Journal of Physics A: Mathematical and General

Phase estimation in interferometry

To cite this article: M Zawisky *et al* 1998 *J. Phys. A: Math. Gen.* **31** 551

J. Phys. A: Math. Gen. **31** (1998) 551–564. Printed in the UK PII: S0305-4470(98)84349-7

Phase estimation in interferometry

M Zawisky†, Y Hasegawa†‖, H Rauch†, Z Hradil‡, R Myška§ and J Peřina§

† Atominstitut der Österreichischen Universitäten, A-1020 Vienna, Austria

‡ Department of Optics, Palacký University, 17. listopadu 50, 772 07 Olomouc, Czech Republic

§ Joint Laboratory of Optics, Palacký University and Czech Acad. Sci., 17. listopadu 50,772 07 Olomouc, Czech Republic

Received 15 May 1997, in final form 16 September 1997

Abstract. An objective theory of phase shift estimation is formulated within quantum mechanics and quantum estimation theory. The general validity of Bayesian analysis even in the limit of very low particle numbers is demonstrated using the information of measured phase-sensitive data. The formalism was applied to the estimation of an unknown phase in neutron interferometry operating near the quantum limit, when only a few particles are registered, and also in the regime of large particle numbers.

1. Introduction

Although interferometry has a history of being a useful technique for research, the quantum limitation of phase measurement must be acknowledged as an open problem in quantum mechanics and one which engages the attention of theoreticians and experimentalists alike [1]. To date only a few experiments have been published concerning phase-shift measurements with few interfering particles limited only by quantum fluctuations. The results of those published were sometimes affected by a particular statistical treatment. For example in cases when only a few data were available, the root-mean-square deviation of phase shift was evaluated instead of the full phase distribution [2, 3]. Alternatively, phase was implicitly assumed as a quantum variable with discrete spectrum. In experiments some registered data were neglected or the phase-shift invariance of resulting distribution was anticipated in operational phase concept [4]. There are also several sophisticated concepts associating phase and phase distribution with marginal distribution of quasidistribution functions in quantum optics [5, 6], even if this treatment sometimes seems to be problematic [7]. As well as already published experimental results accompanied by theoretical models, there are also several contributions analysing the feasible scheme for phase-shift measurements. The ultimate limitation of the accuracy is crucial for the detection of gravitational waves [8]. The improvement beyond the classical limit may be achieved by applying a non-classical signal [9–12]. The papers dealing with such interferometric measurements mainly concentrate on the resolution achievable in the case of a strong signal. For such a treatment there is no need to quantify the content of phase information in the form of phase distribution, since the standard root-mean-square deviation may be used instead. Nevertheless, the careful treatment is necessary even in this semiclassical case. Particularly, the differences between measurements with and without accumulation of counted data may

‖ Current address: Department of Applied Physics, University of Tokyo, Tokyo 113, Japan.

0305-4470/98/020551+14$19.50

appear to be crucial. For example, the estimation of phase shift in interferometer where an equal number particle states interfere, may achieve the resolution up to $1/N$ for multiple measurements [13] whereas a single measurement with the same total particle number N yields the resolution $1/\sqrt{\ln N}$ only [14, 15]. On the other hand, there are theoretical concepts which do not give any realistic instruction on how to get the phase information from the performed measurement [16–21] at all, but they are able to predict the phase distribution corresponding to the given quantum state.

This contribution was inspired by the semiclassical analysis of phase-shift measurement and by estimation theory applied to various kinds of estimation problems in mathematics and contemporary physics [13, 22–30]. Its purpose is to diminish the gap between pure theoretical predictions and possible experimental observations. In comparison with other quantum estimation approaches [31, 32], the realistic feasible measurement (particle counting) will be assumed instead of the optimum detection scheme. The phase is treated as an unknown but non-fluctuating parameter (c-number). The Bayesian estimation of this parameter is objective in the sense that the whole information from the measured phase-sensitive data is used without any prior or additional assumptions about the unknown phase. The phase distribution conditioned by the actual value of phase-shift will be forecast theoretically and verified experimentally even for very few detected particles. Phase estimation based on the evaluation of the likelihood function will be compared with the standard treatments, which fail in the quantum regime, while in the case of high particle numbers all the predictions are comparable. The theoretical analysis will be illustrated on the ideal interferometer in comparison with the approach already used [4] and then used for estimation in neutron interferometry [3, 33, 34], which offers an excellent possibility to operate in the quantum regime when only a few particles may be detected.

This paper is organized as follows. Elements of quantum measurement and estimation are presented in section 2. In section 3 the theory is illustrated on the case of ideal Mach–Zehnder interferometer. Section 4 presents an analysis of realistic measurements with decreased contrast of interference fringes together with the experimental and numerical results, also illustrating a common but faulty and potentially dangerous treatment. The appendix details the incorrect statement 'zero outcome—no information'.

2. Quantum measurement and estimation

Quantum mechanics is a fundamental theory yielding information about the observable quantities of interest. Significantly, such predictions always exhibit a certain inevitable portion of uncertainty connected with the quantum nature of investigated effects. Moreover, the directly measurable variables need not be those, which the experimentalist is interested in. In this case, the desired information must be inferred from the available experimental data and the applied statistical treatment would further increase the uncertainty of measurement. While quantum mechanics predicts outcomes of direct observations, for indirect observations the estimation theory should also be used.

