Collisional properties and many body effects in ultracold systems. Quantum interferometry.

Tomasz Wasak

Promotor
prof. dr hab. Marek Trippenbach
Faculty of Physics
University of Warsaw

Auxiliary Promoter

dr Jan Chwedeńczuk
Faculty of Physics
University of Warsaw

April 2015
Mojej żonie
Gabrieli
Collisional properties and many body effects in ultracold systems. 
Quantum interferometry.

Abstract

In a growing number of experiments, many-body entangled states are created and employed to surpass the shot-noise limit in atom interferometry. The precision is bounded by this limiting value whenever the particles exhibit only classical correlations. However, with a proper choice of the state which is injected into the interferometer, this limit can be overcome. Thus, a certain class of non-classically correlated states is useful for precise metrology.

The main objective of this thesis is to show that the indistinguishability of bosons is a resource that can be employed for ultra-precise interferometry. To support this idea, two systems are considered, where useful quantum correlations arise solely from the indistinguishability of the constituent bosons. The first example is a system, in which pairs of identical particles are scattered from a Bose-Einstein condensate into two well-separated regions, which are later identified with the two arms of an interferometer. In the second case, the interferometer is fed with two independently prepared Bose-Einstein condensates correlated only due to the indistinguishability.

The analysis of these systems is founded on the concept that some entangled states are highly susceptible to interferometric transformations. As is shown, this statement can be precisely quantified using measures of the distinguishability of the probability distributions characterizing those states.

In the course of discussion, the object known as the Fisher information naturally emerges as the quantity unifying all the presented measures. In other words, it is the Fisher information that describes the susceptibility of the state to an interferometric transformation. Moreover, the theorem known as the Cramér-Rao lower bound constitutes the bridge between the distinguishability of states and metrology.

In the next step, a theorem is presented which relates the Fisher information with non-classical correlations. The consequence of this statement is that the sub-shot-noise sensitivity of interferometric phase estimation is a signature of non-classical correlations.

Before injecting a many-body state into the interferometer, one should make sure if this state is usefully entangled. Therefore, it would be helpful to have a simple criterion for entanglement that would signify potential usefulness of the state. In this thesis, such a criterion is provided which is based on the Cauchy-Schwarz inequality. It applies to the system of bosons, which do not possess coherences between different number states.

After setting the theoretical framework, specific experimental scenarios are considered. First, a question is asked whether a collision of two Bose-Einstein condensates can lead to creation of entangled atomic pairs. It is shown that the detailed setup of the collision leads to effects that are destructive for generating entanglement useful for metrology. Therefore, a different scheme is necessary which would
be free of such phenomena.

In the main part of the thesis, a theory is developed, which describes the scattering of atomic pairs from a condensate into two disjointed zones. It is shown, that, contrary to the previous case, the system consists of highly entangled pairs useful for ultra-precise sub-shot-noise interferometry. Moreover, since the Cauchy-Schwarz inequality is violated in this system, the presence of entanglement can be easily verified. This theory is then applied to characterise the realistic case of the twin-atom beam experiment.

Finally, an interferometric scenario is considered in which the initial state consists of two independently prepared Bose-Einstein condensates. It is shown, that, when this state is taken as the input of an interferometer, the device can operate below the shot-noise limit. The conditions for this to happen are presented, together with the scheme of the interferometric phase estimation, such that could benefit from these non-classical correlations. The proposed scheme circumvents the difficult stage of the entangled state preparation. The only condition, which is required, is to have a precise control over the distribution of the total number of atoms in each atomic cloud.

Through this thesis, the link between the state susceptibility to the change and the entanglement is thoroughly exploited. Here, the high susceptibility of states, quantified by the Fisher information, is considered as a potential resource for ultra-precise metrology. Nevertheless, the underlying principle adopted in the thesis may also be applied in other areas of physics.
Acknowledgements

The acknowledgements are usually more personal and are governed by different laws than the rest of the text. Also, the words written here are always addressed to specific people. I thus feel justified to freely choose the language of expression, and that is why most of the thanks below are in Polish.

Szczególne podziękowania chciałbym przekazać mojemu opiekunowi naukowemu prof. Markowi Tripenbachowi, który umożliwił mi rozpoczęcie drogi naukowej przyjmując mnie do swojej grupy badawczej i angażując w liczne projekty. Dziękuję również za to, że przez te wszystkie lata dzielił się ze mną swoim doświadczeniem i wiedzą. Jestem wdzięczny za umożliwienie mi wyjazdów do zagranicznych ośrodków naukowych oraz nieustanne wsparcie mojej pracy naukowej.

Gorąco chciałbym podziękować Janowi Chwedeńczukowi za nasze wspólne niezliczone dyskusje dotyczące niemal każdej dziedziny fizyki oraz innych aspektów życia. Za ogromne wsparcie merytoryczne oraz emocjonalne jakim mnie obdarzył przy pisaniu tej pracy. Za aktywną pomoc, której nie raz mi udzielał, oraz na którą zawsze mogę liczyć.

Chciałbym podziękować Piotrowi Szańkowskiemu za to, że już od czasów studenckich dzieliliśmy się pomysłami, wspieramy oraz aktywnie współpracujemy.

Dziękuję również Pawłowi Ziź za to, że pokazał mi, jak głęboko można zrozumieć problemy fizyczne.

Piotrowi Deuarowi chciałbym podziękować za wspólnie spędzony czas, kiedy przekazywał mi tajniki swojej wiedzy na temat metody symulacji positive-P.

I would also like to thank Robert Bücker for discussions and working together on the twin-atom beam system.

Many thanks go to Augusto Smerzi who, during my stays in Florence, has been giving me advices about how to ask questions that drive one into deeper understanding of physical problems. Also, I am grateful for working together on projects, that always turned out to be intellectual adventures.


Szczególne podziękowania chciałbym złożyć moim rodzicom Bożenie i Ryszardowi Wasak za to, że dzięki nim rozbudziło się moje zainteresowanie fizyką. Również dziękuję za bezwarunkowe wsparcie mnie we wszystkich dziedzinach życia i na wszystkich etapach mojej edukacji.

Na koniec chciałbym podziękować kotu Bardo za naukę cierpliwego patrzenia, słuchania i wytrwałego skupienia. Za wielokrotnie towarzyszenie mi przy komputerze w wieczornych sesjach pisania pracy.
Acknowledgements

1 Introduction
  1.1 Classical and quantum mechanics ........................................... 1
  1.2 Experiments with non-classical correlations .............................. 8
    1.2.1 Particle-wave duality .................................................. 8
    1.2.2 Scattering of indistinguishable particles ............................ 11
  1.3 Conclusions ........................................................................... 16
  1.4 Structure of thesis ............................................................... 17

2 Distinguishability of states ..................................................... 19
  2.1 States distinguishability ........................................................ 20
  2.2 Fidelity as a measure of distinguishability .................................. 23
  2.3 Classical Fisher information .................................................... 28
  2.4 Other measures of distinguishability, $f$-divergences .................... 29
    2.4.1 Kolmogorov distance or trace distance .................................. 29
    2.4.2 Hellinger distance ........................................................ 30
    2.4.3 Kullback-Leibler divergence or relative entropy ..................... 31
    2.4.4 Generalization to the $f$-divergence .................................... 32
  2.5 Properties of the Fisher information ........................................ 35
  2.6 Metrology .............................................................................. 36
    2.6.1 Maximum-likelihood estimation ......................................... 38
    2.6.2 Estimation from the mean .................................................. 40
  2.7 Quantum Fisher information and quantum correlations .................... 43
    2.7.1 Optimal measurement. Quantum Fisher information .................... 44
    2.7.2 Particle entanglement ....................................................... 48
    2.7.3 Two mode interferometry with bosons .................................... 53
    2.7.4 Spin-squeezing and number-squeezing ................................. 55
  2.8 Conclusions ........................................................................... 57

3 Optimal measurements. Simple examples. .................................... 59
  3.1 Introduction ........................................................................... 60
  3.2 Single qubit ............................................................................ 61
    3.2.1 Classical and quantum Fisher information ............................... 62
    3.2.2 Optimal measurements ....................................................... 63
    3.2.3 Estimation from the population imbalance ............................... 64
  3.3 Symmetric Werner states ......................................................... 65
    3.3.1 Two qubits: symmetric Werner States ..................................... 65
    3.3.2 Generalization to higher N ................................................... 69
    3.3.3 Noisy NOON state ............................................................ 71
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4</td>
<td>Optimal measurements for symmetric N-qubit pure states</td>
<td>72</td>
</tr>
<tr>
<td>3.4.1</td>
<td>QFI and the statistical distance</td>
<td>72</td>
</tr>
<tr>
<td>3.4.2</td>
<td>“In-situ” measurements – localized modes</td>
<td>74</td>
</tr>
<tr>
<td>3.4.3</td>
<td>Measurement after expansion</td>
<td>76</td>
</tr>
<tr>
<td>3.5</td>
<td>Conclusions</td>
<td>77</td>
</tr>
<tr>
<td>4</td>
<td>Cauchy-Schwarz inequality and particle entanglement</td>
<td>79</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>79</td>
</tr>
<tr>
<td>4.2</td>
<td>Formulation of the problem</td>
<td>81</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Relation to number squeezing</td>
<td>83</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Generalization to non-fixed N</td>
<td>84</td>
</tr>
<tr>
<td>4.3</td>
<td>CSI vs. Fisher information</td>
<td>85</td>
</tr>
<tr>
<td>4.4</td>
<td>CSI for distinguishable particles</td>
<td>86</td>
</tr>
<tr>
<td>4.5</td>
<td>Conclusions</td>
<td>87</td>
</tr>
<tr>
<td>5</td>
<td>Collisions of Bose-Einstein condensates</td>
<td>89</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction and Outline</td>
<td>90</td>
</tr>
<tr>
<td>5.2</td>
<td>Theoretical model for BEC collision</td>
<td>91</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Collision parameters</td>
<td>92</td>
</tr>
<tr>
<td>5.2.2</td>
<td>BEC wave-function</td>
<td>92</td>
</tr>
<tr>
<td>5.2.3</td>
<td>Positive-P method</td>
<td>94</td>
</tr>
<tr>
<td>5.2.4</td>
<td>Scattering in the absence of bosonic enhancement</td>
<td>94</td>
</tr>
<tr>
<td>5.3</td>
<td>Two-body correlations of scattered atoms</td>
<td>96</td>
</tr>
<tr>
<td>5.4</td>
<td>Number squeezing in a multi-mode system</td>
<td>97</td>
</tr>
<tr>
<td>5.5</td>
<td>Simulation results and analysis</td>
<td>100</td>
</tr>
<tr>
<td>5.5.1</td>
<td>γ = 1.02 case</td>
<td>100</td>
</tr>
<tr>
<td>5.5.2</td>
<td>γ = 0.24 case</td>
<td>103</td>
</tr>
<tr>
<td>5.5.3</td>
<td>Collision of two plane-waves</td>
<td>105</td>
</tr>
<tr>
<td>5.6</td>
<td>Conclusions</td>
<td>106</td>
</tr>
<tr>
<td>6</td>
<td>Bogoliubov theory for atom scattering</td>
<td>109</td>
</tr>
<tr>
<td>6.1</td>
<td>Introduction</td>
<td>110</td>
</tr>
<tr>
<td>6.2</td>
<td>General properties of the scattered particles</td>
<td>111</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Bogoliubov equation for pair scattering</td>
<td>111</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Density and correlations</td>
<td>114</td>
</tr>
<tr>
<td>6.2.3</td>
<td>Number squeezing</td>
<td>115</td>
</tr>
<tr>
<td>6.2.4</td>
<td>Violation of the Cauchy-Schwarz inequality</td>
<td>116</td>
</tr>
<tr>
<td>6.2.5</td>
<td>Entanglement and interferometry</td>
<td>117</td>
</tr>
<tr>
<td>6.3</td>
<td>Application: twin-beam system</td>
<td>120</td>
</tr>
<tr>
<td>6.3.1</td>
<td>Scheme of the experiment</td>
<td>121</td>
</tr>
<tr>
<td>6.3.2</td>
<td>Theoretical description</td>
<td>121</td>
</tr>
<tr>
<td>6.3.3</td>
<td>Numerical results</td>
<td>123</td>
</tr>
<tr>
<td>6.4</td>
<td>Concluding remarks</td>
<td>130</td>
</tr>
</tbody>
</table>
The next experiment we will describe is one which shows one of the beautiful consequences of quantum mechanics. (…) And so you have the unusual result that when particles are identical, a certain new thing happens that doesn’t happen when particles can be distinguished.

Richard P. Feynman\textsuperscript{1}

\section*{1.1 Classical and quantum mechanics}

In the physics known by 1905, fundamentals of classical physics were captured in relatively few equations. With these equations, it was possible to comprehensively understand the realms of classical physics. First, is the full set of Maxwell’s equations describing electromagnetic fields in vacuum as well as their interactions with charges. In the mathematical formalism they include also the conservation of charge, energy and momentum. Next, in the classical physics, we have laws of motion expressing the relationship between forces acting on particles and their response to these forces. This is the Newton’s equation of motion which was then supplemented by Einstein’s relativistic modification. To fully comprehend the classical world, the gravitation has also to be taken into account. At this moment, the classical physics is complete.

Although it might seem that the understanding of Nature within the classical physics was established, there were problems which worried researchers at the break

of the 19th and 20th century. The first sign that something was wrong with fundamental equations appeared in statistical mechanics. The seminal paper on molecular theory of gases was published by Maxwell in 1859. On the basis of kinetic theory of gases he was able to explain a number of phenomena, like internal friction of gases, conduction of heat through a gas or the diffusion of one gas through another. Nevertheless, there was one problem which escaped from the description in the classical physics \[1\]. An explanation of temperature behaviour of specific heats of thermodynamic systems was in disagreement with experiments. Ten years later in the lecture Maxwell stated\(^2\) that “I have now put before you what I consider to be the greatest difficulty yet encountered by the molecular theory.” This was a direct statement that, by comparing the experiment with the rigorous theory, there was something fundamentally wrong in the classical physics and the problem lied in the theory of molecular structure of matter.

A second problem arose 30 years later, again in statistical mechanics, but this time with, what we now call, properties of a gas of photons. This one was resolved by Max Planck in the early years of the 20th century. The solution, together with Einstein’s explanation of photoelectric effect, was a step forward into the description of Nature within quantum mechanics.

During the first quarter of the 20th century scientists gradually accumulated information about the behaviour of particles at the microscopic level. During 1926 and 1927 Schrödinger, Heisenberg and Born obtained a consistent theory of behaviour of matter on the atomic level. Quantum mechanics, which was formulated that years, is the complete theory describing matter and light in all its aspects.

Quantum mechanics succeeded in the areas where old classical mechanics failed to explain the phenomena. It proved a useful framework to develop many areas of modern science, like properties of chemical reactions, description of motions of atoms, sub-atomic and elementary particles. Quantum mechanics also helped in finding a consistent description of theory of gases and condensed-matter physics. The problem of specific heat, which heralded the fall of classical world, was resolved. All these discoveries led to a significant development of many modern technologies.

A natural question that may be asked at this point is: what quantum mechanics brought with its birth? The old concept of separated domains of particles and waves had to be abandoned. In the new theory, these two worlds intertwine, yielding the phenomena in which particles behave like waves and the waves behave like particles. On the very small scale, the behaviour of things is not that anyone could have direct experience about, because human intuition applies only to macroscopic objects. One has to learn about quantum phenomena only through abstract mathematical formalism and by constantly applying it to specific problems. Only through gradual gathering of information about these phenomena one is capable to train the intuition and acquire the ability of gaining insight into quantum mechanics.

What are the concepts introduced by the new formulation of the laws of Nature,
1.1. Classical and quantum mechanics

captured by quantum mechanics and hidden from sight in classical mechanics? This
timey is mathematically complicated, especially if one deals with systems composed
of many particles, where the dimensions of abstract Hilbert spaces become expo-
nentially large, and any analytical and numerical treatments are out of reach. In
describing the phenomena, invoking the intuitions brought from classical grounds,
one sometimes can intuitively understand some behaviours of physical systems. It
is thus interesting to address the problem of a distinction between classical and
quantum aspects of phenomena. Is quantum theory always needed to describe or
understand the processes?

One of the most fundamental laws of quantum mechanics is the principle
of superposition of states. This is how P.A.M. Dirac in his book “The principles of
quantum mechanics” [2] describes its non-classical character:

The non-classical nature of the superposition process is brought out
clearly if we consider the superposition of two states, \( A \) and \( B \), such
that there exists an observation which, when made on the system in
state \( A \), is certain to lead to one particular result, \( a \) say, and when
made on the system in state \( B \) is certain to lead to some different result,
\( b \) say. What will be the result of the observation when made on the
system in the superposed state? The answer is that the result will be
sometimes \( a \) and sometimes \( b \), according to a probability law depending
on the relative weights of \( A \) and \( B \) in the superposition process. It will
never be different from both \( a \) and \( b \).

The intermediate character of the
state formed by superposition thus expresses itself through the probability
of a particular result for an observation being intermediate between the
corresponding probabilities for the original states, not through the result
itself being intermediate between the corresponding results for the original
states.

The superposition principle is used to explain the double slit experiment. The
particle can take one of the two alternative ways to reach a detector. If both slits are
open interference fringes are observed in the probability distribution of positions at
a distant screen. It is important to note one aspect of this scheme. The interference
effect occurs for a single particle, and hence this phenomena is referred to as the
single particle coherence. An analogue of this scheme is often employed to evaluate
the coherence of the cloud of atoms by splitting it into two halves (which corresponds
to a pair of slits), releasing from the trap and overlapping due to their expansions
(this corresponds to the detection at distant screen). If the interference effect is
revealed, it is possible to determine the coherence of the cloud.

The superposition principle can also involve states of many particles. The sim-
plest are the well known Bell states of a pair of two-level particles [3]. Denoting the
two levels with 0 and 1, we can write an example of the state as:

\[
|\psi\rangle = \frac{1}{\sqrt{2}} \left( |0\rangle_1 |1\rangle_2 + |1\rangle_1 |0\rangle_2 \right),
\] (1.1)
Chapter 1. Introduction

where the \( i \)th particle in the 0 or 1 state is denoted by either \( |0\rangle_i \) or \( |1\rangle_i \). The properties of such superposition can hardly be caught by human intuition, because of their absence in the classical physics. The only way is to consider specific physical situations to gradually acquire knowledge about these properties. One prominent example of a phenomenon is the Einstein-Podolsky-Rosen paradox \([4]\). The authors tried to show that quantum mechanics as a theory of physics is not complete. They considered the state of the form \( (1.1) \) in a situation where two particles are far apart. In the thought experiment they concluded that either super-luminal transfer of information is allowed or there are hidden variables that govern the outcomes of the measurements. John Bell in his famous paper \([5]\) found an expression, now called the Bell’s inequality, which has to be satisfied if a theory of hidden parameters may exist. He showed that the idea of additional variables is incompatible with the statistical predictions of quantum mechanics, and any theory which reproduces the outcomes of quantum mechanics has to be non-local. The non-locality, though, does not exclude causality \([6]\). These kind of correlations between particles led to many new applications, especially in quantum information theory and in quantum computing \([7]\).

There is still one more place where superposition is present, but is frequently left unnoticed, because it is hidden in the standard mathematical formalism of second quantization of quantum mechanics. Suppose that we have two particles, which are identical and cannot be distinguished by any means. In that case talking about identity of the particles makes no sense and the etiquettes that are attached to the particles, serving as a tool to number them, are physically meaningless. Quantum mechanics says that only physically allowed states of such particles are those in which superposition appears in a particular manner.

We shall illustrate this idea with the following example. Suppose that each particle can be in one of the two states, called modes and denoted with \( a \) and \( b \), and their wavefunctions are orthogonal. Let us consider the state in which first of the indistinguishable particles is in mode \( a \), denoted with \( |a\rangle_1 \), and the the second particle is in mode \( b \), which we denote as \( |b\rangle_2 \). As, in fact, we are not allowed to say which particle is which, the wavefunction of the composite system is given by:

\[
|n_a = 1, n_b = 1\rangle = \frac{1}{\sqrt{2}} \left( |a\rangle_1 |b\rangle_2 \pm |b\rangle_1 |a\rangle_2 \right),
\]  

(1.2)

where \( n_a \) and \( n_b \) are the occupation numbers of the two modes. The plus sign is for bosons and minus sign is for fermions. Wolfgang Pauli \([8]\) showed that the explanation, that for bosons we have to symmetrize and for fermions antisymmetrize the total wavefunction of particles, is deeply rooted in relativistic quantum mechanics. There, it is natural to employ the formalism of second quantization, in which the field operators are the fundamental objects. This formalism, brought into non-relativistic domain, turns out to be also very useful. However, in the second quantization formalism, the superposition present in the state \( (1.2) \) is hidden, because, when written in terms of the creation operators of the two modes \( \hat{a} \) and \( \hat{b} \), it
1.1. Classical and quantum mechanics

takes a particularly simple form:

\[ |n_a = 1, n_b = 1 \rangle = \hat{a}^\dagger \hat{b}^\dagger |0 \rangle, \]  

(1.3)

where \(|0 \rangle\) is the vacuum. In this formula a product of two operators appears, so one might think that the state is, in some sense, separable. It turns out, though, that the superposition present in Eq. (1.2) is important and has physical consequences.

The first indication that indistinguishability of particles has to be taken into account was recognized by Josiah Willard Gibbs, and after him named the “Gibbs paradox”. He noticed that the increase \(\Delta S\) in the entropy in the process of the mixing of two gases, which are initially separated but held at the same temperature and pressure, behaves unexpectedly. If the two separated gases are composed of different types of particles, the increase is \(\Delta S = N k_B \log 2\), where \(k_B\) is the Boltzmann’s constant and \(N\) is the total number of atoms in the container after the mixing. This value \(\Delta S\) is constant whatever the proportions, even negligible, of gases are. On the other hands, if the two gases are of the same type, there should be no increase of entropy, \(\Delta S = 0\). This discontinuity is the said paradox [9,10]. The paradox is averted if in the definition of entropy one takes into account the indistinguishability of particles of the same kind. Indistinguishable particles simply behave in a different way.

Another problem with indistinguishable particles arises as soon as one considers their scattering on the grounds of quantum mechanics. This can most easily be seen in an experiment in which the probability of scattering in various directions is measured [11]. For simplicity, assume that spins of the particles are the same. The amplitude of scattering at an angle \(\theta\) is denoted by \(f(\theta)\), and the sought probability is given by \(|f(\theta)|^2\). If the colliding particles are of different kinds, then the probability of measuring any particle scattered around the direction \(\theta\) is given by \(|f(\theta)|^2 + |f(\pi - \theta)|^2\). The particles are distinguishable in principle, but in the experiment we do not measure their types. However, if the particles are identical, quantum mechanics introduces a serious modification to this law. In that case, even in principle, the particles cannot be distinguished, and the superposition has to be invoked. As a consequence, the probability of observing the particle scattered in the direction \(\theta\) is given by \(|f(\theta) \pm f(\pi - \theta)|^2\). This amplitude has two signs, the upper referring to bosons, and the lower to fermions. A striking effect appears at the angle \(\theta = \pi/2\), for which many results are possible. For distinguishable particles we have \(2|f(\pi/2)|^2\), for bosons it is \(4|f(\pi/2)|^2\), twice as much as for distinguishable particles. For fermions the probability is zero at that angle. Therefore, the mere fact of being able to distinguish particles introduces important modifications to scattering amplitudes. We note here that for very small collision velocities the amplitude \(f\) (in this case called the s-wave) is independent of the angle, therefore the scattering cross-section for bosons is twice that of distinguishable particles, whereas fermions do not collide at all.

There is one more example where the indistinguishability plays an important role. The experimental scheme for the Hong-Ou-Mandel effect [3,12] is presented...
in Fig. 1.1a. Suppose that a photon is directed onto the beam-splitter which splits the trajectory of a photon into a superposition of two equally possible ways (50 : 50 beam-splitter). First, the photon can be reflected and then goes straight to detector $D_1$ with probability $p_1$. The other possibility is to go through the crystal unaffected and reach detector $D_2$ with probability $p_2$. The probability of detecting photon at $D_1$ or $D_2$, if the beam-splitter is symmetric, is, in both cases, $p_1 = p_2 = 1/2$. The same situation occurs if the photon is directed onto the crystal from the other side, which is depicted in Fig. 1.1b. Let us denote the respective probabilities with $q_1 = q_2 = 1/2$. What will happen if two photons are simultaneously directed onto the beam-splitter in the configuration shown in Fig. 1.1c? We assume here that the photons are prepared independently and never interacted before. If the two particles could possibly be distinguished, the two previous cases could be treated as independent events. Therefore, the probability of detecting two particles in detector $D_1$ would be $p_1q_1 = 1/2 \times 1/2 = 1/4$. Analogically, the probability of detecting two particles in detector $D_2$ would be $p_2q_2 = 1/4$. Finally, the probability of detecting any particle in detector $D_1$ and another in detector $D_2$ would be $p_1q_2 + p_2q_1 = 1/2 \times 1/2 + 1/2 \times 1/2 = 1/2$. Therefore, it is possible to obtain coincident detections between $D_1$ and $D_2$.

Now, an interesting thing happens, when the particles are indistinguishable. It turns out that in reality photons behave in a very different way from what we just described [13]. The problem is that the two alternative ways of detecting photons in $D_1$ and $D_2$ are indistinguishable and, therefore, has to be described by the superposition. When quantum mechanics of photons is used, the probability of coincidence (one photon is detected in $D_1$ and the other in $D_2$) vanishes. Once again, the mere fact of the particles being identical leads to non-trivial correlations.

The conclusion from the foregoing discussion is that the indistinguishability of particles leads to non-trivial correlations between them, even in cases when the
1.1. Classical and quantum mechanics

particles never interacted before. Of course, during the experiment one should give the possibility for the particles to loose their identity and show their identical nature. Only then the correlations can be manifested. If this was not so, in the calculations we should always include all the particles from the whole universe in order not to accidentally drop out some important correlations. This would lead to impossibility of constructing any reasonable physical theory. In case of the example with two photons, they showed their bosonic nature only when they overlapped on the beam-splitter. If the overlap between photons on the crystal was zero, the coincidences would built up, which was shown in the experiment \[12\]. This phenomenon can be used to precisely determine the arrival times on the beam-splitter up to the order of sub-picoseconds.

Now, we reach an important conclusion, that the non-trivial correlations that stem from indistinguishability can be utilized in some devices. This is natural, since the correlations that are revealed by quantum mechanics, and which were unnoticed before, can be useful in some sense to solve different tasks.

Of particular interests are the correlations that involve superposition of many-body states, because they lead to correlations between particles that cannot be explained within the classical theory. The notion of non-classical correlations will be precisely formulated in chapter 2 (see section 2.7). Now, we leave it on the intuitive level. We refer to such states, in which particles exhibit non-classical correlations, as particle entangled or, simply, entangled.

The entangled states are difficult to prepare in experimental schemes, which indicates how fragile they are. Any uncontrolled disturbance or disorder can lead to the drop of non-classical correlations. This might be the reason why we do not observe entanglement in everyday life and only after the discovery of quantum laws of Nature, this idea could be brought to light. However, it is this property of states that can be utilized in some particular manner. Suppose that an entangled state can be precisely controlled by a single parameter, say \(\theta\). It is natural to expect that this state is, in some sense, fragile. It may even happen that it is particularly susceptible to small changes of the parameter \(\theta\). In other words, if this is true, the change of the parameter \(\theta\) by a small amount \(\delta\theta\), can lead to large changes in the state, which cannot be described by a theory of classically correlated particles.

Note that the state of the system is not accessible directly, but is always rendered by some measurements. In quantum mechanics, we deal with probabilistic nature of phenomena, where outcomes of the measurements of physical observables are governed by some probability distributions. If one was able to distinguish two distributions, which describe the states that slightly differ in the value of \(\theta\), one could use this knowledge to infer the value of that parameter. This approach is related to metrology and, in particular, to the estimation of an unknown parameter with the highest precision possible. In the next chapter we analyse thoroughly this perspective starting from the idea of the distinguishability of probability distributions, basing on which we naturally introduce the metrological context and the problem of estimating an unknown parameter. It will turn out that the precision of
estimation is strictly connected to the distinguishability of states via the probability distributions, to the susceptibility of states, and, ultimately, to the presence of non-classical correlations. The class of useful states in metrology is called usefully particle entangled.

In the next section 1.2, we present briefly a few experiments that took this metrological approach. In some of them, entangled states were used in interferometers to infer a value of a parameter with classically unreachable precision.

1.2 Experiments with non-classical correlations

In this section we describe modern experiments with ultracold atomic systems. These systems are attractive from both fundamental and practical points of view, because a vast range of techniques were invented to precisely control their states. The invention of methods that enable the detection of single atoms, like microchannel plate detector [14], single-atom-resolved fluorescence imaging system [15] and light-sheet fluorescence method [16,17], paved the way toward direct measurement of correlations between the particles.

Below, we present the experiments in which the presence of non-classical correlations was observed in many-body atomic systems. There are essentially two schemes of practical implementations of procedures that lead to entangled states. One involves the coherent splitting of a trapped atomic cloud into two separated parts. Then, the interaction between atoms creates the non-classical correlations.

In the other scheme one starts with a single cloud of atoms. It serves as a source of emitted beams of atoms, because, due to a specific choice of the initial state, particles collide and scatter out of the mother-cloud. As the atoms are scattered in pairs, the non-classical correlations are the result of indistinguishability. We already noticed this phenomenon during discussion of the state (1.2). Such states can be employed for interferometry to obtain the precision of phase estimation, which is unattainable with classically-correlated states.

1.2.1 Particle-wave duality

In the first class of experiments a trapped atomic cloud is split into two parts. The non-classicality of the correlations between the particles results from atomic interactions.

The question remains how to detect entanglement. There is a criterion, which is related to the concept of particle-wave duality of matter. Suppose that after the splitting of the initial cloud, which in total consists of \( N \) atoms, we have two smaller parts, which we denote with \( a \) and \( b \). We can measure the number of particles in each part in every experimental run; denote them with \( n_a \) and \( n_b \). It is natural to expect that the two parts behave as two stable bulks of matter, so we demand that the fluctuations of the population imbalance, denoted with \( n = n_a - n_b \), should not be large. We quantify the fluctuations with the variance of \( n \). But what does “large” mean? It should be compared with a quantity that is natural in this system, which
is, of course, the total number of particles $N$. To understand it better, consider the following scenario. Suppose for the moment, that each of the $N$ considered atoms can be independently, with equal probabilities, put into the cloud $a$ or $b$. It is easy to show that in this case, the fluctuations of $n$, quantified with variance $(\Delta n)^2$ is exactly $N$. Therefore, the fluctuations of $n$ should be compared with total number of particles $N$; we quantify them with $\eta^2 = (\Delta n)^2/N$.

Until now, we treated the clouds as composed of particles. We may suspect, though, that the clouds can reveal wave nature. To observe it, the clouds are released from the traps and undergo a free-fall evolution. Their overlap can lead to build up of interference fringes in the particles density, which is analogous to the case of a double-slit experiment. If the fringes are clear, the clouds can be said to exhibit wave nature. The visibility of the fringes can be quantified with the parameter $\nu = (n_{\text{max}} - n_{\text{min}})/(n_{\text{max}} + n_{\text{min}})$, in the same manner as is done in optics, where $n_{\text{max}}$ and $n_{\text{min}}$ are the maximum and minimum values of density.

Now, we can introduce the criterion for entanglement. It can be demonstrated that if the fluctuations of the population imbalance are smaller than the visibility in the interference experiment, $\eta < \nu$, the state is particle entangled. Although the clouds behave as separate parts composed of independent particles, they show their wave nature. Note that a necessary condition for entanglement is $\eta < 1$, because the visibility is bounded from above, $\nu \leq 1$.

From $\eta$ and $\nu$ a single quantity $\xi_s = \eta / \nu$ can be constructed, which is known as the spin-squeezing. The condition for entanglement can now be rephrased as $\xi_s < 1$, and entangled states, detected by this criterion, are called spin-squeezed. In the next chapter, we prove in which situations $\xi_s$ is a measure of entanglement. It turns out that spin-squeezed states are useful for reaching a precision of the interferometric phase estimation that is unattainable by any classically correlated state.

The spin-squeezing for the first time was measured in an atomic system in 2008 [18]. The experiment starts with a trapped atomic cloud of $^{87}\text{Rb}$ which is in the state of Bose-Einstein condensation. After preparing the initial state, an optical lattice is raised at the proper pace, which leads to adiabatic distribution of particles over small number of wells. The two parts of the split clouds is the central pair (see Fig. 1.2a). The resolution in the experiment is so high that it allows to count the atoms in each well. The rate of the rising of the barrier should be such that the fluctuations of the particles between wells are reduced due to the collisions. It is not energetically preferable for particles to hop through the barrier. On the other hand, if the barrier is high so that the tunnelling between sites is reduced, the visibility of fringes drops in the interference experiment.

The next step is to calculate the spin-squeezing parameter. In order to obtain $\eta$, in-situ images are taken from which the fluctuations of the population imbalance are calculated. The histogram of occupation difference $n$ is presented in Fig. 1.2b in the bottom panel ($n$ is denoted there by $\Delta n$ and scaled by $\sqrt{N}/2$). The distribution proves that fluctuations of $n$ are reduced compared to the binominal distribution (red curve), which indicates that $\eta < 1$. 
To determine the visibility \( \nu \), which is the second ingredient of the spin-squeezing, the trap is turned off and the clouds are let to expand freely. When they overlap an absorption imaging system is used to determine the density, from which the visibility of fringes is calculated. Such density profile, in a single experimental shot, is presented in Fig. 1.2c. The bottom plot shows the histogram of the relative phase between the clouds, from which \( \nu \) can be deduced. The best observed spin-squeezing was equal to \( \xi^2_s = -3.8 \text{ dB} = 0.64 < 1 \), which proved that the atoms in the system were non-classically correlated.

A similar approach was also taken in [19] to reach \( \xi^2_s < 1 \). In the experiments we just described, spin-squeezing of motional degrees of freedom was measured. The analogical quantity can be constructed for internal states of particles, like two spin components. In this case, \( \eta \) quantifies fluctuations of the population imbalance of the two components, and \( \nu \) describes the coherence between them.

There are numerous experiments with internal states in which spin-squeezed states were observed. In [20] quantum nondemolition measurement was used to determine that \( \xi^2_s = -3.4 \text{ dB} \) in a large sample of caesium atoms. In [21] observation of \( \xi^2_s = -2.5 \pm 0.6 \text{ dB} < 1 \) was reported in the system of \(^{87}\text{Rb}\) atoms trapped on an atom chip.

The fact that \( \xi_s < 1 \) indicates that the state is useful for metrology was utilized in [22]. Here, an interferometer was constructed which operated at the precision that is out of reach if the system was only classically correlated. However, the notion of usefulness in metrology is more fundamental than the spin-squeezing. In [23] a state of \(^{87}\text{Rb}\) Bose-Einstein condensate was prepared, which, despite having \( \xi_s \geq 1 \), benefited from non-classical correlations in ultra-precise interferometry.
1.2. Experiments with non-classical correlations

In the mentioned experiments, correlations between particles were created by atomic interactions between two separated (externally or internally) ultracold clouds. The demonstration that the correlations are of quantum nature reduced to showing that the visibility of the interference fringes is higher than the reduction of fluctuations between populations of the clouds. In other words, in order to show the non-classical nature of correlations, it was enough to demonstrate that $\xi_s < 1$. Such spin-squeezed states can be also used in interferometers to yield precision that is beyond classical limit. In the next section, we present experiments in which spin-squeezing cannot be used as a signature of entanglement, but the sub-shot noise interferometry survives as the criterion of entanglement.

1.2.2 Scattering of indistinguishable particles

In this section, we present experiments in which many-body entangled states were created by scattering pairs of atoms. Here, the non-classical correlations result from indistinguishability of the colliding particles. This kind of states can be used to surpass classical limits placed on a precision of estimation in interferometers. When the interferometric phase $\theta$ is estimated from a series of measurements, its uncertainty $\Delta \theta$ cannot be lower than a certain value $\Delta \theta_{\text{sn}}$, called the shot-noise limit, provided that only classical correlations are present in the system. The ability to overcome this bound, when $\Delta \theta < \Delta \theta_{\text{sn}}$, demonstrates that the state, apart from exhibiting quantum correlations, is useful for precise metrology.

Experiment with twin matter waves. The creation of entangled states by the scattering was realized in the experiment reported in [24]. It started from the preparation of $2.8 \times 10^4$ optically trapped atoms in a $^{87}\text{Rb}$ Bose-Einstein condensate in a hyperfine state $F = 2, m_F = 0$ (see Fig. 1.3A). The dynamics in the cloud is such that a pair of atoms from the state $m_F = 0$ can collide, and, as a result, one atom is scattered into state $m_F = 1$ and the other into $m_F = -1$. As time passes, more and more pairs are scattered leading to gradual accumulation of atoms in the states $m_F = \pm 1$.

Although the distribution of the total number of atoms may be broad, the atoms are always scattered in pairs. As a consequence, the fluctuations in the difference between occupations of the states $m_F = \pm 1$ should be zero. In fact, this is confirmed in the experiment. To measure the populations of the states, atoms are released from the trap. After that, a current in coils creates a gradient of magnetic field, which is applied to free-falling cloud. This leads to spatial separation of hyperfine components (see Fig. 1.3B), which is analogous to the Stern-Gerlach experiment. Finally, absorption imaging is used to detect the population of the three clouds (registered within white circles in Fig. 1.3C). The analysis of the outcomes demonstrates that the fluctuations of population imbalance is very small and limited only by the detection efficiency. This shows that, indeed, atoms are scattered in pairs.

The entanglement between two scattered atoms comes from the fact that they are identical, so the total wavefunction has to be symmetrical, exactly as in the case
Figure 1.3: The experimental scheme of the interferometer operating with sub-shot-noise precision on internal states of ultracold atoms. (A) Initially atoms in BEC are prepared in an optical dipole trap in the hyperfine state $F = 2, m_F = 0$. (B) To measure the populations of the states $m_F = 0, \pm 1$ atoms are released from the trap. The coils create the gradient of magnetic field, which spatially separates clouds with different $m_F$, in analogy with the Stern-Gerlach experiment with spins. (C) Absorption imaging technique is used to infer the populations of the split clouds (in white circular regions). (D) The chronology of the interferometric scheme. (1): Scattering of atoms populating the states $m_F = \pm 1$. (2): and (4): Coupling the states $|2, -1\rangle$ with $|1, 0\rangle$. (3): Coupling the states $F = 2, m_F = 1$ and $F = 1, F = 1, m_F = 0$ via microwave pulse of variable duration $\tau$, which maps on the interferometric phase $\theta$. All pictures are taken from [24].

of two particles, which is shown in Eq. (1.2). To demonstrate that the particles in the system are non-classically correlated, an interferometer was constructed that benefited from the expected non-classicality.

The interferometric sequence consists of four stages, pictured in Fig. 1.3D. First, pairs of atoms are populating states $|F = 2, m_F = \pm 1\rangle$. Second, a microwave pulse is used to couple states $|2, -1\rangle$ with $|1, 0\rangle$. Then, a pulse of variable duration $\tau$ is used to couple states $|1, 0\rangle$ to $|2, 1\rangle$. Finally, a pulse coupling $|2, -1\rangle$ and $|1, 0\rangle$ is used, after which the populations of $m_F = \pm 1$ is measured. The duration of the pulse $\tau$ maps on the interferometric phase $\theta$, and the sequence of microwave pulses effectively mixes the states $F = 2, m_F = \pm 1$, the time $\tau$ being the duration of mixing.

By measuring the statistics of the difference in the populations of $m_F = \pm 1$ for every $\theta$, it is possible to find the uncertainty $\Delta \theta$. In the experiment the best achievable sensitivity was demonstrated to be $\Delta \theta/\Delta \theta_{\text{sn}} = 0.83(0.1) < 1$, which proves that the particles in the state are entangled and useful for sub-shot noise interferometry.
1.2. Experiments with non-classical correlations

Figure 1.4: The momentum view of the halo of scattered atoms formed in the process of the collision of two Bose-Einstein condensates. (a) Each dot represents a detected atom. The black dots are the atoms on the halo, and yellow-orange structures at the poles are the two BECs. (b) The view of division of the halo into disjointed zones. The red regions situated on the opposite sides exhibit sub-Poissonian fluctuations. All pictures are taken from [25].

Experiment with colliding condensates. Another way of obtaining pairs of particles atoms is by colliding two atomic clouds. This approach was used in [25, 26]. The experiment starts with a single Bose-Einstein condensate of metastable helium atoms confined in an optical dipole trap. The application of two counter-propagating laser beams splits the cloud into two parts moving in opposite directions with large velocities. As a result of a collision, a pair of particles, each from a different cloud, is scattered outside the source. A halo of scattered atoms is formed due to the accumulation of such events. Next, all the particles fall freely to reach a micro-channel plate, where their positions are detected. The single-atom registration system allows for a three-dimensional reconstruction of momenta of the particles, an example of which is shown in Fig. 1.4a.

To demonstrate that atoms are scattered in pairs, the experimenters showed that the fluctuations of the numbers of atoms with opposite momenta are sub-Poissonian, \( \eta < 1 \). To this end, they divide the halo into a number of zones and count the detected atoms that fall into regions that are situated in the opposite sides of the halo (red in Fig. 1.4b). The observed suppression of fluctuations gives \( \eta^2 \approx 0.87(2) \), which demonstrates the sub-Poissonian statistics. This is contrasted with the value of \( \eta^2 \approx 1 \) that is calculated for pairs of zones that are not expected to be correlated (any pair from blue zones in Fig. 1.4b).

It remains to be proved that non-classical correlations are present in the system. One approach is to construct an interferometer and show its sub-shot-noise performance. An alternative way, adapted by the experimentalists, was to use
the non-classicality criterion based on the violation of Cauchy-Schwarz inequality. In its most basic form, it states that the cross-correlation between two quantities $I_1$ and $I_2$ cannot exceed geometric mean of auto-correlations in each quantity, $|\langle I_1 I_2 \rangle| \leq \sqrt{|\langle I_1^2 \rangle \langle I_2^2 \rangle|}$. In other word, this inequality may be reformulated with help of $C$:

$$C \equiv \frac{|\langle I_1 I_2 \rangle|}{\sqrt{|\langle I_1^2 \rangle \langle I_2^2 \rangle|}} \leq 1.$$  

(1.4)

The violation of the Cauchy-Schwarz inequality would indicate that the system evinces a sort of non-classical behaviour. The stronger than classical correlations, $C > 1$, were demonstrated in quantum optics with antibunched photons produced by spontaneous emission, twin-photon beams generated in a radiative cascade, parametric down conversion and four-wave mixing (see references in [26]). In quantum optics it is understood that the Cauchy-Schwarz inequality signals non-classicality of light. Such states cannot be described by statistical mixture of coherent states, i.e. the states that resemble most closely the properties of classical light. However, it is not clear what meaning should be attributed to $C$ in domain of atomic physics. We address this question in more details in chapter 4, where we show that $C > 1$ proves that correlations between particles in a system of bosons are non-classical.

In the experiment with colliding condensates, $C$ is constructed from objects that quantify correlations between opposite zones (cross-correlation) and within these zones (auto-correlations). Depending on the number of division of the halo, $C$ is varying from almost equal 1 up to $1.088(6)$. Interestingly, the violation is very small and the maximum value obtained is reached for the about 600 division of halo. For such a fine partition the number of scattered particles in a pair of zones is also small.

**Experiment with pairs of atoms generated from BEC in optical lattice.**

To overcome the limitation of the previous configuration a different experiment was arranged [27]. In this scheme, an elongated Bose-Einstein condensate of helium atoms is prepared in a one dimensional optical lattice, which serves as a mean to modify the dispersion relation $\varepsilon(k)$ of atoms. Now, if two atoms from the condensate, having the same wavevector $k_0$, collide, they can be scattered out from the initial cloud resulting in a change of their wavevectors. The presence of the lattice allows to satisfy the energy and momentum conservation laws in one dimension, which otherwise is forbidden. As a result, the source emits two beams of atoms with wavevectors $k_1$ and $k_2$.

This reasoning was employed in an experiment, in which the system was prepared in an unstable state. The wavevector $k_0$ of the cloud was set to such a value that initiated the dynamics in which $k_1$ and $k_2$ are populated. By varying the time for which the lattice was on, it is possible to obtain populations up to $10^3$ atoms per beam. An example of a single experimental outcome is shown in Fig. 1.5, which presents two large distinct peaks centred around $k_1$ and $k_2$ rising from a constant background. These are the beams that should manifest non-classical correlations. It was shown that the fluctuations of population difference, between two selected
1.2. Experiments with non-classical correlations

Figure 1.5: A single shot of the integrated (over $x$ and $y$ directions) momentum distribution of atoms after the decay of an initial Bose-Einstein condensate. The main peak $k_0$ is the wavevector of the initial cloud, while $k_1$ and $k_2$ are the two emitted beams. The wavevectors on the horizontal axis are given in the units of $k_{\text{rec}}$ which is the characteristic magnitude related to depth of the lattice. The graphics is taken from [27].

regions close to the beams, are sub-Poissonian, $\eta^2 = 0.91(4) < 1$. Whether non-classical correlations are present in the system remains an open question.

Figure 1.6: The schematic representation of excitation of the cloud and formation of the twin-atom beams. (a) The view of the $y$-axis mode. The initial cloud (red) is excited into the first excited state (blue) by shaking the trap in a particular manner (inset). Two atoms from the excited cloud de-excite into the ground state transferring the energy into the motional states in the $x$-direction. (b) After the emission, the cloud is released from the trap and imaged after a long period of free fall expansion. Transversely excited state is shown in the central part of the figure (blue). The twin-atom beams are centred around wavevectors $\pm k_0$ (red). All pictures are taken from [28].
Experiment with twin-atom beams. The system of an unstable condensate in an optical lattice is an example of a scheme applied in production of large correlated beams. It is based on the process of scattering particles out from a source into two well separated regions. An experiment that was conducted in Vienna [28] is very important in this context.

The $^{87}$Rb Bose-Einstein condensate is trapped in a tight waveguide potential with a shallow axial harmonic trap (the $x$-direction) confinement. The apparatus allows for precise control of the position of the trap in the $y$-axis direction. By shaking the trap in a particular manner, the inversion of population is obtained, in which atoms are occupying the first excited energy level in the $y$-direction (see Fig. 1.6a). In the excited cloud, due to collisions, a pair of atoms can de-excite into the ground state. As a consequence, the excess of the energy is transformed into motional states. The restricted geometry of the system forces atoms to be scattered in opposite directions along $x$-axis. The dynamics leads to emission of twin-atom beams that are intrinsically correlated in their number of atoms.

The time-of-flight fluorescence imaging is used to register the positions of atoms after long time of free expansion. This allows to observe the momentum distribution of particles, which shows two peaks centred around $\pm k_0$ (red in Fig. 1.6b). The single-particle resolution permits determination of fluctuations in the difference between populations of the two beams yielding $\eta^2 = 0.11(2)$. This result proves that particles are populated in pairs in the two beams. An open question remains about the presence of the non-classical correlations, which are expected from the indistinguishability in the same manner as in the experiment with twin matter waves [24] (see also page 11).

1.3 Conclusions

This finishes our discussion on the distinction between classical and quantum correlations. In this chapter, we presented the concept of superposition of states, which is an intrinsic property of quantum mechanics, absent in classical physics. The superposition was a basis for the idea of non-classical correlations between particles. We presented a number of contemporary experiments in which non-classical correlations were demonstrated. To prove that the states were entangled, the context of metrology was used.

