Proposal for a loophole-free Bell inequality experiment

Paul G. Kwiat,[1] Philippe H. Eberhard,[2] Aephraim M. Steinberg,[1] and Raymond Y. Chiao[1]
[1]*Department of Physics, University of California, Berkeley, California 94720*
[2]*Lawrence Berkeley Laboratory, University of California, Berkeley, California 94720*
(Received 2 August 1993)

The nonlocality inherent in quantum mechanics has been tested experimentally, but the nonlocal interpretation of all the relevant experiments can be challenged. All past tests of Bell's inequalities have required supplementary assumptions, without which the experimental results satisfy the inequalities, in agreement with the notion of local realism. There are basically three loopholes that need to be closed; neither the cascade sources nor the parametric down-conversion sources of correlation photons used to date are capable of closing all of them, even if the detectors used had been 100% efficient. We propose a two-crystal down-conversion source, relying on type-II collinear phase matching, which should permit a violation of Bell's inequalities without the need for supplementary assumptions. As the source can produce a true singletlike state, it is also relevant for quantum cryptographic applications.

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

I. INTRODUCTION

It is now well known that quantum mechanics (QM) yields predictions which are inconsistent with the seemingly innocuous concepts of locality and reality. This was first shown by Bell in 1964 [1,2] for the case of two quantum-mechanically "entangled" particles, e.g., particles in a singlet state, which do not possess definite polarizations (or spins) even though they are always orthogonally polarized. As implied by Einstein, Podolsky, and Rosen (EPR) [3], it is straightforward to construct local realistic models that explain certain features predicted by QM (e.g., the total anticorrelation between detectors measuring the same polarization component of the two particles). The quantum-mechanical contradiction with local realism becomes apparent only by considering situations of nonperfect correlations (i.e., measuring the polarization components at intermediate, nonorthogonal angles). More recently, Greenberger, Horne, and Zeilinger [4] and Mermin [5] have shown that QM and local realism are incompatible even at the level of *perfect* correlations, for certain states of three or more particles. Hardy has also presented a clever gedanken experiment using electron-positron annihilation to achieve a contradiction with local realism without the need for inequalities [6] and has recently proposed an optical analog which may allow a feasible experimental implementation [7]. Violations of Bell's inequalities with macroscopic (but nonclassical) states of light have been discussed by Munro and Reid [8] and Franson [9]. Unfortunately, none of these ingenious extensions and generalizations of the work of Bell reduces the experimental requirements for a completely unambiguous test. In fact, all of them seem to mandate even *stronger* constraints on any real experiment than the original two-particle inequalities [10]. One exception is the recent discovery that the detection efficiency requirement can be reduced by employing a state of two particles that are not maximally entangled, i.e., with an unequal superposition of the two terms [11,12].

The situation for experimental tests has seen somewhat less growth. While tests of Bell's inequalities have been extended to new systems, some relying on energy-time or phase-momentum entanglement [13–16], to date *no incontrovertible violation of Bell's inequalities has been observed.* All experiments thus far have required supplementary assumptions (in addition to that of local realism), which although seemingly reasonable, severely reduce the true impact such an experiment might yield. These additional assumptions constitute several loopholes, which can be divided into three general categories: the angular-correlation loophole, the detection loophole, and the spacelike separation loophole. We propose here a setup which should permit for the first time (simultaneous) closure of the first two of these loopholes; we will also discuss briefly how current technologies should allow an extension to close the third as well [17,18]. The source may also find application in quantum cryptography [19,20], as it doubles the signal-to-noise ratio of most previous down-conversion EPR schemes.

In Sec. II we briefly summarize the three experimental loopholes and how they have been manifested in experiments to date. Our proposed source is described in Sec. III, while several potential problems arising from various phase-matching considerations are discussed in Sec. IV, along with an experimental modification to mitigate these problems. The benefits of utilizing a nonmaximal entanglement are presented in Sec. V. Section VI evaluates the detrimental effects of imperfect optical elements. A final analysis and conclusion are given in Sec. VII. A somewhat thorough calculation allowing for various sorts of entangled states, as well as for nonideal polarizing beam splitters, is presented in the Appendix.

II. EXPERIMENTAL LOOPHOLES

The angular correlation loophole was investigated by Clauser and Horne [21] and in detail by Santos [22].

1050-2947/94/49(5)/3209(12)/$06.00

They showed that because of the cosine-squared angular correlation of the directions of photons emitted in an atomic cascade (of the type used in nearly all of the early Bell-inequalities experiments [2,23]), there was an inherent polarization decorrelation, due to the transversality condition. More explicitly, since the photons do not necessarily fly off back-to-back in this three-body decay process, one must detect the photons emitted into a large solid angle in order to have a sufficiently high detection efficiency (see discussion below). But with this source the very polarization correlation which could result in a violation of Bell's inequalities is reduced for noncollinear photons, so that it is strictly impossible to disprove local realism by using a cascade source.

In order to essentially remove the angular-correlation problem, experimenters have switched from cascade sources to those using correlated photons produced in the process of spontaneous parametric down-conversion [14–16,24,25]. These photons can have an angular correlation of better than 1 mrad, although in general they need not be collinear [26]. In the simplest of the down-conversion Bell-inequality experiments [24,25], noncollinear correlated photons were directed through equal path lengths to opposite sides of a 50-50 beam splitter, aligned so that the transmitted mode of one photon coincided with the reflected mode of the conjugate photon, and vice versa (cf. Fig. 1 of [25]). A half-wave plate prior to the beam splitter was used to rotate the polarization of one of the photons (which were initially horizontally polarized) by 90°. The output state of this source (including the down-conversion crystal, wave plate, and beam splitter) was then

$$\begin{aligned}|\Psi\rangle &\approx \tfrac{1}{2}[|H\rangle_3+i|H\rangle_4][i|V\rangle_3+|V\rangle_4]\\ &=\tfrac{1}{2}[|H\rangle_3|V\rangle_4-|V\rangle_3|H\rangle_4\\ &\qquad+i|V\rangle_3|H\rangle_3+i|V\rangle_4|H\rangle_4]\;, \end{aligned} \tag{1}$$

where the subscripts 3 and 4 denote the two output port modes of the beam splitter and $|H\rangle_j$ ($|V\rangle_j$) denotes a single photon in mode j, horizontally (vertically) polarized. Coincidence rates between detectors looking at the two output ports were recorded as a function of the orientation of polarizers at the detectors. In only measuring coincidence rates, the experiments were able to effectively create a singletlike state by discarding the last two terms of (1). However, it should be stressed that because of these discarded terms, the detection efficiency is inherently limited to 50% [unless the detector can differentiate one photon from two (as could be done if high-efficiency single-photon image amplifiers were available)], and no indisputable test of Bell's inequalities is possible [27]. Similar problems arise in the Bell-inequality experiments based on energy-time entanglement [28] and phase-momentum entanglement [29].

Low detection efficiencies have hindered experiments on Bell's inequalities from the outset. The detection loophole basically deals with the fact that with nonunity-efficiency detectors (for simplicity, we "lump" all losses of our interfering particles into the effective efficiency of the detectors [30]), only a fraction of the emitted correlated pairs is detected. If the efficiency is sufficiently low, then it is possible for the subensemble of detected pairs to give results in agreement with quantum mechanics, even though the *entire* ensemble satisfies Bell's inequalities. Due to the nonexistence of adequate detectors, experiments have so far employed an additional assumption, equivalent to the fair-sampling assumption that the fraction of detected pairs is representative of the entire ensemble. However, it should be noted that physicists have invented local, realistic models which can explain all of the observed results [21,22]—these models necessarily violate the fair-sampling-type assumptions. In order to experimentally close this loophole, one must have detectors with sufficiently high single-photon detection efficiencies. Formerly, it was believed that $\sim 83\%$ ($=2\sqrt{2}-2$) was the lower efficiency limit. However, one of us (P.H.E.) has shown that by using a nonmaximally entangled state (i.e., one where the magnitudes of the probability amplitudes of the contributing terms are not equal), one may reduce the detector requirement to $\sim 67\%$, in the limit of no background [11]. Recently, we have measured the absolute single-photon detection efficiency of several detectors [31,32] and observed corrected efficiencies as high as 75%. However, there were additional losses (not corrected for in the above result), which should in principle be avoidable, leading us to believe that efficiencies in excess of 90% may be feasible; efforts toward this end are currently in progress. Combined with our proposal for a down-conversion source in which there is no need to reject half the counts, this should allow us to close the detection and the angular-correlation loopholes simultaneously.

