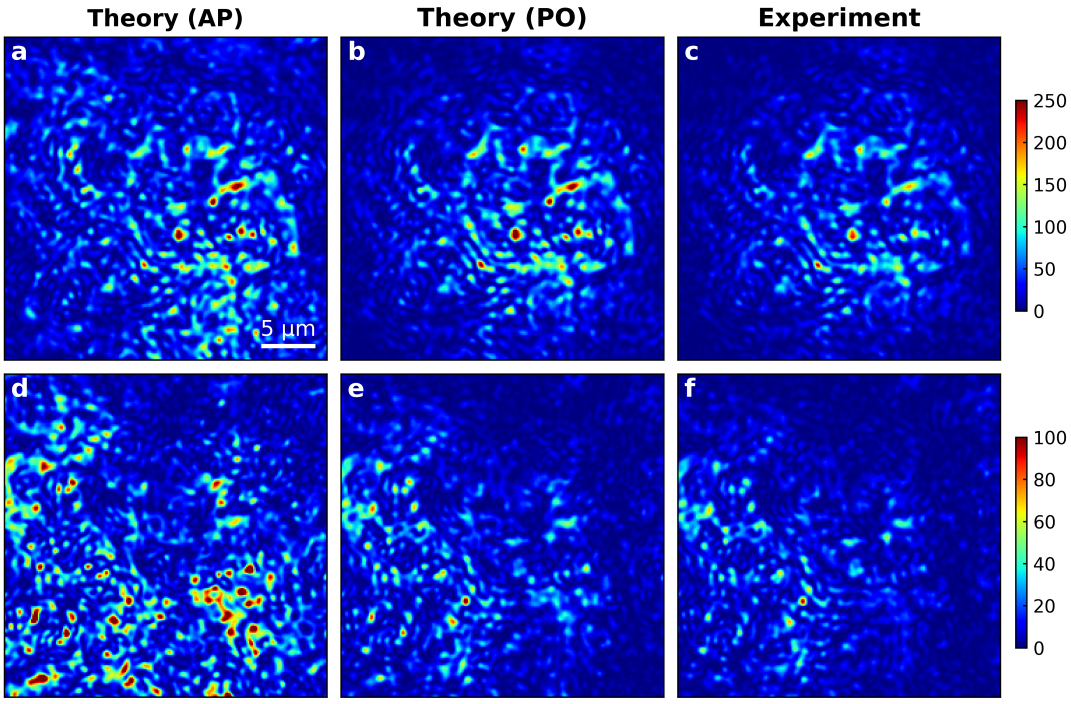
In this section, we show that reflection matrix measurements allow to faithfully predict the outgoing field and its derivatives with respect to the phase and to the position of hidden objects, and we characterize the correlation between the outgoing field distribution and its derivative.

S3.1 – Predicting the outgoing field

In general, illuminating the scattering medium with an arbitrary state $|E^{\mathrm{in}}\rangle$ requires amplitude and phase modulation of the incident field state. However, in the experiments, only the phase of the field can be modulated by the input SLM. We must therefore determine the expression of the phase-only modulated state $|\tilde{E}^{\mathrm{in}}\rangle$ that is experimentally used to illuminate the medium instead of $|E^{\mathrm{in}}\rangle$. Let us define the SLM pattern $|E_{\mathrm{SLM}}\rangle = M|E^{\mathrm{in}}\rangle$, where M is a transformation matrix mapping incident states (expressed in a basis of plane waves) to SLM patterns (expressed in a basis of SLM pixels). The phase-only modulated state can be numerically approximated by $|\tilde{E}^{\mathrm{in}}\rangle = \mathrm{argmin}(\||\tilde{E}_{\mathrm{SLM}}\rangle - M|E^{\mathrm{in}}\rangle\|, |E^{\mathrm{in}}\rangle)$, where $|\tilde{E}_{\mathrm{SLM}}\rangle$ is

the SLM pattern that has the same phase as $|E_{\mathrm{SLM}}\rangle$ but with uniform amplitude. Thus, predicted outgoing states are given by $|E_{\mathrm{ap}}^{\mathrm{out}}\rangle = r|E^{\mathrm{in}}\rangle$ for an amplitude and phase modulation, and by $|E_{\mathrm{po}}^{\mathrm{out}}\rangle = r|\tilde{E}^{\mathrm{in}}\rangle$ for a phase-only modulation.

