Measuring by Darkness? Let there be light!

Rainer Heintzmann^{a,b,c,*}, Jan Becker^a

^aLeibniz Institute of Photonic Technology, Albert-Einstein-Str. 9, 07745 Jena, Germany ^bInstitute of Physical Chemistry, Friedrich-Schiller-University Jena, Helmholtzweg 4, 07743 Jena, Germany ^cAbbe Center of Photonics, Friedrich-Schiller-University Jena, Albert-Einstein Str. 6, 07745 Jena, Germany

Abstract Building on work of York et al.,¹ we analyze a particular aspect of the following optics problem: Is the signal to noise ratio (SNR) of interferometric nulling with a given photon budget infinite? We compare the previously stated expression for the SNR with a different way of estimating an unknown amplitude: interfering with a very strong reference wave of known strength, i.e. optical amplification. Our analysis reveals that optical amplification is in all aspects superior to interferometric nulling. It not only yields more precise estimates, even when a scheme based on rejecting an estimate is used, but in addition optical amplification does not require a rejection of a hypothesized position estimate. We confirm our theoretical prediction with some numerical investigation; visually observing the differences between both methods, as well as stating some statistical analysis.

Keywords: optical amplification, interference, SNR, nulling.

* Correspondence to: heintzmann@gmail.com

1 Introduction

In the micro–publication,¹ York et al. describe an interesting scientific problem relating to measuring an amplitude-phase object using a Mach-Zehnder interferometer (Fig. 1). Their thought– experiment is based on the only source of noise being the stochastic nature of the photons:

$$Y_i = \mathcal{P}\{y_i\} \tag{1}$$

$$y_i = |a_i x + b_i|^2 \tag{2}$$

with Y_i being the measured number of photons in experiment number i, \mathcal{P} describes the processes of drawing from a Poisson distribution with expected number of photons y_i , x being the amplitudephase object to measure ($|x| \leq 1$), b_i being the reference amplitude in the interferometer and a_i being the illumination amplitude before the object in this experiment.



Fig 1 Mach-Zehnder type interferometer as described by York et al.¹ One arm of the interferometer contains the unknown object x + spatial light modulator (SLM) a. The reference arm only a single SLM b. Both SLM's are able to perform a complex modulation of the light distribution, hence change amplitude and phase. In the constructive channel we obtain the signal $y = |ax + b|^2$.

The Poisson process $\mathcal{P}\{y\}$ has the probability distribution:

$$P(Y|y) = \frac{y^Y}{Y!} \exp(-y)$$
(3)

The guessing "game" as described in¹ is to estimate x by a smart choice of a_i and b_i with the additional limitation of a finite illumination budget A, which is given as: $\sum_{i=1}^{N} |a_i|^2 \leq A$ with N being the number of experiments performed. York et al.¹ continue to analyze this problem and arrive at the following expression for the signal-to-noise ratio (SNR):

$$SNR = \frac{|a+b|^2}{|ax+b|} - |ax+b|$$
(4)

which has the surprising property of becoming infinite if the interferometer is adjusted for perfect

configuration: b = -ax. This is the point which our micro-publication is addressing: Is the SNR of such an interferometric measurement for estimating x really infinite ?

2 Nulling or amplification?

The first point to notice in eqn. 2 is that the result of each measurement is independent of the choice of a global phase, as the magnitude squared is calculated. Since both, a_i and b_i , are permitted freely to choose their phase, a single relative phase suffices and we can set a_i to always be real valued without loss of generality. Next we can separate the problem into a real and an imaginary part, which are added in quadrature. We therefore first analyze the real part of the problem by also assuming x and b_i to be real. To gain some insight, we now further assume that all N experiments refer to the same choice of a_i . Hence we can replace eqn. 2 with a single experiment using all the available photon budget at once:

$$y = |\sqrt{Ax} + b|^2 \tag{5}$$

The infinite SNR in a nulling experiment as stated in eqn. 4 does not consider that we may have estimated x wrongly and thus chose a wrong value for a, but nevertheless obtained zero photons in our measurement due to the *limited* overall photon budget.