The history of quantum phase started in the early days of quantum mechanics and Dirac [35] attempted to quantize it using the straightforward analogy of polar decomposition. The difficulties associated with this treatment were addressed by Susskind and Glogower [17] and further detailed by Carruthers and Nieto [18]. Besides this approach, there are many other phase concepts addressing quantum phase. Some of them are equivalent and predict the same phase distribution, some of them not. The waves of renewed interest in quantum phase were inspired by the treatment of Pegg and Barnett [20] based on the construction of the phase operator acting on finite-dimensional Hilbert space, by phase

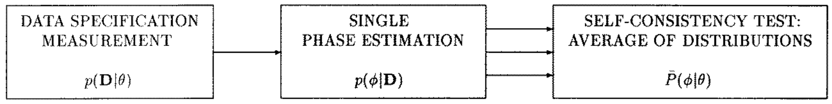

Figure 1. Block scheme of quantum estimation.

distributions obtained via various reconstruction techniques [5], and also by operational definition of phase given by Noh *et al* [4]. An overview of various concepts and methods may be found, for example in [1]. The common feature of all the existing theoretical phase concepts seems to be a missing or improper analysis of realistic phase-sensitive measurement. This is an obvious discrepancy between the theory and the experiment and quantum phase is still assumed as a shaky ground for theoreticians. The approach proposed in this contribution solves this issue in a pragmatic way distinguishing between directly observed variables and those which are inferred from measurements. A typical example is just the problem of phase-shift, which may be registered indirectly via the measurement of phase-sensitive particle numbers in interferometry. The phase-shift θ represents an unknown but non-random c-number displacement parameter in some unitary displacement transformation. The distinction between '*phase measurement*' and '*phase estimation*' may be characterized as follows. In the former case the measurement should be described by a measurable operator with the spectrum confined to the 2π-window. In the latter case any phase-sensitive measurement, from which the phase-shift can be inferred, is allowed. This approach is obviously more suitable for interferometry, where the photocurrents represented by real or natural numbers are registered. On the other hand, the quantum estimation also involves various phase concepts as special cases, provided that the measurement of the corresponding phase operator is considered as a registered 'phase-sensitive' signal.

The specification and measurement of phase-sensitive data represents the first step of the whole estimation problem. In the second step the detected data will be interpreted as a phase-shift prediction $p(\phi|\boldsymbol{D})$. This is the conditional probability of phase-shift ϕ being inferred when the data D is true. There is a wide freedom in the choice of the hypothesis $p(\phi|\boldsymbol{D})$ and therefore the final step should check the self-consistency of the procedure (see the diagram in figure 1).

The phase-sensitive data are specified in the first step. They will be denoted by a variable $\boldsymbol{D}$ representing a matrix of the data $\boldsymbol{D} = \{D_{ij}\}, i = 1, \ldots, n, j = 1, \ldots, m$. This general formulation describes quantum measurement performed simultaneously on m settings of a measuring device and all this measuring procedure, called a *single measurement*, is repeated n times. The term 'simultaneous' does not necessarily mean that all m data of the single measurement are obtained in the same instant of the time, but rather that the information accumulated in this data is used simultaneously. More details about the internal structure of the data will be given in the section detailing the description of the interferometric measurement. For independent measurements, the probability of a result $\boldsymbol{D}$ can be written as a likelihood function

$$p(\boldsymbol{D}|\theta) = \prod_{i,j=1}^{n,m} p_j(D_{ij}|\theta) \tag{1}$$

assembling all the information about the data phase sensitivity.

In the second step the true value of phase-shift will be estimated by means of an estimator ϕ, so-called *inferred phase-shift*. This is the heart of the problem, since in this step one needs to know more than the detected data $\boldsymbol{D}$. For example, this additional information may be the knowledge, how the counted probability (1) depends on the phase-shift. The behaviour of the system in the classical limit may serve for this purpose. However, since the relation between the given phase-shift and detected data is probabilistic, this correspondence of the data to a certain phase-shift (i.e. the estimation strategy) has the form of a hypothesis. Various strategies will be specified explicitly later in this section.

The third step of the whole estimation procedure should reveal how the various estimations are consistent among themselves. Provided that the data sets $\boldsymbol{D}_1, \boldsymbol{D}_2, \ldots$ are available, the experimentalist should decide the purpose of the phase observation. Provided that he wishes to achieve an optimal phase resolution, he has to multiply all obtained hypothesis functions $p(\phi|\boldsymbol{D}_i)$. On the other hand, if he wants to investigate the consistency among the data sets, he has to average over them,

$$\bar{P}(\phi|\theta) = \langle p(\phi|\boldsymbol{D}_i)\rangle_i. \tag{2}$$

In the latter case the obtained distribution converges with increasing number of data sets to the average phase distribution of the *inferred phase-shift* ϕ conditioned by the true value θ,

$$\bar{P}_{\mathrm{mean}}(\phi|\theta) = \sum_{\boldsymbol{D}} p(\phi|\boldsymbol{D})\,p(\boldsymbol{D}|\theta). \tag{3}$$

The sum depletes the spectrum of all possible values of $\boldsymbol{D}$. The statistics of inferred variable is fully determined by the statistics of measured events and by the strategy applied to the evaluation of phase-sensitive data. This result is a counterpart of the theoretical predictions of ideal phase concepts. Various forms of mappings $p(\phi|\boldsymbol{D})$ could be adopted, nevertheless at the end the width of the distribution (3) reveals whether or not such attempts are successful. It is therefore essential to find a reliable correspondence between the detected data and the estimated phase-shift. The right correspondence represents a crucial point of the formulation and one may ask what is optimal. The optimizations of estimation problems remain very complex tasks, solutions of which are not known either in much more simpler cases. However, the following Bayes' estimation has the advantage of its general validity even in the case of very weak fields.

2.1. Bayes' estimation

As is well known [36], the relation between conditioning of two stochastic events may be expressed using Bayes' theorem. Applying the theorem to the quantum estimation, the variable $\boldsymbol{D}$ is a fluctuating quantity whereas the induced phase-shift θ is a given number. On the other hand, the phase estimation ϕ represents a random variable and its statistics may be regarded as a distribution of probability in the sense of degrees of belief [22]. Detection of a particular combination of phase-sensitive data $\boldsymbol{D}$ does not determine exactly the value of the phase shift. Provided that $\boldsymbol{D}$ is detected, the probability of possible values of phase shift can be predicted in a form of a likelihood function. This is the essence of Bayes' approach linking the detected data $\boldsymbol{D}$ and estimated phase ϕ as

$$p(\phi|\boldsymbol{D})\,p(\boldsymbol{D}) = p(\boldsymbol{D}|\phi)\,p(\phi). \tag{4}$$

