Many-port homodyne detection of an optical phase

M. G. Raymer,* J. Cooper,† and M. Beck
Department of Physics and Chemical Physics Institute, University of Oregon, Eugene, Oregon 97403
(Received 23 April 1993)

We present a quantum analysis of the measurement of a relative optical phase using a many-port method that can be implemented using arrays of photodetectors in the limit of a large number of detector elements. Its application to coherent states and states not describable by the semiclassical theory shows that the many-port method has in principle several advantages compared to an eight-port method previously studied. In some cases, especially for very weak fields, the many-port method provides a more faithful characterization (i.e., a smoother distribution) of the distribution of the relative phase than does the eight-port method. The definition used here for the relative phase between two fields is an operational one, not corresponding to any presently known Hermitian phase operator. We discuss the conditions under which our phase definition corresponds to the best estimate in the maximum-likelihood sense. We find that if the many-port data set is analyzed in a way that retains the contributions from zero photon counts, the many-port phase distribution takes on a character that resembles other phase distributions, such as the Pegg-Barnett or Wigner phase distributions, that show very small modulation depth in the case of weak fields.

PACS number(s): 42.50.Wm, 03.65.Bz, 42.50.Ar

I. INTRODUCTION

It is known that using an eight-port homodyne detector allows quantum-limited measurements of the complex amplitude of a field [1] or of the relative phase between two optical fields [2]. As shown in Fig. 1, of the eight ports, two are for the excited input-field modes, which are each split into two beams, one of which is phase shifted by $\pi/4$. The other two input ports are unexcited, and four ports are for the output fields, each of which is detected in a dual balanced-homodyne arrangement. In the case that one of the input fields (the "reference") is strong and coherent, this apparatus measures simultaneously the two quadrature amplitudes of the other ("signal") field. The simultaneous measurement of both quadratures (represented by noncommuting operators) allows a determination of the optical phase. The price that is paid for the simultaneous measurement of noncommuting operators is the introduction of noise due to the vacuum fields that leak in through the two input beam splitters, in general agreement with the uncertainty principle [3,4]. In this case the measured distribution of phases is a marginal distribution of the Q function for the signal [5–9]. The Q function is broader than the Wigner distribution, which can be measured using four-port homodyne detection, in which case no vacuum leaks in [10].

Another seeming difficulty with the eight-port arrangement is that when both input fields are at the few-photon level the measured difference-phase distribution is comprised of a small number of δ-function peaks, corresponding to small integer numbers of photons hitting each detector on different trials. For example, if one input field is in the vacuum and the other is in a number state with one photon per counting time, the measured relative-phase distribution consists of four δ functions,

$$P_{8p}(\phi)=\tfrac{1}{4}[\,\delta(\phi)+\delta(\phi-\pi/2)+\delta(\phi-\pi)+\delta(\phi-3\pi/2)]\;, \tag{1.1}$$

as derived by Noh, Fougères, and Mandel [2], referred to as NFM hereafter. From this distribution the variance of the phase $\Delta\phi^2=\langle(\phi-\langle\phi\rangle)^2\rangle$ is found to be $\Delta\phi^2=\pi^2 5/16=0.313\pi^2$. (The brackets $\langle\ \rangle$ represent an ensemble average.) These results are contrary to the intuitive notion that number states should have uniform phase distributions, with distribution $P(\phi)=1/2\pi$ and variance $\Delta\phi^2=\pi^2/3=0.333\pi^2$. Another quantity that is used to characterize the fluctuations of the phase is the "dispersion" D, defined in terms of the variances of $C=\cos\phi$ and $S=\sin\phi$ as $D=(\Delta C^2+\Delta S^2)^{1/2}$, where $\Delta C^2=\langle(\cos\phi-\langle\cos\phi\rangle)^2\rangle$, etc. [8]. Both a uniform distribution and the distribution in (1.1) yield for the dispersion $D=1$. The mean phase from (1.1) is $3\pi/4$, which is determined by an arbitrary choice of $[0,2\pi]$ for the window of the phase variable.

One might interpret (1.1) as a sampling of the "true," smooth distribution at four discrete phase values, the number being limited in the one-photon case by the number of detectors used at the output of the interferometer. In the case that one field is in the vacuum and the other is in a large-photon number state, the eight-port method yields many more possibilities for distributing photons

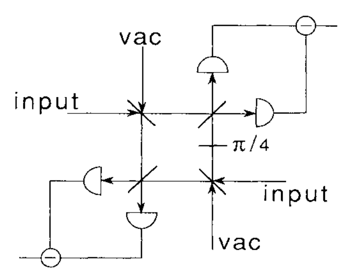

FIG. 1. Eight-port scheme for measuring relative phase between two input fields.

among detectors, yielding a quasiuniform phase distribution, as would be expected.

Another method has been used in experiments to measure phase distributions of optical fields dominated by quantum fluctuations [11]. Spatial intensity fringes are produced by interfering two fields at a slight angle and are recorded by an array of many small detectors. Since the geometry of the setup is well known, the locations of the intensity maxima and minima are used to determine the relative phase. In the case of uncorrelated fields this method yields a uniform, quasicontinuous phase distribution. It has been shown theoretically that for photon-limited fringes with low visibility, this method can be used to make optimum phase estimates in the maximum-likelihood sense, at least for fields described by the semiclassical theory [12]. An analogous many-port technique has been analyzed quantum mechanically in [1], with emphasis on the case of a strong reference field, and with the aim of obtaining complex amplitude rather than phase.

This motivates a more detailed quantum analysis of the array, or many-port, method of phase measurement, and its application to states not describable by the semiclassical theory (due to their lack of a positive-definite Glauber-Sudarshan quasidistribution [13]). Our hypothesis is that a detection scheme with a large number of high-efficiency detectors might provide a more faithful characterization of the distribution of relative phase than does the eight-port method. Two questions are of interest: Does use of many ports increase, decrease, or leave unchanged the amount of vacuum noise that is effectively coupled into the measurement? And, does use of many ports lead to measured distributions that carry more information about the "true phase distribution"? For example, one can measure a distribution in the one-photon case that is essentially smooth?

A related theoretical development is the progress in understanding of the so-called quantum phase distributions—the marginal Wigner distribution [14,15], the marginal Q distribution [5,8,9], the Susskind-Glogower probability-operator-measure (SG-POM) phase distribution [16], or the Pegg-Barnett (PB) Hermitian-phase operator distribution [17]. It is known how to measure the variables described by the Q distribution [6,7], making it possible to build up this distribution by making repeated measurement [1]. It is not known how to measure directly the variables in the Wigner, SG-POM, or PB phase distributions, although all of these distributions can be inferred from complete state measurements obtained using homodyne detection [10,18]. We wish to address whether the many-port scheme can measure directly any of these distributions, or similar distributions.

II. MANY-PORT PHASE MEASUREMENT

Consider the measurement scheme shown in Fig. 2. Two fields are incident from opposite sides of a beam splitter which is at 45° from the faces of two photodetector arrays. We will idealize the arrays as having unity quantum efficiency; limits of present technology will be discussed below. The (positive-frequency) fields incident

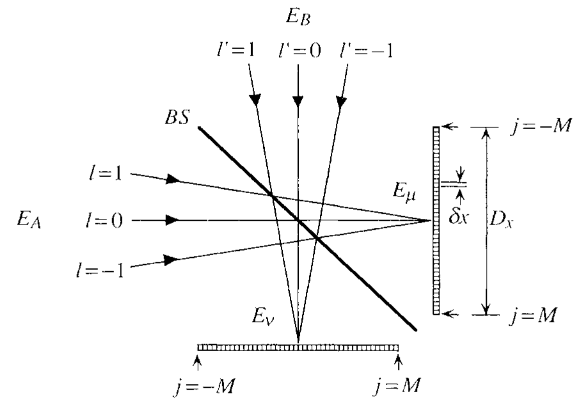

FIG. 2. Many-port scheme for measuring relative phase. Two fields are superposed by a beam splitter (BS) and the resulting fields are detected by two photodetector arrays, each containing $2M+1$ elements.

on the two detectors, labeled μ and ν, are given by

$$\hat{E}_\mu^{(+)}=t\hat{E}_A^{(+)}+r'\hat{E}_B^{(+)}\ ,$$
$$\hat{E}_\nu^{(+)}=t'\hat{E}_B^{(+)}+r\hat{E}_A^{(+)}\ , \tag{2.1}$$

where unitarity of the transformation requires that the complex transmission and reflection coefficients satisfy $rt'^*+tr'^*=0$ [19]. This guarantees that the output fields commute. For simplicity we assume that $|t|=|r|=|t'|=|r'|=(\frac{1}{2})^{1/2}$, corresponding to a lossless 50%-50% beam splitter. These coefficients are written as $t=|t|\exp(i\theta_t), r'=|r'|\exp(i\theta_{r'})$, etc., with the condition $\theta_t+\theta_{t'}-\theta_r-\theta_{r'}=\pm\pi$. Throughout we assume linearly, copolarized fields with uniform intensity distributions filling the aperture of the detectors.

