Precision Optical Measurements Using Weak Values and Nonclassical States

by

Kevin Lyons

Submitted in Partial Fulfillment of the

Requirements for the Degree

Doctor of Philosophy

Supervised by Professor Andrew N. Jordan

Department of Physics & Astronomy

Arts, Sciences and Engineering

School of Arts and Sciences

University of Rochester

Rochester, New York

2017

Table of Contents

Biographical Sketch

The author was born in Albany, New York. He attended Stony Brook University and graduated with a Bachelor of Science degree in Physics and a Bachelor of Science degree in Astronomy. He began doctoral studies at the University of Rochester in 2008, earning a Master of Arts degree in Physics in 2014. He pursued his research in Physics under the direction of Professor Andrew N. Jordan.

The following publications were a result of work conducted during doctoral study:

Justin Dressel, Kevin Lyons, Andrew N Jordan, Trent M Graham, and Paul G Kwiat. Strengthening weak-value amplification with recycled photons. *Physical Review A*, 88(2):023821, 2013

Kevin Lyons, Justin Dressel, Andrew N Jordan, John C Howell, and Paul G Kwiat. Power-recycled weak-value-based metrology. *Physical review letters*, 114(17):170801, 2015

Kevin Lyons, Shengshi Pang, Paul G. Kwiat, and Andrew N. Jordan.

Precision optical displacement measurements using biphotons. *Phys. Rev. A*, 93:043841, Apr 2016

Acknowlegments

I would like to thank my advisor Andrew Jordan for his constant help and guidance throughout my graduate studies. Beyond his aid in my growth as a scientist, I also owe him an enormous debt of gratitude for allowing me to join and work in his group during a period when I needed many special acommodations due to a long-term illness. Without his help, it is likely I would have been unable to finish work on my Master's degree, much less complete my doctoral studies.

I have been very fortunate to collaborate with a number of highly talented physicists in my research. Fellow group members Justin Dressel and Shengshi Pang both made projects I was a part of much stronger than they otherwise would have been, and both have introduced me to physical and mathematical concepts I otherwise would not have known about. I must also thank Justin Dressel in particular for spending a large amount of time and effort helping me get up to speed in my first research efforts in the group. I have also had the opportunity to work with John Howell and Paul Kwiat on several papers and would like to thank both for their ideas, insights, and contributions.

Laura Blumkin has helped me with every possible administrative item since she joined the Department, and has consistently made my life much easier. I am particularly grateful for the assitance related to navigating my graduate studies while dealing with my illness.

Finally, I cannot thank my family enough for their help during my degree work. Even with the arrival of two daughters, I have had the luxury of focusing on physics with very little interruption. This would not have been possible without their help, and is especially due to the extraordinary efforts of my wife Dana.

Abstract

We consider a number of optical precision measurement schemes, and their statistical performance. We place a particular emphasis on the performance of several enhancements to weak value measurements in optical displacement and deflection measurements. In particular, to overcome losses to "postselection" present in any weak value measurement, we examine two separate "recycling" schemes which allow each photon to have many chances of being postselected. We also show that this type of weak value measurement is capable of being used in conjunction with squeezed states of light, allowing for substantial reductions in quantum uncertainty of (say) the position of a an optical beam while also reducing the effects of technical noise. Finally, we examine a simple scheme for using entangled biphoton pairs to approximately achieve Heisenberg scaling in an optical displacement measurement.

Contributors and Funding Sources

List of Tables

List of Figures

Chapter 1

Introduction

Occasionally in physics we are interested in measuring extremely small quantities. If we are interested in pushing the limits of just how small an effect we would like to be able to measure, physics and mathematical statistics will place fundamental constraints on what can be accomplished.

To provide a motivating example which we will discuss at length over the course of this thesis, let's imagine an optical experiment measuring the tilt of a mirror simply by shining a laser source on the mirror and measuring the beam with a position sensing detector. If we first consider the case where this experiment is performed under ideal conditions with no noise and a perfect detector, there are still limits on the minimum resolvable tilt. Quantum mechanics imposes some uncertainty in the position distribution of photons coming from the source. If the detector is composed of individual pixels, some information carried by the photons will be lost compared to the case

of true continuous resolution.

As we will discuss further in this introductory chapter, the problem of finding the minimum resolvable tilt in this experiment is fully determined by the probability distribution of photons coming from the source, and a physical understanding of the parametric form the magnitude of the mirror tilt will have on the distribution. Once that is known, it is straightforward to come up with a way to estimate the mirror tilt given some experimental data, and to determine the approximate error of that estimate.

Of course, no experiment can actually be carried out under noiseless conditions with perfect detectors. On top of the quantum uncertainty discussed above, there will be a number of sources of technical noise, which refers to any type of noise due to experimental conditions. In the simple mirror tilt experiment for instance, we would be likely to experience shifts in any of the experimental components due to vibrations, shifts in the beam due to turbulence, or electronic noise in the detector, to name a few. With that in mind, it is also important to consider how robust a given measurement is to technical noise, and if there are ways to make it more robust to this noise. As we will discuss throughout this work, a technique known as weak value amplification has proven capable of mitigating many forms of technical noise [JMRH14, SDJH09]. This has been examined in detail both theoretically and experimentally since the discovery of weak values by Aharonov, Albert, and Vaidman in 1988 [AAV88].

Another method of countering noise developed over the past several decades

involves the use of nonclassical states of light. There are numerous different nonclassical states capable of reducing the quantum uncertainty in some variable of interest, however the most prominent of these are the so-called squeezed states. These allow the experimenter to decrease the variance in some variable at the expense of an increase in variance of the conjugate variable. Squeezed states have been primarily considered in precision phase measurements, and in gravitational wave detection in particular [Cav81, VCH+05]. They have been used to great effect in precision spatial measurements as well [BFM03, TAB+02, TGB+03], which we will focus on heavily in this work.

In the remainder of this introductory chapter we will briefly discuss several physical and statistical methods and techniques that will be useful throughout the rest of the work. This includes weak values, squeezed states of light, maximum likelihood estimation, and the Cramér-Rao lower bound. In chapter 2, we will discuss an enhancement to weak value amplification which recycles pulses of light to allow greatly improved measurement statistics. This method is indeed an improvement, but requires precision electronics and active optics to function. In chapter 3, this weakness is remedied by using a continuous wave recycling scheme which is entirely passive and actually allows for a larger enhancement over plain weak values than the active scheme. In chapter 4, we analyze the robustness of weak values to several types of technical noise and compare with more traditional approaches. In chapter 5 we analyze a squeezed light scheme which combines nonclassical

light and weak values. This shows that the two are not mutually exclusive, and it is in fact possible to mitigate the effects of technical noise and quantum uncertainty simultaneously. Finally, in chapter 6 we discuss an alternative optical displacement measurement using correlated biphoton pairs. We show that even if the uncertainty in the individual photons is large, it is possible to make the uncertainty in the sum of the positions of the two photons small, which in turn can be used to improve measurement sensitivity.

1.1 Weak values

In a typical quantum measurement scheme, we have some quantum system of interest which we entangle with a pointer which is measured at a detector. We take $\hat{\boldsymbol{A}}$ to be some system quantum variable, and $\hat{\boldsymbol{x}}$ to be a pointer quantum variable. For simplicity, we also take the initial state of the system and pointer to be pure, *i.e.,*

$$|\Psi_0\rangle = |\psi_0\rangle \otimes |\varphi_0\rangle. \tag{1.1}$$

Finally, the interaction Hamiltonian combining the system and meter is given by

$$\hat{\boldsymbol{H}} = \epsilon \hat{\boldsymbol{A}} \otimes \hat{\boldsymbol{x}} \delta(t - t_0), \tag{1.2}$$

where ϵ sets the strength of the interaction, and t_0 is the time where the impulsive interaction takes place. This modifies the state as

$$|\Psi\rangle = \hat{\boldsymbol{U}}|\Psi_0\rangle,$$
$$= \exp\left(-i\epsilon\hat{\boldsymbol{A}}\otimes\hat{\boldsymbol{x}}\right)|\Psi_0\rangle. \tag{1.3}$$

Here we are restricting our discussion to initial product states and Hamiltonians with trivial time dependence for clarity in this short introduction. It is straightforward to remove either restriction, and doing so does not change anything fundamental about our discussion of quantum measurements and weak values. For a more in depth review of both theory and applications, see the review article [DMM+14].

In most cases, we imagine the experimenter makes a direct measurement on the pointer state after this interaction takes place. In a weak value measurement however, an intermediate step is taken first. We make a measurement on the system state, in a process called postselection. That is, we choose a postselection state $|\psi_{ps}\rangle$, and make a projective measurement on the state $|\Psi\rangle$,

$$\frac{|\psi_{ps}\rangle\langle\psi_{ps}|}{\langle\psi_{ps}|\psi_0\rangle}|\Psi\rangle = \frac{\langle\psi_{ps}|\hat{\boldsymbol{U}}|\psi_0\rangle}{\langle\psi_{ps}|\psi_0\rangle}|\psi_{ps}\rangle\otimes|\varphi_0\rangle,$$
$$= \left(1 - i\epsilon A_w\hat{\boldsymbol{x}} + ...\right)|\psi_{ps}\rangle\otimes|\varphi_0\rangle, \tag{1.4}$$

where A_w is the so-called "weak value" given by

$$A_w = \frac{\langle \psi_{ps} | \hat{\boldsymbol{A}} | \psi_0 \rangle}{\langle \psi_{ps} | \psi_0 \rangle}. \tag{1.5}$$

We have truncated the expansion (1.4) to first order in ϵ for clarity only, the weak value effect is still well defined even if higher order terms are not small. Throughout this work we will be principally concerned with very small effects, and this first order approximation is well justified for our systems of interest. The shift in the average value of the pointer variable is then

$$\langle \Psi_0 | \hat{\boldsymbol{X}} | \Psi_0 \rangle - \langle \Psi_{ps} | \hat{\boldsymbol{X}} | \Psi_{ps} \rangle = 2i\epsilon \, \mathrm{Im} A_w \langle \varphi_0 | \hat{\boldsymbol{X}}^2 | \varphi_0 \rangle + ... \tag{1.6}$$

In the limit of very small $\langle \psi_{ps} | \psi_0 \rangle$, the weak value can be made arbitrarily large, and so this pointer shift can be made very large as well. Note however that $|\langle \psi_{ps} | \psi_0 \rangle|^2$ gives the postselection probability (i.e. the probability of a given event being postselected), so fewer probes make it to the detector for larger pointer shifts. In this work we will only be concerned with weak value measurements that have a nonzero imaginary component. As we will discuss in later chapters, this is the essential tradeoff in weak value measurements, and it actually prevents weak value measurements from exceeding the sensitivity of other types of measurements in the absence of noise.

Despite this, weak values have proven remarkably effective in a number of experiments to date. This is primarily due to weak value measurements being robust to many commonly encountered forms of techincal noise. We have

also showed theoretically that weak value measurements can be "recycled" to allow probes which fail postselection to be reused, in principle allowing all probes to be eventually recycled. This effect has been recently demonstrated experimentally.

1.2 Squeezed light

While weak values are very useful in reducing technical noise, they do not allow one to reduce quantum uncertainty while using classical (*i.e.,* coherent) light. It is possible to do this however by using nonclassical light. Here we will treat the specific case of single mode quadrature squeezed light for simplicity, leaving the case of spatial multimode squeezing for chapter 5 and the case of an alternate choice of nonclassical light in chapter 6. This section will largely follow the treatment of squeezed states by [Aga13], other useful general references include [MW95, Lou00].

We begin by defining the so-called "squeezed vacuum state"

$$|\xi\rangle = \hat{\boldsymbol{S}}(\xi)|0\rangle, \tag{1.7}$$

where $\hat{\boldsymbol{S}}(\xi)$ is the unitary "squeeze operator" defined by

$$\hat{\boldsymbol{S}}(\xi) = \exp\left[\frac{1}{2}\left(\xi\hat{\boldsymbol{a}}^{\dagger 2} - \xi^*\hat{\boldsymbol{a}}^2\right)\right]. \tag{1.8}$$

We will now show the field quadratures $\hat{X}$ and $\hat{Y}$ defined as

$$\hat{X} = \frac{1}{\sqrt{2}} \left(\hat{a} + \hat{a}^\dagger \right),$$
$$\hat{Y} = \frac{1}{\sqrt{2i}} \left(\hat{a} - \hat{a}^\dagger \right), \tag{1.9}$$

can be squeezed, in the sense that the variance of one can be reduced at the cost of raising the variance of the other.

For this task, it is convenient to work with the Bogoliubov transformed operator

$$\hat{a}(\xi) = \hat{S}^\dagger \hat{a} \hat{S},$$
$$= \hat{a} \cosh r + \hat{a}^\dagger e^{i\theta} \sinh r, \tag{1.10}$$

and the transformed creation operator given by the Hermitian conjugate. Here r is the real part of the squeeze parameter ξ and θ is $\arg \xi$. This allows us to convert problems using field variables on the squeezed vacuum state to ones using transformed variables on the vacuum. As a very simple example,

$$\left\langle \hat{a}^\dagger \hat{a} \right\rangle = \langle \xi | \hat{a}^\dagger \hat{a} | \xi \rangle,$$
$$= \langle 0 | \hat{S}^\dagger \hat{a}^\dagger \hat{S} \hat{S}^\dagger \hat{a} \hat{S} | 0 \rangle,$$
$$= \langle 0 | \hat{a}^\dagger(\xi) \hat{a}(\xi) | 0 \rangle. \tag{1.11}$$

Using this method to calculate the first two moments of the quadratures (1.9)

in the squeezed vacuum state, a somewhat lengthy calculation yields

$$\mathrm{Var}[\hat{\boldsymbol{X}}] = \frac{1}{2}e^{-2r},$$
$$\mathrm{Var}[\hat{\boldsymbol{Y}}] = \frac{1}{2}e^{2r}, \qquad\qquad (1.12)$$

for an appropriate choice of phase. Hence, the variance of one quadrature can be made arbitrarily small at the cost of making the variance of the other arbitrarily large. This effect is not limited to the quadratures of the field however, and we will see a completely analogous variance reduction to spatial measurements of an optical beam in chapter 5.

1.3   Statistical methods

1.3.1   Signal to noise ratio

The simplest statistic we will use in this work is the signal to noise ratio $\mathcal{R}$, used frequently in the literature [DSJH09, Lou00, SDJH09]. If we have some experimental data $\{X_1, X_2, ..., X_n\}$, the signal to noise is given by

$$\mathcal{R} = \frac{\langle X \rangle}{\sqrt{\mathrm{Var}[X]}},$$
$$\approx \frac{\langle X \rangle}{\sqrt{\langle X^2 \rangle}}, \qquad\qquad (1.13)$$

where $\langle X \rangle$ is assumed to be very close to zero due to the smallness of the parameter we are making inferences about. We then make the argument that

$\langle X \rangle \sim \sqrt{\text{Var}[X]}$ sets the minimum resolvable signal, i.e. $x_{min} \approx \sqrt{\text{Var}[X]}$.

Note this metric is closely related to Student's t statistic commonly used by statisticians [CB02],

$$t = \frac{\langle X \rangle - \mu}{\sqrt{S_n}/n}, \tag{1.14}$$

where $\langle X \rangle$ is the sample mean, μ is the population mean, and S_n is the sample variance. In most cases we cannot know the population mean μ, so $\mathcal{R}$ will follow a t distribution shifted by some unknown amount and scaled by the number of events. This is useful for hypothesis testing, *e.g.* if we wanted to test whether the true population mean is zero. The unknown shift limits the usefulness of signal to noise for making formal statements about the estimates of the parameter. It still finds widespread use however in the literature due to being an intuititive estimator that usually has reliable performance for reason we will discuss in the next section.

1.3.2 Parameter estimation

Typically in a precision measurement of a physical quantity, we are able to fully determine the parametric form of the sampling distribution of our data from physical first principles. This makes precision physical measurements an ideal use case for the statistical method of maximum likelihood estimation. Again having some data $\{X_1, X_2, ..., X_n\}$, the likelihood function is defined

as

$$L(\theta) = \prod_i^n p(X = x_i | \theta),\tag{1.15}$$

where θ is the parameter of interest in the single parameter case, or a vector of parameters in the multivariate case [CB02, Was04]. We will restrict our attention to the univariate case throughout this work. Note that the likelihood function is defined in terms of fixed data, and has only functional dependence on the parameter θ.

To obtain a maximum likelihood estimate (MLE), we simply take a derivative of the likelihood function with respect to θ and set it equal to zero. Solving the resulting equation provides the maximum likelihood estimate. Since it is cumbersome to solve $\partial L / \partial \theta = 0$ in most cases of interest, it is much more common to work with the log-likelihood,

$$\begin{aligned}\ell(\theta) &= \log L(\theta),\\ &= \sum_i^n \log p(X = x_i | \theta),\end{aligned}\tag{1.16}$$

which has the same argmax as (1.15) since $\log(x)$ is a monotonically increasing function of x.

The maximum likelihood estimator provides a very useful way to estimate the value of some parameter. It is also possible in this framework to get reliable estimates of the expected errors of the estimators. To begin, we will

introduce a quantity known as Fisher information [CB02, Was04],

$$\begin{aligned}
\mathcal{I}(\theta) &= -\mathbb{E}\left[\partial_\theta^2 \ln p(x|\theta)\right], \\
&= -\int \mathrm{d}x \; p(x|\theta)\partial_\theta^2 \ln p(x|\theta).
\end{aligned} \qquad (1.17)$$

For any unbiased estimator $\hat{\theta}$ (*i.e.* $\mathbb{E}[\hat{\theta}] = \theta$), the variance of the estimator obeys the Cramér-Rao lower bound,

$$\mathrm{Var}[\hat{\theta}] \geq \frac{1}{\mathcal{I}(\theta)}. \qquad (1.18)$$

If a given estimator saturates the inequality it is said to be "efficient". Most maximum likelihood estimators we will encounter are efficient, and even in cases where they are not, the Cramér-Rao inequality gives a good approximate value of the variance. This in turn allows for a straightforward way to construct confidence intervals or make signal-to-noise type arguments like $\theta_{min} \sim \mathcal{I}(\theta)^{-1/2}$.

1.3.3 A simple example

It is helpful to consider a maximum likelihood estimation example to clarify why it is useful in practice. Say we have calculated the distribution of an optical beam just before a position sensing detector will be of the form

$$p(x|\theta) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\theta)^2}{2\sigma^2}\right), \qquad (1.19)$$

where θ is an unknown displacement of the beam and σ^2 is the variance of the beam distribution.

If n photons are measured at positions x_i, the log-likelihood is given by

$$\ell(\theta) = -\sum_i^n \frac{(x_i - \theta)^2}{2\sigma^2} + \text{const.} \tag{1.20}$$

To find the maximum likelihood estimator we take the derivative of this expression with respect to θ and set it equal to zero,

$$\frac{\partial \ell}{\partial \theta} = \sum_i^n (x_i - \theta),$$

$$\tag{1.21}$$

which yields

$$\hat{\theta} = \frac{1}{n} \sum_i^n x_i. \tag{1.22}$$

That is, the maximum likelihood estimate of θ is equal to the sample average of our data. To obtain the variance in the estimator we simply calculate the Fisher information

$$\begin{aligned}
\mathcal{I}(\theta) &= -\int_{-\infty}^{\infty} \mathrm{d}x \; p(x|\theta) \partial_\theta^2 \log p(x|\theta), \\
&= \int_{-\infty}^{\infty} \mathrm{d}x \; p(x|\theta), \\
&= 1.
\end{aligned} \tag{1.23}$$

Hence, the Fisher information is equal to the inverse spatial variance (also numerically equal to the variance in this case) of the beam at the detector. This tells us the minimum resolvable parameter is approximately given by $\theta_{min} \sim 1$ and an approximate 95% confidence interval is given by $(\langle X \rangle - 2, \langle X \rangle + 2)$.

Chapter 2

Recycled postselected weak measurements

A long standing goal in optics is the development and improvement of precision optical metrology. In the first paper on weak values in 1988 [AAV88, DSS89], Aharanov, Albert and Vaidman suggested that the weak value effect might be used as an amplifier in order to measure (in the case they were considering) the value of a small magnetic field by looking at the anomalously large deflection of a beam of atoms traversing a Stern-Gerlach apparatus. The general validity of this weak value effect was later shown experimentally in an optical context by Ritchie *et al.* [RSH91], who replaced the magnetic spin with transverse polarization, and Brunner *et al.* [BAC$^+$03], who illustrated the pervasiveness of the weak value effect in common optical telecom networks.

More recently, the amplification properties of this weak value effect have been exploited in similar optical systems to precisely measure beam deflection [HK08, DSJH09, SDJH09, THSG11, PF11, HHC$^+$11, ZXLW12], phase shifts [SDW$^+$10a, SDW$^+$10b], frequency shifts [SDJH10], time delays [BS10, SB13, VMRH$^+$13], and even temperature shifts [ES12], by using either polarization or which-path degrees of freedom. Although these experiments can be described using classical wave optics [HSD$^+$10], the analysis using quantum techniques provides additional insight and allows for future extension to cases with no classical counterpart. Moreover, a wave optics approach may be less conceptually transparent due to the many spatio-temporal modes required for the analysis. Hence, we shall continue to use more flexible and compact quantum operator methods in this work as well.

Our theoretical analysis begins with the Rochester setup [DSJH09], where the tilt of a moving mirror within an interferometer is detected from the signal on a split-detector. While this setup has a sub-picoradian resolution with only milliwatts of laser power, there are a number of ways this can be improved to yield even greater sensitivity.

A generic shortcoming of weak-value-related metrological techniques is the fact that only a small fraction of the events are "post-selected", while the vast majority of events are intentionally thrown away. The main goal of the current work is to investigate how this situation can be further improved if those events are recycled. This will be done by taking photons which are not post-selected and reinjecting them back into the interferometer, so that

eventually, every photon can be post-selected in principle. We will see that this strategy does indeed lead to an improvement in the signal-to-noise ratio of the desired parameter, effectively given by the power increase of the split-detection signal. Moreover, since the existing single-pass weak-value amplification already achieves the sensitivity of standard measurement techniques (such as homodyne detection) but with lower technical noise [SDJH09], the improvements from recycling should *exceed* the sensitivity of the standard techniques. We note that because we employ not just a single pass, but many passes of a given photon through the interferometer, the simple weak value formula used in the first paper on the subject [DSJH09] will no longer suffice, and we must develop a theoretical formalism for multiple passes that will account for the amplification of the deflection, as well as the probability of reaching the detector after some number of traversals.

While the recycling scheme is an important advance in its own right (and can be generically applied to all weak value amplification schemes), it also lends itself to further enhancement if combined with other precision metrology techniques currently in use. For example, the inclusion of a spatial filter or parity-flipping element to Zeno-stabilize the beam, or the use of a squeezed reference beam [Cav81, BFM03, TAB+02, TGB+03] could significantly reduce degradation effects and quantum noise, respectively. The recycling technique, therefore, sets the stage for combined weak value/quantum light amplification strategies for future research. Furthermore, though our present work focuses on a novel pulsed recycling method, possible extensions to con-

tinuous wave operation may allow for the use of power recycling [Dre83] and signal recycling [MS91] techniques, both of which are in use in modern gravitational wave detectors [SMH⁺97, VCH⁺05].

This chapter is organized as follows. In Section 2.1 we give a heuristic estimation of the expected gains from a recycling setup based on qualitative power considerations and accessible laboratory conditions. In Section 2.2 we analytically compute the exact recycling solution for a particular setup under the assumption that the beam can stay perfectly collimated. In Section 2.3 we relax this assumption with a numerical treatment including diffraction effects. We summarize our conclusions in Section 6.4.

2.1 Qualitative Arguments

Our baseline for comparison will be the Rochester continuous wave (CW) Sagnac interferometric scheme described in [DSJH09, SDJH09, HSD⁺10]. We wish to improve the detected signal-to-noise ratio (SNR) by using a combination of pulsed laser operation with the same average power output and a design that sends the undetected portion of each pulse back into the interferometer. Such a setup is illustrated in Figure 2.1 for reference. However, before committing to a particular recycling design we can make fairly general estimations about the increases in sensitivity that we expect from any similar recycling scheme.
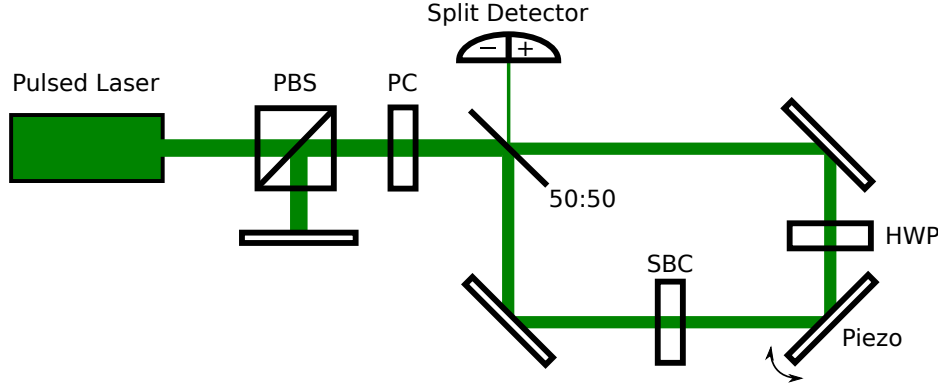
Figure 2.1: Simple cyclic weak measurement scheme. A laser emits a pulse of horizontal (H) polarization through a polarizing beam splitter (PBS), which travels through an active Pockels cell (PC) that rotates the polarization to vertical (V), after which it enters a Sagnac interferometer through a 50:50 beam splitter. Inside the interferometer the combination of half-wave plate (HWP) and Soleil-Babinet compensator (SBC) rotates the pulse polarization back to H with a relative tunable phase shift of ϕ between the clockwise ($\circlearrowright$) and counter-clockwise ($\circlearrowleft$) traveling paths. The piezo-driven mirror imparts a transverse momentum kick k that differs by a sign for the $\circlearrowright$ and $\circlearrowleft$ paths. A split detector is placed at the dark port to measure a resulting pulse deflection. The H-polarized part of the pulse that exits the bright port is rotated again by the active PC back to V before being confined by the PBS and mirror to return the pulse to the interferometer through the PC, now *inactive*.

2.1.1 Characteristic Time Scales

The constraints on how much we can increase the power collected by the dark port detector in order to improve the measurement sensitivity depend crucially on the relative time scales involved, which include:

1. The pulse duration τ being emitted by the laser. For typical lasers this can vary between 1 ns and 5 fs reasonably, which correspond to pulse

Symbol	Description	Estimate
τ	Laser pulse duration	5 fs - 1 ns
T	Laser repetition period	1 ns - 1 s
T_r	Traversal period	1 ns - 10 ns
T_g	Gating time	2 ns
T_p	Minimum inter-pulse spacing	$\tau + T_g$

Table 2.1: Relevant time scales for a recycling experiment.

lengths of 0.3 m and 1.5 μm, respectively.

2. The repetition period $T = 1/f$ of the laser. For typical lasers the repetition rate f can vary from 1 Hz to several GHz reasonably.

3. The traversal period T_r of the interferometer setup. This is determined by the physical size of the setup. As an upper-bound estimate, a 3 m long recycling setup will have a total period of 10 ns.

4. The gating time T_g for adding new pulses to the interferometer. This will determine the minimum inter-pulse spacing $T_p = \tau + T_g$ inside the interferometer. This also must be strictly less than the time between pulses emitted by the laser $T_g < T - \tau$ so that every new laser pulse can be injected. For a Pockels cell, which is a device capable of changing the polarization of light depending on an electric current applied to it, $T_g \approx 2$ ns.

These are summarized in Table 2.1.

We assume in what follows that $\tau < T_p < T_r$, so that at least one pulse can be trapped inside the interferometer. We also assume that the average power

output P of the pulsed laser is equal to the average power of a comparison CW beam. As a result, the average power of each individual pulse will be increased by a factor T/τ from the reference CW beam.

2.1.2 Detector SNR

The SNR is a useful indicator for the sensitivity of the measurement, since a signal producing an SNR of unity indicates the smallest practically resolvable signal [BFM03]. The detected SNR is defined as the ratio of the collected signal to the square root of the variance of that collected signal. Our raw signal $\langle S \rangle$ is a split-detection of the transverse profile of the beam, which measures the difference in photon number collected by each side of the detector, thereby providing information about the horizontal displacement of the beam. For small displacements, the variance of the split-detected signal is well approximated by the second moment, which is in turn proportional to the total photon number for position-uncorrelated photons (see, for example, Ref [BFM03]).