Quantum theory predicts the conditional probability of detecting data $\boldsymbol{D}$ provided that ϕ is given as $p(\boldsymbol{D}|\phi)$ by relation (1). Provided that there is no prior knowledge about phase-shift, i.e. $p(\phi) = 1/2\pi$, Bayes' theorem gives the estimation

$$p_{\mathrm{B}}(\phi|\boldsymbol{D}) = \frac{1}{C_{\boldsymbol{D}}}\,p(\boldsymbol{D}|\phi) \tag{5}$$

with normalization $C_{\boldsymbol{D}} = \int_0^{2\pi} p(\boldsymbol{D}|\phi)\,\mathrm{d}\phi$, $p(\boldsymbol{D}) = C_{\boldsymbol{D}}/2\pi$. As the most important feature, the interpretation of the measurement is not discrete (deterministic) but fuzzy (probabilistic). Registration of $\boldsymbol{D}$ corresponds to any phase-shift ϕ in the whole interval of the length 2π but with different probabilities according to the distribution $p_{\mathrm{B}}(\phi|\boldsymbol{D})$. This is in accordance with the probabilistic nature of quantum theory. Since the dependence of the detected data on the phase-shift is probabilistic, any strict correspondence between inferred phase-shift and data would be incompetent. As is pointed out in [28], the Bayesian methods are applicable to small data sets lying outside the domain of the central limit theorem. This feature is important in quantum technology where the quantum noise connected with irreducible fluctuations dominates. This has been demonstrated in [30] where the phase properties were ascribed to a single interfering particle.

2.2. Maximum likelihood estimation

The maximum likelihood (ML) estimation represents a choice of estimation procedure belonging to point estimation methods [27]. Adopting this method, the detection of data $\boldsymbol{D}$ is interpreted as an observation of phase $\phi_{\boldsymbol{D}}$ maximizing the likelihood distribution (5),

$$p_{\mathrm{ML}}(\phi|\boldsymbol{D}) = \delta(\phi - \phi_{\boldsymbol{D}}). \tag{6}$$

A histogram of such inferred values creates the average phase distribution (2). A theoretical envelope of this histogram may be obtained on the basis of distribution (3). The phase histogram converges to the smooth phase distribution for an increasing number of repetitions n. Theory is simplified significantly in this case since frequency of the particular value D_j detected at the jth position converges to $np_j(D_j|\theta)$. The conditional phase distribution then reads [29],

$$\bar{P}_{\mathrm{ML}}(\phi|\theta) \propto \left\{ \prod_j \prod_{D_j} [p_j(D_j|\phi)]^{p_j(D_j|\theta)} \right\}^n. \tag{7}$$

A multiplicative factor ensures normalization of the distribution over ϕ on the interval of length 2π. The product contains likelihood functions corresponding to all the values D_j appearing at the jth position with non-zero probability. This distribution characterizes the spread of the ML estimations and is tightly related to the quantities used in information theory such as Kullback–Leibler divergence, Shannon entropy or Fisher information [7, 24, 25]. As a consequence of the central limit theorem, the discrete estimator converges to the true value of the phase for a sufficiently high number of registered particles. If this is not true and if there are not enough detected particles, the application of point estimation methods is meaningless, since the convergence is no longer guaranteed.

2.3. Semiclassical estimation

The previous two estimations were inspired by the analysis of the likelihood function, usually considered as an advanced problem. Nevertheless, a method which is slightly simpler available in textbooks of information theory is known as the *method of moments* (MM) [36]. It may occur that a certain combination of moments of detected fluctuating variable(s) enables us to express the desired estimated variable using a deterministic relation, which may further serve for inversion. This is the core of the principle of correspondence in quantum theory. Provided that quantum field is replaced by a classical one, the relation between detected signal and investigated variable may be specified deterministically and this relation may be imposed back onto the quantum field. This relation may be quantized by

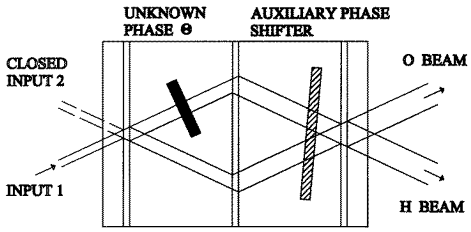

Figure 2. Scheme of Mach–Zehnder neutron interferometer.

replacing the simultaneously measured data by commuting operators and significantly this relation may be extended by definition also to the region of microscopic quantum fields. The method already used in quantum phase [4] may be interpreted in this way. Nevertheless, since the classical relation does not need to be unique, an ambiguity may also appear in a quantum regime [37].

Sometimes it is enough to predict the fluctuation of the phase using a linearized theory. As a phase-sensitive signal, the first moment (mean value) of detected variable $M(\boldsymbol{D}|\theta)$ is usually considered. Fluctuations, both classical or quantum, $\Delta M(\boldsymbol{D}|\theta)$, then determine the uncertainty of estimated variable according to the linearized theory as

$$\Delta\theta \left| \frac{\mathrm{d}}{\mathrm{d}\theta} M(\boldsymbol{D}|\theta) \right| = \Delta M(\boldsymbol{D}|\theta). \tag{8}$$

This technique has been used many times giving satisfactory results [3, 9, 11], although it may fail in some cases. On the other hand, Bayes' statistical analysis of phase-sensitive data developed above always tends to the right phase distribution.

3. Ideal interferometry

The scheme considered here is the Mach–Zehnder interferometer as is generally used in neutron interferometry (figure 2). An unknown phase-shift θ characterizes the path difference between arms of the interferometer. Classically, the measured signal represented by the mean number of particles depends on the induced phase-shift, yielding interference fringes on the outputs. Within quantum mechanics, the phase-shift should be determined by the evaluation of measured phase-sensitive data. Here the number of particles detected on the output ports o and h at m positions of an auxiliary phase-shifter serves for this purpose. For a single measurement ($n = 1$), $\boldsymbol{D}$ represents one detected combination of counting numbers, $\boldsymbol{D} \equiv \left\{N_1^o, N_1^h, \ldots, N_m^o, N_m^h\right\}$. The phase sensitivity is manifested in the counting distribution of possible outputs. Denoting formally the phase-shift transformation induced by the interferometer as $|\psi(\theta)\rangle = \mathrm{e}^{-\mathrm{i}\theta\hat{N}}|\psi\rangle$, $\hat{N}$ is an operator inducing the phase-shift (the difference in the number of particles in both the arms here) [29], the conditional probability of output $\boldsymbol{D}$ when the true phase-shift is θ reads