The fields are assumed to be narrow band (relative to the inverse integration time of the detectors), at the same frequency, and may each be weak or strong [20]. The operators describing their electric fields (in Gaussian units) can be expressed at the position x on the surface of the μ detector as

$$\hat{E}_A^{(+)}(x)=\left[\frac{2\pi\hbar\bar{\omega}}{D_xD_yL}\right]^{1/2}\sum_{l=-M}^{M}\hat{a}_l e^{il\Delta kx}\ ,$$
$$\hat{E}_B^{(+)}(x)=\left[\frac{2\pi\hbar\bar{\omega}}{D_xD_yL}\right]^{1/2}\sum_{l'=-M}^{M}\hat{b}_{l'} e^{il'\Delta kx}\ , \tag{2.2}$$

where $\hat{a}_l$ and $\hat{b}_{l'}$ are boson annihilation operators for plane-wave modes with transverse propagation constant $l\Delta k$, for each integer $l=0,\pm1,\pm2,\ldots,\pm M$, where $\Delta k=2\pi/D_x$, with D_x and D_y being the widths of the detector face in the x and y directions, consistent with periodic boundary conditions. There are analogous expressions for the fields at the ν detector. Each plane wave enters the beam splitter at a different angle determined by its transverse propagation constant. L is a longitudinal (z) quantization length and $\bar{\omega}$ is the mean frequency of the field. One of the modes ($l=S$) in $\hat{E}_A^{(+)}$ will contain a "signal field" and one of the modes ($l'=R$) in

$\hat{E}_B^{(+)}$ will contain a "reference field." We wish to measure the relative phase between these two modes. The number of intensity fringes across the detector face is equal to $|p_0|$, where $p_0=S-R$. This is a parameter that is set by the experimenter by adjusting the angle between the signal and reference fields. The reason for summing over other transverse modes is to allow for unexcited (vacuum) modes to enter the device and possibly alter the statistics (i.e., add "shot noise" due to random photon partitioning at the beam splitter).

Each detector is comprised of N small, adjacent detectors (pixels) of width $\delta x = D_x/N$, labeled by position $x_j = j\delta x$, with $j=0,\pm1,\ldots,\pm M$. In order for the field to spatially resolve each pixel the number of modes, $(2M+1)$ per beam, is chosen to equal the number of pixels N. The height of each pixel is D_y (into the page in Fig. 2), and may be much larger than δx, suitable for detecting straight-line fringes perpendicular to the x dimension. The jth pixel on the μ detector receives a number of photons in a time T represented by the operator

$$\hat{N}_{\mu j} = \frac{cTD_y}{2\pi\hbar\bar{\omega}} \int_{x_j}^{x_{j+1}} \hat{E}_\mu^{(-)}(x_\mu)\hat{E}_\mu^{(+)}(x_\mu)dx_\mu$$
$$= \frac{cT}{2NL}\sum_{l,l'}(\hat{a}_l^\dagger\hat{a}_{l'} + \hat{b}_l^\dagger\hat{b}_{l'} + \hat{a}_l^\dagger\hat{b}_{l'}e^{i\eta} + \hat{b}_l^\dagger\hat{a}_{l'}e^{-i\eta}) \times \exp\left[-\frac{i(l-l')j2\pi}{N}\right], \quad (2.3)$$

where the phase $\eta=\theta_{r'}-\theta_t$ is a property of the beam splitter. The jth pixel on the ν detector receives a number represented by

$$\hat{N}_{\nu j} = \frac{cT}{2NL}\sum_{l,l'}(\hat{a}_l^\dagger\hat{a}_{l'} + \hat{b}_l^\dagger\hat{b}_{l'} - \hat{a}_l^\dagger\hat{b}_{l'}e^{i\eta} - \hat{b}_l^\dagger\hat{a}_{l'}e^{-i\eta}) \times \exp\left[-\frac{i(l-l')j2\pi}{N}\right], \quad (2.4)$$

where we used $\exp[i(\theta_{t'}-\theta_r)] = -e^{i\eta}$. The exponential factors represent the relative phase shifts due to the path difference of the beams at each pixel. Next we compute the sum and difference photon numbers for each pair of pixels with equal j values, $\hat{S}_j = \hat{N}_{\mu j} + \hat{N}_{\nu j}$ and $\Delta\hat{N}_j = \hat{N}_{\mu j} - \hat{N}_{\nu j}$, which are found to be

$$\hat{S}_j = \frac{cT}{NL}\sum_{l,l'}(\hat{a}_l^\dagger\hat{a}_{l'} + \hat{b}_l^\dagger\hat{b}_{l'})\exp\left[-\frac{i(l-l')j2\pi}{N}\right], \quad (2.5)$$

$$\Delta\hat{N}_j = \frac{cT}{NL}\sum_{l,l'}(\hat{a}_l^\dagger\hat{b}_{l'}e^{i\eta} + \hat{b}_l^\dagger\hat{a}_{l'}e^{-i\eta})\exp\left[-\frac{i(l-l')j2\pi}{N}\right]. \quad (2.6)$$

Everywhere in the following, the factor $e^{i\eta}$ will be absorbed into each operator $\hat{b}_{l'}$, since it is just an overall phase shift of the E_B field. In the case of a strong reference field the subtraction removes unwanted classical noise on the reference field [21]. In the case of two weak fields the use of two detectors ensures that all photons are counted.

These operators can be interpreted as follows. $\hat{S}_j$ can be rewritten as an incoherent sum of photon numbers, $\hat{S}_j = (cT/L)[\hat{\Phi}_{Aj}^{(-)}\hat{\Phi}_{Aj}^{(+)} + \hat{\Phi}_{Bj}^{(-)}\hat{\Phi}_{Bj}^{(+)}]$, where $\hat{\Phi}_{Aj}^{(+)}$ is a spatially localized (transversely) photon annihilation operator for the signal beam, for the jth pixel [22],

$$\hat{\Phi}_{Aj}^{(+)} = \left[\frac{1}{N}\right]^{1/2}\sum_{l=-M}^{M}\hat{a}_l e^{ilj2\pi/N}, \quad (2.7)$$

and similarly for the reference beam,

$$\hat{\Phi}_{Bj}^{(+)} = \left[\frac{1}{N}\right]^{1/2}\sum_{l=-M}^{M}\hat{b}_l e^{ilj2\pi/N}. \quad (2.8)$$

These obey the commutation relations $[\hat{\Phi}_{Aj}^{(+)},\hat{\Phi}_{Ak}^{(-)}]=\delta_{jk}$, etc. The difference number, on the other hand, is influenced by interference; for example, its average value depends on the correlations $\langle\hat{a}_l^\dagger\hat{b}_{l'}\rangle$ between the various mode amplitudes. It can be written as $\Delta\hat{N}_j = (cT/L)[\hat{\Phi}_{Aj}^{(-)}\hat{\Phi}_{Bj}^{(+)} + \hat{\Phi}_{Bj}^{(-)}\hat{\Phi}_{Aj}^{(+)}]$. The difference-number operators commute, for different j, and so are simultaneously measurable.

The information concerning the relative phase of the two input fields is contained in the positions of the fringe maxima and minima on the detector's faces. On a single exposure of the detector array to the field, if many photons are detected one can determine the relative phase by fitting a sinusoidal function to the data. This method has been used experimentally to study phase-correlation effects in stimulated Raman scattering where the photon numbers are large and the statistics are thermal-like [11]. By making many measurements on similarly prepared systems, statistical distributions of relative phase were obtained under various conditions.