The total accumulated split-detected signal scales linearly with the average collected energy, which can be factored into the average power at the detector P_d multiplying the collection duration t. The variance will be similarly proportional to $P_d t$, so the SNR will scale as $P_d t / \sqrt{P_d t} = \sqrt{P_d t}$. Hence, the SNR can be increased either by waiting for a longer duration t or by increasing the average power P_d at the dark port detector. Our recycling scheme proposes to increase the average power collected within a fixed du-

ration to increase the sensitivity.

In the original interferometric weak value scheme, the detector collected a power of $P_d = \eta P$, where η was the post-selected fraction of the total laser power P coming from the dark port. If we recycle the unused light, however, the average power P_d collected at the detector after r_T recycling passes in a laser repetition period T will have the modified form,

$$P_d = \sum_{n=1}^{r_T} (1 - \eta)^{n-1} \eta P = (1 - (1 - \eta)^{r_T}) P, \qquad (2.1)$$

where η is the fraction of the input power that exits the dark port of the interferometer after each traversal, and r_T is the number of recycled pulses that hit the detector. Here we have ignored optical losses and detector inefficiencies for clarity. The power collected at the detector after a single traversal is ηP and the SNR scales as $\sqrt{P_d}$, so the net SNR gain factor will be,

$$\sqrt{\frac{P_d}{\eta P}} = \sqrt{\frac{1 - (1 - \eta)^{r_T}}{\eta}}. \qquad (2.2)$$

For a small post-selection probability—such as those used in weak value experiments—then we can expand (2.2) around $\eta = 0$ to find,

$$\sqrt{\frac{P_d}{\eta P}} \approx \sqrt{r_T} \left(1 - (r_T - 1)\frac{\eta}{4}\right) + O(\eta^2). \qquad (2.3)$$

For $\eta(r_T - 1) \ll 1$, then we can neglect the attenuation of the pulse to see an approximate $\sqrt{r_T}$ SNR scaling.

For sufficiently large r_T, however, the factor (2.2) saturates to the constant value $\sqrt{1/\eta}$. This saturation stems from the progressive attenuation of the recycled pulse. Furthermore, the smaller the post-selection probability gets, the larger we can make the possible SNR gain over a single pass. In this limit, however, $P_d \to P$ according to (2.1) and all the photons will be collected. Note that despite the large gain in power (2.2) at the detector, the best SNR that one can obtain still scales according to the standard quantum limit.

The measured signal at the detector may be additionally modified by geometric and propagation effects, which we can encapsulate by an overall factor $\xi(r_T)$ that depends on r_T. The total SNR gain factor over an unrecycled pulse will then be $\xi(r_T)\sqrt{P_d/\eta P}$. For the sake of comparison, we will initially ignore these effects on the signal, so we will approximate $\xi(r_T) \approx 1$ in our qualitative arguments. We will see in Section 2.2 that for a collimated beam $\xi(r_T)$ will approximate unity for small r_T but will eventually converge to zero for large r_T. Corrections to this effect will be discussed in Sections 2.2.2 and 2.2.2, where we will see that one can maintain a measurable signal for a collimated beam by inverting photons about the optical axis on each traversal or Zeno-stabilizing the beam with a spatial filter. We shall also see in Section 2.3 that $\xi(r_T)$ can exceed unity for a carefully chosen pulse divergence, which can compensate for the attenuation effects and recover the approximate $\sqrt{r_T}$ scaling for a much wider range of r_T.

2.1.3 Recycled Pulse Number

We can compute the number of pulses r_T that hit the detector per laser repetition period T from two factors. First, each trapped pulse can traverse the interferometer roughly $r \leq T/T_r$ times each repetition period. Each traversal contributes one additional pulse impact to the detector. Second, one can accumulate a maximum of $p \leq T_r/T_p = T_r/(\tau + T_g)$ pulses that are trapped inside the interferometer. Hence the total number of detector impacts $r_T = pr$ per period T will be bounded by $T/(\tau + T_g)$. Correspondingly, the maximum SNR gain factor (2.2) that we can expect from power considerations will also be bounded entirely by the relative time scales and the post-selection probability.

In practice, not every recycled pulse will contribute constructively to the SNR. Indeed, as shown in (2.2) and as we shall see in Section 2.2, there will be some maximum number $r_{\max}$ of constructive pulse impacts before the SNR saturates or decays. To maximize the SNR gain in such a case, the pulse should be discarded and replaced by a fresh pulse. Hence, the number of practical detector collections r_T will be less than the maximum estimation $r_T = pr_{\max} \leq T/(\tau + T_g)$, so the number of pulses p that can fit inside the interferometer will become important. Both pulse stabilization techniques and diverging lenses can increase the practical range of $r_{\max}$, which we will discuss in Sections 2.2.2, 2.2.2 and 2.3.

2.1.4 Practical Estimates

Using the Pockels cell as a gate, we expect $T_g \approx 2$ ns. Assuming a short pulse $\tau \ll T_g$, then the inter-pulse spacing will be $T_p c \approx T_g c = 0.6$ m. It follows that the maximum number of pulses inside the interferometer will be $p \approx T_r/T_g$. Assuming a large 3m setup, $T_r \approx 10$ ns, so $p \approx 10/2 = 5$ will be a generous upper bound to the number of pulses that we can expect to fit inside any interferometer. For contrast, the smallest setup that fits only a single pulse will be $T_r = T_p$, or $T_p c \approx 0.6$ m in length.

As shown in Section 2.2.2, without loss or stabilization we can expect $r_{\max} \leq 80$ to be an optimistic upper bound for a constructive number of recycling passes. The maximum number of pulse impacts $r_T = p r_{\max}$ per period T that we expect with the largest setup of $p = 5$ pulses is thus $r_T \leq 400$. Therefore, we can expect an SNR gain to span the range from a maximum of $\sqrt{400} = 20$ over a single pass for very small post-selection probability η to $\sqrt{1/\eta}$ for larger η according to (2.2). Since $r_{\max} \leq T/T_r$ and $T_r \approx 10$ ns for the 3 m setup, the laser repetition period must be $T \approx 400$ ns, implying a repetition rate of $f \approx 2.5$ MHz. For contrast, the smallest setup of 0.6m can fit only $p = 1$ pulse, so $r_T \leq 80$. The SNR gain thus ranges from a maximum of $\sqrt{80} \approx 9$ over a single pass to $\sqrt{1/\eta}$. The 0.6 m setup has recycling period $T_r \approx$ 2ns, so must have a laser repetition period $T \leq 160$ ns, or rate $f \geq 6.25$ MHz. These laser specifications should be readily achievable in the laboratory.

2.2 Analytic Results

2.2.1 Sagnac Interferometer

Following the experiment described in [DSJH09, SDJH09, HSD$^+$10], we extend the schematic to pulsed laser operation and pulse recycling. As shown in Fig. 2.1, the addition of a Pockels cell (PC) and polarizing beam splitter (PBS) allows the unused portion of each pulse that exits the bright port of the Sagnac interferometer to be redirected back inside the interferometer to complete multiple traversals. The combination of half-wave plate (HWP) and Soleil-Babinet compensator (SBC) provide a tunable relative phase ϕ between the clockwise ($\circlearrowright$) and counter-clockwise ($\circlearrowleft$) propagating paths of the interferometer, but also flips the net polarization of each pulse. As a result, the PC must be active as each pulse initially enters the bright port and when each pulse exits the bright port again; however, it must be *inactive* as each pulse returns to the bright port after being confined by the PBS and mirror. By injecting new pulses exactly when older pulses exit the bright port, one can minimize the inter-pulse spacing inside the interferometer to roughly a single gating time.

We also briefly note that the HWP and SBC can be removed in favor of a vertical tilt to provide the relative phase ϕ. With this variation, the PC turns on and off only once per repetition period in order to inject a new pulse into the interferometer, as opposed to cycling for every pulse traversal. This variation does not change the inter-pulse spacing inside the interferometer,

however, so it provides no SNR benefits, though it does provide a technical advantage due to the minimized number of PC cycles per laser repetition period T.

2.2.2 Pulse Recycling

Because there is no important interaction between distinct pulses in the recycling scheme, the SNR gains are fundamentally determined by the effects of single pulse recycling. Therefore, we shall consider in some detail what happens to a single pulse profile after r passes through the interferometer under the assumption that the pulse remains collimated. We will relax the collimation assumption numerically in Section 2.3.

Pulse States

Assume the clockwise-propagating state of the pulse in the Sagnac interferometer is denoted $|\circlearrowright\rangle$ and the counter-clockwise-propagating state is denoted $|\circlearrowleft\rangle$. Then the state that enters the interferometer through the 50:50 beam splitter will have the form, $|\psi_+\rangle = \frac{1}{\sqrt{2}} (|\circlearrowright\rangle + i|\circlearrowleft\rangle)$. This will also be the post-selection state for the bright port of the interferometer. Similarly, the post-selection state for the dark port of the interferometer will have the orthogonal form, $|\psi_-\rangle = \frac{1}{\sqrt{2}} (|\circlearrowright\rangle - i|\circlearrowleft\rangle)$. We also define the which-path operator as $\hat{W} = |\circlearrowright\rangle\langle\circlearrowright| - |\circlearrowleft\rangle\langle\circlearrowleft|$.

Assume the initial transverse pulse profile is given by a state $|\varphi\rangle$. We normalize the state of the transverse pulse profile so that its squared norm
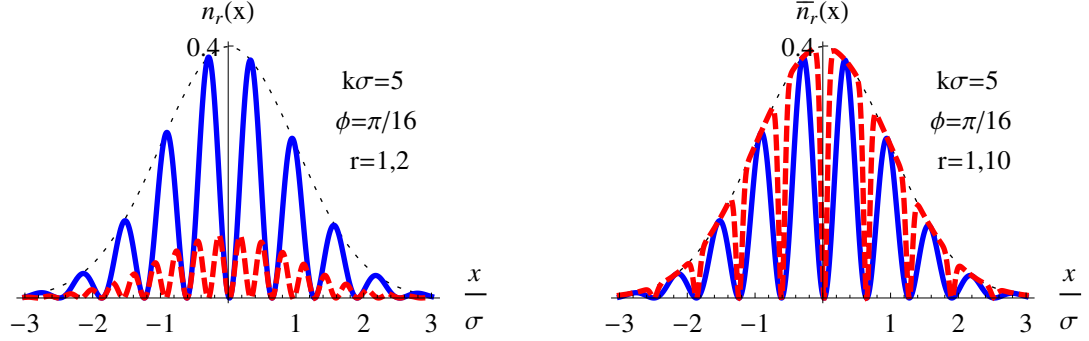
Figure 2.2: Strongly misaligned regime with $\phi < 1 < k\sigma$. Left: the transverse pulse profile $n_r(x)$ that impacts the dark port detector on the first (blue, solid) and second (red, dashed) traversals. Right: the accumulated transverse pulse profile $\bar{n}_r(x)$ on the dark port detector after a single (blue, solid) and ten (red, dashed) traversals. For this regime the interference pattern covers the entire profile, subsequent pulses are strongly attentuated, and the interference of the accumulated profile is slowly filled in.

will encode the average photodetection rate. Hence, measuring a pulse with a detector for the pulse duration τ will produce $N = \tau ||\varphi||^2$ photon impacts upon the detector per pulse on average. This choice of normalization will allow simple computation of the SNR without multi-particle Fock space calculations (e.g., as used in [BFM03]).

The total initial pulse state that enters the interferometer will have the product form, $|\Psi_0\rangle = |\psi_+\rangle|\varphi\rangle$. For simplicity we suppress the polarization of the state and any propagation effects.

The traversal through the interferometer performs three operations on the state. The first is the passage through the SBC and HWP, which creates

a relative phase shift ϕ between the paths that can be described by the unitary operator $\hat{\boldsymbol{U}}_{\text{SBC}} = e^{i\phi\hat{\boldsymbol{W}}/2}$. The second is the tilting piezo mirror, which imparts a transverse momentum kick k to the pulse, described by the unitary operator $\hat{\boldsymbol{U}}_{\text{P}} = e^{-ik\hat{\boldsymbol{W}}\hat{\boldsymbol{x}}}$, where $\hbar = 1$ and the transverse position operator $\hat{\boldsymbol{x}}$ generates a momentum translation k. The third is a generic uniform loss with probability γ, described by a nonunitary loss operator, $\hat{\boldsymbol{L}} = \sqrt{1-\gamma}\hat{\mathbf{1}}$.

The state of the pulse profile as it arrives back at the 50:50 beam splitter after one traversal has the form, $|\Psi_1\rangle = \hat{\boldsymbol{L}}\hat{\boldsymbol{U}}_{\text{P}}\hat{\boldsymbol{U}}_{\text{SBC}}|\Psi_0\rangle$. After the pulse traverses the 50:50 beam splitter, it splits into two paths once more. The dark port projects the photon onto the $|\psi_-\rangle$ state, and the bright port projects the photon onto the $|\psi_+\rangle$ state. Hence, we obtain the following two states in the bright and dark ports, respectively, $|\Psi_\pm\rangle = |\psi_\pm\rangle\left(\hat{\boldsymbol{M}}_\pm|\varphi\rangle\right)$, where we have factored out the measurement operators $\hat{\boldsymbol{M}}_\pm = \langle\psi_\pm|\hat{\boldsymbol{L}}\hat{\boldsymbol{U}}_{\text{P}}\hat{\boldsymbol{U}}_{\text{SBC}}|\psi_+\rangle$ that affect the transverse profile of the pulse in each case. Written out explicitly, these measurement operators are diagonal in the position basis and have a remarkably simple form,

$$\hat{\boldsymbol{M}}_+ = \sqrt{1-\gamma}\cos\left(\phi/2 - k\hat{\boldsymbol{x}}\right), \tag{2.4a}$$

$$\hat{\boldsymbol{M}}_- = i\sqrt{1-\gamma}\sin\left(\phi/2 - k\hat{\boldsymbol{x}}\right), \tag{2.4b}$$

where we have used $\langle\psi_\pm|\hat{\boldsymbol{W}}^n|\psi_+\rangle = (1\pm(-1)^n)/2$.

Number Densities

Using the measurement operators (2.4), the exact pulse state that exits the dark port after r traversals through the interferometer will be $|\Psi_-^r\rangle = |\psi_-\rangle \left(\hat{M}_-(\hat{M}_+)^{r-1}|\varphi\rangle \right)$. Therefore, the number density $n_r(x)$ of photons that hit the dark port detector at a transverse position x on the rth pulse traversal is,

$$n_r(x) = \tau|\langle x|\Psi_-^r\rangle|^2 = n_0(x)(1-\gamma)^r \times \tag{2.5}$$

$$\sin^2\left(\frac{\phi}{2} - kx\right)\cos^{2(r-1)}\left(\frac{\phi}{2} - kx\right),$$

where $n_0(x) = \tau|\langle x|\varphi\rangle|^2$ is the number density for the input pulse.

The total number density $\bar{n}_r(x)$ that accumulates on the dark port detector after r traversals of the pulse will be the sum of the number densities for the r traversals,

$$\bar{n}_r(x) = \sum_{j=1}^{r} n_j(x) \tag{2.6}$$

$$= n_0(x)\frac{(1-\gamma)\left(1 - [(1-\gamma)\cos^2(\phi/2 - kx)]^r\right)}{1 + \gamma\cot^2(\phi/2 - kx)}.$$

Hence the total number of photons that hit the detector after r traversals is $N_r = \int dx\, \bar{n}_r(x)$. Furthermore, if we compare (2.6) to the heuristically estimated detector power (2.1) when $\gamma \to 0$, we see that the spatially resolved version of the post-selection probability is given by $\eta \leftrightarrow \sin^2(\phi/2 - kx)$.

In the limit of an infinite number of trials $r \to \infty$, the final term in (2.6)
vanishes and we are left with the number density,

$$\bar{n}_\infty(x) = n_0(x)\frac{1-\gamma}{1+\gamma\cot^2\left(\phi/2 - kx\right)}. \tag{2.7}$$

For no loss, $\gamma \to 0$, the modulating factor from the measurement vanishes
and the original pulse is *completely recovered*, which is surprising since for
the first pass there is an anomalously large position shift. This means that if
all the photons in a perfectly collimated pulse are collected through repeated
recycling, then the information about the measurement will be erased due
to a progressive smearing—or walk-off—effect of the interference pattern, as
illustrated in Figs. 2.2, 2.3, and 2.4. Such a result indicates that a collimated
pulse should be thrown away or reshaped after a finite number of traversals
in order to maximize the information collected at the detector regarding the
momentum kick k and the induced phase shift ϕ.

Gaussian Pulse

To gain some intuition about the collected number density (2.6), consider an
initial pulse with a zero-mean Gaussian transverse profile.

$$n_0(x) = \frac{N}{\sqrt{2\pi\sigma^2}}e^{-x^2/2\sigma^2}. \tag{2.8}$$

In what follows, we will consider three specific parameter regimes for the
Gaussian pulse:
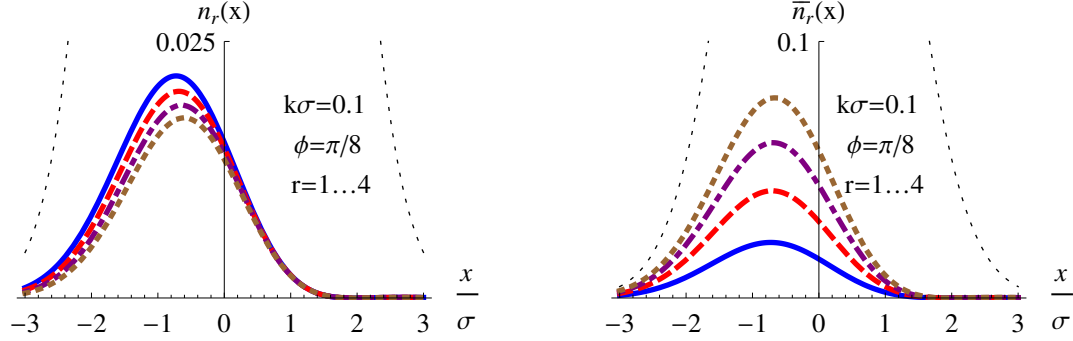
Figure 2.3: Weak-value regime with $k\sigma < \phi < 1$, with parameters chosen to exaggerate the walk-off effect. Left: the transverse single lobed pulse profile $n_r(x)$ that impacts the dark port detector on the first four traversals in order of (blue, solid), (red, dashed), (purple, dot-dashed), and (brown, dotted). Right: the accumulated transverse pulse profile $\bar{n}_r(x)$ on the dark port detector after the first four traversals, with the same color coding. For this regime, the dark port profile resembles a single shifted Gaussian that gradually walks back toward the center on multiple traversals with some attenuation, eventually recovering the original profile. For more realistic parameters, such as those used in [DSJH09], the walk-off effect is smaller per traversal.

1. the strongly misaligned regime $\phi < 1 < k\sigma$

2. the weak-value regime $k\sigma < \phi < 1$

3. the inverse weak-value regime $\phi < k\sigma < 1$

These regimes are illustrated in Figures 2.2, 2.3, and 2.4, respectively.

In the strongly misaligned regime $\phi < 1 < k\sigma$, the profile that exits the dark port on each traversal (2.5) is shown in Fig. 2.2. The interference pattern covers the entire beam profile. On the first pass, the intensity of
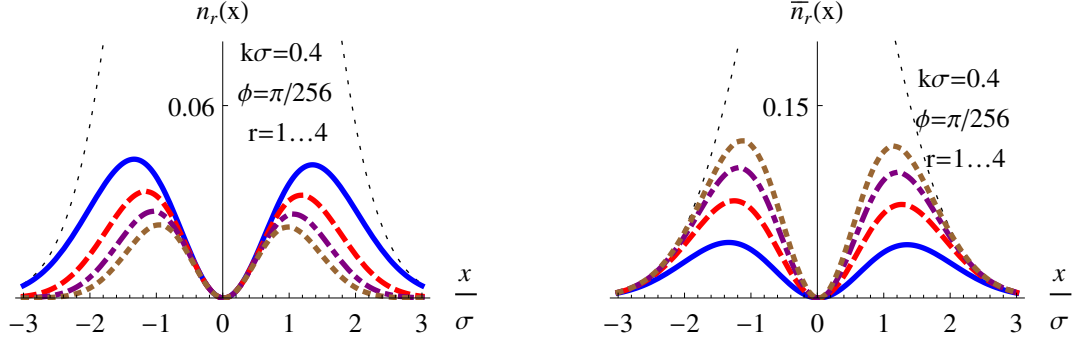
Figure 2.4: Inverse weak-value regime with $\phi < k\sigma < 1$, with parameters chosen to exaggerate the walk-off effect. Left: the transverse double-lobed pulse profile $n_r(x)$ that impacts the dark port detector on the first four traversals in order of (blue, solid), (red, dashed), (purple, dot-dashed), and (brown, dotted). Right: the accumulated transverse pulse profile $\bar{n}_r(x)$ on the dark port detector after the first four traversals, with the same color coding. For this regime, there are two lobes that very gradually walk back toward the center on multiple traversals with some attenuation to eventually recover the original beam profile. For more realistic parameters, such as those used in [SDW$^+$10b, SDW$^+$10a], the walk-off effect is smaller per traversal.

the peaks match the maximum intensity of the beam. Subsequent passes are strongly attenuated due to the small overlap with the complementary interference pattern in the beam that remains inside the interferometer. The accumulated profile $\bar{n}_r(x)$ in (2.6) steadily shrinks the width of the interference dips with increasing traversal number until the entire beam profile is recovered. The strongly misaligned regime is unlikely to be useful in a precision measurement due to the large value of k; we have included it in our discussion for completeness and to emphasize that the single and double

lobes that appear in the other regimes are not simple beam shifts, but appear from an interference effect.

For $k\sigma < \phi < 1$, we obtain the weak-value parameter regime considered in [DSJH09]. The interference pattern in the number density $n_r(x)$ indicated in (2.5) leaves a single displaced peak that resembles a shifted Gaussian that is shown in Fig. 2.3. Subsequent traversals have similar intensities, but progressively walk toward the center with increasing r. The amplified signal comes from the anomalously large shift, so this walk-off degrades the amplification properties of the setup with increasing r.

The walk-off effect arises because the beam that remains inside the interferometer has had a small fraction of light removed by the post-selection from one side, which causes a complementary displacement in the opposing direction. This complementary shift counter-acts the dark-port displacement on subsequent traversals, which makes the output walk back toward the center of the original profile. Hence, after r traversals the accumulated weak-value signal $\bar{n}_r(x)$ (2.6) will resemble r times the intensity of a single traversal, but will also be degraded due to the walk-off effect. The walk-off is shown exaggerated in Fig. 2.3, but is a smaller effect per traversal with more realistic parameters, such as those in [DSJH09]. However, even though the effect per traversal is smaller, for a sufficiently large number of traversals the signal will always be completely erased by this walk-off effect according to (2.7).

For $\phi < k\sigma < 1$, we enter the inverse weak-value regime considered in [SDW$^+$10b, SDW$^+$10a] and originally observed in [RSH91]. The weak-value

assumptions that produce the single peak break down and (2.5) produces the double-lobed profile shown in Fig. 2.4. On multiple traversals the peaks gradually walk back toward the center, similarly to the weak-value regime. However, the forced zero in the center will stabilize the profile, so that after r traversals the accumulated profile $\bar{n}_r(x)$ (2.6) will more closely resemble r times the intensity of a single traversal than in the weak-value regime. The walk-off is shown exaggerated in Fig. 2.4, but is also a smaller effect per traversal with more realistic parameters, such as those used in [SDW$^+$10b, SDW$^+$10a].

These different regimes for weak value amplification measurements are also carefully explored in the recent review paper [KAN12] and were mentioned earlier in [Di 12].

Parity Flips

A simple technique for compensating for the profile erosion on multiple traversals is to invert the profile around the $x = 0$ line so that each new traversal partially cancels the walk-off from the previous traversal. This can be accomplished by introducing a parity-flipping optic represented by a parity operator $\hat{\boldsymbol{P}}_x$ that modifies the profile by replacing $x \rightarrow -x$. This results in a net replacement of the operator $\hat{\boldsymbol{M}}_+ \rightarrow \hat{\boldsymbol{P}}_x \hat{\boldsymbol{M}}_+$ in (2.4). After an *even* number of traversals $2r$, the accumulated number density (2.6) then has the

modified form,

$$\bar{n}_{2r}(x) = n_0(x) \sin^2 \left(\frac{\phi}{2} - kx \right) \left(1 + \cos^2 \left(\frac{\phi}{2} - kx \right) \right) \times$$

$$\frac{1 - \left[\cos^2 \left(\frac{\phi}{2} - kx \right) \cos^2 \left(\frac{\phi}{2} + kx \right) \right]^r}{1 - \cos^2 \left(\frac{\phi}{2} - kx \right) \cos^2 \left(\frac{\phi}{2} + kx \right)}, \tag{2.9}$$

where we have set $\gamma \to 0$ for clarity. Unlike (2.6) where we do not flip the output of the bright port on each traversal, this expression does not yield the original input profile in the limit of large r. Instead it yields,

$$\bar{n}_\infty(x) = n_0(x) \frac{\sin^2 \left(\frac{\phi}{2} - kx \right) \left(1 + \cos^2 \left(\frac{\phi}{2} - kx \right) \right)}{1 - \cos^2 \left(\frac{\phi}{2} - kx \right) \cos^2 \left(\frac{\phi}{2} + kx \right)}, \tag{2.10}$$

which maintains a signal, in contrast with the case of no parity flips on a collimated beam.

Zeno stabilization

Another way to reduce transverse walk-off effects and thereby restore the signal-to-noise ratio to the power-limited scaling of (2.2) in Section 2.1.2 is to utilize the physics of the quantum Zeno effect by using an optical filter to project the transverse profile back into its original state. The advantage of the Zeno stabilization over parity flipping is that the former does not swap the transverse locations of the photons; this may be important when using quantum states of light, e.g., squeezed or entangled states, whose benefits rely on maintaining transverse correlations between the photons.
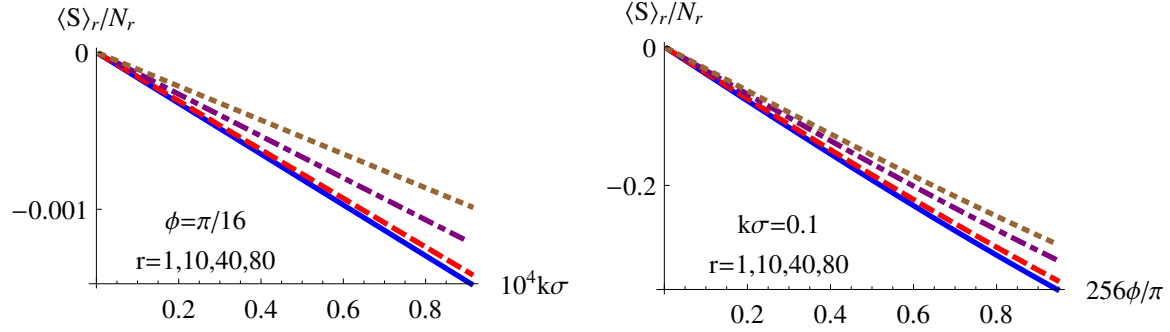
Figure 2.5: Normalized split-detection response for a collimated pulse. Left: response for the weak-value regime as a function of $k\sigma$ with fixed ϕ, where $k\sigma < \phi < 1$, and with parameters consistent with Ref. [DSJH09]. The accumulated signal is shown for traversal numbers $r = 1$ (blue, solid), $r = 5$ (red, dashed), $r = 10$ (purple, dot-dashed), and $r = 20$ (brown, dotted). Right: response for the inverse weak-value regime as a function of ϕ/π with fixed k, where $\phi < k\sigma < 1$, and with parameters consistent with Ref. [SDW$^+$10b, SDW$^+$10a]. The accumulated signal is shown for the same traversal numbers and color coding. Though the walk-off effects at large traversal numbers change the slope in both regimes, the linear response is preserved. Hence, one can calibrate the slope through repeated experiments with a fixed number of traversals per laser period. The slope is negative here since the signal is negative in Eq. (2.20).

