PHYSICAL REVIEW A VOLUME 47, NUMBER 4 APRIL 1993

High-visibility interference in a Bell-inequality experiment for energy and time

P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao
Department of Physics, University of California, Berkeley, California 94720
(Received 19 June 1992)

We report on a two-photon interference experiment proposed by Franson [Phys. Rev. Lett. **62**, 2205 (1989)], in which sinusoidal fringes with visibilities greater than 70.7%, such as those predicted by quantum mechanics, violate a Bell inequality. We observe visibility of 80.4±0.6%, implying a violation of the inequality by 16 standard deviations. Here the elements of reality under consideration are energy and time rather than spin components. Any classical field models describing separate beams in a Franson interferometer are limited to visibilities less than 50%, and hence ruled out as well, without the need for any supplementary assumptions.

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

When Einstein, Podolsky, and Rosen (EPR) introduced their famous *Gedankenexperiment* in 1935 [1], they proposed that quantum mechanics was incomplete, in an effort to rescue locality and an intuitive notion of "reality." The issue remained a philosophical one until Bell proved in 1964 that *any* local hidden variable (LHV) theory that incorporated the seemingly innocuous concepts of locality and reality would be inconsistent with certain predictions of quantum theory [2]. Many experiments have been performed that, with certain reasonable auxiliary assumptions, violated the inequalities Bell derived for LHV models, agreeing instead with quantum mechanics (QM). Unless one believes that Nature is contrived so as to violate the auxiliary assumptions in just such a way as to mimic QM, the conclusion that Nature is nonlocal is inescapable if one accepts EPR reality. This conclusion is so striking that it is important to test QM in as many different realms as possible.

Nearly all experimental tests of the inequalities to date have involved the superposition of polarization (spin) states, along the lines of the Bohm version of the EPR paradox [3–7]. A noteworthy exception is the recent experiment by Rarity and Tapster [8] based on "phase and momentum." The experimental proposal by Franson [9] concerning a Bell inequality for nonpolarization variables has received a fair amount of attention since its appearance several years ago [10–18]. The "Franson experiment" relies on the entanglement of a *continuous* variable, energy, and is thus closely related to the original EPR paradox. Instead of polarizers or Stern-Gerlach analyzers, spatially separated Mach-Zehnder-like interferometers are used to investigate the nonlocal correlations. Initial attempts to perform the experiment using photon pairs produced in spontaneous down-conversion were limited by slow detectors and relatively small interferometers [10,11]. The resulting extra background reduced the fringe visibility so that it was not possible in a single experiment to rule out classical field models, let alone *all* local realistic theories. (A recent paper [17] even seems to imply that at large path-length differences, high visibility would still be unattainable, due to the finite size of the down-converting crystal.) Brendel, Mohler, and Martienssen were the first to succeed in removing the unwanted background [14], reporting a visibility of 87%. However, their arrangement employed a *single-beam* Michelson interferometer, and no conclusions regarding locality could be made. Researchers have also studied the possibility of employing the nonlocal correlations in various communications and cryptography schemes [13,15,16,18]. We have studied elsewhere [19] how the correlations can be used to make dispersion-free, high-resolution time-of-flight measurements.

Via parametric down-conversion in a crystal possessing a $\chi^{(2)}$ nonlinearity, a pump photon at ω_p may be spontaneously converted into a pair of highly correlated photons. This process is subject to phase-matching constraints, and energy conservation, $\omega_p = \omega_1 + \omega_2$. Although the finite size of the crystal contributes to a nonzero bandwidth for each down-converted beam, energy conservation is strictly enforced. Thus, while *each* down-converted photon may have a substantial bandwidth, the *sum* of their frequencies is fixed to within the pump bandwidth, which is negligible in our experiment. The emitted photons are thus described by an energy-entangled state. We select out pairs such that ω_1 and ω_2 are centered at $\omega_p/2$, and essentially equal (to within their bandwidths).

