Fundamental bounds on the precision of classical phase microscopes: supplementary material

Dorian Bouchet,[1,*] Jonathan Dong,[2,*] Dante Maestre,[3,4] and Thomas Juffmann[3,4,†]

[1] *Debye Institute for Nanomaterials Science and Center for Extreme Matter and Emergent Phenomena, Utrecht University, P.O. Box 80000, 3508 TA Utrecht, the Netherlands*
[2] *Laboratoire Kastler Brossel, Ecole Normale Supérieure, Université PSL, CNRS, Sorbonne Université, Collège de France, F-75005 Paris, France*
[3] *University of Vienna, Faculty of Physics, VCQ, A-1090 Vienna, Austria*
[4] *University of Vienna, Max Perutz Laboratories, Department of Structural and Computational Biology, A-1030 Vienna, Austria*

(Dated: January 28, 2021)

This document provides supplementary information to "Fundamental bounds on the precision of classical phase microscopes".

S1. DERIVATION OF THE OPTIMAL CRAMÉR-RAO BOUND FOR ABSORPTION ESTIMATIONS

Instead of estimating the set of parameters $\phi = (\phi_1, \ldots, \phi_p)$, one may want to estimate the set of parameters $n = (n_1, \ldots, n_p)$ in order to characterize an absorptive sample. In the same way that we obtained the Fisher information matrix $\mathcal{J}(\phi)$ for phase estimations, we now consider the Fisher information matrix $\mathcal{J}(n)$ for absorption measurements. In the shot-noise limit and assuming that the values measured by all camera pixels are statistically independent, we obtain

$$[\mathcal{J}(n)]_{ij} = \Delta t \sum_{k=1}^{d} \frac{1}{I_k^{\mathrm{det}}} \left(\frac{\partial I_k^{\mathrm{det}}}{\partial n_i}\right) \left(\frac{\partial I_k^{\mathrm{det}}}{\partial n_j}\right) . \tag{S1}$$

In order to express the derivative of the intensity I_k^{det} with respect to ϕ_j, we can use the following chain rule:

$$\frac{\partial I_k^{\mathrm{det}}}{\partial n_j} = \left(\frac{\partial I_k^{\mathrm{det}}}{\partial \mathcal{A}_j}\right) \left(\frac{\partial \mathcal{A}_j}{\partial n_j}\right) . \tag{S2}$$

This yields

$$\begin{aligned} \frac{\partial I_k^{\mathrm{det}}}{\partial n_j} &= \frac{1}{2\sqrt{\Delta t n_j}} \left[(E_k^{\mathrm{det}})^* h_{kj} e^{i\phi_j} + E_k^{\mathrm{det}} h_{kj}^* e^{-i\phi_j}\right] \\ &= \frac{1}{\sqrt{\Delta t n_j}} \mathrm{Re}\left[(E_k^{\mathrm{det}})^* h_{kj} e^{i\phi_j}\right] . \end{aligned} \tag{S3}$$

The elements of the Fisher information matrix can thus be explicitly expressed as follows:

$$\begin{aligned} [\mathcal{J}(n)]_{ij} =& \frac{1}{n_j} \sum_{k=1}^{d} \frac{1}{|E_k^{\mathrm{det}}|^2} \\ & \times \mathrm{Re}\left[(E_k^{\mathrm{det}})^* h_{ki} e^{i\phi_i}\right] \mathrm{Re}\left[(E_k^{\mathrm{det}})^* h_{kj} e^{i\phi_j}\right] . \end{aligned} \tag{S4}$$

Writing $E_k^{\mathrm{det}} = |E_k^{\mathrm{det}}| e^{i\alpha_k}$, the diagonal elements of the Fisher information matrix are expressed by:

$$[\mathcal{J}(n)]_{jj} = \frac{1}{n_j} \sum_{k=1}^{d} |h_{kj}|^2 \cos^2(\phi_j + \beta_{kj} - \alpha_k) . \tag{S5}$$

Considering that $\cos^2(\phi_j + \beta_{kj} - \alpha_k) \leq 1$ and $\sum_k |h_{kj}|^2 \leq 1$, we have the following inequality:

$$[\mathcal{J}(\phi)]_{jj} \leq \frac{1}{n_j} , \tag{S6}$$

which is saturated if H is unitary and if the following phase-matching condition is satisfied:

$$\phi_j + \beta_{kj} - \alpha_k = m\pi , \tag{S7}$$

where m can be any integer. This must hold for all k, j for which $n_j h_{kj} \neq 0$. By comparing Eq. (S7) to the corresponding one for phase estimation in the main text, it is clear that the condition required to perform optimal phase estimations is complementary to the condition required to perform optimal absorption estimations.