At every round of the recycling with Zeno stabilization, the beam is passed through a spatial filter, so if the beam is in its original profile, it will pass through the filter perfectly; however, if the waveform is distorted, then a photon in that mode will have some probability to be absorbed. In passing through the filter on each traversal, a photon in this mode will only experience a small disturbance to the transverse profile, and the state will tend to "freeze" in its original state with only a small rate of being projected into
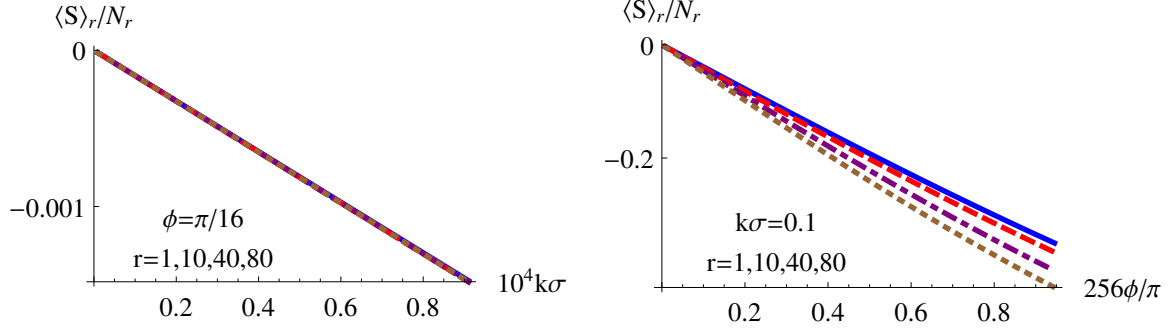
Figure 2.6: Normalized split-detection response for a collimated pulse stabilized by parity flipping, using the same color coding as in Figure 2.5. Left: response for the weak-value regime as a function of $k\sigma$ for fixed ϕ, where $k\sigma < \phi < 1$. All traversal numbers r have the same linear response due to the parity-flip stabilization; Zeno stabilization produces the same result. Right: response for the inverse weak-value regime as a function of ϕ/π for fixed k, where $\phi < k\sigma < 1$. The linear response acquires a steeper slope for larger traversal number using parity-flip stabilization; Zeno stabilization, however, would result in identical slopes for any r.

an orthogonal state (in this case, being absorbed by the filter). We note this technique will work regardless of the nature of the disturbance, provided it is small in each pass.

To see how this works, let us consider the Gaussian transverse state in (2.8) for a single photon ($N = 1$). The corresponding transverse spatial state for this photon has the form

$$\phi_0(x) = \langle x|\phi_0\rangle = \frac{1}{(2\pi\sigma^2)^{1/4}} \exp\left(-\frac{x^2}{4\sigma^2}\right). \qquad (2.11)$$

After one traversal through the interferometer, the state emerging from the bright port according to (2.4) is

$$|\phi_1\rangle = \hat{\boldsymbol{M}}_+|\phi_0\rangle = \sqrt{1-\gamma}\cos(\phi/2 - k\hat{\boldsymbol{x}})|\phi_0\rangle. \tag{2.12}$$

To compute the reshaping probability, we renormalize this state by dividing out its norm

$$\langle\phi_1|\phi_1\rangle = (1-\gamma)\int \mathrm{d}x|\phi_0(x)|^2\cos^2(\phi/2 - kx) \tag{2.13}$$
$$= (1-\gamma)(1 + e^{-2k^2\sigma^2}\cos\phi)/2,$$

to produce the normalized state $|\phi_{1,n}\rangle = |\phi_1\rangle/\sqrt{\langle\phi_1|\phi_1\rangle}$.

If we now make a projective measurement with a spatial filter of the shape $|\langle x|\phi_0\rangle|^2$, the photon will be restored to the state $|\phi_0\rangle$ with a probability $P_1 = |\langle\phi_0|\phi_{1,n}\rangle|^2$. The probability can be calculated from

$$\langle\phi_0|\phi_{1,n}\rangle = \frac{1}{\sqrt{N_1}}\int \mathrm{d}x|\phi_0(x)|^2\cos(\phi/2 - kx), \tag{2.14}$$
$$= \frac{\sqrt{2}e^{-k^2\sigma^2/2}\cos(\phi/2)}{\sqrt{1 + e^{-2k^2\sigma^2}\cos\phi}}.$$

We are interested in the case where both ϕ and $k\sigma$ are less than 1. Consequently, we can expand P_1 to leading order in $k\sigma$ and ϕ,

$$P_1 = 1 - (k\sigma)^4/2 - (k\sigma)^2\phi^2/4 + \ldots, \tag{2.15}$$

where we drop terms of higher order in powers of $(k\sigma)^2$ and ϕ^2. In the weak-value regime where $k\sigma < \phi < 1$, the second term in (2.15) may be dropped. In the inverse weak-value regime where $\phi < k\sigma < 1$, the third term in (2.15) may be dropped. In either case, for repeated cycles consisting of M independent measurements the probability $P_M = P_1^M$ of being projected back into state $\langle x|\phi_0 \rangle$ will decay approximately exponentially as

$$P_M = \exp[-M\Gamma], \tag{2.16}$$

where $\Gamma \approx (k\sigma)^4/2 + (k\sigma)^2\phi^2/4$ is an effective decay rate. We can therefore make $M \sim M_Z = 1/\Gamma$ measurements before a photon is typically absorbed by the reshaping filter. This is the manifestation of the Zeno effect, where by making repeated projections, the state is kept in its initial state for much longer than would happen otherwise.

This Zeno number M_Z is many more cycles that we will be able to make before the detector measures all the photons exiting the dark port. For example, if we chose the exaggerated values $k\sigma = 0.1$ and $\phi = \pi/8$ as in Figure 2.3 then this gives a Zeno number of $M_Z \approx 2.3 \times 10^3$, which is still an order of magnitude larger than we require for the detection physics.

2.2.3 Split-detected signal

In order to measure the transverse momentum kick k or the phase shift ϕ we compare the sides of the transverse profile using a split detector. As outlined
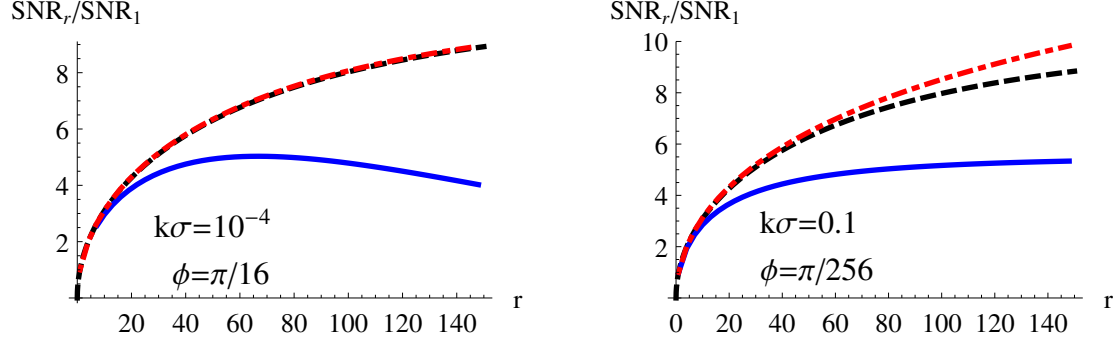
Figure 2.7: SNR gain versus traversal number r for a collimated pulse. Left: the SNR gain for the weak-value regime $k\sigma < \phi < 1$ with $k\sigma = 10^{-4}$ and $\phi = \pi/16$. The uncorrected beam with walkoff (blue, solid) shows clear degradation with traversal number, while the beam corrected with parity flipping (red, dashed) as in Sec. 2.2.2 matches the simple power scaling law exactly (black, dashed) from Eq. (2.2). Note that Zeno stabilization will identically produce this power-scaled SNR by construction. Right: the SNR gain for the inverse weak-value regime $\phi < k\sigma < 1$ with $k\sigma = 0.1$ and $\phi = \pi/256$. The uncorrected double-lobed beam with walkoff (blue, solid) shows somewhat less degradation than the weak-value regime due to the forced zero in the profile; however, the double-lobed beam corrected with parity flipping (red, dashed) manages to exceed the simple power scaling law (black, dashed), and thus the scheme with Zeno stabilization.

in Section 2.1.2, the accumulated split-detected signal after r pulse repetitions of time duration τ is given by the difference of the number densities on each side,

$$\langle S \rangle_r = \int_0^\infty \mathrm{d}x\, \bar{n}_r(x) - \int_{-\infty}^0 \mathrm{d}x\, \bar{n}_r(x). \tag{2.17}$$

To measure the displacement of the pulse, the signal should be subsequently normalized by the total photon number $\langle S \rangle_r / N_r$ in order to extract the averaged behavior.

For small displacements the variance of the raw split detected signal is the second moment to a good approximation,

$$(\Delta S)_r^2 \approx \langle S^2 \rangle_r = \int \mathrm{d}x\, \bar{n}_r(x) = N_r, \tag{2.18}$$

which is the total number of photons that have impacted at the dark port detector. Hence, the SNR that has accumulated after the rth traversal will be given by,

$$\text{SNR}_r = \frac{\langle S \rangle_r}{(\Delta S)_r} = \frac{\langle S \rangle_r}{\sqrt{N_r}}. \tag{2.19}$$

For the zero-mean Gaussian (2.8) these quantities can be computed exactly for the first traversal,

$$N_1 = \frac{(1-\gamma)N}{2}\left(1 - e^{-(2k\sigma)^2/2}\cos\phi\right), \tag{2.20a}$$

$$\langle S \rangle_1 = -\frac{(1-\gamma)N}{2}e^{-(2k\sigma)^2/2}\text{Erfi}(\sqrt{2}k\sigma)\sin\phi, \tag{2.20b}$$

$$\frac{\langle S \rangle_1}{N_1} = -\text{Erfi}(\sqrt{2}k\sigma)\frac{e^{-(2k\sigma)^2/2}\sin\phi}{1 - e^{-(2k\sigma)^2/2}\cos\phi}, \tag{2.20c}$$

$$\text{SNR}_1 = -\sqrt{(1-\gamma)N}\,\frac{\text{Erfi}(\sqrt{2}k\sigma)e^{-(2k\sigma)^2/2}\sin\phi}{\sqrt{2(1 - e^{-(2k\sigma)^2/2}\cos\phi)}}, \tag{2.20d}$$

where $\text{Erfi}(x) = \text{Erf}(ix)/i = (2/\sqrt{\pi})\int_0^x e^{t^2}\mathrm{d}t$ is the imaginary error function.

We now specialize these exact solutions to the two amplification regimes under consideration and indicate numerically how larger traversal numbers behave in each regime.

Weak-value regime

When $k\sigma < \phi < 1$ then we can neglect terms of order $(k\sigma)^2$ in (2.20) to find,

$$N_1 = (1 - \gamma)N\sin^2(\phi/2), \tag{2.21a}$$

$$\langle S \rangle_1 = -\sqrt{\frac{2}{\pi}}(1 - \gamma)Nk\sigma\sin\phi, \tag{2.21b}$$

$$\frac{\langle S \rangle_1}{N_1} = -\sqrt{\frac{2}{\pi}}2k\sigma\cot(\phi/2), \tag{2.21c}$$

$$\mathrm{SNR}_1 = -\sqrt{\frac{2}{\pi}}\sqrt{(1-\gamma)N}(2k\sigma\cos(\phi/2)) \tag{2.21d}$$

$$= -\sqrt{\frac{2}{\pi}}\sqrt{N_1}(2k\sigma\cot(\phi/2)).$$

These linear order solutions correctly match the weak value analyses made in [DSJH09, SDJH09, HSD$^+$10], as expected. Due to the factor $\cot(\phi/2)$ in the normalized split detection $\langle S \rangle_1 / N_1$, setting a known small ϕ provides an amplification factor for measuring an unknown small k. This regime gets its name from the fact that this amplification factor is the imaginary part of the weak value $W_w = \langle\psi_-|\hat{\boldsymbol{W}}|\psi_\phi\rangle/\langle\psi_-|\psi_\phi\rangle = i\cot(\phi/2)$ of the which-path operator $\hat{\boldsymbol{W}}$ with initial state $|\psi_\phi\rangle = \hat{\boldsymbol{U}}_{\mathrm{SBC}}|\psi_+\rangle$ and post-selection state $|\psi_-\rangle$. The normalized signal for this parameter regime is shown in the left plot of

Fig. 2.5 as a function of $k\sigma$, demonstrating the linear response.

We can reproduce the dominant SNR gain factor for small post-selection probability by neglecting the walk-off effects and the power attenuation. To do this, we expand the accumulate profile $\bar{n}_r(x)$ in (2.6) to first order in $k\sigma$ and second order in ϕ to obtain,

$$\bar{n}_r(x) = rc(\gamma, r)n_0(x)\left(-kx\phi + \left(\frac{\phi}{2}\right)^2\right), \tag{2.22}$$

$$c(\gamma, r) = \frac{(1 - \gamma)(1 - (1 - \gamma)^r)}{r\gamma}, \tag{2.23}$$

where $\lim_{\gamma \to 0} c(\gamma, r) = 1$.

The only r-dependence in the number density is in the numeric prefactor $rc(\gamma, r)$, which effectively scales the total photon number $N \to rc(\gamma, r)N$. Using this scaling, the result (2.21) for the split-detector will hold for any r to second order in ϕ and first order in k. Hence, the SNR should scale as $\sqrt{Nrc(\gamma, r)}$ when walk-off and power attenuation effects are neglected. When $\gamma \to 0$, this recovers the dominant $\sqrt{r}$ SNR enhancement factor that we found to zeroth order in the post-selection probability of (2.2) in Section 2.1.2 from power considerations.

However, the walk-off effects and power attenuation combine to reduce the actual SNR below this optimistic level. To see this, consider the solid blue curve in the left plot of Fig. 2.7, which shows the split-detected SNR gain versus traversal number for the weak-value regime. The SNR gain for any sufficiently small $k\sigma$ is universal, but plateaus quickly due to the beam

degradation from the walk-off. Even worse, for sufficiently large traversal number r the signal will eventually decline and then converge to zero due to the erasure effect implied by (2.7), so the SNR gain factor will also correspondingly decay to zero.

The power scaling in (2.2) can be recovered, however, if the walk-off is corrected with the parity flipping method discussed in Section 2.2.2. The normalized signal produced with the parity-flip correction—illustrated in the left plot of Fig. 2.6—has an identical slope for any traversal number, demonstrating the simple power scaling behavior. This correction is illustrated as the dot-dashed red curve in the left plot of Fig. 2.7, which exactly overlaps the power scaling curve illustrated as the dashed black curve. If the walk-off is corrected with Zeno stabilization instead, then the signal slope will be identical for any traversal number by construction, and the SNR gain will also exactly follow the power scaling curve in (2.2).

Inverse weak-value regime

If $\phi < k\sigma < 1$, then the approximation to linear order in $k\sigma$ will break down, as shown earlier in Section 2.2.2. For this regime, we keep linear order in ϕ and second order in $k\sigma$ in (2.20) to find,

$$N_1 = (1 - \gamma)N(k\sigma)^2, \tag{2.24a}$$

$$\langle S \rangle_1 = -\sqrt{\frac{2}{\pi}}(1 - \gamma)Nk\sigma\phi, \tag{2.24b}$$

$$\frac{\langle S \rangle_1}{N_1} = \sqrt{\frac{2}{\pi}} \left(\frac{k\sigma}{3} - \frac{1}{k\sigma} \right) \phi, \tag{2.24c}$$

$$\text{SNR}_1 = \sqrt{\frac{2}{\pi}} \sqrt{(1 - \gamma)N} \left(\frac{5}{6}(k\sigma)^2 - 1 \right) \phi \tag{2.24d}$$

$$= \sqrt{\frac{2}{\pi}} \sqrt{N_1} \left(\frac{5}{6}k\sigma - \frac{1}{k\sigma} \right) \phi.$$

In contrast to the previous approximation, the $1/k$ term in the normalized signal $\langle S \rangle_1 / N_1$ leads to an amplification in measuring an unknown ϕ given a known small k. Indeed, this regime was used in Refs. [SDW$^+$10b, SDW$^+$10a] for exactly this purpose. In the preprint version of Ref. [SDW$^+$10b] it was noted that $\phi \approx 2\text{Im}W_w^{-1}$ is the inverse of the weak value present in the weak-value regime for small ϕ, which motivates our name for this parameter regime; this inverted relationship has also been rediscovered more recently in Ref. [KAN12]. The normalized signal for this parameter regime is shown in the right plot of Fig. 2.5 as a function of ϕ, demonstrating the linear response.

Again, we can reproduce the dominant SNR gain factor for small post-selection probability by neglecting the walk-off effects and the power attenuation, which can be done by expanding $\bar{n}_r(x)$ to second order in $k\sigma$ and first order in ϕ to obtain,

$$\bar{n}_r(x) = rc(\gamma, r)n_0(x) \left(-kx\phi + k^2 x^2 \right), \tag{2.25}$$

with the same $c(\gamma, r)$ as in (2.23).

As with the weak-value regime, the only r-dependence in the number

density is in the numeric prefactor $rc(\gamma, r)$, which effectively scales the total photon number $N \to rc(\gamma, r)N$. Using this scaling, the result (2.24) for the split-detector will hold for any r to second order in k and first order in ϕ. Hence, the SNR will scale as $\sqrt{Nrc(\gamma, r)}$ when walk-off and power attenuation effects are neglected. When $\gamma \to 0$, this also recovers the dominant $\sqrt{r}$ SNR enhancement factor that we found to zeroth order in the post-selection probability of (2.2) in Section 2.1.2 from power considerations.

As before, the walk-off and power attenuation effects reduce the SNR gain below this optimistic level. Consider the solid blue curve in the right plot of Fig. 2.7, which shows the exact split-detected SNR gain versus traversal number for the inverse weak-value regime. As anticipated in Section 2.2.2, the forced zero in the center of the double-lobed profile naturally stabilizes the beam to produce a saturated SNR for more traversals than the weak-value regime. However, the SNR still plateaus relatively quickly before eventually decaying to zero for a sufficiently large number of traversals r without additional stabilization.

Similarly, the degradation from walk-off can be completely reversed by employing the parity flipping technique. The dot-dashed red curve in the right plot of Fig. 2.7 actually *exceeds* the simple power scaling law illustrated as the dashed black curve due to an additional accumulation of momentum information on each subsequent traversal. Moreover, the normalized signal shown in the right plot of Fig. 2.6 shows a corresponding increase in the slope of SNR vs. ϕ with traversal number.

Symbol	Numerical Value
ℓ	1.5 m
σ	1 mm
k	1×10^{-3} m^{-1}
k_0	8×10^6 m^{-1}
γ	0.01
d	1 cm

Table 2.2: Parameters used for numerical computations. ℓ is half the length of the interferometer, k is the momentum kick from the mirror, σ is the input beam width, k_0 is the carrier momentum, γ is the loss per traversal, and d is the half-width of the split-detector.

2.3 Diverging Pulse

By assuming a collimated beam, we have so far neglected beam propagation effects in the analysis, as well as any lens effects that could further change the detection physics. In order to incorporate these effects, we now alter our measurement operators and pursue a numerical approach. We find that these effects may slightly enhance the SNR gains from recycling before saturation due to the finite detector size, but do not fundamentally alter the basic power-scaling behavior. For that reason, it will be sufficient to illustrate only the weak-value regime as an example.

Extending the collimated analysis in Section 2.2 to include beam propagation leads to a replacement of the measurement operators with $\hat{M}_\pm \to \hat{M}'_\pm = \hat{U}_\ell \hat{M}_\pm \hat{U}_\ell$, where $\hat{U}_\ell = \exp(-i\hat{p}^2\ell/2k_0)$, k_0 is the carrier momentum of the pulse, and ℓ is the propagation length from piezo to 50:50 beam splitter [BW59]. The number density (2.5) will then involve the composite
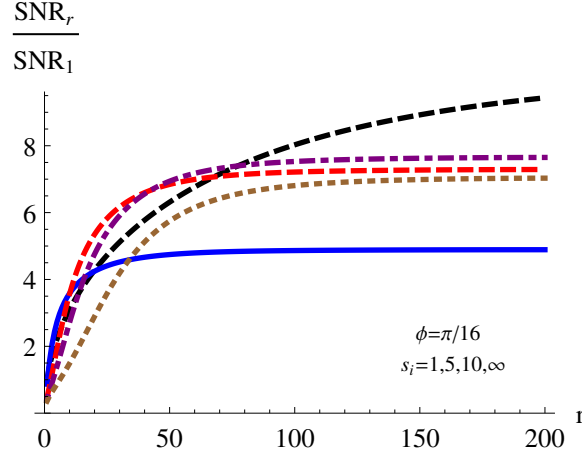
Figure 2.8: The effect of propagation and different diverging lens choices on the SNR of the weak-value regime with $k\sigma = 10^{-6}$ and $\phi = \pi/16$. In order of (blue, solid), (red, dashed), (purple, dot-dashed), and (brown, dotted) we show weak initial diverging lenses with extreme focal lengths $s_i = -1$ m, -5 m, -10 m, and no lens. The (black, dashed) curve shows the scaling given by power considerations in Eq. (2.2). Here SNR_1 refers to the SNR of a single unrecycled pulse with an optimally chosen focal length of $s_i = -0.5$ m, while SNR_r is the accumulated SNR over r traversals for the indicated lens choices.

measurement operator

$$\hat{M}'_- \left(\hat{M}'_+ \right)^{r-1} = \hat{U}_\ell \hat{M}_- \hat{U}_\ell \left(\hat{U}_\ell \hat{M}_+ \hat{U}_\ell \right)^{r-1}. \qquad (2.26)$$

Adding a diverging lens with focal length s_i also modifies the initial state with an operator $\hat{U}_L = \exp(ik_0\hat{x}^2/2s_i)$. Hence, powers of the following modified

operators will appear in the full solution,

$$\hat{\boldsymbol{U}}_\ell \hat{\boldsymbol{M}}_- = i e^{-i\hat{\boldsymbol{p}}^2 \ell/2k_0} \sin\left(\phi/2 - k\hat{\boldsymbol{x}}\right), \tag{2.27a}$$

$$\hat{\boldsymbol{U}}_\ell^2 \hat{\boldsymbol{M}}_+ = e^{-i\hat{\boldsymbol{p}}^2 \ell/k_0} \cos\left(\phi/2 - k\hat{\boldsymbol{x}}\right), \tag{2.27b}$$

$$\hat{\boldsymbol{U}}_\ell \hat{\boldsymbol{U}}_L = e^{-i\hat{\boldsymbol{p}}^2 \ell/2k_0} e^{ik_0\hat{\boldsymbol{x}}^2/2s_i}, \tag{2.27c}$$

which can be simplified using Equations (2.26) and (2.27) from above, an exact recursive simplification can be constructed. Written explicitly in the momentum basis we have

$$\hat{\boldsymbol{U}}_\ell \hat{\boldsymbol{M}}_- \left(\hat{\boldsymbol{U}}_\ell^2 \hat{\boldsymbol{M}}_+\right)^{r-1} = \left(\frac{1}{2}\right)^r \int \mathrm{d}p |p\rangle e^{-i\ell p^2/2k_0} \times$$
$$\sum_{j=-r}^{r} a_j(r,p) e^{ij\phi/2} \langle p + jk|, \tag{2.28}$$

where the functions $a_j(r,p)$ are given by the recursion relation

$$a_j(r,p) = a_{j-1}(r-1,p) e^{-i\ell(p+(j-1)k)^2/k_0} +$$
$$a_{j+1}(r-1,p) e^{-i\ell(p+(j+1)k)^2/k_0}. \tag{2.29}$$

The overall measurement operator is then given simply by multiplying the unitary operators $\hat{\boldsymbol{U}}_\ell$ and $\hat{\boldsymbol{U}}_L$ from the right. Note that each pulse retains the phase and momentum information accumulated from previous traversals.

Given an input Gaussian profile as in (2.8), we find

$$\langle x|\Psi_-^r\rangle = \left(\frac{1}{2}\right)^r \left(\frac{2}{\pi a^2}\right)^{1/4} \beta \times \tag{2.30}$$
$$\int \mathrm{d}p e^{ipx} e^{-i\ell p^2/2k_0} e^{-i\ell(p+k)^2/k_0} \times$$
$$\sum_{j=-r}^{r} a_j(r,p) e^{ij\phi/2} e^{-i\ell(p+jk)^2/2k_0} e^{-\beta^2(p+jk)^2},$$

with

$$\beta = \sqrt{\frac{\sigma^2 s_i}{s_i - 2ik_0\sigma^2}}. \tag{2.31}$$

The solution with no initial diverging lens is given by taking limit $s_i \to \infty$, which modifies Equation (2.30) with the replacement $\beta \to \sigma$. Similarly, setting $\ell = 0$ recovers the collimated solution as expected. The effect of a diverging lens is considered for comparison with the unrecycled experiment in Ref. [DSJH09], where such a lens was able to enhance sensitivity.

Table 2.2 shows the parameters which describe the laser and experimental geometry. Our choice of k_0 corresponds to the 780-800 nm lasers used in [DSJH09, SDJH09, SDW$^+$10b, SDW$^+$10a], and the 3-meter interferometer length ℓ is taken from the generous upper bound estimate discussed in Section 2.1 as a worst case scenario for beam divergence effects. For the weak-value regime of small $k\sigma$ and ϕ such that $k\sigma < \phi < 1$, we found that it was more computationally efficient to expand the sine and cosine functions in (2.27) to second order in k and fourth order in ϕ, in the weak value regime

this produces the expressions,

$$\hat{\boldsymbol{U}}_\ell \hat{\boldsymbol{M}}_- = \int \mathrm{d}p |p\rangle e^{-ip^2\ell/2k_0} \times \tag{2.32a}$$
$$\left[\frac{i\phi}{2} - \frac{i\phi^3}{48} + k\partial_p - \frac{k\phi^2}{8}\partial_p + \frac{ik^2\phi}{4}\partial_p^2 \right] \langle p|,$$

$$\hat{\boldsymbol{U}}_\ell^2 \hat{\boldsymbol{M}}_+ = \int \mathrm{d}p |p\rangle e^{-ip^2\ell/k_0} \times \tag{2.32b}$$
$$\left[1 - \frac{\phi^2}{8} + \frac{\phi^4}{384} + \frac{ik\phi}{2}\partial_p - \frac{ik\phi^3}{48}\partial_p \right.$$
$$\left. + \frac{k^2}{2}\partial_p^2 - \frac{k^2\phi^2}{16}\partial_p^2 \right] \langle p|.$$

These expansions can be numerically iterated more easily than the full solution, using the initial Gaussian profile (2.8) as an input. To test the validity of this truncation, we initially set the interferometer length ℓ to zero so that a comparison could be made with the previously calculated collimated solutions.

We restrict our attention to the SNR gains achieved by recycling a single pulse for r traversals, since adding more pulses leads to a simple scaling of the single pulse result. The SNR gains for different choices of initial diverging lens are shown in Fig. 2.8, where they are compared to the ideal power-scaling curve that we expect from our qualitative considerations given by (2.2). In all cases, the expected gains roughly follow the qualitative power scaling rule for a large number of traversals before saturating due to the beam growing larger than the finite size of the split-detector. Note that the beam divergence mitigates the SNR decay that was observed for the collimated

case, even without beam stabilization due to flipping or Zeno reshaping.

2.4   Remarks

By investigating the optical design shown in Fig. 2.1, we have shown how a single optical pulse can be trapped inside the interferometer until the photons all exit the dark port and are "post-selected," greatly boosting the sensitivity of the precision measurement. The added power accumulated at the detector within a fixed duration of time is the dominant source of sensitivity gain. Further increases are achievable by trapping multiple pulses in the interferometer simultaneously. The number of trapped pulses is limited by the length of the pulses, the gating frequency of current Pockels cells, and the physical size of the interferometer.

We carefully analyzed the case of a collimated beam and showed that repeated post-selections cause a walk-off effect in the recycled pulse, which tends to diminish the SNR. However, we also showed that these walk-off effects can be easily corrected by Zeno reshaping, or by a parity flip, which reflects the beam around its optic axis on each traversal. Somewhat surprisingly, the gains with parity correction can even exceed those expected from the power scaling. Including propagation effects does not destroy the sensitivity gain shown for the collimated case, but instead can produce additional enhancement.

While these sensitivity gains alone are a substantial improvement over

the original idea, the combination of these techniques with other established metrology techniques—such as the use of a squeezed reference beam—could further increase the sensitivity beyond that indicated here. These considerations have motivated our treatment using quantum operator methods, rather than classical wave optics, even though the latter is an equally valid approach for a coherent laser source.