The physical principle behind the metrology is that entangled states are fragile and volatile. Therefore, small changes in a parameter describing the system (an interferometric phase) leads to large changes in the state. This property can be used in interferometers, to obtain classically unattainable precision. On the deeper level, this perspective also touches the problem of distinguishing two states of the system. Only if such discrimination of two neighbouring states is possible, they can be useful in metrology.

All the states in the experiments described above were prepared in a dedicated procedure. Highlighting the underlying idea, that indistinguishability of particles
constituting the system may lead to non-classical correlations, may indicate a different way of preparing systems that could be used for ultra-precise interferometry.

1.4 Structure of thesis

The thesis covers a wide range of problems, which are structured in the following order:

- In chapter 2, fundamental concepts are introduced upon which the thesis is based. First, an idea of states distinguishability is presented, together with the tools to measure to what extent two states are different. It is shown that for a large class of measures an object called the classical Fisher information emerges in a natural way, when considering two neighbouring states. Next, a metrological approach is taken to show that the Fisher information plays a central role in the estimation theory, limiting the best achievable precision of estimating an unknown parameter. The object called the quantum Fisher information is defined as the maximum of the classical Fisher information over all possible measurements allowed by quantum laws of nature. Further in this chapter, the concept of non-classical correlations between particles (particle entanglement) is precisely defined. A theorem is presented that relates the value of the Fisher information to the presence of entanglement and usefulness of a state for sub-shot noise interferometry. Basing on a simple example of a two-mode bosonic system, the spin-squeezing is shown to satisfy conditions for a particle entanglement criterion.

- In chapter 3, simple examples are considered to better understand the concept of the classical and quantum Fisher information, introduced in chapter 2. The optimal measurement strategies, for which one obtains the highest precision possible, are identified for simple interferometric scenarios.

- In chapter 4, we address the question raised in the introducing chapter of the thesis (see discussion on page 14), whether and in what sense the Cauchy-Schwarz inequality (see Eq. (1.4)) is a signature of non-classicality. It is shown that violation of this inequality, after a slight refinement, proves the presence of particle entanglement in a many-body system of bosons. This criterion applies to any quantum system of bosons with either a fixed or a fluctuating number of particles, provided that there is no coherence between different number states.

- In chapter 5, it is examined what influences the usefulness of collisions of Bose-Einstein condensates as a potential source of entangled atomic pairs. This is done by analysing the reduction of fluctuations in the number difference between opposite regions on the halo of scattered atoms. It is shown that the spread in the momentum of initial clouds is the leading limitation, destroying the reduction of fluctuations as the density of scattered pairs grows. It
is indicated that scattered pairs are particularly useful in configurations in which atoms are scattered not uniformly on the halo, but rather into two well-separated regions.

- In chapter 6, the idea that the scattering of atoms into two well-separated regions might give a usefully entangled state, is further developed. The Bogoliubov theory is applied to find the full dynamics of a pair-production process, together with the first and second order correlation functions. Also, it is shown that there are no fluctuations of atom number difference between the regions. The quantum Fisher information is employed to show that the system exhibits non-classical properties, that can be used in ultra-precise sub-shot noise interferometry. The theory is applied to the “twin-atom beams” experiment [28] (see also page 15).

- Finally, in chapter 7, we tackle the problem whether two independently prepared condensates can be useful for metrology. This question is based on the observation that pairs of particles are non-classically correlated only due to indistinguishability of bosons. It is shown that quantum correlations, arising from the sole fact that the particles are identical, are a sufficient resource for sub-shot-noise interferometry. The measurement of the number of atoms in the Mach-Zehnder interferometer is used to show that the system operates with non-classical precision. This result shines new light on the approach to correlating particles by scattering. Highly entangled state are already at hand, but new technological tools has to be developed for shaping the distribution of each clouds alone.

- Various technical explanations and derivations are included in the appendices. Usually, a chapter is supplemented by a proper appendix.
In this chapter we describe the concept of distinguishability of physical states. The analysis will be based on the observation that all systems are described in terms of measurable physical quantities and any description must include these quantities, which are accessible to the observer. The issue of the states’ distinguishability is encountered in statistics in the context of the theory of probability. As we show below, it reduces to the following question: to what extent two probability distributions are different?

In the next sections, we develop some simple tools which quantify the degree of distinguishability. The common property of these measures is that they can be unified within a single quantity called the Fisher information. This is true, when the two states are, in some sense, close to each other. The Fisher information is the central object of this thesis, thoroughly analysed in the following chapters.

The problem of distinguishing the probability distributions can be regarded from

## Chapter 2

### Distinguishability of states

|   |   |
|---|---|---|---|---|---|---|---|
| 2.1 | States distinguishability | 20 |
| 2.2 | Fidelity as a measure of distinguishability | 23 |
| 2.3 | Classical Fisher information | 28 |
| 2.4 | Other measures of distinguishability. \( f \)-divergences. | 29 |
| 2.4.1 | Kolmogorov distance or trace distance | 29 |
| 2.4.2 | Hellinger distance | 30 |
| 2.4.3 | Kullback-Leibler divergence or relative entropy | 31 |
| 2.4.4 | Generalization to the \( f \)-divergence | 32 |
| 2.5 | Properties of the Fisher information | 35 |
| 2.6 | Metrology | 36 |
| 2.6.1 | Maximum-likelihood estimation | 38 |
| 2.6.2 | Estimation from the mean | 40 |
| 2.7 | Quantum Fisher information and quantum correlations | 43 |
| 2.7.1 | Optimal measurement. Quantum Fisher information | 44 |
| 2.7.2 | Particle entanglement | 48 |
| 2.7.3 | Two mode interferometry with bosons | 53 |
| 2.7.4 | Spin-squeezing and number-squeezing | 55 |
| 2.8 | Conclusions | 57 |
another point of view, which leads to a different interpretation of the Fisher information. In this new approach, one assumes that the state of the system, which is initially well known, undergoes some evolution. It is the Fisher information that describes how the state has changed. In other words, it quantifies the susceptibility of the state in response to some action performed on the system.

If the Fisher information is large, so is the susceptibility of the state. In fact, for some types of evolution (such as generated by interferometers), high values of the Fisher information can be linked with the presence of the non-classical correlations in the initial state \[29\]. This means that the Fisher information acquires another important property – it measures the amount of “useful” entanglement.

The ubiquity of the Fisher information suggests it is a powerful and fundamental tool for many mathematical and physical seemingly non-related problems. This observation triggered our interest and provided the motivation for an extensive research presented in this thesis.

This chapter is organized as follows. In section 2.1 we define the problem of states distinguishability as a problem of distinguishing classical probability distributions. In section 2.2 we introduce fidelity as the first measure that quantifies to what extent two distributions are different. In section 2.3 we introduce the classical Fisher information. In section 2.4 we present a few more measures which can be used instead of fidelity to distinguish distributions. These include the Kolmogorov distance, the Hellinger distance and the Kullback-Leibler divergence. Then, we argue that all these measures can be generalized to the quantity called the $f$-divergence. In section 2.5, basing on the knowledge of the fidelity and the $f$-divergence, properties of the Fisher information are inferred in a simple way. In section 2.6 a metrological approach is taken to the problem of distinguishability of states. It is shown that the Fisher information bounds the maximum attainable precision in the process of estimation of an unknown parameter. In section 2.7 we introduce the concept of an optimal measurement, which leads to the object called the quantum Fisher information. In the same section, we juxtapose the separable against the particle-entangled states and show that the value of the quantum Fisher information is related to the non-classical correlations. Also here, we consider a simple example of two-mode interferometry with bosons. Basing on these results, a simple criterion for particle entanglement, known as spin-squeezing, is presented and its natural metrological interpretation is highlighted. Section 2.8 contains conclusions of this chapter.

2.1 States distinguishability

The problem how to distinguish two states can be formulated on the ground of classical mechanics. Here, we introduce this concept by analysing the following example. Consider two particles of the same species with momenta $p_1$ and $p_2$ in one dimension. In classical mechanics, if the detectors are perfect, these momenta can be measured with arbitrary accuracy. We may thus say that if the two momenta are the same, so is the state of the two particles; if they differ, the states are different.
2.1. States distinguishability

as well. This means that the two states, via the measurements of the momenta, can be perfectly distinguished.

The situation becomes more complex if the two momenta are random variables distributed according to probability distributions $P_1(p_1)$ and $P_2(p_2)$. If we measure the momenta of the particles, the outcomes in the first shot of the experiment are $p_1^{(1)}$ and $p_2^{(1)}$. However, basing on these results it is not legitimate to say that the states of the particles are different if $p_1^{(1)} \neq p_2^{(1)}$ and are the same if $p_1^{(1)} = p_2^{(1)}$. This is because the relation between the single outcomes may be accidental. Even if the momenta are different in one experimental shot, then in second run of the measurement it may happen that $p_1^{(2)} = p_2^{(2)}$. We cannot judge about distinguishability of the states solely on the outcomes of the measurement. Instead, we must refer to the probabilities $P_1(p_1)$ and $P_2(p_2)$, and say that they describe the states of the particles. In this way, the problem of distinguishing the states is reduced to the problem of distinguishing the two distributions $P_1$ and $P_2$.

To illustrate this approach, consider the following problem encountered in the statistical mechanics. Take two containers of volume $V$ filled with gas with a great number $N$ of particles, each of mass $m$. The temperatures of the two containers are $T_1$ and $T_2$. How can we distinguish the two thermodynamic states? Stated in other words, how can we distinguish the two temperatures?

The first attempt would be to take a thermometer, a macroscopic object, and put the first (second) system in contact with the thermometer. Assuming infinite accuracy, the thermometer would yield $T_1$ ($T_2$). The comparison of the two temperatures, in principle, would result in the statement whether the two states of equilibrium are the same or different. However, difficulties related to this approach arise: in general the accuracy of the thermometer is limited; also, the thermalization may lead to heating of the system, which could result in the change of the measured temperatures.

Let us now investigate a different procedure. A far less invasive measurement would be to pick up a single particle from each container, and measure its momentum, which in principle, can be done. The two outcomes $p_1$ and $p_2$, according to statistical mechanics, are distributed according to the Maxwell’s probability distributions

$$P_1(p_1) \propto \exp\left[\frac{-p_1^2}{2mk_BT_1}\right] \quad \text{and} \quad P_2(p_2) \propto \exp\left[\frac{-p_2^2}{2mk_BT_2}\right]. \quad (2.1)$$

Thus, distinguishing between the temperatures is equivalent to distinguishing between the two distributions (2.1), such as those pictured in Fig. 2.1. Note that $P_1$ and $P_2$ differ in the fact that for the former, the histogram of the outcomes is more peaked around zero than for the latter. Consequently, the outcomes with large values of $|p|$ are more frequent for $P_2$ than for $P_1$. There is also a small probability of having very large $|p|$ for $P_1$, whereas for $P_2$ this event is highly improbale.

This simple example provides a general framework for the state discrimination in the classical scheme. However, our aim is to transfer these concepts to the realm of the quantum mechanics. This, as we argue below, requires more subtle
The physical states are described by density matrices $\hat{\rho}$, operators that are non-negative with unit trace. Any probability distribution that governs the statistics of outcomes in some measurement is generated from a density matrix $\hat{\rho}$ according to the following formula:

$$p_n = \text{Tr} \left[ \hat{\rho} \hat{E}_n \right].$$

The operators $\hat{E}_n$ characterize the procedure of measurement process with the corresponding $n$th outcome. These operators, forming a set called positive-operator valued measure (POVM), have the following properties. First, the mean of any $\hat{E}_n$ is always non-negative (positive-semidefinite operator), and secondly, they sum up to identity:

$$\sum_n \hat{E}_n = \mathbb{1}.$$  

Consider now two systems in the states represented by density matrices $\hat{\rho}_1$ and $\hat{\rho}_2$. How can we distinguish these two states? The direct access to the operators $\hat{\rho}_1$ and $\hat{\rho}_2$ is not possible, but one is capable of performing measurements on the systems. The probabilities of the outcomes give information about the states. In this way, one is forced to reduce the problem of distinguishing physical states to distinguishing probability distributions. When the measurement is fixed, the outcomes are governed by the two distributions

$$p_n^{(1)} = \text{Tr} \left[ \hat{\rho}_1 \hat{E}_n \right] \quad \text{and} \quad p_n^{(2)} = \text{Tr} \left[ \hat{\rho}_2 \hat{E}_n \right].$$
It should be emphasized that comparison of the two distributions makes sense only if the same measurement is performed on both systems, so \( p^{(1)} \) and \( p^{(2)} \) are generated by the same set \( \hat{E}_n \). However, it is reasonable to change that measurement simultaneously for both systems, as it may give more information about the states. For, let us take different measuring operators and call them \( \hat{\Pi}_m \), with \( m = 0, 1, 2, \ldots \). Now, the respective probabilities are given by:

\[
q_m^{(1)} = \text{Tr} \left[ \hat{\rho}_1 \hat{\Pi}_m \right] \quad \text{and} \quad q_m^{(2)} = \text{Tr} \left[ \hat{\rho}_2 \hat{\Pi}_m \right].
\] (2.5)

It may happen that \( p \)'s for measurement \( \hat{E}_n \) are less distinguishable than \( q \)'s for measurement \( \hat{\Pi}_m \). This will depend on the chosen measure of distinguishability (discussed in detail below) and on the considered states. In this way, we are led to the concept of the optimal measurement, for which two probability distributions are distinguishable the most. We address this idea in section 2.7.1, and further consider in more details in chapter 3.

It is worth mentioning that the idea of combining two simultaneous measurements \( \hat{E}_n \) and \( \hat{\Pi}_m \), in the hope of better distinguishability, is in general forbidden in quantum mechanics. This fact is strictly related to the noncommutativity of the operators representing the physical observables. It is allowed, though, to consider joined measurements of a set of commuting observables, which in general leads to better states discrimination.

We indicated the problems that arise when the ideas of differentiation of states are transferred from classical to quantum realms. Now, we present the first object that will quantify to what extent two probability distributions are different.

### 2.2 Fidelity as a measure of distinguishability

In this section we focus on the problem of the distinguishability of probabilities by introducing a simple measure. All of the forthcoming discussion is based entirely on the standard probability theory, quantum mechanics is not needed for the the moment.

Below we present a measure called the fidelity, denoted by symbol \( f \). Next, we consider an operational approach to differentiation of distributions founded on statistical ground. We shall see that, in a particular manner, it is related to the number of distinguishable states in the space of probability distributions. Further on, we consider two states that are not far apart in this space. In that case, the fidelity \( f \) can be reduced to one parameter called the Fisher information, denoted by \( F_{\text{cl}} \). As FI appears in the context of probability distribution it conventionally acquires attribute “classical” and so the subscript “cl” is placed in the symbol.

Let us consider two probability distributions \( p_n^{(1)} \) and \( p_n^{(2)} \), where the outcomes of some physical quantity \( X \), which is a random variable, are enumerated with index \( n = 1, 2, \ldots \). In other words, if the outcomes are governed by \( p_n^{(1)} \) (or \( p_n^{(2)} \)) the probability of obtaining the realization \( x_n \) of variable \( X \) is given by \( p_n^{(1)} \) (or \( p_n^{(2)} \)). For the sake of simplicity, we, henceforth, consider only outcomes that are countable,
but the inclusion of continuous distributions into the formulas is straightforward. In fact, some of the formulas we shall see below are most easily expressed in the continuous case.

The fidelity is given by the following formula:

$$f\left(p^{(1)}|p^{(2)}\right) = \sum_n \sqrt{p^{(1)}_n p^{(2)}_n}. \quad (2.6)$$

If the two probability distributions are identical, $p^{(1)}_n = p^{(2)}_n$, then $f = 1$. However, if the probabilities are maximally distinguishable, in which case the outcome possible in one state is improbable in the second, the probabilities have disjoint supports. Consequently, for every $n$ the product $p^{(1)}_n p^{(2)}_n$ vanishes, and the fidelity is zero. In fact, it can be shown that $f$ is always between zero and one:

$$0 \leq f\left(p^{(1)}|p^{(2)}\right) \leq 1. \quad (2.7)$$

The first inequality is achieved only for completely distinguishable probabilities, and the second is saturated if and only if they are the same. The second inequality can be simply proved with help of the Cauchy-Schwarz inequality:

$$\left|\sum_n a_n b_n^{*}\right|^2 \leq \sum_m |a_m|^2 \sum_k |b_k|^2. \quad (2.8)$$

Inserting here $a_n = \sqrt{p^{(1)}_n}$ and $b_n = \sqrt{p^{(2)}_n}$, and upon using the normalization condition, $\sum_n p^{(1)}_n = \sum_k p^{(2)}_k = 1$, we obtain upper bound shown in (2.7). As a corollary, we get that the equality is achieved only when $a_n = cb_n$. Inserting here the respective probabilities, and invoking the normalization condition once again, we conclude that $f = 1$ if and only if $p^{(1)} = p^{(2)}$.

The fidelity is related to the Bhattacharyya-Wootters distance [30–32] between two distributions, and is given by the formula

$$B\left(p^{(1)}|p^{(2)}\right) = \arccos\left(f\left(p^{(1)}|p^{(2)}\right)\right) \quad (2.9a)$$

$$= \arccos\left(\sum_n \sqrt{p^{(1)}_n p^{(2)}_n}\right). \quad (2.9b)$$

The distance function $B$ has all the properties of a metric, i.e. non-negativity, symmetry, it satisfies triangle inequality and coincidence axiom. The formula (2.9b) can lead to another interpretation of the formula. It can be viewed as the angle between two normalized vectors $v_1 = \left(\sqrt{p^{(1)}_1}, \sqrt{p^{(1)}_2}, \ldots\right)$ and $v_2 = \left(\sqrt{p^{(2)}_1}, \sqrt{p^{(2)}_2}, \ldots\right)$ with the usual Euclidean scalar product. Thus, it can alternatively be written as $\cos B = v_1 \cdot v_2$.

A deeper interpretation of this metric was given by Wootters in [30]. Consider the following problem of estimating the probabilities $p_i$. The experiment gives only access to the number of occurrences of $i$th outcome after $m$ trials. We denote these
2.2. Fidelity as a measure of distinguishability

Frequencies by $\xi_i$; in the limit of very large $m$, $\xi_i$ approaches $p_i$. If the number of trials is finite but large, the frequencies $\xi_i$ are distributed according to a Gaussian given by:

$$P(\{\xi_i\}) \sim \exp \left[ -\frac{m}{2} \sum_i \frac{(\xi_i - p_i)^2}{p_i} \right].$$ \hspace{1cm} (2.10)

From this relation one can define the region of uncertainty in the probability space around the point $p = (p_1, p_2, \ldots)$. This region is the set of all points $(\xi_1, \xi_2, \ldots)$, for which the exponent in (2.10) is, in absolute value, less or equal to $\frac{1}{4}$. Two points $p$ and $p'$ are said to be distinguishable in $m$ trials if their regions of uncertainty do not overlap. For large $m$, this will be the case if

$$\sqrt{m} \left\{ \sum_i \frac{(p_i - p'_i)^2}{p_i} \right\}^{1/2} > 1.$$ \hspace{1cm} (2.11)

Having defined the regions of uncertainty, Wootters presents the definition of the statistical distance along arbitrary curve between two probabilities $p^{(1)}$ and $p^{(2)}$. It is proportional to the maximum number of mutually distinguishable points along the curve $\ell$ (in probability space), divided by $\sqrt{m}$, in the limit as $m \to \infty$:

$$d(p^{(1)}, p^{(2)}|\ell) = \lim_{m \to \infty} \frac{1}{\sqrt{m}} \left\{ \text{maximum number of mutually distinguishable intermediate points along the curve } \ell \text{ in } m \text{ trials} \right\}.$$ \hspace{1cm} (2.12)

The number of distinguishable points along the curve is of the order of $\sqrt{m}$, thus the factor $1/\sqrt{m}$ is included so that the limit is finite. The statement “along the curve $\ell$” means, that there is a function $p(t)$ of variable $t$ that is differentiable for $0 \leq t \leq 1$, and that satisfies: $p(0) = p^{(0)}$ and $p(1) = p^{(1)}$. An example for three outcomes is shown in Fig. 2.2.

![Figure 2.2](image-url)
Chapter 2. Distinguishability of states

The concept of regions of uncertainty in the probability space is founded on statistical grounds. The introduced measure (2.12) is obtained by essentially counting the number of points, that can be distinguished in $m$ trials, between two distributions in probability space. Wootters goes on to define the statistical distance between two points $p^{(1)}$ and $p^{(2)}$ as the length of the shortest curve between the points. Of course the term "shortest" curve is understood in the sense of the statistical length defined above. Therefore we can write

$$d\left(p^{(1)}, p^{(2)}\right) = \min_{\ell} d\left(p^{(1)}, p^{(2)}|\ell\right). \quad (2.13)$$

Now, we calculate this minimum. We begin by noting that, if we set the curve $\ell$ in the probability space between two points, we can express the region of minimum uncertainty $\Delta t$ (which leads to maximum distinguishability) along the curve in $m$ trials as:

$$\sqrt{m} \left\{ \sum_i \frac{1}{p_i} \left( \frac{dp_i}{dt} \right)^2 \right\}^{1/2} \Delta t = 1. \quad (2.14)$$

Between $t$ and $t + dt$, there are $dt/\Delta t$ of distinguishable regions along the path. Therefore, the number for maximum number of distinguishable intermediate points along the curve $\ell$ is given by the integral:

$$d\left(p^{(1)}, p^{(2)}|\ell\right) = \lim_{m \to \infty} \frac{1}{\sqrt{m}} \int_0^1 dt \frac{d}{dt} \Delta t = \frac{1}{2} \int_0^1 dt \left\{ \sum_i \frac{1}{p_i} \left( \frac{dp_i}{dt} \right)^2 \right\}^{1/2}. \quad (2.15a)$$

One can use the methods of the variational calculus to find the shortest curve connecting two points and minimize the integral. Another approach is based on geometrical considerations [30], giving the outcome

$$d\left(p^{(1)}, p^{(2)}\right) = \arccos\left(f\left(p^{(1)}|p^{(2)}\right)\right), \quad (2.16)$$

which is exactly the Bhattacharyya-Wootters distance presented already in Eq. (2.9).

This reasoning reveals a deeper relation between statistical distinguishability of probabilities in $m$ trial, and the fidelity. Until now the fidelity was introduced as an arbitrary measure that somehow quantifies to what extent two probability distributions are different. An underlying idea, which results in a more profound understanding of this quantity, is related to the regions of uncertainty in large number of trials and statistical distinguishability of the frequencies of outcomes.

A relation to quantum mechanics can be derived as follows. Suppose that the system can be prepared in one of the two pure states $|\psi_1\rangle$ or $|\psi_2\rangle$. Any problem of distinguishability between the two states will also involve the problem of measurement and calculation of probabilities of outcomes. Therefore, the statistical distance $d$ also depends on the measurement. Wootters [30] showed that for the optimal measurement, for which the distinguishability is the highest,

$$\max_{\text{all measurements}} d\left(p^{(1)}, p^{(2)}\right) = \arccos \left( |\langle \psi_1 | \psi_2 \rangle| \right). \quad (2.17)$$
2.2. Fidelity as a measure of distinguishability

This formula states that the maximum statistical distinguishability of two pure states is given by their scalar product.

Properties of fidelity

We present here a few properties of the fidelity, which will be useful in the coming sections where we examine quantum systems and probabilities that describe measurements on the physical systems.

Multiplicativity

The first property of fidelity describes its behaviour in case when two probability distributions $P$ and $Q$ factorise, so that

\begin{align}
P_{n_1, n_2, \ldots, n_N} &= p_{n_1}^{(1)} p_{n_2}^{(2)} \cdots p_{n_N}^{(N)}, \quad (2.18) \\
Q_{n_1, n_2, \ldots, n_N} &= q_{n_1}^{(1)} q_{n_2}^{(2)} \cdots q_{n_N}^{(N)}. \quad (2.19)
\end{align}

In this notation index $n_i$ numbers the $i$th from $N$ outcomes. Omitting the indices that enumerate the outcomes, these formulas can symbolically be written as $P = p^{(1)} \otimes p^{(2)} \otimes \cdots \otimes p^{(N)},$ and analogically for $Q$. In this case, the fidelity between $P$ and $Q$, is the product of fidelities of constituting distributions $p^{(i)}$ and $q^{(i)}$ for each $i$. We say that the fidelity is multiplicative and write:

\begin{align}
f(P|Q) &= f(p^{(1)}|q^{(1)}) f(p^{(2)}|q^{(2)}) \cdots f(p^{(N)}|q^{(N)}). \quad (2.20)
\end{align}

Notice that, for independent variables, the logarithm of fidelity is additive in the following sense:

\begin{align}
\log f(P|Q) = \sum_i \log f(p^{(i)}|q^{(i)}). \quad (2.22)
\end{align}

Assume now that all $N$ probabilities are the same, which can be written as $P = p \otimes \cdots \otimes p = p^\otimes N$. In this case the fidelity takes a particularly simple form of the scaling law:

\begin{align}
f(p^\otimes N|q^\otimes N) = \left(f(p|q)\right)^N. \quad (2.23)
\end{align}

Upon taking logarithm of both sides we find that logarithm of fidelity scales linearly with $N$:

\begin{align}
\log f(p^\otimes N|q^\otimes N) = N \log f(p|q). \quad (2.24)
\end{align}

Joint concavity

The second property of the fidelity is the joint concavity. Imagine the following scenario, when the the two distributions $P$ and $Q$ come from the averaging of some
additional variables. This can be written as

\[ P_i = \sum_n p_n P_i^{(n)}, \quad (2.25a) \]
\[ Q_i = \sum_n p_n Q_i^{(n)}, \quad (2.25b) \]

where \( i \) indexes the outcomes. The \( P_i^{(n)} \) and \( Q_i^{(n)} \) are probabilities conditioned on the \( n \)th outcome, which comes with probability \( p_n \). Let us write symbolically that \( P = \sum_n p_n P^{(n)} \) and analogically for \( Q \). One can prove that the following inequality holds\(^1\):

\[ f\left( \sum_n p_n P^{(n)} | \sum_n p_n Q^{(n)} \right) \geq \sum_n p_n f\left( P^{(n)} | Q^{(n)} \right). \quad (2.26) \]

This result can be interpreted as follows: the loss of information about some variables (expressed in the form of averaging in (2.25)) leads to lower distinguishability of states. The equality in (2.26) is achieved if and only if all \( p_n \)'s except one vanish.

### 2.3 Classical Fisher information

In the previous section a measure of distinguishability of states was introduced, which is the fidelity. As we have seen, its definition is deeply rooted in statistical reasoning and the concept of maximum distinguishable points lying between the two states. In the present section, we investigate how the fidelity behaves when the two probability distributions, which we want to distinguish, are, in some sense, close to each other. It turns out that in the lowest order fidelity is governed by one parameter, which is called the Fisher information. Due to the fact that it is founded on the grounds of the theory of probability, it is also called the classical Fisher information. However, this object can also be regarded within quantum theory, where it acquires a special name — the quantum Fisher information \([29, 33, 34]\). The latter quantity is thoroughly analysed further in this chapter (see section 2.7), and is a central object for this thesis.

The Fisher information also appears in the context of metrology. It places the ultimate limit on convergence in maximum likelihood parameter estimation scheme \([35, 36]\). We shall return to this problem in the section 2.6.

To understand how the classical Fisher information is related to fidelity, consider a probability distribution \( p_i(\theta) \), which depends on the parameter \( \theta \). Now, suppose that one wants to distinguish between two probabilities calculated at \( \theta \) and \( \theta + \delta \theta \), i.e. between \( p_i^{(1)} = p_i(\theta) \) and \( p_i^{(2)} = p_i(\theta + \delta \theta) \). Calculation of the the fidelity to the second order in \( \delta \theta \) reveals the quadratic behaviour in \( \delta \theta \):

\[ f\left( p(\theta + \delta \theta) | p(\theta) \right) = 1 - \frac{1}{8} F_{cl} \delta \theta^2 + \ldots, \quad (2.27) \]

\(^1\) The inequality is based on a very general property satisfied by all concave (convex) functions: assume \( x, y \geq 0 \), if \( h(x) \) is concave (convex), then the function \( (x, y) \mapsto yf(x/y) \) is concave (convex). In the present situation \( h(x) = \sqrt{x} \).
2.4 Other measures of distinguishability.

2.4.1 Kolmogorov distance or trace distance

The Kolmogorov distance $D_K$ is widely used in quantum computation and quantum information theory [7]. In this context, it is known as the trace distance. This object is defined by the following equation:

$$D_K(p|q) = \frac{1}{2} \sum_i |p_i - q_i|.$$  \hfill (2.29)
It turns out that $D_K$ is a metric on probability distributions, which justifies the term “distance”. Moreover, it has a physically motivated meaning, which can be seen from an alternative formulation:

$$D_K(p|q) = \max_A |P(A) - Q(A)| = \max_A \left| \sum_{i \in A} p_i - \sum_{i \in A} q_i \right|. \quad (2.30)$$

In this notation, $P(A) = \sum_{i \in A} p_i$ and $Q(A) = \sum_{i \in A} q_i$ denote the probabilities of event $A$ under the condition that outcomes are distributed according to $p$ and $q$, respectively. Thus, the event $A$ for which the difference $|P(A) - Q(A)|$ is maximum is optimal to examine whether the distributions $p$ and $q$ are different.

The Kolmogorov distance $D_K$ satisfies Fuchs and van de Graaf inequality [37], which relates it to fidelity. The inequality takes the following form:

$$1 - f(p|q) \leq D_K(p|q) \leq \sqrt{1 - f^2(p|q)}. \quad (2.31)$$

We recall that to obtain the Fisher information we considered fidelity of two neighbouring distributions, i.e. $p_i = p_i(\theta + \delta \theta)$ and $q_i = p_i(\theta)$. The same reasoning can be applied to the Kolmogorov distance. Substituting these distributions into Eq. (2.29), we obtain to the lowest order in $\delta \theta$:

$$D_K(p|q) = \frac{1}{2} \sum_i |p_i(\theta + \delta \theta) - p_i(\theta)| = \frac{1}{2} T_K |\delta \theta| + \ldots, \quad (2.32)$$

where

$$T_K = \sum_i |\partial_\theta p_i(\theta)|. \quad (2.33)$$

The Fisher information does not emerge in this measure, contrary to the case of the fidelity. Instead, the object $T_K$ appears. Nevertheless, it is not difficult to prove, that it is related to a lower bound on the Fisher information. To see this, rewrite the squared $T_K$ in the following form:

$$T_K^2 = \left( \sum_i \frac{|\partial_\theta p_i(\theta)|}{\sqrt{p_i(\theta)}} \cdot \sqrt{p_i(\theta)} \right)^2 \leq \left( \sum_i \frac{(\partial_\theta p_i(\theta))^2}{p_i(\theta)} \right) \left( \sum_i p_i(\theta) \right) = F_{cl}. \quad (2.34)$$

We have used here the Cauchy-Schwarz inequality (2.8) and normalization condition for probabilities. Thus, we obtain that squared $T_K$ is a lower bound on the Fisher information; the same result could also have been achieved directly from inequality (2.31).

### 2.4.2 Hellinger distance

The Hellinger distance, denoted by $H$, was introduced in [38] and, for two distributions $p$ and $q$, is given by

$$H^2(p|q) = \frac{1}{2} \sum_i (\sqrt{p_i} - \sqrt{q_i})^2. \quad (2.35)$$
Apart from being a proper distance function, all the properties of $H$ can be reduced to that of the fidelity, since:

\[ H^2(p|q) = 1 - f(p|q). \]  

(2.36)

Therefore, for neighbouring distributions the Hellinger distance reduces to the Fisher information:

\[ H^2(p(\theta + \delta \theta)|q(\theta)) = \frac{1}{8} F_{cl} \delta^2 \theta + \ldots, \]  

(2.37)

where the dots represent higher-order terms in $\delta \theta$. Because of the fact that the Hellinger distance and the fidelity are so closely related, we shall not pause here, but rather move to a different distance, which importance will appear when investigating the metrological context (see section 2.6).

### 2.4.3 Kullback-Leibler divergence or relative entropy

The Kullback-Leibler divergence, known also as relative entropy, for two distributions $p$ and $q$, is given by [39]:

\[ D_{KL}(p|q) = - \sum_i q_i \log \left( \frac{p_i}{q_i} \right). \]  

(2.38)

This measure appears in a general method of inference [40]. To see this, suppose that a large number $m$ of measurements of some physical observable was performed on a system, and the outcomes are governed by a probability distribution $q_i$. Basing on the observed data, we can assign frequencies $p_i$ to these outcomes. However, assigning $p_i$ that are much different from the underlying distribution $q_i$ will become less plausible as $m$ grows. It can be proved that any set of frequencies $p_i$ which is different from $q_i$, is exponentially less likely to be assigned with increasing $m$. This effect is quantified by the formula [35, 41]:

\[ P(p|q) \approx e^{-mD_{KL}(p|q)}. \]  

(2.39)

The above equation should be read as follows: the probability of assigning frequencies $p_i$ in $m$ trials, provided that the outcomes are governed by the distribution $q_i$, is decaying exponentially with $m$. The quantity $D_{KL}(p|q)$ which controls the decay is the Kullback-Leibler divergence between the distributions $p_i$ and $q_i$. Therefore, relative entropy quantifies how dissimilar are the two distributions, and can be used as a measure of distinguishability between the physical states.

It is not immediately clear that the $D_{KL}$ is non-negative, because the sum contains terms of both signs due to the presence of logarithm. Nevertheless, using Jensen inequality for convex function $\varphi$ of random variable $X$:

\[ \varphi(\langle X \rangle) \leq \langle \varphi(X) \rangle, \]  

(2.40)

with the discrete variable $x_i = p_i/q_i$, and the mean to be taken over $q$ distribution, we obtain:

\[ \varphi \left( \sum_i q_i \frac{p_i}{q_i} \right) = \varphi(1) \leq \sum_i q_i \varphi \left( \frac{p_i}{q_i} \right). \]

(2.41)
Inserting a specific function $\varphi(x) = -\log x$, we arrive at

$$D_{KL}(p|q) \geq 0. \quad (2.42)$$

The equality in (2.40) is reached if and only if $X$ is constant (degenerate random variable) or $\varphi$ is linear. Therefore, $D_{KL}(p|q)$ is zero if and only if $p_i/q_i$ is a constant independent of index $i$, which results in $p = q$.

From the definition (2.38), it follows that the relative entropy for probability distributions describing $N$ independent events breaks up into a sum of relative entropies of individual distributions. We say that the Kullback-Leibler divergence is additive. For the distributions given by (2.18), we have

$$D_{KL}(P|Q) = \sum_{i=1}^{N} D_{KL}(p^{(i)}|q^{(i)}). \quad (2.43)$$

If all the distributions are the same, $p^{(i)} = p$ and $q^{(i)} = q$, the following linear scaling can be formulated:

$$D_{KL}(P|Q) = D_{KL}(p^{\otimes N}|q^{\otimes N}) = ND_{KL}(p|q). \quad (2.44)$$

The Kullback-Leibler divergence satisfies the “loss-of-information” inequality, called the joined-convexity, similar to Eq. (2.26). Assuming that the two distributions, between which we want to calculate the divergence, are given by Eq. (2.25), this inequality reads$^2$:

$$D_{KL}\left(\sum_n p_n P^{(n)} \mid \sum_n p_n Q^{(n)}\right) \leq \sum_n p_n D_{KL}\left( P^{(n)} \mid Q^{(n)} \right). \quad (2.45)$$

We saw that the Kullback-Leibler divergence has properties that are analogous to the properties of fidelity and other measures introduced to distinguish two probability distributions. It is not surprising that for two neighbouring probability distributions relative entropy, similarly to fidelity, is proportional to the Fisher information $[30, 42, 43]$:

$$D_{KL}(p(\theta + \delta \theta) | p(\theta)) = \frac{1}{2} F_{\text{cl}} \delta \theta^2 + \ldots, \quad (2.46)$$

where the dots indicate terms that are of higher order than $\delta \theta^2$.

### 2.4.4 Generalization to the $f$-divergence

In the sections 2.4.1 (Kolmogorov distance), 2.4.2 (Hellinger distance) and 2.4.3 (Kullback-Leibler divergence) different quantities used for states differentiation were presented.

In this section, we present a family of quantities that can serve as a mean for distinguishing the states. They are called the $f$-divergences, and were first introduced by Csiszár [44]. In information theory, the $f$-divergences are described as

$^2$ See footnote on page 28. In the present case the convex function $h(x) = -\log x$. 
In statistical terms, the equality is achieved if the transition probability that satisfies the property of information

\[ D_f(p|q) = \sum_i q_i f \left( \frac{p_i}{q_i} \right). \]  

(2.47)

In analogy with the Kullback divergence (see Eq. (2.40) and the subsequent paragraphs), with help of Jensen’s inequality, it can be proved that \( D_f \) is non-negative:

\[ D_f(p|q) \geq f(1) = 0. \]  

(2.48)

The equality sign is achieved if and only if \( p = q \). Therefore, each convex function \( f \), that satisfies additional restriction \( f(1) = 0 \) gives some particular \( f \)-divergence, which in turn quantifies how distinguishable the two distributions are. Note, that there is no symmetry with respect to interchange of the distributions. Nevertheless, such interchange can be included in the transformation of the function itself, replacing \( f(t) \) with \( f^*(t) = tf(1/t) \); then \( D_f(p|q) = D_f(q|p) \).

At this point, we present a very important property of the \( f \)-divergences called the monotonicity [46]. This property is particularly easy derive using continuous variables. Suppose that \( p(x) \) and \( q(x) \) are the two probability densities of the variable \( X \). Assume that \( \kappa(y|x) \) is the transition probability, which describes the probability of random transformation of the value \( x \) to \( y \). Denote the probability densities of variable \( Y \) derived from \( p(x) \) and \( q(x) \) by \( p_\kappa(y) \) and \( q_\kappa(y) \). Then the following inequality is satisfied:

\[ D_f(p_\kappa|q_\kappa) \leq D_f(p|q). \]  

(2.49)

Note that the deterministic transformation described by the function \( y = F(x) \), for example the change of variables, corresponds to \( \kappa(y|x) = \delta(y - F(x)) \), where \( \delta(x) \) is the Dirac delta. The new probability distribution is \( p(y) = \langle \delta(y - F(x)) \rangle \), where the mean is taken over \( p(x) \) distribution. The inequality (2.49) states that in general the transformation of the considered initial variables is accompanied by the loss of information how different the two distributions are. The equality is achieved if \( p_\kappa(x|y) = q_\kappa(x|y) \) for all \( x \) and \( y \).\(^4\) This happens for example when \( F \) is one-to-one mapping, where no information is lost about the distinguishability of distributions.

The \( f \)-divergence also satisfies an inequality which we have already encountered when inspecting the Kullback-Leibler divergence (2.45) and the fidelity (2.26). The joint convexity property states that for two distributions given by Eq. (2.25), the divergence \( D_f \) satisfies an inequality:\(^5\)

\[ D_f \left( \sum_n p_n P^{(n)} \right) \sum_n p_n Q^{(n)} \right) \leq \sum_n p_n D_f \left( P^{(n)} \middle| Q^{(n)} \right). \]  

(2.50)

\(^3\) A divergence \( D(p|q) \) is called separable (or sometimes decomposable) if it is written as a sum of functions of two variables \( p_i \) and \( q_i \), \( D(p|q) = \sum_i D(p_i|q_i) \).

\(^4\) In statistical terms, the equality is achieved if the transition probability \( \kappa \) is induced from a function which is a sufficient statistic with respect to a dichotomous variable \( \xi = \{ p, q \} \).

\(^5\) See footnote on page 28. In the present case the convex function \( h(x) = f(x) \).
The above property reflects “loss-of-information”, because it is a special case of the monotonicity (2.49).

<table>
<thead>
<tr>
<th>$f$-divergence $D_f$</th>
<th>Function $f(t)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kolmogorov distance $D_K$</td>
<td>$</td>
</tr>
<tr>
<td>Hellinger distance $H^2$</td>
<td>$1 - \sqrt{t}$</td>
</tr>
<tr>
<td>Kullback-Leibler divergence $D_{KL}$</td>
<td>$-\log t$</td>
</tr>
<tr>
<td>$\chi^2$-divergence</td>
<td>$t^2 - 1$</td>
</tr>
<tr>
<td>$\alpha$-divergence</td>
<td>$\frac{4}{1-\alpha^2} \left(1 - t^{(1+\alpha)/2}\right)$, for $</td>
</tr>
</tbody>
</table>

Table 2.1: List of the $f$-divergences with the corresponding function $f$ and its name.

The $f$-divergence unifies all the measures that were considered so far. For instance, the fidelity $f(p|q)$ is related to the squared Hellinger distance $H^2(p|q) = 1 - f(p|q)$, which can be written as $H^2 = D_f(p|q)$, with the function $f(t) = 1 - \sqrt{t}$. The common list of the divergences and its names are presented in the table 2.1. Some of them were not discussed here, but are encountered in many problems. For instance, the $\chi^2$-divergence appears in the Chapman-Robinson bound [49], which is related to Cramer-Rao bound that we discuss in details below in the section 2.6.

The $f$-divergence can serve as a convenient yet powerful measure of how two states (described by probability distributions) are different. A particularly interesting is the property (2.49), which states that any manipulation of the variables will in general lead to lower distinguishability. Only if the transformation (data reduction) is properly chosen, the information is not lost and the distinguishability remains the same.

Now, suppose that we wish to distinguish two neighbouring distributions $p(\theta)$ and $p(\theta + \delta \theta)$. The expansion of the $f$-divergence in $\delta \theta$ leads to:

$$D_f(p(\theta + \delta \theta)|p(\theta)) = \sum_i p_i(\theta) f \left( \frac{p_i(\theta + \delta \theta)}{p_i(\theta)} \right) = \frac{f''(1)}{2} F_{cl} \delta \theta^2 + \ldots, \quad (2.51)$$

where dots indicate higher-order terms than $\delta \theta^2$; $f''(1)$ is the second derivative of $f(t)$ calculated at $t = 1$. We used here that $f(1) = 0$, which is the same condition needed also in Eq. (2.48). Note, that the order of arguments in $D_f$ does not matter, because $D_f(p(\theta)|p(\theta + \delta \theta)) = D_{f^*}(p(\theta + \delta \theta)|p(\theta))$ which after expansion, and using $f^{*''}(1) = f''(1)$, is exactly the same as (2.51).

The formula (2.51) is important, because it reveals that the underlying object, common to all measures $D_f$, is the classical Fisher information. The quantity $F_{cl}$ that emerged for the fidelity, the Hellinger distance, and the Kullback-Leibler entropy, turns out to be deeply related to a broad class of measures. These expansions also confirm that from the perspective of a broad class of measures, the Fisher information quantifies the susceptibility of the state. It must be borne in mind that
2.5 Properties of the Fisher information

When considering the susceptibility of the state a context must be presented – to which physical quantity or to which measurement the susceptibility is related.

2.5 Properties of the Fisher information

Until now, the Fisher information appeared as a coefficient in the expansion of a measure that quantifies how dissimilar two probability distributions are. In this section, we briefly describe its main properties. These are: additivity, monotonicity and convexity. We will see that all of them can be derived from the corresponding properties of some specific $f$-divergences.

Additivity

The Fisher information is additive with respect to the joint probability distribution describing independent events. Suppose that the distribution $P(\theta)$ is of the form (2.18). Then we have:

$$F_{cl}[p^{(1)} \otimes \cdots \otimes p^{(N)}] = \sum_{n_1,\ldots,n_N} \frac{1}{P_{n_1,\ldots,n_N}(\theta)} \left( \frac{\partial P_{n_1,\ldots,n_N}(\theta)}{\partial \theta} \right)^2$$

$$= \sum_{i=1}^N \sum_{n_i} \frac{1}{p^{(i)}_{n_i}(\theta)} \left( \frac{\partial p^{(i)}_{n_i}(\theta)}{\partial \theta} \right)^2$$

$$= \sum_{i=1}^N F_{cl}[p^{(i)}]. \quad (2.52)$$

The $F_{cl}[p^{(i)}]$ is the FI of the $i$th single probability distribution $p^{(i)}$. If all the distributions are the same $p^{(i)} = p$, a simple scaling law is obtained,

$$F_{cl}[p^{\otimes N}] = NF_{cl}[p]. \quad (2.53)$$

As the Fisher information is encountered in all the $f$-divergences, we can take a specific one, which has very particular properties. The above properties can be simplify proved with the use of the Kullback-Leibler divergence. Upon expansions of Eq. (2.43) and Eq. (2.44) we immediately arrive at Eq. (2.52) and Eq. (2.53), respectively.

Monotonicity

Suppose that the transformation of the random variable $X$ is performed. The realization $x$ of $X$ is mapped onto $y$ with the conditional probability $\kappa(y|x)$. In the same manner as is described in the section above (see Eq. (2.49)), we denote the probability distribution of obtaining value $y$ as $p_\kappa(y)$. Inserting the distributions $p = p(\theta + \delta \theta)$ and $q = p(\theta)$ into Eq. (2.49), and expanding in $\delta \theta$, the following inequality is obtained:

$$F_{cl}[p_\kappa] \leq F_{cl}[p]. \quad (2.54)$$
This inequality justifies the name for the Fisher “information”, because it quantifies the information about the distinguishability of two neighbouring distributions. Any transformation performed on the given initial distributions may be accompanied with the loss of information.

**Convexity**

From the joined convexity that is satisfied by all the $f$-divergences, a convexity of the Fisher information can be derived. Expanding Eq. (2.50) for two neighbouring distributions that differ in the parameter $\theta$ we obtain the following inequality:

$$F_{cl} \left[ \sum_n p_n P^{(n)}(\theta) \right] \leq \sum_n p_n F_{cl} \left[ P^{(n)}(\theta) \right].$$  \hfill (2.55)

With the presentation of the three properties of the classical Fisher information: additivity (2.52), monotonicity (2.54), and convexity (2.55), we conclude the general discussion on the measures of distinguishability. In the following sections we consider a problem that is frequently found in physics, which is an estimation of an unknown parameter from the experimental data. It turns out that the classical Fisher information is the quantity of the most importance in that field.

### 2.6 Metrology

In this section we consider a general problem of inferring the value of a parameter, on which the state of a system depends, and demonstrate that this task is closely related to the distinguishability of two neighbouring probability distributions.

![Diagram of estimation scheme](image)

**Figure 2.3:** The general estimation scheme. In the first stage, the system is prepared. In the second, an action, depending on an unknown parameter $\theta$, is performed on the system. As a result, the state of the system starts to depend on the parameter. In the third stage, a set of observables is measured. The value of the parameter $\theta$ is estimated with the precision $\Delta \theta$. 
A general estimation scheme consists of four stages. In the first stage (denoted by “state preparation” in Fig. 2.3), the state of the system, given by density matrix \( \hat{\rho} \), is prepared in a dedicated procedure.

In the second stage (“phase imprinting” in Fig. 2.3), the system is subject to an interferometric transformation. This action distorts the state of the system, which acquires a dependence on a parameter \( \theta \), called a phase. As a consequence, the state is given by \( \hat{\rho}(\theta) \).