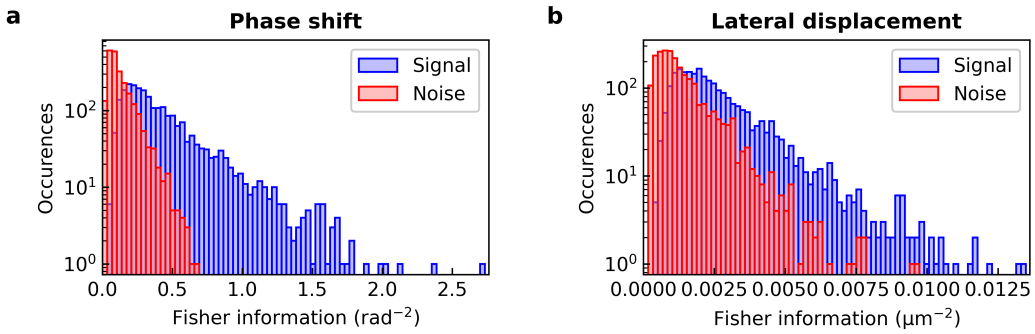
We can test this procedure by measuring the outgoing field when the medium is illuminated using optimal incident states. This characterization is experimentally performed by averaging the outgoing field over 10 measurements, with the neutral density filter ND2 placed in the signal path (measured fractional transmittance $\mathcal{T}' = 0.76 \times 10^{-2}$) to avoid saturation of the camera. The complex correlation coefficients between measured and predicted fields are $0.96 - 0.21\,i$ (for the maximum information state relative to the estimation of the phase shift of the cross-shaped object) and $0.95 + 0.06\,i$ (for the maximum information state relative to the estimation of the lateral shift of the circular object). These values show that we can correctly predict the outgoing states from the measured reflection matrix. Comparing the spatial distributions of the measured intensity to the predicted ones, we can see that these distributions are indeed very similar (Fig. S3). However, due to the presence of noise in reflection matrix measurements, we observe that the total measured intensity is lower by a factor of the order of $\eta_{\mathrm{I}} \simeq 0.80$ when compared to the predicted intensity.



Supplementary Figure S3 | Predicted and measured intensity distributions for maximum information states. **a**, **b**, Predicted spatial distribution of the normalized intensity for amplitude and phase (AP) modulation and phase-only (PO) modulation, when the observable parameter is the phase shift of the cross-shaped object. **c**, Measured spatial distribution of the normalized intensity. **d**–**f** Analogous to **a**–**c** when the observable parameter is the lateral displacement of the circular object.