$$p(\boldsymbol{D}|\theta) = |\langle\boldsymbol{D}|\mathrm{e}^{-\mathrm{i}\theta\hat{N}}|\psi\rangle|^2 \tag{9}$$

with $|\boldsymbol{D}\rangle$ denoting formally the quantum state with the number of particles corresponding to the variable $\boldsymbol{D}$. This measurable quantity plays the crucial role in quantum phase estimation, replacing the notion of classical signal, i.e. the mean number of particles detected at the output. All further considerations will be done for Poissonian statistics [3]. Although neutrons considered in the experiment are fermions, this fact does not influence the photocount statistics. They enter and interfere as single-particle states, demonstrating the

Dirac's statement that 'any particle interferes only with itself' [38]. The Poissonian statistics, characterized by some average number of particles accumulated during the detection time reflects the independence of single neutrons. More formal arguments may be based on the quantization of fermion and boson wavefunctions. Although their annihilation and creation operators fulfill different commutation relations, they both yield the same $SU(2)$ symmetry characterizing an interferometer [12]. The counted distribution (9) is then given by the product (1) with the statistics at the particular positions given as

$$p_j(N_j|\theta) = \mathrm{e}^{-\bar{N}_j(\theta)} \frac{[\bar{N}_j(\theta)]^{N_j}}{N_j!} \tag{10}$$

$\bar{N}_j(\theta)$ being the respective mean number of particles and N_j being the number of particles actually detected at the respective position.

The method will be illustrated with a simple but theoretically worthwhile example of a single mode ideal interferometer with 50/50 lossless beam splitters and one input port. To get unambiguous information about phase-shift in the interval $[0, 2\pi]$, the counting at phase-shifts θ and $\theta - \pi/2$, denoted here by the counting numbers N_1^o, N_1^h and N_2^o, N_2^h, will be considered. These numbers correspond to the outputs N_3, N_4, N_5 and N_6 in the scheme 2 of [4]. The dependence of the average number of counts on the induced phase-shift determines the interference fringes as

$$\bar{N}_1^{o,h}(\theta) = \frac{\bar{N}_{\mathrm{in}}}{2}(1 \pm \cos\theta) \tag{11}$$

$$\bar{N}_2^{o,h}(\theta) = \frac{\bar{N}_{\mathrm{in}}}{2}(1 \pm \sin\theta) \tag{12}$$

$\bar{N}_{\mathrm{in}}$ being the number of particles feeding the open interferometer input port. The Bayes', maximum likelihood, and semiclassical estimations will be detailed here.

A straightforward application of the Bayes' estimation to the case where the phase-shift is inferred after the registration of particle numbers N_1^o, N_1^h and N_2^o, N_2^h yields the inferred conditional distribution (3) for $n = 1$ as

$$\bar{P}_{\mathrm{B}}(\phi|\theta) = \sum_{\{N_j^b\}_{j=1,2}^{b=o,h}} \frac{1}{C_{\{N_j^b\}}} \prod_{\substack{b=o,h \\ j=1,2}} p_{jb}(N_j^b|\phi)\, p_{jb}(N_j^b|\theta) \tag{13}$$

where $C_{\{N_j^b\}} = \int_0^{2\pi} \prod_{j,b} p_{jb}(N_j^b|\phi)\mathrm{d}\phi$. This phase distribution depends only on the mean number of particles counted in the output beams o and h (index b) at the phase-shifter's positions 0 and $\pi/2$ (index j). Further generalization for $n = 2, 3, \ldots$, when the phase-shift is evaluated *after* second, third, ... round of counts, coincides with distribution (13) providing that the total number of particles is conserved. This is a consequence of input Poissonian statistics—the phase distribution (13) depends on the number of samplings n only through the total number of particles $N = n\bar{N}_{\mathrm{in}}$.

The ML estimation predicts a surprisingly simple form of phase distribution for this case. The likelihood distribution (7) may be reduced to a formula

$$\bar{P}_{\mathrm{ML}}(\phi|\theta) \propto \prod_{j,b} [\bar{N}_j^b(\phi)]^{n\bar{N}_j^b(\theta)} \tag{14}$$

$j = 1, 2$ and $b = o, h$ for *any input statistics*. The predicted distribution represents an envelope of histograms of discrete ML values, which are attached to the counted values in the following way. Provided that values $N_{1,2}^{o,h}$ were experimentally counted, the corresponding phase should maximalize the likelihood function (1)

$$p(N_j^b|\phi) \propto (1+\cos\phi)^{N_1^o}(1-\cos\phi)^{N_1^h} \\ (1+\sin\phi)^{N_2^o}(1-\sin\phi)^{N_2^h}. \tag{15}$$

This yields, after a little manipulation the ML equation for ϕ as

$$-N_1^o\sqrt{\frac{1-\cos\phi}{1+\cos\phi}} + N_1^h\sqrt{\frac{1+\cos\phi}{1-\cos\phi}} \\ +N_2^o\sqrt{\frac{1-\sin\phi}{1+\sin\phi}} - N_2^h\sqrt{\frac{1+\sin\phi}{1-\sin\phi}} = 0. \tag{16}$$

The method already used in quantum mechanics [4] is closely related to the semiclassical treatment. A detected data set $\boldsymbol{D} \equiv \{N_1^o, N_1^h, N_2^o, N_2^h\}$ may be interpreted as detection of the phase $\phi_{\boldsymbol{D}}$,

$$\mathrm{e}^{\mathrm{i}\phi_{\boldsymbol{D}}} = \frac{N_1^o - N_1^h + \mathrm{i}(N_2^o - N_2^h)}{\sqrt{(N_1^o - N_1^h)^2 + (N_2^o - N_2^h)^2}}. \tag{17}$$

The discrete phase variable then fluctuates in accordance with the distribution

$$\bar{P}_{\mathrm{MM}}(\phi|\theta) = \sum_{\boldsymbol{D}} p(\boldsymbol{D}|\theta)\delta[\phi - \phi_{\boldsymbol{D}}]. \tag{18}$$