In the third stage (“measurement” in Fig. 2.3), a set of observables is measured. The measurement process is described by a set of POVM operators, \( \{ \hat{E}_i \} \). The statistics of the outcomes, indexed with \( i \), are governed by the probability distribution \( p_i(\theta) = \text{Tr} \left[ \hat{E}_i \hat{\rho}(\theta) \right] \), which, in general, also depends on the value of the unknown parameter. Alternatively, the probability \( p_i(\theta) \) can be written as \( p(i|\theta) \), to emphasize that the probability of observing \( i \)th outcome is conditioned on the value of \( \theta \). Also, it is important to stress that the measurement is pre-chosen, and it does not depend on the phase \( \theta \). It may happen, though, that the measurement is chosen inadequately, so the distribution \( p_i(\theta) \) is independent of \( \theta \), preventing any inference about its value.

In the final stage (“estimation scheme” in Fig. 2.3), from the measured quantities the value of \( \theta \) is estimated with an accuracy \( \Delta \theta \). This step is called an “estimation scheme”, as there are many options to handle the data before assessing the value of \( \theta \). In practice, this stage is undertaken after completing a large number of repetitions of steps 1–3 from Fig. 2.3.

We have defined the problem encountered in metrology as the estimation of some unknown parameter \( \theta \) from a physical system. Now we will focus only on the fourth stage in Fig. 2.3, i.e. how to extract information about \( \theta \) knowing the outcomes of the measurement.

Consider now a specified measurement repeated \( m \) times, with outcomes which we shall denote with \( x_1, x_2, \ldots, x_m \). This is the sample of the data of length \( m \). Each \( x_i \) is a realisation of a random variable \( X \), characterised by probability distribution \( p(x|\theta) \). The variable \( X \) can be continuous, discrete or both; it can be single or multivariate. Here for simplicity we use the notation as if \( X \) was continuous distribution, but the generalization to other cases is straightforward.

The information about \( \theta \) is inferred with help of a function, called an estimator,

\[
\theta_{\text{est}} = \theta_{\text{est}}(x_1, \ldots, x_m).
\] (2.56)

The estimator is a rule which applied to a given sample yields an estimate of the parameter \( \theta \). Of course, \( \theta_{\text{est}} \) is also a random variable, and its distribution depends on the parameter \( \theta \). For each sample the estimate is different. Now, we have to introduce a condition which ensures that \( \theta_{\text{est}} \) leads to proper value of \( \theta \). An estimator is called unbiased if the following condition is satisfied for all values of \( \theta \):

\[
\langle \theta_{\text{est}} \rangle = \theta.
\] (2.57)

Here the mean is taken over distribution \( p(x_1|\theta) \cdots p(x_m|\theta) \) of variable \( (x_1, \ldots, x_m) \).
It means that the average is taken over all samples. Alternatively, the condition (2.57), can be written as

\[ \int dx_1 dx_2 \cdots dx_m p(x_1|\theta) \cdots p(x_m|\theta) \left( \theta_{\text{est}}(x_1, \ldots, x_m) - \theta \right) = 0. \]  

(2.58)

Differentiating with respect to \( \theta \) on both sides, yields

\[ \int dx_1 dx_2 \cdots dx_m p(x_1|\theta) \cdots p(x_m|\theta) \left( \sum_{i=1}^{m} \frac{\partial \log p(x_i|\theta)}{\partial \theta} \right) \left( \theta_{\text{est}}(x_1, \ldots, x_m) - \theta \right) = 1. \]  

(2.59)

Now, to the left hand side we apply the Cauchy-Schwarz inequality, which leads to

\[ \left\langle \left( \theta_{\text{est}}(x_1, \ldots, x_m) - \theta \right)^2 \right\rangle \geq \frac{1}{m F_{\text{cl}}}, \]  

(2.60)

where \( F_{\text{cl}} \) is the the Fisher information calculated for the distribution \( p(x|\theta) \). The left hand side is the mean squared error of the estimation, \( (\Delta \theta)^2 \), therefore this inequality, known as Cramér-Rao lower bound [36], takes a particularly simple form:

\[ \Delta \theta \geq \frac{1}{\sqrt{m F_{\text{cl}}}}. \]  

(2.61)

This inequality places a limit on variance of any unbiased estimator. An estimator which achieves the bound in (2.61) is called efficient, and thus is the best unbiased estimator in the sense of minimum mean squared error.

### 2.6.1 Maximum-likelihood estimation

Does an efficient estimator always exists? It turns out that in general it is not the case. Nevertheless, Fisher constructed [50] an estimator that is asymptotically, for large \( m \), unbiased and efficient. This estimator is called the maximum-likelihood estimation (MLE). Fisher’s theorem states that for large \( m \) the estimated parameter tends to the true value\(^6\) and achieves Cramér-Rao lower bound.

Some rudimentary forms of the maximum-likelihood estimation appeared before Fisher’s works [51]. However, it was Fisher who created the modern version of the maximum likelihood, developing the scheme between 1912 and 1922 [52–56]. In 1921, the term “likelihood” was coined, and in 1922 the “maximum likelihood estimate.”

To understand the method of the maximum-likelihood estimation, we first recall the Kullback-Leibler divergence given by Eq. (2.38). Suppose that the outcomes \( x_1, \ldots, x_m \) of the measurement are distributed according to the probability distribution \( p(x|\theta) \) with parameter \( \theta \). To proceed with any estimation, a model is necessary, which requires the knowledge of the function \( p(x|\phi) \) for all \( \phi \) and \( x \). We inserted here \( \phi \) instead of \( \theta \), because the true value \( \theta \) is unknown; only the functional form

\(^6\) In statistics, this property of an estimator, that for large samples the estimator tends to the true value of an estimated parameter, is called consistency.
$p(x|\phi)$ is given. We may calculate the Kullback divergence and use inequality (2.42), which reads:

$$D_{\text{KL}}\left(p(x|\phi)p(x|\theta)\right) \geq 0. \tag{2.62}$$

Using the explicit formula given by Eq. (2.38), the following result is obtained:

$$-\int dx \, p(x|\theta) \log \left( \frac{p(x|\phi)}{p(x|\theta)} \right) \geq 0. \tag{2.63}$$

Rearranging the terms yields

$$\int dx \, p(x|\theta) \log p(x|\phi) \leq \int dx \, p(x|\theta) \log p(x|\theta) = \beta. \tag{2.64}$$

As the true value of $\theta$ is unknown, right hand side of the above inequality is also unknown; call it $\beta$. Now, we can rewrite Eq. (2.64) to obtain:

$$\langle \log p(x|\phi) \rangle_\theta \leq \beta. \tag{2.65}$$

The subscript $\theta$ indicates that the mean is taken over the distribution with the true value $\theta$. From the properties of the relative entropy, we know that the maximum value of the left hand side, equal to $\beta$, is reached only for $\phi = \theta$. Therefore, taking the maximum of the left hand side with respect to the variable $\phi$ will lead to $\phi = \theta$, and thus to the true value of the parameter. Although, the mean on the left hand side cannot be evaluated as the true distribution is unknown, it can be estimated with the data sample of length $m$. If the sample is large enough, we approximately have:

$$\langle \log p(x|\phi) \rangle_\theta \approx \frac{1}{m} \sum_{i=1}^{m} \log p(x_i|\phi) = \frac{1}{m} \log \left( \mathcal{L}(\phi) \right). \tag{2.66}$$

The term in brackets on the far right hand side is called the likelihood $\mathcal{L}$:

$$\mathcal{L}(\phi) = \prod_{i=1}^{m} p(x_i|\phi). \tag{2.67}$$

Basing on the equations (2.65) and (2.66) we obtain that, asymptotically for large $m$, the estimate $\theta_{\text{est}}$ of $\theta$ in the maximum-likelihood estimation scheme (denoted by $\theta_{\text{MLE}}$) is the maximum of the function $\mathcal{L}(\phi)$:

$$\theta_{\text{est}} \equiv \theta_{\text{MLE}} = \phi \text{ for which } \mathcal{L}(\phi) \text{ is maximum.} \tag{2.68}$$

Alternatively, one can look for the maximum of $\frac{1}{m} \log \mathcal{L}$. The likelihood $\mathcal{L}(\phi)$ is also a function of the sample and thus can be regarded as a random variable. Fisher’s theorem states that for large $m$, the mean value of $\theta_{\text{MLE}}$ is equal to the true value $\theta$, and its variance is equal to the lower bound of Cramér-Rao inequality:

$$\langle (\theta_{\text{MLE}} - \theta)^2 \rangle \equiv \Delta \theta_{\text{MLE}}^2 = \frac{1}{m F_{\text{cl}}}. \tag{2.69}$$
Therefore, the maximum-likelihood estimation asymptotically saturates the Cramér-Rao lower bound. It is thus a powerful method of estimation that leads to the best precision possible. However, the disadvantage of this method is the necessity to use large data samples.

In the next section, we present one more estimation scheme, which is frequently used in physics, called an estimation from the mean. It will lead us to an uncertainty in $\theta$ which is given by the error-propagation formula.

### 2.6.2 Estimation from the mean

Suppose that a full probability distribution $p(x|\theta)$ of a random variable $X$ is not known, but something more simple is available. For the sake of simplicity, let it be the first moment of that distribution, i.e. a functional form of the mean value of $X$ given by:

$$ \int x p(x|\theta) dx \equiv \langle X \rangle_\theta = f(\theta). $$

(2.70)

The subscript reminds that the mean value also depends on $\theta$, in general. If the values $x_1, x_2, \ldots, x_m$ are measured in $m$ trial, the mean value can be approximated by:

$$ t = \frac{1}{m} \sum_{i=1}^{m} x_i. $$

(2.71)

Its statistics is governed by the probability distribution $q(t|\theta)$, which in the limit of large $m$, by the central limit theorem, can be approximated by a Gaussian distribution

$$ q(t|\theta) = \frac{1}{\sqrt{2\pi\sigma^2(\theta)}} e^{-\frac{(t-f(\theta))^2}{2\sigma^2(\theta)}}, $$

(2.72)

where we have used Eq. (2.70). The standard deviation $\sigma_m$ is given by the variance of $t$:

$$ \sigma^2_m(\theta) = \langle (t - \langle t \rangle)^2 \rangle = \frac{\sigma^2(\theta)}{m}, $$

(2.73)

where $\sigma^2(\theta)$ is the variance of $X$.

If we used all the information provided by this distribution, we would obtain, by the Cramér-Rao lower bound for a single outcome (in Eq. (2.61) set $m = 1$), the uncertainty bounded by

$$ (\Delta \theta)^2 \geq \frac{1}{F_{\text{cl}}[q(t|\theta)]}. $$

(2.74)

The Fisher information on the right hand side can be calculated to give

$$ F_{\text{cl}}[q(t|\theta)] = m \frac{f'(\theta)^2 + \frac{2}{m} (\sigma'(\theta))^2}{(\sigma(\theta))^2}, $$

(2.75)

where the prime denotes differentiation with respect to $\theta$. It can be shown that asymptotically for large $m$, the second term in the numerator can be dropped yielding:

$$ F_{\text{cl}}[q(t|\theta)] \approx m \frac{f'(\theta)^2}{\sigma^2(\theta)}. $$

(2.76)
The Cramér-Rao bound can thus be written as
\[ \Delta \theta \gtrsim \frac{1}{\sqrt{m}} \frac{\Delta X}{\partial \langle X \rangle/\partial \theta}, \]  
(2.77)
where in this notation \( \Delta X \) is the standard deviation of variable \( X \). Care must be taken, since in this formula \( \Delta X \), along with \( \langle X \rangle \), depends on the value of \( \theta \). In this way we reduced the problem of estimating \( \theta \) from sample \( x_1, \ldots, x_m \), each coming from distribution \( p(x|\theta) \), to estimating the unknown parameter \( \theta \) from new distribution for \( t \), which is less fluctuating (cf. Eq. (2.73)).

Note that in the formula (2.77), the statistical factor \( 1/\sqrt{m} \) is the same that appeared in the Cramér-Rao lower bound (2.61).

An inequality can now be formulated by noting that the transformation of variables \( (x_1, \ldots, x_m) \mapsto t \) cannot increase the Fisher information due to the monotonicity expressed by Eq. (2.54). Therefore, the following chain of inequalities can be written:
\[ F_{cl}[p(x_1|\theta) \cdots p(x_m|\theta)] \geq F_{cl}[q(t|\theta)] \geq m \frac{1}{(\Delta X)^2} \left( \frac{\partial \langle X \rangle}{\partial \theta} \right)^2. \]  
(2.78)
Consequently, the information obtained from the estimation from the mean is lower than the Fisher information:
\[ \frac{1}{(\Delta X)^2} \left( \frac{\partial \langle X \rangle}{\partial \theta} \right)^2 \leq F_{cl}[p(x|\theta)]. \]  
(2.79)
This is an important inequality. Namely, it often happens that the evaluation of the Fisher information is intractable, especially experimentally. However, it might be possible to independently evaluate the first two moments: the mean \( \langle X \rangle \) and the variance \( (\Delta X)^2 \). With the help of these two objects, one can formulate a simple bound that is given by Eq. (2.79).

A remark is in order: one can consider a new variable \( Y = X^2 \), which is obtained from the initial random variable \( X \). The estimation from mean value of \( Y \) is called estimation from the second moment of \( X \). All the above reasoning can be reformulated for this new variable, which leads to the conclusion that the estimation from mean value of \( Y \) gives
\[ \Delta \theta \gtrsim \frac{1}{\sqrt{m}} \frac{\Delta Y}{\partial \langle Y \rangle/\partial \theta}, \]  
(2.80)
provided that \( m \) is large. Here, \( \langle Y \rangle \) is the mean value of \( Y \) and \( \Delta Y \) is its standard deviation. Now, a similar formula to Eq. (2.79) can be stated:
\[ \frac{1}{(\Delta Y)^2} \left( \frac{\partial \langle Y \rangle}{\partial \theta} \right)^2 \leq F_{cl}[p(x|\theta)]. \]  
(2.81)
Unfortunately, there is no relation between left hand side of (2.79) and (2.81), because there is no mapping between variables \( t = \frac{1}{m} \sum x_i \), which estimates the mean
of $\langle X \rangle$ in $m$ trials, and $t_2 = \frac{1}{m} \sum_i x_i^2$, which estimates the mean of $\langle Y \rangle$ in $m$ trials. The monotonicity of the Fisher information cannot be invoked in this situation.

Now, we present a method of estimation which saturates the inequality (2.77). We assume that the relation (2.70) for the mean is known, and can be inverted, i.e. $g(t) = f^{-1}(t)$. A natural proposal for the estimator is given by

$$\theta_{\text{est}} = g(t), \quad (2.82)$$

where $t$ is given by Eq. (2.71), which, for large $m$, is distributed according to Gaussian $q(t|\theta)$ given by Eq. (2.72). From the construction, $g$ has the property that $g(\langle t \rangle) = \theta$. For large data samples, $t$ should be close to the mean value $\langle t \rangle$. Now we show, that $\theta_{\text{est}}$ is asymptotically unbiased. To see this, we expand $g(t)$ around the mean value up to the second order in $(t - \langle t \rangle)$:

$$\langle \theta_{\text{est}} \rangle = \langle g(t) \rangle = \left\langle g(\langle t \rangle) + g'(\langle t \rangle)(t - \langle t \rangle) + \frac{1}{2} g''(\langle t \rangle)(t - \langle t \rangle)^2 \right\rangle. \quad (2.83)$$

Using the fact that $t$ is normally distributed, we obtain

$$\langle \theta_{\text{est}} \rangle = \theta + \frac{1}{2} g''(\langle t \rangle) \frac{\sigma^2}{m} \xrightarrow{m\to\infty} \theta. \quad (2.84)$$

What remains to be shown is the mean squared error of $\theta_{\text{est}}$. We proceed in the similar manner, and expand up to the second order in $(t - \langle t \rangle)$:

$$\theta_{\text{est}} - \theta = g(t) - t = g'(\langle t \rangle)(t - \langle t \rangle) + \frac{1}{2} g''(\langle t \rangle)(t - \langle t \rangle)^2. \quad (2.85)$$

Now, squaring both sides and taking the average, we obtain

$$(\Delta \theta)^2 \equiv \langle (\theta_{\text{est}} - \theta)^2 \rangle = \langle g'(\langle t \rangle) \rangle^2 \langle (t - \langle t \rangle)^2 \rangle + \frac{1}{4} \langle g''(\langle t \rangle) \rangle^2 \langle (t - \langle t \rangle)^4 \rangle. \quad (2.86)$$

As the distribution of $t$ is Gaussian, the mean on the right hand side can be evaluated, which leads to:

$$(\Delta \theta)^2 = (g'(\langle t \rangle))^2 \frac{\sigma^2}{m} + \frac{3}{4} (g''(\langle t \rangle))^2 \frac{\sigma^4}{m^2}. \quad (2.87)$$

For large values of $m$, the uncertainty in $\theta_{\text{est}}$, up to the order of $1/m$ is given by:

$$(\Delta \theta)^2 = \left( \frac{1}{f'(\theta)} \right)^2 \frac{\sigma^2}{m}. \quad (2.88)$$

where we used the fact that $g$ is the inverse function of $f$; we know that $g(\langle t \rangle) = \theta$ and $(t) = f(\theta)$, and hence $g'(\langle t \rangle) = 1/f'(\theta)$. The precision given by Eq. (2.88) can be rewritten as:

$$\Delta \theta = \frac{1}{m} \frac{\Delta X}{\partial \theta}, \quad (2.89)$$
which is exactly the right hand side of the inequality (2.77). We thus see, that the estimator defined in Eq. (2.82) is asymptotically unbiased and efficient with respect to the distribution \( q(t|\theta) \). However, the estimated precision does not exceed the one from the initial distribution \( p(x|\theta) \), from which \( q(t|\theta) \) was derived. It often happens that it is lower, and this loss of information is manifested by the inequality (2.79).

In this section, we introduced a metrological problem of estimating an unknown parameter \( \theta \). We inferred its value from a data sample basing on the function called the estimator. The precision of the estimation, or in other words the uncertainty of the estimated parameter, measured with the mean squared error, is ultimately given by the Cramér-Rao lower bound. In this inequality, the Fisher information appears, which is related to a limit on the best attainable precision. The bound is achievable in the procedure called the maximum-likelihood estimation scheme, as long as the data is large enough. We also investigated simple scenarios of estimation basing on imperfect knowledge of the underlying probability distribution. An example of such an attempt is an estimation from the mean, where the precision is given by the error-propagation formula.

In the previous sections, the Fisher information was introduced as a measure of how different two neighbouring distributions are. Now, we have seen that the distinguishability of the two distributions can be mapped on the estimation of an unknown parameter. The more the distributions differ in the vicinity of the true value of the parameter, with better precision that value can be estimated.

At this point we close the discussion about the properties of the Fisher information and situations where it can be found. In the following section, we move to the grounds of quantum mechanics, and use the knowledge, that a probability distribution that describes the statistics of outcomes, is a result of a specified measurement.

2.7 Quantum Fisher information and quantum correlations

In this section we introduce an important concept of the optimal measurement. In the previous section we discussed several different measures of distinguishability of probability distributions. Until now, we assumed that these distributions describe the knowledge about outcomes of a specified measurement. Now, we explicitly include the measurement in the mathematical formalism, which entails the dependence of the distinguishability measure on the physical measurement. The optimal measurement is the one which leads to the highest (for instance, for the Kullback-Leibler divergence) or smallest (for example for fidelity) possible value of the measure of distinguishability. In other words, the optimal measurement is the one which implies the best discrimination between the two states in quantum mechanics. Clearly, the choice of the optimal measurement depends on the specific measure.

Due to the fact that a broad class of distinguishability measures applied to neighbouring distributions leads to the classical Fisher information, the respective
optimal measurement will lead to the highest possible value of the classical Fisher information. This value, which is attained for the optimal measurement is called the *quantum Fisher information* and denoted usually by $F_Q$.

### 2.7.1 Optimal measurement. Quantum Fisher information.

**Optimal measurement**

Suppose that we have a physical system that can be in two states, described with two density matrices $\rho_1$ and $\rho_2$. Any statement about the distinguishability between these two states must involve a measurement, which is described with measurement operators $\hat{E}_n$. Here, $n$ indexes the outcomes, distributed according to

$$p^{(1)}_n = \text{Tr}[\hat{\rho}_1 \hat{E}_n] \quad \text{and} \quad p^{(2)}_n = \text{Tr}[\hat{\rho}_2 \hat{E}_n]. \quad (2.90)$$

The *optimal measurement* (always in relation to a chosen measure) is the one which leads to the highest possible distinguishability between the two states $\hat{\rho}_1$ and $\hat{\rho}_2$. The degree to which two states are different is then quantified by the resulting quantum measure. Below, we present this idea in more details taking $f$-divergence as an example.

Let us assume that the distinguishability can be quantified with the $f$-divergence:

$$D_f(p^{(1)}|p^{(2)}) = \sum_n p^{(2)}_n f\left(\frac{p^{(1)}_n}{p^{(2)}_n}\right) = \sum_n \text{Tr}[\hat{\rho}_2 \hat{E}_n] f\left(\frac{\text{Tr}[\hat{\rho}_1 \hat{E}_n]}{\text{Tr}[\hat{\rho}_2 \hat{E}_n]}\right). \quad (2.91)$$

Here, the right hand side shows explicit dependence on the measurement $\hat{E}_n$. The optimal measurement, denoted by $\hat{E}^{\text{opt}}_n$, is the one for which the following inequality is satisfied:

$$D_f\left(\text{Tr}[\hat{\rho}_1 \hat{E}_n] \bigg{|} \text{Tr}[\hat{\rho}_2 \hat{E}_n]\right) \leq D_f\left(\text{Tr}[\hat{\rho}_1 \hat{E}^{\text{opt}}_n] \bigg{|} \text{Tr}[\hat{\rho}_2 \hat{E}^{\text{opt}}_n]\right) \equiv D_{f,Q}(\hat{\rho}_1|\hat{\rho}_2), \quad (2.92)$$

for all measurements $\hat{E}_n$. In other words, the maximum value of the $f$-divergence, maximised over all possible measurements, and called the *quantum $f$-divergence* is given by

$$D_{f,Q}(\hat{\rho}_1|\hat{\rho}_2) \equiv \max_{\{\hat{E}_n\}} \left\{\sum_n \text{Tr}[\hat{\rho}_2 \hat{E}_n] f\left(\frac{\text{Tr}[\hat{\rho}_1 \hat{E}_n]}{\text{Tr}[\hat{\rho}_2 \hat{E}_n]}\right)\right\}. \quad (2.93)$$

It is important to note that the value $D_{f,Q}$ is not only an upper bound on $D_f$ but also it is attainable by a specified measurement $\hat{E}^{\text{opt}}_n$. The $D_{f,Q}$ is physically meaningful only when it can be reached, even in principle, by a real measurement.

Does an explicit formula for $D_{f,Q}$ exists? The answer is that it is known only for two functions – the fidelity (and hence for the Hellinger distance) and the Kolmogorov distance.
2.7. Quantum Fisher information and quantum correlations

An explicit expression for quantum fidelity was derived in [41] and [57], where the following formula was obtained:

\[ f_Q(\hat{\rho}_1|\hat{\rho}_2) \equiv \min_{\{E_n\}} f \left( \text{Tr} \left[ \hat{\rho}_1 \hat{E}_n \right] \bigg| \text{Tr} \left[ \hat{\rho}_2 \hat{E}_n \right] \right) \]  

\[ = \text{Tr} \left[ \sqrt{\hat{\rho}_1^{1/2} \hat{\rho}_2 \hat{\rho}_1^{1/2}} \right]. \]  

(2.94)

Although it is not clear from this equation, quantum fidelity \( f_Q \) is symmetric with respect to change of the density matrices. This property is also obvious from the constructions, as the classical fidelity \( f \) is symmetric with respect to interchange of probability distributions. The formula for the optimal measurement is also known, but has involved form and will not be given here.

With the expression for the quantum fidelity at hand, we can consider the maximization of the Hellinger distance:

\[ H_Q^2 \equiv \max_{\{E_n\}} H^2 \left( \text{Tr} \left[ \hat{\rho}_1 \hat{E}_n \right] \bigg| \text{Tr} \left[ \hat{\rho}_2 \hat{E}_n \right] \right). \]  

(2.96)

With help of the equation (2.36), which relates the Hellinger distance to the fidelity, we obtain that the quantum Hellinger distance is given by:

\[ H_Q^2 = 1 - f_Q(\hat{\rho}_1|\hat{\rho}_2). \]  

(2.97)

In fact, this expression is connected to a definition of the distance between density operators [58]. Uhlmann [59] interpreted this distance as a generalization of the transition probabilities to the mixed states. An explicit formula for the Bures distance was found, which turned out to be [33, 60]:

\[ d_{\text{Bures}}^2(\hat{\rho}_1, \hat{\rho}_2) = 2(1 - f_Q(\hat{\rho}_1|\hat{\rho}_2)). \]  

(2.98)

We present here one more example of a quantum measure. An analytic formula for the quantum Kolmogorov distance was derived in [7], and is given by:

\[ D_{K,Q} \equiv \max_{\{E_n\}} D_K \left( \text{Tr} \left[ \hat{\rho}_1 \hat{E}_n \right] \bigg| \text{Tr} \left[ \hat{\rho}_2 \hat{E}_n \right] \right) = \frac{1}{2} \text{Tr} |\hat{\rho}_1 - \hat{\rho}_2|, \]  

(2.99)

where \( |\hat{A}| = \sqrt{\hat{A}^\dagger \hat{A}} \) is a positive square root of the operator \( \hat{A}^\dagger \hat{A} \). Therefore, \( D_{K,Q} \) is given by the trace-distance between two operators.

Unfortunately, although of great importance, no explicit formula is known for the quantum Kullback-Leibler divergence. However, some lower and upper bounds are known, which can be found in [41, 61].

**Quantum Fisher information**

We now turn to the problem whether an explicit formula for the quantum Fisher information exists. Consider two neighbouring states \( \hat{\rho}_1 = \hat{\rho}(\theta + \delta \theta) \) and \( \hat{\rho}_2 = \hat{\rho}(\theta) \).
If the difference $\delta \theta$ of the parameters is small, we can expand the distance to obtain:

$$D_{f,Q}(\hat{\rho}(\theta + \delta \theta)|\hat{\rho}(\theta)) = \frac{f''(1)}{2} F_Q \delta \theta^2 + \ldots,$$

(2.100)

where the dots represent higher order terms in $\delta \theta$. This formula is analogous to the expansion of $D_{f}$, given by Eq. (2.51), where the classical Fisher information appeared. The object $F_Q$ is called the quantum Fisher information. It is given by the maximum of the classical Fisher information over all possible measurements allowed by quantum mechanics. Therefore, we have

$$F_Q[\hat{\rho}] = \max_{\{E_n\}} F_{cl}[p],$$

(2.101)

where the probability distribution is given by $p_n(\theta) = \text{Tr}[\hat{\rho}(\theta)E_n]$. Notice, that the measurement in (2.101) is independent of the parameter, whose value is only contained in the state.

The Cramér-Rao lower bound (2.61) states then that the optimal measurement is the one, which provides the highest precision possible in the problem of estimation of an unknown parameter. When evaluated for the corresponding measuring operators, the bound takes the following form:

$$\Delta \theta \geq \frac{1}{\sqrt{mF_Q}}.$$  

(2.102)

As the equality can be achieved in the maximum likelihood estimation, the issue of determining an optimal measurement in metrology receives much importance. Unfortunately, this is in general a very difficult task.

The maximization in Eq. (2.101) was performed by Braunstein and Caves in [33]. To understand the explicit formula we have to introduce additional quantities. Suppose that the density matrix $\hat{\rho}(\theta)$ is given in its orthonormal basis:

$$\hat{\rho}(\theta) = \sum_j p_j |j\rangle \langle j|.$$  

(2.103)

Now, upon the small change of $\theta$ by amount $\delta \theta$, eigenvalues $p_j$ change by an amount $dp_j$, and also the eigenvectors change from $|j\rangle$ to $|j'\rangle$. Therefore after the transformation we have

$$\hat{\rho}(\theta + \delta \theta) = \sum_j (p_j + dp_j) |j'\rangle \langle j'|,$$

(2.104)

where the new eigenvectors are given by the transformation generated by the Hermitian operator $\hat{h}$:

$$|j'\rangle = e^{i\delta \theta \hat{h}} |j\rangle.$$  

(2.105)

With that we can write an explicit formula for the quantum Fisher information:

$$F_Q[\hat{\rho}] = \sum_j \frac{1}{p_j} \left( \frac{\partial p_j}{\partial \theta} \right)^2 + \sum_{j,k} \frac{(p_j - p_k)^2}{p_j + p_k} |h_{j,k}|^2,$$

(2.106)
where $h_{j,k}$ are matrix elements of $\hat{h}$ in the orthonormal basis of $\hat{\rho}$, i.e. $h_{j,k} = \langle j | \hat{h} | k \rangle$.

In the second sum terms with both $p_i = 0$ and $p_k = 0$ should be excluded.

In case of pure states when $\hat{\rho} = |\psi\rangle\langle\psi|$, the quantum Fisher information is reduced to the variance of the generator according to the formula:

$$F_Q[|\psi\rangle\langle\psi|] = 4 \left\langle (\Delta \hat{h})^2 \right\rangle_{\psi},$$

(2.107)

where $\Delta \hat{h} = \hat{h} - \langle \hat{h} \rangle_{\psi}$ and the means are taken on the state $|\psi\rangle$. In fact, a general inequality can be written, which bounds the quantum Fisher information from above

$$F_Q[\hat{\rho}] \leq 4 \left\langle (\Delta \hat{h})^2 \right\rangle_{\rho},$$

(2.108)

and the averages of the operators are taken over the state $\hat{\rho}$, and the equality is reached if and only if the state is pure.

The quantum Fisher information can be written in an alternative way using the Symmetric Logarithmic Derivative (SLD) operator $\hat{L}$ defined the relation $\hat{\rho} \hat{L} + \hat{L} \hat{\rho} = 2\hat{\rho}'(\theta)$, where the prime denotes the derivative with respect to the parameter $\theta$. Then, the quantum Fisher information is given by

$$F_Q[\hat{\rho}(\theta)] = \text{Tr}\left[\hat{\rho} \hat{L}^2\right].$$

(2.109)

The optimal measurement for the quantum Fisher information satisfies the relation:

$$\hat{E}_n^{\text{opt}} \hat{\rho}(\theta) = \lambda_n \hat{E}_n^{\text{opt}} \hat{L} \hat{\rho}(\theta),$$

(2.110)

where $\lambda_n$ is a real constant. This condition can always be satisfied with a suitable choice of the operators, and thus the inequality

$$F_{\text{cl}}[\rho(\theta)] \leq F_Q[\hat{\rho}(\theta)],$$

(2.111)

with $p_n(\theta) = \text{Tr}\left[\hat{\rho}(\theta) \hat{E}_n\right]$ can always be saturated with $F_{\text{cl}}$.

We now present two properties of the quantum Fisher information which will be useful in the coming section.

**Convexity.** The first property, called convexity of the quantum Fisher information, is directly related to inequality (2.55), i.e. the convexity of the classical Fisher information. To see this, suppose that the density matrix $\hat{\rho}$ is of the form:

$$\hat{\rho} = \sum_i p_i \hat{\rho}_i,$$

(2.112)

where $p_i$ are some probabilities that sum up to one. Then, after the measurement described by $\hat{E}_n$, the probability of the $n$th outcome is $P_n = \sum_i p_i P_n^{(i)}$, where $P_n^{(i)} = \text{Tr}\left[\hat{\rho}_i \hat{E}_n\right]$. Therefore, we may write

$$F_Q[\hat{\rho}] = \max_{\{E_n\}} F_{\text{cl}}\left[\sum_i p_i P_n^{(i)}\right] \leq \max_{\{E_n\}} \sum_i p_i F_{\text{cl}}\left[P_n^{(i)}\right] \leq \sum_i p_i F_Q[\hat{\rho}_i].$$

(2.113)
The convexity of the quantum Fisher information may compactly be expressed in the form of the following inequality:

\[ F_Q \left[ \sum_i p_i \hat{\rho}_i \right] \leq \sum_i p_i F_Q[\hat{\rho}_i]. \]  

(2.114)

**Additivity.** The second property is related to the additivity exhibited by the classical Fisher information (see formula (2.52)). If the density matrix is of the form

\[ \hat{\rho} = \hat{\rho}^{(1)} \otimes \cdots \otimes \hat{\rho}^{(N)}, \]  

(2.115)

where \( \otimes \) denotes a tensor product of the operators, then the quantum Fisher information is additive with respect to each density matrix:

\[ F_Q\left[ \hat{\rho}^{(1)} \otimes \cdots \otimes \hat{\rho}^{(N)} \right] = \sum_i F_Q[\hat{\rho}^{(i)}]. \]  

(2.116)

This result can be obtained by a straightforward calculation using the formula (2.106). Another way to see this, is to apply the formula for the quantum fidelity (2.95) to a pair of neighbouring states. Using the fact that quantum fidelity satisfies the same multiplicativity relation [37] as the classical counterpart (2.21), its expansion in the small parameter leads to Eq. (2.116).

Here, we close the discussion on the emergence of the quantum Fisher information and its properties. In the next section, we precisely formulate the concept of quantum correlations. The convexity and additivity of the quantum Fisher information will be the crucial properties in establishing a relation between quantum correlations and the quantum Fisher information.

### 2.7.2 Particle entanglement

In this section we introduce the concept of an entanglement of particles. To this end, consider the system that is composed of \( N \) particles. If its state can be written in the following form:

\[ \hat{\rho} = \hat{\rho}^{(1)} \otimes \hat{\rho}^{(2)} \otimes \cdots \otimes \hat{\rho}^{(N)}, \]  

(2.117)

where say that there are no correlations between the particles. In this notation \( \hat{\rho}^{(n)} \) is the density matrix of the \( n \)th particle. Next, if the state can be written in the form of:

\[ \hat{\rho}_{\text{sep}} = \sum_i p_i \hat{\rho}_i^{(1)} \otimes \hat{\rho}_i^{(2)} \otimes \cdots \otimes \hat{\rho}_i^{(N)}, \]  

(2.118)

the particles exhibit only classical correlations and the state \( \hat{\rho}_{\text{sep}} \) is called separable [29]. In this notation, \( p_i \)'s are some probabilities that sum up to one and \( \hat{\rho}_i^{(n)} \) is the \( i \)th density matrix of the \( n \)th particle. On the other hand, if the state of the system cannot be written in the form (2.118), we call it particle entangled or say that the correlations between particles are non-classical. Thus, the state is particle entangled if

\[ \hat{\rho} \neq \sum_i p_i \hat{\rho}_i^{(1)} \otimes \hat{\rho}_i^{(2)} \otimes \cdots \otimes \hat{\rho}_i^{(N)}. \]  

(2.119)
Notice, that the definition of entanglement is constructed by a negation — an entangled state is the state that is not separable. As the separable states are given by construction, we know what types of correlations are in the system. On the other hand, it is not clear what are the properties of entangled states. These can only be discovered by considering specific problems and examples, leading to phenomena like the EPR paradox [4], the Bell inequalities [5, 62], quantum teleportation [63] or superdense-coding [7].

In order to determine if a state is entangled we need a criterion of entanglement. A construction of such criterion should be based merely on this discriminative definition. As a consequence, it should be formulated in the following manner. First, we need a quantity $C$, composed of outcomes obtained in measurements. Second, a general inequality should be established, that for all separable states, the quantity $C$ should be bounded, for instance, from above, i.e. $C_{\text{sep}} \leq 1$. Here, the subscript "sep" indicates that the quantity is calculated with the outcomes of a measurement performed on a separable state. To rule out the possibility that $C \leq 1$ for all states, we need a single example of a state for which $C > 1$. From this a logical conclusion follows that $C$ is the criterion of entanglement. Therefore, if we find that $C > 1$ for a state under consideration, we deduce that the state is entangled. However, obtaining $C \leq 1$ does not imply that the state is separable. In general, $C$ does not provide a clear-cut distinction between separable and entangled states.

Below, we show that the quantum Fisher information is a criterion of entanglement (see also [29]). To this end, assume that a state of the system is separable and given by Eq. (2.118). It is important to note that the Fisher information depends on a transformation of the state. If the quantum Fisher information can be used as a signature of quantumness, the very transformation, which is assumed to be unitary, should not introduce non-classical correlations. As a consequence, it should be of the following form:

$$
\hat{\rho}_{\text{sep}}(\theta) = \hat{U}(\theta)\hat{\rho}_{\text{sep}}\hat{U}^\dagger(\theta)
$$

$$
= \sum_i p_i\hat{\rho}_i^{(1)}(\theta) \otimes \hat{\rho}_i^{(2)}(\theta) \otimes \cdots \otimes \hat{\rho}_i^{(N)}(\theta),
$$

where the unitary operator $\hat{U}(\theta) = \hat{U}_1(\theta) \otimes \hat{U}_2(\theta) \otimes \cdots \otimes \hat{U}_N(\theta)$, and the density matrix of the $n$th particle after transformation is $\hat{\rho}_i^{(n)}(\theta) = \hat{U}_n(\theta)\hat{\rho}_i^{(n)}\hat{U}_n^\dagger(\theta)$. From Eq. (2.120b) we see that the state after transformation is also separable, so the chosen unitary transformation does not produce a state that is outside the set of separable states. Now, let $\hat{h}$ be the generator of the infinitesimal transformation from $\hat{\rho}_{\text{sep}}(\theta)$ to $\hat{\rho}_{\text{sep}}(\theta + \delta\theta)$:

$$
\hat{\rho}_{\text{sep}}(\theta + \delta\theta) = e^{i\hat{h}\delta\theta}\hat{\rho}_{\text{sep}}(\theta)e^{-i\hat{h}\delta\theta},
$$

where $\hat{h} = \sum_n \hat{h}^{(n)}$, and $\hat{h}^{(n)}$ acts on the $n$th particle. From Eqs. (2.104) and (2.105) we observe that the generator of the transformation that appears in the quantum Fisher information is $\hat{h}$ and the probabilities $p_i$ do not change. Now, the quantum
Fisher information can be bounded from above:

\[
F_Q[\hat{\rho}_{\text{sep}}(\theta)] \leq \sum_i p_i F_Q[\hat{\rho}_i^{(1)}(\theta) \otimes \hat{\rho}_i^{(2)}(\theta) \otimes \ldots \otimes \hat{\rho}_i^{(N)}(\theta)].
\] (2.122)

where we used convexity of the quantum Fisher information \((2.114)\). On the right hand side, we use the additivity property of the quantum Fisher information \((2.116)\) which yields:

\[
F_Q[\hat{\rho}_{\text{sep}}(\theta)] \leq N \sum_{n=1}^N \sum_i p_i F_Q[\hat{\rho}_i^{(n)}(\theta)].
\] (2.123)

The quantity \(F_Q[\hat{\rho}_i^{(n)}(\theta)]\) depends on the state \(\hat{\rho}_i^{(n)}(\theta)\) and on the generator of transformation \(\hat{h}_n\). Thus, for given generator, we may maximize it over all possible states of a single particle. Define

\[
\alpha_n \equiv \max_{\hat{\rho}_i^{(n)}} F_Q[\hat{\rho}_i^{(n)}(\theta)].
\] (2.124)

Then, the quantum Fisher information in Eq. \((2.123)\) can be bounded by

\[
F_Q[\hat{\rho}_{\text{sep}}(\theta)] \leq \sum_{n=1}^N \alpha_n.
\] (2.125)

In case of \textit{coherent transformations}, where all generators are the same \(\hat{h}_n = \hat{h}^{(1)}\), \(\alpha_n\) is the same number for all particles, i.e. \(\alpha_n = \alpha\). As a consequence, the quantum Fisher information in Eq. \((2.123)\) can be bounded by

\[
F_Q[\hat{\rho}_{\text{sep}}(\theta)] \leq \alpha N.
\] (2.126)

Therefore, we conclude that for separable states of \(N\) particles, undergoing a coherent transformation, the quantum Fisher information is bounded from above by the number that is proportional to the number of particles in the system. The limit \(\alpha N\), with the quantum Fisher information scaling linearly with the number of particles, is called the \textit{shot-noise limit}. The inequality \((2.126)\) bounds the precision of estimation of the phase \(\theta\) in any estimation scheme:

\[
\Delta \theta \geq \frac{1}{\sqrt{m}} \frac{1}{\sqrt{\alpha}} \frac{1}{\sqrt{N}}.
\] (2.127)

As a consequence if for a for chosen generator we find that

\[
F_Q[\hat{\rho}(\theta)] > \alpha N,
\] (2.128)

we conclude that the state is not separable, and thus it is entangled. Therefore, the quantum Fisher information is a criterion of particle entanglement.

The parameter \(\alpha\) in the shot-noise limit definition is given by the maximum of the quantum Fisher information for a single particle:

\[
\alpha \equiv \alpha_1 = \max_{\hat{\rho}} F_Q[\hat{\rho}^{(1)}(\theta)].
\] (2.129)
2.7. Quantum Fisher information and quantum correlations

The maximization can be analytically calculated, because we know that the maximum is reached for pure states. Denote \( |\psi\rangle \) a pure state after the unitary transformation, then \( \alpha \) is given by the variance of the single-particle generator \( \hat{h}^{(1)} \) (see formula (2.108)),

\[
\alpha = 4 \max_{|\psi\rangle} \left\langle (\Delta \hat{h}^{(1)})^2 \right\rangle_{\psi},
\]

and the maximization is performed over all states \( |\psi\rangle \) of the single particle. This maximum may be calculated explicitly and is given by

\[
\alpha = (h_M - h_m)^2 \text{ for the state } |\psi\rangle = \frac{|h_M\rangle + |h_m\rangle}{\sqrt{2}},
\]

where \( h_M \) and \( h_m \) are maximum and minimum eigenvalues of \( \hat{h}^{(1)} \) with corresponding eigenvectors \( |h_M\rangle \) and \( |h_m\rangle \), respectively.

We know that the quantum Fisher information can be bounded by the variance of the generator of the transformation (cf. 2.108):

\[
F_Q[\hat{\rho}(\theta)] \leq 4 \left\langle (\Delta \hat{h})^2 \right\rangle.
\]

The equality is achieved if and only if \( \hat{\rho} \) is a pure state. In the same manner as above, the maximum of the variance over the states can be calculated analytically, which leads to:

\[
\max_{\hat{\rho}} F_Q[\hat{\rho}] = N^2 (h_M - h_m)^2 \text{ for the state } |\psi\rangle = \frac{|h_M\rangle \otimes N + |h_m\rangle \otimes N}{\sqrt{2}}.
\]

This limiting value is called the Heisenberg limit, and the corresponding uncertainty of precision is given by

\[
\Delta \theta \geq \frac{1}{\sqrt{m}} \frac{1}{\sqrt{\alpha}} \frac{1}{N}.
\]

It is possible to benefit from the non-classical correlations and to improve the limit on precision from Eq. (2.127) to the one given by Eq. (2.134), which is better by a factor of \( \sqrt{N} \).

The quantum Fisher information is the measure of susceptibility of the state when a fixed action is performed on the system composed of \( N \) particles. The action is parametrized by the unitary transformation, which is also related to the generator of the transformation. It is important that these transformations do not create entanglement. Therefore, for separable states the change of the state is limited, which is signified by the existence of the upper bound on susceptibility, i.e. \( F_Q \leq \alpha N \). On the other hand, there exist entangled states (like the one in Eq. (2.133)), for which this bound is violated. This is related to the improved sensitivity of the precise estimation of the phase in the interferometer. Therefore, the quantum Fisher information is a quantity that splits the class of entangled states into two. One class is composed of states for which \( F_Q \leq \alpha N \) (the same as for separable states), and these are not much useful in metrology. The second class
is formed of the states with $F_Q > \alpha N$, which we call *usefully entangled*, because, apart from being entangled, they lead to improved sensitivity.

As a final remark, we may note that the generalization to the state with fluctuating total number of particles is also possible. In this thesis all the systems that are analysed are composed of massive particles. For these, the super selection rules apply \[64–67\] forbidding the creation of states with coherences between different number of particles in the system. Therefore, the density matrix of the separable state of the system with fluctuating number of particles can be written as:

$$\hat{\rho}_{\text{sep}} = \sum_{N=0}^{\infty} P_N \sum_i p(i|N) \hat{\rho}_{i|N}^{(1)} \otimes \hat{\rho}_{i|N}^{(2)} \otimes \cdots \otimes \hat{\rho}_{i|N}^{(N)}, \quad (2.135)$$

where $P_N$ is the probability that there are exactly $N$ particles observed in the system, $p(i|N)$ are some probabilities that sum up (over index $i$) to one; the $i$th density matrix of the $n$th particle, when measured $N$ particles in total, is given by $\hat{\rho}_{i|N}^{(n)}$. For this kind of states, an analysis similar to the above can be performed \[68\], leading to the result that the quantum Fisher information is bounded from above by:

$$F_Q[\hat{\rho}_{\text{sep}}] \leq \alpha \langle N \rangle , \quad (2.136)$$

where the mean number of particles $\langle N \rangle$ is calculated over $P_N$. If the quantum Fisher information beats the shot-noise limit, $F_Q > \alpha \langle N \rangle$, useful entanglement is present in the system.

In the preceding sections we presented a connection between the notion of non-classical correlations between particles and the quantum Fisher information. For particle entangled states the value of the quantum Fisher information is larger than the largest value reached by any classically correlated state. This relation can be viewed in three-fold perspectives.

First, for a fixed difference in a parameter that governs the change of the state, there are non-classical states that are more distinguishable than any state with classical correlations. Second, if the change of the state can be controlled, the non-classical states are more useful for metrology as they lead to more precise parameter estimation. Third, if the change of the state is due to some disturbance, the non-classical states are more volatile than classical ones, which indicates that particle entanglement is fragile and can easily be destroyed.

Because there are different estimation schemes of the parameter, each of them gives different estimate of the uncertainty of that parameter. From monotonicity we know that each such uncertainty bounds the value of the Fisher information from below. As the quantum Fisher information is the highest possible attainable value of the classical Fisher information, we obtain a powerful method for indicating whether a given state exhibits non-classical correlations. Moreover, these correlations can be used in an interferometer to improve the sensitivity of the measurement. Therefore, the class of systems with non-classical correlations is called *usefully entangled*. 
2.7.3 Two mode interferometry with bosons

In this section we consider $N$ indistinguishable particles that occupy two modes ($N$ qubits). The annihilation operators of the modes are denoted with $\hat{a}$ and $\hat{b}$. An orthonormal basis is given in the occupation basis:

$$|n,N-n\rangle = (\hat{a}^\dagger)^n (\hat{b}^\dagger)^{N-n} \sqrt{n!} \sqrt{(N-n)!} |0\rangle,$$

(2.137)

where $|0\rangle$ is the state of vacuum. The state $|n,N-n\rangle$ denotes $n$ particles in $a$ and $N-n$ in $b$.

A pure state of the system can be written as:

$$|\psi\rangle = \sum_{n=0}^{N} C_n |n,N-n\rangle,$$

(2.138)

and the amplitudes satisfy normalization condition, $\sum_n |C_n|^2 = 1$.