Consider the case of coherent states in the signal and reference modes, $|\alpha_S\rangle_S|\beta_R\rangle_R$, where $\alpha_S=|\alpha_S|\exp(i\theta_S)$ and $\beta_R=|\beta_R|\exp(i\theta_R)$. The quantum expectation value (ensemble average) of the photocount number at the pixel labeled a, which is the j_ath pixel on the μ or ν array, is

$$\langle\hat{N}_{j_a}\rangle = \frac{cT}{2NL}\left[|\alpha_S|^2+|\beta_R|^2 + 2f_a|\alpha_S\beta_R|\cos\left[\frac{j_a p_0 2\pi}{N}+\theta_S-\theta_R\right]\right], \quad (2.9)$$

where the symbol $f_a=1$ for pixels on the μ array and $f_a=-1$ for pixels on the ν array. The parameter $p_0=S-R$ determines the period of the fringe modulation (since $l=S$ and $l'=R$ as defined earlier). We will normally consider $p_0=1$ for which there is one period of modulation across each detector array. The mean difference-count number for the jth pair of pixels is given by

$$\langle\Delta\hat{N}_j\rangle = 2\frac{cT}{NL}|\alpha_S\beta_R|\cos\left[\frac{jp_0 2\pi}{N}+\theta_S-\theta_R\right]. \quad (2.10)$$

A convenient way to extract the phase information is to perform a discrete Fourier transform (DFT) on the data set $\{\Delta N_j\}$ corresponding to the measurements of the difference-number operators on a particular trial. This is defined as

$$K(p)=\sum_{j=-M}^{M}\Delta N_j e^{-ijp2\pi/N} . \qquad (2.11)$$

The resulting complex function of the index $p=0,1,\ldots,N(=2M+1)$ will have a peak at the spatial frequency $(p=p_0)$ corresponding to the chosen fringe spacing. The complex phase of the peak value $K(p_0)$ of the transform will be the desired optical phase, that is,

$$\phi=\arg[K(p_0)]=\arctan(K_Y/K_X) , \qquad (2.12)$$

where $K(p_0)=K_X+iK_Y$. In our use of the arctangent of the ratio of two numbers the relative signs of the numbers are used to place the phase in the correct quadrant. For classical fields with small interference visibility this method is optimum in the maximum-likelihood sense, as discussed in Sec. VI below [12].

As an example, consider the case of coherent states. The DFT of the mean of the difference numbers is given in Eq. (2.10) as

$$\langle K(p)\rangle=(cT/L)|\alpha_S\beta_R|e^{i(\theta_S-\theta_R)}\delta_{p,p_0} , \qquad (2.13)$$

where δ_{p,p_0} is a Kronecker delta. This shows that in the classical limit the component $\langle K(p_0)\rangle$ carries the relative-phase information as its complex argument, while other components are zero.

The quantum statistics of the fringe fluctuations are represented by the Fourier transform in operator form,

$$\hat{K}(p)=\sum_{j=-M}^{M}\Delta\hat{N}_j e^{-ijp2\pi/N} , \qquad (2.14)$$

which using Eq. (6) becomes

$$\hat{K}(p)=\frac{cT}{L}\sum_{l=-M}^{M}(\hat{a}_l^\dagger\hat{b}_{l+p}+\hat{b}_l^\dagger\hat{a}_{l+p}) . \qquad (2.15)$$

One can view (2.14) or (2.15) evaluated at $p=p_0$ as the operator measured by this many-port scheme. Note that for the purpose of calculating higher moments of (2.15), terms containing only vacuum modes cannot in general be dropped; they are responsible for the differences between our scheme and the eight-port scheme.

Although there is a certain similarity between our operator $\hat{K}(p)$ and the operator $(\hat{n}_4-\hat{n}_3)+i(\hat{n}_6-\hat{n}_5)$ defined by NFM for the eight-port device, it does not appear that the eight-port operator can be interpreted as a special case of our many-port operator. This is because in the eight-port device each beam reflects from or passes through two beam splitters, imparting a certain set of phase relationships between the beams. In our case, with only one beam splitter, the phase relationships are different.

We can define a phase operator in the following way. Define Hermitian operators by $\hat{K}_X=[\hat{K}(p_0)+\hat{K}^\dagger(p_0)]/2$ and $\hat{K}_Y=[\hat{K}(p_0)-\hat{K}^\dagger(p_0)]/2i$, so that $\hat{K}(p_0)=\hat{K}_X+i\hat{K}_Y$. It can be shown that $\hat{K}_X$ and $\hat{K}_Y$ commute, making it possible to define a relative-phase operator by

$$\hat{\phi}=\arg[\hat{K}(p_0)]=\arctan[\hat{K}_Y/\hat{K}_X] . \qquad (2.16)$$

Note, however, that this operator is not defined for any state for which both $\hat{K}_X$ and $\hat{K}_Y$ equal zero [which means $\hat{K}(p_0)=0$]. This includes the vacuum state and other less likely states with photons arriving only in pairs that hit each of two pixels having the same pixel numbers, so that $\Delta N_j=0$ for all j. So the operator in (2.16) is in general not Hermitian. To overcome a similar limitation NFM proposed to omit from the analysis any data counts for which their (analogous) phase operator is ill defined. This is an acceptable experimental procedure, although it does not apparently give rise to a Hermitian-phase operator. Rather it leads to the "operational" phase measurement. In Sec. III we will adopt this same procedure, while in Sec. VII we will discuss an alternative procedure that retains the zero counts. These provide two new operational definitions of phase.

III. PHASE STATISTICS

To find the statistics of the phase measurements, the statistics for the complete set of pixel counts $\{N_{\mu,j},N_{\nu,j}\}$ must be evaluated. To this end we will make an idealization that the number N of pixels approaches infinity. Therefore each pixel receives either zero or one photon on a given trial. In this limit the probability to detect one photon in the pixel labeled a is $P(1_a)=\langle:\hat{N}_a:\rangle$, where the number operator $\hat{N}_a$ is given by (2.3) or (2.4), and the symbol :: indicates normal ordering of creation and annihilation operators $\hat{\Phi}^{(\pm)}_{A(B)j}$ *at the detectors.* For present purposes this is equivalent to normally ordering the signal and reference operators $\hat{a}_l$ and $\hat{b}_l$. This yields for the a pixel, which is the j_ath pixel on one of the two arrays,

$$\begin{aligned}P(1_a)&=P(1_{f_a,j_a})\\&=\frac{cT}{2NL}\Bigg[\langle\hat{a}_S^\dagger\hat{a}_S\rangle+\langle\hat{b}_R^\dagger\hat{b}_R\rangle\\&\qquad+f_a\langle\hat{a}_S^\dagger\hat{b}_R\rangle\exp\left[-\frac{ip_0j_a2\pi}{N}\right]\\&\qquad+f_a\langle\hat{b}_R^\dagger\hat{a}_S\rangle\exp\left[\frac{ip_0j_a2\pi}{N}\right]\Bigg] , \qquad (3.1)\end{aligned}$$

where again $f_a=1$ (-1) if the pixel is in the μ (ν) array. The joint probability to detect one photon in each of two pixels labeled a and b is $P(1_a,1_b)=\langle:\hat{N}_a\hat{N}_b:\rangle$. The joint probability to detect one photon in each of a particular set of pixels $a,b,c,\ldots,z$ is [23]

$$P(1_a,1_b,1_c,\ldots,1_z)=\langle:\hat{N}_a\hat{N}_b\hat{N}_c\cdots\hat{N}_z:\rangle . \qquad (3.2)$$

For weak fields many trials will result in no photons being observed. Following NFM, one can choose to throw such cases out of the ensemble, since they carry no phase information; then the sample must be accordingly renormalized. Such events occur with probability [23]

$$P(0_a,0_b,\ldots,0_{2N})=\langle:e^{-\hat{N}_a}e^{-\hat{N}_b}\cdots e^{-\hat{N}_{2N}}:\rangle . \qquad (3.3)$$

Also, if all photons should arrive in pairs that hit pixels with the same pixel numbers j then no phase information can be obtained because $K(p_0)=0$, and these cases may also be discarded. Actually, in our scheme with a very large number of pixels, the probability for such events is negligible, so these cases need not be explicitly discarded.

IV. PHASE STATISTICS WITH STRONG REFERENCE FIELD

Consider first the case when the reference mode is in a strong coherent state and the signal mode is weak. Then the terms that dominate the sum in (2.15) are those containing the reference-mode amplitude $\hat{b}_R$, which is replaced by its coherent-state amplitude $\beta_R=|\beta_R|\exp(i\theta_R)$, giving

$$\hat{K}(p_0)\simeq\frac{cT}{L}(\hat{a}^\dagger_{R-p_0}\beta_R+\beta^*_R\hat{a}_S)\ , \tag{4.1}$$

where again, $S=R+p_0$. This expression is identical in form to that found in optical heterodyne detection, in which a frequency-shifted local oscillator (LO) field β_R is interfered with the signal field. A so-called image mode, in the vacuum state, is introduced through its temporal beating with the LO [5]. In the present case the image mode, with index $R-p_0$, is introduced through its spatial beating with the reference field. That is, it contributes to the noise on the interference pattern. Alternatively one may say that the noise is introduced by the random partitioning of signal photons at the beam splitter. In any case, the distribution of measured values of $\hat{K}(p_0)$ is known to be given in this limit by the quasidistribution function known as the Q function [1,5–7]. Replacing the reference field operator by its coherent amplitude, and choosing $\theta_R=0$, we can write

$$\hat{K}(p_0)=\frac{cT|\beta_R|}{\sqrt{2}L}(\hat{X}+i\hat{Y})\ , \tag{4.2}$$

where the Hermitian operators $\hat{X}$ and $\hat{Y}$ are $\hat{X}=\hat{x}_S+\hat{x}_{R-p_0}$ and $\hat{Y}=\hat{y}_S-\hat{y}_{R-p_0}$, where $\hat{a}_n=(\hat{x}_n+i\hat{y}_n)/2^{1/2}$ for both signal and image modes. Because $\hat{X}$ and $\hat{Y}$ commute they are simultaneously measurable. Their joint distribution is given by the Q function of the signal mode [6,7],

$$Q(X,Y)={}_S\langle\alpha_S|\hat{\rho}|\alpha_S\rangle_S\ , \tag{4.3}$$

where $|\alpha_S\rangle_S$ is a signal-mode coherent state with amplitude given by $\alpha_S=X+iY$, and $\hat{\rho}$ is the density operator for an arbitrary state of the signal mode. This function is positive and well behaved.