In our work, we aim to compare the estimation of x by *nulling* (setting $b = -\sqrt{Ax}$) with another scheme where $b \gg 1$. This principle, termed *optical amplification*, is known for example in optical coherence tomography (OCT) as a method to eliminate read noise.² Here we do not analyze it in the context of read noise, rather as a general interferometric technique and compare it with the nulling scheme as suggested by York et al.¹

3 Comparison of both approaches

For a fair comparison it is essential to construct a thought experiment treating the schemes to compare equally. This is not trivial: the nulling scheme (I) is a hypothesis test. It is able to achieve complete experimental agreement, i.e. for a chosen b our measurement does not yield a single detected photon, hence our obtained estimate \hat{x} is consistent with being the correct value x. The optical amplification scheme (II), however, typically detects a number of photons from which a value of x is estimated with a predictable precision. These are two fundamentally different ways of how to infer information on x.

3.1 A new thought experiment

To nevertheless be able to compare both schemes on equal terms we designed the following hypothesis testing scheme: for a given (unknown) x the range $x \in [-1, 1]$ is tested in equal steps S_x searching for $x = \hat{x}$. For the nulling scheme, consistency is given by detecting zero photons, whereas for the amplification scheme we have to define a different consistency rule:

In this case, being consistent means that the detected number of photons (Y) is within a small distance Δy from the predicted number of photons (assuming $\hat{x} = x$), given the chosen value of b. To further simplify our analysis, we approximate Δy to be small enough that there is no significant change of the probability of detecting photons over this range, i.e. the probability of detecting photons in the range $Y = y \pm \Delta y/2$ is approximately given by $\Delta y P(y|y)$. Since we further assumed b to be large, we can approximate the Poisson distribution (eqn. 3) by a Gaussian probability density, with a variance equal to the expectancy:

$$P(Y|y) = \frac{\exp\left(\frac{-|Y-y|^2}{2y}\right)}{\sqrt{2\pi y}}$$
(6)

3.2 Comparing both inference schemes

Now we are able to compare the quality to estimate \hat{x} in our thought experiments. For the nulling scheme, we obtain the probability to measure zero photons according to eqn. 3 as

$$y = |\sqrt{Ax} - \sqrt{A}\hat{x}|^{2} = A|x - \hat{x}|^{2}$$

$$P(Y|y) = \frac{(A|x - \hat{x}|^{2})^{Y}}{Y!} \exp(-A|x - \hat{x}|^{2})$$

$$P(0|y) = \exp(-A|x - \hat{x}|^{2})$$
(7)

Introducing $b' = -\frac{b}{\sqrt{A}}$, this has to be compared to the result obtained in the amplification scheme. Noting that:

$$y = |\sqrt{Ax} - \sqrt{Ab'}|^2 = A|x - b'|$$
(8)

We can express eqn. 6 as:

$$P(Y|y) = \Delta y \frac{\exp(-\frac{|Y-y|^2}{2y})}{\sqrt{2\pi y}} \\ = \Delta y \frac{\exp(-\frac{|A|x-b'|^2 - A|\hat{x}-b'|^2|^2}{2A|\hat{x}-b'|^2})}{\sqrt{2\pi A}|\hat{x}-b'|}$$
(9)

With $Y = A|x - b'|^2$ the *actual* measured value and $y = A|\hat{x} - b'|^2$ our *estimated* measurement result. The exponent can be simplified as follows:

$$-\frac{A|x^2 - 2b'x - \hat{x}^2 + 2b'\hat{x}|^2}{2(\hat{x}^2 - 2b'\hat{x} + b'^2)} \approx -\frac{A|-2b'(x-\hat{x})|^2}{2b'^2} = -2A|x-\hat{x}|^2$$
(10)

Using the fact that $x \ll b'$ and $\hat{x} \ll b'$. With this we can express eqn. 9 as:

$$P(Y|y) \approx \Delta y \frac{\exp\left(-2A|x - \hat{x}|^2\right)}{\sqrt{2\pi A}b'}$$
(11)

In both cases we obtain Gaussian probability distributions. However, if we now compare the two results, we notice a difference in their width. For the nulling scheme we obtain:

$$\sigma_{\text{nulling}} = \frac{1}{\sqrt{2A}} \tag{12}$$

In case of the amplification scheme we get:

$$\sigma_{\text{amplification}} = \frac{1}{2\sqrt{A}} = \frac{\sigma_{\text{nulling}}}{\sqrt{2}} \tag{13}$$

Leading to the interesting result that the estimation based on the amplification scheme seems to be more precise than the estimation based on interferometric nulling.