Now, we introduce unitary transformations that describe actions of various interferometers. These are not allowed to introduce an entanglement into the system, and so they must be built of single particle operations. In such case, the unitary operation can be written in terms of the vector of Pauli operators $\hat{\sigma}$; for example, $\hat{\sigma}_z$ acting on the state $|1,0\rangle$ (a particle in the mode $a$) gives $+1$, and on the state $|0,1\rangle$ (a particle in the mode $b$) gives $-1$. Then, a single particle unitary transformation is given by $\exp(i\frac{\phi}{2} \hat{n} \cdot \hat{\sigma})$, where $\hat{n}$ is a unit vector pointing in the direction of rotation of a qubit, and $\phi$ is an angle of the rotation. As a consequence, a unitary transformation which does not create particle entanglement is

$$\hat{U} = e^{i\frac{\phi}{2} \hat{n} \cdot \hat{\sigma}(1)} \otimes e^{i\frac{\phi}{2} \hat{n} \cdot \hat{\sigma}(2)} \otimes \cdots \otimes e^{i\frac{\phi}{2} \hat{n} \cdot \hat{\sigma}(N)},$$

(2.139)

where $\hat{\sigma}^{(n)}$ acts on the $n$th particle. This can be rewritten as

$$\hat{U}(\phi) = e^{i\phi \hat{n} \cdot \hat{J}},$$

(2.140)

where the total angular momentum operators are given by the generator $\hat{J} = \frac{1}{2} \sum_{j=1}^{N} \hat{\sigma}^{(j)}$. Note, that this is a sum of single particle operators and does not create entanglement. The operation (2.140) can be interpreted as a coherent transformation of qubits. The angular momentum operators can alternatively be written in terms of annihilation and creation operators (Schwinger representation):

$$\hat{J}_x = \frac{\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a}}{2},$$

(2.141a)

$$\hat{J}_y = \frac{\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a}}{2i},$$

(2.141b)

$$\hat{J}_z = \frac{\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b}}{2}.$$  

(2.141c)

The $\hat{J}_z$ can be written in terms of number operators of the modes as $\hat{J}_z = \frac{1}{2}(\hat{n}_a - \hat{n}_b)$, where the occupation number operators are $\hat{n}_a = \hat{a}^\dagger \hat{a}$ and $\hat{n}_b = \hat{b}^\dagger \hat{b}$. From this, a
natural mapping between the occupation number states (2.137) and eigenstates of $\hat{J}_z$ can be established in the form:

$$|J, m_z\rangle \leftrightarrow |n_a, n_b\rangle \quad \text{with} \quad J = \frac{N}{2}, m_z = n - \frac{N}{2} = |n, N - n\rangle. \quad (2.142)$$

With help of angular momentum operators a number of devices known from quantum optics can be described [3]. For example $\exp(i\frac{\pi}{2}\hat{J}_x)$ is the symmetric beamsplitter (50/50), whereas the transformation composed of three rotations, given by $\exp(i\frac{\pi}{2}\hat{J}_x)\exp(i\theta J_y)\exp(-i\frac{\pi}{2}\hat{J}_x) = \exp(i\theta \hat{J}_y)$, is the Mach-Zehnder interferometer with relative phase difference $\theta$ imprinted onto its arms.

Suppose that an interferometric device shown in the second step in Fig. 2.3 is given by $\hat{U}(\theta) = \exp(i\theta \hat{n} \cdot \hat{J})$. The interferometric phase, which we wish to estimate with the highest precision possible, is $\theta$. After the transformation, the pure state $|\psi\rangle$, becomes $|\psi(\theta)\rangle = e^{i\theta \hat{n} \cdot \hat{J}} |\psi\rangle$. The best precision is related to the quantum Fisher information, which for pure states is given by the variance (cf. (2.108)):

$$F_Q = 4 \left[ \langle (\hat{n} \cdot \hat{J})^2 \rangle - \langle \hat{n} \cdot \hat{J} \rangle^2 \right], \quad (2.143)$$

where the averages are calculated on the state $|\psi(\theta)\rangle$, or alternatively $|\psi\rangle$, because the variance in this case does not depend on the phase. Quantum Fisher information is equal to the right hand side in Eq. (2.143) only for pure states. In case of mixed states, right hand side is only an upper bound for the quantum Fisher information.

The condition for the presence of particle entanglement useful for interferometry is given by Eq. (2.136), where $\alpha$ is the single particle property of the generator. This is related, by Eq. (2.131), with the maximal and the minimal eigenvalue of $\hat{n} \cdot \hat{\sigma} / 2$. These are $\pm 1/2$, and therefore $\alpha = 1$. The condition for useful entanglement becomes:

$$F_Q > \langle N \rangle. \quad (2.144)$$

Below, we give two simple examples of states that are particle entangled, which can be employed for sub-shot-noise interferometry.

**NOON state.** The first example is the NOON state of $N$ particles:

$$|\text{NOON}\rangle = \frac{1}{\sqrt{2}} \left( |N, 0\rangle + |0, N\rangle \right). \quad (2.145)$$

This state is of the form (2.133), therefore, if the interferometer $U_z(\theta) = \exp(i\theta \hat{J}_z)$ is applied to the NOON state, the quantum Fisher information achieves its maximum possible value:

$$F_Q = N^2. \quad (2.146)$$
2.7. Quantum Fisher information and quantum correlations 55

**Twin Fock state.** In the *Twin-Fock state* particles are equally distributed between the modes:

\[ |TF\rangle = \left| n_a = \frac{N}{2}, n_b = \frac{N}{2} \right\rangle. \] (2.147)

If this state is taken as the input of the interferometer described with \( \hat{U}_x(\theta) = \exp(i\theta \hat{J}_x) \) or \( \hat{U}_y(\theta) = \exp(i\theta \hat{J}_y) \), then the quantum Fisher information is

\[ F_Q = \frac{N(N + 1)}{2}. \] (2.148)

For large \( N \), this value is as big as the half of the Heisenberg limit.

**2.7.4 Spin-squeezing and number-squeezing**

In the final section of this chapter we present a simple quantity which can serve as a criterion of entanglement between \( N \) bosonic qubits. Note that the precision of a specified estimation scheme is lower than the limit set by the quantum Fisher information. We can invoke the inequality (2.79), which bounds the classical Fisher information, if the estimation is performed from the mean value of the variable \( X \), where for \( X \) we take \( \hat{J}_{n'} \equiv n' \cdot \hat{J} \), where \( n' \) is a unit vector. For example, in case of \( \hat{J}_z \), due to Eq. (2.141c), this amounts to the measurement of the difference between number of atoms in modes \( a \) and \( b \). Therefore, we have the following inequality for generator \( \hat{J}_{n'} \):

\[ \frac{1}{(\Delta \hat{J}_{n'}^2)^2} \left( \frac{\partial \langle \hat{J}_{n'} \rangle}{\partial \theta} \right)^2 \leq F_{\text{cl}} \leq F_Q. \] (2.149)

As \( F_Q > N \) signals the useful entanglement, we have the following criterion of entanglement:

\[ N < \frac{1}{(\Delta \hat{J}_{n'}^2)^2} \left( \frac{\partial \langle \hat{J}_{n'} \rangle}{\partial \theta} \right)^2. \] (2.150)

In this notation \( (\Delta \hat{J}_{n'})^2 \) is the variance of \( \hat{J}_{n'} \) operator. If this inequality is satisfied, then the useful entanglement is present, because then \( F_Q > N \). Below, we reformulate this condition in the alternative form. Since \( \langle \hat{J}_{n'} \rangle = \text{Tr}[\hat{\rho} \hat{J}_{n'}] \), we can differentiate under the trace, which yields

\[ \frac{\partial \langle \hat{J}_{n'}(\theta) \rangle}{\partial \theta} = (n \times n') \cdot \langle \hat{J}(\theta) \rangle, \] (2.151)

where mean angular momentum operator vector is expressed by \( \langle \hat{J}_k(\theta) \rangle = \text{Tr}[\hat{\rho}(\theta) \hat{J}_k] \), with \( k = x, y, z \). Therefore, if the inequality

\[ N \frac{(\Delta \hat{J}_{n'}(\theta))^2}{(n \times n') \cdot \langle \hat{J}(\theta) \rangle} < 1 \] (2.152)

is true, the correlations between particles exhibit non-classical behaviour and the state is usefully entangled.
From this inequality another criterion can be obtained. Suppose that we investigate only the value of that parameter at $\theta = 0$ and regard the vectors $\mathbf{n}$ and $\mathbf{n}'$ as arbitrary. Then

$$N \frac{(\Delta \hat{J}_n')^2}{(\mathbf{n}' \times \mathbf{J}) \cdot \mathbf{n})^2} < 1,$$  \hspace{1cm} (2.153)

where $\langle \mathbf{J} \rangle = \mathbf{J}$. Here, the unit vector $\mathbf{n}$ can be chosen such that the left hand side is as small as possible. It is possible only if $\mathbf{n}$ points in the direction of $\mathbf{n}' \times \mathbf{J}$. Then, the condition for entanglement can be rewritten as:

$$N \frac{(\Delta \hat{J}_n')^2}{(\mathbf{n}' \times \mathbf{J})^2} < 1,$$  \hspace{1cm} (2.154)

We can write this inequality in an equivalent form, by introducing three mutually orthonormal vectors $\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{n}_3$, where we set $\mathbf{n}_1 = \mathbf{n}'$. The condition for the useful entanglement is now

$$\xi_s^2 \equiv \frac{N(\Delta \hat{J}_{n_1})^2}{\langle \hat{J}_{n_2} \rangle^2 + \langle \hat{J}_{n_3} \rangle^2} < 1.$$  \hspace{1cm} (2.155)

The parameter $\xi_s$ is called the spin-squeezing, and if $\xi_s < 1$ the state of the system is called spin-squeezed in the direction of $\mathbf{n}_1$ [69–72]. Note that if the total number of atoms in the system is fluctuating, but the state has no coherences between different number of particles, i.e. it is given by Eq. (2.135), then $N$ in the definition of $\xi_s$ can be replaced with $\langle N \rangle$.

Let us assume now that $\mathbf{n}_1$, $\mathbf{n}_2$ and $\mathbf{n}_3$ are directed along $z$-, $x$- and $y$-axes, respectively. In this case, the spin-squeezing parameter is given by

$$\xi_s^2 = \frac{N(\Delta \hat{J}_z)^2}{\langle \hat{J}_x \rangle^2 + \langle \hat{J}_y \rangle^2}.$$  \hspace{1cm} (2.156)

This quantity can be interpreted as follows. The variance of $\hat{J}_z$ can be written as the variance of $\frac{1}{2} \hat{n}$, where $\hat{n} = \hat{n}_a - \hat{n}_b$ is the difference between number operators in mode $a$ and $b$; it describes the imbalance between populations of the modes. Inserting the definitions of operators $\hat{J}_x$ and $\hat{J}_y$, given in Eqs. (2.141), into Eq. (2.156), we obtain a particularly simple form of $\xi_s$:

$$\xi_s^2 = \frac{\eta^2}{\nu^2}.$$  \hspace{1cm} (2.157)

The parameter $\eta$, called the number-squeezing, is given by the variance of $\hat{n}$:

$$\eta^2 = \frac{\langle (\Delta \hat{n})^2 \rangle}{N}.$$  \hspace{1cm} (2.158)

It relates the fluctuations of populations between the two modes to the Poisson distribution. If $\eta < 1$ we say that the fluctuations of population imbalance are sub-Poissonian and the state is number-squeezed, whereas for $\eta > 1$ they are called
super-Poissonian. It should be emphasized that number-squeezed states do not necessarily indicate the presence of useful entanglement.

The second parameter that appears in Eq. (2.157) is the $\nu^2$, which is given by the following formula:

$$\nu^2 = \frac{4\langle \hat{b}^\dagger \hat{a} \rangle^2}{N^2}. \quad (2.159)$$

The quantity in the numerator is the coherence between the modes $a$ and $b$, the denominator serves as the normalization so that $\nu \leq 1$. To see this, we use the Schwarz inequality for the operators: $|\langle \hat{b}^\dagger \hat{a} \rangle|^2 \leq \langle \hat{n}_b \rangle \langle \hat{n}_a \rangle \leq \frac{1}{4} N^2$. Therefore, $\nu$ quantifies the visibility of the fringes in the interference experiment, when two spatially separated wavepackets of the two modes overlap in the far field. The maximum visibility results in $\nu = 1$, whereas the absence of the fringes leads to $\nu = 0$.

As a final note, one may generalize all these results to systems with fluctuating number of particles (when the super-selection rule applies). In such a case, all the equations starting from (2.152) are valid when for as long as $N$ is replaced with $\langle N \rangle$.

To summarize, the combination of a number-squeezing with a presence of sufficiently high visibility indicates that a state of the system is usefully entangled and can be employed for sub-shot-noise interferometry. The condition that signals non-classical correlations is the inequality $\xi_s = \frac{\eta}{\nu} < 1$. It is sometimes easy to check if $\eta < 1$. However, if it happens that $\eta \leq 1$, spin-squeezing parameter cannot be smaller than 1. This results from $\xi_s = \frac{\eta}{\nu} \geq \eta \geq 1$, where the condition that visibility $\nu \leq 1$ was used. In fact, there is a more severe inequality for visibility: $\nu^2 \leq 1 - x^2$, where $x = \langle \hat{n} \rangle / N$. This leads to the conclusion that spin-squeezing is absent in the system, if the inequality $\eta^2 \geq 1 - x^2$ is satisfied. In this case, useful entanglement can still be present in the system, but the specific criterion $\xi_s < 1$ cannot serve as a signature of non-classical correlations, and one has to refer to a different criterion.

\section{2.8 Conclusions}

In this chapter we introduced the concept of the distinguishability of states, that was analysed on two levels – statistical and quantum one.

To understand to what degree two probability distributions are different, we defined a few measures: the fidelity, the Hellinger distance, the Kolmogorov distance and the Kullback-Leibler divergence. These objects are unified by the $f$-divergence. It turned out that the $f$-divergence satisfies important properties like monotonicity under transformation of variables, and joined convexity. The considerations of the neighbouring probability distributions revealed that the $f$-divergence can be described by a single parameter called the classical Fisher information. This object inherited important information properties: the monotonicity under transformation of variables, the additivity and the convexity.
We also pointed out the connection between the Fisher information and a metro-
 logical problem of estimating an unknown parameter. The uncertainty of the estima-
tion is, according to the Cramér-Rao theorem, bounded from below by the classical 
Fisher information. The best precision can be obtained in the scheme called the 
maximum-likelihood estimation in the limit of very large data samples. We also 
argued that the estimation from the mean of a physical variable implies in general – due to monotonicity – a lower precision. Nevertheless, these schemes provided 
useful bounds for the Fisher information.

In the analysis of the distinguishability of states at the quantum level, we ob-
served that quantum mechanics allows to choose a measurement which is performed 
on the system. Some choices may result in better discrimination of the states than 
the others, so it is important to take the optimal one. This led to the quantum 
measures of distinguishability of states.

This degree of freedom was used to maximize the classical Fisher information 
over all measurements, leading to the quantum Fisher information. We showed that 
the quantum Fisher information is a criterion of entanglement. It indicates there 
are entangled states which are highly susceptible under some transformations. This 
susceptibility can be used in interferometers yielding an improved sensitivity and a 
better precision estimation than any other classically correlated state. In this way, 
we introduced the concept of useful entanglement.
Optimal measurements in phase estimation: simple examples

In this chapter\(^1\), we apply the mathematical formalism, that was developed in the previous chapter, to metrological problems in a simple two-mode system. We identify optimal measurement strategies for phase estimation in different interferometric scenarios. For pure states of a single qubit, we show that optimal measurements form a broad set parametrized with a continuous variable. When the state is mixed this set reduces to merely two possible measurements. For the two-qubit symmetric Werner state, we find the optimal measurement and show that estimation from the population imbalance is optimal only if the state is pure. We also determine the optimal measurements for a wide class of symmetric $N$-qubit Werner-like states. Finally, for a pure symmetric state of $N$ qubits, we finds under which conditions the estimation from the full $N$-body correlation and from the population imbalance is optimal.

---

3.1 Introduction

Quantum interferometry aims at estimating an unknown phase value $\theta$ with the smallest possible uncertainty $\Delta \theta$, taking advantage of quantum correlations in the probe state. The phase estimation protocol generally consists of four steps (see Fig. 2.3), which were discussed in details in chapter 2 in section 2.6 (see also recent reviews on quantum-enhanced phase estimation [73–75]).

First is the preparation of $N$ particles in an input (probe) state $\hat{\rho}_0$, which then undergoes the phase-dependent transformation (determined by the choice of the interferometer). Finally, a measurement, which is in general described by the quantum-mechanical positive-operator valued measures (POVMs), is carried out at the output. Usually, the protocol is repeated $m$ times and the phase shift is inferred using an estimator $\theta_{\text{est}}$, which is a properly-chosen function of the collected results.

There are good and obvious bad choices of the estimator. The good ones are those which are unbiased and minimize the phase sensitivity $\Delta \theta$, defined as the statistical mean square fluctuation of the estimator. The Cramér-Rao theorem sets a lower bound on the sensitivity of any estimator [36, 76], i.e. $\Delta \theta \geq \Delta \theta_{\text{CR}}$. For unbiased estimators and independent measurements we have

$$\Delta \theta_{\text{CR}} = \frac{1}{\sqrt{mF}},$$

(3.1)

where $F$ is the Fisher information (FI), which depends on all the steps of the interferometric sequence. As we argued in section 2.6.1, the Cramér-Rao lower bound (CRLB) from Eq. (3.1) is saturated by the maximum likelihood estimator when the number of measurements tends to infinity.

A particularly important value of the FI is expressed in terms of the number of particles $N$ passing through the interferometer. For separable states the FI is bounded by the shot-noise limit (SNL), i.e. $F \leq N$. In such case, according to the CRLB the phase sensitivity is not better then $\Delta \theta \geq \frac{1}{\sqrt{mN}}$ [77]. The SNL is not fundamental [29, 69, 71, 75, 77, 78] and can be surpassed when the particles are prepared in an entangled state [29]. Moreover, for a fixed $N$, the value of the FI may raise by increasing the amount of particle entanglement in the probe [79, 80]. The ultimate bound is the Heisenberg limit (HL) $F = N^2$ giving $\Delta \theta \geq \frac{1}{N\sqrt{m}}$, reached when the phase is imprinted on the maximally entangled NOON state $|\psi\rangle = \frac{1}{\sqrt{2}}(|N0\rangle + |0N\rangle)$ [77]. Recent experiments acheive sensitivities below the SNL with entangled ions [81–84], photons [85–88], cold [20, 89, 90] and ultracold [18, 19, 24, 91–93] atoms.

Note that once the input state and the interferometric transformation are fixed, the value of the FI can be optimized with respect to all possible POVMs. This maximal value is the quantum Fisher information (QFI) [33, 76], denoted by $F_Q$. In some situations, especially when the input state is pure, the QFI has a simple form [33, 34]. In such case the optimal measurements can be identified: they are related to the projections over the output state and over the orthogonal subspace...
3.2. Single qubit

We begin our analysis with the basic two-mode object, which is a single qubit. Its density matrix is represented by a combination Pauli matrices $\hat{\sigma}$ (hereafter bold symbols will indicate vectors)

$$\hat{\rho}_0 = \frac{1}{2}(\hat{1} + s_0 \cdot \hat{\sigma}).$$

(3.2)

Depending on the length of the vector $s_0$, the state is either pure (if $|s_0| = 1$) or mixed ($|s_0| < 1$). We now consider a rotation of the density matrix $\hat{\rho}_0$ by an angle $\theta$ around the $y$-axis

$$\hat{U}_{\text{mzi}}(\theta) = e^{-i\theta \hat{\sigma}_y}.$$  

(3.3)

The choice of such transformation is dictated by the fact that it is commonly implemented in experiments. Namely, when the two modes are linked with the external degrees of freedom, the operator (3.3) represents the Mach-Zehnder interferometer (MZI) [71,78]. For manipulations between internal degrees of freedom, such transformation is called the Ramsey interferometer. The output density matrix reads

$$\hat{\rho} = \hat{U}_{\text{mzi}}(\theta) \hat{\rho}_0 \hat{U}_{\text{mzi}}^\dagger(\theta) = \frac{1}{2}(\hat{1} + s \cdot \hat{\sigma}),$$

(3.4)

where the three components of the rotated vector are

$$s_x = s_{0,x} \cos \theta + s_{0,z} \sin \theta$$

(3.5a)

$$s_y = s_{0,y}$$  

(3.5b)

$$s_z = s_{0,z} \cos \theta - s_{0,x} \sin \theta.$$  

(3.5c)

Note that we have omitted the explicit dependence of $\hat{\rho}$ and $s$ on the phase $\theta$ – in order to simplify the notation. The task is to determine the value of the angle $\theta$
with the precision maximized with respect to all possible measurements. As we show below, for a single-qubit a full family of optimal measurements can be identified, both for pure and mixed states.

### 3.2.1 Classical and quantum Fisher information

The broadest family of measurement allowed by quantum mechanics is related to POVMs, which include the commonly used projective measurements [76]. A POVM is formed by a complete set of self-adjoint operators with non-negative eigenvalues. For a single qubit, any POVM is represented by the following family of operators

\[ \hat{E}_q = \gamma_q (\hat{1} + q \cdot \hat{\sigma}), \]  

(3.6)
parametrised by \( q \). The conditions

\[ |q| \leq 1, \quad \text{and} \quad \gamma_q > 0, \]  

(3.7)

guarantee that \( \hat{E}_q \) is nonnegative definite, while

\[ \int dq \gamma_q = 1, \quad \text{and} \quad \int dq \gamma_q q = 0, \]  

(3.8)

ensures the completeness of the POVM set, i.e. \( \int dq \hat{E}_q = \hat{1} \). The following trace

\[ p(q|\theta) = \text{Tr}[\hat{\rho} \hat{E}_q]. \]  

(3.9)

is in fact the probability for finding the qubit aligned along the direction \( q \).

The FI, which measures the amount of information about \( \theta \) contained in \( p(q|\theta) \), reads

\[ F = \int dq \frac{1}{p(q|\theta)} \left( \frac{\partial p(q|\theta)}{\partial \theta} \right)^2. \]  

(3.10)

Clearly, through Eq. (3.9), \( F \) depends on the choice of the POVM. For a given probe state and interferometric transformation, some output measurements are better than others as they give a larger FI (for an experimental demonstration see Ref. [94]). The optimal measurements are those which give the maximal value of the FI. The maximization procedure of (3.10) can be performed analytically [33] and the resulting QFI for unitary transformations is

\[ F_Q = 2 \sum_{j,k} \frac{(p_j - p_k)^2}{p_j + p_k} |\langle j | \hat{\sigma}_y | k \rangle|^2. \]  

(3.11)

The ket \( |k\rangle \) denotes the eigenvector of the density matrix \( \hat{\rho} \) with eigenvalue \( p_k \), i.e. \( \hat{\rho} = \sum_k p_k |k\rangle\langle k| \), with \( p_k \geq 0, \sum_k |k\rangle\langle k| = \hat{1} \). The sum in Eq. (3.11) extends over values \( p_j + p_k \neq 0 \). For the matrix \( \hat{\rho} \) from Eq. (3.4) we obtain

\[ F_Q = s_x^2 + s_z^2, \]  

(3.12)

which does not depend on \( \theta \). The maximum \( F_Q = 1 \) is for pure states (\( |s| = 1 \) giving \( F_Q = 4(\Delta \hat{\sigma}_y)^2 \)) with a vanishing component along the rotation, \( s_y = 0 \).
3.2. Single qubit

3.2.2 Optimal measurements

Next, we determine the optimal measurements for which the FI from Eq. (3.10) equals the QFI from Eq. (3.12). As we presented in Eq. (2.110), a POVM is optimal if and only if it satisfies the condition [33]:

$$\hat{E}_q \hat{\rho} = \lambda_q \hat{E}_q \hat{L}_\rho \hat{\rho}$$

(3.13)

with $\lambda_q \in \mathbb{R}$. The super-operator $\hat{L}_\rho$ is the symmetric logarithmic derivative, defined by the relation

$$\frac{\partial \hat{\rho}}{\partial \theta} = \hat{\rho} \hat{L}_\rho + \hat{L}_\rho \hat{\rho} \frac{\partial}{\partial \theta}.$$  

(3.14)

When $\hat{\rho}$ can be inverted (which is not the case of pure states), we multiply Eq. (3.13) by $\hat{\rho}^{-1}$ and get

$$\hat{E}_q (\hat{1} - \lambda_q \hat{L}_\rho) = 0,$$

(3.15)

which is fulfilled if the POVM $\hat{E}_q$ is constructed from the projectors over the eigenstates of $\hat{L}_\rho$ with $\lambda_q$ equal to the inverse of the corresponding real eigenvalue.

In all cases, the optimal measurements are found in two steps. First, we write down the symmetric logarithmic derivative (SLD) in its general form

$$\hat{L}_\rho = \hat{1} \ell + \mathbf{n} \cdot \hat{\sigma}.$$  

(3.14)

Here $\ell$ is a real number and $\mathbf{n}$ is a three-dimensional vector. Using Eqs. (3.4) and (3.14), we obtain the expressions for $\ell$ and $\mathbf{n}$. The result is that $\ell = 0$, while $\mathbf{n} = s_x e_x - s_x e_z \equiv s_\perp$. Here $e_x$, $e_y$ and $e_z$ are the three orthogonal unit vectors and $s \cdot s_\perp = 0$. Therefore, the SLD reads

$$\hat{L}_\rho = s_\perp \cdot \hat{\sigma}.$$  

(3.16)

In the second step, we insert $\hat{L}_\rho$ into Eq. (3.13) and use the general parametrization of the POVM, Eq. (3.6). By comparing the scalar and vector parts and then the real and imaginary parts, we obtain the set of equations

$$\mathbf{q} \cdot (s_\perp \times \mathbf{s}) = 0$$

(3.17a)

$$\lambda_q = \frac{1 + \mathbf{q} \cdot \mathbf{s}}{\mathbf{q} \cdot s_\perp}$$

(3.17b)

$$\mathbf{q} + \mathbf{s} = \lambda_q [s_\perp - \mathbf{q} \times (s_\perp \times \mathbf{s})]$$

(3.17c)

$$\mathbf{q} \times \mathbf{s} = \lambda_q [s_\perp \times \mathbf{s} + \mathbf{q} \times s_\perp].$$

(3.17d)

From Eq. (3.17a) we deduce that $\mathbf{q}$ lies in the plane spanned by vectors $\mathbf{s}$ and $s_\perp$, so it can be written as $\mathbf{q} = q_1 \mathbf{e}_s + q_2 \mathbf{e}_{s_\perp}$. Here, $\mathbf{e}_s$ and $\mathbf{e}_{s_\perp}$ are unit vectors pointing into directions $\mathbf{s}$ and $s_\perp$, respectively. Both vectors are orthogonal to the rotation direction $\mathbf{e}_y$. These observations reduce the set of eight equations (3.17) to

$$\lambda_q = \frac{1 + s q_1}{s_\perp q_2}$$

(3.18a)

$$q_1 + s = \lambda_q s_\perp q_2$$

(3.18b)

$$q_2 s = \lambda_q s_\perp (s - q_1)$$

(3.18c)

$$q_2 = \lambda_q s_\perp (1 - sq_1),$$

(3.18d)
where \( s = |s| \) and \( s_\perp = |s_\perp| \). This set of four equations for the three variables \( q_1, q_2 \) and \( \lambda q \) is non-contradictory when two of these equations are linearly dependent. If the state is pure \((s = 1)\), Eqs. (3.18c) and (3.18d) are equivalent and the solution is \( q_1^2 + q_2^2 = 1 \). Thus for pure states there are several choices of optimal POVM. The obvious choice is to take the projection over the output state \((q_1 = 1, q_2 = 0)\) and over the orthogonal vector \((q_1 = 0, q_2 = 1)\), with \( \gamma_q = 1/2 \). However, we can also consider POVMs with two elements along different directions (still satisfying \( q_1^2 + q_2^2 = 1 \)), or even a continuous set of POVM parametrized by a vector \( q \) which lies on a circle of unit radius.

If the state is mixed \((s < 1)\), Eqs. (3.18c) and (3.18d) are non-contradictory only if \( q_1 = 0 \) and the other two equations give \( q_2 = \pm 1 \). There is only one optimal POVM and its elements are \((\hat{1} \pm e_{s_\perp} \cdot \sigma)/2\). This optimal POVM corresponds to the projection over pure states pointing along the \( \pm e_{s_\perp} \) direction in the Bloch sphere. These are the eigenstates of Eq. (3.16) with (real) eigenvalues \( \pm s_\perp \). In other words, the POVM with elements \((\hat{1} \pm e_{s_\perp} \cdot \sigma)/2\) is necessary and sufficient to saturate the QFI for mixed state while it is only sufficient for pure states (for which the density matrix is non invertible).

In conclusion, there is a substantial difference between the sets of optimal measurements for pure and mixed states. The possible choice of POVM in the continuous set for pure states reduces to just two projection operators for mixed states, as schematically shown in Fig. 3.1.

![Figure 3.1: Schematic representation of the directions of vectors \( q \) for which the POVM set (3.6) is optimal. The vector \( s_\perp \) is orthogonal to \( s \) and both lie in the \( x - z \) plane. For pure states the elements of the optimal POVM can be arbitrary chosen in the unit (gray) circle. For mixed states the optimal POVM set is made only by the two vectors on the circle, \( q = \pm e_{s_\perp} \), indicated by the black dots.](image)

### 3.2.3 Estimation from the population imbalance

Population imbalance is a measurement, which is commonly used in experiments to determine the value of the interferometric phase. Here we show under which conditions this estimation strategy is optimal in the one-qubit case, when it corresponds
3.3 Symmetric Werner states

to the projection over the eigenstates $|\pm\rangle$ of $\hat{\sigma}_z$ ($\hat{\sigma}_z|\pm\rangle = \pm|\pm\rangle$):

$$\hat{E}_+ = |+\rangle\langle+| \quad \text{and} \quad \hat{E}_- = |-\rangle\langle-|.$$  (3.19)

The probability for detecting the qubit in $|\pm\rangle$ is

$$p_\pm = \text{Tr}[\hat{\rho}\hat{E}_\pm] = \frac{1}{2}(1 \pm s_z).$$  (3.20)

According to Eq. (3.10) and using the $\theta$-dependence of the vector $s$ from Eq. (3.4), we obtain that the FI for the estimation from the population imbalance is

$$F_{\text{imb}} = \frac{1}{p_+} \left( \frac{\partial p_+}{\partial \theta} \right)^2 + \frac{1}{p_-} \left( \frac{\partial p_-}{\partial \theta} \right)^2 = \frac{s_x^2}{1 - s_z^2},$$  (3.21)

This equation should be compared with Eq. (3.12). We obtain that $F_{\text{imb}} = F_Q$ when either (a) $s_x^2 + s_z^2 = 1$ or (b) $s_z = 0$. The condition (a) is satisfied with pure states only. Moreover, for any initial pure state $\hat{\rho}_0$ lying in the $x-z$ plane and rotated along the $y$ direction according to Eq. (3.3), the population imbalance is not only optimal but it also gives the maximum possible value $F_{\text{imb}} = 1$, independently from $\theta$. The condition (b) can be fulfilled both by mixed and pure states. However, $F_{\text{imb}} = s_x^2$ depends on $\theta$ and $F_{\text{imb}} = 1$ only for pure states and a specific rotation angle $\tan \theta = s_{0,z}/s_{0,x}$. When the states is mixed, $F_{\text{imb}} < 1$ and for every $\theta$ there is only one orientation of $\hat{\rho}_0$, which gives $s_z = 0$.

This once again shows how the optimal estimation strategies change abruptly when we switch between mixed and pure states. There is a continuum of pure states, which used for the estimation from the population imbalance saturate the QFI, but there is only one such mixed state for each $\theta$.

3.3 Symmetric Werner states

In this section, we extend the investigation of the optimal POVM to higher $N$. However, since general considerations are no longer possible, we restrict our analysis to a particular family of states. We begin with the two-qubit symmetric Werner state. Then, we extrapolate these results to higher number of qubits. Finally, basing on the structure of the Werner states, we construct a model for a a noisy NOON state, for which we determine the optimal measurements.

3.3.1 Two qubits: symmetric Werner States

A general symmetric density matrix of two qubits has eight independent real parameters and is too difficult to treat. We thus restrict our analysis to the “Werner” states [95], which are described with a single real coefficient $0 \leq \alpha \leq 1$:

$$\hat{\rho}_w = \frac{1 - \alpha}{3} \hat{1} + \alpha \hat{H}_{TF}.$$  (3.22)
Here $\hat{\Pi}_{TF}$ is the projection over the twin-Fock (TF) state $|1,1\rangle$, where each mode occupied by one particle. When $\alpha$ varies from 0 to 1, $\hat{\rho}_w$ changes from a complete mixture which is useless for parameter estimation, to a strongly entangled pure TF state, which is known to provide a sub shot-noise sensitivity (SSN) in the MZI [96].

The Werner state written in the mode occupation basis (\( |j,2-j\rangle \), with \( j = 0,1,2 \)) is diagonal

$$\hat{\rho}_w = \begin{pmatrix} \frac{1-\alpha}{3} & 0 & 0 \\ 0 & \frac{1+2\alpha}{3} & 0 \\ 0 & 0 & \frac{1-\alpha}{3} \end{pmatrix}. \quad (3.23)$$

This matrix is always invertible except from the case when the state is pure. Werner states (3.22) are a narrow subset of all possible two spin-\( \frac{1}{2} \) symmetric states, nevertheless – as we show below – they provide valuable insight into the optimal estimation strategies in quantum metrology. We take a generic linear interferometric transformation

$$\hat{U}(\theta) = e^{-i\theta \hat{\mathbf{n}} \cdot \hat{\mathbf{J}}}, \quad (3.24)$$

where $\hat{\mathbf{J}} = \frac{1}{2}\hat{\sigma}^{(1)} + \frac{1}{2}\hat{\sigma}^{(2)}$ is the “composite” pseudo-angular momentum operator, given by the sum of corresponding single-particle Pauli matrices (the upper index labels the particle).

First, we calculate the QFI, using the expression from Eq. (3.11), keeping in mind that the previously used generator of the interferometric transformation $\hat{\sigma}_y$ must be replaced by $\hat{\mathbf{n}} \cdot \hat{\mathbf{J}}$. The outcome is

$$F_Q(\alpha) = \frac{12\alpha^2}{2 + \alpha} \left( n_x^2 + n_y^2 \right) \quad (3.25)$$

Note that the $z$-component of the generator does not contribute to the QFI, because $\hat{\rho}_w$ is invariant upon rotation around the $z$ axis. The QFI depends only on the length of the projection of the vector $\mathbf{n}$ onto the $x$-$y$ plane. Thus, without any loss of generality, in the remaining of this Section we will restrict to the MZI transformation $\hat{U}_{\text{mzi}}(\theta) = e^{-i\theta \hat{J}_y}$. The output density matrix $\hat{\rho}_w(\theta) = \hat{U}_{\text{mzi}}(\theta)\hat{\rho}_w\hat{U}_{\text{mzi}}^\dagger(\theta)$ reads

$$\hat{\rho}_w(\theta) = \begin{pmatrix} \frac{2-\alpha(2-3\sin^2\theta)}{6} & \frac{\alpha \sin 2\theta}{2\sqrt{2}} & \frac{\alpha \sin^2 \theta}{2} \\ \frac{\alpha \sin 2\theta}{2\sqrt{2}} & \frac{1-\alpha(1-3\cos^2\theta)}{3} & \frac{\alpha \sin 2\theta}{2\sqrt{2}} \\ \frac{\alpha \sin^2 \theta}{2} & \frac{\alpha \sin 2\theta}{2\sqrt{2}} & \frac{2-\alpha(2-3\sin^2\theta)}{6} \end{pmatrix},$$

and the QFI reduces to

$$F_Q(\alpha) = \frac{12\alpha^2}{2 + \alpha}. \quad (3.26)$$

As anticipated, if $\alpha = 0$ the state (3.22) is a complete mixture useless for phase estimation and $F_Q(0) = 0$. The SSN sensitivity $F_Q(\alpha) > 2$ is reached when $\alpha > 2/3$. In the extreme case $\alpha = 1$, we obtain the Heisenberg scaling, i.e. $F_Q(1) = 4$. We compare this result with the concurrence [97], the entanglement measure for two qubits, which tells that the state (3.22) is entangled already when $\alpha > \frac{1}{4}$. This
3.3. Symmetric Werner states

Example confirms the known fact [29,98] that not all entangled states provide a SSN sensitivity.

Note, that the value of the QFI can be optimized by allowing independent rotations of both particles. In such case, the non-symmetric part of the Hilbert space must be included. We replace the transformation (3.24) with

\[ \hat{U}(\theta) = e^{-i\frac{\theta}{2}(n_1 \cdot \sigma_1 + n_2 \cdot \sigma_2)}. \] (3.27)

For the Werner state (3.22), the QFI can be calculated analytically for any \( \alpha \) and \( n_1, n_2 \). In a particular case when \( n_1 = e_z \) and \( n_2 = -e_z \) we obtain that

\[ F_Q(\alpha) = \frac{4}{3}(2\alpha + 1), \] (3.28)

which exceeds the SNL \( F_Q(\alpha) > 2 \), when \( \alpha > \frac{1}{4} \). Therefore, if the physically allowed operations distinguish the particles, all entangled two-qubit states of the form (3.22) become useful for SSN metrology.

3.3.1.1 Optimal measurements

In the following, we go back to the symmetric case and find the optimal measurements for which the FI equals the QFI (3.26). In analogy to the single-qubit case, we should now parametrize the POVM similarly as in Eq. (3.6) and find the parameters from the condition Eq. (3.13). However, this procedure gives equations which cannot be solved analytically. Therefore, we restrict to the optimal POVMs obtained from the diagonalization of the SLD, which using Eq. (3.14) is

\[ \hat{L}_{\hat{\rho}_w(\theta)} = -\frac{6i\alpha}{2 + \alpha} [\hat{J}_y, \hat{\rho}_w(\theta)]. \] (3.29)

We rewrite Eq. (3.29) in the matrix form

\[ \hat{L}_{\hat{\rho}_w(\theta)} = C_\alpha \begin{pmatrix} \frac{1}{\sqrt{2}} \sin 2\theta & -\cos 2\theta & -\frac{1}{\sqrt{2}} \sin 2\theta \\ -\cos 2\theta & \sqrt{2} \sin 2\theta & \cos 2\theta \\ -\frac{1}{\sqrt{2}} \sin 2\theta & \cos 2\theta & \frac{1}{\sqrt{2}} \sin 2\theta \end{pmatrix} \] (3.30)

where \( C_\alpha = \frac{6\alpha}{\sqrt{2}(2+\alpha)} \). A direct diagonalization provides the three (normalized) eigenstates

\[ |\Psi_1\rangle = \frac{(\cos \theta - \sin \theta)}{\sqrt{2}} |\psi_-\rangle - \frac{(\cos \theta + \sin \theta)}{\sqrt{2}} |1, 1\rangle \] (3.31a)
\[ |\Psi_2\rangle = \frac{(\cos \theta + \sin \theta)}{\sqrt{2}} |\psi_-\rangle + \frac{(\cos \theta - \sin \theta)}{\sqrt{2}} |1, 1\rangle \] (3.31b)
\[ |\Psi_3\rangle = |\psi_+\rangle, \] (3.31c)
where $|\psi_\pm\rangle = |2.0\rangle \pm |0.2\rangle / \sqrt{2}$. The set of projectors over the eigenstates of $\hat{L}_\rho (\theta)$ forms an optimal POVM. They depend on $\theta$ and have a complicated form but nevertheless do not depend on $\alpha$. This last fact is a consequence of the structure of the Werner state (3.22). In this case the SLD is proportional to the commutator of the projection operator with the generator of the interferometric transformation, see Eq. (3.29). Since $\alpha$ enters in Eq. (3.30) through the prefactor $C_\alpha$, only the eigenvalues of the SLD – contrary to the eigenvectors – depend on $\alpha$.

It is interesting to see explicitly that the FI calculated with the projectors on the states (3.31) saturates the QFI. We first calculate $P_i (\phi) = \langle \Psi_i | e^{-i\phi \hat{J}_y} \hat{\rho}_w e^{i\phi \hat{J}_y} | \Psi_i \rangle$ with $i = 1, 2, 3$:

$$P_1 (\phi) = \beta + \alpha \left[ \sin \left( \theta - \frac{\pi}{4} \right) \sin \phi - \sin \left( \theta + \frac{\pi}{4} \right) \cos \phi \right]^2,$$

$$P_2 (\phi) = \beta + \alpha \left[ \sin \left( \theta + \frac{\pi}{4} \right) \sin \phi + \sin \left( \theta - \frac{\pi}{4} \right) \cos \phi \right]^2,$$

$$P_3 (\phi) = \beta,$$

where $\beta = (1 - \alpha)/3$. For the specific phase value $\phi = \theta$, we obtain

$$F (\theta) = \sum_{i=1}^{3} \frac{1}{P_i (\phi = \theta)} \left( \frac{\partial P_i (\phi)}{\partial \phi} \right)_{\phi=\theta}^2 = \frac{12\alpha^2}{2 + \alpha}, \quad (3.32)$$

as expected. It should be noticed that for $\alpha < 1$, when the density matrix is invertible, the projection over the states (3.31) is the unique optimal POVM. For $\alpha = 1$, there are other POVMs, besides the projection over the states (3.31), for which $F = F_Q$. An example is shown in the following section.

### 3.3.1.2 Estimation from the population imbalance

To this end we again focus on estimation from the measurement of the population imbalance between the two modes. The FI is

$$F_{\text{imb}} = \sum_{j=0}^{2} \frac{1}{p(j|\theta)} \left( \frac{\partial p(j|\theta)}{\partial \theta} \right)_j^2, \quad (3.33)$$

Note that at $\theta = 0$, there is a simple way to reduce the optimal POVM to the simple measurement of population imbalance. Indeed, the following transformation

$$\hat{V} = \exp \left( i \frac{\pi}{2} \hat{J}_x \hat{J}_y + \frac{\hat{J}_y^2}{2} \right) \exp \left( i \frac{\pi}{4} \hat{J}_y \right).$$

applied to the states (3.31) gives $\hat{V}|\Psi_1\rangle = |0, 2\rangle$, $\hat{V}|\Psi_2\rangle = |1, 1\rangle$ and $\hat{V}|\Psi_3\rangle = |2, 0\rangle$. In this way, we obtain the eigenstates of the $\hat{J}_x$ operator, and the optimal measurement is based on the determination of the population imbalance. Nevertheless, to accomplish this we needed an additional operation $\hat{V}$ acting on the state. This transformation is non-local – it correlates the particles, since the product of two angular-momentum operators cannot be written as a sum of operators acting on each qubit independently.
3.3. Symmetric Werner states

where \( p(j|\theta) \) is a probability for finding \( j \) particles in one of the modes and \( 2 - j \) in the other and is given by

\[
p(j|\theta) = \langle j, 2-j| \hat{\rho}_w(\theta) |j, 2-j\rangle. \tag{3.34}
\]

A direct calculation shows that \( p(0|\theta) = p(2|\theta) = \frac{1-\alpha}{3} + \frac{\alpha}{2} \sin^2 \theta \) and \( p(1|\theta) = \frac{1-\alpha}{3} + \alpha \cos^2 \theta \). These, put into Eq. (3.33), give

\[
F_{\text{imb}} = \frac{36 \alpha^2 \sin^2(2\theta)}{[4 - \alpha(1 + 3 \cos(2\theta))][2 + \alpha(1 + 3 \cos(2\theta))]}.
\tag{3.35}
\]

In Fig. 3.2 we plot Eq. (3.35) as a function of \( \theta \) and for different values of \( \alpha \). When \( \alpha = 1 \) the Werner state is pure and we find \( F_{\text{imb}} = 4 \), independently on \( \theta \). In this case \( F_{\text{imb}} \) saturates the QFI and reaches the maximum value attainable for two qubits. For \( \alpha < 1 \), the estimation from the population imbalance is generally non-optimal and depends strongly on \( \theta \).

This is another example, after the single qubit case, of how the estimation strategy, which is optimal for a pure state, immediately deteriorates as soon as the state becomes mixed. It also shows that, for pure states there are measurements that allow to saturate the QFI for any value of \( \theta \). This highly desirable property is lost as soon as the state becomes mixed.

### 3.3.2 Generalization to higher \( N \)

Note that the above considerations can be generalized to higher \( N \) as follows. First, basing on Eq. (3.22), we introduce a symmetric state of \( N \) qubits

\[
\hat{\rho}_w = \frac{1-\alpha}{N+1} \mathbb{I} + \alpha \hat{\Pi}_N.
\tag{3.36}
\]
Here, the unity operator acts in the $N + 1$ dimensional symmetrized space of $N$ qubits. Quite generally, $\Pi_N$ projects onto any state within this space. The density matrix (3.36) undergoes a linear interferometric transformation governed by the evolution operator $\hat{U}(\theta) = e^{-i\theta \hat{h}}$,

$$\hat{\rho}(\theta) = \hat{U}(\theta) \hat{\rho} \hat{U}^\dagger(\theta).$$

(3.37)

Owing to the simple structure of the state (3.36), the QFI is easily evaluated and the outcome is

$$F_Q = \frac{\alpha^2(N + 1)}{2 + \alpha(N - 1)} 4(\Delta \hat{h})^2.$$  

(3.38)

Here the variance of the generator is calculated using the state onto which $\hat{\Pi}_N$ projects. We now consider two distinct cases, one when $\hat{\Pi}_N$ projects onto the twin-Fock state, and the other when $\hat{\Pi}_N$ projects onto the NOON state. For these two examples, we determine the set of optimal measurements using the SLD, analogically to the approach presented in the previous section.

The state (3.36) has a similar structure to (3.22) – it is a mixture of identity and the projection operator. Therefore, the derivation of the SLD is also analogical, and we obtain that

$$\hat{L} = \beta \left[ \hat{h}, \hat{\Pi}_N(\theta) \right],$$

(3.39)

with $\beta = \frac{2\alpha(N+1)}{2+\alpha(N-1)}$ and $\hat{\Pi}_N(\theta) = \hat{U}(\theta) \hat{\Pi}_N \hat{U}^\dagger(\theta)$. Since the evolution is generated by the $\hat{h}$ operator, we write the SLD as

$$\hat{L} = \hat{U}(\theta) \left( \beta \left[ \hat{h}, \hat{\Pi}_N \right] \right) \hat{U}^\dagger(\theta).$$

(3.40)

The optimal measurements are found in two steps. First, we evaluate the eigenstates of the matrix, which is the outcome of the above commutator $[\hat{h}, \hat{\Pi}_N]$. Then, we rotate the resulting states using the evolution operator. Clearly, this procedure yields the optimal projectors, which do not depend on $\alpha$, just as was the case of the two-particle Werner state.

The optimal POVM is constructed from those rotated eigenstates and projections onto all the remaining orthogonal states from the $(N + 1)$-dimensional space. Nevertheless, these remaining states do not contribute to the Fisher information, since their trace with the density matrix $\hat{\rho}(\theta)$ vanishes. Therefore, to find the maximal Fisher information, it is sufficient to determine only the projection operators onto the eigenstates of the SLD.

### 3.3.2.1 Twin-Fock state

First, we apply this procedure to find the optimal measurements for the “noisy” twin-Fock state. In this case, the operator $\hat{\Pi}_N$ projects onto the state $\left| \frac{N}{2}, \frac{N}{2} \right\rangle$. If the generator is equal to $\hat{h} = \hat{J}_y$, the commutator $[\hat{h}, \hat{\Pi}_N]$ in Eq. (3.40) is proportional
3.3. Symmetric Werner states

to

$$\hat{C} = \begin{pmatrix} 0 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix},$$

(3.41)

expressed in the basis \( |\frac{N}{2} - 1, \frac{N}{2} + 1\rangle, |\frac{N}{2}, \frac{N}{2}\rangle \) and \( |\frac{N}{2} + 1, \frac{N}{2} - 1\rangle \). We immediately obtain the three eigenvectors, which are \( \vec{v}_1 = \left( \frac{1}{\sqrt{2}}, \frac{i}{\sqrt{2}}, -\frac{1}{\sqrt{2}} \right) \), \( \vec{v}_2 = \left( \frac{1}{2}, -\frac{1}{\sqrt{2}}, \frac{i}{2} \right) \) and \( \vec{v}_3 = \left( \frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right) \). The optimal measurements are the projections onto \( \vec{v}_1, \vec{v}_2 \) and \( \vec{v}_3 \) and the remaining \( N - 2 \) orthogonal states, rotated by the operator \( e^{-i\hat{J}_y} \).