The final loophole concerns the spacelike separation of the different parts of the experiment. Clearly, no claims about nonlocality can be made if the predetector analyzers are varied so slowly that a signal traveling at the speed of light could carry the analyzer-setting information back to the source or to the other analyzer before a pair was produced or detected. [In standard practice, the detectors are positioned on opposite sides of the source, with the correlated particles directed accordingly (e.g., by mirrors, if necessary); then sufficient separation of the source and detectors guarantees sufficient separation between the detectors.] To close this loophole, the analyzers' settings should be changed *after* the correlated pair has left the source [33]. Only one Bell-type experiment, that of Aspect *et al.* [23], has made any attempt at all to address this locality condition, but even in that experiment the loophole remains. Although the experiment used rapidly varying analyzers, the variation was not random, and it has been argued that the time of the polarization switching was not sufficient to disprove a causal connection between the analyzer and the source [34,35]. Moreover, the inequality used in [23] included coincidence rates when the polarizers were removed. This removal was certainly not performed with any alacrity. Below we will discuss briefly how this loophole might be closed in an advanced version of our proposed scheme, but this is not the central topic of our article.

III. PROPOSED SOURCE

As mentioned above, even with unity-efficiency detectors [27], the down-conversion schemes used until now are inadequate for a completely unambiguous test of Bell's inequalities, because they must perforce discard counts. A schematic of our proposed source is shown in Fig. 1(a). Two nonlinear crystals are simultaneously pumped by a coherent pump beam to induce spontaneous parametric down-conversion; the pumping intensity can be independently varied at each crystal. The crystals are cut for type-II collinear degenerate phase matching (i.e., the down-converted photons are collinear and orthogonally polarized, with spectra at roughly twice the pump wavelength). For example, we envisage using 1-cm-long crystals of beta-barium borate (BBO), pumped by the 325-nm line from a He-Cd laser incident at 54° to the optic axis.

For clarity, we first assume a monochromatic pump beam (at frequency $2\omega_0$) and a single-mode treatment of the down-converted photons. Then the state after the crystals is

$$|\Psi\rangle = \sqrt{1-|A|^2}|\text{vac}\rangle + \frac{A}{\sqrt{1+|f|^2}}(|H,V\rangle_{\text{crystal 1}} + f|H,V\rangle_{\text{crystal 2}}) , \quad (2)$$

where we have omitted higher-order terms (for the very unlikely case in which more than one pump photon down-converts; by reducing the pump intensity, the contribution of these terms can be made as small as desired). A includes the down-conversion efficiency into the modes we are considering and also the pump field strength; f represents a possible attenuation of the pump beam incident on crystal 2. The state (2) describes a photon pair [one photon polarized horizontally (H), the other vertically (V)] originating with probability amplitude $A/\sqrt{1+|f|^2}$ in crystal 1 and with probability $Af/\sqrt{1+|f|^2}$ in crystal 2; it does not include that part of the pump beam which was not down-converted. [Physically, this implies that we have removed the unconverted pump photons, as with a cutoff filter before the detectors (see Fig. 1(b)]. We now combine the modes from the two crystals at a polarizing beam splitter. For an ideal polarizing beam splitter, incident p-polarized light (horizontal in Fig. 1) is completely transmitted, while incident s-polarized light [vertical (out of the plane of the paper) in Fig. 1] is completely reflected; therefore, one photon of each pair will be directed to output port 3, while the conjugate photon is directed to output port 4. (In Sec. VI the case of a nonideal beam splitter is examined; a general calculation is given in the Appendix.) Including a phase shift $\delta = 2\omega_0 \Delta x / c$ (where Δx, the difference in path lengths, may be varied by moving one of the mirrors slightly) between the two nonvacuum terms of (2), we then have

$$|\Psi\rangle \approx (|V\rangle_3 |H\rangle_4 + f e^{i\delta} |H\rangle_3 |V\rangle_4) , \quad (3)$$

where we have omitted the (predominant, but uninteresting) vacuum term and the prefactor $A/\sqrt{1+|f|^2}$. For the balanced case ($f=1$), and for $\delta=180°$, (3) reduces to the familiar singletlike state. (The case where $f \neq 1$ is discussed in Sec. V; Sec. VI examines the case $\delta \neq 180°$.) Note that this is different from Eq. (1), which additionally contains noncoincidence terms that must be intentionally discarded to prepare a singletlike state. Consequently, such a scheme may find application in quantum cryptog-

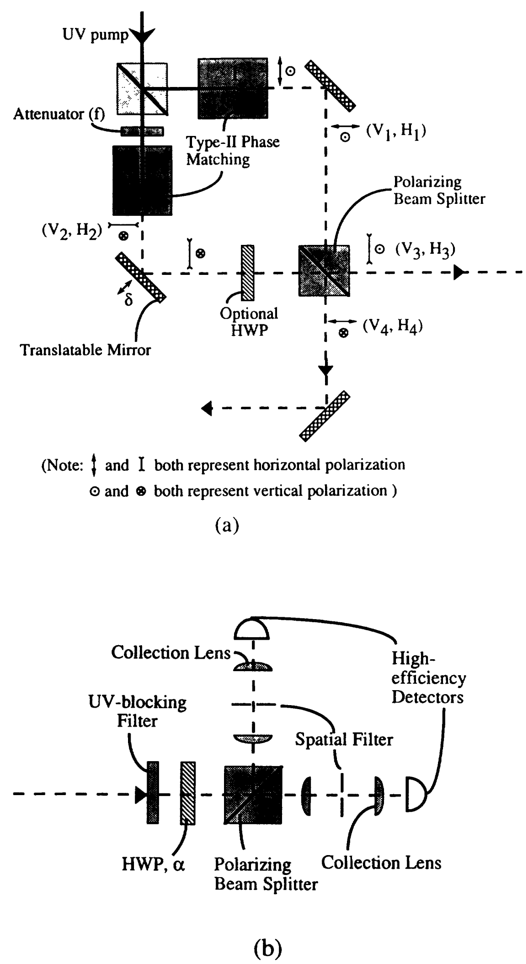

FIG. 1. Schematic of an arrangement in which a loophole-free test of Bell's inequalities is feasible. (a) An ultraviolet pump photon may be spontaneously down-converted in either of two nonlinear crystals, producing a pair of orthogonally polarized photons at half the frequency. One photon from each pair is directed to each output port of a polarizing beam splitter. When the outputs of both crystals are combined with an appropriate relative phase δ, a true singlet- or tripletlike state may be produced. By using a half-wave plate to effectively exchange the polarizations of photons originating in crystal 2, one overcomes several problems arising from nonideal phase matching (see Sec. IV). An additional mirror is used to direct the photons oppositely to separated analyzers. (b) A typical analyzer, including a half-wave plate (HWP) to rotate by α the polarization component selected by the analyzing beam splitter, and precision spatial filters to select only conjugate pairs of photons. In an advanced version of the experiment, the HWP could be replaced by an ultrafast polarization rotator (such as a Pockels or Kerr cell) to close the spacelike-separation loophole.

raphy. It has previously been suggested to use the nonlocal correlations manifest in an entangled state such as (3) to allow two "collaborators" to share a secret, random key [19,20]. As discussed earlier, though, previous EPR schemes have mandated the discarding of up to 50% of the correlated pairs.

With the above source of correlated particles, one can now perform a polarization test of Bell's inequalities [36]. Polarization analysis is performed using an additional polarizing beam splitter after each output port of the interferometer and examining one or both channels of each analyzer with high efficiency detectors [see Fig. 1(b)]. "Rotation" of these analyzers can be effectively accomplished by using a half-wave plate before each one to rotate the polarization of the light. The alignment of the entire setup is obviously somewhat critical. Fortunately, there are simple tests which can be made to verify the integrity of the system. For example, if a half-wave plate (oriented with its axis at 22.5° to the horizontal) is used after crystal 2 to rotate the polarizations by 45°, then the new state of light from that crystal will be $|\Psi\rangle_{\text{crystal 2}} \approx [(|H_2\rangle + |V_2\rangle)/\sqrt{2}][(|H_2\rangle - |V_2\rangle)/\sqrt{2}] = (|H_2,H_2\rangle - |V_2,V_2\rangle)/2$. By blocking the light from the other crystal, one can easily verify this state at the analyzers—either both photons go out on port 3 or both on port 4, and there should be no coincidence counts (assuming no background noise) between ports 3 and 4. More parameters can be checked with half-wave plates after *both* crystals.

If the detectors are far separated from each other and from the source, and one uses some rapid, random means to rotate the light before the analyzers (such as a Pockels or Kerr cell, whose voltage is controlled by a random signal), then one can close the spacelike separation loophole. The signal could be derived, for instance, from the decay of a radioactive substance, or even from the arrival of starlight. Note that since the down-converted photons are emitted within tens of femtoseconds of one another [37] (unlike the photons in an atomic cascade), the limiting time factors will be the detector resolution (expected to be less than 10 ns) and the switching time (which can also be on the order of nanoseconds) [38].