3.3 Estimating from each individual measurement

However, this leaves us with the problem that the overall success rate of our amplification hypothesis test is fairly low, since Δy has to be chosen much smaller than the width of the probability distribution to render our approximation in eqn. 9 valid. Yet, if we chose a different estimation scheme, using a large reference b, but estimating \hat{x} by directly from each measurement as:

$$y = A|x - b'|^2$$

$$\rightarrow \hat{x} = b' - \sqrt{y/A}$$
(14)

for large positive b'. This yields a probability distribution equal to the hypothesis testing scheme (eqn. 9), with the only difference that in each case a solution is found leading to an estimate and there is no experiment which gets rejected by the hypothesis test. This is in all respects superior compared to the hypothesis testing scheme using interferometric nulling. Not only is an estimation possible every single time, but even for the cases where the experimental outcome is not rejected, its estimate is on average closer to the true value x.

4 Numerical results

To verify our theoretical findings, we did some numerical evaluation of both inferring schemes.

4.1 General description of our numerical evaluation

In total we performed three different experiments: for three different photon budgets (10, 40 and 100) we inferred the unknown x using both strategies. For each photon budget, we inferred 10 different values of x, which have been chosen randomly within the unit circle, by having amplitude and phase being uniformly distributed. For each x we run the respective inferring scheme 10 times to get some statistical significant. Note that we have implemented the algorithm in an iterative way $(a_i = 1)$, in contrast to our derivation in section sec. 2. In each iteration the algorithm calculates a probability map, indicating the likelihood of the true x being found at that specific position. For the nulling approach the value of b has been chosen according to $b = -\sqrt{A}\hat{x}$, with \hat{x} being the current (best) estimate. In case of the amplification scheme, b has been chosen according to |b| >> 1. The phase of b changes in steps of $\pi/2$ for each iteration, so that there is always an alternating between the real and the imaginary axis . For more detail have a look in the code we provide with this publication (see sec. 7).

4.2 Graphical comparison between both strategies

In Figure 2 we show the results of both algorithms for five different iteration steps. the parameters used were: A = 40; a = 1; $|b| = \sqrt{10}$. A red cross indicates the true position of x and the yellow ring the current estimate \hat{x} . The blue background indicates the probability map, telling us the likelihood of the true x to be found here. A higher probability is indicated by a darker color. Both algorithms start with the same probability distribution (blue); a constant as there is no information yet leaving all possible x with the same change of being the true value. From there on we describe both algorithms separately.

Nulling:

Already in the beginning of the iterations a photon is measured; hence indicating that the current estimate can not be the correct one. As the intention of the nulling scheme is to achieve perfect interferometric canceling, overall there wont be many photons being measured. Hence an update of the probability map happens rarely (also see video *Interference-Inference-nulling.avi* in sec. 7). When a photon is measured the probability map changes (Fig. 2, top), as we must include a zero probability at the position of our current estimate \hat{x} . The more photons are measured the more irregularly shaped the respective probability distribution gets. The final guess ended up close to the correct value of x. We show the residuals throughout the iteration in Fig. 3.

Amplification:

The main difference to the nulling approach is that in each iteration, we always obtain a new estimate \hat{x} . Therefore also the probability map changes in every iteration, but stays in an ellipsoidal shape. Following the maximum of the distribution reveals some *wiggly*-motion (see video

Interference-Inference-amplification.avi in sec. 7). It is possible that a new estimate actually is worse than the previous one, which comes from the randomness of the measurement process. Nevertheless, on average, we end up much closer to the true value x, when all illumination photons have been used, as indicated in Fig. 3 and Tab. 1.



