Lecture 4

Fisher information and entropy
atomic “twin beams”
Mutual information through PIAs and PSAs
information

If we want to look at the transfer of information through different media and to quantify its degradation, we need to define a few types of information.

Information is related to entropy.

Different “brands” of information are more useful for different applications; ex., Fisher information is used in interferometry (and parameter estimation in general)
entropy

\[ S = -k_B \sum_i p_i \log p_i = -k_B Tr(\rho \log \rho) \]

if \( p_i \) is a constant, \( 1/\Omega \), (all states are equally probable) then \( S = -k_B \log \Omega \)

if the log is base \( e \) the units are “nats” (from “natural logarithm”)

**Shannon entropy**

\[ H = - \sum_i p_i \log_2(p_i) \]

log base 2 – the unit is “bits”
mutual information

\[ \text{MI} = I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x) p(y)} \right) \]

This quantifies the amount of information shared by two systems.
In some of the measurements that follow we will use both a single-quadrature MI as well as an MI for a beam of light (defined to include information in both of the two orthogonal quadratures).
quantum mutual information

von Neumann entropy $S_V(\rho) = -Tr[\rho \log(\rho)]$

two classically equivalent expressions for mutual information:

\[ I(\rho) = S_V(\rho_1) + S_V(\rho_2) - S_V(\rho) \]
\[ J(\rho) = S_V(\rho_2) - S_V(\rho_2|\rho_1) \]

quantum discord $D = I(\rho) - J(\rho)$
(maximized over all projective measurements onto the basis)
For a Gaussian state, this can be expressed in terms of variances and covariances that can be extracted from a series of homodyne measurements on the state.

Discord is, in some scenarios, a more faithful measure of nonclassicality, as it is sensitive to correlations that are ignored by measures like the inseparability or the logarithmic negativity.
Fisher information

\[ I(\theta) = E \left[ \left( \frac{\partial}{\partial \theta} \log f(X; \theta) \right)^2 \right] = \int \left( \frac{\partial}{\partial \theta} \log f(x; \theta) \right)^2 f(x; \theta) \, dx, \]

\[ F(\phi) = \int_\Lambda dx \, p(x|\phi) \left[ \partial_\phi \log p(x|\phi) \right]^2 \]

From Wikipedia:
“Fisher information is widely used in optimal experimental design. Because of the reciprocity of estimator-variance and Fisher information, minimizing the variance corresponds to maximizing the information.”

The Fisher information tells one the amount of information in the measurements and allows one to extract a parameter to a degree depending on how rapidly the distribution of measurements changes as the parameter changes.
Fisher information (cont.)

The Fisher information is the quantity to maximize in order to get the best (smallest) sensitivity in estimating the phase of an interferometer.

This information tells us about the distributions of possible measurements, given that a parameter has a certain value. The easier it is to tell two distributions apart, the larger the information distance is between them.
Cramér-Rao bound

\[ F(\phi) = \int_\Lambda dx \, p(x|\phi) \left[ \partial_\phi \log p(x|\phi) \right]^2 \]

Fisher information

\[ \langle \Delta^2 \phi \rangle \geq \frac{1}{MF(\phi)} \]

variance of the phase estimator or sensitivity

where \( M \) is the number of repeated measurements of a system

This is the variance and thus the standard deviation of the measurement scales as \( \sqrt{M} \).

For a classical system (of \( N \) un-entangled particles) \( F \sim N \) and the standard deviation of the measurement again scales as \( \sqrt{N} \).

Gaussian states

coherent states have symmetric distributions

squeezed states have asymmetric distributions

non-Gaussian states with negative-going Wigner functions that do not represent classical probability functions

figures from “A guide to Experiments in Quantum Optics” by H. Bachor and T. Ralph and “Quantum Optics” by G.S. Agarwal
Fisher information for Gaussian states

If \( p(x|\phi) \) is a Gaussian function with variance \( \sigma^2 \), the Fisher information is

\[
F(\phi) = \frac{(\partial_\phi X)^2 + 2(\partial_\phi \sigma)^2}{\sigma^2}
\]

The observable or estimator is \( X \).

If \( \langle X \rangle \) varies rapidly with \( \phi \), then the first term usually dominates. This is the usual condition – for example the intensity out of your interferometer varies with the phase.

The second term says that you can get some extra information by looking at how the width of the distribution changes.