IV. NONIDEAL DOWN-CONVERSION CONSIDERATIONS

In practice, one must take into account several other features of the down-conversion process. First, the down-converted photons will have a finite bandwidth, so that the members of each pair may have different frequencies (though energy conservation still constrains the *sum* of the frequencies to be equal to the pump frequency). We investigate the effect of this by letting the frequency of the horizontally-polarized photons be $\omega_0+\omega$ and the frequency of the vertically polarized photons be $\omega_0-\omega$. We will always assume that the output spectra of the two crystals are identical, i.e., we do not need to further label our frequencies by the crystal number. Furthermore, because of the strong energy correlations, the phase $\delta=(\omega_0+\omega)\Delta x/c+(\omega_0-\omega)\Delta x/c=2\omega_0\Delta x/c$ is independent of the spread in frequency. Then (3) becomes

$$|\Psi\rangle \approx \int d\omega\, A(\omega)(|V_{\omega_0-\omega}\rangle_3|H_{\omega_0+\omega}\rangle_4 + fe^{i\delta}|H_{\omega_0+\omega}\rangle_3|V_{\omega_0-\omega}\rangle_4) . \quad (4a)$$

Note that the frequency of the light at each port is different in the two terms. Unless the bandwidth function $A(\omega)$ describing the amplitude for down-conversion production of the pair $|H_{\omega_0+\omega}\rangle|V_{\omega_0-\omega}\rangle$ is symmetric [i.e., $A(\omega)=A(-\omega)$], the photon color reaching a given detector can serve to *label* from which crystal a pair originated—the first and second terms of (4a) arise from photons from crystals 1 and 2, respectively. To see this, consider the extreme case that $A(\omega)=\delta(\omega-\Omega)$; (4a) becomes

$$|\Psi\rangle \approx |V_{\omega_0-\Omega}\rangle_3|H_{\omega_0+\Omega}\rangle_4+fe^{i\delta}|H_{\omega_0+\Omega}\rangle_3|V_{\omega_0-\Omega}\rangle_4 . \quad (4b)$$

In principle, a precise frequency measurement at either detector could determine the *definite* polarizations of the photons. The distinguishability of the two terms of (4b) destroys their coherence even if no frequency measurement is actually made [39]; this weakens the correlations, making it impossible to violate a Bell inequality.

This difficulty can be avoided, however, by inserting a half-wave plate after one of the crystals (e.g., between crystal 2 and the polarizing beam splitter), oriented with its optic axis at 45° to the horizontal. This will effectively exchange the polarizations of the photons originating in crystal 2

$$(|H_{\omega_0+\omega},V_{\omega_0-\omega}\rangle_{\text{crystal 2}} \Longrightarrow |V_{\omega_0+\omega},H_{\omega_0-\omega}\rangle_{\text{crystal 2}}) ,$$

so that (3) instead becomes

$$|\Psi\rangle \approx \int d\omega\, A(\omega)(|V_{\omega_0-\omega}\rangle_3|H_{\omega_0+\omega}\rangle_4 + fe^{i\delta}|H_{\omega_0-\omega}\rangle_3|V_{\omega_0+\omega}\rangle_4) , \quad (5)$$

which shows that photon color no longer yields which-crystal information; interference persists, and violation of a Bell inequality is possible.

For a plane-wave pump, the phase-matching constraints imply that with careful spatial filtering, one can in principle collect *only* conjugate pairs of photons from the crystals (i.e., essentially no unpaired photons, aside from stray light). Once we allow a more realistic, Gaussian-mode pump, then this is no longer possible. For identical, finite-sized collection irises, there will always exist situations where one photon is detected while the other is not (even aside from the problem of inefficient detectors). This effect is mitigated by collecting over a larger solid angle. Figure 2 shows a plot of inherent collection efficiency versus the collection angle of the iris in units of the pump beam divergence angle. We see that in order to keep the losses less than 2%, we must employ irises which accept light out to 30 times the pump divergence angle. For example, if we employed a 325-nm pump with a beam waist radius ($1/e^2$) of 3.5 mm, then we would need to accept all half-angles up to 1 mrad. [In

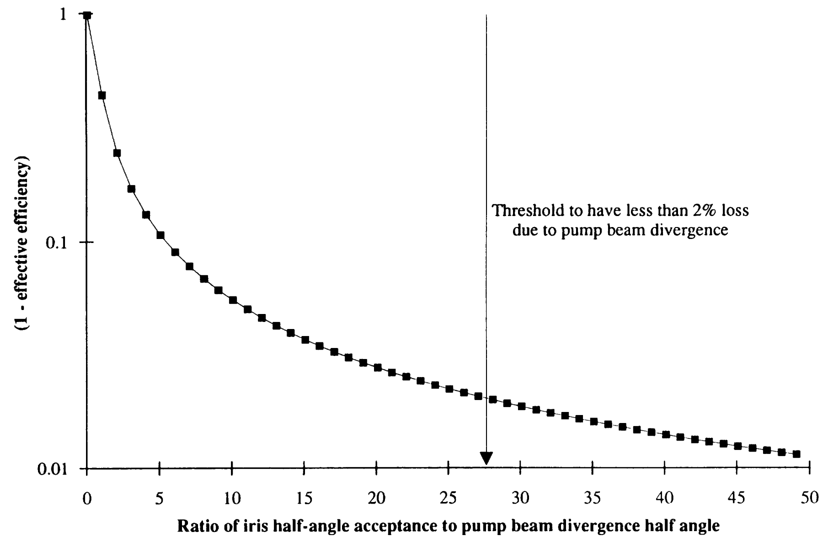

FIG. 2. There is no way to guarantee that both photons will pass through irises of identical size (the optimal situation), due to the Gaussian nature of the pump beam, which causes a slight spread in the angular correlation of the pair. The unavoidable losses may be reduced by using larger collection irises, however: When the acceptance angle of the irises is more than 30 times the divergence angle of the pump beam, losses from this effect are less than 2%.

practice, this could be accomplished by use of a precision spatial filter system in each output port; see Fig. 1(b).] If we assume that the length difference of the two paths is the minimum required to give a $\delta=180°$ phase shift (i.e., $\Delta x=\lambda_{\rm pump}/4$), then the correction factor to this phase shift for including a finite solid angle of detection is at most 10^{-6}. This is completely negligible compared to the wave-front distortion from available optics, discussed in Sec. VI. Moreover, we will see there that one is much better off using $\delta\approx 0$ for other reasons anyway, in which case there is no correction from a finite detection solid angle.

A third practical consideration is the effect of walkoff in the nonlinear crystals. While the birefringence of the nonlinear crystal is essential for achieving phase matching, it also results in a relative *displacement* of the two down-converted photons: they propagate in the same direction after exiting the crystal, but are separated by a distance $d=L\tan\rho$, where L is the propagation distance *inside* the crystal and ρ is the intracrystal angle between the ordinary and extraordinary beams [40]. For a typical crystal length of 1 cm, and a typical ρ of 4° (appropriate for BBO, pumped at 54° to the optic axis), this separation is 0.7 mm, which can be a significant fraction of the beam width (cf. our previous example of a 3.5-mm beam radius). Consequently, after the polarizing beam splitter, the *position* of a detected photon partially labels its origin, degrading coherence. Consider, for example, horizontal polarization corresponding to the ordinary (undeviated) mode (i.e., in Fig. 1, let the plane defined by the pump beam and the crystal optic axis be perpendicular to the plane of the paper) and vertical polarization corresponding to the extraordinary mode, *higher* by d. After the polarizing beam splitter, the two photon modes which travel to a given detector are therefore separated by the amount d. Just as in the situation of asymmetric spectra discussed above, this distinguishes the two otherwise coherent processes. Remarkably, insertion of an extra half-wave plate after one of the crystals to rotate the polarization by 90° avoids this problem, in addition to solving the finite-bandwidth problem. Photons from either crystal exiting port 3 of the beam splitter would be initially extraordinary polarized; photons exiting port 4 of the beam splitter would be initially ordinary polarized. (Note that we could have arranged the optic axes of the crystals to lie in the plane of the figure. The wave plate will be an efficacious measure as long as the setup possesses mirror symmetry about the plane perpendicular to the figure, as defined by the input and output beam splitters [41].)

A similar situation occurs due to *longitudinal* walkoff: After propagation through some length of the birefringent down-conversion crystal, one of the down-converted photons (extraordinary polarized) will "pull ahead" of its conjugate (ordinary polarized). In the absence of the half-wave plate, one could in principle determine from which crystal a given pair of photons originated by examining the timing of the coincident detection—detector 3 going off first would indicate that a given crystal was the parent; detector 4 first would indicate the other crystal. This distinguishability of contributing paths removes quantum interference, just as in the transverse walkoff situation considered above. Fortunately, the extra half-wave plate also removes this longitudinal walkoff effect. With the wave plate, an infinitely fast detector looking at the *originally* extraordinary-polarized photons would *always* trigger before the detector looking at the *originally* ordinary-polarized photons; hence interference remains. In some sense, this longitudinal-walkoff compensation was already implicit in the discussion leading from Eq. (4) to (5), if one interprets $A(\omega)$ as containing the phases acquired by the down-converted photons in propagating through the birefringent crystal.