The probability density function for the difference phase is an average over a δ function,

$$\begin{aligned}P^{(Q)}_\phi(\phi)&=\int dX\,dY\,Q(X,Y)\delta(\phi-\arctan(Y/X))\\&=\int_0^\infty rdr\,Q(X[r,\phi],Y[r,\phi])\ ,\end{aligned} \tag{4.4}$$

where X and Y are expressed in polar coordinates. This marginal distribution has been discussed previously [5,9], and is the same distribution obtained in the strong-field case using the eight-port scheme [7].

In this case we find no fundamental advantage to using a many-port detector over the eight-port detector. In both cases an irreducible amount of noise, represented here by the image mode, is introduced. This noise is a necessary consequence of designing an apparatus that is capable of simultaneously measuring both quadratures of the signal field [4,6]. Simultaneous X,Y measurement is necessary if one is to make phase measurements on single trials rather than inferring phase distributions from an ensemble of single-quadrature measurements, as was done by Smithey *et al.* in [10] and [18].

A specific example is the case where the signal field is in a coherent state $|\alpha_S\rangle_S$, where the amplitude is written $\alpha_S=|\alpha_S|\exp(i\phi_c)=(\bar{x}+i\bar{y})/2^{1/2}$. In this case the Q function is

$$Q(X,Y)=\frac{1}{2\pi}\exp\left[\frac{(X-\bar{x})^2}{2}\right]\exp\left[-\frac{(Y-\bar{y})^2}{2}\right]\ . \tag{4.5}$$

If the coherent state is very weak, so that $|\alpha_S|\ll1$, then the phase distribution (4.4) is given, to first order in the amplitude, by

$$P^{(Q)}_\phi(\phi)=\frac{1}{2\pi}[1+\pi^{1/2}|\alpha_S|\cos(\phi-\phi_c)]\ . \tag{4.6}$$

Note that, according to this definition, as the signal amplitude becomes smaller, the phase distribution becomes more nearly uniform, indicating that the relative phase between the reference and signal fields becomes less well defined.

V. PHASE STATISTICS WITH TWO WEAK FIELDS

When the photon numbers are of the order of one in both fields, there are differences between the many-port and the eight-port approach.

A. One-photon number state

An illustrative case is when the signal mode is in a one-photon number state and the reference mode is in the vacuum. If we choose $cT=L$ then the photon will be detected with unity efficiency some time during the observation interval. One and only one pixel will register a photon. Label this pixel by its array label f_a and its pixel number j_a. All difference numbers will be zero except for one, denoted by $\Delta N_{f_a,j_a}$, which will have value 1 (for $f_a=1$) or -1 (for $f_a=-1$), with equal probability. The sum in (2.11), for $p_0=1$, thus gives $K(1)=f_a\exp(-ij_a2\pi/N)$ for the peak of the Fourier transform. The phase on that particular trial is then determined to be

$$\phi_{f_a,j_a}=\arg[K(1)]=-\frac{j_a2\pi}{N}+\left[\frac{\pi}{2}\right](1-f_a)\ . \tag{5.1}$$

This can be interpreted simply as the phase of an interference pattern that has maximum intensity at the position at which the photon was detected. Because there are N pixels per array, there are N distinct possible measured phase values, uniformly spread between $-\pi$ and π, all with equal probability. In the limit of many pixels the distribution becomes uniform, $P(\phi_n)=1/N$, as expected for a number state. Recall that for this state the eight-port method yields (1.1), which is not uniform. In this sense the many-port scheme yields the "true" (or at least more sensible) distribution for this state, and thus is better than the eight-port scheme. In the limit of many pixels the phase variance is $\Delta\phi^2=\pi^2/3=0.333\pi^2$ and the dispersion squared is $D^2=\Delta C^2+\Delta S^2=1$.

B. Split-photon state

Another interesting example to apply this method to is that resulting from splitting a one-photon state at a 50%-50% beam splitter, and introducing the two output modes into the signal and reference ports in Fig. 2. The output state is written, in a two-mode Fock basis, as $|\psi\rangle\rangle=C_{01}|0\rangle_S|1\rangle_R+C_{10}e^{i\phi_C}|1\rangle_S|0\rangle_R$ [2], where the C_{ij} are real and positive, and ϕ_C is a classical phase shift introduced into one beam by, say, a piece of glass. We might wish to determine this phase shift. For $cT=L$, one and only one pixel will detect a photon, and from (3.1) the probability for each pixel (labeled a) is (for $p_0=1$)

$$P(1_{f_a,j_a})=\frac{1}{2N}\left[C_{01}^2+C_{10}^2+2f_aC_{01}C_{10}\cos\left[\frac{j_a2\pi}{N}+\phi_c\right]\right]. \tag{5.2}$$

Given that the photon is detected in the (f_a,j_a) pixel, the measured phase will again be given by $\phi_{f_a,j_a}=-j_a2\pi/N+(\pi/2)(1-f_a)$. The probability for observing this phase value is given by substituting this phase expression into (5.2), giving

$$P_\phi(\phi_{f_a,j_a})=\frac{1}{2N}[C_{01}^2+C_{10}^2+2C_{01}C_{10}\cos(\phi_{f_a,j_a}-\phi_c)], \tag{5.3}$$

which is independent of f_a, that is, which array detected the photon. The two cases $f_a=\pm1$ can be combined to give the measured phase distribution

$$\begin{aligned}P_\phi(\phi_n)&=P_\phi(\phi_{1,j_a})+P_\phi(\phi_{-1,j_a})\\&=\frac{1}{N}[C_{01}^2+C_{10}^2+2C_{01}C_{10}\cos(\phi_n-\phi_c)],\end{aligned} \tag{5.4}$$

where ϕ_n takes on N values uniformly spread from $-\pi$ to π. Because N is large this is a smooth phase distribution. Its peak shifts linearly with changes in the classical phase shift ϕ_c.

The calculated variance of the phase depends in general on the position of the mean phase relative to the 2π-wide window over which the distribution is defined [8,17]. In contrast, the dispersion does not depend on this choice of window [8]. For the examples given in this paper we will assume that the mean phase has been determined first, and the phase window has been shifted to place the mean in the center of the window (which we take to be $[-\pi,\pi]$), making the mean phase zero.

Consider the case of the equally split photon, where $C_{01}=C_{10}=1/2^{1/2}$, and the phase distribution is

$$P_\phi(\phi_n)=\frac{1}{N}[1+\cos(\phi_n-\phi_c)]. \tag{5.5}$$

The choice $\phi_c=0$ places the mean phase at zero. In the limit of many pixels the phase variance is found to be $\Delta\phi^2=\pi^2/3-2=0.131\pi^2$ and the dispersion squared is $D^2=\Delta C^2+\Delta S^2=\frac{1}{4}$. These results can be compared to those obtained using the eight-port detector: $\Delta\phi^2=\pi^2/8=0.125\pi^2$ and $D^2=\Delta C^2+\Delta S^2=\frac{3}{4}$ [2]. The difference in the phase variances from the two schemes results from the values of the higher-order moments of the cosine and sine variables, given in Table I. These in turn result from the detailed shape of the phase distributions, which are not the same in the two schemes. The distribution in the eight-port case can be inferred from the cosine and sine moments to be, for $\phi_c=0$,

$$P_{8p}(\phi)=\tfrac{1}{4}\delta(\phi+\pi/2)+\tfrac{1}{2}\delta(\phi)+\tfrac{1}{4}\delta(\phi-\pi/2), \tag{5.6}$$

which is very different than Eq. (5.5), as illustrated in Fig. 3.