3.3.3 Noisy NOON state

We now exploit the algebraic properties of the Werner states to construct a simple model of a symmetric NOON state of \( N \) particles subject to decoherence. A pure NOON state is \( |\psi\rangle_{\text{NOON}} = \frac{1}{\sqrt{2}}(|N0\rangle + |0N\rangle) \). Upon the influence of the environment, it might deteriorate in a complicated way, depending on the type of decoherence. Here, we will assume that the interaction with the environment transforms it into the mixture (3.36), where the projection is onto \( |\psi\rangle_{\text{NOON}} \). Choosing the generator to be the pure phase-imprint, i.e. \( \hat{h} = \hat{J}_z \), we obtain that the commutator \([\hat{h}, \hat{\Pi}_N]\) is proportional to

$$\hat{C} = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix},$$

(3.42)

expressed in the basis \( |N0\rangle \) and \( |0N\rangle \). Optimal measurements are projections onto \( |\psi_\pm\rangle = \frac{1}{\sqrt{2}} (1, \pm i) \) plus all the other states which do not contribute to the Fisher information, rotated by the operator \( e^{-i\theta \hat{J}_z} \).

Figure 3.3: Schematic representation of the two steps performed to calculate the QFI. The input state \( |\psi_0\rangle \), which enters the interferometer, is transformed by a unitary evolution operator \( \hat{U}(\theta) \), giving the output state \( |\psi\rangle \). To calculate the speed at which the state changes, and thus the statistical distance, we make a further infinitesimal rotation \( \hat{U}(d\theta) \) to obtain \( |\tilde{\psi}\rangle \).
3.4 Optimal measurements for symmetric \( N \)-qubit pure states

We now show that for generic pure symmetric states of \( N \) qubits, some optimal estimation strategies can be found using the notion of the statistical distance.

3.4.1 QFI and the statistical distance

In [33] it is shown how the QFI from Eq. (3.11) is related to the statistical distance between two neighbouring states [30]. To be more precise, the QFI emerges as the coefficient of the leading (second order) term in the Taylor expansion of the Bures distance between the probe state and the phase-shifted one. This result is valid for pure and mixed states and for any parameter-dependent transformation. In the following, we use the concept of the statistical distance to provide a general condition under which the FI obtained from the population imbalance measurement saturates the QFI. To this end, we follow the scheme pictured in Fig. 3.3. First, we consider a pure two-mode input state \( |\psi_0\rangle \), which is transformed by a unitary evolution operator \( \hat{U}(\theta) = e^{-i\theta \hat{J}_n} \), where \( n \) is an arbitrary direction in the three-dimensional space. As a result, we obtain a \( \theta \)-dependent output state, which expanded in some orthogonal basis \( \{ |j\rangle \} \) reads

\[
|\psi\rangle = e^{-i\theta \hat{J}_n} |\psi_0\rangle = \sum_{j=0}^{N} \sqrt{p_j} e^{i\varphi_j} |j\rangle = \sum_{j=0}^{N} C_j |j\rangle ,
\]

(3.43)

with \( C_j \equiv \langle j | \psi \rangle \). The neighbouring state is found by applying a further infinitesimal transformation

\[
|\tilde{\psi}\rangle = e^{-i\theta \hat{J}_n} |\psi\rangle
\]

(3.44)

and is approximated as

\[
|\tilde{\psi}\rangle \approx \sum_{j=0}^{N} (1 - i\theta \eta_j^{(n)} \hat{J}_n) C_j |j\rangle = \sum_{j=0}^{N} (1 - i\theta \eta_j^{(n)}) C_j |j\rangle ,
\]

\[
= \sum_{j=0}^{N} \tilde{C}_j |j\rangle .
\]

where

\[
\eta_j^{(n)} C_j = \langle j | \hat{J}_n | \psi \rangle .
\]

(3.45)

Notice that this equation unequivocally defines \( \eta_j^{(n)} \) only if \( C_j \neq 0 \). The state (3.44) can be alternatively written as

\[
|\tilde{\psi}\rangle = \sum_{j=0}^{N} \sqrt{p_j + dp_j} e^{i(\varphi_j + d\varphi_j)} |j\rangle ,
\]

(3.46)
3.4. Optimal measurements for symmetric $N$-qubit pure states

where the probability- and phase-increments are

$$dp_j = |\tilde{C}_j|^2 - |C_j|^2 = 2 \text{Im} \eta_j^{(n)} |C_j|^2 d\theta$$  \hspace{1cm} (3.47a)

$$e^{id\varphi_j} = \frac{\tilde{C}_j}{|C_j|} C_j = e^{-i \text{Re} \eta_j^{(n)} d\theta}.$$  \hspace{1cm} (3.47b)

The distance between two neighbouring states is equal to

$$ds_{\text{ps}}^2 = 1 - |\langle \psi | \tilde{\psi} \rangle|^2$$ \hspace{1cm} (3.48a)

$$= \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4 \left[ \sum_{j=0}^{N} p_j d\varphi_j^2 - \left( \sum_{j=0}^{N} p_j d\varphi_j \right)^2 \right]$$ \hspace{1cm} (3.48b)

$$\equiv \sum_{j=0}^{N} \frac{dp_j^2}{p_j} + 4 \Delta^2 d\varphi.$$ \hspace{1cm} (3.48c)

Finally, the QFI can be interpreted as the speed at which the state changes upon the infinitesimal increment of the parameter $\theta$. By using Eqs. (3.48c) with $dp_j$ and $d\varphi_j$ obtained from Eqs. (3.47a) and (3.47b), respectively, we find

$$F_Q = \frac{ds_{\text{ps}}^2}{d\theta^2} = 4 \sum_{j=0}^{N} |C_j|^2 \left( \text{Im} \eta_j^{(n)} \right)^2 +$$ \hspace{1cm} (3.49a)

$$4 \sum_{j=0}^{N} |C_j|^2 \left( \text{Re} \eta_j^{(n)} \right)^2 - 4 \left( \sum_{j=0}^{N} |C_j|^2 \text{Re} \eta_j^{(n)} \right)^2.$$ \hspace{1cm} (3.49b)

The above result is equivalent to the well known expression for the QFI for pure states

$$F_Q = 4(\Delta \hat{J}_n)^2.$$ \hspace{1cm} (3.50)

However, as will become evident below, for the purpose of finding the optimal measurements, it is more convenient to keep the QFI in the form of Eq. (3.49), i.e. as a sum of two non-negative parts – the change of the probability $p_j$ and the variance of the phase increment $d\varphi_j$. We thus arrive at the main result of this section: the FI calculated from projective measurements on the basis $\{|j\rangle\}$ saturates the QFI if and only line (3.49b) is equal to zero. In this case, no information about $\theta$ is carried by the phases $\varphi_j$, which are not witnessed by the projection measurement $|j\rangle$ ($j$) and thus do not contribute to probabilities $p_j$. The line (3.49b) vanishes if and only if

$$\text{Re} \eta_j^{(n)} = \langle \psi | \hat{J}_n | \psi \rangle, \hspace{1cm} \forall j \text{ such that } C_j \neq 0.$$ \hspace{1cm} (3.51)

To demonstrate this result, we use the Cauchy-Schwarz inequality

$$\left( \sum_{j=0}^{N} |C_j|^2 \text{Re} \eta_j^{(n)} \right)^2 \leq \sum_{j=0}^{N} |C_j|^2 (\text{Re} \eta_j^{(n)})^2.$$ \hspace{1cm} (3.52)
When this inequality is saturated, the line (3.49b) vanishes. This happens if and only if \( \text{Re} \eta_j^{(n)} = c \), where \( c \) is a real number which does not depend on \( j \). To find \( c \), we multiply both sides of Eq. (3.45) by \( C_j^* \), sum over \( j \) and obtain \( c = \langle \psi | \hat{J}_n | \psi \rangle \).

In particular, if one of the states in the basis \( \{|j\rangle\} \) is equal to \( |\psi\rangle \) (and thus \( C_j = 1 \) for \( |j\rangle = |\psi\rangle \) and \( C_j = 0 \) otherwise), then Eq. (3.51) is always satisfied. In this case, we recover a well known result [34]: for pure states, any POVM consisting of projectors over the probe state and over any basis on the orthogonal subspace allows to saturate the QFI.

### 3.4.2 “In-situ” measurements – localized modes

In this section, we focus on the measurement of the number of particles in the two output modes. This corresponds to take the basis \( \{|j\rangle\} \) given by the mode occupation basis, where \( |j\rangle \) denotes a Fock state with \( j \) particles in the left and \( N - j \) in the right arm. In this case, the coefficient \( \eta_j^{(n)} \) is a result of acting with \( \hat{J}_n \) on a ket \( |j\rangle \) and is equal to

\[
\eta_j^{(x)} = \frac{1}{2} \frac{\alpha_j C_{j+1} + \alpha_{j-1} C_{j-1}}{C_j} \tag{3.53a}
\]

\[
\eta_j^{(y)} = \frac{1}{2i} \frac{\alpha_j C_{j+1} - \alpha_{j-1} C_{j-1}}{C_j} \tag{3.53b}
\]

\[
\eta_j^{(z)} = j - \frac{N}{2}, \tag{3.53c}
\]

where \( \alpha_j = \sqrt{(j+1)(N-j)} \). By referring to Eq. (3.49), we identify two optimal measurements performed “in-situ”, when the particles remain trapped in the two arms of the interferometer and their spatial mode functions do not overlap.

#### 3.4.2.1 Estimation from the full correlation

As a first example, we consider the phase estimation from the full \( N \)-body probability

\[
p_N(r|\theta) = \frac{1}{N!} \langle \psi | \hat{\Psi}^\dagger(x_1) \ldots \hat{\Psi}^\dagger(x_N) \hat{\Psi}(x_N) \ldots \hat{\Psi}(x_1) | \psi \rangle \equiv \langle \psi | \hat{G}(r) | \psi \rangle. \tag{3.54}
\]

of finding particles at positions \( r = (x_1 \ldots x_N) \). The two-mode field operator is \( \hat{\Psi}(x) = \psi_a(x)\hat{a} + \psi_b(x)\hat{b} \) and the wave-packets are separated in two arms of the interferometer, for instance by imposing \( \psi_a(x) = 0 \) for \( x < 0 \) and \( \psi_b(x) = 0 \) for \( x > 0 \). The \( \theta \)-dependence of the probability \( p_N(r|\theta) \) comes from the state \( |\psi\rangle \) from Eq. (3.43), which is used to calculate the average value of the operator \( \hat{G}(r) \).

The estimation sequence relies upon detecting positions of \( N \) atoms in \( m \gg 1 \) experiments. If the phase is obtained from the maximum likelihood estimator, according to the Fisher theorem, its sensitivity is given by

\[
\Delta^2 \theta = \frac{1}{m} \frac{1}{F_N}, \tag{3.55}
\]
where $F_N$ is the FI equal to

$$F_N = \int d\mathbf{r} \frac{1}{p_N(\mathbf{r}|\theta)} \left( \frac{\partial p_N(\mathbf{r}|\theta)}{\partial \theta} \right)^2.$$  \hfill (3.56)

In order to calculate $F_N$ we first evaluate the derivative of the probability (3.54),

$$\partial_\theta p_N(\mathbf{r}|\theta) = i \langle \psi | \hat{J}_n \hat{G}(\mathbf{r}) | \psi \rangle - i \langle \psi | \hat{G}(\mathbf{r}) \hat{J}_n | \psi \rangle =$$

$$2 \text{Im} \langle \psi | \hat{G}(\mathbf{r}) \hat{J}_k | \psi \rangle = 2 \text{Im} \sum_{j,j'=0}^{N} C_j^* C_{j'} \eta_j^{(n)} \langle j | \hat{G}(\mathbf{r}) | j' \rangle.$$  \hfill (3.57)

The FI is therefore equal to

$$F_N = 4 \int d\mathbf{r} \frac{\left[ \text{Im} \sum_{j,j'=0}^{N} C_j^* C_{j'} \eta_j^{(n)} \langle j | \hat{G}(\mathbf{r}) | j' \rangle \right]^2}{\text{Im} \sum_{j,j'=0}^{N} C_j^* C_{j'} \langle j | \hat{G}(\mathbf{r}) | j' \rangle}.$$  \hfill (3.58)

We now define $\Omega_\mu$ by saying that $\mathbf{r} \in \Omega_\mu$ when $x_1 \ldots x_\mu < 0$ and $x_{\mu+1} \ldots x_N > 0$. Using this definition we obtain

$$F_N = 4 \sum_{\mu=0}^{N} \binom{N}{\mu} \int_{\Omega_\mu} d\mathbf{r} \frac{\left[ \text{Im} \sum_{j,j'=0}^{N} C_j^* C_{j'} \eta_j^{(n)} \langle j | \hat{G}(\mathbf{r}) | j' \rangle \right]^2}{\text{Im} \sum_{j,j'=0}^{N} C_j^* C_{j'} \langle j | \hat{G}(\mathbf{r}) | j' \rangle}$$

where the combinatory factor is due to the symmetry of the state and stands for all possible choices of $\mu$ particles out of a set of $N$. When $\mathbf{r} \in \Omega_\mu$, then for separated wave-packets $\hat{G}(\mathbf{r}) | n \rangle \propto \frac{\nu^{(N-\mu)!}}{N!} | n \rangle \delta_{n\mu}$ and the above integral gives

$$F_N = 4 \sum_{j=0}^{N} |C_j|^2 \left( \text{Im} \eta_j^{(n)} \right)^2.$$  \hfill (3.59)

We notice that this expression is equal to the first line of the QFI, see (3.49a).

Therefore, estimation from the $N$-body probability of trapped particles is optimal only if the other terms in line (3.49b) vanish, which requires $\text{Re} \eta_j^{(n)} = 0$ for all $j$. According to Eq. (3.53), this condition is satisfied only for the rotations around $x$ and $y$ axes. For the rotations around $x$-axis, $\text{Re} \eta_j^{(x)} = 0$ if $C_j = i^\phi a_j$, while for rotations around the $y$-axis the (necessary and sufficient) condition is $C_j = e^{i\phi} a_j$, where $a_j \in \mathbb{R}$ and $\phi$ is a common phase. In particular for $\phi = 0$, the measurement is optimal, when all $C_j$’s are real. Since the elements of the Wigner rotation matrix – which transforms the input state $|\psi\rangle$ into the output state $|\psi'\rangle$ – are all real, we conclude that if the input state of the MZI has real coefficients, the estimation from $p_N$ is optimal. For rotations around the $z$ axis, $\text{Im} \eta_j^{(z)} = 0$ and thus $F_N = 0$, because the simple phase imprint $e^{-i\phi J_z}$ requires further mode mixing to provide information about $\theta$.

---

3 L.C. Biedenharn and J.D. Louck, *Angular Momentum in Quantum Physics: Theory and Appli-
3.4.2.2 Estimation from the population imbalance

Although phase estimation from the $N$-body probability is optimal for a wide class of states and rotations around $x$ and $y$, it has one major flaw – it is impractical, since it requires sampling of a vast configurational space. We now show, that the same value of the FI as in Eq. (3.59) is obtained, when the phase is estimated from a simple population imbalance measurement. The probability of having $j$ atoms in the mode $a$ and $N-j$ in $b$ is

$$p(j|\theta) = |\langle j|\psi\rangle|^2 = |C_j|^2.$$  \hfill (3.60)

Similarly as in Eq. (3.57), its derivative reads

$$\partial_\theta p(j|\theta) = 2|C_j|^2 \text{Im} \eta_j^{(n)}. $$ \hfill (3.61)

Therefore, the FI calculated with (3.60)

$$F_{\text{imb}} = \sum_{j=0}^{N} \frac{1}{p(j|\theta)} \left( \frac{\partial p(j|\theta)}{\partial \theta} \right)^2 = 4 \sum_{j=0}^{N} |C_j|^2 \left( \text{Im} \eta_j^{(n)} \right)^2 $$ \hfill (3.62)

is equal to (3.59). In consequence, the QFI from Eq. (3.49) is saturated with the same family of states for the $x$ and $y$ rotations as in the case of the estimation from $p_N(r|\theta)$. Note that this result has been obtained independently in Ref. [99] (see also considerations in Ref. [100] for a specific class of probe mixed states). Furthermore, if the input state of the Mach-Zehnder interferometer (i.e. a rotation around the $y$ axis) has real coefficients, the estimation from the population imbalance is optimal for all values of $\theta$. If the coefficients $C_j^{(0)}$ are imaginary, it is still possible to saturate the QFI, at specific phase values, provided that the condition $\text{Re} \eta_j^{(y)} = 0$ for all $j$ is satisfied. Note that less general conditions for the saturation of the QFI, independently if the phase shift, and with population imbalance measurement, has been given in Ref. [101] (see also [102]).

3.4.3 Measurement after expansion

As argued above, when the interferometer rotates the state around the $z$-axis, giving a sole phase-imprint, further manipulation is necessary to exchange the information about the phase between the two modes. Here we assume, that this operation is realized by letting the two mode functions $\psi_a(x)$ and $\psi_b(x)$ expand and form an interference pattern. In such situation, the two modes cannot be distinguished...
anymore, and it is not possible to define a proper population imbalance operator. Instead, one must estimate $\theta$ in some different way. For instance, estimation from the least-squares fit of the one-body probability $p_1(x|\theta) = \frac{1}{N} \langle \psi | \hat{\Psi}^\dagger(x) \hat{\Psi}(x) | \psi \rangle$ to the interference pattern, although gives sub shot-noise sensitivity when the input state $|\psi_0\rangle$ is phase-squeezed [103], is never optimal [104].

Nevertheless, the optimal measurement can be identified and it is the $N$-body FI from Eq. (3.56) which saturates the bound of the QFI under following additional assumptions [105]. First, the information between the two modes must be fully exchanged. This means, that the envelopes of $\psi_a(x)$ and $\psi_b(x)$ fully overlap and the functions only differ by the phase. This is true if initially $\psi_a(x)$ and $\psi_b(x)$ are of the same shape but are separated in space and then expand to reach the far-field regime. Another requirement is that the coefficients of the initial state $C_j^{(0)}$ are real and possess the symmetry $C_j^{(0)} = C_{N-j}^{(0)}$. States having these properties naturally appear in the context of quantum interferometry with ultra-cold gas trapped in the double-well potential. Namely, the ground state of the symmetric two-mode Bose-Hubbard Hamiltonian for every ratio of the interaction strength $U$ to the tunnelling rate $J$ has real and symmetric coefficients $C_j^{(0)}$.

According to Eq. (3.43), the rotation around the $z$ axis transforms the state into
\begin{equation}
|\psi\rangle = e^{-i\theta \hat{J}_z} |\psi_0\rangle = \sum_{j=0}^{N} C_j |j\rangle,
\end{equation}
where $C_j = C_j^{(0)} e^{-i\theta (j - \frac{N}{2})}$. As argued in detail in [105], the FI from Eq. (3.56), calculated under the aforementioned assumptions, is
\begin{equation}
F = 4 \sum_j |C_j|^{2} \eta_z^{(j)} = 4(\Delta \hat{J}_z)^2 = F_Q,
\end{equation}
where $\eta_z^{(j)}$ was defined in Eq. (3.53c). This shows that the estimation from the $N$th body correlation in the far field is optimal.

### 3.5 Conclusions

In this chapter, we have identified optimal measurements in various two-mode interferometric systems. We have first considered the qubit, which is the simplest – but nevertheless illustrative – testbed for the theory. There is a clear picture emerging from the analysis of the single qubit: while for pure probe states there is a continuous set of optimal estimation strategies (i.e. optimal POVMs for which the classical Fisher information is equal to the quantum Fisher information), they reduce to a specific POVM when the probe state is mixed. As a direct consequence, while for pure states it is possible to saturate the quantum Fisher information independently of the phase $\theta$, this is not possible for mixed states. Indeed, in this case, the optimal POVM strongly depends on the parameter $\theta$. This overall picture has been confirmed in the two-qubit case from the analysis of the Werner states. In
addition, in the case of a single pure qubit state and two-qubit pure Werner states, we have shown that the measurement of the population imbalance among the two output modes of the interferometer is optimal and saturates the quantum Fisher information independently of $\theta$. We have further extended the analysis to the general case of pure $N$-qubit states. We have found the conditions on the state and interferometric transformation for which the classical Fisher information obtained from the population imbalance measurement or $N$th order correlation function can saturate the quantum Fisher information, independently of the phase shift.

The discussion on the strategies for precise estimation in quantum interferometry shows that finding an optimal measurement is in general a difficult task. In the following chapter, we present a very simple particle entanglement criterion, which is based on the violation of the Cauchy-Schwarz inequality.
In this chapter\footnote{The chapter is based on the work: T. Wasak, P. Szańkowski, P. Ziń, M. Trippenbach, J. Chwedeńczuk, Phys. Rev. A 90, 033616 (2014).}, we address the question raised in the discussion of the experiment with colliding Bose-Einstein condensates (see page 14): Are the particles that are emitted onto the collisional halo entangled? In this experiment\footnote{[26]}, the violation of the Cauchy-Schwarz inequality was observed and interpreted as a signature of non-classicality. However, this criterion is directly taken from quantum optics, and it is not clear in what sense it detects quantum correlations in a system of massive particles. Here, we show that the observation of this effect proves the presence of particle entanglement in a many-body system of bosons. Our derivation is based on the analogy between the coherent states of electromagnetic fields and separable states of massive particles. The presented argument applies to any quantum system of bosons with either a fixed or a fluctuating number of particles, provided that there is no coherence between different number states.

\section{Introduction}

The challenging question whether, and in what sense, the pulse of light is not classical, is one of central issues of quantum optics. This problem was addressed by Glauber and Sudarshan in their studies on coherence in the context of correlation functions\footnote{[106, 107]}. They employed the coherent states $|\Phi\rangle$ defined by the relation $\hat{E}(\Phi)(x)|\Phi\rangle = \Phi(x)|\Phi\rangle$, where $\hat{E}(\Phi)(x)$ is the positive-frequency part of the
electromagnetic field $\hat{E}(x)$, and expressed the density matrix using the so-called P-representation:

$$\hat{\rho} = \int \mathcal{D}\Phi |\Phi\rangle\langle\Phi| P(\Phi).$$

(4.1)

The symbol $\mathcal{D}\Phi$ denotes the integration measure over the set of complex fields $\Phi$. The state of light is considered classical if an outcome of any measurement of the normally ordered correlation function can be explained in terms of classical electromagnetic fields. It happens when the P-representation can be interpreted as a probability distribution, i.e. when $P(\Phi)$ is normalized and for all positive-defined functions $F(\Phi) > 0$ a relation

$$\int V \mathcal{D}\Phi P(\Phi) F(\Phi) > 0$$

(4.2)

holds for any volume $V$. Otherwise, the field is said to be quantum.

Once the electromagnetic field is quantized, photons can be treated on a more equal footing with other particles. It is then reasonable to ask the question about correlations between individual particles and in this context the concept of particle entanglement emerges [97, 108]. The quantum provenance of entanglement is underlined by the word “paradox” often used to describe some highly counter-intuitive phenomena such as the Einstein-Podolsky-Rosen paradox [4] and the related Schrödinger’s cat problem. Apart from fundamental aspects, systems of entangled particles have applications in quantum information [7], teleportation [109] or ultra-precise metrology [29, 110]. Entanglement is also related to the concept of non-locality [111] and recently a versatile witness of the multi-partite entanglement has been introduced [112]. Finally, the entanglement might also contribute to the extreme efficiency of the energy transfer in the process of photosynthesis [113].

Much as a fascinating consequence of quantum mechanics, entanglement is also elusive. It is not simple to entangle particles on demand, because this requires complicated experimental strategies and it is difficult to protect them from the destructive influence of the environment, which inevitably leads to decoherence [114–116]. Finally, even if a carefully prepared state reaches detectors fairly intact, it is often not clear which quantity should be measured to witness entanglement.

This last difficulty is related to the very definition, which states that an entangled state of $N$ particles is such that is not separable, meaning that it cannot be written as a mixture of product states [69, 95, 117]

$$\hat{\rho}_{\text{sep}} = \sum_i p_i \hat{\rho}_i^{(1)} \otimes \ldots \otimes \hat{\rho}_i^{(N)}.$$  

(4.3)

Here $p_i$’s are non-negative weights that add up to unity. The consequence of this indirect definition is that to detect entanglement we usually refer to some bounds achievable by separable states. A good example is the two-mode quantum interferometer, which utilizes a collection of $N$ qubits in state $\rho$ to determine an unknown phase $\theta$. If the precision of the parameter estimation $\Delta\theta$ is better than shot-noise
\[ \Delta \theta = N^{-1/2} \] (the smallest error attainable with separable states), then \( \hat{\rho} \) is entangled \([29]\). However, the argument cannot be reversed.

Below in this chapter, we show that the recently observed violation of the CSI by the second order correlation function \([26]\) is a consequence of the particle entanglement in a system of many bosons. We base our derivation, presented in Section 4.2, on the analogy of the classical states of electromagnetic field, such as in Eq. (4.1) with separable states of \( N \) bosons. In Section 4.2.1 we show what is the relation of the CSI to the commonly measured squeezing of the relative number counts between two regions. In Section 4.2.2 we show that the CSI applies also to systems of bosons with fluctuating number of particles, as long as coherences between different number states are absent. Such states, according to the superselection rule, are the only allowed for massive particles \([118]\). In Section 4.3 we compare the CSI criterion for particle entanglement with the approach of the Fisher information known from quantum interferometry. Finally, in Section 4.4 we show that the CSI does not apply to systems of distinguishable particles as a criterion for particle entanglement. We conclude in Section 4.5.

### 4.2 Formulation of the problem

The discussion begins by stating that a separable pure state of \( N \) identical bosons \( |\phi; N\rangle \) must be a product of \( N \) identical single-particle orbitals \( |\phi\rangle \), i.e.

\[ |\phi; N\rangle = |\phi\rangle^\otimes N. \] (4.4)

When the bosonic field operator \( \hat{\Psi}(x) \) acts on the state (4.4), the result is

\[ \hat{\Psi}(x) |\phi; N\rangle = \sqrt{N} \phi(x) |\phi; N - 1\rangle, \] (4.5)

which is a fixed-\( N \) counterpart of the property of a coherent state of light \( |\Phi\rangle \). In Eq. (4.5), \( \phi(x) \) is a single-particle function determining the spatial properties of the system.

The general separable state of \( N \) identical bosons is a mixture of states given in Eq. (4.4) \([119–122]\), and takes the following form\(^2\):

\[ \hat{\rho} = \int \mathcal{D}\phi \ |\phi; N\rangle\langle\phi; N| \mathcal{P}(\phi). \] (4.6)

Here \( \mathcal{D}\phi \) denotes the integration over the complex field \( \phi \) and \( \mathcal{P}(\phi) \) is the probability distribution, which means it is normalized and its integral with every \( F(\phi) \geq 0 \) over any volume \( \mathcal{V} \) is non-negative, i.e.

\[ \int_{\mathcal{V}} \mathcal{D}\phi \mathcal{P}(\phi) F(\phi) \geq 0. \] (4.7)

---

\(^2\) The proof, that the separable state takes this form, is also included in appendix A. There, we additionally provide a simple example for two qubit system.
There is a direct analogy between the P-representation from Eq. (4.1) and $P(\phi)$ from Eq. (4.6). Recall, that if the former does not satisfy condition (4.2), the electromagnetic field is not classical. Analogically, for $N$ indistinguishable bosons, if condition (4.7) is not fulfilled, the density matrix cannot be written as a statistical mixture (4.6), meaning that correlations between the particles are genuinely quantum.

In the domain of quantum optics, numerous criteria were introduced to verify if the state’s P-representation satisfies condition (4.2). Among them we find the Cauchy-Schwarz inequality (CSI) for the second order correlation function (for experimental realization see [123]) and the related number squeezing between two regions [124–126]. Recently, the CSI has been used to characterize various phenomena such as the analogue of the Hawking radiation [127] or phonon excitations in planar microcavities [128]. The CSI can also indicate if the P-representation is partially negative, which signifies entanglement in a two-party system [129]. Here we show to which extent the violation of the CSI applies to systems of $N$ identical bosons as a probe of particle entanglement.

For separable states (4.6), the second-order correlation function, calculated using Eq. (4.5) is

$$G^{(2)}(x, x') = \left\langle \hat{\Psi}^\dagger(x) \hat{\Psi}^\dagger(x') \hat{\Psi}(x') \hat{\Psi}(x) \right\rangle = N(N-1) \int \mathcal{D}\phi P(\phi)|\phi(x)|^2|\phi(x')|^2. \quad (4.8)$$

For the purpose of the following, considerations we introduce two regions $a$ and $b$ having volumes $V_a$ and $V_b$ and an integrated second-order correlation function (4.8):

$$G_{ij}^{(2)} = \int_{V_i} dx \int_{V_j} dx' G^{(2)}(x, x') = N(N - 1) \int \mathcal{D}\phi P(\phi) I_i(\phi) I_j(\phi), \quad (4.9)$$

where $I_i(\phi) = \int_{V_i} dx |\phi(x)|^2$ with $i$ being either $a$ or $b$. We now apply the CSI for
4.2. Formulation of the problem

non-negative functions $f_i(\phi) = \sqrt{P(\phi)} I_i$, 

$$\int D\phi \, f_a(\phi) \, f_b(\phi) \leq \sqrt{\int D\phi f^2_a(\phi)} \sqrt{\int D\phi f^2_b(\phi)}, \quad (4.10)$$

which gives 

$$C \equiv \frac{G_{ab}^{(2)}}{\sqrt{G_{aa}^{(2)} G_{bb}^{(2)}}} \leq 1. \quad (4.11)$$

On the other hand, take $N$ identical bosons equally occupying two modes $\hat{a}$ and $\hat{b}$ in a particle-entangled twin-Fock state of $|\psi_{tf}\rangle = \frac{N}{2} |\frac{N}{2}, \frac{N}{2}\rangle$. In this case $G_{aa}^{(2)} = \langle \hat{a}^\dagger \hat{a}^\dagger \hat{a} \hat{a} \rangle$ and analogical expressions hold for $G_{bb}^{(2)}$ and $G_{ab}^{(2)}$, giving $C = 1 + \frac{2}{N-2}$, which violates the CSI. To summarize, we have shown that:

(a) the CSI is satisfied by all separable states of identical bosons, and

(b) there exists an entangled state, by which the CSI is violated.

Therefore, the CSI is a criterion for particle entanglement.

4.2.1 Relation to number squeezing

A broad family of entangled states, which violate the CSI, can be identified using the number-squeezing parameter defined as a variance of the population imbalance operator between the two regions, $\eta^2 = \frac{\langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2}{n_{\text{tot}}}$. Here $\hat{n} = \hat{n}_a - \hat{n}_b$ and $n_{\text{tot}} = \langle \hat{n}_a \rangle + \langle \hat{n}_b \rangle$ with $\hat{n}_{a/b} = \int_{V_{a/b}} dx \hat{\Psi}^\dagger(x) \hat{\Psi}(x)$. This parameter can be expressed in terms of the local- and cross-correlations as follows

$$\eta^2 = 1 + \frac{G_{aa}^{(2)} + G_{bb}^{(2)} - 2G_{ab}^{(2)} - \langle \hat{n} \rangle^2}{n_{\text{tot}}} \quad (4.12)$$

If we are dealing with a balanced state with $\langle \hat{n} \rangle = 0$ and $G_{aa}^{(2)} = G_{bb}^{(2)}$, Eq. (4.12) gives

$$\eta^2 = 1 + 2 \frac{(1 - C) G_{aa}^{(2)}}{n_{\text{tot}}}. \quad (4.13)$$

The system is number squeezed ($\eta^2 < 1$) when $C > 1$, so the number-squeezing is equivalent to the violation of the CSI and is a signature of particle entanglement. If the state is not balanced, one cannot link the number squeezing with the CSI, due to presence of the non-vanishing term $\langle \hat{n} \rangle^2$. To confirm this we take a separable state of $N$ particles in a pure state $|\phi_0\rangle \otimes^N$, divided into two unequal parts. We obtain $\eta^2 = 1 - \frac{1}{N} \langle \hat{n} \rangle^2 < 1$, so the number squeezing is present without entanglement. Clearly, the CSI is more universal than the number squeezing parameter, because its violation always implies entanglement of identical bosons and its construction does not require any assumptions about the two regions. Nevertheless, the relation between the integrated correlation functions (4.9) and $\eta^2$ from Eq. (4.12) is a strong
suggestion that the violation of the CSI is more likely to manifest in systems, where the fluctuations between the regions are reduced rather than enhanced. In line with this argument, the CSI criterion does not detect entanglement of the NOON state $\frac{1}{\sqrt{2}}(|N,0\rangle + |0,N\rangle)$, which has maximal fluctuations between the modes, i.e. $\eta^2 = N$, while $C = 0$.

For two-mode systems, another criterion for particle entanglement is the spin-squeezing quantified by the parameter $\xi_s^2 [70, 71, 130]$ (see also section 2.7.4 in chapter 2). It is equal to $\xi_s^2 = \eta^2/\nu^2$, where the quantity $\nu^2$ is related to the coherence between the modes. The system is particle entangled and useful for sub shot-noise interferometry when $\xi_s^2 < 1 [29, 69]$. Since $\nu^2 \leq 1$ always holds, then $\xi_s^2 \geq \eta^2$. Thus, for balanced systems, $\xi_s^2 < 1$ implies $C > 1$, and the system is particle entangled. As a consequence, when $C < 1$, then also $\xi_s^2 > 1$. Therefore, the CSI signifies if a balanced two-mode state is potentially valuable for ultra-precise metrology. For unbalanced systems, there is no straightforward connection of $\eta^2$ to $C$, and hence $\xi_s^2$ and $C$ are not simply related in this case.

### 4.2.2 Generalization to non-fixed $N$

The CSI criterion from Eq. (4.11) applies to systems with fixed number of particles, such as an array of trapped bosonic ions [131] or pairs of photons post-selected from the parametric down conversion signal [132, 133]. However, for large systems, the number of particles is usually hard to control and differs between experimental realizations. In this context, it is relevant to extend above considerations to the situation when $N$ fluctuates. In such case, in the absence of coherence between states with different $N$’s, the density matrix of a separable state reads

$$\hat{\rho} = \sum_{N=0}^{\infty} p_N \int D\phi |\phi;N\rangle\langle\phi;N| P_N(\phi).$$  \hspace{1cm} (4.14)$$

Here, $p_N$ is the probability for having $N$ particles in the system, while $P_N(\phi)$ for each $N$ satisfies condition (4.7). The CSI calculated with the density matrix (4.14) involves the integrals (4.9) averaged with $p_N$, thus whether the CSI is violated or not depends on the properties of

$$\langle P_N(\phi) \rangle = \sum_{N=0}^{\infty} p_N N(N-1) P_N(\phi).$$  \hspace{1cm} (4.15)$$

This is much different from the fixed-$N$ case, where the violation of the CSI is related to the characteristics of a single $P(\phi)$. By inspecting the formula (4.15), we note that even if certain $P_N(\phi)$ is partially negative, the CSI can overlook entanglement present in this $N$-particle sector if the overall $\langle P_N(\phi) \rangle$ is positive. In other words, from the point of view of the CSI, the separable part of the density matrix can overshadow the entangled component.
4.3 CSI vs. Fisher information

The CSI criterion can be compared with yet another method of detecting entanglement in many-body systems known from quantum metrology. The Quantum Fisher information (QFI), denoted here by $F_Q$, provides a lower bound for the precision $\Delta \theta$ of the estimation of an unknown parameter $\theta$ in a series of $m$ experiments, $\Delta \theta \geq \frac{1}{\sqrt{m}} \frac{1}{\sqrt{F_Q}}$. The value of $F_Q$ is determined by the properties of the state $\hat{\rho}$ and the transformation, which introduced the $\theta$-dependence in the system. A particularly important case appears when $\theta$ is the relative phase between two modes $\hat{a}$ and $\hat{b}$ of $\hat{\rho}$, imprinted by an interferometer represented by a unitary transformation $e^{-i\hat{n} \cdot \hat{J}}$. Here $\hat{n} \cdot \hat{J}$ is a product of a unit vector and the vector of angular momentum operators $\hat{J}_x = (\hat{a}^\dagger \hat{b} + \hat{b}^\dagger \hat{a})/2$, $\hat{J}_y = (\hat{a}^\dagger \hat{b} - \hat{b}^\dagger \hat{a})/2i$ and $\hat{J}_z = (\hat{a}^\dagger \hat{a} - \hat{b}^\dagger \hat{b})/2$. Such interferometric transformation addresses each particle independently, so it cannot entangle them, and $F_Q \leq N$ holds for all separable $\hat{\rho}$ [29]. Consequently, all two-mode states for which $F_Q > N$ must be entangled. Typically it is not even necessary to find $\Delta \theta$ to estimate the value of $F_Q$. Often some quantity $\chi < F_Q$ is measured, for instance the inverse of the spin-squeezing parameter [18,20,21,93,134–136]. If $\chi > N$, then also $F_Q > N$ meaning that the system is particle-entangled.

Differently from the interferometric criterion, the CSI from Eq. (4.11) does not require any assumptions about the modal structure of $\hat{\rho}$. The metrological approach is inevitably related to an interferometric transformation, so to experimentally confirm the entanglement one is usually bound to use this interferometer [24]. On the contrary, the CSI criterion is not linked with any transformation, and is derived solely from the properties of the integrated correlation function (4.9).

The metrological approach seems more powerful than the violation of the CSI, because it is sensitive to a wider spectrum of entangled states – due to the freedom of choice of the interferometric apparatus. This can be illustrated by considering the bare phase imprint, which is represented by the transformation with $\hat{n} = (0,0,1)$. Such interferometer gives for the NOON state $F_Q = N^2$, while the CSI criterion overlooks the entanglement of this state, as argued above. Another advantage of the QFI can be shown using a twin-Fock state $|\psi_{tf}\rangle$ introduced above, passing through the Mach-Zehnder interferometer represented by the interferometric transformation with $\hat{n} = (0,1,0)$. In this case $F_Q = N + \frac{N^2}{2}$ showing that the state is strongly entangled and the correction to the no-entanglement limit $F_Q = N$ is quadratic in $N$. In contrast, the CSI gives $C = 1 + \frac{2}{N^2}$ so the deviation from the classical limit is negligible for large $N$.

The above example underlines an important difference between the QFI approach and the CSI criterion. The former, although usually difficult to implement, exploits information about the whole density matrix. The latter, much easier to check experimentally, is based solely on the second-order correlation function. Therefore, when $N$ is large, some knowledge about the non-classical relations between the particles, contained in higher order correlations, is lost.

One can overcome some of these limitations using the CSI calculated with higher
order integrated correlation function between \( n/2 \) particles in region \( a \) and other \( n/2 \) in \( b \). When \( N \) is large and \( n \ll N \), the coefficient \( C \) for a twin-Fock state is \( C \simeq 1 + \frac{1}{2} \frac{n^2}{N} \), displaying stronger deviation from unity as \( n \) grows. Using this approach, one might increase the sensitivity of entanglement detection for instance in cold-atom systems, where ultra-precise measurements of spatial correlations up to sixth order were recently reported [137].

We underline that identical bosons can be entangled simply as a result of symmetrization. Such type of entanglement has been disregarded as a useful resource for quantum information protocols [138,139]. However, it has been recently demonstrated [140] that entanglement of identical particles can be mapped onto the mode entanglement useful for quantum information. As an illustration, consider a pair of identical bosons in a twin-Fock state \(|1, 1\rangle\) of two far separated modes. This state is particle-entangled only as a result of symmetrization. However, it is not mode entangled thus when the particles are far apart and the modes can be addressed separately, this type of entanglement is of no importance. Nevertheless, if the particles are brought together and simultaneously pass a beam-splitter, the indistinguishability comes into play and the state is transformed into the NOON state \( \frac{1}{\sqrt{2}} (|2, 0\rangle + |0, 2\rangle) \), which is both mode- and particle-entangled. The beam-splitter, which acts on modes non-locally, can be regarded as a device, which extracts useful correlations between the modes from a state which is only particle-entangled.

Following this example, it seems feasible to create large ensembles of entangled identical particles “for free”. Is it, for instance, possible to construct an entangled twin-Fock state \(|\frac{N}{2}, \frac{N}{2}\rangle\) by taking two independent Bose-Einstein condensates (BECs) of \( \frac{N}{2} \) identical bosons each? In principle yes, however it might be extremely difficult to achieve in experiment. For large \( N \) it is impossible to control the number of atoms in each condensate from shot to shot. A realistic state of two BECs is a mixture of states with different numbers of particles fluctuating at least at the shot-noise level. These fluctuations preclude the number-squeezing, thus rule out the violation of the CSI and render the state useless for applications, such as quantum metrology. Although some amount of entanglement due to symmetrization might be present in such a mixture, it is drastically reduced with respect to the ideal pure twin-Fock state. Therefore, even when dealing with identical bosons, a dedicated experimental procedure is necessary to create usefully entangled many-body states. We address these issues in more details in chapter 7.

### 4.4 CSI for distinguishable particles

Finally, we demonstrate that the indistinguishability of particles is crucial to establish relation between the violation of the CSI and entanglement. To this end, we calculate the second-order correlation function for a separable state (4.3) without imposing the indistinguishability of particles

\[
G^{(2)}(x, x') = \sum_{n \neq m=1}^{N} \text{Tr} \left[ \rho_{\text{sep}} \hat{\Pi}^{(n)}(x) \otimes \hat{\Pi}^{(m)}(x') \right].
\]  

(4.16)
Here, \(\hat{\Pi}^{(n)}(x)\) projects the \(n\)-th particle onto the position state \(|x\rangle\), while the sum ensures that all possible combinations of one particle being at position \(x\) and the other at \(x'\) contribute to the correlation function. Using Eq. (4.3) we obtain that

\[
G^{(2)}(x, x') = \sum_i p_i \sum_{n \neq m = 1}^N P_i^{(n)}(x)P_i^{(m)}(x'),
\]

(4.17)

where \(P_i^{(n)}(x) = \text{Tr}[\hat{\rho}_i^{(n)} \hat{\Pi}^{(n)}(x)]\) is the one-body probability for finding the \(n\)-th particle in state \(\hat{\rho}_i^{(n)}\) at position \(x\). If particles are identical, then these probabilities do not depend on indices \(n\) and \(m\). In this case, the sum over \(n \neq m\) gives the coefficient \(N(N-1)\) and after integrating \(x\) over volume \(V_i\) and \(x'\) over \(V_j\), we obtain the discrete version of Eq. (4.9). However, if particles are not identical, then the sum over \(n \neq m\) gives

\[
G^{(2)}(x, x') = \sum_i p_i f_i(x, x'),
\]

(4.18)

where \(f_i(x, x')\) does not factorize into a product of functions of \(x\) and \(x'\). In consequence, no such relation as in Eq. (4.11) can be established.

We now provide a simple example, which illustrates that the indistinguishability of particles is crucial for the violation of the CSI to be used as an entanglement probe. Consider two particles occupying two modes \(\phi^{(a)}\) and \(\phi^{(b)}\) in a Werner state [95]:

\[
\hat{\rho}_w = \frac{1 - p}{4} \mathbb{1} + p |\psi_1\rangle \langle \psi_1|,
\]

(4.19)

where \(|\psi_1\rangle = \frac{1}{\sqrt{2}}(|\phi^{(a)}_1, \phi^{(b)}_2\rangle + |\phi^{(a)}_2, \phi^{(b)}_1\rangle)\) and \(0 \leq p \leq 1\). The identity operator \(\mathbb{1}\) is spanned by the triplet of bosonic vectors \(|\psi_1\rangle, |\psi_2\rangle = |\phi^{(a)}_1, \phi^{(a)}_2\rangle, |\psi_3\rangle = |\phi^{(b)}_1, \phi^{(b)}_2\rangle\) and a fermionic singlet \(|\psi_4\rangle = \frac{1}{\sqrt{2}}(|\phi^{(a)}_1, \phi^{(b)}_2\rangle - |\phi^{(a)}_2, \phi^{(b)}_1\rangle)\). Therefore \(\hat{\rho}_w\) does not represent a pair of indistinguishable bosons apart from \(p = 1\). Nevertheless, the second order correlation function and the CSI can be still calculated using Eq. (4.19). The result is that \(G^{(2)}_{aa} = \text{Tr}[\hat{\rho}_w |\psi_2\rangle \langle \psi_2|] = \frac{1-p}{4}\) and symmetrically \(G^{(2)}_{bb} = \frac{1-p}{4}\) while \(G^{(2)}_{ab} = \frac{1+p}{2}\). By combining these expressions we obtain \(C = 2\frac{1+p}{1-p} > 1\) for all \(0 \leq p \leq 1\), although \(\hat{\rho}_w\) is entangled only when \(p > \frac{1}{3}\), according to the Peres-Horodecki criterion [117,141]. This confirms that indeed the violation of the CSI does not imply entanglement of distinguishable particles.

### 4.5 Conclusions

In conclusion, in this chapter we demonstrated that the analogue of the Glauber-Sudarshan P-representation can be employed to show that the violation of the Cauchy-Schwarz inequality for the second-order correlation function proves entanglement in any system of identical bosons. The Cauchy-Schwarz inequality condition is not as powerful as those inferred from quantum metrology, but usually is much
simpler to implement. It could be used as a direct test of entanglement for instance in systems where the number-squeezing is likely to be present. Among those are the twin-beam setups [27, 28], halos of particles scattered in the Bose-Einstein condensates collisions [25,142–145] and many other correlated systems. After the Cauchy-Schwarz inequality violation is demonstrated, the next step towards applications of systems of entangled particles can be made.

In the following chapter, we investigate usefulness of the collisions of Bose-Einstein condensates as a potential source of non-classically correlated atomic pairs. The whole analysis is based on the number-squeezing parameter $\eta^2$, which, in the presence of symmetry, is, as we now know, directly related to the Cauchy-Schwarz inequality.
Chapter 5

Tradeoffs for number-squeezing in collisions of Bose-Einstein condensates

Contents

5.1 Introduction and Outline ........................................... 90
5.2 Theoretical model for BEC collision ............................... 91
  5.2.1 Collision parameters ........................................ 92
  5.2.2 BEC wave-function .......................................... 92
  5.2.3 Positive-P method ........................................... 94
  5.2.4 Scattering in the absence of bosonic enhancement ............ 94
5.3 Two-body correlations of scattered atoms .......................... 96
5.4 Number squeezing in a multi-mode system ......................... 97
5.5 Simulation results and analysis ................................... 100
  5.5.1 $\gamma = 1.02$ case .......................................... 100
  5.5.2 $\gamma = 0.24$ case .......................................... 103
  5.5.3 Collision of two plane-waves ................................ 105
5.6 Conclusions .......................................................... 106

This chapter\textsuperscript{1} is devoted to the analysis of factors that influence the usefulness of collisions of Bose-Einstein condensates as a potential source of entangled atomic pairs. We show that the non-monochromaticity of the mother clouds is typically the leading limitation on number squeezing, and that the squeezing becomes less robust to this effect as the density of pairs grows.