For a Gaussian distribution, there is nothing more; it is completely characterized by knowledge of the mean and variance of the distribution.

If your detection scheme is non-Gaussian (photon counting) then you cannot use this formula!
“noise on the noise”

If \( \langle X \rangle = 0 \), however, the “signal” is now the noise, and the “noise” is the “noise on the noise”... and the second term is vital!

How does this work?

signal level \( \langle X \rangle \)

suppose the signal level stays at zero but the noise changes

Here the noise or \( \sigma \) varies with the phase. As it changes with time it tells us information about \( \phi \). How good this information is (“the noise on the noise”) tells us how sensitive the measurement is.
twin-beam atom interferometer example where this analysis is useful

spinor condensate in an optical trap

collisional interaction:

$m = 0$ BEC

“pump”

$m = +1$ atoms

$m = -1$ atoms

differential light shifts allow phase-shifting in trap

observation techniques

Make population measurements by Stern-Gerlach technique:

Atomic homodyne technique: couple atoms from the $m_F = 0$ “reservoir” (or LO) into the state that you want to observe before expansion and interfere the coherent matter waves.
atom interferometers

note:
This is NOT a “matter wave interferometer”
and does not rely on having a BEC or a coherent matter wave.

It is a spin-wave interferometer and can be constructed using a thermal cloud of atoms as well!
reversal of spinor dynamics

By applying a microwave phase pulse to the system, we can reverse the dynamics even after 24 ms where the evolved fraction is >50%
extracting the Fisher info. from data

The distribution at $\pi$ is the reference. Differences between the distributions are taken at different settings and used to calculate a “distance” (squared Hellinger distance) between the distributions.

The mean doesn’t change much, but the variance does!

M. Oberthaler group:
spinor 4WM

a number of groups have made interesting measurements on matter-wave 4WM in Rb spinor BECs

Klempt/Ertmer, Stamper-Kurn, Bongs/Sengstock, Oberthaler, Chapman


recent experiments:
“Twin Matter Waves for Interferometry Beyond the Classical Limit,”
Where Fisher information is particularly useful...

Certain nonlinear interactions create squeezing at short interaction times, but then go “beyond” squeezing. The variance grows even while the measurement distribution gets ever more non-classical.

These are non-Gaussian distributions!

extracting the Fisher info. from data

The blue distribution at 0 degrees is the reference. Differences between the distributions are taken at different settings and used to calculate a “distance” (squared Hellinger distance) between the distributions.

\[ d^2_H(\theta) = \frac{1}{2} \sum \left[ \sqrt{P_z(\theta)} - \sqrt{P_z(0)} \right]^2 \]

Fisher information and entanglement of non-Gaussian spin states
useful for “over-squeezing”

quantum enhanced sensitivity - even in the absence of squeezing!

Figure from: H. Strobel, et al., Science 345, 424 (2014)
other correlations

Bell tests with atoms
Many-body system with collective measurements on a
spin system.
A spin-squeezed state is generated.
The non-Gaussian operation here is atom number counting.
A special Bell inequality is derived for this many body system.

R. Schmied, et al., “Bell correlations in a BEC,”
Science 352, 441 (2016) (Treutlein, Sangouard)
First we will look at the transmission of information through a dispersive system.

We will generate our information as spontaneous noise on a beam, but we will generate it in correlated “twin beams” so that we can have a reference beam.

Our dispersive medium will be a 4WM cell.

We will look at the sum of the mutual information in the two quadratures of the Gaussian beams generated in the 4WM process.
quantum correlations through a fast/slow light cell

Generate quantum-correlated information into twin beams by spontaneous emission.
quantum correlations through a fast/slow light cell

Send one of the twin beams through a second 4wm cell that is tuned to amplify the light only very slightly, but has anomalous dispersion in the region of interest.
quantum correlations through a fast/slow light cell

- Cross-correlation trace -> advancement or delay.

- Intensity difference trace -> squeezing.
quantum correlations through a fast/slow light cell

intensity correlations

- Cross-correlation trace -> advancement or delay.
- Intensity difference trace -> squeezing.
quantum correlations through a fast/slow light cell

advance / delay and squeezing

- Cross-correlation trace -> advancement or delay.
- Intensity difference trace -> squeezing.
PIA noise model

quantum correlations through a fast/slow light cell
Intensity correlation results for vacuum quadrature squeezing

results for correlated and anti-correlated joint quadratures
entanglement results
Quantum mutual information advance/delay

-The advanced mutual information peak is advanced by the same amount as the intensity correlations.
-The advancement is always below the leading edge of the reference.
-The delay, for the same degradation, is able to escape the trailing edge of the reference.
(These are cw measurements!)  