C. Weak coherent states

Another interesting case is when weak coherent states, $|\alpha_S\rangle_S|\beta_R\rangle_R$, are incident in the signal and reference ports, where $\alpha_S=|\alpha_S|\exp(i\theta_S)$ and $\beta_R=|\beta_R|\exp(i\theta_R)$. In this special case, which has been treated semiclassically before [12], all pixels receive photons independently. The joint probabilities are then relatively simple. The probability for a single pixel (labeled f_a,j_a) to receive a photon is, from (3.1),

$$P(1_{f_a,j_a})=\frac{\overline{m}}{2N}\left[1+f_aV\cos\left[\frac{j_a2\pi}{N}+\phi_c\right]\right], \tag{5.7}$$

where $\overline{m}=\overline{n}cT/L$ is the mean number of photoelectrons detected, $\overline{n}=|\alpha_S|^2+|\beta_R|^2$ is the mean number of photons in the field, $V=2|\alpha_S\beta_R|/\overline{n}$ is the visibility of the fringe pattern, and $\phi_c=\theta_S-\theta_R$ is the classical phase difference of the coherent states. The sum of (5.7) over all pixels is equal to $\overline{m}$. The joint probability for exactly two pixels (labeled f_a,j_a and f_b,j_b) to each receive a photon is, from (3.2), (2.3), and (2.4),

$$\begin{aligned}P(1_{f_a,j_a},1_{f_b,j_b})&=\langle:\hat{N}_{f_a,j_a}\hat{N}_{f_b,j_b}:\rangle\\&=P(1_{f_a,j_a})P(1_{f_b,j_b}),\end{aligned} \tag{5.8}$$

showing that they are independent. We will neglect all events with more than two photons detected, since the field is assumed very weak. We will also neglect events in which a pair of photons hit pixels with equal pixel number, but on different arrays, since the number of pixels is assumed to approach infinity. Here we will adopt the approach of NFM in which all zero-count events are discarded from the data ensemble, with the rationale being that they provide no phase information. (How these zero counts may be incorporated in a consistent fashion will be discussed in Sec. VII.) Thus here we need to renor-

TABLE I. Properties of weak-field difference-phase distributions from a many-port scheme (present work, Sec. V) and an eight-port scheme (from Ref. [2]). Operational phase definitions are used, in which events containing no phase information (such as zero photon counts) are thrown out of the ensemble. Abbreviations are mean relative phase: $\langle\phi\rangle$; phase variance: $\Delta\phi^2=\langle(\phi-\langle\phi\rangle)^2\rangle$; cosine and sine: $C=\cos\phi$, $S=\sin\phi$; moments of sine and cosine: $\langle C^n\rangle,\langle S^n\rangle$. The dispersion squared is $D^2=\Delta C^2+\Delta S^2$, where $\Delta C^2=\langle(C-\langle C\rangle)^2\rangle$, etc. (We have assumed $cT/L=1$.)

State	Many-port detection	Eight-port detection
Equally split photon,	$\langle\phi\rangle=0$; $\Delta\phi^2=\pi^2/3-2=0.131\pi^2$	$\langle\phi\rangle=0$; $\Delta\phi^2=\pi^2/8=0.125\pi^2$
$\|0\rangle_S\|1\rangle_R+e^{i\phi_c}\|1\rangle_S\|0\rangle_R$	$D^2=\frac{3}{4}$	$D^2=\frac{3}{4}$
$\phi_c=0$	$\langle C^2\rangle=\langle S^2\rangle=\frac{1}{2}$	$\langle C^2\rangle=\langle S^2\rangle=\frac{1}{2}$
	$\langle C^2\rangle=\begin{cases}(n-1)!!/n!! & (n\text{ even})\\ n!!/(n+1)!! & (n\text{ odd})\end{cases}$	$\langle C^n\rangle=\frac{1}{2}$
	$\langle S^2\rangle=\begin{cases}(n-1)!!/n!! & (n\text{ even})\\ 0 & (n\text{ odd})\end{cases}$	$\langle S^n\rangle=\begin{cases}\frac{1}{2} & (n\text{ even})\\ 0 & (n\text{ odd})\end{cases}$
Equal, weak coherent states, $\|\alpha\rangle_S\|\alpha\rangle_R$. $\bar{n}=2\|\alpha\|^2$ is the mean photon number.	$\langle\phi\rangle=0$; $\Delta\phi^2=\pi^2/3-2-0.296\bar{n}$ $D^2=\frac{3}{4}-0.137\bar{n}$ $\langle C\rangle=\frac{1}{2}[1+\bar{n}(4/\pi-1)]$ $\langle S\rangle=0$	$\langle\phi\rangle=0$; $\Delta\phi^2$ is not known $D^2=\frac{3}{4}-0.083\bar{n}$ $\langle C\rangle=\frac{1}{2}$ $\langle S\rangle=0$

malize the resulting distributions by

$$\begin{aligned}1-P(0_a,0_b,\ldots,0_{2N})&=\sum_a P(1_{f_a,j_a},1_{f_b,j_b})\\&\quad+\sum_{a,b}P(1_{f_a,j_a})P(1_{f_b,j_b})\\&=\bar{m}+\bar{m}^2\ .\end{aligned}\tag{5.9}$$

Then the resulting phase distribution is given by

$$P_\phi(\phi_n)=[P_1(\phi_n)+P_2(\phi_n)]/(\bar{m}+\bar{m}^2)\ ,\tag{5.10}$$

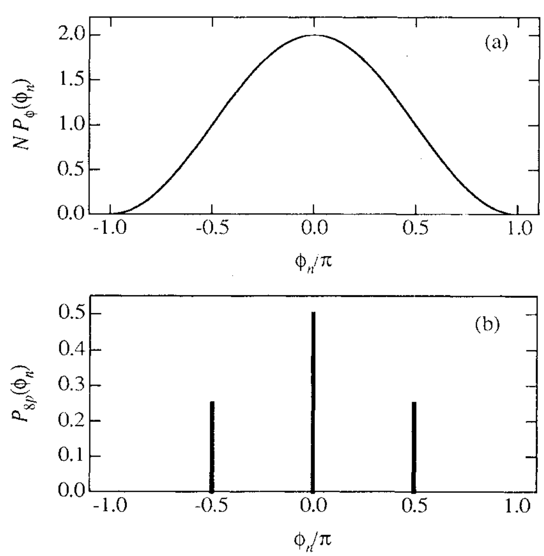

FIG. 3. Relative-phase distributions for the equally split photon state obtained in (a) the present many-port scheme, Eq. (5.5), normalized by the number of detector elements N; (b) the eight-port scheme of Ref. [2], Eq. (5.6), with height of bar representing weight of δ function.

where $P_1(\phi_n)$ is the contribution from single-count observations and $P_2(\phi_n)$ is the contribution from double counts.

The calculation for the single counts is identical to that leading to (5.4), and yields

$$P_1(\phi_n)=\frac{\bar{m}}{N}[1+V\cos(\phi_n-\phi_c)]\ ,\tag{5.11}$$

where $\phi_n=n2\pi/N$ and n ranges from $-M$ to $+M$ by integers, taking on N values. For double-count events the DFT takes on values given by

$$K(1)=f_a\exp\left[-\frac{ij_a2\pi}{N}\right]+f_b\exp\left[-\frac{ij_b2\pi}{N}\right]\ .\tag{5.12}$$

The probability for each double-count event is given by (5.8). Inferring the phase from the DFT requires consideration of several cases, depending on whether both photons hit the same array or each hits a different array. The calculation, described in the Appendix, yields for the double-count phase distribution

$$\begin{aligned}P_2(\phi_n)=\frac{\bar{m}^2}{N}\Bigg[&1+\frac{4V}{\pi}\cos(\phi_n-\phi_c)\\&+\frac{V^2}{2}\cos(2\phi_n-2\phi_c)\Bigg]\ .\end{aligned}\tag{5.13}$$

When the single- and double-count events are combined, according to (5.10), the resulting phase distribution is given by

$$\begin{aligned}P_\phi(\phi_n)=\frac{1}{N}\Bigg\{&1+\left[1+\bar{m}\left[\frac{4}{\pi}-1\right]\right]V\cos(\phi_n-\phi_c)\\&+\frac{1}{2}\bar{m}V^2\cos(2\phi_n-2\phi_c)\Bigg\}\ ,\end{aligned}\tag{5.14}$$

where we made use of the smallness of $\overline{m}$. The last term, arising from double counts, has the effect of narrowing the distribution slightly, compared with the first two terms. Note that when the visibility is unity this phase distribution retains a large depth of modulation, resulting in a smaller variance than obtained with a uniform distribution.

Again assuming that the mean value of phase is chosen to be zero by shifting the window, the variance is calculated, yielding

$$\Delta\phi^2=\frac{\pi^2}{3}-2\left[1+\overline{m}\left[\frac{4}{\pi}-1\right]\right]V+\frac{\overline{m}V^2}{4}\ , \tag{5.15}$$

which for equal-amplitude beams ($V=1$) is $\Delta\phi^2=\pi^2(0.131-0.030\overline{m})$. The dispersion squared is

$$D^2=1-\frac{1}{4}\left[1+2\overline{m}\left[\frac{4}{\pi}-1\right]\right]V^2\ , \tag{5.16}$$

which for $V=1$ is $D^2=\frac{3}{4}-0.137\overline{m}$. The mean values of cosine and sine are given by

$$\langle\cos(\phi_n)\rangle=\frac{1}{2}\left[1+\overline{m}\left[\frac{4}{\pi}-1\right]\right]V\cos(\phi_c)\ ,$$
$$\langle\sin(\phi_n)\rangle=\frac{1}{2}\left[1+\overline{m}\left[\frac{4}{\pi}-1\right]\right]V\sin(\phi_c)\ . \tag{5.17}$$

Comparing to the analogous result for the dispersion in the eight-port scheme given in Table I, we see that the dispersion in the present many-port scheme is smaller, indicating a slightly better average phase measurement when using many ports.