Chapter 3

Continuous wave recycled weak value measurements

In this chapter, we consider an alternative recycling technique initially proposed by Drever [Dre83]. By placing a partially transmitting mirror at the bright port of an interferometer to form a resonant cavity, one is able to more efficiently use the input laser power by increasing the total power inside the interferometer. In the context of weak value-based experiments, the photons that would have previously exited the interferometer through the bright port now destructively interfere, being trapped inside the interferometer, so that, in the absence of loss, the entire input beam eventually exits the (formerly) dark port of the interferometer.

With the weak value deflection measurement inside the cavity, the full beam experiences the amplified deflection, which can then be detected with
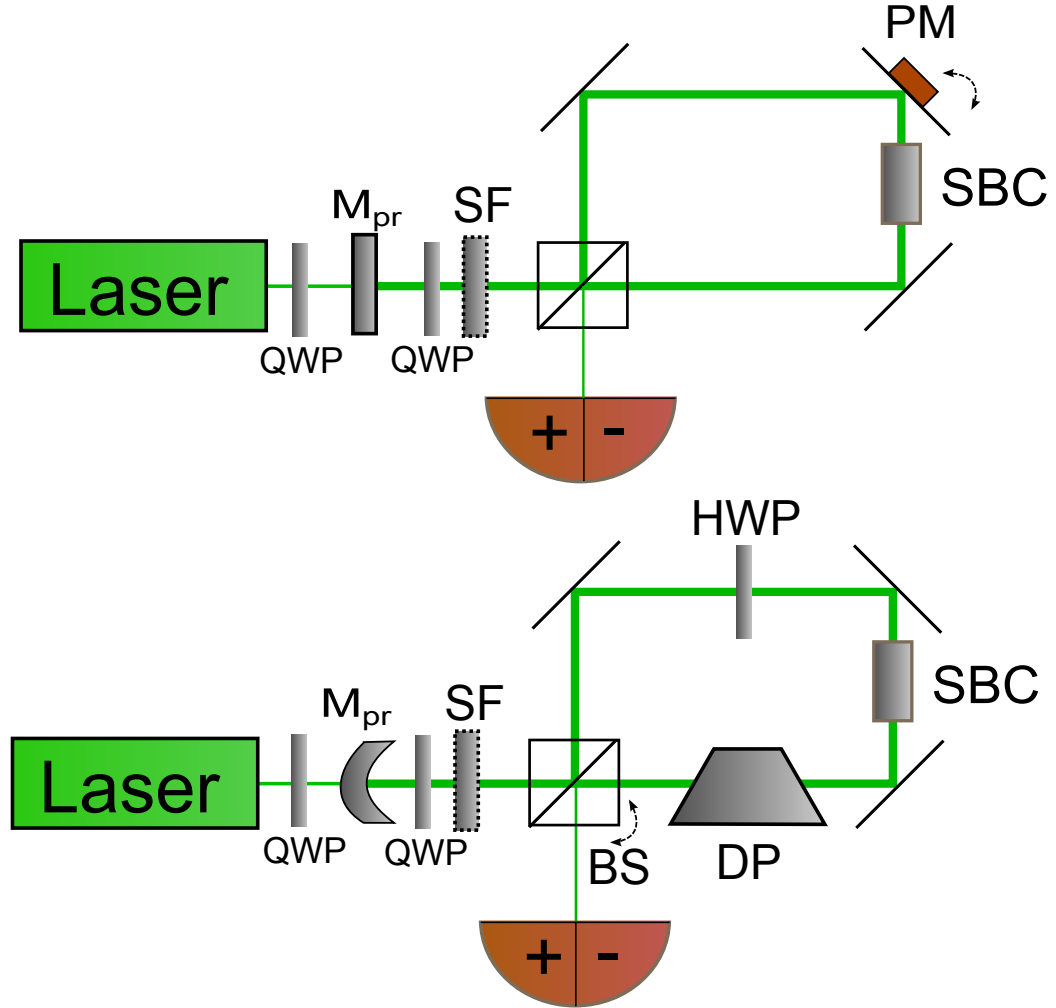
Figure 3.1: Schematics of weak value-based metrology with power recycling cavities. (a) Plane parallel cavity. A partially reflecting mirror (M_{pr}), forms an optical cavity in combination with a Sagnac interferometer, causing all the light to exit to the detector. The tunable phase difference between the counter-propagating arms is controlled by a Soleil-Babinet Compensator (SBC). A piezo-driven mirror (PM) imparts a small deflection to the beam, that is measured at a split detector. A spatial filter (SF) refreshes the beam profile on each pass. (b) Confocal cavity. A symmetric confocal resonator is established with the curved recycling mirror. Here the 50/50 beamsplitter (BS) is tilted instead of the mirror (which is now at the focus). A Dove prism (DP) inside the interferometer corrects for the profile flip when the Gaussian beam passes through its focus. The quarter waveplates (QWP) together with the SBC ensure the polarization optics leads to the same displacement at the detector on every pass.

a position sensitive detector, such as a split detector. Since the single-pass, postselected, weak value measurement displays the same SNR as the ideal focusing deflection technique [SDJH10, JMRH14, VMRA$^+$14], the power-recycling enhancement permits the amplification of the SNR itself by the large weak value factor, thus surpassing both techniques. We focus in this chapter on the continuous wave recycling of power to increase the precision of the weak value technique. A related pulsed recycling scheme, using a Pockels cell to trap a pulse inside the interferometer, was proposed in Ref. [DLJ$^+$13]. One important difference between the two ideas is that the pulsed scheme relies on the Pockels cell and polarization optics to trap the pulse in the interferometer, whereas here, it is the destructive interference of the reflected field that causes all the light to exit to the detector.

3.1 Weak Value Amplification

We first briefly review the interferometic weak value setup of Ref. [DSJH09] using a Sagnac interferometer. A phase difference ϕ between the counter-propagating modes of the interferometer may be controlled with a Soleil-Babinet Compensator (SBC) to break the clockwise/counterclockwise symmetry. A piezo-driven deflecting mirror is placed at the symmetry point of the interferometer to induce a small transverse momentum kick k, the parameter we wish to precisely measure. This parameter can be measured using a split detector placed at the output dark port.

We consider a continuous wave laser with a transverse Gaussian profile $E_0(x) = (N^2/2\pi\sigma^2)^{1/4} \exp(-x^2/4\sigma^2)$ that we have normalized to the average number of available photons N per unit time. The split detector gives a signal as the difference between the number of photons on the left versus right side of the detector, $S = N_R - N_L$, per unit time. The signal of the split detector at the dark port is linear in k for the interferometric weak value measurement (assuming $k\sigma \ll \phi/2 \ll 1$) [DSJH09, DLJ$^+$13],

$$\langle S \rangle \approx 2\sqrt{\frac{2}{\pi}}\, N_{\text{det}}\, \frac{2k\sigma}{\phi}, \tag{3.1}$$

where N_{det} is the total number of photons that are detected out of the initial beam. From a quantum mechanical perspective, we note that the large factor of $2/\phi$ is related to the weak value of the which-path operator, $\hat{\boldsymbol{W}}$, given by $\langle \psi_f | \hat{\boldsymbol{W}} | \psi_i \rangle / \langle \psi_f | \psi_i \rangle = -i \cot(\phi/2) \approx -2i/\phi$. Here, $\hat{\boldsymbol{W}} = |\circlearrowright\rangle\langle\circlearrowright| - |\circlearrowleft\rangle\langle\circlearrowleft|$ is defined with the two orthogonal circulating states $|\circlearrowright\rangle$ and $|\circlearrowleft\rangle$ of the Sagnac interferometer, and $|\psi_i\rangle$ and $|\psi_f\rangle$ are the pre- and post-selected states of the interferometer, defined by the phase ϕ and the selection of the port to measure. The weak value effectively amplifies the kick k for each collected photon. The variance of this signal is limited by the detected shot noise N_{det}, so the SNR $\mathcal{R}$ is given by

$$\mathcal{R} \equiv \frac{\langle S \rangle}{\sqrt{N_{\text{det}}}} \approx 2\sqrt{\frac{2}{\pi}}\, \sqrt{N_{\text{det}}}\, \frac{2k\sigma}{\phi}. \tag{3.2}$$

3.2 Continuous-wave power recycling

To introduce power recycling, we modify the previous setup of Ref. [DSJH09] by adding a partially transmitting mirror, illustrated in Fig. 3.1(a), to make the Sagnac interferometer a resonant optical cavity. In the absence of the cavity, the detected number of photons $N_{\text{det}} = p\,N$ is small when the postselection probability $p \approx (\phi/2)^2$ of the dark port is also small. The resulting factor of $\sqrt{p}$ in Eq. (3.2) thus exactly cancels the amplification of the large weak value factor, and precisely recovers the SNR that one expects from a traditional beam deflection measurement with all the light [SDJH09], but still with enhanced robustness to some types of technical noise [JMRH14, VMRA^{+}14]. However, the addition of the cavity will permit the entire input beam to exit the (formerly) dark port and be detected with the amplified deflection. As we will soon show, including the cavity ideally makes $N_{\text{det}} = N$, so the large weak value $(2/\phi)$ directly amplifies the SNR itself in Eq. (3.2), giving a large prefactor to the standard quantum limit scaling. We note this variation also turns the probabilistic weak value method into a deterministic one with respect to the output port.

To see why such a resonant cavity will permit the entire beam to be detected, consider an initial amplitude E_0 that is incident on a cavity formed by two partially transmitting mirrors with transmission, t_1 and t_2, and reflection, $r_1 = \sqrt{1 - t_1}$ and $r_2 = \sqrt{1 - t_2}$, coefficients. For light inside the cavity, each round trip adds a phase θ, which depends on the geometry of the cavity

and the wavelength of light, giving a geometric series for the amplitude,

$$E_{\text{cav}} = t_1[1 + r_1 r_2 e^{i\theta} + (r_1 r_2 e^{i\theta})^2 + \cdots] E_0,$$

$$= \frac{t_1}{1 - r_1 r_2 e^{i\theta}} E_0. \tag{3.3}$$

The light reflected back towards the laser is similarly geometric, with the amplitude E_r being a superposition of amplitudes from light directly reflected from the first mirror and multiple reflections inside the cavity,

$$E_r = \left[-r_1 + \frac{t_1^2 r_2 e^{i\theta}}{1 - r_1 r_2 e^{i\theta}}\right] E_0. \tag{3.4}$$

If $\theta = 2\pi n$, where n is any integer, and $r_1 = r_2 \equiv r$, then the reflected amplitude is exactly zero, so the power leaving the cavity through the second mirror becomes equal to the input laser power. This condition is known as impedance matching [SMH$^+$97]. In this case the light intensity inside the cavity is amplified above the input laser intensity by a gain factor

$$G \equiv \frac{|E_{\text{cav}}|^2}{|E_0|^2} = \frac{t^2}{(1 - r^2)^2} = \frac{1}{t^2} = \frac{1}{T}, \tag{3.5}$$

equivalent to the inverse transmission probability T.

We will soon show that these general results may be applied to our weak-value amplification cavity formed from one partially transmitting mirror and the partially transmitting dark port of the interferometer. We effectively replace the transmission T in (3.5) by the postselection probability of exiting

the interferometer, $T \to p \approx (\phi/2)^2$. Thus, the large power gain inside the inteferometer allows for the small postselection probability to give the entire input beam out of one port, and no light out of the other port, boosting the SNR of Eq. (3.2) by $1/\sqrt{p}$.

3.3 Resource counting

The relative advantages of one technique to another should specify the resources given as a constraint. We note that if the resource is taken to be the number of *detected* photons, then there is already an advantage in the single-pass weak value experiment over the direct deflection experiment (not considering technical noise sources) [DSJH09, SDJH09]. If instead the resource is the total number of photons entering the interferometer, the power-recycled proposal gives an advantage $1/\sqrt{p}$ over the single pass experiment for many cycles. Perhaps a fairer way of counting resources is the number of times an interaction takes place. In this proposal, the enhancement in sensitivity is due to an effective power increase, but we stress that signal does not accumulate for multiple passes. The total number of interactions with the unknown parameter is NM, where N is the total photon number, and M is the number of times a photon enters the interferometer. While M is a stochastic variable in our system, on average it is $1/p$, giving N/p as the total number of interactions, whose square-root will be shown to determine $\mathcal{R}$ (3.12). Another interesting technique using multiple interactions

where the signal scales linearly with the number of cycles, while keeping the noise constant is a simple example of *signal recycling* [MS91]. In this case recycled photons accumulate additional momentum kicks on each traversal thereby enhancing the signal. This separate technique can be used together with power recycling as a complimentary method, and was already incorporated in this kind of measurement via an optical lever by the Kasevich group [HHC$^+$11].

Recycling with flat mirrors.—The remaining requirement for the weak value technique to still work in the presence of the cavity is for the transverse position profile at the beamsplitter to be preserved, so that the enhanced deflection remains with each traversal through the interferometer. We now calculate the transverse profile for a beam confined by flat cavity mirrors by adapting the operator approach used in [DLJ$^+$13]. While all intensities in this setup may be calculated from classical wave optics, it is convenient to adopt a quantum state analysis. The flat mirror approach is a reasonable approximation if the roundtrip distance of the cavity multiplied by the finesse is much less than the Rayleigh length of the beam. Later we will discuss a more realistic cavity design that uses a curved mirror to confine a Gaussian beam.

To determine the steady state beam profile at the detector, we introduce the "system" state $|\psi\rangle$ spanned by the orthogonal circulating modes $|\circlearrowright\rangle$ and $|\circlearrowleft\rangle$ of the Sagnac interferometer, and the "meter" state $|\varphi\rangle$ which represents the transverse profile of the beam, with the position amplitude for a single

photon given by $\langle x|\varphi_0\rangle = E(x)/\sqrt{N}$. The total state in the interferometer is then the tensor product $|\Psi\rangle = |\psi\rangle|\varphi_0\rangle$. In what follows, $\hbar = 1$.

The beam will experience two distinct effects inside the interferometer that depend on the path. First, the tilted mirror couples the system and meter by imparting a momentum kick k to the transverse beam depending on the path taken, which modifies the state with the unitary operator $\hat{U}_{\mathrm{PM}} = e^{ik\hat{x}\hat{W}}$, depending on the which-path operator. Second, the SBC produces a net phase shift ϕ between the circulating modes, corresponding to the unitary operator $\hat{U}_{\mathrm{SBC}} = e^{-i\phi\hat{W}/2}$. To take into account small but constant losses, we introduce the nonunitary operator $\hat{L} = \sqrt{1-\gamma}\,\hat{1}$, where γ is the probability of loss per traversal from all optical imperfections. Note that the "loss" of the cavity to the detector via the beamsplitter is treated separately.

After entering the interferometer through the 50:50 beamsplitter, the path state becomes an equal superposition of circulating modes $|\psi_+\rangle = \frac{1}{\sqrt{2}}\left(|\circlearrowleft\rangle + i|\circlearrowright\rangle\right)$, which is also the projection state for the bright port. The dark port is correspondingly described by the orthogonal state $|\psi_-\rangle = \frac{1}{\sqrt{2}}\left(|\circlearrowleft\rangle - i|\circlearrowright\rangle\right)$. Since $\hat{U}_{\mathrm{PM}}$ and $\hat{U}_{\mathrm{SBC}}$ are the only non-trivial actions on the system Hilbert space, it is convenient to combine their effects with the projection onto the output ports of the interferometer, which produces measurement operators $\hat{M}_\pm = \langle\psi_\pm|\hat{U}_{\mathrm{PM}}\hat{U}_{\mathrm{SBC}}|\psi_+\rangle$ given by

$$\hat{M}_+ = \cos\left(\phi/2 - k\hat{x}\right), \qquad \hat{M}_- = i\sin\left(\phi/2 - k\hat{x}\right), \qquad (3.6)$$

where $\hat{\boldsymbol{x}}$ is the position operator, so $\hat{\boldsymbol{M}}_{\pm}$ are diagonal in the position basis. Here we have used $\langle\psi_{\pm}|\hat{\boldsymbol{W}}^n|\psi_+\rangle = (1 \pm (-1)^n)/2$.

Zeno refreshing.—After many traversals, transverse beam degradation tends to diminish the signal as discussed in [DLJ+13]. One strategy to solve this problem is to introduce a Gaussian spatial filter as shown in Fig. 3.1 [1]. Although a spatial filter is not essential for a successful power recycling scheme, we treat it here because it offers a more straightforward analysis compared to other arrangements. The filter acts as a projection back onto the initial state, which can be implemented with an additional projection operator, $\hat{\boldsymbol{\Pi}} = |\varphi_0\rangle\langle\varphi_0|$, so that the (normalized) state after the filter is again given by $|\varphi_0\rangle$. The overlap between the one-pass transverse state and the initial transverse state is close to 1, indicating that the vast majority of the time, a photon will pass unimpeded through the filter. This may be interpreted as a Zeno effect, refreshing the transverse profile, which is ideally lossless in the small k limit as we will now see. The probability of exiting the beam splitter toward the recycling mirror is given by

$$P_+ = |\hat{\boldsymbol{M}}_+|\varphi_0\rangle|^2 = (1/2)[1 + \cos\phi\exp(-2k^2\sigma^2)], \qquad (3.7)$$

leaving the pre-filter normalized state as $|\varphi'\rangle = \hat{\boldsymbol{M}}_+|\varphi_0\rangle/\sqrt{P_+}$. The proba-

---

[1]It should be noted this is an ideal Gaussian filter (as could be well approximated by coupling into a single-mode optical fiber) rather than a typical lens-pinhole spatial filter which would suffer unacceptable diffraction and absorption losses.

bility of surviving the filter is given by $P_Z = |\langle \varphi_0 | \varphi' \rangle|^2$,

$$
\begin{aligned}
P_Z &= (1/P_+) \left(\int dx\, \varphi_0^2(x) \cos(\phi/2 - kx) \right)^2, \\
&= \frac{\cos^2(\phi/2)}{\sinh(k^2\sigma^2) + \cos^2(\phi/2)e^{-k^2\sigma^2}}, \\
&\approx 1 - (\phi/2)^2 k^2\sigma^2 - k^4\sigma^4/2 + \dots,
\end{aligned} \tag{3.8}
$$

where the last approximation is taken in the weak value parameter range, $k\sigma \ll \phi/2 \ll 1$. We note that the exact expression for P_Z correctly equals one when $k = 0$ for any value of ϕ, while the filter loss is $k^2\sigma^2\phi^2/4$ to leading order.

This small amount of loss per cycle can be incorporated into the loss from the imperfect optics (*e.g.* unwanted reflection and absorption events in the cavity) as $\gamma \to \gamma + k^2\sigma^2\phi^2/4$, which we assume to be small compared to 1. With the filter included, the transverse beam profile is refreshed every cycle, making the calculations of the many-cycle case straightforward. The steady state amplitude exiting the detection port is given by the sum of amplitudes from all traversal numbers,

$$
\begin{aligned}
|\varphi\rangle &= t\sqrt{1-\gamma}\hat{M}_- \sum_{n=0}^{\infty} \left(r\sqrt{(1-\gamma)P_+} \right)^n |\varphi_0\rangle, \\
&= \frac{it\sqrt{1-\gamma}\sin(\phi/2 - k\hat{x})}{1 - r\sqrt{(1-\gamma)P_+}} |\varphi_0\rangle,
\end{aligned} \tag{3.9}
$$

where $P_+ \approx \cos^2(\phi/2)$ is the probability of exiting the recycling mirror port (3.7). Similarly, the steady state amplitude of light reflected back towards

the laser is

$$|\varphi_r\rangle = \left[-r + \frac{t^2\sqrt{1-\gamma}\cos(\phi/2)}{1 - r\sqrt{(1-\gamma)P_+}} \right] |\varphi_0\rangle, \tag{3.10}$$

which yields an impedance matching condition of $r = \sqrt{(1-\gamma)P_+}$. Notice that this choice stops any light leaking back through the recycling mirror in the steady state, regardless of the losses involved (as long as the beam remains phase-coherent, and its spectral and spatial profiles correctly realign back at the input confining mirror). In the limit of small loss, this corresponds to setting the mirror transmission amplitude $t \approx \phi/2$.

Using an initial Gaussian profile as originally considered, the average split detector signal is still given by Eq. (3.1), but now the total number of detected photons is

$$N_{\text{det}} = N \int_{-\infty}^{\infty} dx \, |\langle x|\varphi\rangle|^2 \approx N \left(1 - \frac{4\gamma}{\phi^2} \right), \tag{3.11}$$

giving all of them, minus losses. Therefore, the detector has an SNR of

$$\mathcal{R} \approx 4\sqrt{\frac{2}{\pi}} \sqrt{N} \frac{k\sigma}{\phi} \left[1 - \frac{2\gamma}{\phi^2} \right]. \tag{3.12}$$

From the spatial filter, the minimum loss for ideal optics is $\gamma \approx k^2\sigma^2\phi^2/4$, which produces an overall (negligible) loss factor of $(1 - k^2\sigma^2/2)$ in the SNR. As predicted after Eq. (3.2), the SNR has been increased by the weak value factor of $(2/\phi) = 1/\sqrt{p}$ from the SNR of the single-pass weak value setup (and thus the traditional deflection setup) when the loss $\gamma \ll \phi^2/4$ is small.

3.4   Power recycling with curved mirrors

While the recycling analysis with flat mirrors is straightforward, this cavity geometry is on the borderline between stable and unstable [ME10]. In practice, the beam will be a diffracting Gaussian beam with a curved phase front rather than the collimated beam treated above. To confine such a Gaussian beam in a stable configuration, the radius of curvature of the recycling mirror must match the radius of curvature of the phase fronts to ensure proper phase cancellation at the mirror. This geometry is sketched in Fig. 3.1(b). The remaining flat mirrors in our setup have no effect on the confinement properties. The resonant cavity is characterized by its *finesse*, the typical number of bounces before the beam decays. In order to maximize the gains from the recycling procedure, the finesse should exceed the inverse probability to exit the interferometer to the detector. Placing the beam focus at the far mirror ensures that the symmetry of the interferometer paths is not disturbed by the changing beam waist, enabling proper interference of the clockwise and counter-clockwise propagating curved phase fronts at the beam splitter. This geometry defines a symmetric confocal cavity, which has well-established properties. The confocal cavity lies at the other stability extreme of the plane parallel geometry [ME10].

The physics of this cavity is similar to that of the collimated analysis, with a few important differences. The beam will achieve its minimum waist σ_0 at the far mirror and its maximum waist σ at the recycling mirror, where

it should match the spatial and spectral profiles of the input beam. If the coordinate z along the optical axis is measured from the minimum waist at the symmetry point, and the maximum cavity length between that point and recycling mirror is ℓ, then the beam width inside the cavity is given by $\sigma(z) = \sigma_0 \sqrt{1 + z^2/\ell^2}$, where ℓ is also equal to the Rayleigh range, and $\sigma(\ell) = \sqrt{2}\,\sigma_0 \equiv \sigma$ is the input beam waist. Putting the transverse mirror momentum kick k at the focus of the cavity does not yield a sensitive response, so we put the momentum kick instead on the beam splitter, as was done in the experiments of Ref. [VMRA$^+$14]. The transmitted beams acquire no momentum kick, while the reflected beams acquire a momentum kick k. The presence of a focus in the cavity gives two new effects. The first is that a Gouy phase appears from the focus, giving an additional phase factor of π in both beams [ME10]. The second more important effect is that passing through the focus flips the tilt of the phase front, so the effective transverse momentum kick from the beam splitter is inverted when the expanding beam returns to the beam splitter, $k \to -k$. If left uncorrected, this momentum kick is undone by the additional k momentum kick from the second reflection to the detector. We can compensate for this effect by adding a Dove prism inside the interferometer, which provides a transverse parity flip to restore the previous phase front, recovering similar weak value physics to the collimated case. The only significant difference from the previous analysis concerns the changing width of the beam $\sigma(z)$. The choice of cavity geometry will set the width σ in Eq. (3.12).

3.5 Remarks

By including a power-recycling mirror in a continuous wave interferometric weak value amplification setup, we are able to maintain the large pointer shift associated with previous weak value amplification experiments while acquiring all of the input light in principle. Our main result is that the SNR (or, equivalently, the Fisher information about the desired parameter) is boosted by the weak value factor, which can be made large in principle, limited only by the fidelity of the optics and the finesse of the cavity. We have given two different cavity geometries to realize this proposal, but other stable geometries giving similar physics also exist.

In this work, we have focused on the interferometric implementation of the optical weak value effect to propose the use of the power recycling technique. However, the same basic idea may be applied to other experimental realizations of the same, such as the polarization-based version [HK08], where the postselection is accomplished with a polarizing beam splitter, and the other output beam is reinjected into the experiment.

Power recycling is only one of the techniques used in precision interferometric measurements. There are several others which may be able to be combined with our setup as well. As for further improvements in sensitivity, we have already discussed the possibility of recycling the signal. Future work may focus on the combination of this technique and quantum light metrological approaches such as using squeezed and entangled states [Cav81, TAB$^+$02].

Chapter 4

Inverse weak value measurements under technical noise

In this chapter, we will thoroughly examine the ability of weak value amplification to improve measurement sensitivity. As we will show, the principal reason is that they allow certain types of technical noise or other experimental limitations to be overcome, while still achieving the same sensitivity as ideal traditional measurement schemes [SDJH09, JMRH14, KG14, VMRA$^+$15, AEdMF$^+$15, TSS16, HBL16]. Previous theoretical work demonstrating the effectiveness of weak value amplification in reducing the negative effects of technical noise has been in what we term the *weak value regime*. Namely, a small but known phase parameter is used in the system to amplify

a very small and unknown interaction parameter (*i.e.,* one coupling system and meter quantum dynamical variables). It has also been recently proven that in the case of systematic noise on the meter, weak value amplification can suppress this systematic contribution in comparison to a direct measurement method [PABJ16]. Further improvements in precision can be obtained by recycling the non-selected events [DLJ+13, LDJ+15, BGJK15, WTH+16].

In this chapter we consider the opposite case where we have a small but known interaction parameter which is used to amplify a very small and unknown phase. We refer to this as the *inverse weak value regime*, following previous work [SDW+10b, DLJ+13, LDJ+15, KAN12].

The chapter is organized as follows: in Sec. 4.1, we briefly review weak measurements in a two level system, and compare basic results in the weak value and inverse weak value regimes. In Sec. 4.2, we calculate the Fisher information for weak measurements in the inverse weak value regime, and compare the result with traditional phase estimation schemes. In Sec. 4.3, we treat the case of uncorrelated additive Gaussian technical noise, and in Sec. 4.4 we treat angular jitter and diffraction effects in optical phase measurements. In Sec. 4.5, we extend our analysis to the case of precision frequency measurements in the inverse weak value regime, and demonstrate that such a measurement will have the same robustness to technical noise as phase measurements.

4.1 Inverse weak value

Here we briefly recall the usual weak value amplification process in a two level system, and then discuss how this relates to the process of amplification by the so-called inverse weak value.

We take an initial "system" state $|i\rangle$ and "meter" state $|\varphi\rangle$ to be in a product state, *i.e.*, $|\Psi\rangle = |i\rangle \otimes |\varphi\rangle$. An effective interaction Hamiltonian

$$H_1 = g_1(t)\hat{\boldsymbol{\sigma}}_z \otimes \hat{\boldsymbol{x}}, \tag{4.1}$$

entangles the system and meter weakly, with $g(t) \ll 1$ for some short time interval. Here $\hat{\boldsymbol{\sigma}}_z$ is the usual Pauli z spin operator which acts on the system state space, and $\hat{\boldsymbol{x}}$ is taken to be the (transverse) position operator acting on the meter state space. If the intended system postselection state $|f\rangle$ is orthogonal to $|i\rangle$, it is also necessary to introduce a second interaction in order to create some small but nontrivial transition probability between the initial and final system states. A simple choice would be

$$H_2 = g_2(t)\hat{\boldsymbol{\sigma}}_z \otimes \hat{\boldsymbol{1}}, \tag{4.2}$$

which only has a nontrivial action on the system space. Qualitatively, we see that $g_1(t)$ generates a displacement in the basis of the operator conjugate to $\hat{\boldsymbol{x}}$, with the sign of the displacement opposite for each of the two basis vectors, and $g_2(t)$ generates a relative phase between the two basis vectors of

$\hat{\boldsymbol{\sigma}}_z$. The action of these two operators on the state $|\Psi\rangle$ is straightforwardly given by the unitary operators

$$\hat{\boldsymbol{U}}_1 = \exp(-ik\hat{\boldsymbol{\sigma}}_z \otimes \hat{\boldsymbol{x}}),$$
$$\hat{\boldsymbol{U}}_2 = \exp(-i\frac{\phi}{2}\hat{\boldsymbol{\sigma}}_z \otimes \hat{\boldsymbol{1}}), \tag{4.3}$$

where k and ϕ are effective interaction parameters found by integrating the above Hamiltonians over the relevant time interval.