Previous experiments [20,21] have demonstrated that the photons also have very strong temporal correlations, so that if one photon is detected at time t, its conjugate will be detected within the two-photon correlation time τ_{TPC}. This is of the same order as the single-photon coherence time τ_c, usually determined in practice by filters before the detectors, and is ~ 120 fs for our 10-nm filters. Note that we now have a situation very similar to that originally proposed by EPR. They described a system of two particles in a simultaneous eigenstate of the operators $k_1 + k_2$ and $x_1 - x_2$. If we make the transformations $k \rightarrow \omega/c$ and $x \rightarrow ct$, we essentially have this state. The signal and idler energies sum to a constant, and the difference in their times of emission is nearly zero. Bell has shown, however, that since there is a positive-definite Wigner function that can describe these properties, no *direct* violation of a Bell inequality is possible for these observables [22]. Consequently, in order to investigate an inequality based on energy and time, it is necessary to produce a state more akin to the spin singlet state of the Bohm version of the EPR paradox. This is what the present experiment aims to do.

We first give a simplified analysis of this experiment based on Feynman's notion of interference of indistinguishable processes. Let each conjugate photon enter a modified Mach-Zehnder interferometer (MZ); see Fig. 1. Each interferometer ($j=1,2$) has a short path of length S_j and a long path of length L_j. If the optical imbalance $\Delta L_j = L_j - S_j$ is less than the coherence length $c\tau_c$ of the incident photon, then fringes will be visible in single-event (singles) rates as the long arm is moved slightly. Henceforth we restrict our discussion to the case where $\Delta L_j \gg c\tau_c$, in which there are no single-event fringes. Fringes can nevertheless be observed in the rate of coincident detections between the two detectors [23]. It becomes helpful to regard each MZ as consisting of two optical delay lines in parallel, s ("short") and l ("long"). Then for any incident photon pair, there are four processes (s-s, l-l, s-l, and l-s) leading to pair detection. In the s-l and l-s processes, however, the two photons exit their respective MZ's having acquired a relative time lag large with respect to τ_{TPC}. These processes, which account for one-half of the emitted photon pairs, are therefore distinguishable from each other as well as from the s-s and l-l processes. According to the Feynman rules for interference, they therefore constitute a noninterfering background. If the difference *between* the path-length differences ($\Delta L \equiv \Delta L_1 - \Delta L_2$) is greater than $c\tau_{\mathrm{TPC}}$, then the l-l's are also distinguishable from the s-s's, and no interference will result. We thus further restrict our discussion to the case $\Delta L \ll c\tau_{\mathrm{TPC}}$. Then the l-l and s-s coincidence processes are indistinguishable from each other, because the absolute time of emission of the pair from the crystal is undetermined (for a cw pump). By means of post-selection using fast coincidence counters, the l-s and s-l counts are rejected, effectively reducing the output state to

$$|\Psi\rangle_{\mathrm{out}} \approx \tfrac{1}{2}[\,|\Psi_s\rangle_1|\Psi_s\rangle_2 + |\Psi_l\rangle_1|\Psi_l\rangle_2] \;, \tag{1}$$

which is an entangled state similar to the familiar singlet state. The relative phase $\Delta\phi$ of the interfering s-s and l-l processes is the sum of the relative phases acquired by the individual photons, i.e.,

$$\begin{aligned}\Delta\phi &= \omega_1 \Delta L_1/c + \omega_2 \Delta L_2/c \\ &= \frac{\omega_1+\omega_2}{2c}(\Delta L_1 + \Delta L_2) + \frac{\omega_1-\omega_2}{2c}(\Delta L_1 - \Delta L_2) \\ &\approx \frac{\omega_p}{2c}(\Delta L_1 + \Delta L_2) \;, \end{aligned} \tag{2}$$

since $\Delta L_1 - \Delta L_2$ is arranged to be small relative to the inverse bandwidth of ω_1 and ω_2. The rate of "true" coincidences is thus proportional to