VI. MAXIMUM-LIKELIHOOD PHASE ESTIMATES

As pointed out by Walkup and Goodman, Eq. (2.12) for the phase inferred from a given set of pixel counts is not necessarily the best estimate in the sense of maximum-likelihood theory [12]. If the state of the fields is known *a priori* to within a few unknown parameter values, then the optimum phase estimator can be found as follows. This method was developed for the case of coherent states (or other states leading to Poisson photoelectron statistics) by Walkup and Goodman. Denote the set of unknown parameters by $\{\xi_k\}$, and the joint probability of pixel counts, given in (3.2) by $P(\{1_{f,j}\}|\{\xi_k\})$, which is conditional on the parameter values. The question is, given an observed set of pixel counts ("data set"), what is the most likely set of parameters that would give rise to that result? The parameter of most interest to us is the phase. In this formulation the phase must be understood as a parameter that specifies the state of the field, such as a classical phase shift $\theta_S-\theta_R$, rather than a more abstract quantity such as the Pegg-Barnett Hermitian-phase operator [17].

To find the most probable parameter set, differentiate the joint probability to solve for the parameter set that jointly maximizes the probability of observing that particular data set:

$$\frac{\partial}{\partial\xi_k}P(\{1_{f,j}\}|\{\xi_k\})=0,\quad k=1,2,3,\ldots\ . \tag{6.1}$$

As an example, consider the case of the split photon, of Sec. V B, where the conditional probability is given by (5.2). Then Eq. (6.1), with ϕ_c identified as ξ_1, gives

$$\frac{\partial}{\partial\phi_c}P(1_{f_a,j_a}|\phi_c)=-\frac{f_a}{N}C_{01}C_{10}\sin\left[\frac{j_a2\pi}{N}+\phi_c\right]=0\ . \tag{6.2}$$

If the (positive) coefficients C_{01},C_{10} are known through knowledge of the beam splitter used to split the one-photon state, then (6.2), along with the additional condition that the second derivative must be negative to ensure a maximum rather than a minimum, uniquely determines the maximum-likelihood estimate of ϕ_c. This gives $\phi_c=-j_a2\pi/N+(\pi/2)(1-f_a)$, exactly the same as inferred in Sec. V using the complex argument of $K(1)$. So for the split-photon state (or any other single-count observation) the complex argument of $K(1)$ is the maximum-likelihood estimator.

Another example is the pair of weak coherent states, with known values of mean photon numbers. The maximum-likelihood-phase calculation for single-photon counts is identical to that for the split photon, and again leads to the same phase as inferred by using the complex argument of $K(1)$. The case of double counts is more interesting. The needed conditional probability is best expressed in the form in (A4) in the Appendix, where the pixel numbers are relabeled using $S=(j_a+j_b)\pi/N$ and $D=(j_a-j_b)\pi/N$. The needed derivative is then

$$\begin{aligned}\frac{\partial}{\partial\phi_c}P(1_{f_a,j_a}|\phi_c)\propto&-(f_a+f_b)V\sin(S+\phi_c)\cos D\\&-(f_a-f_b)V\cos(S+\phi_c)\sin D\\&-f_af_bV^2\sin(2S+2\phi_c)=0\ .\end{aligned} \tag{6.3}$$

This relation, along with the condition that the second derivative must be negative, determines the maximum-likelihood estimate of ϕ_c. Unfortunately the equation cannot always be solved in closed form. In the special case that the fringe visibility is small, the third term can be neglected, and then the solutions are found to be exactly the same as given in (A2a)–(A2d), which were inferred by using the complex argument of $K(1)$. The other case that can be solved is when the visibility V equals 1. Then additional solutions to (6.3) are found, but they are found to correspond to minimum (actually zero) probability, rather than maximum. Again the maximum probabilities can be shown to agree with those phases given in (A2a)–(A2d). In these two limits, small or large visibility, the complex argument of $K(1)$ is the maximum-likelihood estimator in the case of double counts.

In general, however, the complex argument of $K(1)$ is not always the best phase estimate. This presents the possibility to refine the phase inference by solving (6.1) numerically on each measurement. This might be useful in experimental data analysis, although the theoretical predictions in this case are not easy to obtain.

Another useful consequence of adopting a maximum-likelihood interpretation of the phase measurement is that it provides a means to incorporate consistently the results of the zero counts (those events where no photon was observed in the integration window) into the measurement of the phase distribution. Following the above logic, the probability of obtaining a zero count is independent of the phase; therefore from a zero-count event one can only conclude that all phase values are equally likely—all phase values correspond to a solution of (6.1) in this case. In an experiment, then, one is led to assign arbitrarily some random value of the phase as being the result of measurement—a uniform random number generator can be used for this purpose. For example, if many zero counts occurred, with no other events being present, a uniform distribution would be built up, and we would conclude, correctly, that the phase distribution of the vacuum state is uniform. In this interpretation the observation of zero photons in a prescribed time is just as meaningful as the observation of any other number of photons. This idea will be exploited in the next section.

VII. COMPARISON TO OTHER QUANTUM PHASE DISTRIBUTIONS

In considering the relative-phase distribution (5.14) for two weak, equal-amplitude coherent states (in which case $V=1$), shown in Fig. 4(a), one might question the origin

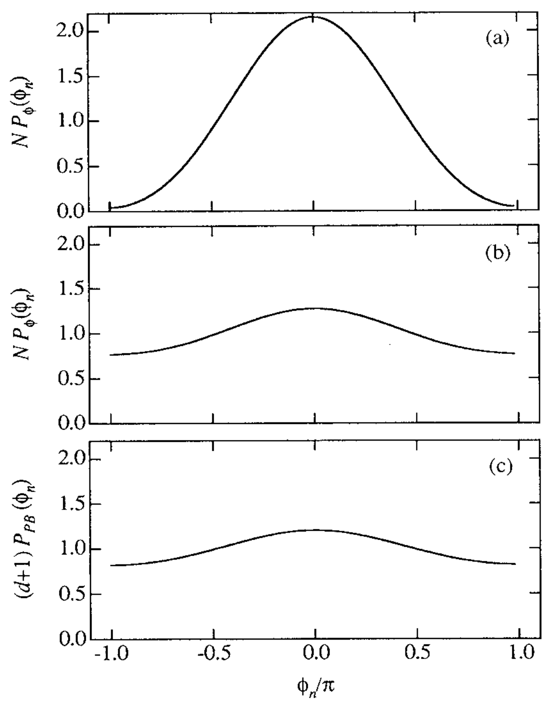

FIG. 4. Relative-phase distributions for equal-amplitude, weak coherent states obtained in (a) the many-port scheme with zero counts discarded [Eq. (5.14), normalized by the number of detector elements]; (b) the many-port scheme with zero counts included [Eq. (7.2)]; (c) the two-mode Pegg-Barnett distribution [Eq. (7.6), normalized by the dimension of the truncated Hilbert space, $d+1$]. Parameter values are $V=1$, and $\bar{m}=\bar{n}=2|\alpha_S|=2|\beta_R|=0.2$.

of the large depth of modulation of this function, and its resulting small variance, in contrast to that obtained with some other quantum phase distributions—the marginal Wigner distribution [7], the marginal Q distribution [Eq. (4.4)] [5,7,9], the SG-POM phase distribution [16], or the Pegg-Barnett Hermitian-phase operator distribution [17]. In particular, these distributions become nearly uniform as the states approach being in the vacuum state (as $\bar{m}$ approaches zero). For example, the marginal Q-function phase distribution for a strong reference field and a weak coherent signal field was derived in (4.6), and was seen to be nearly uniform. This occurs because the Q function for the signal state subtends a large angular width from the origin (in X,Y phase space), due to the large component of the vacuum state in a coherent state with less than one photon on average. Based on this picture, it would be expected that the relative-phase distribution between *two* weak coherent fields would be even less well defined than that for one weak and one strong field.

In contrast to this expectation, (5.14) retains full modulation in the limit of two weak fields. The origin of this behavior is in the discarding of the observations that yielded zero counts, corresponding to the vacuum, [or more generally, $K(p_0)=0$] with the rationale that these cases give no information about the classical phase shift.