Finally, even with a plane-wave pump and a crystal cut for collinear phase matching at degeneracy, there is no way to prevent *vector* phase matching at small angles, for slightly different conjugate colors. In fact, a detailed calculation for BBO has shown that even at degeneracy, there is a vector phase-matching solution, in addition to the collinear solution (see Fig. 3). However, because the angles of the conjugate ordinary and extraordinary modes (with respect to the pump beam) are nearly equal in magnitude (for small deviations from collinearity), the net effect on the phase shift δ is completely negligible compared to the effect of imperfect optical elements, which we discuss later. Also, as with the walkoff problem, a half-wave plate after crystal 2 can prevent the angle from distinguishing the interfering processes.

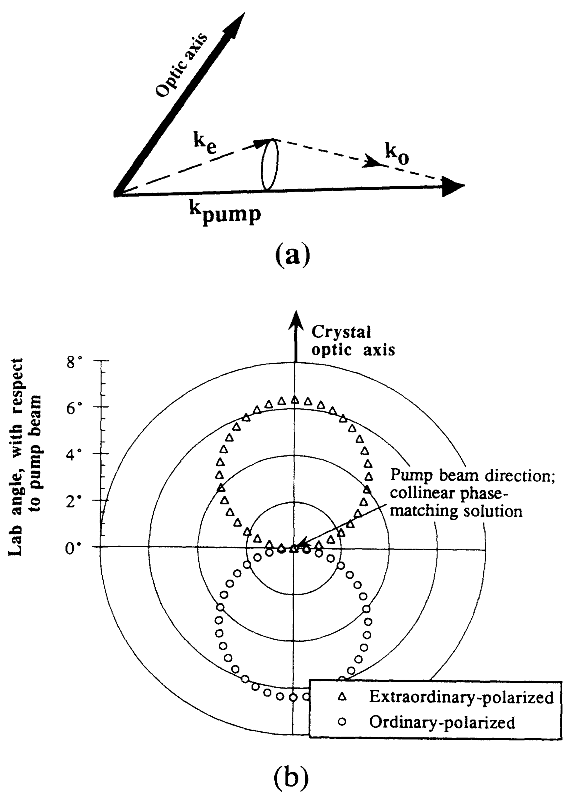

FIG. 3. Spontaneous down-conversion into noncollinear modes occurs due to type-II *vector* phase matching, even when the angle between the crystal optic axis and the pump beam is such as to allow *collinear* phase matching for equal-frequency conjugate photons. (a) A schematic representation of the situation where collinear and vector phase matching are *simultaneously* possible (angles are exaggerated for clarity). For the case of a negative uniaxial crystal (such as BBO), in which the extraordinary index of refraction is less than the ordinary index, the extraordinary beam tends to lie closer to the optic axis than does the ordinary beam. (b) A plot showing the angular deviation from collinearity of degenerate down-converted pairs (at 54° with respect to the BBO optic axis, which points up and out of the page), and to the collinear phase-matching solution. The left axis indicates the angular scale of the four concentric equal-angle contours. The points correspond to the transverse components (i.e., perpendicular to the pump vector) of the photons' momenta—in other words, the plot is essentially what one would see looking into the pump beam, as shown in (a). The ordinary- and extraordinary-polarized photons of a given conjugate pair lie diametrically opposed, on either side of the origin.

V. NONMAXIMAL ENTANGLEMENT

We have until now ignored another parameter at our disposal, namely, the relative amplitude of the terms from the two crystals, governed in (3) by the factor f. It has recently been shown that in the limit of no background [42], one may reduce the required detector efficiency η from 83% to 67% by using a nonmaximally entangled state [11], i.e., $|f|\neq 1$ in (2)–(5). To gain some insight into this, we start with a standard Bell inequality, derived by Clauser and Horne [2,21]:

$$B \leq 0 \ , \tag{6a}$$

where

$$B \equiv n_{ss}(\alpha_1,\beta_1)+n_{ss}(\alpha_2,\beta_1)+n_{ss}(\alpha_1,\beta_2) - n_{ss}(\alpha_2,\beta_2)-n_{3s}(\alpha_1)-n_{4s}(\beta_1) \ , \tag{6b}$$

and $n_{3s}(\alpha_1)$ and $n_{4s}(\beta_1)$ are the singles count rate at the "s" channels of the α analyzer (in output port 3) and β analyzer (in output port 4), respectively (we use here a notation appropriate for an ideal polarizing beam splitter analyzer which reflects all s-polarized light and transmits all p-polarized light). The n_{ss}'s are coincidence rates between the s channels of the two analyzers. The parameters α_i and β_j are essentially the angles of the analyzers (though in practice, they might be the angles of half-wave plates used to rotate the polarization of the incident light). Because the singles rates vary as η (e.g., $n_{3s}=\eta p_{3s}$, where p_{3s} is the singles rate for *unity* efficiency), while the coincidence rates vary as η^2 [i.e., $n_{ss}(\alpha_i,\beta_j)=\eta^2 p_{ss}(\alpha_i,\beta_i)$, where p_{ss} is the coincidence rate for *unity* efficiency], we must have

$$\eta \geq \frac{p_{3s}(\alpha_1)+p_{4s}(\beta_1)}{p_{ss}(\alpha_1,\beta_1)+p_{ss}(\alpha_2,\beta_1)+p_{ss}(\alpha_1,\beta_2)-p_{ss}(\alpha_2,\beta_2)} \tag{7}$$

in order to violate (6). A straightforward calculation (see Appendix) shows that, in the case of perfect polarizing beam splitters, $p_{3s}(\alpha_1)=|A|^2[\cos^2\alpha_1+|f|^2\sin^2\alpha_1]$. If $|f|$ is close to 1, then the singles rate is essentially a constant, independent of α_1; however, if we make f small (by attenuating the pump before crystal 2), then choosing α_1 close to 90° will substantially reduce the contribution of $p_{3s}(\alpha_1)$ to (7) [43]. A similar argument applies for $p_{4s}(\beta_1)$, which is reduced for β_1 close to 0°. Nevertheless, under these conditions, it is still possible to find values of α_2 and β_2 such that $B>0$, even for η as low as 67%. The tradeoff is that the actual magnitude of B is reduced accordingly, however. For example, with $f=0.311$, the maximum value of B is only 0.074, to be contrasted with the maximum value of 0.207 when $f=1$. For this reason, background levels must be kept low for the method to be effective. We hope that background levels of 1% may be achievable [42].

A secondary effect of background counts is to place a lower limit on the number of "accidental" coincidences. Since these arise from *noncorrelated* photons (or dark counts), they will reduce any violation of (6). Accidental coincidence events result when only one member of a given conjugate pair is detected, simultaneously (within the coincidence resolution time) with another count (from a photon from a different pair or a stray-light photon or a dark count), or from simultaneous detection of two background counts. The former process is proportional to the rate of correlated-pair production, so in the absence of background [42], one can in principle make the accidental coincident rate as small as desired by reducing the pumping intensity to the two crystals (in practice, one eventually runs into another limitation: the system will thermally drift over long times). In the presence of a background rate B, the minimum accidental coincidence rate will be $B^2\Delta T$, where ΔT is the coincidence resolution time.

VI. IMPERFECT OPTICAL ELEMENTS

Thus far we have neglected the effects of imperfect analyzers and an imperfect recombining beam splitter. A general treatment is difficult (see the Appendix); however, two special cases can be easily discussed. With no background, a maximally entangled ($f=1$) singletlike ($\delta=0°$ or 180°) state, and perfect beam splitters in the interferometer, the net detection efficiency with imperfect analyzers must satisfy the following relation:

$$\eta(|R_p|^2+|R_s|^2)\left[\sqrt{2}\left[\frac{|R_p|^2-|R_s|^2}{|R_p|^2+|R_s|^2}\right]^2+1\right]>2\ , \tag{8}$$

where R_s is the reflection amplitude for s-polarized light (ideally $|R_s|=1$) and R_p is the reflection amplitude for p-polarized light (ideally $|R_p|=0$). (We have implicitly assumed a lossless beam splitter here.) This result is in agreement with that derived by Clauser and Horne [2,21] and yields the well-known 83% "limit" for perfect analyzers.