We also develop a simple model that explains the relationship between density correlations and the number squeezing. The model allows one to estimate the squeezing from the second order correlation function, and shows that the multi-mode nature of the scattering must be taken into account to understand the behaviour of the pairing. We analyse the impact of the Bose enhancement on the number squeezing, by introducing a simplified low-gain model. We conclude that as far as the squeezing is concerned the preferable configuration occurs when atoms are scattered not uniformly but rather into two well separated regions.

5.1 Introduction and Outline

A supersonic collision of two Bose-Einstein condensates (see discussion on page 13) is a source of strongly correlated atomic pairs, which may be potentially used to create spatially-separated entangled states of massive particles [25, 142, 145–149]. Such states could be used to extend the study of the Einstein-Podolsky-Rosen paradox [97, 150], local realism [133] and Bell inequality tests [151] into a regime where rest mass is non-negligible. In the context of quantum metrology, usefully entangled states allow one to surpass the shot-noise limit (known also as the Standard Quantum Limit) – the maximum parameter estimation precision allowed by classical physics [29, 110].

Scattering of atomic pairs can lead to reduced fluctuations of the relative population between two regions of opposite momenta. This effect, called the number-squeezing and a closely related violation of the Cauchy-Schwartz inequality have been recently observed in experiments [25–28]. Number squeezing, if accompanied by sufficiently high mutual coherence between the regions, is indicative of spin squeezing [70, 71]. Spin squeezed states, which are known to be usefully entangled from a quantum metrology point of view [29, 71, 130] (see also section 2.7.4 in chapter 2), have been recently engineered in a number of experiments [18, 20, 21, 93, 135, 136]. In addition, the perfectly number-squeezed “twin-Fock” state, which is not spin-squeezed but is nevertheless strongly entangled, has been recently observed by Lücke et al [24]. Also pair-production setups were invented, where scattering is preferentially directed into only several spatial modes. These then contain more pairs per mode, and it is more convenient to bin them and possibly process further [151]. In [143–145] a four-wave-mixing type of process between two species of atoms was used to populate two counter-propagating clouds. In [28] a BEC was transferred into the first excited state of a trapping potential and subsequent two-body collisions created a “twin-beam” system of correlated pairs. Finally, another approach used an optical lattice to allow correct phase-matching conditions into only a few selected modes [27].

In some recent experiments, where atoms scatter into two well defined regions, a halo of overlapping spontaneously scattered modes and other imperfections such as spatial inhomogeneity of the mother clouds [152, 153] might limit the amount of number squeezing. An understanding of the main limitations and tradeoffs involved is important for future experiments.

We consider in detail the number difference squeezing between atoms with opposite momenta in a collision of two non-monochromatic BECs [25, 149]. We simulate the collision using the positive-P method in the Bogoliubov approximation [154]. We also introduce a simple yet intuitive model, which relates the number squeezing to the second order correlations between the scattered atoms and demonstrate its validity in a wide range of parameters. The main conclusion of the discussion presented in this chapter, is that the non-monochromatic nature of the mother clouds is the main limiting factor to strong number-squeezing in the scattering halo. We argue that in the presence of many competing modes, smaller cloud densities are ad-
5.2 Theoretical model for BEC collision

Vantageous, because mode-mixing effects are destructive for the number-squeezing and they become more pronounced with higher cloud density.

The chapter is organized as follows. In Section 5.2 we introduce a model, which describes the creation of pairs in BECs collisions. We discuss the relevant physical parameters and develop an alternative low-gain model, useful for simulations when the bosonic amplification of pair scattering can be neglected. In Section 5.3 we calculate the second order correlation function of the scattered atoms and explain how it consists of two parts – co-linear and back-to-back momentum correlations, like in [155]. In Section 5.4 we calculate the number-squeezing parameter and using a simple Gaussian model relate it to the second-order coherence of the system. In Section 5.5 we present the numerical results in the regime of both high- and low-density of the mother BECs and compare these results with the model. In the course of this analysis, the factors affecting the number squeezing become apparent. We conclude in Section 5.6. Some technical details of the calculations and numerics presented in the main text are discussed in the Appendix B.

5.2 Theoretical model for BEC collision

A single stationary BEC can be split into a superposition of two counter-propagating wave-packets by means of Bragg scattering [156]. In a center-of-mass reference frame, the average velocity of each cloud is \( \pm v_{\text{rec}} \) – a recoil velocity due to an absorption and subsequent emission of a Bragg photon. If the relative velocity \( 2v_{\text{rec}} \) is approximately above the speed of sound \( c = \sqrt{g\rho/m} \), i.e. when the Mach number

\[
\text{Ma} = 2v_{\text{rec}}/c > 1,
\]

elastic collisions of particles from the two clouds lead to scattering of atomic pairs out of the BECs. Here \( \rho = N/(4/3\pi R_{\text{TF}}^3) \) is the average density of an isotropic condensate in the Thomas-Fermi approximation with radius \( R_{\text{TF}} \).

The dynamics of the scattering is governed by the energy- and momentum-conservation laws, \( v_1^2 + v_2^2 \simeq 2v_{\text{rec}}^2 \) and \( v_1 + v_2 \simeq 0 \), where index \((1,2)\) labels the pair members. The equalities are only approximate, due to the momentum spread of the initial BEC and finite duration of the collision. These conservation laws dictate that atoms are scattered onto a shell of radius \( \nu_{\text{rec}} \) centered around zero, called the scattering halo. This phenomenon was observed in many experiments [25,26,142,145–148]. Atoms are usually registered after a long time of free expansion, when their positions approximately correspond to the momentum distribution just after the collision.

Since particles scatter in pairs, there is an expectation that the measured population difference between two opposite regions may fluctuate below the shot-noise limit. In the idealized case of scattering from two plane waves, these fluctuations are suppressed down to zero, in analogy to the simplest model of parametric down-conversion. The study, presented in this chapter, takes on the task of generalizing this simple picture and calculating the number-squeezing in condensate collisions
assuming a realistic shape and time evolution of the source, and including the time-dependent interactions with the mean field after the scattering [157].

5.2.1 Collision parameters

To focus on the essential features of the process, we consider the simplest case of the initial condensate prepared in the ground state of a spherical trap. Depending on the non-linearity and the duration of the collision, the scattering of atoms can be either spontaneous or enter the Bose enhanced regime. Therefore, we introduce a dimensionless parameter

$$\gamma = \frac{t_{\text{col}}}{t_{\text{nl}}}$$

to distinguish between these two possible scenarios. Here, $t_{\text{col}} = R_{\text{TF}}/v_{\text{rec}}$ is the duration of the scattering process, while $t_{\text{nl}} = \hbar/\left(\frac{1}{3} g \bar{\rho}\right)$ is the rate of the two-body collisions. The one-third in the denominator approximately accounts for the fact that collisions between the atoms take place mostly in the center of the trap, where the density is high. It has been demonstrated [157–161] that when $\gamma \gtrsim 1$, the system enters the stimulated regime.

We perform numerical simulations using parameters of metastable $^4$He atoms, i.e. $m = 6.65 \times 10^{-27}$ kg, $a_s = 7.5 \times 10^{-9}$ m and take $v_{\text{rec}} = 10$ cm/s. To address both spontaneous and stimulated regimes we consider the following cases. In the first, which we call the dilute gas case, the sample consists of $N = 17540$ atoms in the initial BEC, and $\omega = 2\pi \times 928$ s$^{-1}$. Here $Ma = 13.12$ and $\gamma = 0.24$, hence the scattering is spontaneous all along. In the opposite dense gas case, we take $N = 74360$ and $\omega = 2\pi \times 1911$ s$^{-1}$, which gives $Ma = 6.56$ and $\gamma = 1.02$, so the Bosonic enhancement becomes significant. In both cases, we simulate the scattering process until $t = 1.7 t_{\text{coll}}$ – a time at which the collision is completed. In Fig. 5.1, cuts through the halo density are shown for the dense case at the end of the collision. The result was obtained using the positive-P numerical method, which is discussed in detail below. Note that although the ensemble average of the momentum distribution is symmetric around zero, isolated density grains are present in a single realization, and they show an evident and massive lack of this symmetry. This is related to increased fluctuations of the population imbalance between regions with opposite momentum, and will be investigated further below.

5.2.2 BEC wave-function

We model the initial trapped BEC by a c-number wave-function $\phi_0(x)$ which is a solution of the stationary Gross-Pitaevskii equation,

$$\mu \phi_0(x) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + V_{\text{trap}}(x) + g|\phi_0(x)|^2 \right] \phi_0(x). \quad (5.1)$$

Here $m$ is the atomic mass, $\mu$ is the chemical potential, $g = 4\pi \hbar^2 a_s/m$ the interaction strength related to s-wave scattering length $a_s$, and $V_{\text{trap}}(x) = \frac{1}{2} m \omega^2 x^2$ is the harmonic trapping potential with a frequency $\omega$. The collision is triggered by a pair
5.2. Theoretical model for BEC collision

Figure 5.1: Outcome of a single realization of a collision of two BECs in the dense $\gamma = 1.02$ case (in a simulation of (5.4, 5.5)). Shown are cross-sections through the density of atoms after the end of the collision ($t = 1.6t_{\text{coll}}$), in the $k_z - k_y$ momentum plane (top), and $k_x - k_y$ plane (bottom). Clearly, a halo of scattered atoms is forming, nevertheless the atoms at the opposite regions of the halo are poorly aligned, which implies the absence of number squeezing in the system. Data scaled to units of $[k_0]^{-3}$. The white areas are saturated at this colour scale, the bold (thin) black contours are at 0.1 (0.01) of the peak condensate density.

of brief Bragg pulses shined onto the condensate [156]. Consequently, a superposition of two counter-propagating, mutually coherent atomic clouds is created

$$\phi(x, t = 0) = \phi_0(x) e^{ik_0 z} + e^{-ik_0 z}/\sqrt{2} \equiv \phi_+ (x) + \phi_- (x),$$  \hspace{1cm} (5.2)

where $k_0 = mv_{\text{rec}}/\hbar$ is the wave-vector associated with the recoil velocity. After the pulses are applied, the trap is switched off, and the two condensates begin to move apart, activating the collision process.
5.2.3 Positive-P method

To describe the scattered atoms we introduce a bosonic operator $\hat{\delta}(x, t)$. In the spirit of the Bogoliubov approximation, we use the linearized equations of motion for the field $\hat{\delta}$, assuming that the number of scattered atoms is a small fraction of the condensate population and the self-interaction of $\hat{\delta}$ can be neglected,

$$ih \partial_t \hat{\delta}(x, t) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + 2g|\phi(x, t)|^2 \right] \hat{\delta}(x, t) + g\phi^2(x, t) \hat{\delta}^\dagger(x, t), \quad (5.3)$$

The coherent mean field part $\phi(x, t)$ during the collision evolves according to the time-dependent Gross-Pitaevskii equation,

$$ih \partial_t \phi(x, t) = \left[ -\frac{\hbar^2}{2m} \nabla^2 + g|\phi(x, t)|^2 \right] \phi(x, t). \quad (5.4)$$

To solve the coupled equations (5.3) and (5.4), we use the stochastic positive-P method [154], where instead of directly solving Eq. (5.3) for $\hat{\delta}$ we sample the distribution of two complex fields $\psi(x, t)$ and $\tilde{\psi}(x, t)$. The Bogoliubov equation (5.3) corresponds to a pair of stochastic Itô equations,

$$ih \partial_t \psi(x, t) = \left( -\frac{\hbar^2}{2m} \nabla^2 + 2g|\phi(x, t)|^2 \right) \psi(x, t) \quad (5.5a)$$

$$+ g\phi^2(x, t)\tilde{\psi}(x, t)^* + \sqrt{i\hbar g} \phi(x, t) \xi(x, t), \quad (5.5b)$$

$$ih \partial_t \tilde{\psi}(x, t) = \left( -\frac{\hbar^2}{2m} \nabla^2 + 2g|\phi(x, t)|^2 \right) \tilde{\psi}(x, t) \quad (5.5c)$$

$$+ g\phi^2(x, t)\psi(x, t)^* + \sqrt{i\hbar g} \phi(x, t) \tilde{\xi}(x, t). \quad (5.5d)$$

Here $\xi(x, t)$ and $\tilde{\xi}(x, t)$ are independent, real stochastic Gaussian noise fields with zero mean and second moments given by $\langle \xi(x, t)\tilde{\xi}(x', t') \rangle = 0$ and

$$\langle \xi(x, t)\xi(x', t') \rangle = \langle \tilde{\xi}(x, t)\tilde{\xi}(x', t') \rangle = \delta(x - x')\delta(t - t'). \quad (5.6)$$

Within the stochastic positive-P method, any physical quantity can be obtained with the mapping $\hat{\delta} \rightarrow \psi$ and $\hat{\delta}^\dagger \rightarrow \tilde{\psi}^*$, and replacing quantum averages of normally-ordered operators by stochastic averages [162]. Equations (5.5) recover the exact quantum dynamics of Eq. (5.3) in the limit of an infinite number of samples.

5.2.4 Scattering in the absence of bosonic enhancement

The main goal of this study is to investigate how the number squeezing between two regions of the halo is affected by various phenomena which occur during the collision. Various phenomena influence the dynamics in a complicated way and it seems very advantageous if we could, at least theoretically, “turn on/off” some of them to isolate their effects. For instance, the impact of the BEC mean-field on the scattered particles can be easily “controlled” by including or excluding the second term of the right-hand-sides in lines (5.5a) and (5.5c). Similarly, the mean-field
repulsion in the evolution of the source BEC can be controlled by setting \( g = 0 \) by hand in Eq. (5.4). One can further simplify the dynamics by modelling the two counter-propagating condensates with non-expanding Gaussians, substituted for \( \phi(x, t) \) into Eqs (5.5).

Finally, we propose a simple perturbative model, which describes the condensate collision in the absence of bosonic stimulation. We start by introducing a hierarchy of fields

\[
\hat{\delta}(x, t) = \sum_{j=0}^{\infty} \hat{\delta}^{(j)}(x, t). \tag{5.7}
\]

The lowest order term \( \hat{\delta}^{(0)}(x, t) \) is the solution of the free equation (i.e. without the additional particle creation term in the Bogoliubov field)

\[
i\hbar \partial_t \hat{\delta}^{(0)}(x, t) = H_0(x, t)\hat{\delta}^{(0)}(x, t), \tag{5.8}
\]

where \( H_0(x, t) = -\frac{\hbar^2}{2m} \nabla^2 + 2g|\phi(x, t)|^2 \). The higher terms of expansion (5.7) evolve according to

\[
i\hbar \partial_t \hat{\delta}^{(j)}(x, t) = H_0(x, t)\hat{\delta}^{(j)}(x, t) + g\phi^2(x, t)\hat{\delta}^{(j-1)}(r, t). \tag{5.9}
\]

In this approach, the bosonic enhancement appears only in the higher order fields and can be excluded by restricting the dynamics only to the two lowest ones, namely \( j = 0 \) and \( j = 1 \).

The set of two coupled equations for \( j = 0 \) and \( j = 1 \) can be formally solved by replacing the operators with complex stochastic fields, i.e. \( \hat{\delta}^{(j)}(x, t) \rightarrow \delta^{(j)}(x, t) \) and \( \hat{\delta}^{(j)}(x, t) \rightarrow \delta^{(j)}(x, t)^\ast \). Then, the c-number equivalent of (5.8) and (5.9) is solved numerically from the initial conditions consisting of \( \delta^{(1)}(x, 0) = 0 \) and setting \( \delta^{(0)}(x, 0) \) as a random complex Gaussian field with zero mean and the variances \( \langle \delta^{(0)}(x, 0)\delta^{(0)}(x', 0) \rangle = \delta(x - x') \) and \( \langle \delta^{(0)}(x, 0)\delta^{(0)}(x', 0) \rangle = 0 \). It is important to note that – contrary to the positive-P equations (5.5) – the stochasticity is introduced only through the initial conditions, very much like in the truncated Wigner method, but with twice the variance.

With the solutions \( \delta^{(0)}(x, t) \) and \( \delta^{(1)}(x, t) \) at hand, one can reproduce the observables. For instance the lowest order correlation function is given by

\[
\langle \hat{\delta}^\dagger(x, t)\hat{\delta}(x', t) \rangle = \overline{\delta^{(1)}(x, t)^\ast \delta^{(1)}(x', t)}, \tag{5.10}
\]

where the over-bar denotes averaging over the ensemble of initial conditions.

In the following Section, we analyse some formal properties of the Bogoliubov equation (5.3) and introduce the second order correlation function of scattered atoms, which, as will be argued in Section 5.4, determines the amount of number-squeezing in the system.
5.3 Two-body correlations of scattered atoms

The normalized second order correlation function of scattered atoms in momentum space is defined as

\[ g^{(2)}(k, k') = \frac{\langle \hat{\delta}^\dagger(k) \hat{\delta}(k') \hat{\delta}(k') \hat{\delta}(k) \rangle}{\langle \hat{\delta}^\dagger(k) \hat{\delta}(k) \rangle \langle \hat{\delta}(k') \hat{\delta}(k') \rangle}. \]  

(5.11)

Henceforth, we omit the explicit time dependence from the operator \( \hat{\delta}(k, t) \). Note that since the Bogoliubov equation of motion (5.3) is linear and the initial state of scattered atoms is a vacuum, with the help of Wick’s theorem we can write

\[ g^{(2)}(k, k') = 1 + \frac{|G^{(1)}(k, k')|^2 + |M(k, k')|^2}{G^{(1)}(k, k)G^{(1)}(k', k')} \]. \( (5.12) \)

Here, \( G^{(1)} \) is the one-body density matrix of the scattered atoms defined as

\[ G^{(1)}(k, k') = \langle \hat{\delta}^\dagger(k) \hat{\delta}(k') \rangle, \]  

(5.13)

and \( M \), the anomalous density \([159, 161]\), is

\[ M(k, k') = \langle \hat{\delta}(k) \hat{\delta}(k') \rangle. \]  

(5.14)

In order to find a natural interpretation of the components of (5.12) we make some further simplifications in our model, that are used only in this Section. First, consider the following simplified model \([158, 159, 161]\) of the collision dynamics, described by the equation

\[ i\hbar \partial_t \hat{\delta}(x, t) = -\frac{\hbar^2 \nabla^2}{2m} \hat{\delta}(x, t) + 2g \phi_+ (x, t) \phi_- (x, t) \hat{\delta}^\dagger(x, t), \quad (5.15) \]

Here, a pair of atoms is taken from counter-propagating condensates \( \phi_{\pm} \) – defined in Eq. (5.2) – and placed in the field of scattered atoms. As compared with Eq. (5.3), this model neglects the impact of the mean field of the colliding BECs on the scattering process, as well as terms proportional to \( \phi_2 \) that are strongly non energy-conserving in the halo.

Now let us make a second simplification regarding the internal dynamics of the two colliding wave-packets. In general the two functions \( \phi_{\pm}(x, t) \) evolve according to the Eq. (5.4), but for the sake of the present considerations we neglect the non-linear term and use the equations of motion of free expansion. In this case, \( \phi_{\pm}(k, t) = \phi_{\pm}(k)e^{-i\hbar \phi_{\pm}(x,t)/2m}. \) Such a “reduced Bogoliubov” model with these two approximations has been widely used previously to investigate the dynamics of the collision \([158–161, 163–169]\), and was investigated in some detail in [157].

Equation (5.15) can be solved up to the first order in the perturbative regime, to obtain at times long after the collision

\[ M(k, k') = \frac{mg}{\kappa^2 \hbar^2} \int d^3k_1 d^3k_2 \phi_- (k_1) \phi_+ (k_2) \times \]

(5.16)

\[ \times \delta(k_1^2 + k_2^2 - k^2 - k'^2) \delta^{(3)}(k_1 + k_2 - k - k'). \]
The anomalous density $M(k, k')$ can be interpreted as the probability amplitude for having one atom with momentum $k$ and the second with $k'$. These two momenta come from a coherent superposition of probability amplitudes describing elementary collision events (energy and momentum conservation laws are satisfied) between atoms from the BECs that come with probability amplitudes $\phi_{\pm}$. Since the condensate functions $\phi_{\pm}$ in (5.16) are localized around $\pm k_0$, the anomalous density $M(k, k')$ is non-vanishing only when $k \simeq -k'$ and $|k| \simeq k_0$. In this sense, $M$ describes scattered atomic pairs, correlated for opposite momenta. Related arguments have been presented in \cite{155}.

Using similar arguments, we obtain a useful relation that is valid within Bogoliubov theory in the perturbative regime

$$G^{(1)}(k, k') = \int d^3k'' M^*(k, k'')M(k'', k'). \tag{5.17}$$

As we argued above, the first anomalous density gives the contribution to the integral when $k \simeq -k''$ and the second when $k'' \simeq -k'$, hence the one-body density matrix is non-zero only if $k \simeq k'$ and $|k| \simeq k_0$. In conclusion, the scattered atoms are described either by the co-linear part for $k \simeq k'$, 

$$g^{(2)}_{cl}(k, k') = g^{(2)}(k, k' \approx k) \approx 1 + \frac{|G^{(1)}(k, k')|^2}{G^{(1)}(k, k)G^{(1)}(k', k')} \tag{5.18}$$

or the back-to-back part when their wave-vectors are almost opposite $k \simeq -k'$,

$$g^{(2)}_{bb}(k, k') = g^{(2)}(k, k' \approx -k) \approx 1 + \frac{|M(k, k')|^2}{G^{(1)}(k, k)G^{(1)}(k', k')} \tag{5.19}$$

In our numerical simulations we have seen that the above interpretation of $M$ and $G^{(1)}$ is valid to a high degree of accuracy for a wide range of parameters, even when the assumptions introduced above are not fully valid. Therefore, throughout the rest this work, we use the division of the second order correlation function into the co-linear (5.18) and back-to-back part (5.19).

At this stage we are ready to introduce the number-squeezing parameter and show how it relates to the $g^{(2)}$ correlation function in various relevant regimes.

## 5.4 Number squeezing in a multi-mode system

In this section we define the number-squeezing parameter and show how it is related to the second-order coherence of the system.

Atoms are registered (and counted) in two bins, say $a$ and $b$, encompassing volumes $V_{a/b}$ in momentum space. The corresponding atom-number operators read

$$\hat{n}_{a/b} = \int_{V_{a/b}} d^3k \hat{\delta}^\dagger(k)\hat{\delta}(k). \tag{5.20}$$
We introduce the number-difference operator \( \hat{n} = \hat{n}_a - \hat{n}_b \) and define the number-squeezing parameter as follows

\[
\eta^2 = \frac{\Delta^2 \hat{n}}{\bar{n}},
\]

where \( \bar{n} = \langle \hat{n}_a \rangle + \langle \hat{n}_b \rangle \) is the total number of particles in both bins. States that have sub-Poissonian population imbalance fluctuations \( \eta^2 < 1 \) are called number squeezed. As was demonstrated in chapter 4, in the symmetric case \( \langle \hat{n}_a \rangle = \langle \hat{n}_b \rangle \), number squeezing is equivalent to violation of the Cauchy-Schwartz inequality [26], which implies the presence of non-classical correlations between bosons in the system.

Using Eq. (5.20) and the definition of \( g^{(2)} \) from Eq. (5.11), we obtain that

\[
\eta^2 = 1 + \frac{G_{aa} + G_{bb} - 2G_{ab}}{\bar{n}},
\]

where the \( G_{ij} \) stands for a two-fold integral

\[
G_{ij} = \int_{V_i} \int_{V_j} d^3k d^3k' g^{(2)}(k, k') n(k) n(k').
\]

Here \( n(k) = \langle \hat{\delta}^\dagger(k) \hat{\delta}(k) \rangle \). Note that if the two regions \( a \) and \( b \) are located on the opposite sides of the halo, \( G_{aa} \) and \( G_{bb} \) depend on the co-linear correlation function \( g^{(2)}_{cl} \), while \( G_{ab} \) is a functional of \( g^{(2)}_{bb} \).

To handle the integrals (5.23) and evaluate the number-squeezing parameter \( \eta^2 \) for typical situations, let us model the normalized co-linear and back-to-back correlation functions by Gaussians

\[
g^{(2)}_{cl}(k, k') = 1 + h_{cl} \prod_{i=z,t,r} e^{-\frac{(k_i - k'_i)^2}{2(\sigma_{cl}^i)^2}},
\]

for \( k \simeq k' \) and

\[
g^{(2)}_{bb}(k, k') = 1 + h_{bb} \prod_{i=z,t,r} e^{-\frac{(k_i + k'_i)^2}{2(\sigma_{bb}^i)^2}},
\]

for \( k \simeq -k' \). This way we only need to extract the amplitudes \( h_{cl/bb} \) and the widths \( \sigma_{cl/bb}^i \) from the results of the numerical simulations, similarly to the analysis of the Cauchy-Schwartz violation in [26]. The Gaussian form is a very good match to the calculated and also the experimental correlation shapes.

The product runs over three orthogonal directions where the axis \( z \) is along the collision direction, while \( r \) and \( t \) are orthogonal to each other and lie in the \( x-y \) plane, corresponding to radial (\( r \)) and tangential (\( t \)) directions. Additionally, we model the density \( n(k) \) in (5.23) in the following way. We assume that the bin widths \( L_z \) and \( L_t \) (in the \( z \) and \( t \) directions respectively) are small enough for the density to be practically constant. This assumption is in our case well satisfied. On the other hand, the density quickly decays in the \( r \) direction. To account for this
5.4. Number squeezing in a multi-mode system

Figure 5.2: The number squeezing parameter $\eta^2$ in the dense $\gamma = 1.02$ case as a function of time for three different bin volumes given in $k_0^3$ units. The solid lines are from a numerical solution of the full model (5.4, 5.5). The dashed lines are predictions of the Gaussian correlation model (5.26). The dash-dotted horizontal line denotes the shot-noise level $\eta^2 = 1$.

drop relatively simply, we model the density in this direction with a step function, centered around the peak of the halo. The width $w_r$ is deduced from a Gaussian fit to the radial profile of the halo density obtained numerically.

As shown in the Appendix B.3, the assumptions introduced above lead to the approximate expressions

$$
\eta^2 = 1 + \bar{n} \left( \frac{\sigma_{cl}^2}{\sigma_{bb}^2} \left( \frac{h_{cl}}{f_{cl}^2} \frac{f_{cl}^2}{f_{bb}^2} - h_{bb} \frac{f_{bb}^2}{f_{bb}^2} \right) \right). 
$$

(5.26)

Here we have introduced the function

$$
f_{cl/bb} = \frac{1}{u_i^2} \left[ u_i \sqrt{\frac{\pi}{2}} \text{erf}(u_i \sqrt{2}) - \frac{1}{2} \left( 1 - e^{-2u_i^2} \right) \right].
$$

(5.27)

and the normalized bin widths are $u_{cl/bb}^z = L_{z,t} / 2\sigma_{cl/bb}^z$ in the $z,t$ directions, while in the radial direction the limited density manifests itself by the width $u_{cl/bb}^r = \min(L_r, w_r) / 2\sigma_{cl/bb}^r$. The limiting behaviour of $f$ is $1 - u^2 / 3$ for small bins $u$ much narrower than the correlations, and $(1/u) \sqrt{\pi/2}$ for large bins $u \rightarrow \infty$. The above parametrization is convenient, though elaborate. However, when bin widths tend to infinity, we obtain a particularly useful expression

$$
\eta^2 = 1 + (2\pi)^{3/2} \bar{n} \left[ h_{cl} \sigma_{cl}^2 \sigma_{cl}^2 - h_{bb} \sigma_{bb}^2 \sigma_{bb}^2 \right].
$$

(5.28)

Hence, in general, in the multi-mode system, number squeezing depends on the number of particles in the bins $\bar{n}$, but also depends on the correlation amplitudes $h_i$ and
the widths of the correlation functions $\sigma_1^{cl/bb}$. The situation simplifies dramatically when the bins are very small. In this case we get

$$\eta^2 = 1 + \frac{n}{2}(h_{cl} - h_{bb}),$$

and the quantum state is number-squeezed as long as $h_{cl} < h_{bb}$.

To place this in context, an additional important remark is in order. If we consider a pure two-mode pair production model governed by the Hamiltonian $\hat{H} = \hat{a}\hat{b} + \hat{a}^\dagger\hat{b}^\dagger$ we obtain simply $\eta^2 \equiv 0$. Therefore we see that in a multi-mode system one cannot use the intuitions from the two-mode model to predict such quantities as the number-squeezing parameter, even when the bins being considered are very small.

![Figure 5.3](image.png)  
Figure 5.3: The number squeezing parameter $\eta^2$ in the dense $\gamma = 1.02$ case as a function of bin volume (log scale) calculated at three different scattering times (in units of $t_{col}$). The solid lines are from a numerical solution of the full model (5.4, 5.5). The dashed lines are predictions of the Gaussian correlation model (5.26). The dash-dotted horizontal line denotes the shot-noise level $\eta^2 = 1$.

### 5.5 Simulation results and analysis

In this Section we perform a systematic study of the number squeezing parameter in condensate collisions and identify the physical phenomena, which have the largest impact on $\eta^2$.

#### 5.5.1 $\gamma = 1.02$ case

We begin with the case of a dense mother cloud, with $\gamma = 1.02$, so that the bosonic enhancement comes into play at some stage during the collision. In Fig. 5.2 we plot $\eta^2$ as a function of time for three different bin volumes $V$ (for details how the bins are...
chosen, refer to Appendix B.1). The solid curves, which are a result of the simulation of the full positive-P Equations (5.5) are compared with the analytical prediction (5.26) based on correlation properties. The latter requires both the height and the widths of the second order correlation function as input, and these are extracted from the simulation. The Figure shows very good agreement between that model and a “direct” evaluation of the number squeezing by counting the number of particles in the two bins. Note that apart from early times, \( \eta^2 > 1 \) so opposite bins do not reveal number squeezing.

To check to what extent the number squeezing is a result of a particular choice of bins, in Fig. 5.3 we plot \( \eta^2 \) as a function of the bin volume \( V \) at three times. Again, the agreement between the model (5.26) and the simulation is very good. This Figure confirms that the absence of number squeezing in the scattering halo is a general feature regardless of binning details.

\[
\begin{align*}
\eta^2 & \quad \text{as a function of time for three different bin volumes (in } k_0^2 \text{ units), as predicted by the reduced Bogoliubov model (RBM) (5.15,5.30).} \\
& \quad \text{The dash-dotted horizontal line denotes the shot-noise level } \eta^2 = 1.
\end{align*}
\]

Figure 5.4: The number squeezing parameter \( \eta^2 \) in the dense \( \gamma = 1.02 \) case as a function of time for three different bin volumes (in \( k_0^2 \) units), as predicted by the reduced Bogoliubov model (RBM) (5.15,5.30). The dash-dotted horizontal line denotes the shot-noise level \( \eta^2 = 1 \).

In the past, it has been conjectured that the quantum correlations in the halo in such experiments are degraded by interactions with the mean field, or due to the time-variation of the source cloud. To identify which process is responsible for such dramatic loss of number squeezing at later times, we first compare the results of the full positive-P method with a maximally reduced Bogoliubov Method (RBM). The evolution of the colliding condensates is simplified to a counter-propagation of the two initial clouds with fixed shape, and additionally, in equations (5.5) we include only the pair production term, so it simplifies to (5.15). Both the mean field self-interaction of the BECs and its impact on the scattered particles are neglected. Free kinetic dispersion is also suppressed, so that the equations of motion for the
condensate field are
\[ i\hbar \partial_t \phi_{\pm}(x, t) = \frac{\hbar^2 k_0}{2m} (\mp 2i \partial_x - k_0) \phi_{\pm}(x, t). \] (5.30)

The condensates do not spread and the scattering process is maximally simplified. Figure 5.4 shows the number squeezing parameter as a function of time as predicted by the RBM. Although \( \eta^2 \) does not grow as strongly as in Fig. 5.2, still the atom-difference fluctuations surpass the shot-noise limit. Therefore, it is \textit{neither} the mean-field interaction nor the spreading of the BECs that have the major impact on the number squeezing parameter.

Next, we simulate the condensate collision using the numerical method which \textit{a priori} excludes the bosonic enhancement, introduced in Sec. (5.2.4), see Eq. (5.9). In Fig. 5.5 we compare the results obtained in the full positive-P simulation and the non-enhanced method. We plot \( \eta^2 \) as a function of time for three different bin volumes. Although the growth of the number squeezing parameter is less violent in the absence of bosonic enhancement, still \( \eta^2 \) is above the shot-noise limit. As expected, for short times \( t \lesssim 0.1 \) the outcomes of the two methods agree very well – the system is still in the spontaneous regime.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig5.5.png}
\caption{The number squeezing parameter \( \eta^2 \) in the dense \( \gamma = 1.02 \) case as a function of time for three different bin volumes given in \( k_0^3 \) units. The solid lines are from a numerical solution of the full model (5.4, 5.5). The dashed lines are predictions of the model without Bose enhancement (5.8,5.9). The dash-dotted horizontal line denotes the shot-noise level \( \eta^2 = 1 \).}
\end{figure}

Finally, in Fig. 5.6 we draw the peak value of the back-to-back correlation function, namely the \( h_{bb} \) defined in Eq. (5.25). As indicated by equations (5.26) and (5.28), the number squeezing is more likely to occur for high \( h_{bb} \), the widths of the correlation functions being the other factor. We see that in all three methods (full positive-P, RBM, and non-enhanced) give \( h_{bb} < h_{cl} \equiv 1 \) at long times, a limit below which (5.29) indicates that small bins can never be number-squeezed.
5.5. Simulation results and analysis

Figure 5.6: The peak height of the back-to-back correlation function $h_{bb}$ as a function of time for the dense case $\gamma = 1.02$. The solid line is from a numerical solution of the full model (5.4, 5.5), the dashed line from the model without Bose enhancement (5.8, 5.9), while the dotted line comes from the reduced Bogoliubov model (RBM) (5.15, 5.30). The dot-dashed line shows the border value of $h_{bb} = 1$ when the back-to-back and collinear peaks are equal, and small-bin number-squeezing disappears.

5.5.2 $\gamma = 0.24$ case

Figure 5.7: The number squeezing parameter $\eta^2$ in the dilute $\gamma = 0.24$ case as a function of time for three different bin volumes given in $k_0^3$ units. The solid lines come from a numerical solution of the full model (5.4, 5.5). The dashed lines are predictions of the Gaussian correlation model (5.26). The dash-dotted horizontal line denotes the shot-noise level $\eta^2 = 1$. 
Here we investigate the number squeezing parameter in the dilute case, when
\( \gamma = 0.24 \) and the bosonic enhancement is absent. First, in Fig. 5.7 we plot \( \eta^2 \) as a
function of time resulting from the positive-P method (5.5) and from the Gaussian
correlation model (5.26). The agreement is satisfactory, although the predictions
of the model are very noisy. This is a result of the small number of scattered
atoms. When the signal is low, it is difficult to extract the widths and peak values
of the correlation functions that enter into the model. Nevertheless, we observe a
major difference between the dense and the dilute case. In the latter, the bins are
number-squeezed, irrespectively of their volume and the time.

We confirm that the system is indeed in the spontaneous regime, by comparing
in Fig. 5.8 the number squeezing parameter as predicted by equations (5.5) and the
non-enhanced method (5.9). We do not observe any major discrepancy between
these two results, and conclude that the system is indeed in the low-gain regime.

![Figure 5.8](image)

Figure 5.8: The number squeezing parameter \( \eta^2 \) in the dilute \( \gamma = 0.24 \) case as a
function of time for three different bin volumes given in \( k_0^3 \) units. The solid lines
result from a numerical solution of the full model (5.4, 5.5). The dashed lines are
predictions of the model without Bose enhancement (5.8,5.9). The dash-dotted
horizontal line denotes the shot-noise level \( \eta^2 = 1 \).

To understand why the number squeezing is present in the \( \gamma = 0.24 \) case, in
Fig. 5.9 we plot the peak of the back-to-back correlation function as a function of
time. We see, that it is far from reaching the border value of \( h_{bb} = h_{cl} \equiv 1 \). Also, we
compare this value with the one predicted by the non-enhanced method (5.9) and
find excellent agreement. The number squeezing parameter depends not only on
the peak values of the correlation functions, but also on their widths. In Fig. 5.10
we compare the widths of the correlation function for \( \gamma = 1.02 \) and \( \gamma = 0.24 \) as a
function of time. We notice that in both situations the back-to-back and co-linear
widths are comparable to each other. In the dense case, the back-to-back widths
are slightly larger than the co-linear, which should favour \( \eta^2 < 1 \) for large bins, as
indicated by Eq. (5.28).
Figure 5.9: The peak height of the back-to-back correlation function $h_{bb}$ as a function of time for the dilute case $\gamma = 0.24$. The solid line is from a numerical solution of the full model (5.4, 5.5), the dashed line from the model without Bose enhancement (5.8, 5.9). The dot-dashed line denotes the border value of $h_{bb} = 1$ when the back-to-back and collinear peaks are equal.

The lack of number squeezing for any bin size for long times of the $\gamma = 1.02$ case, must be therefore attributed to the drop of the peak height $h_{bb}$. This drop is due to the non-monochromatic nature of the parent BECs. Their momentum spread leads to a non-zero width of the back-to-back correlation function, which in turn results in scattering into not exactly opposite momentum modes. The pair of atoms can end up in non-opposite bins. Nevertheless, when the number of scattered atoms is low – as in the dilute gas case or at early times in the dense gas case – there is a high probability of finding a single or a few correlated pairs in the opposite regions, which is related to a high value of $h_{bb}$. Crucially, the probability of having some uncorrelated pairs in the opposite bin is low, because the neighbouring bins are mostly empty. However, when the number of scattered particles grows, the chance that neighbouring bins are empty becomes small, and uncorrelated atoms spill over into the opposite bin. In this way, the $\eta^2$ fluctuations grow, since there is a significant amount of uncorrelated pairs in the opposite bins. Figure 5.1, which is an outcome of a single collision in the dense gas case, is an illustration of this scenario. There are some clearly visible regions, where the atoms form a single large speckle, while on the other side of the halo two distinct speckles are present.

5.5.3 Collision of two plane-waves

To confirm the conjecture formulated in the previous paragraph, we simulate a collision of two monochromatic plane waves. We use the same parameters as in the $\gamma = 1.02$ dense case, but replace the initial BECs with monochromatic peaks in momentum space at $k = (0, 0, \pm k_0)$ (and replace $|\phi|^2$ in the evolution equation (5.4))
Figure 5.10: The widths of the co-linear (left column) and back-to-back (right column) correlation functions for the dense (top row) and dilute case (bottom row) as a function of time. Fitted to the numerical solution of the full model (5.4, 5.5). The solid black line is the width in the axial direction $\sigma_z$, the dotted red line is in the tangential direction to the halo center $\sigma_t$, while the dashed blue line is in the radial direction $\sigma_r$. The dot-dashed violet line is the Gaussian fitted half-width of the halo density $w_r$, which narrows in time due to energy-time uncertainty principle. All widths are given in $k_0$ units.

with the mean density). This mean density is chosen equal to the peak density in the usual $\gamma = 1.02$ case, i.e. $(10/7)\bar{\rho}$. In Fig. 5.11 we plot $\eta^2$ as a function of time for three different bin volumes. In all cases, the number squeezing is near perfect ($\eta^2 \approx 0.03$), despite a huge total number of scattered atoms $\sim 10^7$. The residual slightly nonzero value of $\eta^2$ is induced by the presence of a sea of short-lived particles from virtual scattering events. The RBM, which does not include this non-resonant effect, gives $\eta^2 = 0$ within statistical accuracy.

We also plot the peak height of the back-to-back function as a function of time, see Fig. 5.12. At long times it becomes indistinguishable from the border value of unity, but at this stage the number of atoms in the halo and the bin occupation $\bar{n}$ is very large. However, from the model (5.29) one sees that the minimal, fully squeezed value of $\eta^2 = 0$ corresponds to $h_{bb} = 1 + 2/\bar{n}$, which is extremely close to unity, so that this remains consistent with the observed strong number squeezing in Fig. 5.11.

5.6 Conclusions

This chapter was devoted to a systematic study of the number squeezing parameter between two regions of opposite momenta in the scattering halo formed during colli-
5.6. Conclusions

Figure 5.11: The number squeezing parameter $\eta^2$ in the dense $\gamma = 1.02$ case as a function of time for several different bin volumes given in $k_0^3$ units. Here, the BECs are replaced with the plane-waves. The solid lines come from a numerical solution of the full model (5.4, 5.5), while the dashed green line from a reduced Bogoliubov model (RBM) simulation (5.15, 5.30). The dash-dotted horizontal lines denote the shot-noise and zero levels.

Figure 5.12: The peak height of the back-to-back correlation function $h_{bb}$ as a function of time for the dense case $\gamma = 1.02$ and two colliding plane-waves. The solid line is from a numerical solution of the full model (5.4, 5.5), while dot-dashed line denotes the border value of $h_{bb} = 1$ when the back-to-back and collinear peaks are equal.

We have shown that the number squeezing depends strongly on the bin size, density of the mother clouds, mean-field interactions, and above all on the spectral purity of the mother clouds. In the dilute case, the number squeezing is evident, since the number of scattered pairs is low. There-
fore, once a single atom is detected at momentum $k$, there is a high probability of finding one (and just one) at $k' \approx -k$. Such a setup can be useful for investigation of the foundations of quantum mechanics, and indeed most Bell inequality tests in quantum optics have been carried out with weak sources where the granularity of the boson field becomes visible.

On the other hand, having only a single – or a few – correlated pairs is unattractive from a quantum interferometry point of view. The dense, or “squeezing” [170] regime is preferable. However, we find here that when the amount of scattered atoms is high, $\eta^2$ grows rapidly. By performing additional simulations with plane-waves instead of Gaussian condensates, we have related this behaviour to the non-monochromaticity of the colliding condensates. As the sources are not point-like in momentum space and the back-to-back correlation function has a non-zero range, the scattered pairs are not perfectly aligned. This in turn results in imperfect correlation between opposite bins, which is additionally amplified by the increased number of stray atoms that enter them from neighbouring regions. The non-monochromatic nature of the source clouds appears then to be the effect that underlies most of the degradation of pair correlations and number squeezing in the halo.

Our results have potential application for the setup and analysis of existing and future experiments for the production of correlated atomic pairs from ultracold atom sources, as these are almost always multi-mode, i.e. non-monochromatic to an appreciable degree. Quasicondensate and 1D phase-fluctuation aspects can be tamed by an appropriate choice of the counting bins [27,28,144,171,172]. However, additional broad spontaneous halos are observed in some mode-selective experiments, including directed-beam [27], dressed state [173], and four-wave mixing setups [143,147].

The effect of the width of the mother cloud on short-time behaviour was analysed by Ögren and Kheruntsyan for Bose-Einstein condensates collisions [169] and molecular dissociation [152,153], with improvement of squeezing as the cloud broadens. An important point we demonstrate here is that while the short-time behaviour can be squeezed both at low and high densities, any squeezing is lost if the number of particles grows too far. Thus, as clouds become denser, the onus on achieving or maintaining monochromaticity grows. Conversely, if squeezing is lost for a given geometry, it should be recoverable if the density is reduced sufficiently so that a significant proportion of measurements are free of stray unpaired atoms.

Preferable conditions for number squeezing are satisfied when atoms scatter into well separated regions, because the pool of atoms that go into a broad halo is strongly reduced, and the likelihood of unpaired stray atoms in the measuring bins is lessened. Such conditions have been demonstrated recently in twin beam experiments [28], four-wave mixing experiments with two-component matter waves [143–145], or with an optical lattice that selects only a pair of phase-matched modes [27] by modifying the dispersion relation. Thus, in the next chapter, we develop a theory that describes these kind of systems.
The conclusions reached in the previous chapter stated that atomic collisions can potentially be used as a source of entangled particles, provided that they are scattered into two well separated regions. In the introduction, we presented the experiment with twin-atom beams (see page 15), in which such scattering scenario was procured. However, there is a difficulty in detecting the entanglement and proving whether the system can be useful for interferometry. Therefore, in this chapter, we develop the framework within Bogoliubov theory that describes the system in which atoms are scattered from a Bose-Einstein condensate into two separated regions. We find the full dynamics of the pair-production process, calculate the first and second order correlation functions, and show that the system is ideally number-squeezed. We calculate the Fisher information to show how the entanglement between atoms from the two regions changes in time. We also provide a simple expression for the lower bound on the useful entanglement in the system in terms of the average number of scattered atoms and the number of modes they occupy. We then apply the

\[^{1}\text{The chapter is based on the work: T. Wasak, P. Szańkowski, R. Bücker, J. Chwedeńczuk, M. Trippenbach, New J. Phys. 16, 013041 (2014).}\]
theory to the twin-atom beam experiment [28]. The only numerical step of our semi-analytical description can be easily solved and does not require implementation of any stochastic methods.

6.1 Introduction

In recent years, systems where strong correlations between particles are induced by pair-wise scattering, have attracted much attention. In the canonical example, which is the parametric down-conversion, photon pairs are generated during the propagation of a laser beam through a non-linear medium. The resulting pairs of photons are entangled, and can serve as a probe of fundamental properties of quantum mechanics [97, 133], such as the Einstein-Podolsky-Rosen paradox, or violation of the Bell inequalities [97, 133, 150]. On the other hand, entanglement can be exploited in practical applications, such as teleportation [109, 174] or metrology beyond the Shot-Noise Limit (SNL) [29, 110].

In this latter context, recent experiments with entangled states of atoms were a major breakthrough [18]. In [18, 134, 175], two-body interactions were utilized to prepare non-classical squeezed states of atoms trapped in a double-well potential, which implies presence of many-body entanglement [69]. A similar idea was exploited to generate squeezing in the internal [20–22] degrees of freedom. In [89, 176], squeezing of a large spin of a collection of two-level atoms was achieved, using an intense laser field interacting with particles trapped in an optical cavity.

Simultaneously, a substantial experimental effort was put in order to generate entangled pairs of atoms scattered out of a Bose-Einstein Condensate (BEC). In [25, 26, 142, 146], a collision of two BECs lead to weak scattering of correlated atomic pairs onto a three-dimensional sphere of initially unoccupied modes. Moderate number-squeezing between the opposite regions of the halo, and the related violation of the Cauchy-Schwarz inequality were experimentally demonstrated [25, 26], which, according to recent findings, proves entanglement between identical bosons [177]. Alternatively, pair-production schemes were developed, where only few modes are strongly populated in a stimulated process, making the system somewhat easier to handle. Stimulated four-wave-mixing processes have been implemented using different spin states of atoms [143–145] or Bragg scattering [147, 148]. Also, dynamic instabilities in moving optical lattices, populating modes with opposite quasi-momenta, have been used [27, 178]. In [28], a BEC was transferred into the first excited state of a trapping potential and subsequent two-body collisions created a “twin-beam” system, where stronger-than-classical correlations could directly be observed.

Analogous schemes have been implemented in internal atomic states, building upon spin-changing collisions [24, 179, 180]. Furthermore, in [24] it was shown that particles scattered in this process into a pair of \( m_F = \pm 1 \) Zeeman sub-levels are usefully entangled from the metrological point of view.