By retaining these zero-count cases in the data analysis it is possible to obtain experimentally a phase distribution that is similar to the above-mentioned quantum phase distributions. Simply replace (5.10) by

$$P_\phi(\phi_n)=P_0(\phi_n)+P_1(\phi_n)+P_2(\phi_n)\ , \tag{7.1}$$

where the inferred phase distribution from the zero counts is uniform, consistent with no information gained, i.e., $P_0(\phi_n)=[1-(\bar{m}+\bar{m}^2)]/N$, and the single-count and double-count distributions are given in (5.11) and (5.13). Equation (7.1) yields (with $cT=L$)

$$P_\phi(\phi_n)=\frac{1}{N}\left\{1+2|\alpha_S\beta_R|\left[1+\left[\frac{4}{\pi}\right](|\alpha_S|^2+|\beta_R|^2)\right]\times\cos(\phi_n-\phi_c)+2|\alpha_S\beta_R|^2\cos(2\phi_n-2\phi_c)\right\}\ , \tag{7.2}$$

to fourth order in the amplitudes. As seen in Fig. 4(b), this distribution becomes nearly constant in the limit of weak fields, in contrast to that in (5.14), Fig. 4(a).

To check the degree of similarity of (7.2) to the Pegg-Barnett distribution, it is necessary to define a two-mode relative-phase distribution based on the PB phase eigenstates, defined for the signal (S) mode and reference (R) modes, as [17,24]

$$\begin{aligned}|\phi_{S,k}\rangle_S&=\frac{1}{(d+1)^{1/2}}\sum_{n=0}^{d}\exp(in\phi_{S,k})|n\rangle_S\ ,\\ \phi_{S,k}&=\phi_{S,0}+k2\pi/(d+1)\ ,\\ |\phi_{R,l}\rangle_S&=\frac{1}{(d+1)^{1/2}}\sum_{m=0}^{d}\exp(im\phi_{R,l})|m\rangle_R\ ,\\ \phi_{R,l}&=\phi_{R,0}+l2\pi/(d+1)\ ,\end{aligned} \tag{7.3}$$

where $k,l=0,1,\ldots,d$, and $d+1$ is the dimension of the Hilbert space (which goes to infinity after the calculation), and $|n\rangle_S, |m\rangle_R$ are number states for signal and reference modes, respectively. While the difference phase $\phi_{S,l}-\phi_{R,m}$ can take on values between -2π and 2π, values greater than π or less than $-\pi$ can be mapped back into the $[-\pi,\pi]$ interval to give a distribution normalized over this interval. Given a two-mode state $|\Psi\rangle\rangle$, the probability for the difference phase to equal a value ϕ_n is given by [24]

$$P_{\rm PB}(\phi_n)=\sum_{k,l=0}^{d} |_S\langle\phi_{S,k}|_R\langle\phi_{R,l}|\psi\rangle\rangle|^2 \times\delta(\phi_n, M(\phi_{S,k}-\phi_{R,l}))\ , \tag{7.4}$$

where $\delta[i,j]$ represents the Kronecker delta, $\delta(i,j)=\delta_{ij}$, and $M(x)$ stands for an operation that maps x into the $[-\pi,\pi]$ interval, i.e.,

$$M(x)=\begin{cases} x+2\pi, & -2\pi<x<-\pi \\ x, & -\pi<x<\pi \\ x-2\pi, & \pi<x<2\pi\ . \end{cases} \tag{7.5}$$

In calculating (7.4) it is useful to note that the periodicity of the states in (7.3) means that the M operation can be formally dropped in (7.4) and the correct result is easily obtained. For the two-mode coherent state $|\psi\rangle\rangle=|\alpha_S\rangle_S|\beta_R\rangle_R$ this gives, to fourth order in the amplitudes,

$$P_{\rm PB}(\phi_n)=\frac{1}{d+1}\left\{1+2|\alpha_S\beta_R|\left[1-\left[1-\frac{1}{2^{1/2}}\right](|\alpha_S|^2+|\beta_R|^2)\right]\cos(\phi_n-\phi_c)+|\alpha_S\beta_R|^2\cos(2\phi_n-2\phi_c)\right\}. \tag{7.6}$$

This result, shown in Fig. 4(c), is quite similar to (7.2) that is obtained using the many-port detector, but there are significant differences, in particular the factor 2 multiplying the last term in (7.2). We conclude that the many-port scheme can measure a phase distribution which has qualitative features similar to the above-mentioned quantum distributions, but that it does not measure precisely the two-mode PB distribution.

There is no Hermitian-phase operator corresponding to the distribution (5.14) because the operation of $\hat{\phi}$ in (2.16) is not defined for all states. Similarly, there does not appear to be a Hermitian-phase operator corresponding to the distribution in (7.2). An operator $\hat{O}$ can be defined via its action on the two-mode vacuum $|0\rangle\rangle$ (signal and reference) and on all other states,

$$\hat{O}|\psi\rangle\rangle=\begin{cases} e^{ir}|\psi\rangle\rangle, & |\psi\rangle\rangle=|0\rangle\rangle \\ \dfrac{\hat{K}_X+i\hat{K}_Y}{(\hat{K}_X^2+\hat{K}_Y^2)^{1/2}}|\psi\rangle\rangle, & \langle\langle 0|\psi\rangle\rangle=0\ , \end{cases} \tag{7.7}$$

where r is a random number, uniform on $[-\pi,\pi]$, and it is implicit that all modes other than signal and reference are in the vacuum. If this operator were unitary, i.e., $\hat{O}^\dagger\hat{O}|\psi\rangle\rangle=\hat{O}\hat{O}^\dagger|\psi\rangle\rangle=|\psi\rangle\rangle$ for any state, even those with vacuum components, then a phase operator $\hat{\phi}$ that is Hermitian over the entire Hilbert space could be defined via $\hat{O}=e^{i\hat{\phi}}$. This operator is not unitary, however, since each time it operates it returns a different random value of r, which means $\hat{O}^\dagger\hat{O}|\psi\rangle\rangle=e^{-ir'}e^{ir}|\psi\rangle\rangle\neq|\psi\rangle\rangle$. Therefore we may not associate (7.2) with a Hermitian-phase operator; rather (7.2) and (5.14) both must be viewed as resulting from "operational" definitions of phase, analogous to that employed by NFM. Hradil has considered a similar definition, but with r being a nonrandom number [8]. In this case $\hat{O}$ in (7.7) is Hermitian, but the physical interpretation is not clear.

VIII. CONCLUSIONS

We have found that the many-port homodyne detector, which has been implemented (in the strong-field case) using a beam splitter and two array detectors [11], has in principle several favorable features as compared to the eight-port detector [1,2]. Because of its many possibilities for measurement outcomes it provides a quasicontinuous relative-phase distribution, even in the case of very weak fields. We have described a data reduction method, the discrete Fourier transform, that provides a definition of a (non-Hermitian) relative-phase operator (2.16) that is an alternative to that defined by NFM [2]. In the case of a strong, coherent reference field, the DFT effectively eliminates the noise associated with all vacuum modes excepting one (the image mode). The coupling in of this mode is intimately connected with the simultaneous measurement of two quadrature amplitudes, which is necessary for phase measurement.

In the case that both signal and reference fields are weak, all of the vacuum-mode operators must be retained, as they are responsible for the random partitioning of the photons among the many detectors that produces the quasicontinuous distributions. If the zero-count cases are thrown out of the data set then phase distributions with large modulation are produced, even in cases where other quantum phase distributions, such as the Pegg-Barnett or Wigner, show very small modulation depth. We point out that the smallness of the modulation of these other distributions is connected with the contribution from the vacuum component of the state. Furthermore if the many-port data set is reanalyzed in a way that retains the uniform contributions from zero counts, the many-port phase distribution takes on a character that resembles the above-mentioned quantum phase distributions. Of course, whether or not to retain these zero-count contributions is a matter of choice dictated by the type of information one is trying to obtain about the field. If one wants to measure a phase distribution similar to the marginal Wigner or PB distributions, one should retain the zero counts. For example, this method can distinguish between a split-photon state and a split-photon-plus-vacuum state, i.e., $C_{00}|0\rangle_S|0\rangle_R+C_{01}|0\rangle_S|1\rangle_R+C_{10}|1\rangle_S|0\rangle_R$. If one wants to make the best estimate of the mean phase (determined by some classical phase shifter) then it might be better to throw out the zero counts since they contain no relevant information in this context.

In either case, discarding or retaining the zero counts, it does not appear to be possible to define a Hermitian-phase operator crresponding to the measured distribution. Our scheme measures an operational phase distribution, similar in spirit to that defined by NFM.

We have shown that in the case of single counts on arbitrary states the operator (2.16) for relative phase is the best estimator in the maximum-likelihood sense. The same result was found for double counts on states with low visibility or with unity visibility. A method for obtaining the best estimate for arbitrary states was found, although it may not be easy to carry out in all cases.