Fig 2 Graphical comparison of both inferring schemes; *nulling* (top) and *amplification* (bottom) for 5 different inferring steps. Each graph shows the true value x (red), the current estimation \hat{x} (yellow) and the corresponding probability map (blue; darker = higher probability). The following parameters were used: A = 40; a = 1; $|b| = \sqrt{10}$. See also Video 1 & Video 2 for a more traceable visualization.

4.3 Statistical analysis of our numerical results

For each experiment we have calculated the residuals: $\varepsilon = x - \hat{x}$. To analyze the statistical behavior we show the mean $(\bar{\varepsilon})$ and the standard deviation of the mean $(\sigma_{\bar{\varepsilon}})$, for each iteration by taking the 10 · 10 numerical experiments into account. The result is presented in Fig. 3 a), where we show $|\varepsilon|$. In the first half of the inferring game, on average, the nulling scheme outperforms our amplification approach. However, towards higher iteration number we observe a clear enhancement inferring using b >> 1 (see small inlet). Also note that the uncertainty in latter is reduced by a factor ≈ 1.2 (also see Tab. 1). In Fig. 3 b & c) we show the standard deviation of the probability map along real and imaginary axis, for each iteration.



Fig 3 Comparing the performance of the nulling (blue) and amplification (magenta) scheme for each iteration. We show the mean residual ($\bar{\varepsilon}$; line) \pm one std. dev. deviation of the mean ($\sigma_{\bar{\varepsilon}}$), which we obtained by evaluating the inferring of 10 randomly chosen (amplitude & phase both uniformly distributed), each with 10 different tries. a) shows $|\varepsilon|$. b) & c) indicate the evolution of the probability map's variance along the real and imaginary axis. The small inlets show a zoomed region of each plot. The following parameters were used: A = 100; a = 1; $|b| = \sqrt{10}$.

Throughout the whole inferring process, the amplification algorithm results in a much smaller probability distribution, hence a more accurate estimation of x. This becomes also clear when looking at Fig. 2; overall the amplification algorithm tries to find a solution in the direct neighborhood, while in the nulling scheme an updated estimate \hat{x} might lie opposite to the previous.

Photon budget A		Nulling	Amplification	Nulling / Amplification
10	Ē	0.3048	0.2767	1.1017
	$\sigma_{\bar{\varepsilon}}$	0.0226	0.0156	1.4455
	$\bar{\sigma_r}$	0.2400	0.1916	1.2526
	$\bar{\sigma_i}$	0.2532	0.2081	1.1066
40	Ē	0.1455	0.1315	1.1066
	$\sigma_{\bar{\varepsilon}}$	0.0114	0.0066	1.7428
	$\bar{\sigma_r}$	0.1264	0.1003	1.2605
	$\bar{\sigma_i}$	0.1329	0.1088	1.2213
100	Ē	0.1051	0.0861	1.2204
	$\sigma_{\bar{\varepsilon}}$	0.0076	0.0044	1.7300
	$\bar{\sigma_r}$	0.0839	0.0659	1.2740
	$\bar{\sigma_i}$	0.0872	0.0711	1.2266

Table 1 Mean ($\bar{\varepsilon}$) and the standard deviation of the mean ($\sigma_{\bar{\varepsilon}}$) of the residuals at the last iteration. Also the standard deviation of the probability distribution along the real ($\bar{\sigma}_r$) and the imaginary ($\bar{\sigma}_i$) axis is shown. All quantities are given for three different photon budgets (A). In all scenarios the amplification strategy beats the nulling scheme, as can be seen from the ratio > 1 in the last column.

As expected, all quantities decrease when a higher photon budget is available as it is possible to obtain more information about the sample. When comparing both inferring schemes, we note that in all scenarios the amplification strategy gives better results, as predicted in our theoretical section. Note that this is not only true for a low photon budget. Even with higher illumination, the amplification strategy shows better results than the nulling approach.