To summarize, we can write the following chain of inequalities for absorption estimations:

$$\mathrm{Var}(\hat{n}_j) \geq [\mathcal{J}^{-1}(n)]_{jj} \geq [\mathcal{J}(n)]_{jj}^{-1} \geq n_j , \tag{S8}$$

where the optimal bound simply corresponds to the variance of shot-noise limited measurements in a direct imaging configuration of an absorptive sample with known phase contrast.

Note that we have treated phase and absorption estimations separately. In general, the number of unknowns cannot exceed the number of measurements for the Fisher information matrix to be well-conditioned. Precisely determining both phase and absorption entails estimating $2p$ parameters, and the Fisher information matrix can be well-conditioned only if the number d of measurements satisfies $d \geq 2p$. This can be achieved either in a single shot (i.e. in off-axis interferometry) or with multiple acquisitions (i.e. in phase shifting interferometry), as will be discussed in the following section.

* These authors contributed equally.
† Corresponding author: thomas.juffmann@univie.ac.at

S2. CRAMÉR-RAO BOUND FOR PHASE-SHIFTING AND OFF-AXIS INTERFEROMETRY

We now study the case of a simple phase-shifting interferometric scheme in which the object field is successively interfered with $N \geq 3$ external reference plane waves that are phase-shifted by $2u\pi/N$, where $u = 0, \ldots, N-1$. We assume an ideal detection scheme so that H is the identity matrix of size p, resulting in a diagonal Fisher information matrix. We also suppose that the intensity of the reference wave is high everywhere in the field of view, so that $\alpha_k \simeq \arg(E_k^{\mathrm{ref}})$. Then, considering that the N successive measurements are statistically independent, we can add the Fisher information associated with each measurement. Taking an integration time of $\Delta t/N$ for each measurement to keep the total photon number incident on the sample constant, we can use Eq. (8) of the main text to calculate the diagonal elements of the Fisher information matrix, which reads

$$[\mathcal{J}(\phi)]_{jj} = \frac{4n_j}{N} \sum_{u=0}^{N-1} \sin^2(\phi_j + 2u\pi/N) \; . \tag{S9}$$

Making use of trigonometric identities, we readily obtain

$$[\mathcal{J}(\phi)]_{jj} = 2n_j - \frac{n_j}{N} \sum_{u=0}^{N-1} \left[e^{2i(\phi_j + 2u\pi/N)} + \text{c.c.} \right] \; , \tag{S10}$$

where c.c. denotes the complex conjugate of the preceding term. Computing the geometric series, this yields

$$[\mathcal{J}(\phi)]_{jj} = 2n_j - \frac{n_j}{N} \left[e^{2i\phi_j} \left(\frac{1 - e^{i4\pi}}{1 - e^{i4\pi/N}} \right) + \text{c.c.} \right] \; . \tag{S11}$$

Since $N \geq 3$, we finally obtain

$$[\mathcal{J}(\phi)]_{jj} = 2n_j \; . \tag{S12}$$

The resulting Cramér-Rao bound is $1/(2n_j)$, which is two times worse than the optimal Cramér-Rao bound for phase estimations. Note, however, that the optimal scheme for phase estimations yields no absorption information. In contrast, the diagonal elements of the Fisher information matrix relative to absorption estimations is non-zero for the phase-shifting interferometric scheme. Starting from Eq. (S5), it can be easily shown that

$$[\mathcal{J}(n)]_{jj} = \frac{1}{2n_j} \; . \tag{S13}$$

The resulting Cramér-Rao bound is then $2n_j$, which is two times larger than the optimal Cramér-Rao bound for absorption estimations.