Since $[\hat{\boldsymbol{U}}_1, \hat{\boldsymbol{U}}_2] = 0$, we can simply combine the operators as $\hat{\boldsymbol{U}} \equiv \hat{\boldsymbol{U}}_1\hat{\boldsymbol{U}}_2$ by adding the exponents.

Beginning with our state $|\Psi\rangle$, we find that in the diagonal basis of $\hat{\boldsymbol{x}}$, we can represent the postselected state as

$$\langle x|\Psi'\rangle = \langle x| \otimes \frac{|f\rangle\langle f|}{\sqrt{p_f}}\hat{\boldsymbol{U}}|i\rangle \otimes |\varphi\rangle,$$
$$= \frac{-i\langle f|\hat{\boldsymbol{\sigma}}_z|i\rangle}{\sqrt{p_f}}\sin\left(kx + \frac{\phi}{2}\right)\langle x|\varphi\rangle|f\rangle, \tag{4.4}$$

where p_f is the postselection probability. From this point we may ignore the system state and focus on the meter after postselection. The postselection probability is calculated directly as

$$p_f = \int_{-\infty}^{\infty} dx \ |\langle x|\varphi\rangle|^2 \left|\langle f|\hat{\boldsymbol{U}}|i\rangle\right|^2. \tag{4.5}$$

In the specific case of a Gaussian initial meter state with variance σ^2, this

yields

$$p_f = \frac{1 - e^{-2k^2\sigma^2}\cos\phi}{2}. \tag{4.6}$$

In the *weak value regime*, we have $k\sigma$ small compared to $\frac{\phi}{2}$ for values of x where $|\langle x|\varphi\rangle|$ is not very small. For a concrete example we take $|i\rangle = 2^{-1/2}(|0\rangle + |1\rangle)$, and $|f\rangle = 2^{-1/2}(|0\rangle - |1\rangle)$. In the weak value regime this yields $p_f \approx |\langle f|\hat{U}_2|i\rangle|^2 = \sin^2\frac{\phi}{2}$, and $\langle x\rangle \approx k\sigma^2\cot\frac{\phi}{2}$. Hence, a small known phase parameter $\phi/2$ is used to amplify $\langle x\rangle$, which in turn allows for an improved estimation of the unknown parameter k.

In the *inverse weak value regime*, the situation is reversed, where one wishes to use a small known parameter k to amplify an unknown phase parameter $\phi/2$. This does not change Eq. (4.4), but there is no simple way to calculate the postselection probability without assuming a particular initial meter state. Also note that in this regime, there is certain to be an eigenvalue x which yields zero amplitude upon postselection, *i.e.,* where $kx + \phi/2 = 0$. Taking the same initial and final system states as above, and again assuming a Gaussian initial meter state with variance σ^2, we find

$$p_f \approx k^2\sigma^2,$$
$$\langle x\rangle \approx \frac{\phi}{k}. \tag{4.7}$$

4.2 Fisher information

We now turn our attention to estimating the unknown parameter ϕ in the inverse weak value regime for a Gaussian initial meter state. Hence, postselected events will be sampled from the probability distribution

$$p(x|\phi) = \frac{|\langle f|\hat{\boldsymbol{\sigma}}_z|i\rangle|^2}{p_f} \sin^2\left(kx + \phi/2\right) |\langle x|\varphi\rangle|^2, \tag{4.8}$$

where the postselection probability p_f is given by Eq. (4.6). As in the previous section, we will choose orthogonal initial and final states so that $\langle f|\hat{\boldsymbol{\sigma}}_z|i\rangle = 1$. The metric we will use to determine the sensitivity of the estimation scheme is the Fisher information, defined by

$$\mathcal{I}(\phi) = -\int_{-\infty}^{\infty} \mathrm{d}x\; p(x|\phi)\partial_\phi^2 \ln p(x|\phi),$$
$$= \frac{e^{4k^2\sigma^2} - 1}{(e^{2k^2\sigma^2} - \cos\phi)^2}. \tag{4.9}$$

If there are ν total events, then $p_f\nu$ events will be postselected. Since the Fisher information scales linearly with the number of independent detected events, we are interested in the quantity

$$\mathcal{I}_\nu(\phi) = p_f\nu\mathcal{I}(\phi),$$
$$\approx \frac{\nu}{2}\left(1 + e^{-2k^2\sigma^2}\right), \tag{4.10}$$

where the approximation is from expanding to first order in ϕ (*i.e.* $\cos\phi \approx 1$).
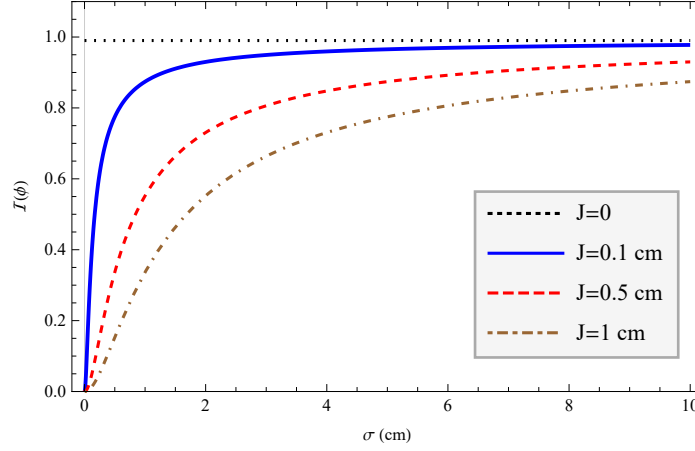
Figure 4.1: Fisher information for estimation of the parameter ϕ under the effect of additive Gaussian white noise, for different noise amplitudes J as a function of the initial state standard deviation σ. The product $k\sigma$ is held constant in this plot to keep the postselection probability (4.6) constant.

This expression approaches ν for small $k\sigma$. Since $\mathcal{I}_\nu(\phi) = \nu$ for more conventional phase estimation schemes such as balanced homodyne detection (or any scheme with Poissonian uncertainties), the inverse weak value approach can approximately recover the full Fisher information with a small subensemble of all events. The same result has been found in the weak value regime [JMRH14].

4.3 Uncorrelated Gaussian noise

For the case of uncorrelated Gaussian technical noise, such as might be the case with beam jitter in an optical experiment, the measured signal position s will be the sum of a position x drawn from the distribution (4.8) and a

random displacement ξ drawn from a Gaussian distribution with zero mean

and variance J^2. Hence, $s = x + \xi$, and

$$
\begin{aligned}
p(s|\phi) &= \frac{1}{\sqrt{2\pi J^2}} \int \mathrm{d}\xi \int \mathrm{d}x \ e^{-\xi^2/2J^2} p(x|\phi)\delta(s - x - \xi), \\
&= \frac{1}{\sqrt{2\pi J^2}} \int \mathrm{d}x \ \exp\left(\frac{-(s-x)^2}{2J^2}\right) p(x|\phi), \\
&= \frac{1}{\sqrt{2\pi(J^2 + \sigma^2)}} \exp\left[\frac{-s^2}{2(J^2 + \sigma^2)}\right] \Bigg(1- \\
&\quad \cos\left[\frac{2ks\sigma^2}{J^2 + \sigma^2} + \phi\right] \exp\left[\frac{-2J^2k^2\sigma^2}{J^2 + \sigma^2}\right]\Bigg).
\end{aligned} \tag{4.11}
$$

It is difficult to directly calculate the Fisher information of this distribu-

tion for an estimator of the parameter ϕ. If we expand to linear order in ϕ

and $k\sigma$ with $\phi < k\sigma < 1$ however, the integration is straightforward. For

clarity of the final result we also expand to first order in J/σ, however our

conclusions are not dependent on this truncation. The resulting information

is given by

$$
\mathcal{I}(\phi) = \frac{1}{k^2\sigma^2}\left(1 - \sqrt{\frac{\pi}{2}}\frac{J}{\sigma}\right). \tag{4.12}
$$

If we choose $\sigma \gg J$, this expression tends to

$$
\mathcal{I}(\phi) = \frac{1}{k^2\sigma^2}. \tag{4.13}
$$

Recall $k^2\sigma^2$ is equal to the postselection probability (4.7) in the inverse weak

value regime. We also note this is equal to Eq.(4.9) when expanded to lowest

nonvanishing order in ϕ and $k\sigma$. Hence, for a sufficiently large choice of σ the Fisher information for ν independent events is given by

$$\mathcal{I}_\nu(\phi) = p_f \nu \mathcal{I}(\phi),$$

$$\approx \nu. \tag{4.14}$$

Remarkably, it is possible to approximately recover the full Fisher information in the presence of additive Gaussian technical noise, again using only a small subensemble of total events, provided the Gaussian width σ is sufficiently large.

4.4   Angular jitter and diffraction

We now consider propagation effects (*e.g.* diffraction), and an angular jitter of an optical beam which we model as an additional tilted mirror as in Fig. 4.2. We can model the unnormalized state at the detector as

$$\langle x|\varphi\rangle = \langle x|\hat{U}_{\ell_2}\hat{U}_k\hat{U}_\phi\hat{U}_{\ell_1}\hat{U}_q|\varphi_0\rangle,$$

$$= \langle x|e^{-i\hat{\boldsymbol{p}}^2\ell_2/2k_0}\sin(k\hat{\boldsymbol{x}} + \phi/2)e^{-i\hat{\boldsymbol{p}}^2\ell_1/2k_0}e^{iq\hat{\boldsymbol{x}}}|\varphi_0\rangle, \tag{4.15}$$

where $\hat{U}_\ell$ is a quadratic phase in momentum space which describes propagation of the beam front by a distance ℓ, and $|\varphi_0\rangle$ is the initial Gaussian meter state. If we take the angular jitter to be normally distributed with zero mean
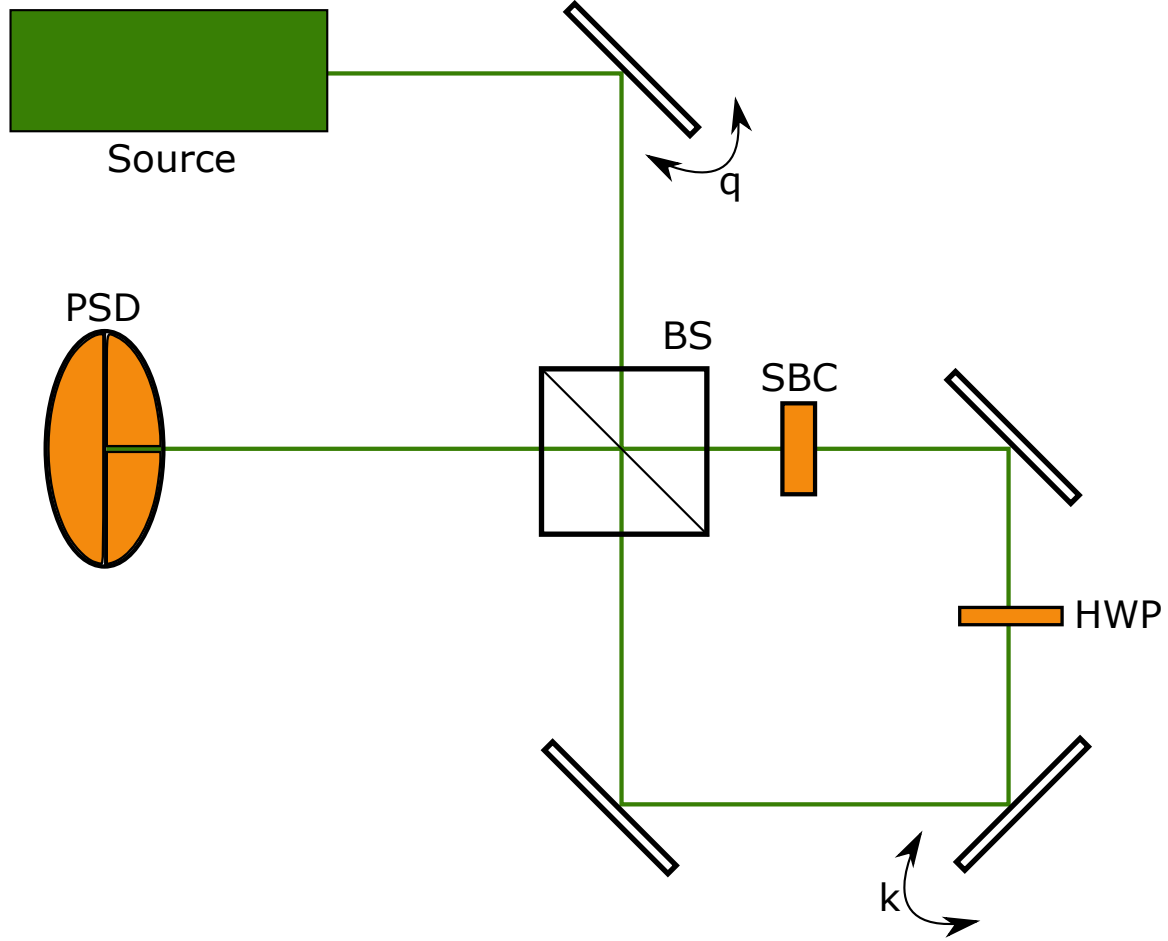
Figure 4.2: An optical two-level weak value amplification setup. Angular jitter noise is modeled by the first mirror which imparts a random momentum shift q to the beam. The beam then passes through a beam splitter (BS) and acquires a path-dependent phase at the Soleil-Babinet compensator (SBC) with half-wave plate (HWP), and a path-dependent momentum shift k at the titled signal mirror inside the Sagnac interferometer. Postselected light is measured at the position sensing detector to determine the unknown phase shift ϕ imparted by the SBC. The distance between the source and the symmetric point of the interferometer is ℓ_1 and the distance between the symmetric point and the position-sensitive detector (PSD) is ℓ_2.

and variance Q^2, the overall distribution is given by

$$p(x, q|\phi) = \mathcal{N} \exp\left(\frac{-q^2}{2Q^2}\right) |\langle x|\varphi\rangle|^2, \tag{4.16}$$

where $\mathcal{N}$ is a normalization constant. Note that in the limit of a collimated beam (i.e., $\ell_1, \ell_2 \to 0$), the momentum kick q will have no effect on the measured profile as it only appears as an overall phase factor in the wavefunction. The Fisher information for an estimator of the parameter ϕ is then

$$\mathcal{I}(\phi) = -\int dx \; dq \; p(x, q|\phi) \partial_\phi^2 \ln p(x, q|\phi). \tag{4.17}$$

This integral is very complicated for the general state (4.15), but becomes tractable with a few simplifying assumptions. Namely, we expand to linear order in ϕ and $k\sigma$ in the inverse weak value regime so that $\phi < k\sigma \ll 1$. We also expand to linear order in kQ, but make no assumption about the relative size of kQ to the other parameters. Finally, since ℓ_2 is completely under the experimenter's control, and can be made arbitrarily small, we take $\ell_2 = 0$. This yields a per-event Fisher information of

$$\mathcal{I}(\phi) \approx \frac{4k_0^2}{k^2} \cdot \frac{4k_0^2\sigma^2 + \ell_1^2\left(\frac{1}{\sigma^2} - 4Q^2\right)}{\left(\frac{\ell_1^2}{\sigma^2} + 4k_0^2\sigma^2\right)^2}. \tag{4.18}$$

In the limit of $k_0\sigma \gg \ell_1/\sigma$ and $k_0\sigma \gg \ell_1 Q$, this reduces to

$$\mathcal{I}(\phi) = \frac{1}{k^2\sigma^2}, \tag{4.19}$$

which is simply the inverse of the postselection probability. Hence, the total Fisher information for ν independent events becomes

$$\mathcal{I}_\nu(\phi) = p_F \nu \mathcal{I}(\phi),$$
$$\approx \nu. \tag{4.20}$$

Amplification in the inverse weak value regime under the effect of angular jitter noise allows for recovery of the full noiseless Fisher information for large beam width σ, as in the case of additive Gaussian white noise considered in Sec. 4.3. This is confirmed numerically using the full state (4.15), with results shown in Fig. 4.3.

4.5 Frequency measurements in the inverse weak value regime

It is straightforward to extend the above results for precision phase estimation to frequency estimation. Here we show an optical frequency measurement scheme which is robust to certain kinds of additive technical noise.

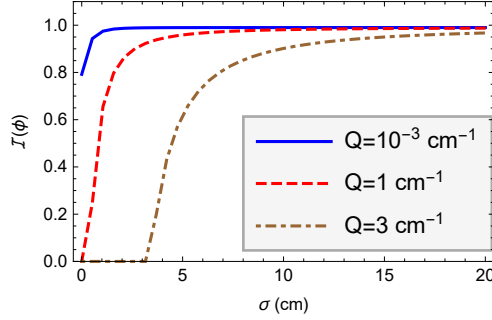We can convert the phase measurements above to a frequency shift mea-

Figure 4.3: Fisher information as a function of beam width for two different amplitudes of angular jitter, with $k\sigma$ held at a constant value of .1. The numeric values of the parameters are $\phi = 10^{-3}$, $k_0 = 10^5 cm^{-1}$, $\ell_1 = 10^5 cm$, $\ell_2 = 10^2 cm$. Evidently larger choices of σ allow one to overcome angular jitter, even when $k \ll Q$.

surement simply by replacing the SBC and HWP in Fig. 4.2 with a dispersive prism, to arrive at Fig. 4.4. If we take the optical axis to be the z-direction and the momentum shift from the tilted mirror in the x-direction, we choose the prism to cause a small frequency dependent momentum shift $k_p(\omega)$ in the y-direction. We will abbreviate this as k_p in what follows for the sake of notational simplicity. If our detector only measures shifts in the x-direction, the effect of the prism will be essentially the same as introducing a frequency dependent phase.

Before the detector, the two-dimensional distribution becomes

$$p(x, y | k_p) = \frac{1}{p_f} \sin^2 (kx + k_p y) \left| \langle x, y | \varphi \rangle \right|^2 , \tag{4.21}$$

where $\left| \langle x, y | \varphi \rangle \right|^2$ is an uncorrelated two-dimensional Gaussian distribution with zero mean and the same variance in both directions equal to σ^2 (*i.e.,* the
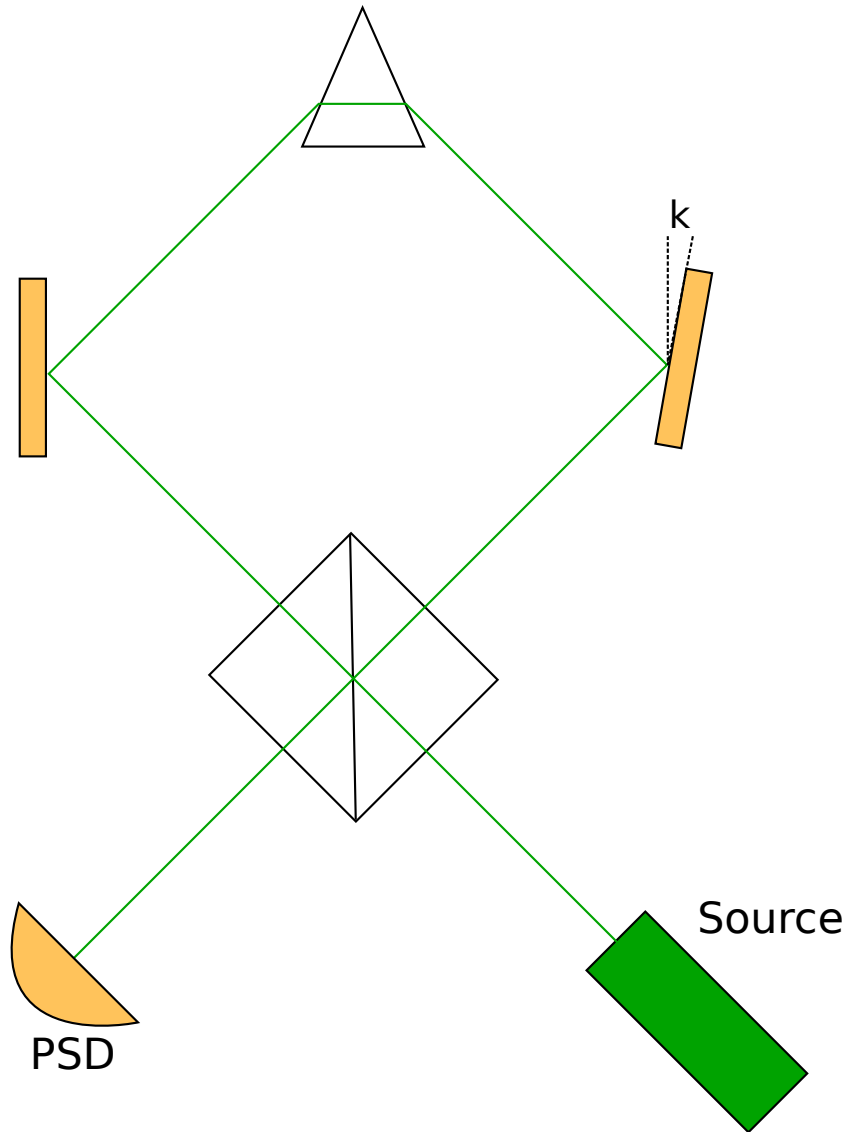
Figure 4.4: An inverse weak value-based frequency measurement scheme as discussed in Sec. 4.5. Inside the Sagnac interferometer we introduce a path-dependent momentum shift k with a titled mirror to set the postselection probability. A prism introduces a frequency-dependent momentum shift in a direction orthogonal to the titled mirror.

covariance matrix for the initial state is proportional to the identity). To first order in k_p the postselection probability is unchanged from Eq. (4.6). If we are only measuring in the x-direction, the distribution becomes

$$
\begin{aligned}
p(x|k_p) &= \int_{-\infty}^{\infty} \mathrm{d}y \; p(x, y|k_p), \\
&\approx \frac{1}{p_f} \sin^2 (kx + k_p\sigma) \left| \langle x|\varphi \rangle \right|^2 .
\end{aligned}
\tag{4.22}
$$

Hence, the above analysis for phase measurements hold for this type of frequency measurement with the change $\phi \to 2k_p\sigma$, where the factor of two is simply due to our definitions of the relevant interaction Hamiltonians in Sec. 4.1.

The Fisher information per detected event for estimation of the parameter k_p to lowest nonvanishing order in $k\sigma$ is equal to

$$
\mathcal{I}(k_p) \approx \frac{2\sigma^2}{k^2\sigma^2},
\tag{4.23}
$$

where the expression is no longer dimensionless since the parameter being estimated is not dimensionless. The total Fisher information for ν events is then

$$
\begin{aligned}
\mathcal{I}_\nu(k_p) &= p_f \nu \mathcal{I}(k_p), \\
&\approx 2\sigma^2 \nu.
\end{aligned}
\tag{4.24}
$$

To translate this to a minimum frequency resolution we note that

$$k_p \approx k_0 \theta,$$
$$\approx k_0 \frac{\partial \theta}{\partial \omega} \Delta \omega, \tag{4.25}$$

where k_0 is the wavenumber, $\Delta \omega$ is the frequency shift relative to some reference frequency, and θ is the angle of deflection relative to the optical axis. Since the estimator for $\Delta \omega$ differs from the one for k_p by only a factor, it follows directly the Fisher information is given by

$$\mathcal{I}_\nu(\Delta \omega) = \left[k_0 \frac{\partial \theta}{\partial \omega} \right]^2 \mathcal{I}_\nu(k_p),$$
$$= 2 \left[k_0 \sigma \frac{\partial \theta}{\partial \omega} \right]^2 \nu. \tag{4.26}$$

Hence, the resolution of a precision frequency measurement can be made large for large σ and for large $\partial \theta / \partial \omega$.

4.6 Remarks

As in the weak value regime, precision phase measurements in the inverse weak value regime are robust to several types of technical noise. In particular, the type of measurements considered here can approach the full noiseless Fisher information under additive Gaussian technical noise and angular jitter noise if one is free to increase the variance σ^2 of the initial state arbitrarily.

Our results do not crucially rely on a specific experimental geometry, and will generalize well to a wide range of weak value based phase measurements.

Chapter 5

Squeezed light enhanced weak values

The precision of a beam deflection measurement is limited both by fundamental sources of uncertainty derived from quantum mechanical considerations, as well as unavoidable sources of technical noise present in any laboratory setting.

The possibility of reducing the quantum uncertainty of an observable using squeezed states of light was recognized in the early 80s. In a particularly significant paper along these lines, Caves [Cav81] showed that the precision of interferometric phase measurements could be improved by using squeezed states to reduce quantum vacuum fluctuations. Precision phase measurements beyond the so-called "shot-noise limit" or "standard quantum limit" were experimentally realized later in the decade [XWK87] with the

generation of sufficiently strong squeezed light using optical parametric amplifiers (OPAs). This technique was extended both theoretically and experimentally later to improve the sensitivity of beam deflection measurements [TAB+02, TGB+03].

Using both squeezed light and weak value amplification in concert allows us to simultaneously reduce the quantum uncertainty of the transverse position observable and technical noise encountered in the lab.

5.1 Squeezed Weak Value Measurement

Using the basic configuration shown in Fig. 5.1, we may combine the output of the dark port of a Rochester setup weak-value measurement with squeezed vacuum states. Coherent light prepared in a Gaussian mode from the laser is deflected by the piezo-mirror in the Sagnac interferometer. The squeezed light source creates a squeezed vacuum prepared in a "flipped" Gaussian mode described below. Though the coherent mode is deflected and the squeezed mode is not, this will only trivially affect the measurement provided the beam deflection is small compared to the beam width.

The effect of squeezed light on split-detected beam deflection measurements has been theoretically examined in detail by Barnett et al. in Ref. [BFM03]. In that paper, they consider both coherent and squeezed vacuum states of light.

The coherent state $|\alpha\rangle$ can be written with the help of the displacement
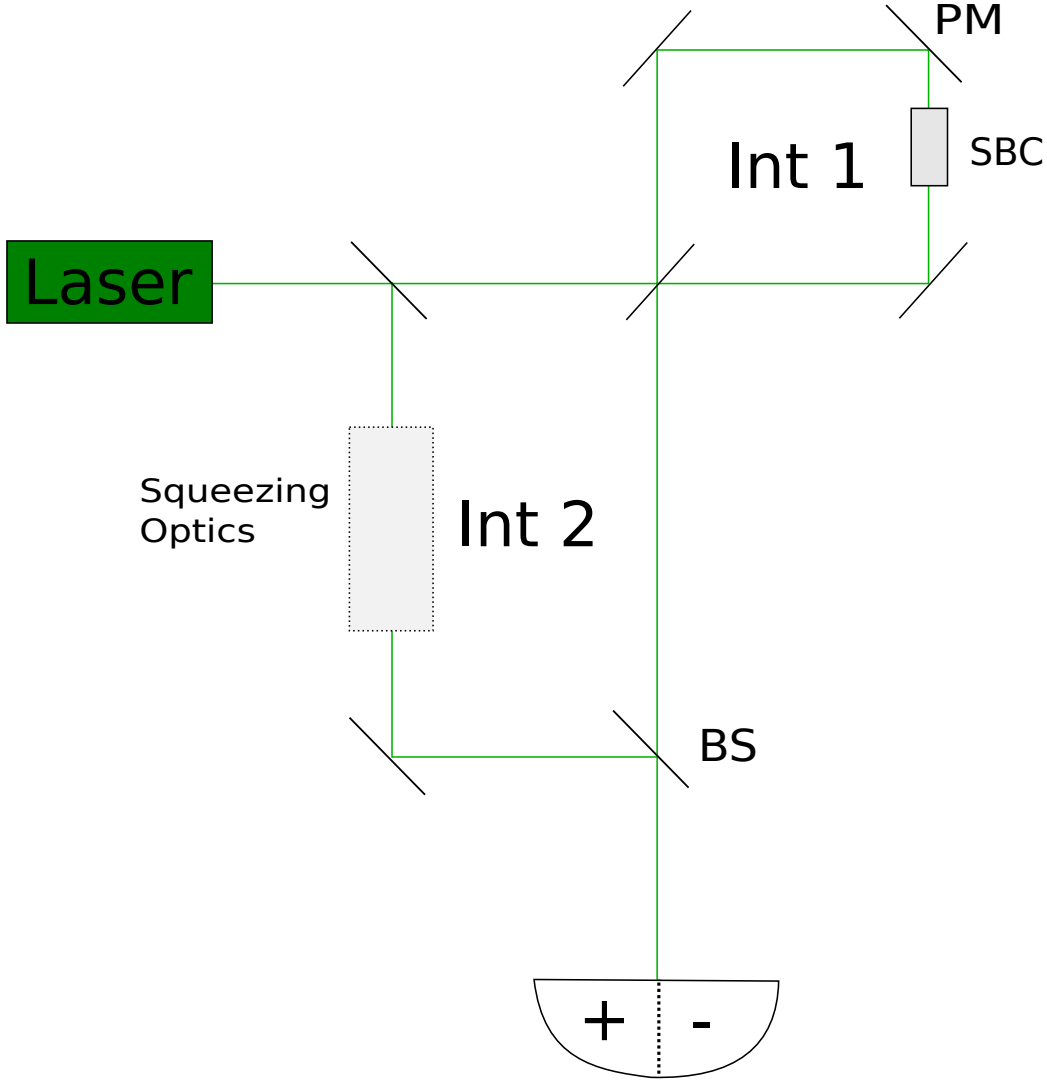
Figure 5.1: Basic setup for reducing the quantum noise of a weak value amplification measurement using squeezed light. Coherent light from the laser is split into two beams- one travels through the Rochester setup WV interferometer (Int 1) with deflection-amplified photons leaving the dark port, the other travels through "squeezing optics" which create a flipped mode in a squeezed vacuum state as in [TAB$^+$02, TGB$^+$03] in Int 2. The beams are combined using a highly reflective beamsplitter (BS) ($R > 90\%$) and are then incident on a split detector.

operator $\hat{\boldsymbol{D}}(\alpha) = \exp(\alpha\hat{\boldsymbol{a}}^\dagger - \alpha^*\hat{\boldsymbol{a}})$ as

$$|\alpha\rangle = \hat{\boldsymbol{D}}(\alpha)|0\rangle, \tag{5.1}$$

where $|0\rangle$ is the vacuum state. The complex number α satisfies the eigenvalue equation $\hat{\boldsymbol{a}}|\alpha\rangle = \alpha|\alpha\rangle$ and hence the average number of photons in a coherent state is given by $\langle\alpha|\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{a}}|\alpha\rangle = |\alpha|^2$.