S3.2 – Predicting the derivative of the outgoing field

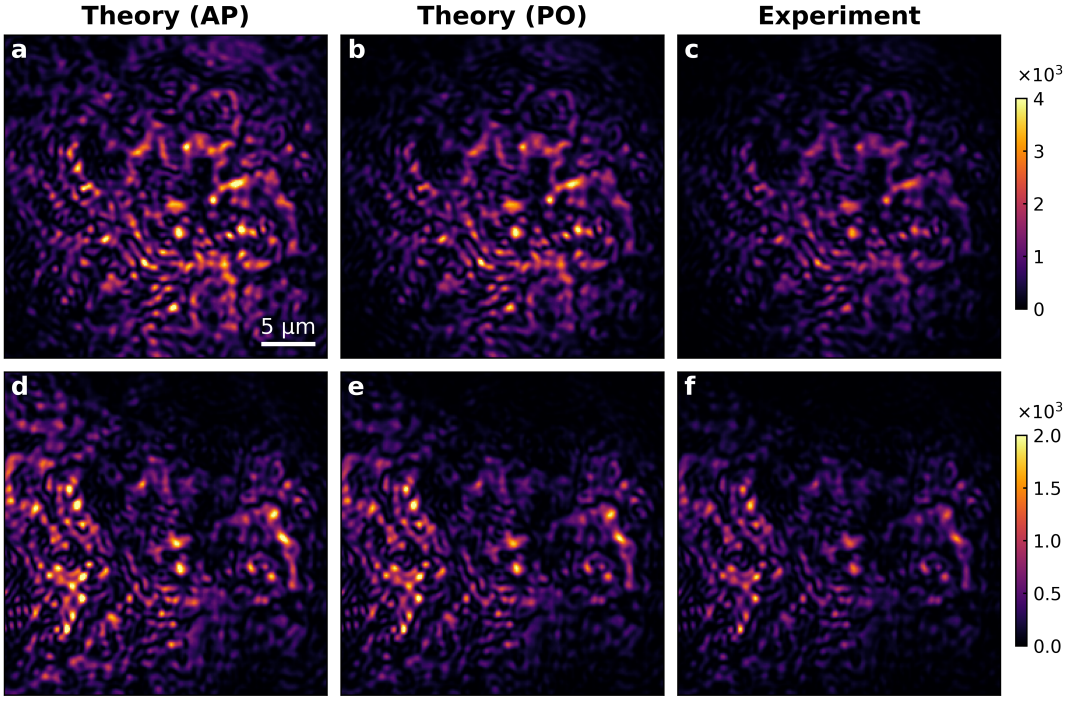
In order to faithfully assess the Fisher information in the experiments, we must ensure that $\partial_\theta r$ is correctly estimated using the finite-difference scheme $\partial_\theta r \simeq [r(\theta_0 + \Delta\theta) - r(\theta_0 - \Delta\theta)]/(2\Delta\theta)$. To this end, the change in the measured outgoing states generated by parameter variations of $\pm\Delta\theta$ must be larger than the level of noise in the measurements. In Fig. S4a, we consider the case in which the observable of interest is the phase shift induced by the cross-shape object, and we show the distribution of the estimated Fisher information for each incident plane wave used to generate the reflection matrix (blue histogram). For comparison purpose, we show the distribution of noise estimates (red histogram), obtained for each plane wave by taking two identical measurements of the outgoing state for $\varphi = \varphi_0$. We clearly observe that the signal is significantly larger than the noise, with a signal mean value of $0.41\,\mathrm{rad}^{-2}$ and a noise mean value of $0.13\,\mathrm{rad}^{-2}$. Similarly, in Fig. S4b, we consider the case in which the observable of interest is the lateral displacement of the circular object, and we show the distribution of the estimated Fisher information for each incident plane wave used to generate the reflection matrix along with the distribution of noise estimates. In this case, the signal is also larger than the noise, with a signal mean value of $0.0027\,\mu\mathrm{m}^{-2}$ and a noise mean value of $0.0014\,\mu\mathrm{m}^{-2}$.



Supplementary Figure S4 | Histograms of Fisher information and measurement noise. **a**, Histogram of Fisher information for the 2437 plane waves used to construct the reflection matrix, along with the histogram of associated measurement noise, when the observable parameter is the phase shift of the cross-shaped object. **b**, Analogous to **a** when the observable parameter is the lateral displacement of the circular object.