Next we consider the case of perfect analyzers but an imperfect recombining polarizing beam splitter (in the interferometer); specifically, we consider a lossless splitter whose reflection and transmission amplitudes are related by $|r_s|=|t_p|$ and $|r_p|=|t_s|$ (i.e., the nonidealities for the two polarizations are equal). As before, we keep $f=1$ and no background. We find that our results depend crucially on whether $\delta=0°$ or 180°. In the former case, both the singles rates and the coincidence rates are independent of $|r_s|=|t_p|$; consequently, the value of B is also. In other words, even a nonpolarizing 50-50 beam splitter could be used as the recombiner. The reason is that under these conditions there is quantum interference which prevents the two photons from exiting the same port of the beam splitter [44]; see Appendix. However, for the $\delta=180°$ (singletlike) case, the coincidence rates, and therefore the value of B, depend strongly on $|r_s|=|t_p|$. This increases the required detection efficiency to achieve a violation ($B>0$):

$$\eta>\frac{2}{(1+\sqrt{2})(|r_s|^2-|r_p|^2)^2} \tag{9}$$

is required to close the detection loophole. A plot of the right-hand side of (9) is shown in Fig. 4, along with the minimum required efficiencies for $\delta=0°$ and for $f=0.608$. It is immediately apparent that one is much better off choosing $\delta=0°$. Of course, any phase shift equal to an integer multiple of 360° is equivalent. However, to mitigate effects from unequal path lengths, one should strive to operate in a white-light configuration, with exactly equal optical path lengths in the two arms.

Ideally, one would like completely flat, homogeneous, optical elements to avoid wave-front distortion inside the interferometer. The effect of such distortion is to change the value of the relative phase shift δ for parts of the beam. The total detected rates will then be averaged over a range of δ values (note that there will also be an averaging over δ if there is an uncorrected temporal *drift* of the optical path lengths). To estimate the effect of this, we have calculated the value of B, as function of δ. Typical results are shown in Fig. 5. We see that the value is not a strong function of the phase deviation from $\delta=0°$. Therefore, it should suffice to specify a $\lambda/20$ flatness for the optics within the interferometer (a feasible requirement over a clear aperture of 1 cm), allowing a phase shift $\delta=0\pm10°$.

VII. CONCLUSION

For $\lambda/20$-flatness optics, a background level of 1%, and custom-selected polarizing beam splitters (with an extinction ratio of at least 500:1), numerical calculation predicts that a violation should be possible choosing $\delta\approx0$, $\alpha_1=99°$, $\alpha_2=58°$, $\beta_1=9°$, $\beta_2=-32°$, and $f=0.55$, as long as the net detection efficiency is greater than 82.6%. Naturally, all optics would be antireflection coated to minimize reflection losses; including a 0.25% loss for each interface and the 2% loss from the Gaussian nature of the beam, this means that the bare detector efficiency needs to be at least 88.6%, which may be achievable in light of our recent measurements [31,32]. Of course, for a safety margin, one would like it to be even higher, if possible.

In conclusion, we have investigated a source of EPR-correlated particles, which makes use of a two-crystal interferometer. If one employs type-II phase matching and a polarizing beam splitter to combine the outputs of the two crystals, a true singletlike or tripletlike state may be produced without the need of discarding counts. Using a half-wave plate in one of the interferometer arms to exchange the roles of the ordinary and extraordinary polarizations removes difficulties arising from finite bandwidth, walkoff, and vector phase-matching considerations. With the setup described herein it should be possible to produce an indisputable violation of a Bell inequality. We have examined the effects of background, imperfect polarizing beam splitters and phase-distorting optics. A loophole-free experiment still seems feasible, although detectors need to be improved somewhat.

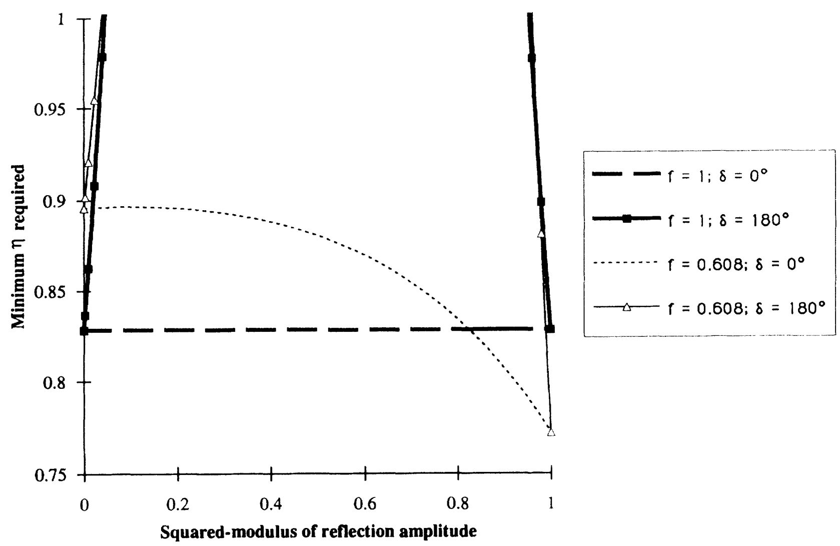

FIG. 4. When $\delta=180°$, the value of B is strongly dependent on the recombining polarizing beam splitter, placing strong constraints on the efficiency required to achieve a violation of $B \leq 0$. Unless the amount of "cross talk" [i.e., the percentage of p-polarized light that exits the s channel (equal to $|r_p|^2$) and the percentage of s-polarized light that exits the p channel (equal to $|t_s|^2$)] is less than 4%, it is not possible to violate Bell's inequality at all. (We also assume a lossless beam splitter and that the cross talk between polarization modes is symmetric, i.e., $|r_p|^2=|t_s|^2$. A general calculation is given in the Appendix.) However, if $\delta=0°$, the value B is independent of $|r_p|^2=|t_s|^2$ for $f=1$, and not a strong function for $f\neq 1$. Consequently, the requirements on detection efficiency are not nearly so severe.

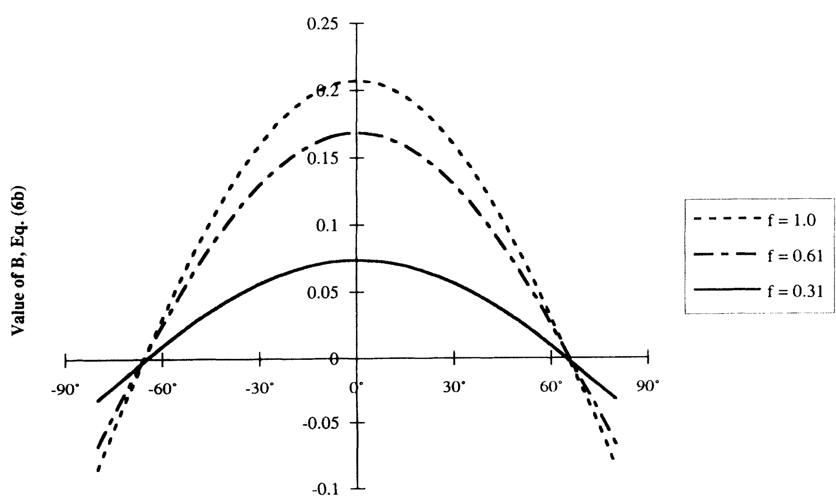

FIG. 5. The dependence of B [Eq. (6b)] on δ, for various values of f (assuming $\eta=1$, no background, and perfect polarizing beam splitters). Wave-front distortion arising from imperfect optics implies that the true value of B will be a sort of weighted average over a range of angles. Due to the weak dependence at small deviations from the ideal phase shift $\delta=0$, this should not pose a serious problem if the deviations can be kept less than $\pm 10°$.

ACKNOWLEDGMENTS

One of us (P.H.E.) is supported by the U.S. Department of Energy, Contract No. DE-AC03-76SF00098. The rest are supported by the Office of Naval Research under Grant No. N00014-90-J-1259.

APPENDIX

We present here a calculation of the various coincidence and singles rates expected from the setup shown in Fig. 1, when nonideal polarizing beam splitters and nonunity efficiency detectors are allowed. We start with the state out of the two crystals, as given in (2):

$$|\Psi\rangle=\sqrt{1-|A|^2}|\text{vac}\rangle + \frac{A}{\sqrt{1-|f|^2}}(|H,V\rangle_{\text{crystal 1}}+f|H,V\rangle_{\text{crystal 2}}) . \tag{A1a}$$

Henceforth we will drop the vacuum term, as well as the prefactor A; physically, we must filter out the unconverted pump beam, which would otherwise give rise to an overwhelming background at our detectors. We write the state (A1a) in terms of photon creation operators:

$$|\Psi\rangle\approx\frac{1}{\sqrt{1+|f|^2}}(\hat{a}^\dagger_{H,1}\hat{a}^\dagger_{V,1}+f\hat{a}^\dagger_{H,2}\hat{a}^\dagger_{V,2})|0\rangle , \tag{A1b}$$

where the subscript *letters* denote polarization and the subscript *numbers* denote spatial mode. Our strategy will be to transform the creation operators $\hat{a}^\dagger_{\lambda,1}$ and $\hat{a}^\dagger_{\lambda,2}$ into operators appropriate for the modes reaching the detectors. To this end, we state at the outset the transformation rules for a half-wave plate and a general beam splitter.