5 Discussion

Our numerical results show that indeed the optical amplification algorithm gives better estimates than the previously suggested nulling scheme. We not only observe better estimations of x, also our prediction accuracy is higher. The probability distributions, generated in each iteration of both algorithms, are always more narrow in the amplification scheme (see Fig.3 b & c). However, we observe a difference in the width along the real and imaginary axis. It turns out, that our predicted factor of improvement ($\sqrt{2} \approx 1.4142$; eq. 13) is not completely reached in our numerical results (see Tab. 1). An explanation for this might, that our choice of $|b| = \sqrt{10}$ only partially resembles the assumption in our theoretical derivation (b >> 1).

Figure 3 a) suggests that there is a "break-even point" from which on optical amplification provides a smaller residual than the nulling scheme. However, this point is not fixed to the iteration nr. ≈ 40 , as it merely depends on the the cost per round (*a*). Choosing a small *a* enables to reach a higher number of total iterations. But comes at the cost of having reduced accuracy in the first iteration steps, which our optical amplification methods first has to overcome. Overall this means that there might be an ideal point for each number of available photons (*A*), for which one quickly reaches the break-even point, while having enough iterations left to further improve the current estimation. A characterisation of these points requires further investigations. Nevertheless, we have clearly shown that a *better* algorithm than the nulling scheme, suggested in,¹ can be found. It may be argued that a smarter algorithm than subsequent testing of possible candidates \hat{x} might also achieve an estimate for x in every single case. However, this modified nulling procedure will never be able to beat the hypothesis testing scheme in precision for the cases where no photons were detected. In such cases the nulling scheme is always "stuck" by assuming that its current estimate is already the correct solution, yet it is still on average further away from the true x than what can be achieved by the amplification scheme III, because: $\sigma_{nulling} > \sigma_{amplification}$.

However, there is an important aspect, which we did not analyze due to the initial assumption of a real-valued b: in the nulling scheme, the hypothesis $(x = \hat{x})$ is confirmed simultaneously for both the real and the imaginary part of the equation. Whereas the amplification scheme only yields information about the real part in both, the hypothesis-testing and the continuous-estimation approaches. A simple way to accommodate for this disadvantage would be to alternate b between large real and large imaginary values similar to the suggestion in.¹ This would half the photon budget along each direction, yielding a worsening of the estimate by $\sqrt{2}$, which would then modify eqn. 13 into $\sigma_{\text{amplification}} = \sigma_{\text{nulling}}$. Nevertheless we argue that the amplification scheme (III) is still superior, since it is not a hypothesis testing scheme but yields a new estimate in every test.

6 Conclusion

We have shown that the SNR for interferometric compensation (nulling) for the practical case of estimating x is, in contrast to the analysis of York et al.,¹ not infinite. Furthermore there is also a significantly better scheme for estimating x, which is simply choosing a very large (ideally infinite) value of b (or b'). The numerical evaluation we have done is in good agreement with our theoretical finding and might lead to an even better inference algorithm in the future.

7 Supplementary material

With this publication we also publish the following two Python scripts:

- Interference-Inference-graphics.py: The script shows the inferring process for a single experiment. Besides a visualization of the unit circle, the estimated x-values and the probability map, there is also an updating graph showing some of the statistical quantities changing in each iteration.
- 2. <u>Interference-Inference-statistics.py</u>: This script performs several experiments (for a given photon budget) and evaluates some statistical quantities, shown in Tab. 1. The only output in the end is given in the editor, no graphical visualization or other user information.

Note that both scripts import the file game.py, which is provided by Andrew York under:

https://zenodo.org/record/1463273

Additionally we also provide two videos visualizing the inferring process for both strategies which are named: *Interference-Inference-nulling.avi* and *Interference-Inference-amplification.avi*.

Acknowledgments

We would like to thanks Andrew York for posing and advertising this interesting problem and his enthusiastic talks about this topic on conferences.

References

- 1 A. G. York and S. R. Varma, "Andrewgyork/adaptive_interference_inference: Pre-print," 2018. [Online]. Available: https://zenodo.org/record/1463273
- 2 P. E. Andersen, A. Bjarklev, and A. Tycho, "Optical amplification in coherent optical frequency modulated continuous wave reflectometry," May 31 2005, uS Patent 6,900,943.