J. Clark, et al., Nat. Phot. 8, 515 (2014)
Noise to the rescue?

Noise is added to the signal, which degrades the mutual information.

The noise is just enough, according to the 4WM PIA (phase insensitive amplifier) model, to prevent any “excess” information from arriving too early.

On the other hand, with a similar degradation of the mutual information we seem to have no problem in getting the mutual information to arrive late!
What about a PSA?

A phase-sensitive amplifier (PSA) is a different type of amplifier:

at the cost of phase sensitivity you can perform noiseless amplification of a (one quadrature) signal!

We can build either a PIA or a PSA with 4WM... and they have similar gain lines.

If you can amplify noiselessly with a PSA... does this imply problems when it comes to fast light???
Recall that the phase of the injected beam, with respect to those of the pumps, will determine whether the beam will be amplified or de-amplified.

One can ask how the signal in a single quadrature is advanced or delayed as a function of this phase.

\[ \phi_s = 2\phi_0 - \phi_- - \phi_+ \]
PSA experiment

(a) Diagram of PSA experiment setup.

(b) 5P_{1/2} level transitions for PIA (source) tunings:
- \( \delta \uparrow \)
- \( F=2 \rightarrow F=3 \)
- \( v_{HF} \)
- \( v_1 \) and \( v_c \) for PIA source.

(c) 5P_{1/2} level transitions for PSA tunings:
- \( \delta \uparrow \)
- \( F=2 \rightarrow F=3 \)
- \( v_1 \), \( v_p \), and \( v_2 \) for PSA.

PIA (source) tunings and PSA tunings are shown separately.
experiment

Insert a signal through the PSA and compare to a reference beam that does not pass through the PSA.

We can measure the cross-correlation between the two beams and find the apparent advance or delay caused by introducing the PSA.
Or we can determine the Mutual Information as a function of advance or delay and find its maximum.

Theory tells us what the relative phase of the input is for a measured gain.

=> this allows us to plot the measured gain vs. delay or input phase
time traces and joint probability distributions to determine MI

twin beams  

split thermal beams  

split coherent beams
mutual information versus time

\[ \text{MI} = I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p(x, y) \log \left( \frac{p(x, y)}{p(x) p(y)} \right) \]

some PSA phases give an advance and some a delay in the MI

red ~ amplified
blue ~ deamplified
mutual information in classical and quantum-correlated beams

Coherent beams are shot-noise limited, and shot noise is random and uncorrelated; no MI. A thermal beam split on a beamsplitter results in classical correlations. Twin beams are thermal beams with sub-shot-noise correlations.
PSA “Dispersion”

sine-wave modulation (correlation peak)

pulses with comparable bandwidth (correlation peak)

thermal noise from 4WM (MI peak)
theory

The PSA behavior doesn’t really look like “dispersion”. We get a phase advance or delay, but it is deterministic with gain...

We assume a simple model that reproduces the main features of the experiment.
- A carrier with two sidebands (a modulated beam) are introduced into the gain medium and the all experience a gain G.
- The lower sideband experiences a small loss and a small phase shift as well.
- The gain/loss, etc. is distributed through the medium
theoretical model fits

$G = 1.8$  
$G = 2.8$

identical PSA conditions

sine wave input

twin beam input
maximum mutual information versus advance/delay

peak PSA gain $\approx 1.8$

peak PSA gain $\approx 2.8$

red points correspond to $G \gtrsim 1$, blue points to $G \lesssim 1$

At maximum deamplification the PSA is noiseless; the MI is preserved (signal is not mixed with the orthogonal quadrature).
PSA action on a signal

If you project the amplified output back onto the original quadrature axis you will see a reduced SNR. No excess “noise” gets into the system but the quadratures get mixed.
PSA summary

• The PSA does not display “dispersion” in the usual way; it mixes quadratures unless the signal is along the maximum amplification or deamplification quadrature.

• This causes an apparent phase advance or delay.

• The output MI curve never escapes the reference boundary; the peak shifts but also shrinks to keep the MI bounded.