The transformation matrix $H_j(\theta)$ for a half-wave plate (in spatial mode j) with its axis at an angle θ with respect to the horizontal is

$$H_j(\theta)=\begin{bmatrix}\cos2\theta & \sin2\theta\\ \sin2\theta & -\cos2\theta\end{bmatrix} , \tag{A2a}$$

so that

$$\begin{aligned}H_j(\theta)\hat{a}^\dagger_{V,j}&=\hat{a}^\dagger_{V,j}\cos2\theta+\hat{a}^\dagger_{H,j}\sin2\theta ,\\ H_j(\theta)\hat{a}^\dagger_{H,j}&=\hat{a}^\dagger_{V,j}\sin2\theta-\hat{a}^\dagger_{H,j}\cos2\theta .\end{aligned} \tag{A2b}$$

[We assume for simplicity that the wavelengths of the fields are equal to the design wavelength of the wave plate. This approximation is very good, given that the bandwidth of our photons (for the 0.05° half-angle iris acceptance calculation in Sec. IV) is less than 1 nm and zero-order wave plates have a broad performance window.] The mode transformation for a lossless symmetric beam splitter with input ports 1 and 2 and outport ports 3 and 4 is

$$\begin{bmatrix}\hat{a}^\dagger_{\lambda,2}\\ \hat{a}^\dagger_{\lambda,1}\end{bmatrix}=\begin{bmatrix}t_\lambda & r_\lambda\\ r_\lambda & t_\lambda\end{bmatrix}\begin{bmatrix}\hat{a}^\dagger_{\lambda,3}\\ \hat{a}^\dagger_{\lambda,4}\end{bmatrix} , \tag{A3}$$

where t_λ and r_λ are the transmission and reflection amplitudes, respectively. We need to distinguish each polarization component λ to allow for the possibility of a polarizing beam splitter. In what follows, we use r_p to denote the reflection amplitude for p-polarized light (ideally equal to 1) and r_s to denote the reflection amplitude for s-polarized light (ideally equal to 0). For each polarization the amplitudes satisfy the standard relations, derivable from unitarity: $|r_\lambda|^2+|t_\lambda|^2=1$ and $\arg(r_\lambda)-\arg(t_\lambda)=\pm\pi/2$. Without loss of generality, let $t_\lambda=|t_\lambda|$ and $r_\lambda=i|r_\lambda|$.

Using (A3) and including a relative phase shift δ between the two paths (associated, say, with a variable path length in the interferometer arm with crystal 2), the state after the recombining beam splitter can be written

$$|\Psi\rangle\approx(A_{33}\hat{a}^\dagger_{V,3}\hat{a}^\dagger_{H,3}+A_{44}\hat{a}^\dagger_{V,4}\hat{a}^\dagger_{H,4} + A_{34}\hat{a}^\dagger_{V,3}\hat{a}^\dagger_{H,4}+A_{43}\hat{a}^\dagger_{V,4}\hat{a}^\dagger_{H,3})|0\rangle , \tag{A4}$$

where

$$A_{33}=\frac{1}{\sqrt{1+|f|^2}}(r_pr_s+fe^{i\delta}t_pt_s) , \tag{A5a}$$

$$A_{44}=\frac{1}{\sqrt{1+|f|^2}}(t_pt_s+fe^{i\delta}r_pr_s) , \tag{A5b}$$

$$A_{34}=\frac{1}{\sqrt{1+|f|^2}}(r_st_p+fe^{i\delta}r_pt_s) , \tag{A5c}$$

and

$$A_{43}=\frac{1}{\sqrt{1+|f|^2}}(r_pt_s+fe^{i\delta}r_st_p) . \tag{A5d}$$

Note that the terms with coefficients A_{33} and A_{44} correspond to processes where both photons exit the same port of the interferometer. Only the A_{34} and A_{43} terms will contribute to coincidences. Next we include the effect of a half-wave plate in port 3, with its axis at angle $\theta_3\equiv\alpha/2$ to the horizontal, and a half-wave plate in port 4, with its axis at angle $\theta_4\equiv\beta/2$:

$$\begin{aligned}|\Psi'\rangle&=H_3(\theta_3)H_4(\theta_4)|\Psi\rangle\\ &=[A_{33}(\hat{a}^\dagger_{V,3}\cos\alpha+\hat{a}^\dagger_{H,3}\sin\alpha)(\hat{a}^\dagger_{V,3}\sin\alpha-\hat{a}^\dagger_{H,3}\cos\alpha)+A_{44}(\hat{a}^\dagger_{V,4}\cos\beta+\hat{a}^\dagger_{H,4}\sin\beta)(\hat{a}^\dagger_{V,4}\sin\beta-\hat{a}^\dagger_{H,4}\cos\beta)\\ &\quad+A_{34}(\hat{a}^\dagger_{V,3}\cos\alpha+\hat{a}^\dagger_{H,3}\sin\alpha)(\hat{a}^\dagger_{V,4}\sin\beta-\hat{a}^\dagger_{H,4}\cos\beta)+A_{43}(\hat{a}^\dagger_{V,4}\cos\beta+\hat{a}^\dagger_{H,4}\sin\beta)(\hat{a}^\dagger_{V,3}\sin\alpha-\hat{a}^\dagger_{H,3}\cos\alpha)]|0\rangle .\end{aligned} \tag{A6}$$

Finally, we include the effects of the final analyzers, polarizing beam splitters with reflection, and transmission amplitudes R_λ and T_λ, respectively. (Note that we use uppercase for the analyzing beam splitters and lowercase [cf. (A5)] for the recombining beam splitter of the interferometer.) The 16 terms of (A6) become a total of 64!

$$|\Psi'\rangle=[A_{33}\{\cos\alpha\sin\alpha(R_s^2\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{V,3s}+T_s^2\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{V,3p}+2R_sT_s\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{V,3p}-R_p^2\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{H,3s}-T_p^2\hat{a}^\dagger_{H,3p}\hat{a}^\dagger_{H,3p}-2R_pT_p\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{H,3p})$$
$$+(\sin^2\alpha-\cos^2\alpha)(R_sR_p\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{H,3s}+T_sT_p\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{H,3p}+R_sT_p\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{H,3p}+T_sR_p\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{H,3s})\}$$
$$+A_{44}\{\cos\beta\sin\beta(R_s^2\hat{a}^\dagger_{V,4s}\hat{a}^\dagger_{V,4s}+T_s^2\hat{a}^\dagger_{V,4p}\hat{a}^\dagger_{V,4p}+2R_sT_s\hat{a}^\dagger_{V,4s}\hat{a}^\dagger_{V,4p}-R_p^2\hat{a}^\dagger_{H,4s}\hat{a}^\dagger_{H,4s}-T_p^2\hat{a}^\dagger_{H,4p}\hat{a}^\dagger_{H,4p}-2R_pT_p\hat{a}^\dagger_{H,4s}\hat{a}^\dagger_{H,4p})$$
$$+(\sin^2\beta-\cos^2\beta)(R_sR_p\hat{a}^\dagger_{V,4s}\hat{a}^\dagger_{H,4s}+T_sT_p\hat{a}^\dagger_{V,4p}\hat{a}^\dagger_{H,4p}+R_sT_p\hat{a}^\dagger_{V,4s}\hat{a}^\dagger_{H,4p}+T_sR_p\hat{a}^\dagger_{V,4p}\hat{a}^\dagger_{H,4s})\}$$
$$+\{A_{34}\cos\alpha\sin\beta+A_{43}\sin\alpha\cos\beta\}(R_s^2\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{V,4s}+T_s^2\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{V,4p}+R_sT_s[\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{V,4p}+\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{V,4s}])$$
$$+\{-A_{34}\cos\alpha\cos\beta+A_{43}\sin\alpha\sin\beta\}(R_sR_p\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{H,4s}+T_sT_p\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{H,4p}+R_sT_p\hat{a}^\dagger_{V,3s}\hat{a}^\dagger_{H,4p}+T_sR_p\hat{a}^\dagger_{V,3p}\hat{a}^\dagger_{H,4s})$$
$$+\{A_{34}\sin\alpha\sin\beta-A_{43}\cos\alpha\cos\beta\}(R_sR_p\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{V,4s}+T_sT_p\hat{a}^\dagger_{H,3p}\hat{a}^\dagger_{V,4p}+T_sR_p\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{V,4p}+R_sT_p\hat{a}^\dagger_{H,3p}\hat{a}^\dagger_{V,4s})$$
$$+\{-A_{34}\sin\alpha\cos\beta-A_{43}\cos\alpha\sin\beta\}(R_p^2\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{H,4s}+T_p^2\hat{a}^\dagger_{H,3p}\hat{a}^\dagger_{H,4p}+R_pT_p[\hat{a}^\dagger_{H,3s}\hat{a}^\dagger_{H,4p}+\hat{a}^\dagger_{H,3p}\hat{a}^\dagger_{H,4s}])]|0\rangle\ . \tag{A7}$$