4. Evaluation of experimental data

In neutron interferometry the estimation of an unknown phase difference between both beam paths plays the crucial role [3, 33, 34] and therefore we tried to generalize the methods of phase inference. The visibility of the interference fringes is reduced by several factors, but the statistical theory of phase measurement developed above is still valid. A common technique of phase measurement is to scan an auxiliary phase-shifter over m discrete positions $\Delta_j,\ j = 1, \ldots, m$ and to count neutrons in the output beams o and h within fixed time intervals, see figure 2. The smallest unit of detected data is given by particle numbers $N_1, \ldots, N_m$, which fluctuate according to the Poissonian distribution with the average number of particles given as

$$\bar{N}_j(\theta) = \bar{N}(1 + V\cos(\Delta_j + \theta)). \tag{19}$$

The mean intensity $\bar{N}$ of one output beam and the visibility V of the interference fringes may be assumed as known (for example from the previous measurements), or alternatively, as *a priori* unknown parameters, estimated together with the phase-shift on the basis of the performed counting. In the former case the Bayes' estimation is given straightforwardly by relation (5). In the latter case the probability density for the inferred phase-shift ϕ is given as the function

$$p(\phi|\{N_j\}) \propto \int \mathrm{d}\bar{N} \int \mathrm{d}V \prod_j p(N_j|\phi) \tag{20}$$

normalized on the phase interval $(0, 2\pi)$. In our case the interference pattern is counted at nine equidistant positions of an auxiliary phase-shifter and such a data set consisting of values $N_1, \ldots, N_9$ represents a single measurement. This single measurement is the smallest unit for phase inference in our experiments, which is analysed individually and independently from all the other data. By means of such a measurement the unknown overall phase-shift θ may be determined without ambiguity in the whole interval $(0, 2\pi)$. The single measurement was repeated several hundred times for the purpose of testing the predictions of estimation theory. Experiments were performed at the Atominstitut in Vienna with its 250 kW TRIGA reactor neutron source. The intensity of the input beam was constant during all measurements, which enabled a simple adjustment of the mean particle numbers $\bar{N}$ by the measurement time (equal for all nine measurement positions). Each data series, characterized by $\bar{N}$, consisted of hundreds of single measurements and the visibility V of the data series differed from 0.37 to 0.47.

The phase evaluation process is shown in the series of figures 3(*a*)–(c). The structure corresponding to the separate steps of the estimation procedure sketched in figure 1 is the following. The left-upper panel shows a randomly selected counted event illustrating step 1. The right-upper panel shows its interpretation as a phase-shift estimation corresponding to step 2. All events counted repeatedly form an interference fringe characterized by its visibility and total average number of particles as the left-bottom panel presents. The inferred phase-shift distribution conditioned by the true phase-shift and corresponding to an 'average' single measurement is displayed in the right-bottom panel demonstrating step 3 of the estimation procedure.

The set with the lowest particle number was chosen as $\bar{N} = 0.25$ neutrons for the single measurement in figure 3(*a*). The smallest data unit for phase estimation then consisted of 9×0.25 neutrons on average and phase-shift was inferred from the single measurements in 1721 single runs. Similarly, figure 3(*b*) is characterized by the average number of counted particles in single experiment 9×4.15, each measurement was performed 455 times and figure 3(*c*) demonstrates the behaviour of phase estimation with average particle number 9×71.3 repeated 87 times. Remarkably, figure 3(*c*) indicates that the phase-shift drifted during the long running experiment resulting in an asymmetric average phase distribution. Fitting its shape by a superposition of two Gaussian curves yields the value of phase drift about 25°. Note that this conclusion is not obvious from the common analysis of interference fringe.

Figures 4(*a*) and (*b*) demonstrate a proper and improper manipulation with phase information. The measurement corresponding to $\bar{N} = 0.25$ was used as the data set for figure 4(*a*). Figure 4(*b*) shows results of measurement with 12 times higher average particle number but 12 times smaller number of repetitions, hence with the same total particle number. The full curve with a grey region plotted against the number of single experiments n shows the mean value and 68.3% confidence interval of Bayes' estimation as a product of the likelihood functions of registered single events. Similarly the broken curve with dotted curves displays the mean value and confidence interval of the ML estimation. Here the phase with the ML φ_i is determined separately for each single measurement and then all these estimates are used to evaluate the preferred phase [25]

$$e^{i\varphi} \propto \sum_i e^{i\varphi_i}. \tag{21}$$

The preferred phase is a meaningful analogy of the mean value of variable for the finite interval. The width of this 'classical' estimation is then scaled as $1/\sqrt{n}$ in analogy with the standard root-mean-square deviation. Figures 4 shows the difference between results of Bayes' and ML estimation. The Bayes' estimation is spread around the true value within the width of the confidence interval. For Poissonian statistics the precision of the Bayes' estimation depends only on the total number of particles $N = n\bar{N}_{\mathrm{in}}$. In this case the multiple measurement with very weak input gives the same phase prediction as the single measurement with equivalent total particle number. On the other hand, the repeatedly evaluated ML estimation fails for weak particle numbers. The single number acquired