Below, we develop a theoretical model for the generic type of experiments, where
6.2 General properties of the scattered particles

6.2.1 Bogoliubov equation for pair scattering

Our theoretical description of the pair-production process starts with a many-body Hamiltonian with contact two-body interactions

\[
\hat{H} = \int dr \left[ \frac{\hbar^2 \nabla^2}{2m} + V(r) \right] \hat{\Psi}(r) + \frac{g}{2} \int dr \hat{\Psi}^\dagger(r) \hat{\Psi}^\dagger(r) \hat{\Psi}(r) \hat{\Psi}(r). \tag{6.1}
\]
Here, \( V(\mathbf{r}) \) is an external trapping potential and \( g = \frac{4\pi\hbar^2 a}{m} \) is the strength of the two-body interactions, \( a \) is the scattering length, \( m \) is the atomic mass and the field operator \( \hat{\Psi}(\mathbf{r}) \) satisfies the bosonic commutation relations. To derive the Bogoliubov equation, we first find the c-number (mean field) wave function of the BEC using the Gross-Pitaevskii Equation (GPE)

\[
\hat{H}_\text{bog} = \frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) + g|\psi(\mathbf{r})|^2 \psi(\mathbf{r}). \tag{6.2}
\]

We then write the field operator as a sum of the c-number part and the Bogoliubov correction, \( \hat{\Psi}(\mathbf{r}) = \psi(\mathbf{r}) + \hat{\delta}(\mathbf{r}) \) and insert this expression into (6.1). By keeping only the terms up to quadratic in \( \hat{\delta} \) we obtain the Bogoliubov Hamiltonian

\[
\hat{H}_\text{bog} = \int d\mathbf{r} \delta^{\dagger}(\mathbf{r}) \left( -\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) + 2g|\psi(\mathbf{r})|^2 \right) \delta(\mathbf{r}) + g \int d\mathbf{r} \left( \delta^{\dagger}(\mathbf{r}) \delta(\mathbf{r}) \psi^2(\mathbf{r}) + \text{h.c.} \right). \tag{6.3}
\]

The resulting Bogoliubov equation of motion is linear

\[
i\hbar \partial_t \delta(\mathbf{r}, t) = \left( -\frac{\hbar^2 \nabla^2}{2m} + V(\mathbf{r}) + 2g|\psi(\mathbf{r})|^2 \right) \delta(\mathbf{r}, t) + g \psi^2(\mathbf{r}) \delta^{\dagger}(\mathbf{r}, t). \tag{6.4}
\]

Usually, a numerical solution of this equation is found in a following way. The field operator is expanded in a basis of wave-functions \( \bar{\varphi}_i(\mathbf{r}) \) which match the geometry of the scattering problem

\[
\hat{\delta}(\mathbf{r}, t) = \sum_i \hat{\varphi}_i(\mathbf{r}) \hat{a}_i(t). \tag{6.5}
\]

This expression is inserted into Eq. (6.4), the resulting equation is multiplied by \( \bar{\varphi}_j^{\ast}(\mathbf{r}) \) and the outcome is integrated by sides over the whole space. In effect, what we obtain is an equation of motion, which, through the matrices \( \hat{A} \) and \( \hat{B} \), couples the evolution of the \( j \)-th operator \( \hat{a}_j(t) \), with (in general) all other operators

\[
i\hbar \partial_t \hat{a}_j(t) = \sum_k A_{jk} \hat{a}_k(t) + \sum_k B_{jk} \hat{a}_k^{\dagger}(t). \tag{6.6}
\]

This equation is linear – a consequence of the linearity of the Bogoliubov equation (6.4) – so the general solution of (6.6) reads

\[
\hat{a}_i(t) = \sum_j C_{ij}(t) \hat{a}_j(0) + \sum_j S_{ij}(t) \hat{a}_j^{\dagger}(0), \tag{6.7}
\]

where the matrices \( \hat{C} \) and \( \hat{S} \) satisfy \( \hat{C} \hat{C}^{\dagger} - \hat{S} \hat{S}^{\dagger} = \mathbb{1} \) and \( \hat{C} \hat{S}^{\dagger} - \hat{S} \hat{C}^{\dagger} = 0 \). Later, we will apply this method to solve the Bogoliubov dynamics of the twin-beam production. However, we will show in the following, that in cases where the detailed form of the Hamiltonian (6.3) drives the scattering of atomic pairs into opposite regions (as indeed happens in twin-beam experiments), the basic properties of the system can be deduced analytically if an appropriate set of mode functions \( \varphi_i(\mathbf{r}) \) is chosen.

Let us denote the two separate regions into which the particles are scattered by L (left) and R (right). Particles populate L and R in a process of elastic scattering,
so the regions are usually separated in momentum space. From this point of view, it is convenient to switch to the space of wave-vectors $\mathbf{k}$ and decompose the field operator as follows

$$\hat{\delta}(\mathbf{k}, t) = \sum_i \varphi_{\mathbf{R}}^{(i)}(\mathbf{k}, t) \hat{a}_R^{(i)}(t) + \sum_i \varphi_{\mathbf{L}}^{(i)}(\mathbf{k}, t) \hat{a}_L^{(i)}(t). \quad (6.8)$$

The operators $\hat{a}_{\mathbf{R}/\mathbf{L}}^{(i)}(t)$ annihilate a particle in a mode characterized by the time-dependent wave function $\varphi_{\mathbf{R}/\mathbf{L}}^{(i)}(\mathbf{k}, t)$, which is localized in the right/left region in momentum space. We underline, that this kind of separation is also present in position space after expansion of the cloud, or after application of a Stern-Gerlach pulse in internal-state experiments, respectively. Moreover, the vector $\mathbf{k}$ might denote the quasi-momentum, if the scattering takes place in an optical lattice.

Formally, the only difference between the formulation (6.5) and (6.8) is the splitting of the field operator into the R and L modes. However, for a linear equation of motion such as (6.4), there exists a unique basis of mode functions for which the evolution equations of the mode pairs decouple from each other:

$$\hat{a}_R^{(i)}(t) = c_i(t) \hat{a}_R^{(i)}(0) + s_i(t) \hat{a}_L^{(i)}(0) \quad (6.9a)$$

$$\hat{a}_L^{(i)}(t) = c_i(t) \hat{a}_L^{(i)}(0) + s_i(t) \hat{a}_R^{(i)}(0), \quad (6.9b)$$

where $|c_i(t)|^2 - |s_i(t)|^2 = 1$. This form of the Bogoliubov equation has a simple physical interpretation: atoms scatter pair-wise into opposite regions, and the total field operator (6.8) is a sum of independent mode pairs, which are squeezed in their relative population fluctuations, as will be explained in detail below.

Although the diagonal form (6.9) is much clearer than (6.7), it is not obvious at the moment how this particular basis (6.8) can be found. This is done in two steps, applying the procedure of the Bloch-Messiah reduction [160, 181, 182]. First, using Equations (6.8) and (6.9), we evaluate the one-body density matrix (first-order correlation function) and obtain

$$G^{(1)}(\mathbf{k}_1, \mathbf{k}_2; t) \equiv \langle \hat{\delta}^{\dagger}(\mathbf{k}_1, t) \hat{\delta}(\mathbf{k}_2, t) \rangle = \sum_i n_i \left( \varphi_{\mathbf{R}}^{(i)}(\mathbf{k}_1, t) \varphi_{\mathbf{R}}^{(i)}(\mathbf{k}_2, t) + \varphi_{\mathbf{L}}^{(i)}(\mathbf{k}_1, t) \varphi_{\mathbf{L}}^{(i)}(\mathbf{k}_2, t) \right), \quad (6.10)$$

where $n_i = |s_i(t)|^2$. Note that $n_i$ is the average occupation of the $i$-th eigen-mode, and that a pair of modes $\varphi_{\mathbf{R}}^{(i)}(\mathbf{k}, t)$ and $\varphi_{\mathbf{L}}^{(i)}(\mathbf{k}, t)$ is degenerate (has the same eigenvalue $n_i$) due to the assumed symmetry between the left and the right region. Since we are using the Heisenberg picture, the average value in Eq. (6.10) and all equations that follow are calculated in the initial vacuum state of scattered atoms.

In any practical approach, if the basis (6.8) is not known a priori, this step first requires a numerical evaluation of the density matrix (6.10) in any convenient basis (6.5), and subsequent diagonalization. Once this is done, then according to Eq. (6.10), the basis functions $\varphi_{\mathbf{R}/\mathbf{L}}^{(i)}(\mathbf{k}, t)$ are the momentary eigen-functions of the one-body density matrix (natural orbitals). However, a second step is necessary to
fully determine the functions \( \varphi_{R/L}^{(i)}(k, t) \), because the density matrix – contrary to the field operator (6.8) – is insensitive to global phases of the mode functions. To retrieve this additional information, we calculate the anomalous density

\[
M(k_1, k_2; t) \equiv \langle \hat{\delta}(k_1, t)\hat{\delta}(k_2, t) \rangle = \sum_i \sqrt{n_i(n_i + 1)} \times \left( \varphi_{R}^{(i)}(k_1, t)\varphi_{L}^{(i)}(k_2, t) + \varphi_{R}^{(i)}(k_2, t)\varphi_{L}^{(i)}(k_1, t) \right),
\]

(6.11)
multiply it by sides with the eigen-functions of the density matrix and integrate over space. As a result, we retrieve the information about the phases and obtain the full form of the mode functions \( \varphi_{R/L}^{(i)}(k, t) \) of the diagonal basis.

To summarize, we have outlined the structure of the solution of the Bogoliubov equation for cases where atoms are scattered into two opposite regions. We will now show that the extra step, which is the transition from the “numerical approach” (6.5) to the diagonal basis (6.8), allows to easily determine the basic properties of the system of scattered atoms, like its density or higher correlation functions.

### 6.2.2 Density and correlations

The simplest observable characterizing the pair-production process is the density

\[
\rho(k; t) \equiv G^{(1)}(k, k; t) = \sum_i n_i \left( \left| \varphi_{R}^{(i)}(k, t) \right|^2 + \left| \varphi_{L}^{(i)}(k, t) \right|^2 \right)
\]

(6.12)

which is, consistently with our derivation, localized in the two opposite regions. By integrating the above function over space, we obtain the information about the expected number of scattered atoms as a function of time

\[
\langle \hat{N} \rangle = \int d\mathbf{k} \rho(\mathbf{k}; t) = 2 \sum_i n_i.
\]

(6.13)

Additional information about the system is carried by the correlations between the scattered particles. The probability of simultaneous detection of two atoms at momenta \( \mathbf{k}_1 \) and \( \mathbf{k}_2 \) can be obtained from the normalized second-order correlation function

\[
g^{(2)}(\mathbf{k}_1, \mathbf{k}_2; t) = \frac{\langle \hat{\delta}^\dagger(\mathbf{k}_1, t)\hat{\delta}^\dagger(\mathbf{k}_2, t)\hat{\delta}(\mathbf{k}_2, t)\hat{\delta}(\mathbf{k}_1, t) \rangle}{\rho(\mathbf{k}_1; t)\rho(\mathbf{k}_2; t)}.
\]

(6.14)

According to the Wick’s theorem, this function can be written in terms of the one-body density matrix (6.10) and the anomalous density (6.11) as follows

\[
g^{(2)}(\mathbf{k}_1, \mathbf{k}_2; t) = 1 + \frac{|G^{(1)}(\mathbf{k}_1, \mathbf{k}_2; t)|^2 + |M(\mathbf{k}_1, \mathbf{k}_2; t)|^2}{\rho(\mathbf{k}_1; t)\rho(\mathbf{k}_2; t)}.
\]

(6.15)

The transition from Eq. (6.14) to (6.15) might seem an unnecessary complication, however we will argue that it allows for a simple and intuitive interpretation of the second-order correlation function. According to Eq. (6.10), the density matrix is non-vanishing only when \( \mathbf{k}_1 \) and \( \mathbf{k}_2 \) are both either in the right or left region, so
6.2. General properties of the scattered particles

$|G^{(1)}|^2$ governs the Hanbury-Brown and Twiss (HBT) type of local correlations. On the other hand, as can be seen from Eq. (6.11), the anomalous density is non-zero only when $k_1$ and $k_2$ are in the opposite regions, so it describes the cross-correlations between the two members of the scattered pair. Clearly, this simple interpretation of the second order correlation function as a sum of local- and opposite-momentum correlations would have been much more difficult if we had not applied the diagonalization procedure and the Wick’s theorem.

6.2.3 Number squeezing

Another property characterizing the scattering process are the fluctuations of the population imbalance between the two regions. If these fluctuations are suppressed below the properly defined shot-noise level, the system is number squeezed, which proves that atoms scatter in pairs rather than independently to the left and to the right region. A quantitative description of the number squeezing involves the left and right atom number operators defined as the integrals of the density operators over the corresponding volumes, i.e.

$$\hat{N}_{R/L} = \int_{R/L} d\mathbf{k} \hat{\delta}^{\dagger}(\mathbf{k},t)\hat{\delta}(\mathbf{k},t).$$  \hspace{1cm} (6.16)

The population imbalance operator is then simply defined as $\hat{n} = \hat{N}_R - \hat{N}_L$ and using Eq. (6.8) we obtain

$$\hat{n} = \sum_i \left( \hat{a}^{(i)\dagger}_R(0)\hat{a}^{(i)}_R(0) - \hat{a}^{(i)\dagger}_L(0)\hat{a}^{(i)}_L(0) \right).$$  \hspace{1cm} (6.17)

The number squeezing factor is defined as

$$\eta^2 = \frac{\Delta^2 \hat{n}}{\langle \hat{N} \rangle},$$  \hspace{1cm} (6.18)

where $\Delta^2 \hat{n} = \langle \hat{n}^2 \rangle - \langle \hat{n} \rangle^2$ is the variance of the population imbalance operator. If the fluctuations between the two regions are suppressed below the shot-noise level defined as $\eta^2 = 1$, the system is called “number-squeezed”. In our case, since $\hat{n}$ does not depend on time and the initial state is a vacuum, we obtain that $\eta^2 \equiv 0$. Therefore, the two-region Bogoliubov system is perfectly number-squeezed, as anticipated in the previous section.

The ideal number-squeezing is a result of clear separation of the two scattering regions. In such case, it is natural to define the local atom-number operators (6.16) and the population imbalance operator (6.17). It is important to note that not all systems, where particles are scattered in pairs are perfectly number squeezed. For instance, when two Bose-Einstein condensates collide, they produce a halo of atoms due to two-body elastic scattering into the initially unoccupied modes [142, 147], which was analysed in chapter 5. In this system however, there is no simple way to
define two separate regions. One can instead measure the number of atoms in two bins lying on the opposite sides of the halo. Moderate number-squeezing of the atom number difference between these bins has been observed experimentally \cite{25}, but it is impossible to reach the limit $\eta^2 = 0$, which was demonstrated in chapter 5. In contrast, the twin matter wave configurations \cite{24, 27, 28, 143, 145}, are ideal sources of correlated atomic pairs occupying two well-defined areas.

### 6.2.4 Violation of the Cauchy-Schwarz inequality

Apart from the number squeezing, the twin-region system can be characterized by another expression, which is called the Cauchy-Schwarz inequality. As we know, it relates the strength of the local and opposite correlations to witness the pair-scattering process (see chapter 4). Following \cite{26}, we define averaged second-order correlations as

$$G^{(2)}_{\mu\nu} \equiv \int_\mu d\mathbf{k}_1 \int_\nu d\mathbf{k}_2 \langle \hat{\delta}^\dagger(\mathbf{k}_1, t) \hat{\delta}(\mathbf{k}_2, t) \hat{\delta}(\mathbf{0}, t) \hat{\delta}(\mathbf{0}, t) \rangle$$

(6.19)

where $\mu, \nu \in \{R, L\}$. In the symmetric case, the Cauchy-Schwarz inequality $G^{(2)}_{RL} \leq (G^{(2)}_{RR} G^{(2)}_{LL})^{1/2}$ can now be re-written as

$$\int_R d\mathbf{k}_1 \int_L d\mathbf{k}_2 |M(\mathbf{k}_1, \mathbf{k}_2; t)|^2 \leq \int_R d\mathbf{k}_1 \int_R d\mathbf{k}_2 |G^{(1)}(\mathbf{k}_1, \mathbf{k}_2; t)|^2.$$  

(6.20)

Using expressions (6.10) and (6.11) we obtain

$$G^{(2)}_{RL} = \left( \sum_i n_i \right)^2 + \sum_i n_i(n_i + 1)$$

(6.21)

$$G^{(2)}_{RR} = \left( \sum_i n_i \right)^2 + \sum_i n_i^2$$

(6.22)

thus the Cauchy-Schwarz inequality reads

$$\sum_i n_i(n_i + 1) \leq \sum_i n_i^2,$$  

(6.23)

which is true only for all $n_i = 0$. As soon as particles start to scatter into the two regions, the Cauchy-Schwarz inequality is clearly violated. The degree of violation is quantified with (see Eq. (4.11) on page 83):

$$C = \frac{G^{(2)}_{RL}}{G^{(2)}_{RR}}.$$  

(6.24)

When $C \leq 1$ the system is in the “classical” regime, while $C > 1$ signify correlations which are stronger than allowed by the classical physics. In our case, this coefficient reads

$$C = 1 + \frac{\sum_i n_i}{(\sum_i n_i^2 + \sum_i n_i^2)}.$$  

(6.25)
Clearly, always \( C > 1 \), because it is a sum of unity and a non-negative part. For high mode populations \( n_i \), the second term, which is inversely proportional to the number of scattered particles tends to zero, restoring the classical limit. Nevertheless, as demonstrated with photons in [183], the confidence by which the Cauchy-Schwarz inequality can be violated in the presence of classical noise still increases with more strongly populated modes.

In chapter 4, we have demonstrated that the violation of the Cauchy-Schwarz inequality is a proof of particle entanglement in all systems of identical bosons. The relation holds for every case, when either the number of particles is fixed or fluctuates from shot to shot, as happens in systems described by the Bogoliubov theory, for as long as coherences between different number states are absent. However in the high-gain regime, it is the Fisher information, which is the quantity more sensitive to particle entanglement then the Cauchy-Schwarz criterion, as we show in the following section. The Fisher information quantifies the potential for sub-shot-noise interferometry, and increases with rising mode population, in spite of the decreasing “granularity” of the matter wave [160] that leads to all second-order correlation functions approaching equal values.

### 6.2.5 Entanglement and interferometry

We now show that atoms occupying the two regions are entangled, and could be used as an input of a quantum interferometer operating below the shot-noise level. We first recall how the precision of the phase estimation is related to the entanglement of input states using as an example the standard two-mode Mach-Zehnder Interferometer (MZI). Then, we extend these concepts to the case, where the interferometer operates between two regions, each having a multi-mode structure determined by the Bogoliubov equations.

When speaking about two-mode interferometers, it is convenient to introduce a set of three operators (see also the two mode example in chapter 2, Eq. (2.141) on page 53):

\[
\begin{align*}
\hat{J}_x &= \frac{1}{2} \left( \hat{a}_R^\dagger \hat{a}_L + \hat{a}_L^\dagger \hat{a}_R \right) \\
\hat{J}_y &= \frac{1}{2i} \left( \hat{a}_R^\dagger \hat{a}_L - \hat{a}_L^\dagger \hat{a}_R \right) \\
\hat{J}_z &= \frac{1}{2} \left( \hat{a}_R^\dagger \hat{a}_R - \hat{a}_L^\dagger \hat{a}_L \right)
\end{align*}
\]

which obey the same commutation relations as the angular momentum operators. The MZI, which is an interferometric device, where the imprint of the phase \( \theta \) onto the input state is preceded and followed by a pair of symmetric beam-splitters, can be represented by a unitary evolution operator \( \hat{U}(\theta) = e^{-i \theta \hat{J}_y} \). If the phase is estimated in a series of \( \nu \gg 1 \) measurements performed on the output state, the precision of the phase estimation is limited by the Cramer-Rao Lower Bound
(CRLB) \cite{36,76} (see Eq. (2.102)),
\[ \Delta \theta \geq \frac{1}{\sqrt{\nu}} \frac{1}{\sqrt{F_Q}}. \] (6.27)

Here, \( F_Q \) is the Quantum Fisher Information (QFI), which is related to the unitary transformation \( \hat{U}(\theta) \). For pure states transformed by the MZI it is equal to \( F_Q = 4 \Delta^2 \hat{J}_y \), where the variance is calculated in the input state of the interferometer \cite{33}. The CRLB states, that if \( \theta \) is determined using any possible type of measurement and estimator, then the precision \( \Delta \theta \) is bounded as in Eq. (6.27).

As we know from chapter 2, apart from providing a lower bound for the error of the phase estimation, the \( F_Q \) is an entanglement measure. Namely, when the input state has an average number of \( \langle \hat{N} \rangle \) particles, then if \( F_Q > \langle \hat{N} \rangle \), the state is particle-entangled \cite{68}.

We now show, that a natural extension of the two-mode picture allows to employ the concept of the QFI as an entanglement measure also in our multi-mode system of interest. To this end, we introduce the following analogue of the two-mode angular momentum operators (6.26),
\[ \hat{J}_x = \frac{1}{2} \int \frac{dk}{R} \left( \hat{\delta}^\dagger (k) \hat{\delta}(-k) + \hat{\delta}^\dagger (-k) \hat{\delta}(k) \right) \] (6.28a)
\[ \hat{J}_y = \frac{1}{2i} \int \frac{dk}{R} \left( \hat{\delta}^\dagger (k) \hat{\delta}(-k) - \hat{\delta}^\dagger (-k) \hat{\delta}(k) \right) \] (6.28b)
\[ \hat{J}_z = \frac{1}{2} \int \frac{dk}{R} \left( \hat{\delta}^\dagger (k) \hat{\delta}(k) - \hat{\delta}^\dagger (-k) \hat{\delta}(-k) \right) , \] (6.28c)

where we dropped the explicit time-dependence of the \( \hat{\delta}(k,t) \) to simplify the notation. Also, for simplicity, we choose the well-separated regions R and L to be localized symmetrically on the opposite sites of \( k = 0 \). The construction of these operators, which satisfy the same commutation relations as (6.26), is based on the analogy between the two-mode systems and the twin-beam configuration. In the former case, the operators connect the right and left modes, while in the latter the left and right sub-spaces. Such a definition (6.28) is meaningful only in situations, where the system consists of two well-separated sub-systems.

Using the decomposition of the field operator into the set of independent modes, Equations (6.8) and (6.9), the above integrals yield, that each angular momentum operator is a sum of operators acting on each mode independently, that is
\[ \hat{J}_x = \sum_i \frac{1}{2} \left( \hat{a}^\dagger_R (i) \hat{a}_L (i) + \hat{a}^\dagger_L (i) \hat{a}_R (i) \right) \equiv \sum_i \hat{J}_x^{(i)} \] (6.29a)
\[ \hat{J}_y = \sum_i \frac{1}{2i} \left( \hat{a}^\dagger_R (i) \hat{a}_L (i) - \hat{a}^\dagger_L (i) \hat{a}_R (i) \right) \equiv \sum_i \hat{J}_y^{(i)} \] (6.29b)
\[ \hat{J}_z = \sum_i \frac{1}{2} \left( \hat{a}^\dagger_R (i) \hat{a}_L (i) - \hat{a}^\dagger_L (i) \hat{a}_R (i) \right) \equiv \sum_i \hat{J}_z^{(i)} . \] (6.29c)
These expressions show again that it is natural to describe the two-region system using the diagonal basis (6.8). In this language, the angular momentum operators are simply a sum of operators acting on each pair of modes independently, which vastly simplifies the further analysis.

To establish a direct relation between the two-mode and two-region case, we now assume that the system is transformed in the multi-mode analogue of the Mach-Zehnder interferometer. As outlined above, to demonstrate the presence of useful entanglement between atoms in the left and in the right, it is necessary to calculate the QFI. Using Eq. (6.29a) we obtain that

\[ F_Q = 4 \Delta^2 \hat{J}_y = 4 \left\langle \left( \sum_i \hat{J}_y^{(i)} \right)^2 \right\rangle = 4 \sum_i \left\langle \left( \hat{J}_y^{(i)} \right)^2 \right\rangle + 4 \sum_{i \neq j} \left\langle \hat{J}_y^{(i)} \hat{J}_y^{(j)} \right\rangle. \]  

(6.30)

Since the construction of the basis (6.8) explicitly assumes that each mode is independent from all others, the second term in the last equality is

\[ 4 \sum_{i \neq j} \left\langle \hat{J}_y^{(i)} \hat{J}_y^{(j)} \right\rangle = 0, \]

because the symmetry between the \( R \) and \( L \) regions implies that \( \left\langle \hat{J}_y^{(i)} \right\rangle = 0 \) for all \( i \). Therefore we obtain that the QFI is equal to

\[ F_Q = 4 \sum_i \left\langle \left( \hat{J}_y^{(i)} \right)^2 \right\rangle = 4 \sum_i n_i^2 + 2 \left\langle \hat{N} \right\rangle, \]  

(6.31)

where the last equality comes directly from the substitution of (6.9) into the definition of the \( \hat{J}_y^{(i)} \) operator. Also, we used \( \left\langle \hat{N} \right\rangle = 2 \sum_i n_i \), according to Eq. (6.13). Clearly \( F_Q > \left\langle \hat{N} \right\rangle \), so the system is entangled. Moreover, one can refer the QFI to the ultimate bound for the precision of the parameter estimation, which is the Heisenberg limit. For a system with fluctuating number of particles, this upper bound is equal to \( \left\langle \hat{N}^2 \right\rangle \). Using (6.9) again, we obtain that [68]

\[ \left\langle \hat{N}^2 \right\rangle = \left\langle \left( \sum_i \left( \hat{a}_R^{(i)^\dagger} \hat{a}_R^{(i)} + \hat{a}_L^{(i)^\dagger} \hat{a}_L^{(i)} \right) \right)^2 \right\rangle = 8 \sum_i n_i^2 + \left\langle \hat{N} \right\rangle. \]  

(6.32)

For a large number of scattered particles, when \( \left\langle \hat{N} \right\rangle \ll \sum_i n_i^2 \), we obtain \( F_Q \simeq \frac{1}{2} \left\langle \hat{N}^2 \right\rangle \). The value of the QFI, which is only one-half smaller than the Heisenberg Limit is a clear indication of very strong entanglement present in the system in the high-gain regime. At intermediate times, \( F_Q < \frac{1}{2} \left\langle \hat{N}^2 \right\rangle \) due to mode competition, which has a negative impact on the entanglement as witnessed by the QFI. To picture this, consider a “frustrated case”, where all atoms scatter uniformly into \( M \) pairs of modes, so that all \( n_i \equiv n \) are equal. In this case, the number of scattered atoms is simply \( \left\langle \hat{N} \right\rangle = 2nM \), and the QFI is \( F_Q = 4n^2M + 2 \left\langle \hat{N} \right\rangle \). The QFI normalized to the SNL is

\[ \frac{F_Q}{\left\langle \hat{N} \right\rangle} = 2 + \frac{\left\langle \hat{N} \right\rangle}{M}. \]  

(6.33)
When, on average, there is less than a particle per a set of modes, i.e. \( \langle \hat{N} \rangle_M \ll 1 \), the QFI surpasses the SNL only by a factor of 2, a natural reminiscence of atoms being scattered in pairs. Equation (6.33) is a simple yet intuitive estimation of the lower bound of useful entanglement in terms of the number of scattered atoms and occupied modes.

Although usually it is impossible to directly measure the QFI, one can employ some other estimation scheme, which gives \( \Delta \theta \) close to the CRLB from Eq. (6.27). For instance, in a two-region setup as in [24], the entanglement was detected through the phase sensitivity in a two-step protocol. First, the particles from the two modes were mixed using a radio-frequency pulse, which can be represented by the unitary operation

\[
\hat{U}_{\text{mix}}(\theta) = e^{-i\theta \hat{J}_x},
\]

(6.34)

where the \( \hat{J}_x \) operator is defined in Eq. (6.28a). Then, the population imbalance between the two regions was measured for different values of \( \theta \) and the phase sensitivity was estimated using the error propagation formula basing on the second moment of the population imbalance operator. Introducing the operator \( \hat{A} = \hat{J}_z \) the relation reads

\[
\Delta^2 \theta = \frac{1}{m} \left( \frac{\langle (\Delta \hat{A})^2 \rangle}{\partial(\langle \hat{A} \rangle / \partial \theta)} \right)^2.
\]

(6.35)

The drop of \( \Delta^2 \theta \) below the shot-noise level in the experiment of [24] signalled entanglement between the scattered particles. A similar method could be employed for the twin-beam experiment [28], where the particle mixing operation should be implemented using for instance an optical Bragg pulse rather than a radio-frequency field.

The problem of finding the optimal measurement, which saturates the CRLB (6.27), is one of the central issues of quantum estimation theory. For the twin-beam setup and in absence of decohering influence of the environment, the estimation scheme presented above is optimal for small mixing angles. However, as soon as the noise starts to affect the system, one must seek for another phase determination protocols. Nevertheless, usually it is very difficult to find a truly optimal scheme, which would be easy to implement in the laboratory.

### 6.3 Application: twin-beam system

We now apply the above formalism to the twin-beam system of [28], that we shortly presented in the introduction (see page 15). First, we describe the physical mechanism which leads to the creation of the two correlated beams. As shown below, some basic information about the dynamics of the pair production allow to construct a simple one-dimensional Bogoliubov model, which can be easily solved numerically.
6.3. Application: twin-beam system

6.3.1 Scheme of the experiment

We begin by describing the procedure employed in [28] to produce correlated atom pairs. First, an almost pure Bose-Einstein Condensate (BEC) of \( N_0 \approx 800 \) \(^{87}\)Rb atoms with scattering length equal to \( a = 5.3 \) nm was created at temperature \( T \approx 25 \) nK. The cloud was trapped in a slightly anharmonic potential, which for our purposes can well be approximated by

\[
V(r) \approx \frac{1}{2}m\omega_x^2x^2 + \frac{1}{2}m\omega_y^2y^2 + \frac{1}{2}m\omega_z^2z^2,
\] (6.36)

where atomic mass is equal to \( m = 1.44 \times 10^{-25} \) kg, and the frequency \( \omega_x = 2\pi \times 16.3 \) Hz is much smaller than \( \omega_y = 2\pi \times 1.83 \) kHz and \( \omega_z = 2\pi \times 2.50 \) kHz, so the BEC is strongly elongated along the \( x \)-axis.

After the BEC was created, the trapping potential was shaken in a controlled way, so the atoms were transferred to the first excited state along the \( y \)-direction. In order to achieve the maximal transfer efficiency, the shaking was optimized using quantum optimal control theory [184]. Afterwards, binary collisions transferred particle pairs to the ground state of the potential, and the excess energy \( 2\hbar \omega_y \) was converted into back-to-back movement of the two atoms along \( x \). Momentum conservation ensured, that their wave vectors had equal lengths \( k_0 \approx \sqrt{2m\omega_y}/\hbar \) and point in opposite directions. Small corrections to the value of \( k_0 \) may arise from an effective mean-field potential, as will be discussed below.

6.3.2 Theoretical description

Neglecting thermal phase fluctuations along the elongated direction \( x \), which is valid at very low temperatures only [185], the condensate wave function acting as a source for the pair-production can be found by solving the stationary GPE

\[
\mu \psi(r) = \left(-\frac{\hbar^2\nabla^2}{2m} + V(r) + g|\psi(r)|^2\right)\psi(r),
\] (6.37)

where \( \mu \) is the chemical potential. This function can be evaluated numerically, by referring to the description of the experiment from the previous section, and noting that after the shaking of the trap, the BEC is in the first excited state \( n_y = 1 \) along the \( y \) axis and in the ground state \( n_z = 0 \) along \( z \). However, this can be approximated by an analytical expression, as argued below.

First note, that since the characteristic energies \( \hbar \omega_y \) and \( \hbar \omega_z \) are large, and the number of atoms in the BEC is small, the non-linear term can be safely neglected in evaluation of the eigenstates along \( y \) and \( z \). As a result, assuming that the total wave-function \( \psi(r) \) separates in three directions (which has been confirmed numerically), we obtain

\[
\psi(r) = \phi(x) \times \psi_{n_y=1}^{(ho)}(y) \times \psi_{n_z=0}^{(ho)}(z),
\] (6.38)
where the functions $\psi_{n_y=1}(y)$ and $\psi_{n_z=0}(z)$ are the eigen-states of the one-dimensional harmonic potential in $y$ and $z$ correspondingly. The function $\phi(x)$ is found by inserting the above expression into Eq. (6.37) and integrating out the orthogonal directions. As a result, we obtain an effective equation

$$(\mu - \varepsilon_\perp) \phi(x) = \left( -\frac{\hbar^2}{2m} \frac{\partial^2}{\partial x^2} + \tilde{g} |\phi(x)|^2 + \frac{1}{2} m \omega_x^2 x^2 \right) \phi(x), \quad (6.39)$$

where zero-point energy equals $\varepsilon_\perp = \frac{3}{2} \omega_y + \frac{1}{2} \omega_z$ and the non-linearity reads

$$\tilde{g} = g \left[ \int dy \left( \psi_{n_y=1}(y) \right)^4 \right] \times \left[ \int dz \left( \psi_{n_z=0}(z) \right)^4 \right] = \frac{3}{8 \pi} \frac{g}{a_{ho,y} a_{ho,z}}. \quad (6.40)$$

Here $a_{ho,i} = \sqrt{\frac{\hbar}{m \omega_i}}$ are the harmonic oscillator lengths for $i = y, z$. Since the trap is shallow in the $x$ direction, the solution of the stationary GPE (6.39) can be well approximated by the Thomas-Fermi (TF) formula [186]

$$\phi(x) = \sqrt{\frac{\tilde{\mu}}{\tilde{g}}} \sqrt{1 - \frac{x^2}{R_{TF}^2}}, \quad (6.41)$$

where the effective chemical potential is

$$\tilde{\mu} = \mu - \varepsilon_\perp = \left( \frac{3\tilde{g} N_0 \sqrt{m \omega_x^2}}{4\sqrt{2}} \right)^{2/3} = 492 \text{ Hz} \times h, \quad (6.42)$$

leading to a TF radius of $R_{TF} = \sqrt{\frac{2\tilde{\mu}}{m \omega_x^2}} = 20.75 \mu m$.

Within the approximation of neglecting thermal phase fluctuations, we have fully determined the wave-function of the BEC, which we insert into the Bogoliubov Hamiltonian (6.3). Next, we expand the field operator $\hat{\delta}(r)$ in an orthonormal basis. Along the $y$ and $z$ directions, it is natural to use the eigen-states of the harmonic oscillator as the basis functions, since it matches the geometry of the source BEC. Along the $x$ direction, we use a plane-wave basis, and get

$$\hat{\delta}(r,t) = \sum_{n_y,n_z} \int \frac{dk}{2\pi} e^{ikx} \psi_{n_y}^{(ho)}(y) \psi_{n_z}^{(ho)}(z) \hat{\delta}(k,n_y,n_z,t). \quad (6.43)$$

Since the atom pairs are emitted into the ground state along $y$ only (which is ensured by the anisotropy and anharmonicity of the potential), the sum over the eigen-states can be safely truncated at $n_y = 0$ and $n_z = 0$. This reduces the dynamics of the pair-production to one-dimensional problem along the $x$ axis, with the orthogonal directions frozen out, i.e.

$$\hat{\delta}(r,t) \simeq \psi_{n_y=0}^{(ho)}(y) \psi_{n_z=0}^{(ho)}(z) \sum_k \frac{e^{ikx}}{\sqrt{L}} \hat{a}_k(t), \quad (6.44)$$
where $L$ is the quantization volume. We insert this field operator into Eq. (6.3), evaluate the spatial integrals and upon the change of variables $\hat{a}_k(t)e^{i\hat{\mu}t} \rightarrow \hat{a}_k(t)$ obtain

$$\hat{H}_{\text{bog}} \simeq \sum_k \left( \frac{\hbar^2 k^2}{2m} - (\hat{\mu} + \hbar\omega_y) \right) \hat{a}^\dagger_k(t)\hat{a}_k(t) + \sum_{k,k'} \left( 2f_{k-k'} \hat{a}^\dagger_k(t)\hat{a}_{k'}(t) + f_{k+k'} \left( \hat{a}^\dagger_k(t)\hat{a}^\dagger_{k'}(t) + h.c. \right) \right),$$

where $f(q) = \frac{2}{\pi} \int dx e^{-iqx} \phi^2(x)$. We solve the resulting Bogoliubov equation numerically, by taking the exponent of the evolution matrix, and find the matrices $\hat{C}$ and $\hat{S}$ as defined in Eq. (6.7).

Using the above Hamiltonian, one can also analytically determine $k_0$, i.e. the position of the central peak. To this end, we employ a two-mode approximation by replacing the function $f(q)$ with a Dirac delta, and obtain the Bogoliubov equation

$$i\hbar \partial_t \hat{\delta}_k(t) = \frac{\hbar^2}{2m} \left( k^2 - k_0^2 \right) \hat{\delta}_k(t) + \frac{2}{3} \hat{\mu} \hat{\delta}^\dagger_{-k}(t),$$

where $k_0$ is shifted with respect to the harmonic excitation energy due to the mean-field repulsion and reads

$$k_0 = \sqrt{\frac{2m}{\hbar^2} \left( \hbar \omega_y - \frac{\hat{\mu}}{3} \right)} = 5.35 \, \mu\text{m}^{-1}.$$ (6.47)

This result is in good agreement with the experimentally measured position of the peak density, i.e. $k_{0,\exp} = 5.55(5) \, \mu\text{m}^{-1}$, where the error arises mainly from the mean-field fluctuations.

### 6.3.3 Numerical results

In this section, we display the most important characteristics of the twin-beam system, starting from the solution of the eigen-problem of the density matrix (6.10). In Fig. 6.1 we plot the first four eigen-values of the density matrix, as a function of time. The inset shows the total number of scattered atoms $\langle \hat{N} \rangle$ normalized to the occupation of the BEC, as a function of time. The Bogoliubov approximation is valid for as long as $\langle \hat{N} \rangle \ll N_0$, so we interrupt the simulation at $t = 1.2\, \text{ms}$, when $\langle \hat{N} \rangle \approx 15\%N_0$. For longer times, when the depletion of the BEC cannot be neglected, an atom-number conserving method, such as the one introduced in [187] must be used.

In Fig. 6.2 we plot the first four eigen-vectors of $G^{(1)}$ localized in the right half-space, i.e. $|\varphi^{(1)}_{\nu}(k)|^2$ with $\nu = 1, 2, 3, 4$, calculated at an early time $t = 0.1\, \text{ms}$ and at $t = 1.2\, \text{ms}$. Due to the time-energy uncertainty relation, the eigen-modes localize around $k = k_0$ at later times.

This can be seen even more clearly, by plotting the density $\rho(k; t)$ at these two instants, as shown in Fig. 6.3 (dashed lines). At $t = 0.1\, \text{ms}$, two broad beams start
to form on top of the uniform density. Later, at $t = 1.2 \text{ ms}$, strongly localized peaks clearly dominate over the flat background. On top of these curves, we plot the normalized second-order correlation function as defined in Eq. (6.14), with one of the arguments set equal to the resonant wave-vector $k_0$, i.e. $g^{(2)}(k_1, k_2 \equiv k_0; t)$. At $t = 0.1 \text{ ms}$, the cross-correlation, which is governed by the anomalous density, is very large, i.e. $g^{(2)}(-k_0, k_0; 0.1 \text{ ms}) \approx 40$. This is a characteristic property of the Bogoliubov system in the low-occupation regime [188], and indicates strong violation of the Cauchy-Schwarz inequality (6.20). Also, for this early time, the width of both $g^{(2)}$ peaks are much more narrow than the beam size. This is consistent with the results shown in Fig. 6.1, where at early times many eigen-mode pairs of the density matrix are almost equally occupied. At later times, when a single pair of modes start to become dominant, the width of the peak in $g^{(2)}$ and the system size approach each other. While this corresponds to beams that are single-mode with respect to their local one-body properties, the local averaged correlation function as introduced in Eq. (6.19) reaches the limit of $G^{(2)}_{\mu\mu} \approx 2(\sum_i n_i)^2 \equiv \frac{1}{2} \langle \hat{N} \rangle^2$, exceeding the number fluctuations of a coherent state by a factor of two.

The full correlation function $G^{(2)}(k_1, k_2)$ at $t = 1.2 \text{ ms}$, shown in Fig. 6.4, reveals both the local and cross correlations. In typical experiments, however, such a function is difficult to measure, and one uses collinearly integrated functions of the
6.3. Application: twin-beam system

Figure 6.2: The modulus square of the first four eigen-vectors localized in the right sub-space, i.e. $|\varphi_R^{(i)}(k)|^2$ with $i = 1, 2, 3, 4$. The solid black lines are results of diagonalization of the density-matrix at $t = 0.1\,\text{ms}$ while the dashed red lines at $t = 1.2\,\text{ms}$. The figure shows how due to the time-energy uncertainty relation, the eigen-vectors narrow in the course of time around the central wave-vector $k_0$.

The modulus square of the first four eigen-vectors is shown in the figure. The solid black lines represent the results at $t = 0.1\,\text{ms}$, while the dashed red lines show the results at $t = 1.2\,\text{ms}$. This demonstrates how the eigen-vectors narrow around the central wave-vector $k_0$ due to the time-energy uncertainty relation.

The type

$$g_{cl}^{(2)}(\delta k; t) = \frac{\int G^{(2)}(k, k + \delta k; t)dk}{\int \rho(k; t)\rho(k + \delta k; t)dk} \quad (6.48)$$

where the integrals run over an appropriately chosen momentum region [26, 142]. For symmetric, non-local correlations, back-to-back integration of the type

$$g_{bb}^{(2)}(k_+; t) = \frac{\int G^{(2)}(k, k_+ - k; t)dk}{\int \rho(k; t)\rho(k_+ - k; t)dk} \quad (6.49)$$

is used. The corresponding normalized functions for our system $g_{cl}^{(2)}(\delta k; t), g_{bb}^{(2)}(\delta k; t)$ are shown in Fig. 6.5 at $t = 1.2\,\text{ms}$. For both local (solid) and non-local (dotted) functions, correlations peaks, which do not span the entire populated range (grey area) and follow a Gaussian shape, are clearly present.

In the next step we take towards future comparison with experiments, we present the results not in momentum space, but rather using real-space data calculated after
some finite time $\tau$ of ballistic expansion. Only in the limit of $\tau \to \infty$ (far field), the real-space data is equivalent to the initial momentum space distribution (if the expanding clouds are sufficiently dilute, so that the mean-field repulsion can be safely neglected). In [28], the expansion time was $\tau = 46$ ms, which was sufficient to resolve the twin-beam peaks. Nevertheless, the system was not fully in the far-field regime yet, which may have some impact on the correlation functions. As shown in Fig. 6.5, the finite expansion time affects the back-to-back peak at $(k_0, -k_0)$ much more strongly than the collinear HBT peak, leading to smearing of the measured $g^{(2)}_{b+b}(k_+; t)$ (dash-dotted line) over the entire size of the twin-beam packets. This observation is consistent with some previous numerical results [189].

The different behaviour of these two types of correlations in the near field can be explained using following semi-classical arguments. First, consider a single body function $G^{(1)}(k_1, k_2)$, which according to Eq. (6.15) determines the local second-order correlations. The time, when the system enters the far-field regime, corresponds to an instant, when the measured density in position space samples the momentum distribution. If a particle is propagating freely, starting from $x_0$, for a time $\tau$ and having velocity $v$, then it is registered at $x_0 + v\tau$, which recovers the initial momentum if $x_0$ can be neglected when compared to the characteristic velocity spread $\sigma_v\tau$. From the width of momentum distribution shown in Fig. 6.3, we obtain
6.3. Application: twin-beam system

that $\sigma_v \approx 290 \, \mu m/s$ while the position density profile gives $x_0 \approx 6.3 \, \mu m$. Therefore, $\sigma_v \tau / x_0 \approx 2$, which suggests that $G^{(1)}$ after $\tau = 46$ ms of ballistic expansion is approximately in the far-field regime.

On the other hand, the anomalous density $M(k_1, k_2)$ is a two-body function, which governs the opposite correlations between a pair of atoms. If two atoms are scattered from the source at position $\hat{x}$ with velocities $v_1$ and $v_2$, then after the ballistic expansions time $\tau$ their total coordinate is $x_1 + x_2 = (v_1 + v_2)\tau + 2\hat{x}$. In the far-field regime, the position measurement samples momentum correlations meaning that the first term $v_+ = v_1 + v_2$ dominates over $2\hat{x}/\tau$. The characteristic range of $v_+$ can be estimated from Fig. 6.5 giving $v_+ \approx 146 \, \mu m/s$. The range of allowed positions of the collisions is given by the size of the mother cloud, thus $\hat{x} \approx R_{1f} \approx 21 \mu m$. Therefore, we obtain that $v_+ \tau / 2\hat{x} \approx 0.16$, which shows that the anomalous density after $\tau = 46$ ms – and in consequence the opposite momentum correlations – is not in the far-field regime. We conclude that the broadening of the opposite correlations results mainly from the spread in the positions of the scattering events.
Chapter 6. Bogoliubov theory for atom scattering

Note that although at every instant of the evolution, the field operator $\hat{\delta}(k, t)$ can be written as a sum of independently squeezed modes, at very early times the division between the right and left modes is unjustified, because the two peaks are not yet fully separated. However, at $t = 0.1 \text{ ms}$ when the density distribution is broad, the number of scattered atoms is $\langle \hat{N} \rangle \approx 2$. Therefore, the system at such early time is hardly accessible experimentally so the quantum state of much less then a single particle is not of interest. As soon as the two peaks are well-formed, at $t \approx 0.3 \text{ ms}$, with $\langle \hat{N} \rangle \approx 8$ scattered atoms, all the general considerations from Sec. 6.2 apply.

Finally, in Fig. 6.6 we plot the QFI from Eq. (6.31) as a function of time and normalized to the Heisenberg Limit, i.e. $F_Q/\langle \hat{N}^2 \rangle$. Instead of interrupting the simulation at 1.2 ms, where the scattered fraction of atoms becomes non-negligible and particle number conservation is strongly violated, we extend the calculation up to 7 ms, when the number of scattered atoms significantly exceeds 15% of $N_0$. This is done solely to illustrate that, once the population of one of the modes dominates, $F_Q \rightarrow \frac{1}{2} \langle \hat{N}^2 \rangle$, as argued in Sec. 6.2.5. Note that the dominance of a single mode pair at long times is also predicted by the number-conserving theory [187], justifying this proceeding. Indeed, in the inset, we show the number of pairs of right/left modes which have an occupation bigger or equal to 10% of the largest mode. This approximately tells, how many modes are significantly occupied in the system. At early times, there are over 100 pairs of modes. At 1.2 ms there are still 5 significantly occupied pairs, and only around 4.2 ms a single pair of modes starts to dominate.
At the same time the QFI approaches its upper bound.

![Graph](image)

Figure 6.6: The Quantum Fisher Information as a function of time, normalized to $\langle \hat{N}^2 \rangle$. The horizontal gray dashed line denotes the best possible value for the Bogoliubov system, which is achieved in a regime, where only a single pair of left/right modes is relevant. The vertical gray dashed line denotes the time $t = 1.2$ ms, when the Bogoliubov simulation should be interrupted. In the inset, we show the number of pairs of left/right modes that have at least 10% of occupation of the largest modes. We see that only around 4.2 ms, the two-mode approximation is valid, as denoted by the horizontal dashed line.