Similar results can also be obtained for an off-axis interferometric imaging scheme. In such scheme, the object field is interfered with a tilted plane wave, such that the period of the resulting interference pattern is smaller than the smallest features encoded in the object wave. Due to this oversampling, each phase value ϕ_j can be estimated from values measured by q camera pixels associated with different values of α_k ranging from 0 to 2π. This yields an averaging effect similar to what was exposed earlier in the case of the phase-shifting interferometric scheme, resulting in a Cramér-Rao bound of $1/(2n_j)$ for the estimation of ϕ_j (phase estimations) and of $2n_j$ for the estimation of n_j (absorption estimations).

S3. OPTIMAL PHASE MICROSCOPES WITH INCOMPLETE KNOWLEDGE OF PHASE OBJECTS

Optimal detection schemes (with or without external reference) require prior knowledge of the entire sample in order to effectively turn a strongly-contrasted object into a weakly-contrasted object. If the prior knowledge is incomplete, the phase-matching condition will then only be approximately satisfied. To determine the influence of such incomplete knowledge, we consider an optimal detection scheme with a uniform intensity distribution ($n_j = 1$ in the whole field of fiew), with a high-intensity reference beam ($|E_k^{\mathrm{ref}}|^2/|E_k^{\mathrm{obj}}|^2 = 100$ in the entire field of view) and with a phase such that $\arg(E_k^{\mathrm{ref}}) = \pi/2 + \phi_k + W_k$, where W_k follows a centered Gaussian distribution of variance σ_{g}^2. We then vary the standard deviation of the noise from 0 to 1 rad and, for each value of σ_{g}, we plot the histogram of the CRB obtained within the field of view. As can be seen in Fig. S1, close to optimal estimations can be performed for $\sigma_{\mathrm{g}} < \pi/8$, with only a few occurrences characterized by a sub-optimal CRB. Note that it is this regime in which PCM works efficiently; spatial light interference

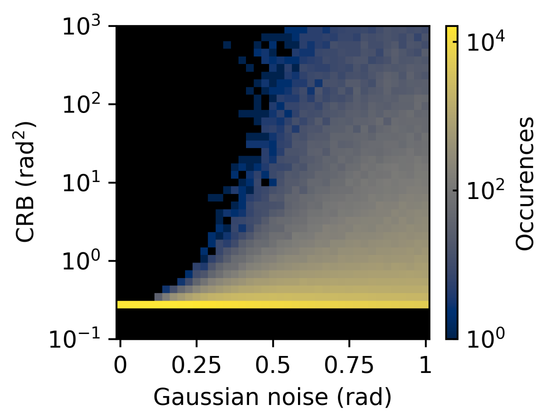

FIG. S1. Stack of histograms of CRB as a function of the standard deviation of a Gaussian uncertainty regarding the true phase distribution in the object plane. These calculations were performed for a 128×128 phase object and assuming an ideal detection setup with a high-intensity optimally-shaped reference beam. No occurrences were observed in the black areas of the figure.

microscopy (SLIM) is then a potential option as an initialization before using an optimal scheme such as Low-phi to perform dynamic precise phase estimations.

S4. INFORMATION LOSS AND SINGULAR FISHER INFORMATION MATRICES

We now discuss important effects that arise when we consider that the finite numerical aperture (NA) of the imaging system blocks all high spatial frequencies in the image plane, and that the finite size of the phase mask prevents phase estimations for low spatial frequencies (only absorption estimations could be performed for such frequencies). Taking into account these two effects, the Fisher information matrix of phase microscopes is necessarily singular, and the associated Cramér-Rao bound therefore cannot be calculated without further analysis. In order to understand why a direct inversion of the Fisher information matrix is not possible, it is relevant to use the Fourier basis for representing the parameters ϕ. Thus, we consider the estimation of a new set of parameters $\xi = W\phi$, where W is the (unitary) discrete Fourier transform (DFT) matrix which is used to approximate the Fourier transform operator. The Fisher information matrix for the parameters ξ, calculated as $\mathcal{J}(\xi) = W\mathcal{J}(\phi)W^\dagger$, has zeros for both low and high spatial frequencies, resulting in a singular Fisher information matrix.