We can verify that this signal-to-noise ratio is sufficient to faithfully estimate $\partial_\varphi r$ and $\partial_x r$ by measuring the derivative of the outgoing field when the optimal incident state is used to illuminate the medium, and by comparing it to the derivative of the field predicted using the derivative of the reflection matrix. This characterization is experimentally performed by averaging the derivative of the outgoing field over 10 measurements, with ND2 in the signal path to avoid saturation of the camera. The complex correlation coefficients between measured and predicted derivative of the field are $0.97 - 0.22\,i$ (for the maximum information state relative to the estimation of the phase shift of the cross-shaped object) and $0.99 + 0.00\,i$ (for the maximum information state relative to the estimation of the lateral shift of the circular object). These values show that we can correctly predict the derivative of the outgoing field from reflection matrix measurements. Comparing the spatial distributions of

the measured Fisher information per unit area to the predicted ones, we can see that these distributions are very similar (Fig. S5). However, the total measured Fisher information is lower by a factor of the order of $\eta_{\mathrm{F}} \simeq 0.75$ when compared to the predicted Fisher information. This difference is largely explained by the observed difference in the predicted and measured outgoing intensity (for which a factor $\eta_{\mathrm{I}} \simeq 0.80$ was measured).



Supplementary Figure S5 | Predicted and measured distributions of the Fisher information per unit area for maximum information states. a, **b**, Predicted spatial distribution of the normalized Fisher information per unit area for amplitude and phase (AP) modulation and phase-only (PO) modulation, when the observable parameter is the phase shift of the cross-shaped object. **c**, Measured spatial distribution of the normalized Fisher information per unit area. **d**–**f** Analogous to **a**–**c** when the observable parameter is the lateral displacement of the circular object.

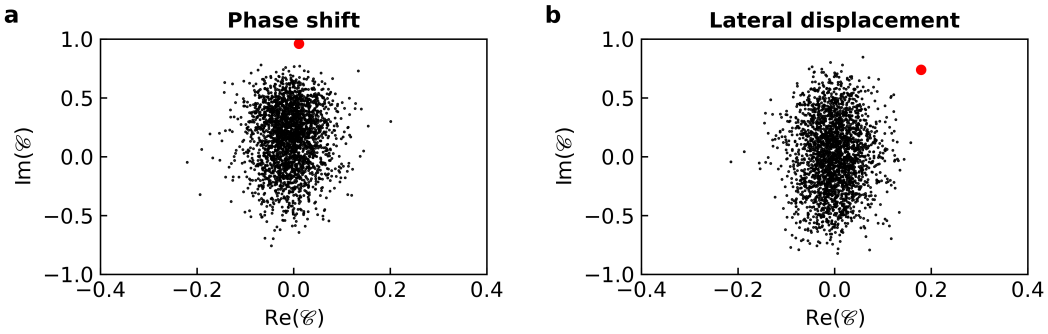
S3.3 – Correlation between the outgoing field state and its derivative

In order to characterize the correlation between $|E^{\mathrm{out}}\rangle$ and $|\partial_\theta E^{\mathrm{out}}\rangle$, we calculate the following complex correlation coefficient:

$$\mathscr{C}_\theta = \frac{\langle E^{\mathrm{out}}|\partial_\theta E^{\mathrm{out}}\rangle}{\|E^{\mathrm{out}}\| \cdot \|\partial_\theta E^{\mathrm{out}}\|} \,. \tag{S47}$$

This correlation coefficient is a relevant quantify to assess whether the Fisher information is enclosed in variations of the global phase of the outgoing state as expected for a unitary S-matrix ($\mathscr{C}_\theta \simeq \pm i$), or whether it rather enclosed in the state's intensity variations ($\mathscr{C}_\theta \simeq \pm 1$) or speckle decorrelation ($\mathscr{C}_\theta \simeq 0$).

We first calculate the correlation coefficient for each plane wave used to construct the reflection matrix, when the observable parameter is the phase shift φ induced by the cross-shaped object (Fig. S6a) and when it is the lateral position x of the circular object (Fig. S6b). On average, the measure correlation coefficients are equal to $\mathscr{C}_\varphi = -0.01 + 0.16\,i$ and $\mathscr{C}_x = 0.00 + 0.01\,i$, respectively. For the maximum information states, the measured correlation coefficients reach $\mathscr{C}_\varphi = 0.01 + 0.96\,i$ and $\mathscr{C}_x = 0.18 + 0.74\,i$, respectively. Thus, in both experiments, we observe that the average correlation coefficient for plane wave is close to zero, while the correlation coefficient for the maximum information state is close to the imaginary unit. This observation is likely to reflect the invariance property of maximum information states in the limit of a unitary S-matrix, in the same way as the principal modes of a multimode fiber are invariant (to first order) in their output profile with respect to parameter variations except for a global phase shift[14–16]. However, in our case, the measured reflection matrices are not unitary, thereby explaining the deviations from this property that are observed in the experiments.