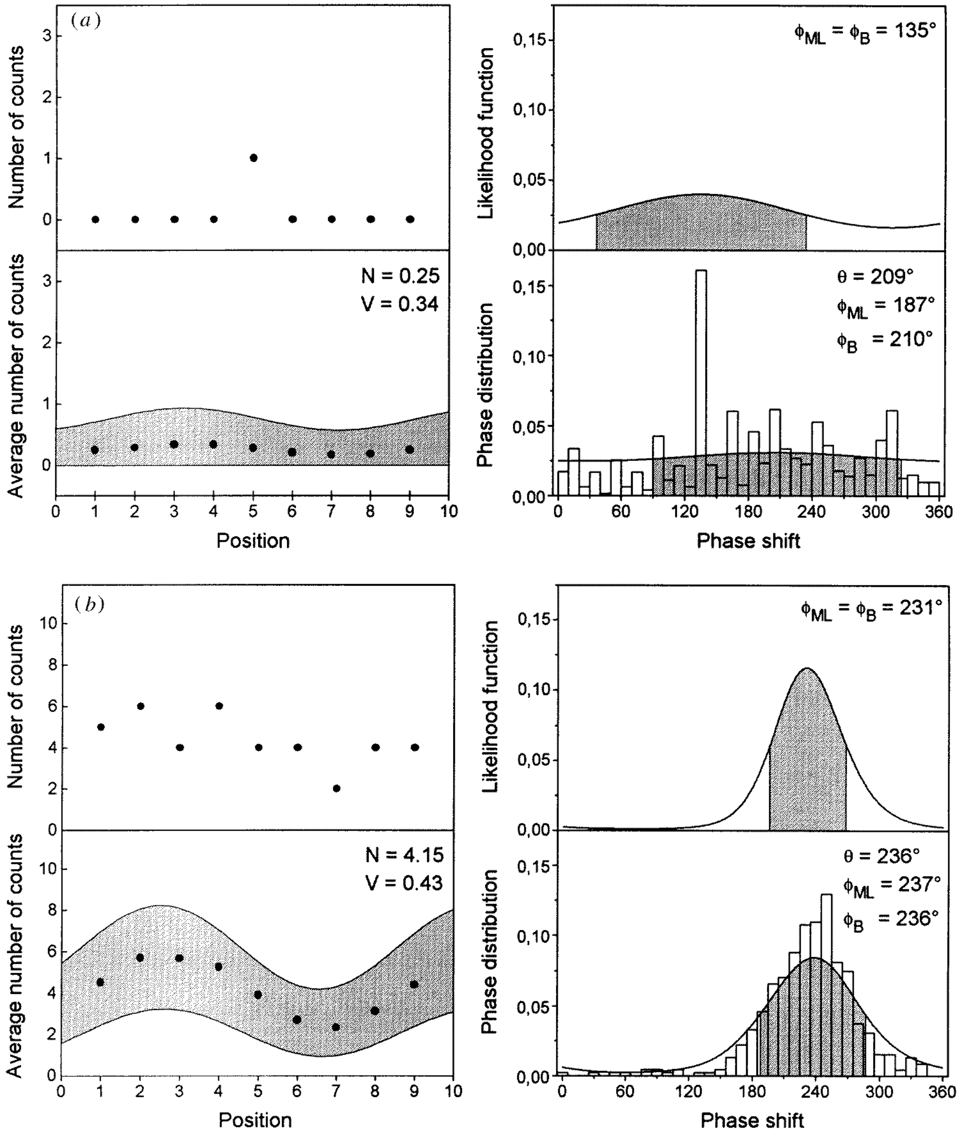

Figure 3. Phase evaluation of a single measurement. The left-upper panels show particle numbers counted at nine positions of phase-shifter (step 1). The right-upper panels show appropriate Bayesian estimates (step 2) including the 68.3% confidence intervals (grey regions). The left-bottom panels show the interference fringes as means of all single measurements. The right-bottom panels show the conditional phase distributions $\bar{P}(\phi|\theta)$ corresponding to average single measurements for both the maximum likelihood (histograms) and Bayes' (smooth curves) estimations (step 3). (*a*) $\bar{N} = 0.25, V = 0.34$. (*b*) $\bar{N} = 4.15, V = 0.43$. (*c*) $\bar{N} = 71.3, V = 0.47$.

from each single measurement cannot describe all information stored in the shape of a non-Gaussian likelihood function. These differences disappear in the case of higher particle numbers plotted in figure 4(*b*). Significantly, the content of phase information may always be defined for Bayes' estimation, while point estimations become moot in the case of weak

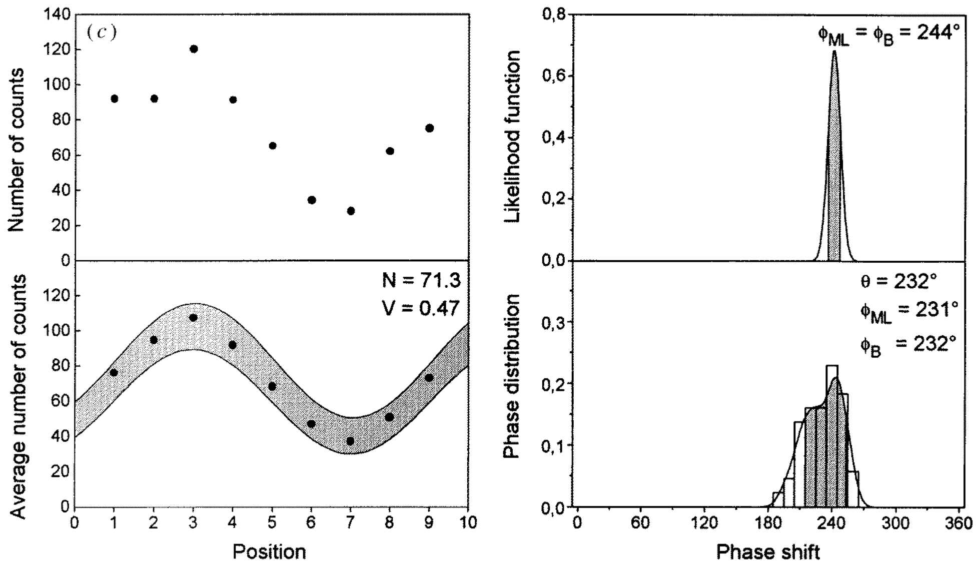

Figure 3. (Continued)

fields. The true value of the phase-shift characterizes the measured sample. Its value was determined as the ML phase corresponding to the all available data collected during repeated runs.

Figure 5 shows the sharpening of the phase information for increasing total number of particles used in experiment. The estimation method corresponds to Bayes' estimation and the same data as in figure 4(a) were used.