Most of the above discussion is based on the assumption that we have a nearly infinite ensemble of measurements on identically prepared systems for study. In the case that we have a small number of measurements possible (due, for example, to the state existing for only a short time, as in a gravity-wave detector) the question arises whether the many-port scheme is more useful for estimating a classical phase shift than is the eight-port scheme. If the variance of the distributions is taken as a measure of the typical error in such an estimate from a single measurement, then the many-port and eight-port schemes are seen (Table I) to be roughly equivalent.

The present analysis assumes detector arrays with negligible internal noise and quantum efficiency approaching unity. While current technology does not provide these ideal properties, recent advances in charge-coupled device (CCD) detectors are impressive. With back-illumination of thinned CCD materials with antireflection coatings, commercially available detectors have quantum efficiency up to 80% in the 650–850 nm wavelength region and root-mean-square noise per pixel readout as low as about four electrons. (High quantum efficiency and low noise are not achievable simultaneously, as they depend differently on temperature.) Thus while these detectors are not quite good enough to carry out the experiments described here with very weak fields, they approach the quantum regime where considerations of the type considered here begin to be important. It is also worth noting that quantum effects can be seen in photodetection even when many photons are detected; see, for example, [25], where one million photons produced by parametric amplification show nonclassical correlation effects. Phase measurements in this regime could be made with array detectors now available. On the other hand, in this many-photon regime we have found that essentially the same results are obtained using the eight-port scheme, which can be implemented using less expensive photodetectors.

ACKNOWLEDGMENTS

We wish to acknowledge several helpful discussions with W. Schleich. This research was supported by the National Science Foundation, Grant No. PHY 8921709-01.

APPENDIX: DOUBLE-COUNT EVENTS FOR COHERENT STATES

To find the phase distribution arising from the subensemble of double-count events it is useful to rewrite (5.12) as

$$K(1)=2\cos\left[D+\frac{(f_a-f_b)\pi}{4}\right]\times\exp\left[-i\left[S+\frac{(f_a+f_b-2)\pi}{4}\right]\right], \tag{A1}$$

where $S=(j_a+j_b)\pi/N$ and $D=(j_a-j_b)\pi/N$. Then careful consideration of four separate cases shows that the complex argument ϕ of $K(1)$ is given by the following.

For $f_a=f_b=1$,

$$\phi=\begin{cases}-S & \text{if } \cos D>0\\ -S-\pi & \text{if } \cos D<0\ .\end{cases} \tag{A2a}$$

For $f_a=f_b=-1$,

$$\phi=\begin{cases}-S-\pi & \text{if } \cos D>0\\ -S & \text{if } \cos D<0\ .\end{cases} \tag{A2b}$$

For $f_a=-f_b=1$,

$$\phi=\begin{cases}-S-\pi/2 & \text{if } \sin D>0\\ -S-3\pi/2 & \text{if } \sin D<0\ .\end{cases} \tag{A2c}$$

For $-f_a=f_b=1$,

$$\phi=\begin{cases}-S-3\pi/2 & \text{if } \sin D>0\\ -S-\pi/2 & \text{if } \sin D<0\ .\end{cases} \tag{A2d}$$

If the variables S and D are treated as continuous variables, bounded by $-\pi$ and π, then in the limit $N\to\infty$ the distribution for phase is given by an average over a δ function,

$$P_2(\phi')=\sum_{f_a,f_b}2\int dS\,dD\,P_{f_a,f_b}(S,D)\delta(\phi'-\phi)\ , \tag{A3}$$

where ϕ is given by the four cases above in Eqs. (A2a)–(A2d) and the joint probability for obtaining values of f_a,f_b and S,D is given by (5.8) [with (5.7)], which is rewritten as

$$P_{f_a,f_b}(S,D)=\left[\frac{\overline{m}}{4\pi}\right]^2\{1+(f_a+f_b)V\cos(S+\phi_c)\cos D$$
$$-(f_a-f_b)V\sin(S+\phi_c)\sin D+\tfrac{1}{2}f_af_bV^2[\cos(2S+2\phi_c)+\cos(2D)]\}\ . \quad \text{(A4)}$$

The 2 in (A3) results from a Jacobian and the integral is over the domain

$$\int_{-\pi+D}^{\pi-D}dS\int_0^{\pi}dD+\int_{-\pi}^{0}dD\int_{-\pi-D}^{\pi+D}dS\ . \quad \text{(A5)}$$

A fairly tedious integration of (A3) leads to (5.13) after the phase is converted back to a discrete variable. Note that in the integration the regions corresponding to pairs of photons hitting pixels with equal pixel number (i.e., $D=0$) are of measure zero and need not be discarded explicitly.

*Electronic address: raymer@oregon.uoregon.edu

†Permanent address: Joint Institute for Laboratory Astrophysics and the Department of Physics, University of Colorado and the National Institute of Standards and Technology, Boulder, CO 80309.

[1] N. G. Walker and J. E. Carroll, Opt. Quantum Electron. **18**, 355 (1986); N. G. Walker, J. Mod. Opt. **34**, 16 (1987).

[2] J. W. Noh, A. Fougères, and L. Mandel, Phys. Rev. Lett. **67**, 1426 (1991); Phys. Rev. A **45**, 424 (1992); **46**, 2840 (1992).

[3] V. B. Braginsky and F. Y. Khalili, *Quantum Measurement* (Cambridge University Press, Cambridge, England, 1992).

[4] E. Arthurs and J. L. Kelly, Jr., Bell Syst. Tech. J. **44**, 725 (1965).

[5] J. H. Shapiro and S. S. Wagner, IEEE J. Quantum Electron. **QE-20**, 803 (1984).

[6] Y. Lai and H. A. Haus, Quantum Opt. **1**, 99 (1989).

[7] M. Freyberger and W. Schleich, Phys. Rev. A **47**, R30 (1993); M. Freyberger, K. Vogel, and W. Schleich, Quantum Opt. **5**, 65 (1993); Phys. Lett. A **176**, 41 (1993); U. Leonhardt and H. Paul, Phys. Rev. A **47**, 2460 (1993).

[8] Z. Hradil, Quantum Opt. **4**, 93 (1992).

[9] W. Schleich, A. Bandilla, and H. Paul, Phys. Rev. A **45**, 6652 (1992).

[10] D. T. Smithey, M. Beck, A. Faridani, and M. G. Raymer, Phys. Rev. Lett. **70**, 1244 (1993); D. T. Smithey, M. Beck, J. Cooper, M. G. Raymer, and A. Faridani, Phys. Scr. T **48**, 35 (1993).

[11] D. T. Smithey, M. Belsley, K. Wedding, and M. G. Raymer, Phys. Rev. Lett. **67**, 2446 (1991); M. Belsley, D. T. Smithey, K. Wedding, and M. G. Raymer, Phys. Rev. A **48**, 1514 (1993).

[12] J. F. Walkup and J. W. Goodman, J. Opt. Soc. Am. **63**, 399 (1973).

[13] D. Walls, Nature (London) **306**, 141 (1983); L. Mandel, Phys. Scr. **T12**, 34 (1986).

[14] W. Schleich, R. J. Horowicz, and S. Varro, Phys. Rev. A **40**, 7405 (1989).

[15] G. S. Agarwal, S. Chaturvedi, K. Tara, and V. Srinivasan, Phys. Rev. A **45**, 4904 (1992).

[16] J. H. Shapiro and S. R. Shepard, Phys. Rev. A **43**, 3795 (1991).

[17] D. T. Pegg and S. M. Barnett, Phys. Rev. A **39**, 1665 (1989).

[18] M. Beck, D. T. Smithey, J. Cooper, and M. G. Raymer, Opt. Lett. **18**, 1259 (1993); M. Beck, D. T. Smithey, and M. G. Rayner, Phys. Rev. A **48**, R890 (1993).

[19] Z. Y. Ou, C. K. Hong, and L. Mandel, Opt. Commun. **63**, 118 (1987); Z. Y. Ou and L. Mandel, Am. J. Phys. **57**, 66 (1989).

[20] We would also expect the present results to apply to the case where both fields are in the form of short pulses, as long as their temporal profiles are the same and their phases do not vary over their duration. See, for example, Ref. [11]; and S. J. Kuo, D. T. Smithey, and M. G. Raymer, Phys. Rev. A **43**, 4083 (1990).

[21] H. P. Yuen and V. W. S. Chan, Opt. Lett. **8**, 177 (1983).

[22] P. D. Drummond and S. J. Carter, J. Opt. Soc. Am. B **4**, 1565 (1987).

[23] R. J. Glauber, Phys. Rev. **130**, 2529 (1963); **131**, 2766 (1963); P. L. Kelly and W. H. Kleiner, *ibid.* **136**, A316 (1964).

[24] S. M. Barnett and D. T. Pegg, Phys. Rev. A **42**, 6713 (1990).

[25] D. T. Smithey, M. Beck, M. Belsley, and M. G. Raymer, Phys. Rev. Lett. **69**, 2650 (1992).