As a check of this rather unwieldy expression, one can readily confirm that the normalization is as expected from (A4) (assuming lossless analyzing beam splitters):

$$\langle\Psi'|\Psi'\rangle=|A_{33}|^2+|A_{44}|^2+|A_{34}|^2+|A_{43}|^2\ . \tag{A8}$$

From (A7) we can now calculate the various rates which constitute Bell's inequalities, using the standard Glauber theory of photodetection [45]. For instance, the rate $n_{3s}(\alpha)$ of detection events at a polarization-*insensitive* detector (with efficiency η) at the s channel of the analyzer in port 3 (described by the parameter α) is given by

$$n_{3s}(\alpha)\propto\eta\langle\Psi'|\mathbf{E}_{3s}^{(+)}\cdot\mathbf{E}_{3s}^{(-)}|\Psi'\rangle$$
$$\propto\eta\langle\Psi'|(\hat{a}^\dagger_{H,3s}\varepsilon^*_{H,3s}+\hat{a}^\dagger_{V,3s}\varepsilon^*_{V,3s})\cdot(\hat{a}_{H,3s}\varepsilon_{H,3s}+\hat{a}_{V,3s}\varepsilon_{V,3s})|\Psi'\rangle$$
$$=\eta\langle\Psi'|\hat{a}^\dagger_{H,3s}\hat{a}_{H,3s}+\hat{a}^\dagger_{V,3s}\hat{a}_{V,3s}|\Psi'\rangle=\eta(|\hat{a}_{H,3s}|\Psi'\rangle|^2+|\hat{a}_{V,3s}|\Psi'\rangle|^2)\ . \tag{A9}$$

Using (A7) in (A9), one finds, after many applications of the canonical commutation relations and much algebra,

$$n_{3s}(\alpha)=\eta\{|A_{33}|^2(|R_s|^2+|R_p|^2)$$
$$+|A_{34}|^2(\cos^2\alpha|R_s|^2+\sin^2\alpha|R_p|^2)$$
$$+|A_{43}|^2(\sin^2\alpha|R_s|^2+\cos^2\alpha|R_p|^2)\}\ . \tag{A10a}$$

Similarly, one can calculate the singles rate for the p channel:

$$n_{3p}(\alpha)=\eta\{|A_{33}|^2(|T_s|^2+|T_p|^2)$$
$$+|A_{34}|^2(\cos^2\alpha|T_s|^2+\sin^2\alpha|T_p|^2)$$
$$+|A_{43}|^2(\sin^2\alpha|T_s|^2+\cos^2\alpha|T_p|^2)\}\ . \tag{A10b}$$

Identical expressions hold for $n_{4s}(\beta)$ and $n_{4p}(\beta)$, when the substitutions $3\leftrightarrow4$ and $\alpha\to\beta$ are made. (Note that, in the numerical calculations which were performed, an additional variable background was added; this was *not* taken to rely on η, and so could include dark counts, as well as counts from stray light.) (A10) is very general, and therefore rather complicated. The results simplify greatly if one considers the special case of ideal polarizing beam splitters ($|r_s|=|R_s|=|t_p|=|T_p|=1$, $|r_p|=|R_p|=|t_s|=|T_s|=0$), for then A_{33} and A_{44} vanish. Moreover, $|A_{34}|^2=1/(1+|f|^2)$ and $|A_{43}|^2=|f|^2/(1+|f|^2)$, so that we have

$$n_{3s}(\alpha)=\frac{\eta}{1+|f|^2}\{\cos^2\alpha+|f|^2\sin^2\alpha\}\ , \tag{A11a}$$

$$n_{3p}(\alpha)=\frac{\eta}{1+|f|^2}\{\sin^2\alpha+|f|^2\cos^2\alpha\}\ . \tag{A11b}$$

The various coincidence rates between detectors in positions i and j can be calculated in like fashion, using

$$n_{i,j}(\alpha,\beta)\propto\eta^2\langle\Psi'|\mathbf{E}_i^{(+)}\mathbf{E}_j^{(+)}\mathbf{E}_j^{(-)}\mathbf{E}_i^{(-)}|\Psi'\rangle$$
$$\propto\eta^2\langle\Psi'|:(\hat{a}^\dagger_{H,i}\hat{a}_{H,i}+\hat{a}^\dagger_{V,i}\hat{a}_{V,i})(\hat{a}^\dagger_{H,j}\hat{a}_{H,j}+\hat{a}^\dagger_{V,j}\hat{a}_{V,j}):|\Psi'\rangle\ , \tag{A12}$$

where again we have assumed polarization-insensitive detectors with efficiency η. For example,

$$n_{3s,4s}(\alpha,\beta)\equiv n_{s,s}(\alpha,\beta)=\eta^2\{|A_{34}\cos\alpha\sin\beta+A_{43}\sin\alpha\cos\beta|^2|R_s|^4$$
$$+(|A_{34}\cos\alpha\cos\beta-A_{43}\sin\alpha\sin\beta|^2+|A_{34}\sin\alpha\sin\beta-A_{43}\cos\alpha\cos\beta|^2)|R_sR_p|^2$$
$$+|A_{34}\sin\alpha\cos\beta+A_{43}\cos\alpha\sin\beta|^2|R_p|^4\}\ . \tag{A13}$$

As with the single-event rates, this fairly complicated form simplifies greatly if, for example, the analyzers (described above by R_s and R_p) are ideal. Then only the first term remains:

$$n_{s,s}(\alpha,\beta)=\eta^2|A_{34}\cos\alpha\sin\beta+A_{43}\sin\alpha\cos\beta|^2 . \tag{A14}$$

If we further specialize to the familiar case of an equal superposition of the contributions from the two crystals (i.e., $f=1$), and consider an ideal recombining beam splitter, we are simply left with

$$n_{s,s}(\alpha,\beta)=\frac{\eta^2}{2}|\cos\alpha\sin\beta+e^{i\delta}\sin\alpha\cos\beta|^2 , \tag{A15}$$

which yields the familiar result $(\eta^2/2)\sin^2(\alpha-\beta)$ when $\delta=180°$.

However, as discussed in Sec. VI, it is preferable to use $\delta=0$ for an implementable, loophole-free test of Bell's inequalities, because the required detection efficiency is much less affected by nonidealities of the recombining polarizing beam splitter than when $\delta=180°$. In fact, by examining the structure of A_{33} and A_{44} [Eqs. (A5a) and (A5b), respectively] one can readily see the destructive interference that prevents both photons from exiting the same port of the recombining beam splitter, under the conditions $f=1$, $|r_s|=|t_p|$ and $|r_p|=|t_s|$, and $\delta=0$. [Remember that $\arg(r_\lambda)-\arg(t_\lambda)=\pm\pi/2$.]