One important conclusion from these low counting number measurements is that there is no lower limit for phase inference. Any single measurement is giving useful phase information provided non-zero mean particle numbers and non-zero visibility. Even the so-called zero measurements, when no neutron was detected by the detector during a single run, lead to some phase information, see the appendix. These measurements therefore cannot be discarded. The information accumulated in low particle number measurements should be, nevertheless, used properly. As demonstrated above, improper use of point estimations may lead to incorrect results in this limit. On the other hand, the Bayesian estimation describes the quantum regime well even when very few particles are detected. The only disadvantage of the Bayesian inference is the high computation time, especially when the mean particle number and visibility are also unknown and counting numbers are high. However, for high counting numbers the classical techniques seem to be sufficient.

5. Conclusions

The statistical analysis of quantum-phase estimation in interferometers was given free of any additional assumptions which cannot be verified experimentally. The probability distribution of inferred phase shift was fully determined only by the measured dependence of counted distribution on the induced phase shift and the treatment may be applied to any measurement of phase-sensitive variable. The content of phase information may be evaluated even in the cases when other methods fail, for example in the case of weak fields.

In the case of Bayesian analysis all information on the phase-sensitive data is conserved

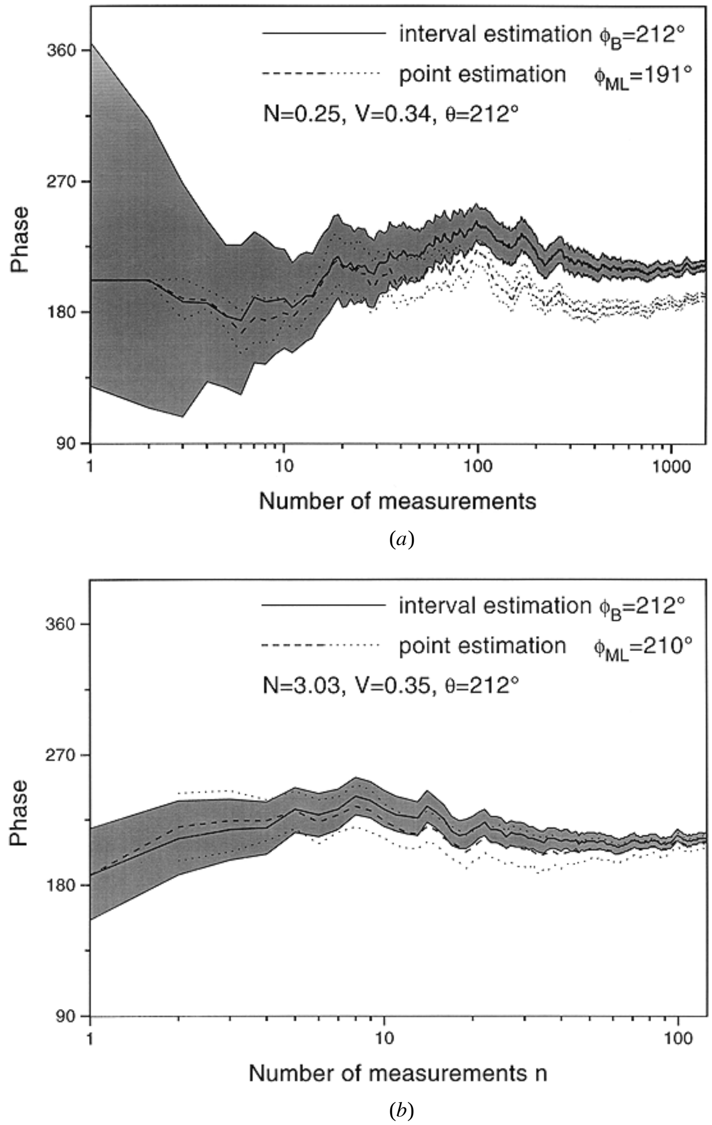

Figure 4. Comparison of Bayes' (interval) estimation with the ML (point) estimation. Phase uncertainty (confidence interval) is plotted against the number of single measurements n used for evaluation. Part (a) corresponds to the data evaluated in figure 3(a), whereas (b) corresponds to the 12 times higher particle numbers in single measurements but the same total particle number.

without discarding any zero or ambiguous data. For practical purposes the shape of the averaged distribution (2) can be used to verify the reproducibility and stability of the measurement. The Bayesian analysis represents an appropriate tool to reveal all information on single detection events releasing an experimentalist from waiting and collecting many data before beginning the analysis. This important feature of quantum estimation can be applied to all quantum measurements where interactive control or fast gain in knowledge is necessary. Our experiments demonstrated that quantum estimation theory and in particular the Bayesian analysis will become an important tool in neutron interferometry, which often

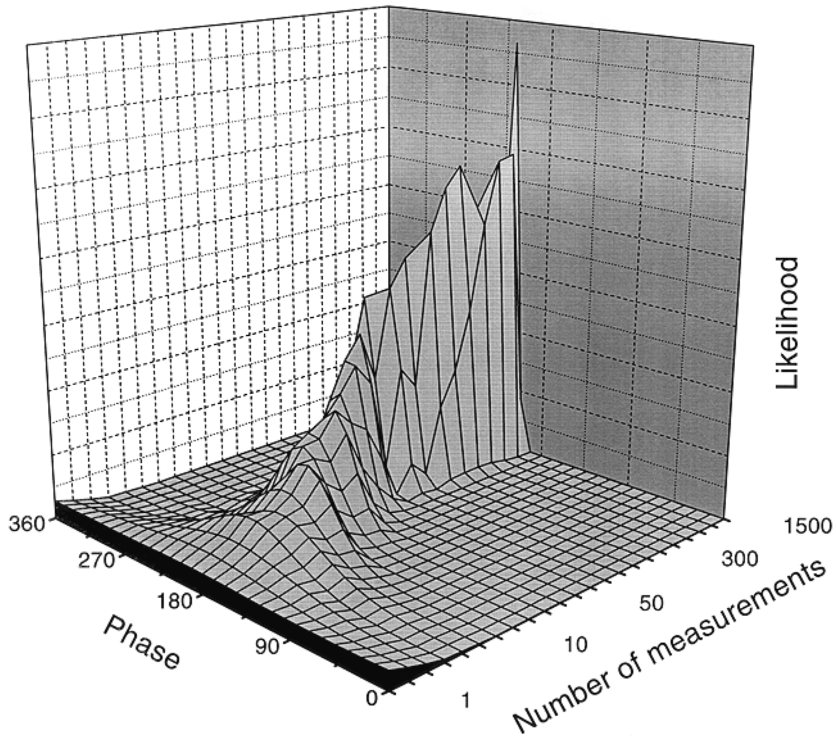

Figure 5. Improvement of the phase resolution with increasing number of particles used for phase evaluation. The likelihood function is plotted as a function of inferred phase and number of single measurements. Data correspond to the same measurement as in figure 3(a).

deals with low neutron numbers and where phase drifts during the measurements have to be analysed in more detail.