$$|1+e^{i\Delta\phi}|^2 = 2 + 2\cos\left[\frac{\omega_p}{2c}(\Delta L_1 + \Delta L_2)\right] \;, \tag{3}$$

which displays 100% visibility due to the strong correla-

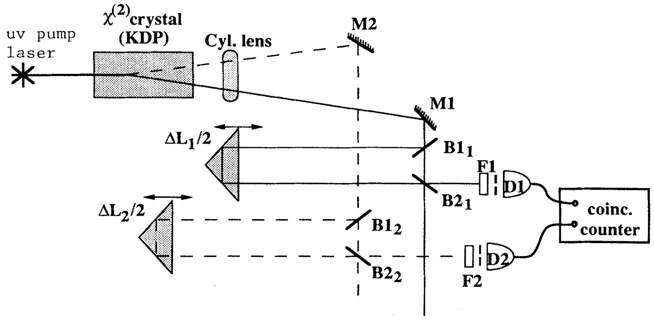

FIG. 1. Experimental setup.

tions between conjugate photons in the entangled state (1) [24], leading to a violation of a Bell inequality. If, on the other hand, the coincidence gate window is set larger than the time lags $\Delta L_j/c$, the *nearly* coincident s-l and l-s pairs are also detected, reducing the visibility to 50% [10,11].

A schematic of our experimental setup is shown in Fig. 1. The pump beam is produced by an argon-ion laser in single-mode operation at $\lambda = 351.1$ nm. The 2-mm beam is attenuated to approximately 90 mW, and then enters a 10-cm potassium dihydrogen phosphate (KDP) crystal. The crystal is oriented with its optic axis at an angle of 50.6° with respect to the uv beam; the degenerate ($\lambda = 702$ nm) signal and idler beams emerge on opposite sides of a cone whose half-opening angle is 2.1°. A 40-cm–focal-length cylindrical lens is used to collimate the beams in the vertical direction. Spatial profiles performed in coincidence showed that 1.5 m from the crystal, the signal beam conjugate to an idler beam selected by a small pinhole had a size on the order of 4 mm^2 and a divergence angle on the order of 1 mrad.

After the lens the conjugate beams traverse similar optics, so we will describe only the path of beam 1. Bending mirror $M1$ directs the beam into an unbalanced MZ, formed by two 50-50 beam splitters ($B1_1$ and $B2_1$) and a translatable right-angle prism. The prism was translated by means of a Burleigh Inchworm piezoelectric system, nominally capable of 4-nm single steps [25]. The position of the prism was monitored via a Heidenhain optical encoder, with a 0.1-μm resolution. The optical path-length difference in each interferometer was approximately 63 cm, much smaller than the pump coherence length ($c\tau_{\mathrm{pump}} > 6$ m), but much larger than the coherence length ($c\tau_c \approx 36$ μm) of the down-converted light. Following a filter and an adjustable iris, the beam was focused onto detector $D1$, a customized silicon avalanche photodiode module (EG&G SPCM-200-PQ). (We have measured the detection efficiency to be greater than 50% at $\lambda = 702$ nm, but at present the filter and iris severely limit the effective efficiency.) We measured the time resolution (10–90% region) of the devices to be 1.1 ns. The outputs of the two detectors were fed into a time-to-amplitude converter with 100-ps resolution. This was operated with a time window of 1.46 ns, thereby eliminating nearly all contribution from the l-s and s-l coincidences, which were displaced by ± 2.1 ns relative to the s-s and l-l (due to the

63-cm path-length differences).

Typical results (see Fig. 2) displayed sinusoidal coincidence fringes with a visibility $V = 80.4 \pm 0.6\%$, while no fringes were discernible in the single-event rates. The less-than-unity visibility even with the short gate window is due to some combination of the following effects:

(i) Imperfect alignment of the bending mirrors, interferometers, and irises, such that s-s and l-l are not entirely indistinguishable.

(ii) Inevitable loss of visibility due to diverging input beams and large path-length imbalance. There were generally about twice as many s-s coincidences as l-l's (this ratio depended critically on alignment), which should reduce the maximum visibility to 94% [26]. (In a later experiment, to be described elsewhere, we compensated for this by using neutral density filters in the short paths.)