[1] J. S. Bell, Physics **1**, 195 (1964).
[2] J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978), and references therein.
[3] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).
[4] D. M. Greenberger, M. A. Horne, and A. Zeilinger, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe,* edited by M. Kafatos (Kluwer, Dordrecht, 1989), p 73.
[5] N. D. Mermin, Phys. Rev. Lett. **65**, 1838 (1990).
[6] L. Hardy, Phys. Rev. Lett. **68**, 2981 (1992).
[7] L. Hardy, Phys. Rev. Lett. **71**, 1665 (1993).
[8] W. J. Munro and M. D. Reid, Phys. Rev. A **47**, 4412 (1993).
[9] J. D. Franson, Phys. Rev. A **48**, 4610 (1993).
[10] For example, although Braunstein and Mann have shown that the detection efficiency requirement can be reduced to 71%, this is only in the limit of no background and a large number of entangled particles [S. L. Braunstein and A. Mann, Phys. Rev. A **47**, R2427 (1993)].
[11] P. H. Eberhard, Phys. Rev. A **47**, R747 (1993).
[12] In [7], Hardy also discusses non-maximally-entangled states, for use in a demonstration of nonlocality without inequalities. However, in the absence of supplementary assumptions, his scheme would require detector efficiencies greater than 98%.
[13] J. D. Franson, Phys. Rev. Lett. **62**, 2205 (1989).
[14] J. Brendel, E. Mohler, and W. Martienssen, Europhys. Lett. **20**, 575 (1992).
[15] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Phys. Rev. A **47**, R2472 (1993).
[16] J. G. Rarity and P. R. Tapster, Phys. Rev. Lett. **64**, 2495 (1990).
[17] Hardy has also proposed a scheme which employs two nonlinear crystals, in a different geometry than that discussed here [L. Hardy, Phys. Lett. A **161**, 326 (1992)]. Specifically, he proposes directing the down-converted photons from one crystal through a second crystal, such that one cannot determine from which crystal a pair originated. However, as he points out, his scheme requires near-perfect alignment to function adequately (the best results reported so far with such a scheme displayed only 30% visibility [X. Y. Zou, L. J. Wang, and L. Mandel, Phys. Rev. Lett. **67**, 318 (1991)]: to violate a Bell's inequality, one needs greater than 71% visibility, assuming 100% detection efficiency and no background). The fact that we use photons which are (very nearly) collinear reduces our alignment difficulties considerably.
[18] A very different proposal has been made by Edward Fry, in *Proceedings of the International Conference on Lasers '93,* edited by C. P. Wang (STS Press, McLean, VA, 1993), using dissociated mercury dimers as the correlated particles. The advantage is that detection efficiencies of 95% are possible by photoionizing the atoms and detecting the photoelectrons. Although it may be that this scheme is a viable one, it is still untested; it should be noted that none of Bell's inequality experiments to date have relied on entangled atoms.
[19] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).
[20] A. K. Ekert, J. G. Rarity, P. R. Tapster, and G. M. Palma, Phys. Rev. Lett. **69**, 1293 (1992).
[21] J. F. Clauser and M. A. Horne, Phys. Rev. D **10**, 526 (1974).
[22] E. Santos, Phys. Rev. A **46**, 3646 (1992).
[23] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982); A. Aspect, P. Grangier, and G. Roger, *ibid.* **49**, 91 (1982); and references therein.
[24] Y. H. Shih and C. O. Alley, Phys. Rev. Lett. **61**, 2921 (1988).
[25] Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988).
[26] In contrast to the cascade situation, the polarizations remain well defined even for noncollinear down-converted pairs. One consequence is that extra means must be used to prepare a *polarization*-entangled state. (Due to energy conservation, the photons are automatically produced in an *energy*-entangled state [13–15].)
[27] Hardy has pointed out that one could use the full state (1) if one's detectors could reliably distinguish between one- and two-photon states [e.g., if a two-photon detection resulted in a pulse twice as high as a single-photon detection, cf. M. D. Petroff, M. G. Stapelbroek, and W. A. Kleinhans, Appl. Phys. Lett. **51**, 406 (1987)], thereby allowing one to account for the cases in which both photons went the same way (private communication).
[28] In these tests as well, half of the counts are discarded (electronically) to eliminate a noninterfering background

[14,15].

[29] Although in this scheme [16] there is in principle no necessity of discarding half of the counts, in practice the small irises needed for high visibility fringes severely reduce the effective detection efficiency.

[30] In experiments performed with γ-ray photons (produced from positronium annhilation, for instance) the actual detector efficiencies are very high, but available polarizers are very poor [2].

[31] P. G. Kwiat, A. M. Steinberg, R. Y. Chiao, P. H. Eberhard, and M. D. Petroff, Phys. Rev. A **48**, R867 (1993).

[32] P. G. Kwiat, A. M. Steinberg, R. Y. Chiao, P. H. Eberhard, and M. D. Petroff, Appl. Opt. (to be published).

[33] Actually, this leaves a factor of 2 safety margin. Ideally, one would desire that the backward light cones of the detection apparatuses and the particle source not overlap, lest it be argued that the apparatus settings share some common cause. Practically, this is impossible (since all must lie within the future light cone of *this* paper), and the best one could achieve would be to base the analyzer settings on random signals from astronomical objects outside each other's light cones (e.g., quasars on opposite sides of the universe). This compromise would rule out all local theories short of "superdeterminism," which amounts to the claim that there is no nonlocality since all events are predetermined, but they are predetermined to precisely mimic a nonlocal theory. Such a metaphysical viewpoint is clearly outside the scope of science.

[34] J. D. Franson, Phys. Rev. D **31**, 2529 (1985).

[35] A. Zeilinger, Phys. Lett. A **118**, 1 (1986).

[36] Using the two-crystal source, one could in principle perform a Franson-type experiment [13–15], without the need to electronically discard counts that arise from the noninterfering processes where one photon travels the short path in its unbalanced Mach-Zehnder interferometer and the other photon travels the long path [28]. To do this, one simply uses polarizing beam splitters in the interferometers, oriented to prevent these processes, i.e., force one long path to be associated with horizontal polarization and the other long path to be associated with vertical polarization. Half-wave plates in two of the paths can be used to rotate all polarizations to vertical, say, restoring coherence. A standard nonpolarizing beam splitter can then be used to recombine the paths.

[37] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. Lett. **71**, 708 (1993).

[38] Franson has argued that the pump coherence time might also come into play [34]. However, this is also of the nanosecond time scale.

[39] The relationship between which-way information and interference has been the subject of much recent discussion [M. O. Scully, B.-G. Englert, and H. Walther, Nature **351**, 111 (1991); P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Phys. Rev. A **45**, 7729 (1992); S. M. Tan and D. F. Walls, *ibid.* **47**, 4663 (1993); P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, *ibid.* **49**, 61 (1994); and references therein].

[40] V. G. Dmitriev, G. G. Gurzadyan, and D. N. Nikogosyan, in *Handbook of Nonlinear Optical Crystals,* edited by A. E. Siegman, Springer Series in Optical Sciences Vol. 64, (Springer-Verlag, New York, 1991).

[41] Actually, it is sufficient that the vectors $\mathbf{k}_1 \times \mathbf{n}_1$ and $\mathbf{k}_2 \times \mathbf{n}_2$ (where $\mathbf{k}_i$ and $\mathbf{n}_i$ are the pump beam and optic axis directions, respectively, associated with crystal *i*) be mirror symmetric about this plane.

[42] By "background" we mean all detector counts not arising from our correlated photons, i.e., stray light, detector dark counts, etc.

[43] Note that $p_{3p}(\alpha)$ varies as $\sin^2\alpha + |f| \cos^2\alpha$, so any attempt to simply average the signals from the *p* and *s* channels, for example, by adding (6) to a similar inequality for the *s*-channel counts, as in the CHSH inequality [J. F. Clauser, R. A. Holt, M. A. Horne, and A. Shimony, Phys. Rev. Lett. **23**, 880 (1969)], will remove the benefit of unbalancing the entanglement.

[44] This is, in some sense, the complementary effect to one in some one-crystal experiments [see, for instance, C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987); and [37]], in which the two-photons from a down-conversion crystal always take the same port of a 50-50 beam splitter.

[45] R. J. Glauber, Phys. Rev. **130**, 2529 (1963).

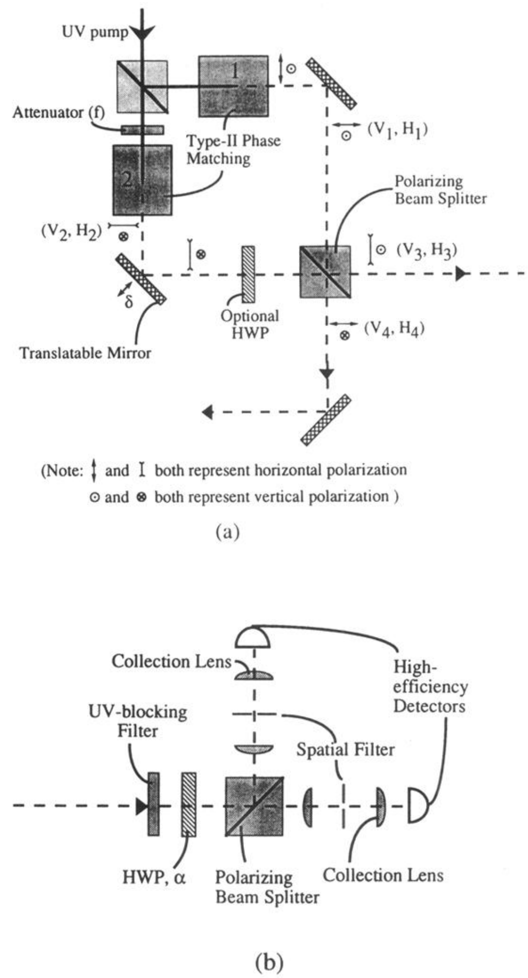

FIG. 1. Schematic of an arrangement in which a loophole-free test of Bell's inequalities is feasible. (a) An ultraviolet pump photon may be spontaneously down-converted in either of two nonlinear crystals, producing a pair of orthogonally polarized photons at half the frequency. One photon from each pair is directed to each output port of a polarizing beam splitter. When the outputs of both crystals are combined with an appropriate relative phase δ, a true singlet- or tripletlike state may be produced. By using a half-wave plate to effectively exchange the polarizations of photons originating in crystal 2, one overcomes several problems arising from nonideal phase matching (see Sec. IV). An additional mirror is used to direct the photons oppositely to separated analyzers. (b) A typical analyzer, including a half-wave plate (HWP) to rotate by α the polarization component selected by the analyzing beam splitter, and precision spatial filters to select only conjugate pairs of photons. In an advanced version of the experiment, the HWP could be replaced by an ultrafast polarization rotator (such as a Pockels or Kerr cell) to close the spacelike-separation loophole.