We wish to present some remarks about validity of the Bogoliubov approximation. Note that the approach neglects the secondary collisions between the scattered particles and the atoms from the source cloud. When a scattered atom propagates through the BEC, the number of such events that are experienced by each particle can be estimated as follows. First, a cross-section for the collision of two indistinguishable bosons is related to the scattering length by the formula $\sigma = \frac{8\pi a^2}{\sqrt{\pi}}$. The mean free path is not smaller than $l_{\text{free}} = \frac{1}{n\sigma}$, where $n$ is the peak density of the BEC. The number of times a scattered particle, which is travelling through a BEC, collides with the atoms from the cloud is given by the ratio of the size of the condensate, which is equal to $2R_{\text{tf}}$, to the mean free path. Therefore, we obtain that the number of secondary collisions is equal to $N_{\text{col}} = 2R_{\text{tf}}/l_{\text{free}} = 16\pi a^2 n R_{\text{tf}}$. By plugging in the experimental numbers, we obtain $N_{\text{col}} = 0.38$, which well justifies the use of the Bogoliubov approximation.
6.4 Concluding remarks

This chapter was devoted to a development of a simple Bogoliubov model describing twin-atom beam experiments similar to Refs. [27,28]. Due to the restricted geometry of the trapping potential, the dynamics is one-dimensional. As a consequence, the final step of our method can be easily solved numerically without the need for any stochastic method. Furthermore, basic information about the scattered particles can directly be drawn from the general properties of the solution of the Bogoliubov equations. In this way, we can quantitatively characterize the mode structure and correlation functions of the scattered atoms. Also, quite generally, we can show that the population imbalance between the two beams is ideally squeezed and that the system is strongly entangled. These general observations can be applied to most recent experiments, where the atomic pairs are scattered into two well-separated regions. Finally, using the notion of the Quantum Fisher Information, we have derived a simple lower bound for the useful entanglement of the system. This expression employs only the average number of scattered particles and the number of occupied modes.

Having understood the fundamental properties of few-mode twin beams, further steps can be made to take into account more specific issues of experimental implementations. A general feature of strongly elongated Bose gases at realistic temperatures, such as the source cloud in [28], is the quasi-condensation [185], where the coherence along the $x$-axis is limited due to thermal phase fluctuations. Although these fluctuations do not alter the general considerations of section 6.2, they affect the emission dynamics [187], and also might have influence on the spatial properties of both the density and the correlation functions [189]. Also, strong depletion of the source state in the experiment [28] precludes a quantitative comparison with the Bogoliubov theory.

In future, the presented method could be applied to the analysis of some more complicated schemes, where atoms are scattered into two separated regions, such as the Rarity-Tapster-type experiments [190]. Also, according to our results, the two-region state could be used as an input to the Mach-Zehnder-like interferometer, similarly to the experiment with twin matter waves [24].

The fundamental reason why pairs of scattered atoms are useful for sub-shot-noise interferometry is that the constituting particles are indistinguishable. This observation raises a reflection that the scattering process itself is the mean to obtain a system in which of major importance are the correlations originating only from the indistinguishability. In the next chapter, we address the issue whether the extreme scenario of two independently prepared Bose-Einstein condensates can be employed for quantum-enhanced metrology.
Chapter 7

Interferometry with independently prepared Bose-Einstein condensates

Contents

7.1 Introduction .............................................. 132
7.2 Derivation of the ultimate bound ......................... 133
7.3 Impact of disorder ....................................... 135
7.4 Precision with the measurement of populations ............ 136
7.5 Summary .................................................. 138

In the previous chapter we developed a general theory that describes systems in which entangled particles are scattered into two separated regions. The non-classical correlations in such systems turn out to be useful for ultra-precise interferometry. The scattering in these systems are used to obtain two groups of atoms that are correlated due to the indistinguishability of particles. This observation leads to a reflection that two independently prepared Bose-Einstein condensates are also non-classically correlated. Therefore, in this chapter\(^1\), we tackle the problem whether such a system can be useful for metrology. We show that quantum correlations arising from the indistinguishability of bosons are a sufficient resource for the sub-shot-noise interferometry. To this end, we consider an interferometer that operates on two independently prepared Bose-Einstein condensates with fluctuating numbers of particles. We calculate the sensitivity obtained from the measurement of the number of atoms and compare it with the ultimate achievable bound. Our main conclusion is that even in the presence of major atom number fluctuations, an interferometer operating on two independently prepared condensates can give very high precision. These observations indicate a new possibility for an interferometer operating below the shot-noise limit.

Chapter 7. Interferometry with independent BECs

7.1 Introduction

The preparation of probes of usefully entangled particles [72,191] has been the center of attention in the field of quantum interferometry in recent years. In the case of optical interferometers, the entanglement between the photons is most commonly generated in the process of parametric down-conversion [192, 193]. The correlated pairs of photons are a result of a non-linear interaction of the incident laser pulse with a crystal. For atomic interferometers, the useful particle entanglement has been achieved by means of two-body interactions present in ultra-cold systems. Usually, such correlations are associated with the spin-squeezing of a two-mode sample [19, 20, 22, 194–199]. Alternatively, in a process which resembles the down-conversion, the interactions drive the scattering of pairs of entangled atoms from a coherent source, such as a Bose-Einstein condensate (BEC) [24, 26–28, 200].

In all these cases, the correlated many-body state is prepared in a dedicated procedure. Below we show that the particle entanglement that arises solely from the indistinguishability of bosons, which is one of the fundamental laws of quantum mechanics, might be a resource for the sub shot-noise (SSN) interferometry. Our inquiry is supported by an effect [3], which has been observed in the experiment [13]. When a pair of bosons is put into two modes $a$ and $b$, in the second quantization the state reads

$$|\psi\rangle = |1\rangle_a |1\rangle_b , \quad (7.1)$$

which implies there is no entanglement between the modes. However, from the particle point of view, the wave function of the first ($1^{\text{st}}$) and the second ($2^{\text{nd}}$) boson will read $|\psi\rangle = \frac{1}{\sqrt{2}} (|1^{\text{st}}\rangle_a \otimes |2^{\text{nd}}\rangle_b + |1^{\text{st}}\rangle_b \otimes |2^{\text{nd}}\rangle_a)$. This is a state, where the entangled parties are the particles and the non-classical correlation is a result only from the indistinguishability. Naturally, when the two modes are spatially separated, it is impossible to observe this entanglement. The situation changes when the two particles simultaneously pass through a beam-splitter. Then, the bosonic statistics comes into play and at the output is a NOON state, which in the second quantization reads

$$|\psi\rangle = \frac{1}{\sqrt{2}} (|2\rangle_a |0\rangle_b + |0\rangle_a |2\rangle_b) . \quad (7.2)$$

The beam-splitter extracts the entanglement between the modes from the initial particle entanglement due to the indistinguishability [201]. On the other hand, it acts on each particle independently, so it does not change the amount of particle entanglement. This means that a twin-Fock (TF) state (7.1) is just as particle entangled as the NOON state (7.2). The presence of strong entanglement between the modes makes the NOON state an ideal candidate for SSN metrology if $a$ and $b$ are identified with the two arms of an interferometer.

This example shows that (at least for two particles) particle entanglement due to the indistinguishability is a viable resource for quantum interferometry. The ultimate bound of the sensitivity is the Heisenberg limit (HL): $\Delta \theta \propto \frac{1}{N}$, which is $\sqrt{N}$ better then the shot-noise limit (SNL), $\Delta \theta \propto \frac{1}{\sqrt{N}}$ ($N$ is the number of particles).
Therefore, the gain from quantum correlations is of growing importance when \( N \) increases. These observations motivate our inquiry: under which conditions are two groups of bosons, which come from two independent sources and entangled solely due to indistinguishability, useful for metrology? Or in other words: under which conditions is the entanglement due to indistinguishability a sufficient resource for SSN sensitivity, eliminating the need for a dedicated entangling procedure?

### 7.2 Derivation of the ultimate bound

The interferometric contrast is best when each cloud forms a BEC (or an equivalently coherent collection of bosons). In an ideal case, condensates are pure states of \( \frac{N}{2} \) particles, which together form a TF state

\[
|\psi\rangle_{\text{TF}} = \left| \frac{N}{2} \right\rangle_a \left| \frac{N}{2} \right\rangle_b .
\]

Such state, passing through the Mach-Zehnder interferometer (MZI), can potentially give the sensitivity

\[
\Delta \theta \propto \sqrt{2} N ,
\]

only \( \sqrt{2} \) worse than the HL \([96, 202-204]\). This is an ultimate precision which can be reached with two independent groups of bosons. However, a realistic BEC is not a state with a fixed number of particles. Rather, it exhibits atom number fluctuations from shot to shot and therefore must be described by a mixture

\[
\hat{\rho} = \left( \sum_{N_a=0}^{\infty} P_a(N_a) |N_a\rangle \langle N_a| \right) \otimes \left( \sum_{N_b=0}^{\infty} P_b(N_b) |N_b\rangle \langle N_b| \right) .
\]

Here \( P_i(N_i) \) is a probability for having \( N_i \) particles in the \( i \)th condensate. These probabilities contain complete information whether the state (7.4) is useful for quantum metrology. It is the main goal of the present discussion to extract this information, and we begin by rewriting Eq. (7.4) in the following way:

\[
\hat{\rho} = \sum_{N=0}^{\infty} \sum_{n=-\frac{N}{2}}^{\frac{N}{2}} P_a \left( \frac{N}{2} + n \right) P_b \left( \frac{N}{2} - n \right) \times \left| \frac{N}{2} + n, \frac{N}{2} - n \right\rangle \langle \frac{N}{2} + n, \frac{N}{2} - n\right| .
\]

Here \( n = \frac{1}{2}(N_a - N_b) \) is the atom number difference, \( N = N_a + N_b \) and \( \left| \frac{N}{2} + n, \frac{N}{2} - n \right\rangle \equiv \left| \frac{N}{2} + n\right\rangle_a \left| \frac{N}{2} - n\right\rangle_b \). This form of the density matrix is useful because the two-mode interferometric transformations are generated by the angular momentum operators \( \hat{J}_x = \frac{1}{2}(\hat{a}^\dagger \hat{b} + \hat{a} \hat{b}^\dagger) \), \( \hat{J}_y = \frac{1}{2}(\hat{a}^\dagger \hat{b} - \hat{a} \hat{b}^\dagger) \), and \( \hat{J}_z = \frac{1}{2}(\hat{a}^\dagger \hat{a} - \hat{b} \hat{b}) \), which do not couple states (7.5b) with different \( N \).

In order to assess the interferometric usefulness of the state (7.5) we refer to the quantum Fisher information (QFI, denoted here by \( F_Q \)). It is a quantity, which tells what is the ultimate sensitivity \( \Delta \theta \), optimized over all possible estimation
schemes. Thus the value of $F_Q$ depends only on the state $\hat{\rho}$ and the $\theta$-dependent transformation representing the interferometer [205]. According to the Cramér-Rao lower bound, the relation between the sensitivity and the QFI is [206]

$$\Delta\theta \geq \frac{1}{\sqrt{m\sqrt{F_Q}}} , \quad (7.6)$$

where $m$ is the number of measurements. In line with the definition of the SNL, the values of $F_Q > N$ are attainable only for particle entangled states. When the total number of particles fluctuates (as is the case of the state (7.5)), the SNL is redefined to $F_Q = \bar{N}$, with $\bar{N}$ being the average number of particles in the system [68].

A linear interferometer is represented by an evolution operator $\hat{U} = \exp \left[ -i\theta \hat{J}^{(n)} \right]$, where $\hat{J}^{(n)}$ is a scalar product of a unit vector $\vec{n}$ and a vector ($\hat{J}_x, \hat{J}_y, \hat{J}_z$) of the angular momentum operators. For such transformation the QFI does not depend on $\theta$ [205], and reads

$$F_Q = 2 \sum_{i,j} \frac{(\lambda_i - \lambda_j)^2}{\lambda_i + \lambda_j} |\langle j_i | \hat{J}^{(n)} | j_j \rangle|^2 . \quad (7.7)$$

Here, $\lambda_i$ is the $i$th eigenvalue (with a corresponding eigenvector $|i\rangle$) of $\hat{\rho}$ and $\hat{J}^{(n)}_{ij} = \langle i | \hat{J}^{(n)} | j \rangle$. For illustration, we take $\hat{J}^{(n)}_{ij} = \hat{J}_z$, i.e. $\hat{U} = \exp \left[ -i\theta \hat{J}_z \right]$. Such a mode-mixing operation has been implemented in various atomic interferometers [19, 22–24, 198]. For this transformation, the $F_Q$ calculated using (7.5) is

$$F_Q = \frac{1}{2} \sum_{N=0}^{\infty} \sum_{n=-N/2}^{N/2} \left[ \frac{(\lambda_{n}^{(N)} - \lambda_{-n}^{(N)})^2}{\lambda_{n}^{(N)} + \lambda_{-n}^{(N)}} a_{n}^{(N)} + \frac{(\lambda_{n}^{(N)} - \lambda_{n+1}^{(N)})^2}{\lambda_{n}^{(N)} + \lambda_{n+1}^{(N)}} a_{n+1}^{(N)} \right] , \quad (7.8)$$

where $a_{n}^{(N)} = \left( \frac{N}{2} + n \right) \left( \frac{N}{2} - n + 1 \right)$ and $\lambda_{n}^{(N)} = P_a \left( \frac{N}{2} + n \right) P_b \left( \frac{N}{2} - n \right)$. This expression allows for a quick estimate whether the state $\hat{\rho}$ is useful for quantum metrology. Although we used $\hat{J}^{(n)} = \hat{J}_z$, any transformation in the $x$-$y$ plane gives the same QFI for the state (7.5). For instance, when $\hat{J}^{(n)} = \hat{J}_y$, the following results would apply to the MZI.

We now show that Eq. (7.8) can take a particularly meaningful form under the following assumptions. When the probabilities $P_{a/b}$ change smoothly over the $n \pm 1$ increment, the difference between the two neighbouring eigenvalues in Eq. (7.8) can be approximated with a derivative. Consequently, sums change into integrals. When $P_{a/b}$ are peaked around the average number of atoms in each BEC, then $a_{n}^{(N)}$ is a slowly varying function of $n$ and can be approximated with $\frac{N^2}{1}$. As a result, Eq. (7.8) simplifies to

$$F_Q \approx \frac{1}{8} \int_{0}^{\infty} dN N^2 \int_{N/2}^{N/2} \frac{1}{\lambda_{n}^{(N)}} \left( \frac{\partial \lambda_{n}^{(N)}}{\partial n} \right)^2 , \quad (7.9)$$
Figure 7.1: Left panel: The QFI (solid lines) and the Fisher information (dashed lines) for the population imbalance measurement as a function of the width $\sigma$ for $\bar{N} = 750$ particles normalized to the SNL and averaged over a statistical ensemble of disorders $\xi$, with statistical properties described in the text. In all cases, the shades show the related statistical widths. From the bottom to the top: $\varepsilon = 0$, 0.3, 1. For $\varepsilon = 0.3$ and $\varepsilon = 1$, short horizontal black lines at large $\sigma$ indicate the limiting values of the QFI calculated using Eq. (7.12). Right panel: The exemplary probability distributions for $\sigma = 50$.

These integrals can be estimated assuming that the probabilities are peaked around $\frac{1}{2} \bar{N}$ each, with the spread equal to $\sigma$, giving a universal scaling

$$F_Q \propto \frac{\bar{N}^2}{\sigma^2}. \quad (7.10)$$

For instance, for Gaussian $P$’s [207] we have $F_Q = \frac{\bar{N}^2}{3\sigma^2}$. This means that the state (7.5) is likely to provide SSN sensitivity when the atom number fluctuations in each BEC are sub-Poissonian [208]. To achieve this, one must be able to precisely control the process of condensation, i.e. manipulation of a single coherent cloud. Note also that the quality of state (7.5) is limited by the purity of each condensate, since there is no coherence between the two input arms of an interferometer.

### 7.3 Impact of disorder

It is important to verify how model dependent the result (7.10) is. In other words, if the assumption that the probabilities $P$ are smooth functions broke down, would it still be possible to reach the SSN sensitivity with independently prepared BECs? To address this question, one should try out many conceivable $P_{a/b}$’s, each featuring different characteristic structures, and calculate the QFI. To avoid such a tedious procedure, we rather proceed as follows. We model $P_{a/b}$ by a smooth envelope $P$ together with a randomly chosen realization of stochastic process $\xi(N_i)$ on top,
namely,
\[ P_{a/b}(N_i) \rightarrow P(N_i)[1 + \varepsilon \xi(N_i)]. \]  
(7.11)

Here, \( \varepsilon \in [0, 1] \) is the amplitude of the disorder, while \( \xi \) is of the order of unity. By varying \( \varepsilon \) and the range of correlations of \( \xi \), one can control the average size and the amplitude of characteristic structures superimposed onto the initial smooth function. The condition that \( P_{a/b} \) are equal could be satisfied if the two condensates were produced using some standardized experimental methods. Nevertheless, the main conclusion that follows is valid also when \( P_{a/b} \) are disturbed by independent disorders and have different envelopes.

For illustration, we numerically calculate the QFI using Eq. (7.8) with the Gaussian envelope (peaked around \( \frac{1}{2} \bar{N} \) with a width \( \sigma \)) and the disturbance obtained by drawing random numbers from the interval \([-1, 1]\) independently for each \( N_i \). The process \( \xi \) generated in such a way is stationary, with a zero mean and is characterized by the correlation function \( \xi(N_i)\xi(N_j) = \kappa(|N_i - N_j|) = \frac{1}{3} \delta_{N_i - N_j, 0} \). The result is shown in Fig. 7.1 for \( \varepsilon = 0.3 \) and \( \varepsilon = 1 \) as a function of the width \( \sigma \) with \( \bar{N} = 750 \) particles. Apparently, by introducing the disorder the ill-effect of particle number fluctuations has been reduced. Moreover, the larger amplitude of the disturbance gives higher QFI.

To provide a quantitative explanation of this unintuitive behaviour, we refer to Eq. (7.8). The value of the QFI depends on the difference between the neighbouring eigenvalues of the density matrix. Thus violent jumps between the adjacent \( \lambda \)'s increase the QFI. This is also evident in the continuous limit (7.9) and is embodied by the derivative of \( \lambda \) over \( n \). These observations explain why in the presence of the disorder the value of the QFI grows. In line with these arguments, a disorder which is more smooth (i.e., when the range of the correlation function \( \kappa \) extends over the neighbours) should have less beneficial influence on the QFI.

Moreover, in Fig. 7.1 we observe that while in the absence of the disturbance the value of the QFI drops monotonically with \( \sigma \), in presence of disorder it reaches some constant for major atom number fluctuations. This limiting value can be obtained analytically by averaging the QFI over different realizations of a stationary process \( \xi \). Although, according to Eq. (7.10), the contribution from the smooth envelope tends to zero for large \( \sigma \), the noisy part adds a \( \sigma \)-independent term (for details, see the appendix C) so that
\[ F_Q \propto \varepsilon^2 \bar{N}^2[\kappa(0) - \kappa(1)]. \]  
(7.12)

For the case used in Fig. 7.1, we obtain \( F_Q \overset{\sigma \gg 1}{\longrightarrow} \frac{\varepsilon^2}{\pi} \bar{N}^2 \) which is in excellent agreement with the numerical results. Also note that expression (7.12) confirms that a smoother disorder, when \( \kappa(1) \approx \kappa(0) \), gives a smaller contribution to the QFI.

### 7.4 Precision with the measurement of populations

We now demonstrate that a double-BEC system can give SSN sensitivity of the phase estimation from the measurement of BECs populations. In the experiment
we have in mind, the number of atoms is measured in the output arms of the interferometer. For each total number of atoms $N$, the atom number difference $n$ is calculated and the phase is deduced from the probability of obtaining $n$ and $N$, given $\theta$, which reads

$$p_N(n|\theta) = \text{Tr} \left[ \hat{\rho}(\theta) \left\langle \frac{N}{2} + n, \frac{N}{2} - n \right| \left\langle \frac{N}{2} + n, \frac{N}{2} - n \right\rangle \right],$$

(7.13)

where $\hat{\rho}(\theta)$ is the density matrix (7.5) propagated with the operator $e^{-i\omega J_x}$. For this particular estimation scheme, the QFI in Eq. (7.6) must be replaced with the Fisher

Figure 7.2: The probability distribution $p_N(n|\theta)$ from Eq. (7.13) calculated with Gaussian $F_\sigma$ with mean $\bar{N} = 100$ atoms and width equal to $\sigma = 0.1$ (a), $\sigma = 10$ (b), and $\sigma = 20$ (c). The panels show $p_{N=100}(n|\theta)$ as a function of $n$ (radial variable) and $\theta \in [0, 2\pi]$ (polar variable). The insets are enlargements of the regions marked with orange squares. When $\sigma$ grows [from (a) to (c)], fine structures of the probability vanish, giving a smaller value of the Fisher information (7.14). However, when the noise is added to the probability, according to Eq. (7.11), the fine structures are restored, as shown in panel (d) for $\sigma = 20$ and $\epsilon = 1$. 

we have in mind, the number of atoms is measured in the output arms of the interferometer. For each total number of atoms $N$, the atom number difference $n$ is calculated and the phase is deduced from the probability of obtaining $n$ and $N$, given $\theta$, which reads

$$p_N(n|\theta) = \text{Tr} \left[ \hat{\rho}(\theta) \left\langle \frac{N}{2} + n, \frac{N}{2} - n \right| \left\langle \frac{N}{2} + n, \frac{N}{2} - n \right\rangle \right],$$

(7.13)

where $\hat{\rho}(\theta)$ is the density matrix (7.5) propagated with the operator $e^{-i\omega J_x}$. For this particular estimation scheme, the QFI in Eq. (7.6) must be replaced with the Fisher
Chapter 7. Interferometry with independent BECs

information, which reads

$$ F = \sum_{N=0}^{\infty} \sum_{n=-N/2}^{N/2} \frac{1}{p_N(n|\theta)} \left( \frac{\partial p_N(n|\theta)}{\partial \theta} \right)^2. \quad (7.14) $$

Figure 7.1 compares the QFI with the Fisher information (7.14) optimized over the angle $\theta$ for a smooth Gaussian and distorted probabilities. Although the estimation from the populations is not optimal (it does not saturate the bound of the QFI), it still provides SSN sensitivity. The Fisher information (7.14) can be measured by detecting the distance between two neighbouring probabilities $p_N(n|\theta)$ and $p_N(n|\theta + \delta\theta)$, as recently experimentally demonstrated in [23]. Indeed, for small $\delta\theta$ the Hellinger distance [209] between two probabilities is

$$ 1 - \sum_{N=0}^{\infty} \sum_{n=-N/2}^{N/2} \sqrt{p_N(n|\theta)p_N(n|\theta + \delta\theta)} \simeq \frac{(\delta\theta)^2}{8} F. \quad (7.15) $$

Implementation of this interferometric scheme would confirm that non-classical correlations due to indistinguishability are a resource for SSN interferometry.

In this view, the Fisher information is related to the distinguishability of the probability distributions. A quantitative picture can be obtained by plotting $p_N(n|\theta)$ as a function of $n$ and $\theta$ with fixed $N$. Figure 7.2 shows such probabilities for different widths $\sigma$ of Gaussian $P_{a/b}$ and $\bar{N} = 100$ particles and $N = 100$. When $\sigma$ grows, fine structures of $p_{N=100}(n|\theta)$ diminish, and render this probability at some $\theta$ less and less distinguishable from its neighbour at $\theta + \delta\theta$ [72]. However, this trend is reversed by the addition of noise. The fine structures are restored, and in consequence the Fisher information grows, according to Eq. (7.12).

7.5 Summary

To summarize, we have demonstrated that the particle entanglement between two independently prepared Bose-Einstein condensates might be a sufficient resource for sub-shot-noise metrology. Such an input state is created in a process of condensation, which increases the coherence of each cloud, potentially giving high interferometric signal. Also, since initially there is no coherence between the two arms of the interferometer, the input state is not affected by the inter-mode decoherence and its quality is limited only by the purity of each condensate.

We have derived the ultimate bound for the sensitivity of the parameter estimation. For smooth atom number distributions, it is possible to reach the sub-shot-noise sensitivity when fluctuations in each cloud are sub-Poissonian. This is a robust result, as for more erratic distributions, the sensitivity below the sub-shot-noise can be reached even for vast atom number fluctuations. Importantly, the precision of the phase estimation from the measurement of the number of particles exploits the particle entanglement in such system, opening the possibility for an experimental realization. This observation, in the context of recent measurement of the Fisher
information in a many-body system [23], opens the way towards a construction of a new type of an interferometer operating below the shot-noise limit. Our formulation provides all the necessary tools to evaluate the interferometric efficiency of any double-BEC configuration.
Before we conclude, let us present a brief summary. We began by considering a general concept of distinguishability of states, which was then applied in the quantum-enhanced interferometry. By analysing realistic systems, we were led to the conclusion, that indistinguishability of particles in systems of bosons, which is responsible for the existence of non-classical correlations, is a resource that can be used for ultra-precise interferometry.

To support this idea, we considered two systems, where useful quantum correlations arise solely from the indistinguishability of the constituent bosons. The first example was a system, in which pairs of identical particles were scattered from a Bose-Einstein condensate into two separated regions identified with the two arms of an interferometer. In the second case, the interferometer is fed with independently prepared Bose-Einstein condensates correlated only due to the indistinguishability.

However, the way to the final results was not direct and straightforward. We noted that the entanglement is the property of physical systems which is fragile and volatile. This means that entangled states are susceptible to some actions. To exploit this property, we took the approach, where physical states are characterized by the probability distributions, which govern outcomes of some measurements. To analyse the problem of entanglement from such perspective, we needed precise tools to distinguish those distributions. Therefore, we introduced measures, which determine to what extent two probability distributions are different. Based on this reasoning, the Fisher information naturally emerged as the quantity unifying all the presented measures. This happened under the assumption that the two distributions were close to each other. In other words, the Fisher information quantified the susceptibility of states. The Cramér-Rao lower bound on the estimator constituted the bridge between the distinguishability of states and metrology.

In the next step, we proved a theorem relating the Fisher information with non-classical correlations. This relation manifests under the condition that an interferometric transformation is performed on the system without introducing entanglement among the particles. It follows that the sub-shot-noise sensitivity of an interferometric phase estimation is a signature of non-classical correlations and usefulness of the state for metrology.

Later, we focused on systems of bosons. We proved that in such systems the violation of the Cauchy-Schwarz inequality is a criterion of particle entanglement, provided that the state does not possess superposition between different number states. As the criterion is simple to construct and easy to obtain in modern experi-
ment, it may serve as a starting point to characterize the state, before employing it for interferometry.

After setting the theoretical framework, we considered realistic experimental scenarios. First, we tackled the question, whether a collision of two Bose-Einstein condensates can lead to reduced fluctuations on the selected regions of the halo of scattered atoms. If this was so, a further scheme would be necessary to use the scattered particles in the interferometer. However, it turned out that the configuration forced by the collision in the free space, leads to moderate reduction of the fluctuations. This effect turns out to be destructive for generation of entangled groups of atoms. Therefore, a different scheme would be preferable in which particles are scattered into two separated regions.

In the main part of the thesis, we developed a theory of scattering of atoms from a condensate into disjointed zones. It was shown, that this process is a source of pairs of atoms that are highly entangled and useful for ultra-precise sub-shot-noise interferometry. Moreover, the Cauchy-Schwarz inequality is violated in such a system, so the presence of the entanglement can be easily verified. We applied the theory to the twin-atom beam experiment, and showed that the resulting two groups of atoms are non-classically correlated and useful for metrology.

Finally, inspired by the observation that entanglement comes from indistinguishability of particles, we proposed the scenario in which the initial state of atoms consists of two independently prepared Bose-Einstein condensates. When this state is injected into an interferometer, it can operate with the precision below the shot noise limit. We found the conditions, when this situations takes place, and also proposed the estimation scheme that benefits from such non-classical correlations. The proposed scheme circumvents the difficult stage of the preparation of the entangled state. The only condition, which is required, is to have a precise control over the distribution of the total number of atoms in each atomic cloud.

After summarizing the content of this thesis, we move to the concluding remarks. The above results can be interpreted on two levels. From the practical side, we discovered which types of system configurations are the most preferable when it comes to application of the particle entanglement for interferometry. The systems in which particles are scattered from a common coherent source into separated regions turned out to be particularly useful for ultra-precise metrology. The other interesting system is composed of two independently prepared condensates. It is also useful for sub-shot-noise interferometry, and moreover, its preparation eliminates the need for an entangling procedure. The presented analysis suggests a new way of constructing quantum-enhanced interferometers.

On the conceptual level, we learned that quantum mechanics embodies in its formalism states that are more susceptible than the classical ones. These form a subclass of all particle entangled states. Here, the high susceptibility of states, quantified by the Fisher information, is considered as a potential resource for ultra-precise metrology. Nevertheless, the underlying principle adopted in the thesis may also be applied in other areas of physics.
Appendices
Appendix A

Cauchy-Schwarz inequality and particle entanglement

In this appendix, we derive a general expression for a separable state of $N$ identical bosons and then discuss the importance of the indistinguishability assumption for the CSI criterion.

A.1 Separable state of $N$ identical bosons

Here we prove that the separable states of $N$ identical bosons has a form of Eq. (4.6) as referred to in the main text (see page 81).

The calculation starts from the general form of the density matrix of a separable state of $N$ particles, which is a mixture of $N$-particle product states, i.e.

$$\hat{\rho} = \sum_i p_i \hat{\rho}_i^{(1)} \otimes \cdots \otimes \hat{\rho}_i^{(N)}. \tag{A.1}$$

Here, $\hat{\rho}_i^{(j)}$ is the density matrix of the $j$th particle, while $p_i$ are the statistical weights of the mixture. This $N$-body density matrix can be rewritten as

$$\hat{\rho} = \sum_i P_i |\psi_i\rangle \langle \psi_i|, \tag{A.2}$$

where each $N$-particle ket $|\psi_i\rangle$ is a product of $N$ single-particle pure states

$$|\psi_i\rangle = |\phi_i^{(1)}\rangle \otimes \cdots \otimes |\phi_i^{(N)}\rangle. \tag{A.3}$$

For indistinguishable bosons, only states, which are symmetrized with respect to the particle interchange, are permitted. These states span the bosonic subspace $H_B$ of full $N$-body Hilbert space $H$. We introduce the operator $\hat{\Pi}_B$, which projects onto $H_B$ and note that if $\hat{\rho}$ describes a separable state of $N$ identical bosons, it must be unaltered by the action of $\hat{\Pi}_B$. In particular the equality

$$\text{Tr}[\hat{\rho}] = \text{Tr}\left[\hat{\Pi}_B \hat{\rho} \hat{\Pi}_B \right] = \sum_i P_i \text{Tr}\left[\hat{\Pi}_B |\psi_i\rangle \langle \psi_i|\right] \tag{A.4}$$

is fulfilled if $\text{Tr}\left[\hat{\Pi}_B |\psi_i\rangle \langle \psi_i|\right] = 1$ for all $i$, meaning that each $|\psi_i\rangle$ belongs to $H_B$. The only pure state, which is symmetric and separable is a product of $N$ identical single-particle states $|\phi_i\rangle$, so Eq. (A.3) simplifies to

$$|\psi_i\rangle = |\phi_i\rangle^\otimes N \equiv |\phi_i; N\rangle. \tag{A.5}$$
Appendix A. CS inequality and particle entanglement

The general separable state of indistinguishable bosons is a mixture of such states and reads

\[ \hat{\rho} = \sum_i P_i |\phi_i; N\rangle \langle \phi_i; N| \]  

(A.6)
or when the set of states $|\phi_i; N\rangle$ is indexed with a continuous parameter:

\[ \hat{\rho} = \int D\phi |\phi; N\rangle \langle \phi; N| \mathcal{P}(\phi), \]  

(A.7)

where symbol $D\phi$ denotes the measure of the integration over the set of states $|\phi; N\rangle$. This expression coincides with Eq. (4.6) of the main text.

A.2 Example: two qubits

In this example we apply the above formalism to determine the general form of the separable state of two identical qubits.

Consider a separable state of two qubits

\[ \hat{\rho} = \sum p_j \hat{\rho}_j^{(1)} \otimes \hat{\rho}_j^{(2)}, \]  

(A.8)

where the density matrix of each qubit can be represented using the set of Pauli matrices $\vec{\hat{\sigma}} = (\hat{\sigma}_x, \hat{\sigma}_y, \hat{\sigma}_z)$

\[ \hat{\rho}_j^{(i)} = \frac{1}{2} \left( \mathbb{1} + \vec{\hat{\sigma}}_j^{(i)} \cdot \vec{s}_j^{(i)} \right). \]  

(A.9)

This operator is a valid density matrix when the length of the vectors $\vec{s}_j^{(i)}$ satisfies $|\vec{s}_j^{(i)}| \leq 1$.

We introduce the triplet of bosonic states

\[ |\psi_1\rangle = |\uparrow\rangle^{(1)} \otimes |\uparrow\rangle^{(2)} \]  

(A.10a)
\[ |\psi_2\rangle = |\downarrow\rangle^{(1)} \otimes |\downarrow\rangle^{(2)} \]  

(A.10b)
\[ |\psi_3\rangle = \frac{1}{\sqrt{2}} \left( |\uparrow\rangle^{(1)} \otimes |\downarrow\rangle^{(2)} + |\downarrow\rangle^{(1)} \otimes |\uparrow\rangle^{(2)} \right). \]  

(A.10c)

where $\hat{\sigma}_z^{(i)} |\uparrow\rangle^{(i)} = \pm |\downarrow\rangle^{(i)}$ and the operator

\[ \hat{\Pi}_B = |\psi_1\rangle\langle\psi_1| + |\psi_2\rangle\langle\psi_2| + |\psi_3\rangle\langle\psi_3|, \]  

(A.11)

which projects onto $\mathcal{H}_B$. The trace of the density matrix (A.8) projected onto $\mathcal{H}_B$ is

\[ \text{Tr} \left[ \hat{\Pi}_B \hat{\rho} \right] = \frac{1}{4} + \frac{3}{4} \sum_j p_j \vec{s}_j^{(1)} \cdot \vec{s}_j^{(2)}. \]  

(A.12)

It is equal to unity only when $\vec{s}_j^{(1)} \cdot \vec{s}_j^{(2)} = 1$ for each $j$, which means that $\vec{s}_j^{(1)} = \vec{s}_j^{(2)} \equiv \vec{s}_j$ and $|\vec{s}_j|^2 = 1$. Therefore, the states $\hat{\rho}_j^{(i)}$ are pure and identical for both qubits, and hence the bosonic density matrix reads:

\[ \hat{\rho} = \sum p_j (|\vec{s}_j\rangle\langle\vec{s}_j|)^{\otimes 2}, \]  

(A.13)

which is consistent with Eq. (A.6).
Tradeoffs for number-squeezing in collisions of Bose-Einstein condensates

B.1 Division of the halo into bins

To calculate the number-squeezing parameter we follow a similar procedure to that used in recent experiments [25, 210]. We take an annular “washer-shaped” region matched to the position and width of the halo, that excludes regions near the condensates, and divide it into zones. The matched dimensions of the washer-shaped region in the various calculations are shown in Table B.1.

This annular region is then divided into equi-sized bins by making a series of equally-spaced cuts in the axial ($z$), radial, and tangential directions. Then number of cuts is varied to achieve a gradation from large bin sizes covering an angle of $\pi/2$ in the $k_x - k_y$ plane (8 bins in total, 4 pairs), to bin sizes comparable with the dimensions of a single correlation volume, then on to small bins the size of a single computational lattice point. The progression of bin dimensions is that we first reduce the radial bin size to approximately the correlation length $\sigma_r$, then the tangential and axial ($z$) sizes to their correlation lengths in that order. Then we repeat reductions in the same order down to single-computational-lattice-point volumes. One example is shown in Fig. B.1 for the first $\gamma = 1.02$ dense case in the table.

Number squeezing between opposite bins is calculated, then averaged over all pairs to obtain the displayed values of $\eta^2$. One subtlety should be mentioned: at small bin sizes, mean occupation can vary appreciably as a result of mismatches

| Figures | $\gamma$ | small radius | large radius | maximum $|k_z|$ |
|---------|---------|--------------|--------------|-------------|
| 5.2,5.3,5.5,5.6,B.1 | 1.02 | 0.671 | 1.029 | 0.333 |
| 5.4,5.6 (RBM) | 1.02 | 0.821 | 1.179 | 0.333 |
| 5.7,5.8,5.9 | 0.24 | 0.826 | 1.138 | 0.100 |
| 5.11,5.12 | 1.02 | 0.900 | 1.100 | 0.333 |

Table B.1: The dimensions of the annular region divided into bins for $\eta^2$ calculations, in units of $k_0$. 
between bin shapes and the positions of discrete lattice points on the computational lattice; some small bins miss all lattice points altogether. Such an effect does not appear experimentally, so to exclude potentially disruptive contributions from atypically discretized bins we exclude some bin pairs from consideration. The excluded bins are those for which either bin has an ensemble average occupation less than half or more than twice the mean bin occupation (when averaged over all bins).

B.2 Calculation of halo correlations

The correlation properties are averaged in a similar way to that performed in experiments [25, 210]. That is, we calculate the mean correlation function in a region $R_j$ as a function of inter-particle distance $\Delta k$ in various directions $j$ in the following way:

$$g^{(2)}_{cc/bb}(\Delta k) = \frac{\int_{R_j} d^3k \langle \delta^\dagger(k) \delta^\dagger(k+\Delta k) \delta(k) \delta(k+\Delta k) \rangle}{\int_{R_j} d^3k \, n(k)n(k+\Delta k)}. \quad (B.1)$$

This region is similar to the annular region used to calculate number-squeezing (as shown in Table B.1 and explained above in Appendix B.1) for the $j = z$ direction, but a reduced volume for the other directions $j = t, r$, where the reduction consists of additional restriction to within $k_0/3$ of the x and y axes. However, the washer shape is much wider radially, i.e. it has the same average radius as in the table, but a radial width of $0.9k_0$.

The averaging method used here is as employed in experiments because it is convenient for analysis of data consisting of detected particle positions. It effectively weights the contribution to the correlations proportionally to the product of the densities at the two points $k$ and $k + \Delta k$, which is approximately the local halo
density squared. Thus, it takes into account primarily the most relevant, dense, part of the halo. Then, Gaussian fits are made to obtain the peak values $h$’s and the widths $\sigma$’s. During the fitting, points with excessive statistical uncertainty, or for distances $\Delta k_j$ so large that the correlation function $g^{(2)}$ begins to rise, are excluded.

**B.3 Gaussian model for halo correlations**

We use Gaussian $g^{(2)}$ functions as in Equations (5.24) and (5.25) and approximate the halo by a step function in the radial direction, $k_r = \sqrt{k_x^2 + k_y^2}$. The density is $\rho_0$ when $|k_r - k_h| < w_h$ and zero otherwise. Here, $k_h$ is the halo mean radius, and $w_h$ the radial halo half-width taken to be $\sqrt{\pi/2}$ times the standard deviation of a Gaussian fit to the true radial profile of the halo density. We also take the bins to be centered radially at $k_h$ so the effective integration range in the radial direction extends from $-q_r$ to $+q_r$, where $q_r = \min(L_r/2, w_h)$. Therefore, the integrals of the correlation functions read

$$G_{aa/ab} = \rho_0^2 \int_{-L_z/2}^{L_z/2} dk_z dk'_z \int_{-L_t/2}^{L_t/2} dk_t dk'_t \int_{-q_r}^{q_r} dk_r dk'_r g^{(2)}_{cl/bb}(k, k'),$$

which – using Equations (5.18) and (5.19) – gives

$$G_{aa/ab} = \bar{n}^2 \left[ 1 + h_{cl/bb} f(u_{z/\sigma_{cl/bb}}) f(u_{t/\sigma_{cl/bb}}) f(u_{r/\sigma_{cl/bb}}) \right].$$

Here

$$f(u) = \frac{1}{u^2} \left[ \sqrt{\frac{\pi}{2}} u \text{erf}(u \sqrt{2}) - \frac{1}{2} \left( 1 - e^{-2u^2} \right) \right]$$

is a function of the normalized bin widths

$$u_{z/\sigma}^{cl/bb} = \frac{L_{z/\sigma_{cl/bb}}}{2\sigma_{z/\sigma_{cl/bb}}} \quad u_{t/\sigma}^{cl/bb} = \frac{q_r}{\sigma_{t/\sigma_{cl/bb}}}. $$

This expression is inserted into Eq. (5.22) to obtain the estimates of $\eta^2$ on the basis of correlation and density measurements.
Interferometry with independently prepared Bose-Einstein condensates

C.1 QFI with disordered probabilities

In this appendix we derive the expression for the quantum Fisher information with disordered probabilities. Following Eq. (7.11) on page 136 from the main text, we take

$$P_{a/b}(N_i) \rightarrow P(N_i)[1 + \varepsilon \xi(N_i)],$$

(C.1)

where $\varepsilon$ is the amplitude, while the intensity of the noise is of the order of unity. We take $\xi$ to be a stationary process so that its average, $\langle \xi(N) \rangle$, is constant, which, without loss of generality, can be set equal to zero. The correlation function depends only on the modulus of the argument difference, i.e.,

$$\overline{\xi(N_i)\xi(N_j)} = \kappa(|N_i - N_j|).$$

(C.2)

Though the envelope $P(N_i)$ is a smooth function of $N_i$ with a width $\sigma$, the whole probability distribution might be erratic, because the disorder introduces rapid “jumps.” The characteristic distance between these jumps, called the correlation length $N_{\text{corr}}$, is described by the range of the correlation function $\kappa$,

$$\kappa(\Delta N) \xrightarrow{\Delta N > N_{\text{corr}}} 0.$$  

(C.3)

In the lowest order in $\varepsilon$, the QFI reads

$$F_Q = F_Q^{(0)} + \varepsilon^2 \sum_{N=0}^{\infty} N^2 \sum_{n=-N/2}^{N/2} \Delta \xi_n(N) \times \left[ (\lambda_n^{(N)} - \lambda_{n-1}^{(N)}) + \Delta \xi_n^{(N)} \lambda_n^{(N)} \right].$$

(C.4)

Here, $\lambda_n^{(N)} = P\left(\frac{N}{2} + n\right)P\left(\frac{N}{2} - n\right)$, while $F_Q^{(0)}$ is the QFI in the absence of disorder and

$$\Delta \xi_n^{(N)} = \xi\left(\frac{N}{2} + n + 1\right) - \xi\left(\frac{N}{2} + n\right) +$$

$$+ \xi\left(\frac{N}{2} - n + 1\right) - \xi\left(\frac{N}{2} - n\right).$$

(C.5a)
The limiting value, to which the QFI tends as $\sigma$ grows is given by the average of Eq. (C.4) over trajectories of the process $\xi$ and reads

$$F_Q = F_Q^{(0)} + \varepsilon^2 \sum_{N=0}^{N/2} N^2 \sum_{n=-N/2}^{N/2} [2[\kappa(0) - \kappa(1)] \quad (C.6a)$$

$$+ \kappa(2n + 2) - 2\kappa(2n + 1) + \kappa(2n)] \lambda_n^{(N)}. \quad (C.6b)$$

It is now important to verify how the two terms proportional to $\varepsilon^2$ scale with $\sigma$. First, we take the continuous limit and change the variables $\tilde{n} \equiv n/\sigma$ and $\tilde{N} \equiv N/\sigma$. Since the probabilities $P$ are peaked around $\frac{1}{2}\tilde{N}$, we obtain that a dominating contribution to the QFI at large $\sigma$ is

$$F_Q = F_Q^{(0)} + \text{const} \times \varepsilon^2 \tilde{N}^2 [\kappa(0) - \kappa(1)], \quad (C.7)$$

where the constant comes from the double integration of the $\lambda$ function. Since the term in line (C.6b) in the continuous limit is the second derivative of the correlation function, it will scale inversely with $\sigma$ and is negligible for large atom number fluctuations.

So far we have considered only the lowest order expansion of the QFI in $\varepsilon$. Note that the higher-order contributions, after taking the ensemble average, will be proportional to the integrals of higher order correlation functions. All these functions have a characteristic range, which is of the order of $N_{\text{corr}}$ [211], which is assumed to be small compared to $\tilde{N}$. Therefore, these correlation functions will effectively act as Dirac delta’s on top of the broad envelope set by $P$'s. As a result, the corresponding integrals will scale inversely with $\sigma$, giving a negligible contribution to the QFI for large atom-number fluctuations. This means that Eq. (C.7) is a universal expression for the QFI at large $\sigma$ for disordered potentials (C.1).

Finally, we note that for a Gaussian envelope $P$ used in the main text, i.e.,

$$P(N) \propto \exp \left[ -\frac{(N - \frac{\bar{N}}{2})^2}{2\sigma^2} \right], \quad (C.8)$$

we obtain

$$F_Q \approx \left( \frac{\bar{N}}{2\sigma} \right)^2 + \frac{\bar{N}^2 \varepsilon^2}{2} [\kappa(0) - \kappa(1)], \quad (C.9)$$

which is in excellent agreement with the numerical results presented in Fig. 7.1.


Bibliography


[72] L. Pezzé and A. Smerzi. Entanglement, nonlinear dynamics, and the

Inf. 7, 125 (2009).

[74] V. Giovannetti, S. Lloyd and M. Maccone. Advances in quantum

[75] L. Pezzé and A. Smerzi. in: Proceedings of the International School of
Physics “Enrico Fermi”, Course 188 (Societá Italiana di Fisica, Bologna and
IOS Press, Amsterdam, 2014).

[76] C. Helstrom. Quantum Detection and Estimation Theory (Academic Press,
1976).


[78] B. Yurke, S. McCall and J. Klauder. SU(2) and SU(1,1) interferometers.

Weinfurter, L. Pezzé and A. Smerzi. Fisher information and multiparticle


[81] D. Leibfried, M. D. Barrett, T. Schaetz, J. Britton, J. Chiaverini,
W. M. Itano, J. D. Jost, C. Langer and D. J. Wineland. Toward
Heisenberg-limited spectroscopy with multiparticle entangled states. Science

[82] D. Leibfried, E. Knill, S. Seidelin, J. Britton, R. B. Blakestad,
J. Chiaverini, D. B. Hume, W. M. Itano, J. D. Jost, C. Langer,
R. Ozeri, R. Reichle and D. J. Wineland. Creation of a six-atom


[84] T. Monz, P. Schindler, J. Barreiro, M. Chwalla, D. Nigg,
(2011).

Beating the standard quantum limit with four-entangled photons. Science
316, 726 (2007).


[153] M. Ögren and K. Kheruntsyan. Role of spatial inhomogeneity in disso-

dynamics of condensate collisions using the positive-$P$ representation. Phys.

[155] K. Mølmer, A. Perrin, V. Krachmalnicoff, V. Leung, D. Boiron,
A. Aspect and C. Westbrook. Hanbury brown and twiss correlations in

[156] M. Kozuma, L. Deng, E. Hagley, J. Wen, R. Lutwak, K. Helmer son,
S. Rolston and W. Phillips. Coherent splitting of bose-einstein condensed
atoms with optically induced bragg diffraction. Phys. Rev. Lett. 82, 871
(1999).

effects on the scattered atoms in condensate collisions. Eur. Phys. J. D 65,

[158] P. Ziń, J. Chwedeńczuk, A. Veitia, K. Rzązewski and M. Trippen-
bach. Quantum multimode model of elastic scattering from Bose-Einstein


97, 170404 (2006).

D. Boiron and C. Westbrook. Pair correlations of scattered atoms from
A 78, 053605 (2008).

[162] P. Deuar and P. Drummond. Gauge $P$ representations for quantum-
(2002).

[163] R. Bach, M. Trippenbach and K. Rzązewski. Spontaneous emission of
(2002).

[164] V. Yurovsky. Quantum effects on dynamics of instabilities in Bose-Einstein


List of publications


This work was supported by the Foundation for Polish Science International PhD Projects Programme co-financed by the EU European Regional Development Fund.