Acknowledgments

The support from the East-West programm of the Austrian Academy of Sciences and from the grant VS96028 of the Czech Ministry of Education are gratefully acknowledged.

Appendix. Phase information of zero outcome

A common experimental myth, formulated in brevity as '*zero outcome—zero information*', will be corrected here. The truth of this statement depends strongly on a particular measurement set-up.

Let the counts in the o beam, according the scheme described above, be detected. The nine positions Δ_j are mutually shifted by $\pi/4$, hence $\Delta_1 = \Delta_9$. The output statistics (10) is Poissonian with the dependence of the mean particle number on the unknown phase described by (19). Having detected $N_j^o = 0$ at all of nine positions, the likelihood function (5) reads

$$p_{\mathrm{B}}(\phi|\{0\ldots0\}) \propto \exp\left[-\bar{N}V\sum_{j=1}^{9}\cos(\Delta_j+\phi)\right]. \tag{22}$$

Due to a non-uniform distribution of the measurement positions Δ_j (the position $\Delta_1 = \Delta_9$ is counted twice), the likelihood function is not uniform but can be simplified to

$$p_{\mathrm{B}}(\phi|\{0\ldots 0\}) \propto \exp[-\bar{N}V\cos(\Delta_1+\phi)]. \tag{23}$$

Hence the detection of very zeros in this set-up does yield some phase information. In contrast omitting detection at Δ_9 position tends to a uniform likelihood function with no phase information.

References

[1] For an up-to-date overview, see: Lynch R 1995 *Phys. Rep.* **256** 367
[2] Gerhardt H, Büchler U and Liftin G 1974 *Phys. Lett.* **49A** 119
[3] Rauch H, Summhammer J, Zawisky M and Jericha E 1990 *Phys. Rev.* A **42** 3726
[4] Noh J W, Fougères A and Mandel L 1991 *Phys. Rev. Lett.* **67** 1426
Noh J W, Fougères A and Mandel L 1992 *Phys. Rev.* A **45** 424
[5] Beck M, Smithey D T and Raymer M G 1993 *Phys. Rev.* A **48** R890
[6] Smithey D T, Beck M, Cooper J and Raymer M G 1993 *Phys. Rev.* A **48** 3159
[7] Hradil Z, Myška R, Opatrný T and Bajer J 1995 *Phys. Rev.* A **53** 3738
[8] Hollenhorst J N 1979 *Phys. Rev.* D **19** 1669
[9] Caves C M 1981 *Phys. Rev.* D **23** 1693
[10] Bondurant R S and Shapiro J H 1984 *Phys. Rev.* D **30** 2548
[11] Yurke B, McCall S L and Klauder J R 1986 *Phys. Rev.* A **33** 4033
[12] Yurke B 1986 *Phys. Rev. Lett.* **56** 1515
[13] Holland M J and Burnett K 1993 *Phys. Rev. Lett.* **71** 1355
[14] Peřina J, Hradil Z and Jurčo B 1995 *Quantum Optics and Fundamentals of Physics* (Dordrecht: Kluwer) p 93
[15] Hradil Z and Myška R 1996 *Acta Phys. Slov.* **46** 405
[16] Sanders B C and Milburn G J 1995 *Phys. Rev. Lett.* **75** 2944
[17] Susskind L and Glogower J 1964 *Physics* **1** 49
[18] Carruthers P and Nieto M M 1968 *Rev. Mod. Phys.* **40** 411
[19] Newton R G 1980 *Ann. Phys., NY* **124** 327
[20] Pegg D T and Barnett S M 1989 *Phys. Rev.* A **39** 1665
[21] Shapiro J H and Shepard S R 1991 *Phys. Rev.* A **43** 3795
[22] Kendall M G and Stuart A 1961 *Advanced Theory of Statistics* vol 2 (London: Charles Griffin)
[23] Walkup J F and Goodman J W 1973 *J. Opt. Soc. Am.* **63** 399
[24] Helstrom C W 1976 *Quantum Detection and Estimation Theory* (New York: Academic)
[25] Rao R C 1973 *Linear Statistical Inference and its Applications* (New York: Wiley) p 175
[26] Holevo A S 1982 *Probabilistic and Statistical Aspects of Quantum Theory* (Amsterdam: North-Holland)
[27] Lane A S, Braunstein S L and Caves C M 1993 *Phys. Rev.* A **47** 1667
[28] Jones K R W 1994 *Phys. Rev.* A **50** 3682
[29] Hradil Z 1995 *Phys. Rev.* A **51** 1870
[30] Hradil Z, Myška R, Peřina J, Zawisky M, Hasegawa Y and Rauch H 1996 *Phys. Rev. Lett.* **76** 4295
[31] Osaki M, Ban M and Hirota O 1996 *Phys. Rev.* A **54** 1691
[32] Paris M G A 1997 *Phys. Lett.* A **225** 23
[33] Bonse U and Rauch H (ed) 1979 *Neutron Interferometry* (Oxford: Clarendon)
Zawisky M, Rauch H and Hasegawa Y 1994 *Phys. Rev.* A **50** 5000
[34] Rauch H and Werner S A 1998 *Neutron Interferometry* (Oxford: Clarendon) to be published
[35] Dirac P A M 1927 *Proc. R. Soc.* A **114** 243
[36] Barlow R 1987 *Statistics* (New York: Wiley)
[37] Barnett S M and Pegg D T 1993 *Phys. Rev.* A **47** 4537
[38] Dirac P A M 1947 *The Principles of Quantum Mechanics* (Oxford: Clarendon)