Supplementary Figure S6 | Correlation coefficient between the outgoing field and its derivative. a, Complex correlation coefficient between the outgoing field and its derivative when the observable parameter is the phase shift induced by the cross-shaped object. Black points represent values of the complex correlation coefficient for the 2437 plane waves used to construct the reflection matrix, and the red point represent the value of the complex correlation coefficient for the optimal incident state. **b**, Analogous to **a** when the observable parameter is the lateral displacement of the circular object.

References

1. Kay, S. *Fundamentals of Statistical Processing, Volume I: Estimation Theory* (Prentice Hall, 1993).

2. Mandel, L. & Wolf, E. *Optical Coherence and Quantum Optics* (Cambridge University Press, 1995).

3. Helstrom, C. W. Quantum detection and estimation theory. *J. Stat. Phys.* **1**, 231–252 (1969).

4. Braunstein, S. L., Caves, C. M. & Milburn, G. J. Generalized Uncertainty Relations: Theory, Examples, and Lorentz Invariance. *Ann. Phys. (N. Y.)* **247**, 135–173 (1996).

5. Demkowicz-Dobrzański, R., Jarzyna, M. & Kołodyński, J. Chapter Four - Quantum Limits in Optical Interferometry. In *Progress in Optics*, vol. 60, 345–435 (Elsevier, 2015).

6. Ambichl, P. *et al.* Focusing inside Disordered Media with the Generalized Wigner-Smith Operator. *Phys. Rev. Lett.* **119**, 033903 (2017).

7. Lerosey, G., Rosny, J. d., Tourin, A. & Fink, M. Focusing Beyond the Diffraction Limit with Far-Field Time Reversal. *Science* **315**, 1120–1122 (2007).

8. Xu, X., Liu, H. & Wang, L. V. Time-reversed ultrasonically encoded optical focusing into scattering media. *Nat. Photonics* **5**, 154–157 (2011).

9. Judkewitz, B., Wang, Y. M., Horstmeyer, R., Mathy, A. & Yang, C. Speckle-scale focusing in the diffusive regime with time reversal of variance-encoded light (TROVE). *Nat. Photonics* **7**, 300–305 (2013).

10. Bosch, J., Goorden, S. A. & Mosk, A. P. Frequency width of open channels in multiple scattering media. *Opt. Express* **24**, 26472–26478 (2016).

11. Zhou, E. H., Ruan, H., Yang, C. & Judkewitz, B. Focusing on moving targets through scattering samples. *Optica* **1**, 227–232 (2014).

12. Ma, C., Xu, X., Liu, Y. & Wang, L. V. Time-reversed adapted-perturbation (TRAP) optical focusing onto dynamic objects inside scattering media. *Nat. Photonics* **8**, 931–936 (2014).

13. Ruan, H. *et al.* Focusing light inside scattering media with magnetic-particle-guided wavefront shaping. *Optica* **4**, 1337–1343 (2017).

14. Fan, S. & Kahn, J. M. Principal modes in multimode waveguides. *Opt. Lett.* **30**, 135–137 (2005).

15. Carpenter, J., Eggleton, B. J. & Schröder, J. Observation of Eisenbud–Wigner–Smith states as principal modes in multimode fibre. *Nat. Photonics* **9**, 751–757 (2015).

16. Xiong, W. *et al.* Spatiotemporal Control of Light Transmission through a Multimode Fiber with Strong Mode Coupling. *Phys. Rev. Lett.* **117**, 053901 (2016).