The squeezed vacuum state $|\zeta\rangle$ can be obtained in a similar manner by using the "squeeze operator" $\hat{\boldsymbol{S}}(\zeta) = \exp(\frac{1}{2}(\zeta^*\hat{\boldsymbol{a}}^2 - \zeta\hat{\boldsymbol{a}}^{\dagger\,2}))$,

$$|\zeta\rangle = \hat{\boldsymbol{S}}(\zeta)|0\rangle. \tag{5.2}$$

Writing the complex squeeze parameter as $\zeta = re^{i\nu}$, the average number of photons in a squeezed vacuum state is given by $\langle\zeta|\hat{\boldsymbol{a}}^\dagger\hat{\boldsymbol{a}}|\zeta\rangle = \sinh^2 r$.

Using the Gaussian and "flipped" modes given by

$$u_0(x) = \frac{1}{\sqrt{\pi^{1/2}\sigma}}e^{-x^2/2\sigma^2},$$
$$u_1(x) = \frac{1}{\sqrt{\pi^{1/2}\sigma}}e^{-x^2/2\sigma^2}\mathrm{sign}(x), \tag{5.3}$$

it is possible to beat the standard quantum limit by using squeezed light to reduce noise due to vacuum fluctuations. If the u_0 mode is prepared in a coherent state and the u_1 mode is prepared as a squeezed vacuum then the

variance in photon counts between the two sides of a split detector is

$$\mathcal{N}^2 = \sinh^2 r + |\alpha|^2 [\cosh 2r$$

$$- 2\cos(2\theta - \nu)\sinh r \cosh r], \qquad (5.4)$$

where $\theta = \arg \alpha$. The variance is a function of relative phase between the two modes, and takes on its minimum value for $\theta = \nu/2$. The split-detected signal is given by

$$\mathcal{S} = \frac{2|\alpha|^2 d}{\pi^{1/2}\sigma}, \qquad (5.5)$$

where d is the deflection distance.

For the optimal relative phase between the coherent and squeezed beams the SNR of a displacement measurement is then

$$\mathrm{SNR} = \frac{2|\alpha|^2 d}{\pi^{1/2}\sigma(|\alpha|^2 e^{-2r} + \sinh^2 r)^{1/2}}. \qquad (5.6)$$

From Eq. (5.6) we see that for moderate squeezing $(\sinh^2 r \ll |\alpha|^2)$

$$\mathrm{SNR} = \frac{2d\sqrt{|\alpha|^2}}{\pi^{1/2}\sigma e^{-r}}, \qquad (5.7)$$

which represents an improvement in SNR by a factor of e^r over the standard quantum limit. Minimizing the denominator of Eq. (5.6) reveals that maximal squeezing can be achieved with the choice $r \approx \ln(4|\alpha|^2)/4$, or equiv-

alently, $\sinh^2 r \approx \sqrt{|\alpha|^2}/2$. For maximal squeezing, the SNR scales as

$$\text{SNR} = \frac{2\sqrt{2}d(|\alpha|^2)^{3/4}}{\pi^{1/2}}, \tag{5.8}$$

which improves SNR scaling by a factor of $(|\alpha|^2)^{1/4}$.

5.1.1 Dark Port Output

To ensure the analysis from Ref. [BFM03] holds, we must verify that the output of the dark port of the interferometer is a coherent beam in the Gaussian u_0 mode.

For the longitudinal mode, it is straightforward to show that when the appropriate operating point for the SBC is chosen such that $\phi = 0$ corresponds to perfectly bright and dark ports the relevant annihilation operator is given by

$$\hat{\boldsymbol{a}}_{dark} = \sin\left(\frac{\phi}{2}\right)\hat{\boldsymbol{a}}_{in}. \tag{5.9}$$

Hence, the annihilation operators describing the dark port output and WV interferometer input only differ by a factor. This does not modify the form of Eq. (5.1), so a coherent input beam will yield a coherent output beam. Note that taking the adjoint of Eq. (5.9) and relating the number operators for the input and dark port output correctly recovers the post-selection probability $P_{ps} = \sin^2(\phi/2)$.

To verify the Gaussian shape of the dark port transverse mode, we examine the effect of our weak measurement on the beam profile. We consider the "system" state $|\psi\rangle$ which describes the path of photons in the WV interferometer and the "meter" state $|\varphi\rangle$ describing the transverse beam. The overall state is then given by the tensor product $|\Psi\rangle = |\psi\rangle|\varphi\rangle$.

With kets $|\circlearrowright\rangle$ and $|\circlearrowleft\rangle$ describing clockwise and counterclockwise travel in the interferometer, we take the input state to be $|\psi_+\rangle = \frac{1}{\sqrt{2}}\left(|\circlearrowright\rangle + i|\circlearrowleft\rangle\right)$. The system state describing post-selection is the the orthogonal $|\psi_-\rangle = \frac{1}{\sqrt{2}}\left(|\circlearrowright\rangle - i|\circlearrowleft\rangle\right)$. It is also convenient to introduce the which-path operator $\hat{W} = |\circlearrowright\rangle\langle\circlearrowright| - |\circlearrowleft\rangle\langle\circlearrowleft|$.

While traversing the WV interferometer photons will experience two distinct effects, a transverse momentum shift due to the piezo mirror and a phase shift due to the SBC. The momentum shift is described by the unitary operator $\hat{U}_{PM} = e^{-ik\hat{x}\hat{W}}$ and the phase shift is described by the unitary operator $\hat{U}_{SBC} = e^{i\phi\hat{W}/2}$. Post-selected photons are then described by the state $|\Psi_-\rangle = \left(\hat{M}_-|\varphi\rangle\right)|\psi_-\rangle$, with $\hat{M}_- = \langle\psi_-|\hat{U}_{PM}\hat{U}_{SBC}|\psi_+\rangle = i\sin(\phi/2 - k\hat{x})$.

Taking the input of the WV interferometer to be in the zeroth mode of the basis $u_0(x)$, we are able to calculate the dark port profile $\langle x|\varphi\rangle$ as

$$\langle x|\varphi\rangle = \langle x|\hat{M}_-|\varphi_0\rangle,$$
$$= A\sin(\phi/2 - kx)u_0(x), \tag{5.10}$$

where $|\varphi_0\rangle$ is the input state and

$$A = \sqrt{\frac{2}{1 - e^{-k^2\sigma^2}\cos\phi}} \tag{5.11}$$

is a normalization constant. The projection of the WV output state on to the initial state is then

$$\begin{aligned}\langle\varphi_0|\varphi\rangle &= A \int \mathrm{d}x\, u_0(x)^2 \sin(\phi/2 - kx), \\ &= \sqrt{\frac{2e^{-\frac{1}{2}k^2\sigma^2}\sin^2\phi/2}{1 - e^{-k^2\sigma^2}\cos\phi}}, \\ &\approx \sqrt{\frac{2\sin^2\phi/2}{1 - \cos\phi}} = 1.\end{aligned} \tag{5.12}$$

For our purposes a reasonable upper bound on $k\sigma$ is given by $k\sigma < 10^{-4}$ or equivalently, $k^2\sigma^2 < 10^{-8}$, hence our approximation $e^{k^2\sigma^2} \approx 1$ is valid for all parameters of experimental interest. Any deviations from a pure $u_0(x)$ mode due to the WV measurement are therefore negligible.

5.1.2 Signal Losses

In both the coherent and squeezed modes the signal will be diminished due to the mixing beamsplitter. If we take the beamsplitter to transmit with a probability T and reflect with a probability R, we may estimate the effect on our signals.

The squeezed mode consists of pairs of correlated photons which must

both be reflected at the beamsplitter in order to contribute to the noise reduction. The probability that both photons from a given pair are reflected is given by R^2. There is also a small probability of single photons from a formerly entangled pair reaching the detector given by $1 - R^2 - T^2 = 2RT$. This is a small effect and will be ignored in following SNR calculations.

The coherent mode has no such correlations, and hence is simply reduced in amplitude by a factor T. If the number of photons sent into the interferometer is N_{in}, then the number of coherent mode photons reaching the detector is given by $N_{out} = P_{ps}TN_{in}$.

Using the experiment in Ref. [DSJH09] for an estimate, approximately 10^{-5} W of power was post-selected using a few mW of input power. Including further attenuation due to the mixing beamsplitter we would expect our signal incident on the detector to be in the range of $10^{-6} - 10^{-7}$ W. Taking a low end estimate of the detector responsivity (i.e., the current generated per unit beam power) of .1 A/W, and a high end estimate of detector dark current (i.e. the current due to the bias voltage across a photodiode) of 10 nA, the beam will be dim enough for the photocurrent it generates to be comparable to the dark current.

5.2 SNR Gains

Only taking quantum fluctuations into account and ignoring constants, the SNR of a split-detected deflection measurement at the standard quantum
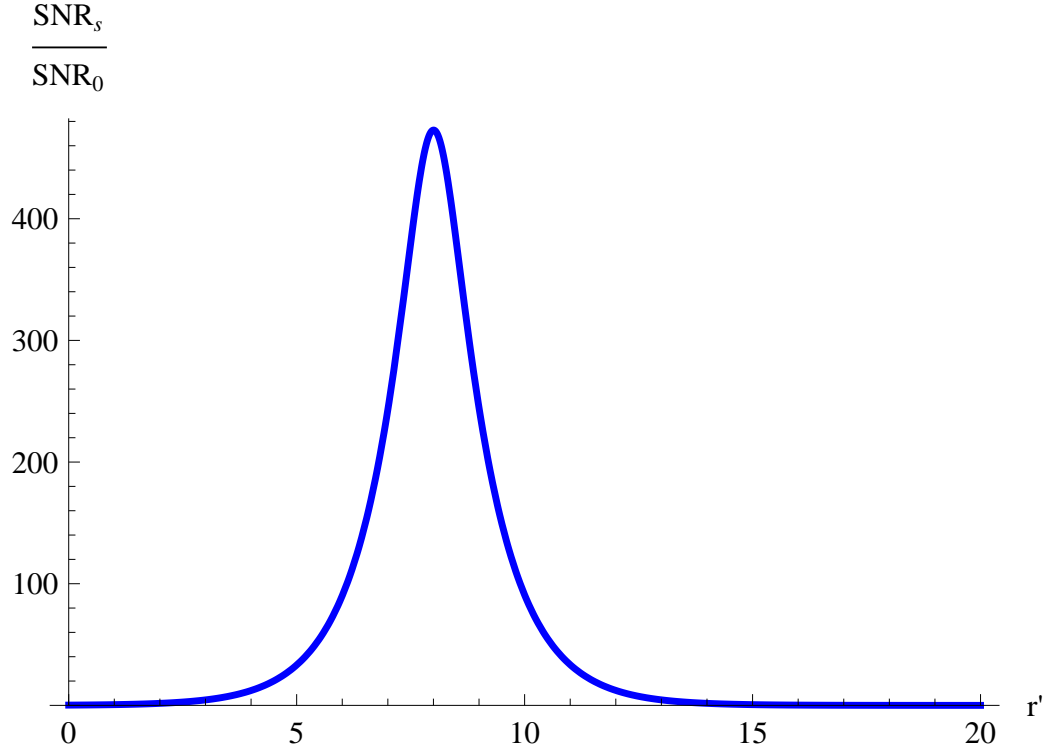
Figure 5.2: SNR gain as a function of squeezing parameter amplitude r', taking the coherent mode to be a 100 μW signal at 780 nm similar to Ref. [DSJH09], and a mixing BS with $T = .05$. The squeezed source power required to maximize SNR is approximately $5\,\mu$W, which is substantially less than the squeezed signals generated in [VCH$^+$05, VMC$^+$08].

limit can be written as [BFM03]

$$\text{SNR}_0 \propto \sqrt{|\alpha|^2 d}. \tag{5.13}$$

Including the squeezed flipped mode, and taking into account the mixing beamsplitter the SNR scaling is given by

$$\text{SNR}_s \propto \frac{T|\alpha|^2 d}{(\sinh^2 r' + T|\alpha|^2 e^{-2r'})^{1/2}}, \tag{5.14}$$

where r' is the modified squeezing amplitude chosen so that $\sinh^2 r' = R^2 \sinh^2 r$. The gain in SNR caused by squeezing is then given by

$$\frac{\text{SNR}_s}{\text{SNR}_0} = \frac{T\sqrt{|\alpha|^2}}{(\sinh^2 r' + T|\alpha|^2 e^{-2r'})^{1/2}}. \tag{5.15}$$

In the limit of moderate squeezing ($\sinh^2 r' \ll T|\alpha|^2$), this yields the condition that

$$r' > -\frac{\ln T}{2} \approx 1.5, \tag{5.16}$$

$$\sinh^2 r' > 4.5, \tag{5.17}$$

for squeezing to overcome the losses in signal due to the mixing beamsplitter. Curves showing SNR gains for arbitrary squeezing strengths are shown in Fig. 5.2.

5.2.1   Technical Noise Effects

As in Ref. [SDJH09], we take the measured signal $x = d + \eta_i + \xi(t)$ where d is the deflection distance, η_i is the quantum uncertainty per photon, and $\xi(t)$ is the technical noise contribution. The technical noise is modeled as a zero-mean white noise process such that $\langle \xi(t)\xi(t') \rangle = S_\xi^2 \delta(t - t')$. The variance of the time-averaged signal $\bar{x}$ is $(\Delta \bar{x})^2 = (1/N^2) \sum_{i,j=1}^{N} \langle \eta_i \eta_j \rangle + (1/t^2) \int_0^\infty \mathrm{d}t'\mathrm{d}t'' \langle \xi(t')\xi(t'') \rangle$. For a coherent beam with no squeezing, the variance per photon due to quantum uncertainty is $\langle \eta_i \eta_j \rangle = \sigma^2 \delta_{ij}$.

The measured signal of a weak value amplified deflection is then

$$\langle \bar{x} \rangle = \mathcal{A}d \pm \sqrt{\frac{\sigma^2}{P_{ps}|\alpha|^2 t} + \frac{S_\xi^2}{t}},$$

$$= \frac{1}{\sqrt{P_{ps}}} \left(\mathcal{A}\sqrt{P_{ps}}d \pm \sqrt{\frac{\sigma^2}{|\alpha|^2 t} + \frac{P_{ps}S_\xi^2}{t}} \right), \qquad (5.18)$$

with the amplification factor $\mathcal{A}$ and the post-selection probability P_{ps} given by

$$\mathcal{A} = \frac{2k_0 \sigma^2}{\ell} \cot(\phi/2),$$

$$P_{ps} = \sin^2(\phi/2). \qquad (5.19)$$

Hence, without squeezing the technical noise becomes trivial compared to the quantum uncertainty at very low post-selection probabilities.

With squeezing, the shot noise variance is reduced by a factor of e^{-2r}.

According to Eq. (5.6), the choice $r = \ln(|\alpha|^2)/4$ maximizes squeezing which in turn yields $e^{-2r} = 1/\sqrt{|\alpha|^2}$. This modifies Eq. (5.18) as

$$\langle \bar{x} \rangle = \frac{1}{\sqrt{P_{ps}}} \left(\mathcal{A}\sqrt{P_{ps}}d \pm \sqrt{\frac{\sigma^2}{\sqrt{P_{ps}}(|\alpha|^2)^{3/2}t} + \frac{P_{ps}S_\xi^2}{t}} \right). \tag{5.20}$$

Evidently, there is a trade-off between minimizing technical noise with weak value amplification and minimizing quantum uncertainty with squeezing. As the post-selection probability is decreased the amount of squeezing we are able to perform is also reduced. Minimizing the uncertainty in $\langle \bar{x} \rangle$ with respect to P_{ps} we find the most precise measurement possible can be made with a post-selection probability given by

$$P_{ps} = \left(\frac{\sigma^2}{2(|\alpha|^2)^{3/2}S_\xi^2} \right)^{2/3}. \tag{5.21}$$

This expression exceeds the maximum possible value of $P_{ps} = 1$ for sufficiently low technical noise power. Physically, this reflects the fact that if the technical noise is less than the quantum uncertainty even with maximal squeezing, a more precise measurement can be made without post-selection than with it.

5.3 Remarks

By combining the output of a weak value amplified measurement using the Rochester setup with squeezed vacuum states, we are able to improve the

sensitivity of a beam deflection measurement beyond a level which could be achieved by either technique alone. The intrinsic uncertainty in the transverse position of the beam is reduced through squeezing, while technical noise is reduced through the post-selection process with weak value amplification.

Interestingly, while we have specifically studied the effect of squeezing the output of a Rochester setup weak value measurement here, the general concept may be applied to a wider class of weak measurement schemes. As long as the quantum state of the output of an optical measurement is known, in principle one can use squeezed light to reduce noise due to interactions with the vacuum.

Chapter 6

Entangled biphoton displacement measurements

In this chapter we consider an alternative scheme for using quantum correlations to enhance the measurement of a spatial deflection or displacement. By using spatially entangled biphoton pairs, such as those created using parametric downconversion [SH15, LE04, FMV09], we show it is possible to substantially reduce noise while using a simple experimental setup. Namely, we find that correlations between photons in each pair allow for the average position of both photons to be determined more precisely than the individual positions of each photon. With a judicious choice of measurement scheme, this allows for an enhancement in the determination of a small displacement parameter. We note that fundamental measurement limits and the validity of Heisenberg's uncertainty relation using schemes with entangled probes

and detectors were considered previously by Di Lorenzo [DL13] and Bullock and Busch [BB14]. They find that it is possible to improve the sensitivity of a measurement using entanglement, which is in agreement with our analysis of a specific measurement scheme here. It is also interesting to consider that while continuous momentum and position correlations as quantum resources are relatively new in metrology, they have been at the center of our understanding of entanglement from a very early point in the development of quantum theory. In particular, they were considered in the seminal work by Einstein, Podolsky, and Rosen [EPR35], and entangled pairs of the form we consider in this chapter may properly be thought of as "EPR pairs".

The chapter is organized as follows: in Sec. 6.1 we discuss the entangled biphoton quantum state, and correlated position-position probability distribution which is used throughout the chapter, along with a brief description of a possible experimental implementation. Generic enhancements over schemes using coherent states with the same mode are also discussed [1]. In Sec. 6.1.1 and Sec. 6.1.2 we extend the treatment to cover the reasonable cases of split-detection and N-pixel detectors, respectively. In particular, these sections demonstrate the enhancement over uncorrelated photons and the robustness of measurements to pixelation of the detection scheme. Sec. 6.1.3 discusses

---

[1]When comparing entangled and uncorrelated photons in this work, we use the same transverse spatial mode for each case. Namely, we do not consider differences in the minimum waist of a classical beam versus the minimum uncertainty of the average position of a SPDC biphoton pair for a given pump wavelength. This allows us to quantify enhancements due to entanglement in a consistent way, and also avoids certain technical limitations such as saturation or burning of the detector for focused beams of light.
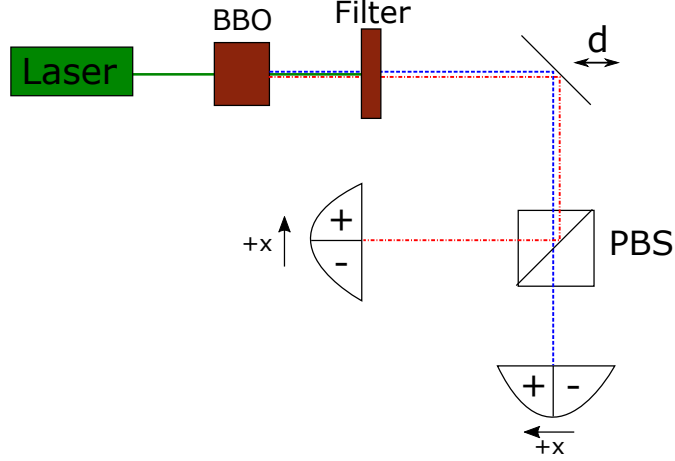
Figure 6.1: Schematic for a biphoton displacement measurement. Photons from a laser source are converted into correlated pairs in a colinear type-II spontaneous parametric down conversion (SPDC) process (*i.e.*, one in which two photons of orthogonal polarization are generated), for example with a BBO crystal. Uncorrelated short wavelength photons (green solid line) are filtered out, while the horizontally polarized (red line) and vertically polarized (blue line) photons pass through to a movable mirror. The shift d of the mirror from the origin displaces the optical beam, and is the small, unknown parameter being measured with this apparatus. Each photon in the pair is detected separately at a position-sensitive detector placed at the output ports of a polarizing beam splitter (PBS). By using coincidences between each detector it is possible to discard spurious events caused by lossy optics or imperfections in the PBS.

the case of very strong spatial correlation, and how this affects the scaling of the measurement resolution with the number of independent events. This also provides insight into the limits of the measurement scheme. In Sec. 6.4 we give our concluding remarks.
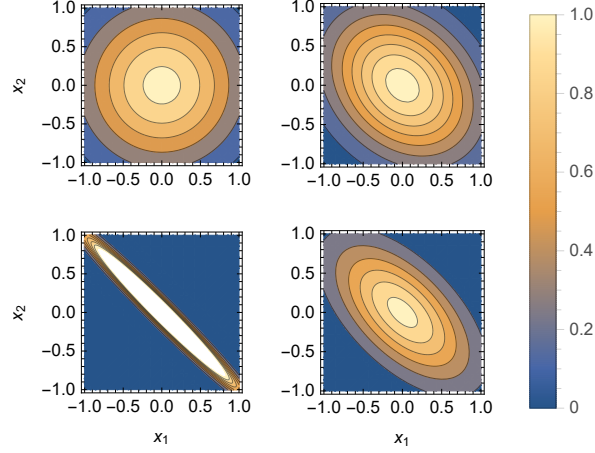
Figure 6.2: The probability distribution (6.2) as a function of x_1 and x_2. Clockwise from the upper left we have increasing correlation with $\epsilon/\sigma = 1$, 0.75, 0.5, and 0.1. As expected, x_1 and x_2 are increasingly anti-correlated with decreasing ϵ.

6.1 Biphoton displacement measurement

As a simple, concrete model of a biphoton state exhibiting spatial correlations useful for displacement measurements, we consider the setup in Fig. 6.1, which uses spontaneous parametric down conversion (SPDC) to generate the desired state. In SPDC, single "pump" photons are converted into two photons (typically referred to as the "signal" and "idler" modes), which are entangled due to constraints set by conservation of energy and conservation of momentum.

The transverse state of biphotons created in SPDC can be approximated

in the position basis as [SH15]

$$\int dx_1 dx_2 |x_1, x_2\rangle\langle x_1, x_2|\psi(d)\rangle =$$
$$\frac{1}{\sqrt{\pi\sigma\epsilon}} \int dx_1 dx_2 \exp\left(-\frac{(x_1 - x_2)^2}{4\sigma^2}\right)$$
$$\times \exp\left(-\frac{(x_1 + x_2 - 2d)^2}{4\epsilon^2}\right) \hat{a}^\dagger(x_1)\hat{a}^\dagger(x_2)|0, 0\rangle, \qquad (6.1)$$

where d is the transverse displacement being measured, and x_1 and x_2 are the transverse position variables for each of the two photons, σ is the pump beam waist, and ϵ is a parameter which describes the spatial correlation between the photons.

The probability distribution follows simply from taking the norm squared of this state,

$$p(x_1, x_2|d) = \frac{1}{\pi\sigma\epsilon} \exp\left(\frac{-(x_1 - x_2)^2}{2\sigma^2}\right) \times$$
$$\exp\left(\frac{-(x_2 + x_1 - 2d)^2}{2\epsilon^2}\right), \qquad (6.2)$$

which is visualized in Fig. 6.2. Note that an entangled pair of photons is necessary here to produce the correlated position distribution (6.2). If one uses some classical resources like coherent states, such a correlated distribution cannot be generated with this particular setup, since the PBS is a linear optical element and the photons at the two outputs are uncorrelated if the input photons are in a coherent state. However, the results in this chapter

depend on this distribution only, rather than the specific optical physics or measurement scheme. Since we will work in this fixed basis throughout this chapter, the results of our analysis applies equally to any quantum or classical system which produces the distribution (6.2), *i.e.,* for a judicious choice of measurement scheme it should be possible to replicate our results here using classically correlated beams.

For this profile, correlations set the constraint that if a photon is measured at position x_1, the probability distribution of the second photon is peaked at the position $x_2 = 2d - x_1$. In other words, the position of each photon is mirrored about the point $x = d$ up to a small uncertainty set by the parameter ϵ. Hence, the smallness of ϵ determines the "strength" of the entanglement.

In this chapter we will use Fisher information [LM01] as the metric used to determine the overall measurement sensitivity for our biphoton distribution in various detector arrangements. From statistics, Fisher information is a measure of how sensitively the distribution depends on the parameter d. The Fisher information per photon pair is simply calculated (for perfect detector resolution) as

$$
\begin{aligned}
\mathcal{I}(d) &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mathrm{d}x_1 \mathrm{d}x_2 \; p(x_1, x_2|d) \left(\partial_d \ln p(x_1, x_2|d) \right)^2 , \\
&= \frac{4}{\epsilon^2}
\end{aligned}
\tag{6.3}
$$

Note that the overall width of the input Gaussian profile σ does not appear in

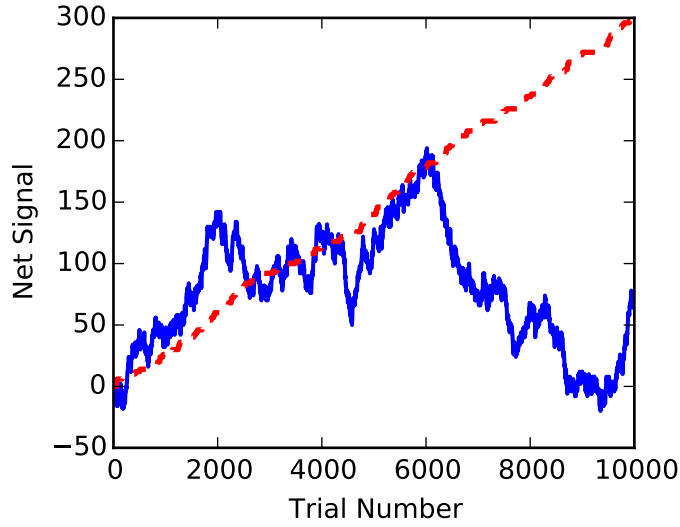Figure 6.3: Net signal for biphotons after 10,000 independent events for split-detection. The net signal here is analogous to the net steps from the origin in a random walk, with the probabilities of two steps backward, two steps forward, and zero steps per trial given by Eq. (6.12). The solid blue signal represents uncorrelated photon pairs, while the red dashed line represents entangled photons with $\epsilon/\sigma = 0.01$.

this result. Comparing with (6.2) we see this amounts to a statement that the sensitivity of our measurement is not determined by the individual photon positions but rather the sum of their positions. This is in stark contrast to the case of uncorrelated photons, where only the mode profile width σ determines the Fisher information, $\mathcal{I}_0(d) = 4/\sigma^2$. The result (6.3) can also be recovered using the quantum Fisher information given by [BCM96]

$$\mathcal{I}_Q(d) = 4 \left(\langle \partial_d \psi | \partial_d \psi \rangle - |\langle \partial_d \psi | \psi \rangle|^2 \right). \tag{6.4}$$

Since the state of biphoton (6.1) has no d-dependent relative phases, it can be proven that the quantum Fisher information $I_Q(d)$ is equal to the classical Fisher information (6.3) in this case (See section 6.3 for the proof).