As an example, we study the PCM configuration described in the manuscript but with a finite NA and a finite phase mask. To this end, we suppose that the area of a pixel in the detection plane is $\lambda_0^2/4$ where λ_0 is the wavelength of the incident light, so that the field is sampled at the Nyquist frequency. We further assume to use an Olympus microscope objective with a numerical aperture $\mathrm{NA} = 0.75$ and a $\times 40$ magnification, along with a phase disc whose radius corresponds to the width of a commonly-used Ph2 phase ring. The finite numerical aperture of the microscope objective effectively blocks all spatial frequencies $k/k_0 > 0.75$ where k is the magnitude of the transverse component of the wavevector and $k_0 = 2\pi/\lambda_0$. Furthermore, the effect of the phase ring is to shift all spatial frequencies $k/k_0 < 0.07$ by $\gamma = \pi/2$. The resulting distribution for the diagonal elements of $\mathcal{J}(\xi)$ is shown in Fig. S2a. Note that a zero in a diagonal element of $\mathcal{J}(\xi)$ implies that a whole row and a whole column of $\mathcal{J}(\xi)$ is zero. This demonstrates that the Fisher information is suppressed for high spatial frequencies (due to the finite NA of the detection system) as well as for low spatial frequencies (due to the finite size of the phase mask). In the limit of an idealized setting with infinite NA and with a phase mask assumed to be infinitely small, this has the well-known consequence that only a uniform global phase factor cannot be determined from the intensity distribution in the camera plane, as discussed in the manuscript.

However, in order to continue the analysis, we can assume that the value of ξ_j associated with the missing spatial frequencies are provided as *a priori* information, or that it is known that the object does not feature such frequencies. In this case, we can reduce the dimensions of $\mathcal{J}(\xi)$ and W by removing the lines and columns associated with these parameters. This procedure results in the construction of the truncated matrices $\tilde{\mathcal{J}}(\xi)$ and $\tilde{W}$, which satisfy $\tilde{\mathcal{J}}(\xi) = \tilde{W}\mathcal{J}(\phi)\tilde{W}^\dagger$. The resulting Fisher information matrix $\tilde{\mathcal{J}}(\xi)$ is now invertible, which allows us to write the following Cramér-Rao inequality

$$\mathrm{Var}(\hat{\phi}_j) \geq \left[\tilde{W}^\dagger \tilde{\mathcal{J}}^{-1}(\xi)\tilde{W}\right]_{jj} . \tag{S14}$$

Note that, due to the addition of *a priori* information, the Cramér-Rao bound can be smaller than the fundamental limit expressed by $1/(4n_j)$ in certain regions of the sample.

In the numerical results shown in Fig. 4b of the manuscript, the NA was assumed to be infinite, and the phase disc was assumed to be one pixel in size, so the Fisher information matrix had to be truncated by one column and one line before inversion. It is interesting to compare these results to the more realistic case of a finite size phase-disc and limited NA. Fig. S2b shows the Cramér-Rao bound obtained by considering the Fisher information matrix whose diagonal elements are shown in Fig. S2a, and by truncating it to remove all zero spatial frequencies. Remarkably, the larger phase mask and the underlying assumption of *a priori* knowledge about these frequencies lead to a reduced average Cramér-Rao bound. Moreover, we can recognize a Fourier-filtered version of the cameraman on the map in Fig. S2b, which confirms that low and high spatial frequencies are suppressed from the measurements.

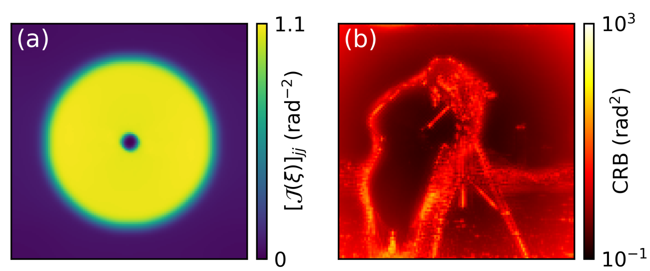

FIG. S2. (a) Diagonal elements of the Fisher information matrix $\mathcal{J}(\xi)$, expressed in the Fourier basis. (b) Resulting Cramér-Rao bound obtained by truncating $\mathcal{J}(\xi)$ for low and high spatial frequencies.