The Cramèr-Rao lower bound [LM01] defines a relationship between the Fisher information and the variance of a statistical estimator $\hat{d}$,

$$\mathrm{Var}(\hat{d}) \geq \frac{1}{\mathcal{I}(d)}. \tag{6.5}$$

Using the Cramèr-Rao lower bound to determine the minimum variance per biphoton pair and setting the minimum resolvable parameter d_{min} equal to the square root of the variance, for ν independent measurements we have

$$d_{min} \sim \frac{\epsilon}{2\sqrt{\nu}}, \tag{6.6}$$

for an efficient estimator (*i.e.,* an estimator which saturates the Cramèr-Rao

lower bound). Note that in this case each biphoton pair is an independent event, so there are 2ν total photons in a given experimental run. In the limit $\epsilon \to \sigma$, Eq. (6.2) is separable, and the distribution is identical to that of two uncorrelated photons. Hence, any advantage over the case of unentangled probes comes from a distribution with $\epsilon < \sigma$. For a given input beam width σ, the enhancement in Fisher information due to entanglement is then

$$\frac{\mathcal{I}(d)}{\mathcal{I}_0(d)} = \frac{\sigma^2}{\epsilon^2}, \tag{6.7}$$

where the comparison is between entangled biphotons and uncorrelated pairs of photons. For a comparison with single uncorrelated photons one can simply insert a factor of two due to the linearity of Fisher information (We note the uncorrelated position variance of a single photon is $\mathrm{Var}[x] = \sigma^2/2$ in this notation). Sample split-detection signals highlighting the advantage of entanglement are shown in Fig. 6.3.

The exact relation of σ and ϵ can be further controlled by the placement of lenses before the polarizing beam splitter. Quantum mechanics dictates that the small variance of $x_1 + x_2$ implies that the transverse wavenumber sum of the photons, $k_{x,1} + k_{x,2}$, has a large variance, since they obey an uncertainty principle, $\sigma_{x_+} \sigma_{k_+} \geq 1/4$ [SH15]. The smallest variance ϵ of $x_1 + x_2$ is determined by the pump beam width at the nonlinear crystal, and the smallest variance σ of $x_1 - x_2$ is given by $\sigma_{\min} = \sqrt{9w\lambda_p/10\pi}$, where w is the width of the crystal, and λ_p is the wavelength of the pump beam [SH15].

With the exception of results from expansions requiring $\epsilon/\sigma \ll 1$, one can simply replace ϵ with σ in expressions throughout this work for generalization to uncorrelated pairs of photons. Eq. (6.7) also provides a useful way to compare resources between the entangled and classical experiments. Because Fisher information scales linearly with independent events, if we can achieve a given measurement precision with ν entangled photons, the classical equivalent will require $2\nu(\sigma^2/\epsilon^2)$ independent photons to achieve the same precision.

The average position of the two photons $\hat{d} = (x_1 + x_2)/2$ is an efficient estimator for the parameter d, which is easily verified by direct calculation of the variance,

$$\mathbb{E}[\hat{d}] = d, \tag{6.8}$$

$$\mathbb{E}[\hat{d}^2] = \frac{\epsilon^2}{4} + d^2. \tag{6.9}$$

Hence, the variance per photon pair is $\epsilon^2/4$ and saturates the Cramèr-Rao lower bound. Interestingly, if one could reduce the value of ϵ to arbitrarily low values, d_{min} would become arbitrarily small for even a single biphoton pair. In the specific case of SPDC production, uncertainty in the birth zone of pairs leads to some minimum ϵ.

We can also understand the effect of the detection scheme by considering the marginal probability distribution obtained by integrating over either x_1

or x_2 in Eq. (6.2),

$$p(x|d) = \sqrt{\frac{2}{\pi(\epsilon^2 + \sigma^2)}} \exp\left(\frac{-2(x-d)^2}{\epsilon^2 + \sigma^2}\right), \qquad (6.10)$$

where subscripts have been dropped due to the symmetry of the distribution (6.2). The Fisher information for the marginal distribution of each single photon if its twin is not measured is then

$$\mathcal{I}_\mathrm{m}(d) = \frac{4}{\epsilon^2 + \sigma^2}, \qquad (6.11)$$

where the subscript m in $\mathcal{I}_\mathrm{m}(d)$ indicates that it is the Fisher information of marginal distribution.

Unlike Eq. (6.3), the information is bounded in the limit $\epsilon \to 0$, and in the limit $\epsilon \to \sigma$ we recover the same result as for uncorrelated photons drawn from the full distribution (6.2). Clearly the enhanced measurement sensitivity is not due only to the correlation between photons, but also the ability to detect the correlations for each event.

6.1.1 Split-detection

The preceding results provide insight into the maximum achievable precision for a detector with perfect (*i.e.,* continuous) position resolution. Here we show that biphoton correlations can provide a benefit with a relatively simple split-detection scheme as well. In a split-detection experiment one creates a

detector out of two pixels and then uses the difference in counts between the two pixels to indicate the magnitude and direction of an optical beam shift. A proposed setup is given in Fig. 1. There, two split detectors are used together with a polarizing beam splitter (PBS), so that only events are counted where there is a coincidence detection event, so photons are simultaneously detected in both split-detectors as coincidences, and other (background) events are discounted. For the presentation in the rest of the chapter, we will discuss the results of a single split detector, which is theoretically the same, but not as technically easy to implement since both photons of a biphoton pair can both land on the same side of the detector. To translate between the setups, both photons landing on the left $(-)$ or right $(+)$ side of the single-detector setup corresponds to the $++$ or $--$ two-detector events, whereas one photon landing left and the other right, corresponds to the two-detector $-+$ or $+-$ events. The corresponding net signal is then the average of the split detector signals. In Sec. 6.1.2 we will extend this discussion to the more general case of a position-sensitive detector with N pixels.

For the case of split-detection with a single detector, we introduce the probabilities $P(-2|d), P(0|d)$, and $P(2|d)$, which are the probabilities of two photons landing on the left half of the detector, one photon on each half of

the detector, and two photons on the right half of the detector, respectively:

$$P(-2|d) = \int_{-\infty}^{0} \int_{-\infty}^{0} dx_1 dx_2 \ p(x_1, x_2|d),$$

$$P(2|d) = \int_{0}^{\infty} \int_{0}^{\infty} dx_1 dx_2 \ p(x_1, x_2|d),$$

$$P(0|d) = 1 - P(-2|d) - P(2|d). \tag{6.12}$$

We ignore gaps between pixels throughout this work. These integrals do not have closed form solutions, so we consider the limit $d \ll \epsilon$, where the probabilities can be expressed as

$$P(\pm 2|d) \approx \frac{1}{4} + \frac{1}{2\pi} \arctan\left(\frac{\epsilon}{2\sigma} - \frac{\sigma}{2\epsilon}\right) \pm d\sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}},$$

$$P(0|d) \approx \frac{1}{2} - \frac{1}{\pi} \arctan\left(\frac{\epsilon}{2\sigma} - \frac{\sigma}{2\epsilon}\right). \tag{6.13}$$

The resulting Fisher information for this discrete distribution is

$$\mathcal{I}(d) \approx \frac{16}{(\epsilon^2 + \sigma^2)\left(\pi + 2\arctan\left(\frac{\epsilon}{2\sigma} - \frac{\sigma}{2\epsilon}\right)\right)}. \tag{6.14}$$

Now, let us compute the correlation in the biphotons. A typical measure for the correlation between two variables X and Y is the correlation coefficient defined as

$$\xi = \frac{\text{Cov}[X, Y]}{\sqrt{\text{Var}[X]\text{Var}[Y]}}. \tag{6.15}$$

From the joint distribution of a biphoton pair, Eq. (6.2), it can be derived that the correlation coefficient between the positions of two entangled photons is

$$\xi = \frac{\epsilon^2 - \sigma^2}{\epsilon^2 + \sigma^2}.$$
(6.16)

This yields an explicit relation between the Fisher information $\mathcal{I}(d)$ and the correlation coefficient ξ:

$$\mathcal{I}(d) \approx \frac{16}{\left(\epsilon^2 + \sigma^2\right)\left(\pi + 2\arcsin\xi\right)}.$$
(6.17)

By setting $\xi = 0$ (*i.e.*, $\epsilon = \sigma$) in Eq. (6.16), we find the Fisher information for uncorrelated photon pairs under split-detection,

$$\mathcal{I}_0(d) \approx \frac{16}{\pi\left(\epsilon^2 + \sigma^2\right)} = \frac{8}{\pi\sigma^2},$$
(6.18)

which is smaller than the perfect resolution case (6.3) by a factor of $2/\pi$. The increase (or decrease) in Fisher information due to entanglement is determined by the correlation coefficient:

$$\frac{\mathcal{I}(d)}{\mathcal{I}_0(d)} \approx \frac{\pi}{\pi + 2\arcsin\xi}.$$
(6.19)

This equation characterizes the relation between the boost in Fisher information and the correlation of the entangled photons. It shows that when

$\xi < 0$, the Fisher information $\mathcal{I}(d)$ will be larger than that without correlation $\mathcal{I}_0(d)$. This implies an advantage of using biphotons with negative correlation (*i.e.,* spatially anticorrelated biphotons) in split-detection. It is widely known that correlation can enhance the precision of parameter estimation in the quantum metrology community. Eq. (6.19) can be perceived as a counterpart of that in the split-detection scheme.

Meanwhile, Eq. (6.19) also shows that when $\xi > 0$, $\mathcal{I}(d)$ becomes smaller than $\mathcal{I}_0(d)$ instead. This contrast with the case of negative correlations implies that different types of correlations can have different effects on the Fisher information, and not all types of correlation in the biphotons are favorable to the performance of split-detection, even when the "strength" of the correlation is the same.

An important topic in quantum metrology is identifying and characterizing useful quantum resources for enhancing the sensitivity of measurements. The above analysis shows that the correlation coefficient between photons of a biphoton pair determines the gain in the Fisher information for detecting a beam displacement. This provides an explicit criterion for identifying useful correlation for split-detection with our setup.

In the limiting case $\xi \to -1$ (*i.e.,* $\epsilon \to 0$), the Fisher information can be extremely large, which is in accordance with Fig. 6.4, while in the limiting case $\xi \to 1$ (*i.e.,* $\sigma \to 0$), the Fisher information decreases to only half of that of uncorrelated photons. Note that in the limit $\epsilon \to 0$, we need $d \to 0$ as well, otherwise the assumption $d \ll \epsilon$ will be violated.

We now investigate the maximum likelihood estimator for split-detection. To begin, we introduce the binary random values X_1 and X_2 which can take on the value 0 or 1, and are mutually exclusive (*i.e.*, only one of them can take on the value 1 for a given event). X_1 represents the outcome where two photons land on the positive half of the detector and X_2 represents both landing on the negative half. The relevant joint probabilities reduce to quantities we have already calculated above, namely $P(X_1 = 1, X_2 = 0|d) = P(2|d)$ and $P(X_1 = 0, X_2 = 1|d) = P(-2|d)$ so we are able to directly calculate the variance of the scaled split-detection estimator

$$\hat{d} = \sqrt{\frac{\pi(\sigma^2 + \epsilon^2)}{8}}\,(X_1 - X_2)\,. \tag{6.20}$$

Computing the variance yields precisely the inverse of Eq. (6.14) and so saturates the Cramèr-Rao lower bound. Hence, the split-detection estimator is efficient for measuring the displacement of biphoton pairs.

6.1.2 N-pixel detector

For the case of a position-sensing detector with N pixels, we may proceed similarly to the case of split-detection above. We will slightly change notation for convenience, with P_{ij} being the probability of one photon each landing on pixels i and pixel j (with the case of $i = j$ being the case of both photons on the same pixel), and the dependence on the parameter d is implied, as opposed to the explicit conditional probability notation used above. We note

that similar calculations have been performed for the cases of weak value amplification and uncorrelated Gaussian beams [KG14, KM15]. If the pixels have a width Δ, the probabilities may be calculated as

$$
\begin{aligned}
P_{i,j\neq i} &= \int_{(i-1)\Delta}^{i\Delta} \mathrm{d}x_1 \int_{(j-1)\Delta}^{j\Delta} \mathrm{d}x_2 p(x_1, x_2|d) \\
&+ \int_{(i-1)\Delta}^{i\Delta} \mathrm{d}x_2 \int_{(j-1)\Delta}^{j\Delta} \mathrm{d}x_1 p(x_1, x_2|d),
\end{aligned} \tag{6.21}
$$

where the case of $i = j$ is equal to the first term alone. If we label the bins from 1 to N, the joint probability distribution for a single event can be written as

$$
p(X_{11}, X_{12}, ...|d) = P_{11}^{X_{11}} P_{12}^{X_{12}}..., \tag{6.22}
$$

where the random variables X_{ij} are again binary with mutually exclusive outcomes for a single event. The indices (i, j) are also subject to the constraint $i \leq j$ to avoid double-counting outcomes. After some calculations analogous to the split-detection case, one finds the Fisher information per biphoton event is equal to

$$
\frac{\mathcal{I}(d)}{\nu} = \sum_{i=1}^{N} \sum_{j \geq i}^{N} \frac{1}{P_{ij}} (\partial_d P_{ij})^2. \tag{6.23}
$$

Results for the Fisher information due to our distribution incident on a ten-pixel and fifty-pixel detector compared to split-detection (two pixels) and to
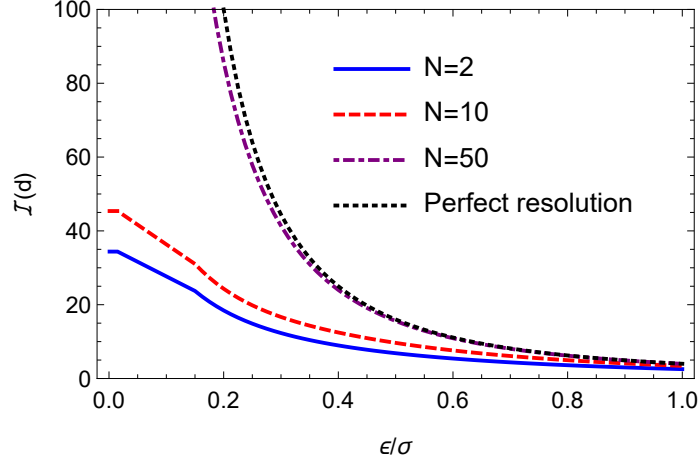
Figure 6.4: A comparison of the Fisher information (in arbitrary units) due to biphoton pairs incident on detectors with different numbers of pixels N, using the numerical value $\sigma = 1$ and $d = 0.05$. For each curve, the information for uncorrelated photons is given by the value of the curve for $\epsilon/\sigma = 1$. The dotted black line represents perfect spatial resolution calculated by Eq. (6.3), the solid blue curve represents split-detection ($N = 2$), the red dashed curve represents a ten-pixel detector, and the case of fifty pixels is given by the purple dot-dashed curve.

perfect spatial resolution are shown in Fig. 6.4. For these computations we use a detector width of 10σ, and hence a pixel width of $10\sigma/N$, where N is the number of pixels. In general, it is possible to improve the resolution by using non-homogeneous pixel widths, but for simplicity we will not treat that case here.

6.1.3   Scaling with independent events

From the above analysis (*e.g.,* Eqs. (6.3)-(6.6)), it is clear that for nonzero ϵ the minimum resolvable parameter scales as $\nu^{-1/2}$ for ν independent events

(*i.e.,* the usual shot-noise limit), with the enhancement appearing as a prefactor, as expected. However, taking the above treatment for the split-detection in the limit $\epsilon \to 0$ yields interesting results (keeping in mind that ϵ does have a minimum value it can take (see the discussion following Eq. (6.7))). For simplicity we rewrite Eq. (6.2) as

$$p(x_1, x_2|d) = \sqrt{\frac{2}{\pi\sigma^2}} \exp\left(-\frac{(x_1 - x_2)^2}{2\sigma^2}\right) \delta(x_1 + x_2 - 2d). \qquad (6.24)$$

The Fisher information for this probability distribution is infinite, implying that the statistical noise is completely suppressed and the estimation process becomes deterministic for the case of perfect spatial resolution, in agreement with Eq. (6.3). One can verify this by noting that the variance of the estimator $\hat{d} = (x_1 + x_2)/2$ is exactly zero. For the case of split-detection, we can exactly solve the integrals in Eq. (6.12). Without loss of generality, we assume $d > 0$;

$$P(2|d) = \mathrm{erf}\left(\frac{\sqrt{2}d}{\sigma}\right),$$

$$P(-2|d) = 0,$$

$$P(0|d) = 1 - P(2|d). \qquad (6.25)$$

We note that these probabilities form a Bernoulli distribution (see, *e.g.,* [Cah13, LM01]) with $P(2|d)$ as the probability of a "success" and $1 - P(2|d)$ as the probability of a "failure". It follows immediately from the properties of a

Bernoulli experiment with ν total trials for very small d:

$$\mathcal{I}(d) \approx \sqrt{\frac{8}{\pi\sigma^2}} \cdot \frac{\nu}{d}. \tag{6.26}$$

Interestingly, the information increases as the parameter d becomes smaller, and diverges as $d \to 0$. This can be understood intuitively from Eq. (6.25), noting the outcome with one photon incident on each half of the detector is certain for the case of $d = 0$. Therefore, events where two photons land on the same side of the detector give a great deal of information about the parameter's value being different from 0.

Eq. (6.26) clarifies the boost in measurement precision of a small parameter due to extremely strong correlations. Comparing the case of uncorrelated photons under split detection (6.18), we note that the enhancement from entanglement can become arbitrarily large for an arbitrarily small parameter d.

Another interesting enhancement to the measurement is the increase of the resolution. The resolution is the minimal resolvable parameter by the measurement with a given signal-to-noise ratio (SNR) $\mathcal{R}$. Since the Fisher information is the inverse of the minimum variance of the estimate, the resolution is related to the Fisher information via the SNR,

$$\mathcal{R} \leq \frac{d_{\min}}{\sqrt{\min \mathrm{Var}[\hat{d}]}} = d_{\min}\sqrt{\mathcal{I}(d_{\min})}, \tag{6.27}$$

where the minimization before $\mathrm{Var}[\hat{d}]$ is over all possible unbiased estimators, and the second equality results from the definition of Fisher information. A standard choice for the minimum SNR required to resolve a parameter is unity. Hence, by substituting Eq. (6.26) into Eq. (6.27), the minimum resolvable d turns out to be

$$d_{min} \approx \sqrt{\frac{\pi \sigma^2}{8}} \cdot \frac{1}{\nu}. \tag{6.28}$$

Eq. (6.28) implies that the resolution reaches the Heisenberg-limited scaling when $\epsilon \to 0$.

This is an interesting result. As is widely known, the standard quantum limit for the resolution scales as $\nu^{-\frac{1}{2}}$. Typical enhancements over this scaling require large numbers of entangled quantum resources, *e.g.*, squeezed states of light, in order for the resolution to reach the Heisenberg-limited scaling [BFM03]. This contrasts with our setup, which only requires pairs of entangled photons, so there must be some substantially different mechanism from that of the other schemes for increasing the scaling of the resolution, as we now discuss.

It can be seen from the definition of SNR (6.27) that the key of improving the resolution is to increase the Fisher information. In most cases, the SNR is linear with the minimal resolvable parameter, so a large N entangled photon state is used in order to increase the scaling of the Fisher information with respect to the number of photons, which results in an improved scaling of

the measurement resolution. This is how the resolution in many optical metrology schemes is improved by using quantum resources like squeezed states.

However, in our protocol, when the parameter d is decreased, the Fisher information increases. This changes the scaling of the Fisher information with respect to d, and similarly makes the SNR nonlinear with d. Hence, even though the scaling of the Fisher information with respect to the number of photons is unchanged, the resolution still can be significantly enhanced, and reach the Heisenberg-limited scaling as shown by Eq. (6.28).

The above analysis also implies that the measurement precision, which is characterized by the minimum variance of the estimate, is not always equivalent to the resolution of the parameter to be measured, although they look quite similar and are sometimes used interchangeably to quantify the metrological performance. If the Fisher information is independent of the parameter, these two measures for the measurement are equivalent, for a given SNR; but if the Fisher information has dependence on the parameter, they can be rather different, as they are in the present case. As our analysis of the biphoton displacement scheme for $\epsilon \to 0$ shows, the measurement precision is characterized by the Fisher information (6.26) which scales proportionally to ν, as we would expect for a classical experiment. However, the minimum resolvable parameter (6.28) yields Heisenberg scaling proportional to ν^{-1}. This is fundamentally due to the dependence of the Fisher information on the parameter d.
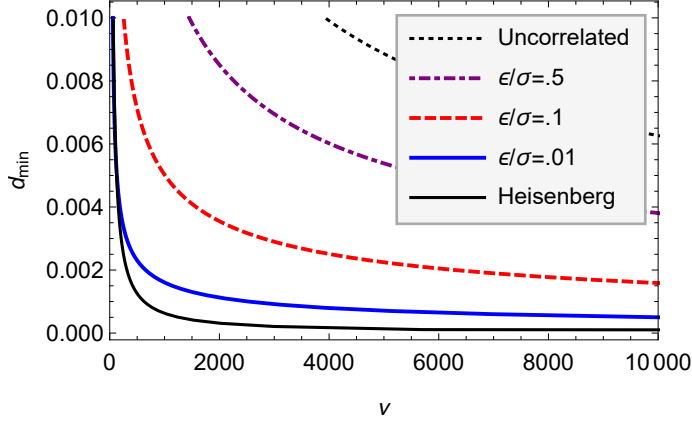
Figure 6.5: The number of events needed to achieve a SNR of 1 under split-detection as a function of the parameter d (in arbitrary units) for different values of ϵ/σ. From top to bottom the curves represent $\epsilon/\sigma = 1$ (*i.e.,* uncorrelated), 0.5, 0.1, 0.05, and 0 (*i.e.,* perfect correlations). As expected from Eq. (6.33), Heisenberg scaling can only be observed for numbers of events much less than σ/ϵ, e.g., the $\epsilon/\sigma = 0.01$ curve only matches the Heisenberg curve for $\nu <\sim 100$.

6.1.4   Small ν scaling

In practice, because one cannot physically create a spatial profile which exactly has one photon hit a detector deterministically at one point conditional on the measurement of another entangled photon (*i.e.,* $\epsilon = 0$), the preceding analysis with delta function correlations is overly optimistic.

As a simple model to consider more realistic imperfections, suppose the probabilities of "success" and "failure" may instead be expanded as

$$P(\text{"success"}|d) \approx \alpha + \beta d,$$

$$P(\text{"failure"}|d) \approx 1 - \alpha - \beta d, \tag{6.29}$$

with constant α and β for some small but nonzero α; α here is effectively a noise term indicating the probability of a success even when the parameter is zero, while β determines the degree to which a change in the parameter correspondingly changes the probability of success. Comparing with Eq. (6.13), we see we can determine α and β simply by matching terms of the same order in d. The Fisher information of this probability distribution with respect to d is

$$\mathcal{I}(d) \approx \frac{\beta^2 \nu}{(\alpha + \beta d)(1 - \alpha - \beta d)}. \tag{6.30}$$

Then the minimum resolvable parameter $d_{\min}$ can be derived from Eq. (6.27) as

$$d_{\min} = \frac{1 - 2\alpha + \sqrt{(1 - 2\alpha)^2 + 4\alpha(1 - \alpha)\nu}}{2\beta\nu}, \tag{6.31}$$

where $\nu \gg 1$ is assumed. As in Sec. 6.1.3, we see the Fisher information is proportional to ν as we would expect for any experiment of ν independent events, but the dependence of the Fisher information on the parameter leads to interesting behavior in the scaling of the minimum resolvable parameter.

The scaling of $d_{\min}$ gradually transitions from ν^{-1} to $\nu^{-1/2}$ as α increases from zero. Therefore, the Heisenberg-limited scaling ν^{-1} is the limiting scaling of the resolution when $\alpha \ll 1$. Hence, the condition to approximately

achieve Heisenberg scaling is

$$\alpha \ll \frac{1}{4\nu}, \tag{6.32}$$

or equivalently for the case of $\epsilon \ll \sigma$,

$$\nu \ll \frac{\pi\sigma}{4\epsilon}. \tag{6.33}$$

When ν increases to beyond this regime, the resolution of the measurement will scale as $\nu^{-1/2}$. This has been confirmed numerically, shown in Fig. 6.5. Note this result also implies that when we introduce any imperfection at all, the scaling will always be standard quantum limit in the asymptotic limit of large ν (but with a small prefactor). This conclusions applies generically to added dephasing imperfections in Heisenberg scaling schemes [RBJD13, DDKG12, JTT$^+$15].

It is straightforward to show that an N-pixel detection scheme also reaches Heisenberg scaling in the limit of very strong entanglement. In the same notation as Sec. 6.1.2, we expand the probabilities to first order in d,

$$P_{ij} \approx \alpha_{ij} + \beta_{ij}d. \tag{6.34}$$

Inserting these probabilities into Eq. (6.23), we see immediately that if even a single term in the sum satisfies $a_{ij} = 0$, the Fisher information in the limit of very small d scales as d^{-1} and hence the minimum resolvable parameter

scales as ν^{-1} for small ν.

We stress that if one has higher resolution, as can be achieved with more pixels, then the measurement precision can be increased. Since the perfect detector limit of the entangled biphoton case scales as the standard quantum limit (6.6) but with a prefactor of ϵ, we have the following ordering of precision from greatest to least given a fixed number of photons: (1) Standard quantum limit scaling of the biphoton case with a perfect resolution detector, (2) Heisenberg scaling of the biphoton case with a split (2-pixel) detector, (3) Standard quantum limit scaling of independent photons with a perfect resolution detector, (4) Standard quantum limit scaling of independent photons with a split (biphoton) resolution detector. This indicates that the scaling behavior is a secondary consideration to the Fisher information.

6.2 Estimation from marginal distributions

In this section, we study an alternative approach for estimating the displacement of biphoton beam using split-detection. We consider the maximum likelihood estimation for the marginal distribution of each photon in a biphoton pair under split-detection and average the two estimates. We show here this procedure can also attain the Cramér-Rao bound when the estimates of the marginal distributions are evenly averaged.

According to Eq. (6.2), the marginal distribution of either photon in a

biphoton pair is

$$p(x|d) = \sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}} \exp\left(-\frac{2(d-x)^2}{\sigma^2 + \epsilon^2}\right), \tag{6.35}$$

where the subscripts 1, 2 are dropped since the marginal distributions for the two photons are the same due to the symmetry of $p(x_1, x_2|d)$.

For split-detection, the probability to find either photon of a biphoton pair at the left or right side of the detector is, respectively,

$$p_- = \int_{-\infty}^{0} p(x|d)\mathrm{d}x = \frac{1}{2}\left(1 - \mathrm{erf}\left(\frac{\sqrt{2}d}{\sqrt{\sigma^2 + \epsilon^2}}\right)\right) \approx \frac{1}{2} - \sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}}d,$$

$$p_+ = \int_{0}^{+\infty} p(x|d)\mathrm{d}x = \frac{1}{2}\left(1 + \mathrm{erf}\left(\frac{\sqrt{2}d}{\sqrt{\sigma^2 + \epsilon^2}}\right)\right) \approx \frac{1}{2} + \sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}}d,$$

$$\tag{6.36}$$

where the approximation assumes $d \ll \sqrt{\sigma^2 + \epsilon^2}$.

The Fisher information for the marginal distribution of either photon is

$$\mathcal{I}_{\mathrm{m}} = \frac{8}{\pi\left(\sigma^2 + \epsilon^2\right)}. \tag{6.37}$$

The subscript m in $\mathcal{I}_{\mathrm{m}}(d)$ indicates it is the Fisher information of marginal distribution.

For the maximum likelihood estimation (MLE),

$$N_- \partial_d \log p_- + N_+ \partial_d \log p_+ = 0, \tag{6.38}$$

where N_- and N_+ are the numbers of the photons that arrive at the left or right side of the detector, respectively, so the estimator for the marginal distribution is

$$\hat{d} = \sqrt{\frac{\pi(\sigma^2 + \epsilon^2)}{8}} \frac{N_+ - N_-}{N_+ + N_-}. \tag{6.39}$$

When there are N pairs of biphotons in total, $N = N_+ + N_-$, and the distribution of N_+, N_- is

$$p(N_+, N_-) = \binom{N}{N_+} p_+^{N_+} p_-^{N_-}. \tag{6.40}$$

The variances of N_+, N_- and the covariance between them can be obtained:

$$\text{Var}[N_\pm] = N p_+ p_-, $$
$$\text{Cov}[N_+, N_-] = -N p_+ p_-, \tag{6.41}$$

so the variance of $\hat{d}$ is

$$\begin{aligned}\text{Var}[\hat{d}] &= \frac{\pi(\sigma^2 + \epsilon^2)}{8N^2}(\text{Var}[N_+] + \text{Var}[N_-] - 2\text{Cov}[N_+, N_-]) \\ &\approx \frac{\pi(\sigma^2 + \epsilon^2)}{8N}. \end{aligned} \tag{6.42}$$

The inverse of $\text{Var}[\hat{d}]$ is exactly the Fisher information (6.37) (scaled by N), so it verifies that the estimator $\hat{d}$ reaches the Cramér-Rao bound.

Since we have two photons from a biphoton pair, we can do some average of the estimates from these two photons. We denote the estimators of the two photons are $\hat{d}_1$ and $\hat{d}_2$ respectively, each satisfying the distribution (6.35). If

we average them with some weights α_1 and α_2, then the total estimator is

$$\hat{d}_{\text{tot}} = \alpha_1 \hat{d}_1 + \alpha_2 \hat{d}_2, \ \alpha_1 + \alpha_2 = 1. \tag{6.43}$$

The variance of this total estimator is

$$\text{Var}[\hat{d}_{\text{tot}}] = \alpha_1^2 \text{Var}[\hat{d}_1] + \alpha_2^2 \text{Var}[\alpha_2] \\ + 2\alpha_1 \alpha_2 \text{Cov}[\hat{d}_1, \hat{d}_2]. \tag{6.44}$$

We have obtained the $\text{Var}[\hat{d}_{1,2}]$ above, and the covariance $\text{Cov}[\hat{d}_1, \hat{d}_2]$ is

$$\text{Cov}[\hat{d}_1, \hat{d}_2] = \mathbb{E}[\hat{d}_1 \hat{d}_2] - \mathbb{E}[\hat{d}_1]\mathbb{E}[\hat{d}_2] \\ = \mathbb{E}[\hat{d}_1 \hat{d}_2] - d^2, \tag{6.45}$$

where we have used the fact $\mathbb{E}[\hat{d}_1] = \mathbb{E}[\hat{d}_2] = d$. From here we must calculate $\mathbb{E}[\hat{d}_1 \hat{d}_2]$.

From the joint spatial distribution (6.2), we can work out the joint distribution for the split-detection results, which can be denoted as p_{++}, p_{+-}, p_{-+}, p_{--}. Then $p_{--} = P(-2|d)$, $p_{++} = P(2|d)$, and $p_{+-} + p_{-+} = P(0|d)$, which were defined in Eq. (6.12), and

$$p_+ = p_{++} + p_{+-} = p_{++} + p_{-+}, \\ p_- = p_{--} + p_{-+} = p_{--} + p_{+-}, \tag{6.46}$$

according to the definition of marginal distribution and the symmetry of the

biphoton distribution.

$\mathbb{E}[\hat{d}_1\hat{d}_2]$ can be expanded as

$$
\begin{aligned}
\mathbb{E}[\hat{d}_1\hat{d}_2] &= \frac{\pi(\sigma^2 + \epsilon^2)}{8N^2}\mathbb{E}[(N_+^{(1)} - N_-^{(1)})(N_+^{(2)} - N_-^{(2)})] \\
&= \frac{\pi(\sigma^2 + \epsilon^2)}{8N^2}\mathrm{E}[(N_{++} + N_{+-} - N_{-+} - N_{--}) \\
&\quad \times (N_{++} + N_{-+} - N_{+-} - N_{--})],
\end{aligned}
\tag{6.47}
$$

where N_{++} is the number of times that both photons of a biphoton pair hit the right side of the detector, and similarly for N_{+-}, N_{-+}, N_{--}. The joint probability distribution for N_{++}, N_{+-}, N_{-+}, N_{--} is a multinomial distribution

$$
\begin{aligned}
&p(N_{++}, N_{+-}, N_{-+}, N_{--}) \\
&= \frac{N!}{N_{++}!N_{+-}!N_{-+}!N_{--}!}p_{++}^{N_{++}}p_{+-}^{N_{+-}}p_{-+}^{N_{-+}}p_{--}^{N_{--}},
\end{aligned}
\tag{6.48}
$$

so,

$$
\begin{aligned}
\mathbb{E}[N_{ij}] &= Np_{ij}, \\
\mathbb{E}[N_{ij}N_{i'j'}] &= N(N-1)p_{ij}p_{i'j'} + Np_{ij}\delta_{ij,i'j'},
\end{aligned}
\tag{6.49}
$$

where $i, j, i', j' = +$ or $-$. Therefore,

$$
\begin{aligned}
\mathbb{E}[\hat{d}_1\hat{d}_2] &= \frac{\pi(\sigma^2 + \epsilon^2)}{8N^2}[N(N-1)(p_{++} + p_{+-} - p_{-+} - p_{--}) \\
&\quad \times (p_{++} + p_{-+} - p_{+-} - p_{--}) + N(p_{++} - p_{+-} - p_{-+} + p_{--})] \\
&= \frac{\pi(\sigma^2 + \epsilon^2)}{8N}[(N-1)(p_+ - p_-)^2 + 2(p_{++} + p_{--}) - 1],
\end{aligned}
\tag{6.50}
$$

where Eq. (6.46) has been used at the second equality.

The $(p_+ - p_-)^2$ term can be dropped in Eq. (6.50) because it is proportional to d^2 according to (6.36), and the $-d^2$ term can also be dropped in Eq. (6.45) for the same reason, so,

$$\text{Cov}[\hat{d}_1, \hat{d}_2] \approx \frac{\pi(\sigma^2 + \epsilon^2)}{8N}[2(p_{++} + p_{--}) - 1]. \tag{6.51}$$

It can be derived by direct calculation that

$$
\begin{aligned}
p_{++} &= \frac{1}{4} + \frac{\arcsin \xi}{2\pi} + d\sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}}, \\
p_{--} &= \frac{1}{4} + \frac{\arcsin \xi}{2\pi} - d\sqrt{\frac{2}{\pi(\sigma^2 + \epsilon^2)}}.
\end{aligned}
\tag{6.52}
$$

Thus,

$$\text{Cov}[\hat{d}_1, \hat{d}_2] = \frac{(\sigma^2 + \epsilon^2)\arcsin \xi}{4N}. \tag{6.53}$$

The result of $\text{Var}[\hat{d}_{\text{tot}}]$ is therefore,

$$
\begin{aligned}
\text{Var}[\hat{d}_{\text{tot}}] &= \frac{\pi(\sigma^2 + \epsilon^2)}{8N}(\alpha_1^2 + \alpha_2^2 + \frac{4}{\pi}\alpha_1\alpha_2 \arcsin \xi) \\
&= \frac{\pi(\sigma^2 + \epsilon^2)}{8N}[1 - 2\alpha_1\alpha_2(1 - \frac{2}{\pi}\arcsin \xi)],
\end{aligned}
\tag{6.54}
$$

where we have used $\alpha_1 + \alpha_2 = 1$.

What is the minimum value of $\text{Var}[\hat{d}_{\text{tot}}]$? Since $\frac{2}{\pi}\arcsin \xi - 1 \leq 0$, $\text{Var}[\hat{d}_{\text{tot}}]$ is minimized when $\alpha_1\alpha_2$ is maximized. Using $\alpha_1 + \alpha_2 = 1$, we have $\alpha_1\alpha_2 \leq \frac{1}{4}$.

The equality holds when $\alpha_1 = \alpha_2 = \frac{1}{2}$. So, the minimum variance of $\hat{d}_{\text{tot}}$ is

$$\text{Var}[\hat{d}_{\text{tot}}]_{\min} \approx \frac{\pi(\sigma^2 + \epsilon^2)}{8N}\left(\frac{1}{2} + \frac{1}{\pi}\arcsin\xi\right). \tag{6.55}$$

This coincides with the Fisher information $\mathcal{I}(d)$ in Eq. (6.17) of the main manuscript. It implies that the optimal precision of split detection of biphotons can be attained via estimation from the two marginal distributions followed by properly averaging the two estimates.

6.3 Fisher information: quantum vs. classical

In this section, we study when the quantum Fisher information of a parameter-dependent state can be achieved by projective measurements. Suppose the state of interest is $|\psi_d\rangle$. If we want to estimate d by performing a projective measurement on this state, along some d-independent basis $\{|e_1\rangle, \cdots, |e_n\rangle\}$, where n is the dimension of the system, we can expand the state $|\psi_d\rangle$ along that basis as

$$|\psi_d\rangle = \sum_{k=1}^{n} \sqrt{p_k}\,\mathrm{e}^{i\theta_k}|e_k\rangle, \tag{6.56}$$

where p_k, θ_k, $k = 1, \cdots, n$ depend on d. Then the probability of obtaining the k-th result is p_k, and the classical Fisher information of estimating d by

such a projective measurement is

$$\mathcal{I}_C = \sum_{k=1}^{n} \frac{(\partial_d p_k)^2}{p_k}. \tag{6.57}$$

The quantum Fisher information of $|\psi_d\rangle$ is the maximum Fisher information of estimating d over all possible POVMs (not only projective measurements) on $|\psi_d\rangle$, and it can be obtained as

$$\mathcal{I}_Q = 4(\langle \partial_d \psi_d | \partial_d \psi_d \rangle - |\langle \psi_d | \partial_d \psi_d \rangle|^2). \tag{6.58}$$

Substituting Eq. (6.56) into (6.58), the quantum Fisher information of $|\psi_d\rangle$ is

$$\mathcal{I}_Q = 4\Big(\sum_{k=1}^{n} \Big| \frac{\partial_d p_k}{2\sqrt{p_k}} + \mathrm{i}\sqrt{p_k}\partial_d\theta_k \Big|^2 - \Big| \sum_{k=1}^{n} \sqrt{p_k}\Big(\frac{\partial_d p_k}{2\sqrt{p_k}} + \mathrm{i}\sqrt{p_k}\partial_d\theta_k \Big) \Big|^2 \Big)$$
$$= \sum_{k=1}^{n} \frac{(\partial_d p_k)^2}{p_k} + 4\Big[\sum_{k=1}^{n} p_k(\partial_d\theta_k)^2 - \Big| \sum_{k=1}^{n} p_k\partial_d\theta_k \Big|^2 \Big]. \tag{6.59}$$

Hence,

$$\mathcal{I}_Q - \mathcal{I}_C = 4\Big[\sum_{k=1}^{n} p_k(\partial_d\theta_k)^2 - \Big| \sum_{k=1}^{n} p_k\partial_d\theta_k \Big|^2 \Big]. \tag{6.60}$$

Applying the Cauchy-Schwartz inequality,

$$\Big| \sum_{k=1}^{n} p_k\partial_d\theta_k \Big|^2 \leq \sum_{k=1}^{n} p_k(\partial_d\theta_k)^2; \tag{6.61}$$

the equality holds only when

$$c\sqrt{p_k} = \sqrt{p_k}\partial_d\theta_k, \forall k, \tag{6.62}$$

where c is a constant. So, when $\mathcal{I}_Q = \mathcal{I}_C$, the solution to θ_k is $\theta_k = cd + \gamma_k$, where γ_k is a constant independent of d for each k. In this case,

$$e^{i\theta_k} = e^{icd}e^{i\gamma_k}. \tag{6.63}$$

Note that e^{icd} is a global phase of $|\psi_d\rangle$ which can be dropped, and $e^{i\gamma_k}$ is a phase independent of d. Therefore, the quantum Fisher information is equal to the classical Fisher information only when the state has no parameter-dependent relative phases in the basis of the measurement.

6.4 Remarks

By using spatially entangled biphoton pairs, it is possible to reduce the quantum noise intrinsic to optical metrology schemes. While this is true in the case of split-detection, which is a typical detector setup for optical displacement measurements, we have shown that an even larger benefit can be obtained as one improves spatial resolution. The ultimate physical limitation for this method is the position uncertainty in the birth zone of the biphoton. We have also seen that it is possible to change the scaling in precision for split detection measurement between standard quantum limit and Heisenberg if

one uses entangled biphotons instead of independent photons for the same mode profile, for relatively small photon number. This advantage comes from the fact that the sum of the biphoton positions is determined more precisely than for uncorrelated photons.

The detection and estimation scheme used in the chapter is to count the number of times that the biphoton arrive at the left side, right side or at different sides of the detector and then extract the information of the beam displacement from the counts. We proposed an experiment using coincidence counting and two split detectors to implement this theory. This method is statistically optimal for this experimental setup. A possible alternative approach is to consider the distribution of each single photon, estimate the displacement from it, and then average the estimates from the two photons. This approach also utilizes the correlation of the entangled photons, since the two estimates are correlated by the photon correlation. We have shown this approach also saturates the Cramér-Rao bound, and therefore matches the precision of the split-detection measurement considered here.

Chapter 7

Conclusion

In this thesis we have investigated several approaches to precision optical measurements. In particular, we have discussed how weak value amplification is capable of increasing the expectation value of a pointer variable (like the transverse position of an optical beam) without bound. However, weak values are not able to increase the sensitivity over more traditional measurement schemes when considering a noiseless experimental environment. As discussed in Chapter 4, when one considers different sources of technical noise, even a relatively simple weak value amplification experiment will outperform direct measurements.

We have also shown several techniques useful in other metrological contexts can be used to significantly enhance the original weak value experiments. In Chapters 2 and 3, we considered two different recycling schemes which allow a boost in sensitivity due to an effective power increase. In the

case of continuous wave recycling, we showed it is actually possible to obtain an arbitrarily precise measurement if system losses can be ignored. In Chapter 5 we demonstrated that it is also possible to use squeezed light to reduce quantum noise in addition to technical noise advantages already afforded by weak value amplification.

Finally, we have discussed the ability of spatially entangled biphoton pairs to reduce the uncertainty in an optical displacement measurement well below the level of the uncertainties in the position of the individual photons. Overall, we have shown it is possible to design optical metrological experiments which are robust to many types of noise, and which can achieve precision far superior to traditional measurement schemes.

Bibliography

[AAV88]   Yakir Aharonov, David Z Albert, and Lev Vaidman. How the result of a measurement of a component of the spin of a spin-1/2 particle can turn out to be 100. *Phys. Rev. Lett.*, 60(14):1351 – 1354, April 1988.

[AEdMF+15]   G. Bié Alves, B. M. Escher, R. L. de Matos Filho, N. Zagury, and L. Davidovich. Weak-value amplification as an optimal metrological protocol. *Phys. Rev. A*, 91:062107, Jun 2015.

[Aga13]   G. S. Agarwal. *Quantum Optics*. Cambridge University Press, 2013.

[BAC+03]   Nicolas Brunner, Antonio Acin, Daniel Collins, Nicolas Gisin, and Valerio Scarani. Optical telecom networks as weak quantum measurements with post-selection. *Phys. Rev. Lett.*, 91:180402, 2003.

[BB14]   Thomas J. Bullock and Paul Busch. Focusing in arthurs-kelly-

type joint measurements with correlated probes. *Phys. Rev. Lett.*, 113:120401, Sep 2014.

[BCM96]   S. L. Braunstein, C. M. Caves, and G.J. Milburn. Generalized uncertainty relations: theory, examples, and Lorentz invariance. *Ann. Phys.*, 247:135, 1996.

[BFM03]   S. M. Barnett, C. Fabre, and A. Maˆitre. Ultimate quantum limits for resolution of beam displacements. *Eur. Phys. J. D*, 22(3):513–519, 2003.

[BGJK15] Courtney Byard, Trent Graham, Andrew Jordan, and Paul G Kwiat. Pulse recycling and weak value metrology. In *Frontiers in Optics*, pages JW2A–69. Optical Society of America, 2015.

[BS10]   Nicolas Brunner and Christoph Simon. Measuring Small Longitudinal Phase Shifts: Weak Measurements or Standard Interferometry? *Phys. Rev. Lett.*, 105:010405, 2010.

[BW59]   Max Born and Emil Wolf. *Principles of Optics*. Pergamon Press, New York, 1959.

[Cah13]   K. Cahill. *Physical Mathematics*. Cambridge University Press, 2013.

[Cav81]   Carlton M. Caves. Quantum-mechanical noise in an interferometer. *Phys. Rev. D*, 23(8):1693–1708, 1981.

[CB02] George Casella and Roger L. Berger. *Statistical Inference.* Wadsworth, 2002.

[DDKG12] Rafał Demkowicz-Dobrzański, Jan Kołodyński, and Mădălin Guţă. The elusive heisenberg limit in quantum-enhanced metrology. *Nature communications*, 3:1063, 2012.

[Di 12] Antonio Di Lorenzo. Full counting statistics of weak measurement. *Phys. Rev. A*, 85(3):032106, 2012.

[DL13] Antonio Di Lorenzo. Correlations between detectors allow violation of the heisenberg noise-disturbance principle for position and momentum measurements. *Phys. Rev. Lett.*, 110:120403, Mar 2013.

[DLJ+13] Justin Dressel, Kevin Lyons, Andrew N Jordan, Trent M Graham, and Paul G Kwiat. Strengthening weak-value amplification with recycled photons. *Physical Review A*, 88(2):023821, 2013.

[DMM+14] J. Dressel, M. Malik, F. M. Miatto, A. N. Jordan, and R. W. Boyd. Understanding Quantum Weak Values: Basics and Applications. *Rev. Mod. Phys.*, 86:307, 2014.

[Dre83] R.W.P. Drever. In *Quantum Optics: Experimental Gravity and Measurement Theory*, pages 525–566. Plenum, New York, 1983.

[DSJH09] P Ben Dixon, David J Starling, Andrew N Jordan, and John C Howell. Ultrasensitive Beam Deflection Measurement via Interferometric Weak Value Amplification. *Phys. Rev. Lett.*, 102:173601, 2009.

[DSS89] I. M. Duck, P. M. Stevenson, and E. C. G. Sudarshan. The sense in which a "weak measurement" of a spin-1/2 particle's spin component yields a value 100. *Phys. Rev. D*, 40(6):2112–2117, 1989.

[EPR35] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Phys. Rev.*, 47:777–780, May 1935.

[ES12] Patrick Egan and Jack A. Stone. Weak-value thermostat with 0.2 mK precision. *Opt. Lett.*, 37:4991–4993, 2012.

[FMV09] M. V. Fedorov, Y. M. Mikhailova, and P.A. Volkov. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 42(17):175503, 2009.

[HBL16] Jérémie Harris, Robert W Boyd, and Jeff S Lundeen. Weak value amplification can outperform conventional measurement in the presence of detector saturation. *arXiv preprint arXiv:1612.04327*, 2016.

[HHC+11] J. M. Hogan, J. Hammer, S.-W. Chiow, S. Dickerson, D. M. S. Johnson, T. Kovachy, A. Sugarbaker, and M. A. Kasevich. Precision angle sensor using an optical lever inside a Sagnac interferometer. *Opt. Lett.*, 36:1698–1700, 2011.

[HK08] Onur Hosten and Paul Kwiat. Observation of the Spin Hall Effect of Light via Weak Measurements. *Science*, 319:787, 2008.

[HSD+10] J. C. Howell, D. J. Starling, P. B. Dixon, P. K. Vudyasetu, and A. N. Jordan. Interferometric weak value deflections: quantum and classical treatments. *Phys. Rev. A*, 81:033813, 2010.

[JMRH14] A. N. Jordan, J. Martinez-Rincón, and J. C. Howell. Technical advantages for weak value amplification: When less is more. *Phys. Rev. X*, 4:011031, 2014.

[JTT+15] Andrew N Jordan, Jeff Tollaksen, James E Troupe, Justin Dressel, and Yakir Aharonov. Heisenberg scaling with weak measurement: a quantum state discrimination point of view. *Quantum Studies: Mathematics and Foundations*, 2(1):5–15, 2015.

[KAN12] Abraham G. Kofman, Sahel Ashhab, and Franco Nori. Nonperturbative theory of weak pre- and post-selected measurements. *Physics Reports*, 520(2):43 – 133, 2012. Nonperturbative theory of weak pre- and post-selected measurements.

[KG14] George C. Knee and Erik M. Gauger. When amplification with weak values fails to suppress technical noise. *Phys. Rev. X*, 4:011032, Mar 2014.

[KM15] G.C. Knee and W.J. Munro. Fisher information vs. signal-to-noise ratio for a split detector. *Phys. Rev. A*, 92:012130, 2015.

[LDJ+15] Kevin Lyons, Justin Dressel, Andrew N Jordan, John C Howell, and Paul G Kwiat. Power-recycled weak-value-based metrology. *Physical review letters*, 114(17):170801, 2015.

[LE04] C. K. Law and J. H. Eberly. Analysis and interpretation of high transverse entanglement in optical parametric down conversion. *Phys. Rev. Lett.*, 92:127903, Mar 2004.

[LM01] R. J. Larsen and L. M. Marx. *An Introduction to Mathematical Statistics and Its Applications*. Prentice Hall, 2001.

[Lou00] Rodney Loudon. *The Quantum Theory of Light*. Oxford Science Publications, 2000.

[LPKJ16] Kevin Lyons, Shengshi Pang, Paul G. Kwiat, and Andrew N. Jordan. Precision optical displacement measurements using biphotons. *Phys. Rev. A*, 93:043841, Apr 2016.

[ME10] Peter W. Milonni and Joseph H. Eberly. *Laser Physics*. Wiley, 2010.

[MS91]   B. J. Meers and K. A. Strain. Wave-front distortion in laser-interferometric gravitational-wave detectors. *Phys. Rev. D*, 43:3117–3130, 1991.

[MW95]   Leonard Mandel and Emil Wolf. *Optical Coherence and Quantum Optics*. Cambridge University Press, 1995.

[PABJ16]   Shengshi Pang, Jose Raul Gonzalez Alonso, Todd A. Brun, and Andrew N. Jordan. Protecting weak measurements against systematic errors. *Phys. Rev. A*, 94:012329, Jul 2016.

[PF11]   Marcel Pfeifer and Peer Fischer. Weak value amplified optical activity measurements. *Opt. Exp.*, 19:16508–16517, 2011.

[RBJD13]   Bhaskar Roy Bardhan, Kebei Jiang, and Jonathan P Dowling. Effects of phase fluctuations on phase sensitivity and visibility of path-entangled photon fock states. *Physical Review A*, 88(2):023857, 2013.

[RSH91]   N W M Ritchie, J G Story, and Randall G Hulet. Realization of a measurement of a "weak value". *Phys. Rev. Lett.*, 66(9):1107 – 1110, 1991.

[SB13]   G. Strübi and C Bruder. Measuring Ultrasmall Time Delays of Light by Joint Weak Measurements. *Phys. Rev. Lett.*, 110:083605, 2013.

[SDJH09] David J. Starling, P. Ben Dixon, Andrew N. Jordan, and John C. Howell. Optimizing the signal-to-noise ratio of a beam-deflection measurement with interferometric weak values. *Phys. Rev. A*, 80:041803, 2009.

[SDJH10] David J. Starling, P. Ben Dixon, Andrew N. Jordan, and John C. Howell. Precision frequency measurements with interferometric weak values. *Phys. Rev. A*, 82:063822, 2010.

[SDW+10a] David J. Starling, P. Ben Dixon, N. S. Williams, Andrew N. Jordan, and John C. Howell. Continuous phase amplification with a Sagnac interferometer. *Phys. Rev. A*, 82:011802(R), 2010.

[SDW+10b] David J. Starling, P. Ben Dixon, N. S. Williams, Andrew N. Jordan, and John C. Howell. Continuous phase amplification with a Sagnac interferometer. *Phys. Rev. A*, 82:011802(R), 2010.

[SH15] J. Schneeloch and J. C. Howell. Introduction to the Transverse Spatial Correlations in Spontaneous Parametric Down-Conversion through the Biphoton Birth Zone. 2015.

[SMH+97] D. Schnier, J. Mizuno, G. Heinzel, H Lück, A. Rüdinger, R. Schilling, M. Schrempel, W. Winkler, and K. Danzmann. Power recycling in the Garching 30 m prototype interferometer

for gravitational-wave detection. *Phys. Lett. A*, 225:210–216, 1997.

[TAB⁺02] N. Treps, U. Andersen, B. Buchler, P. K. Lam, A. Maitre, H.-A. Bachor, and C. Fabre. Surpassing the Standard Quantum Limit for Optical Imaging Using Nonclassical Multimode Light. *Phys. Rev. Lett.*, 88(20):203601, 2002.

[TGB⁺03] N. Treps, Nicolai Grosse, Warwick P. Bowen, Claude Fabre, Hans-A. Bachor, and Ping Koy Lam. A Quantum Laser Pointer. *Science*, 301:940, 2003.

[THSG11] Matthew D. Turner, Charles A. Hagedorn, Stephan Schlamminger, and Jens H. Gundlach. Picoradian deflection measurement with an interferometric quasi-autocollimator using weak value amplification. *Opt. Lett.*, 36:1479–1481, 2011.

[TSS16] Juan P Torres and Luis José Salazar-Serrano. Weak value amplification: a view from quantum estimation theory that highlights what it is and what isnt. *Scientific reports*, 6:19702, 2016.

[VCH⁺05] Henning Vahlbruch, Simon Chelkowski, Boris Hage, Alexander Franzen, Karsten Danzmann, and Roman Schnabel. Demonstration of a Squeezed-Light-Enhanced Power- and

Signal-Recycled Michelson Interferometer. *Phys. Rev. Lett.*, 95:211102, 2005.

[VMC$^+$08] Henning Vahlbruch, Moritz Mehmet, Simon Chelkowski, Boris Hage, Alexander Franzen, Nico Lastzka, Stefan Goßler, Karsten Danzmann, and Roman Schnabel. Observation of Squeezed Light with 10-dB Quantum Noise Reduction. *Phys. Rev. Lett.*, 100:033602, 2008.

[VMRA$^+$14] G. I. Viza, J. Martinez-Rincón, G. B. Alves, A. N. Jordan, and J. C. Howell. Experimentally Quantifying the Advantages of Weak-Values-Based Metrology. *arXiv:1410.8461*, 2014.

[VMRA$^+$15] Gerardo I Viza, Julián Mart´inez-Rincón, Gabriel B Alves, Andrew N Jordan, and John C Howell. Experimentally quantifying the advantages of weak-value-based metrology. *Physical Review A*, 92(3):032127, 2015.

[VMRH$^+$13] Gerardo I. Viza, Julián Martinez-Rincón, Gregory A. Howland, Hadas Frostig, Itay Shromroni, Barak Dayan, and John C. Howell. Weak Values Technique for Velocity Measurements. 2013.

[Was04] Larry Wasserman. *All of Statistics.* Springer, 2004.

[WTH$^+$16] Yi-Tao Wang, Jian-Shun Tang, Gang Hu, Jian Wang, Shang Yu, Zong-Quan Zhou, Ze-Di Cheng, Jin-Shi Xu, Sen-Zhi Fang,

Qing-Lin Wu, et al. Experimental demonstration of higher precision weak-value-based metrology using power recycling. *Physical Review Letters*, 117(23):230801, 2016.

[XWK87] Min Xiao, Ling-An Wu, and H. J. Kimble. Precision Measurement beyond the Shot-Noise Limit. *Phys. Rev. Lett.*, 59(3):278, 1987.

[ZXLW12] Xinxing Zhou, Zhicheng Xiao, Hailu Luo, and Shuangchun Wen. Experimental observation of the spin Hall effect of light on a nanometal film via weak measurements. *Phys. Rev. A*, 85:043809, 2012.