(iii) Finite size of irises, accepting light from various paths that acquire slightly different relative phases in the interferometers. As evidence for this effect, the approximate fringe visibility improved from 63% to 75% to 80% when the irises' sizes were reduced sequentially from ≈ 1.5 mm to ≈ 0.8 mm to ≈ 0.4 mm in diameter.

(iv) Time averaging over slow drifts in laser frequency (the *instantaneous* linewidth of the pump should be negligible) and/or air temperature in the MZ's. From the fit in Fig. 2 one finds that the period of the fringes is ≈ 282 nm over the 1700-s run, differing from the expected value of 351 nm. This is consistent with the quarter-fringe drift over 10 min observed in a separate stability test. As a worst-case estimate, we treat the drift as a random walk, finding a phase-diffusion coefficient of about $5°/s^{1/2}$, and a visibility reduction of about 1.3%. In runs where the counting times were ten times shorter, the observed fringe spacing was 348 ± 1 nm, in much closer agreement with the expected value.

As has now been discussed in many papers [10,27,28], 50% is the maximum visibility possible in a classical field approach to this sort of experiment. Our results therefore preclude any classical description, without the need to include results from other experiments. In general, the derivation of a *testable* Bell inequality requires making additional reasonable assumptions [4,29]. Once this is done, the inequality can be violated when the coincidence rate varies sinusoidally as a function of the difference of the "parameter settings" at the two spatially separated analyzers, as in the QM prediction Eq. (3). (Typically, these settings are polarizer orientations, but in our experiment, one is $+\Delta L_1$ and the other is $-\Delta L_2$.) This violation occurs whenever the visibility exceeds $1/\sqrt{2}$, or about 70.7%. The fringes we have observed therefore imply a violation of the Bell inequality by 16 standard deviations. (The restriction that the rate vary *sinusoidally* with the sum of the two path-length differences is essential. It is fairly easy to concoct LHV models that predict 100%-visibility *triangular* fringes, or fringes proportional to the *product* of separate sinusoidal functions of the two parameters. These models do not violate the Bell inequality.) When observations are only made at two of the four output ports, reasonable assumptions must also be made about the rates at the unused ports, based on the symmetry of the experiment. Using three detectors, and varying the phases in *both* interferometers, we have since performed further experiments that support these assumptions. (Elsewhere, we will publish data that show visibilities greater than 90%, and a direct violation of the inequality, although by fewer standard deviations.)

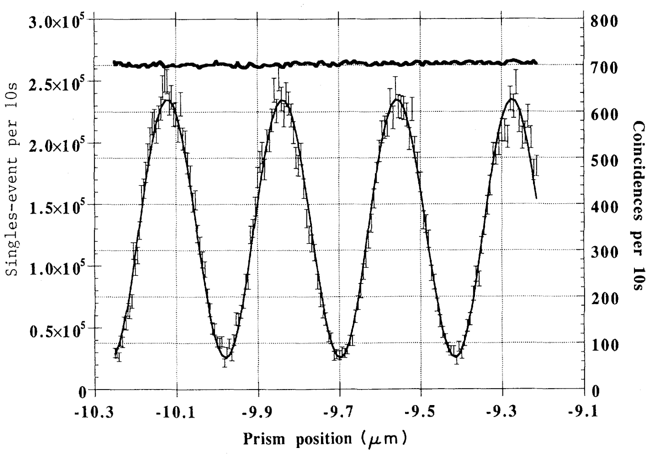

FIG. 2. Constant single-event rate (left axis) and coincidence fringes [right axis, with the accidental rate of approximately 10 (counts/10 s) subtracted]. For this run, the irises were at their minimum size (~0.4 mm diameter). The position data are interpolated from the Heidenhain encoder's 0.1-μm-resolution output. The solid curve is a sinusoidal fit, and has a visibility of 80.4%.

There are several possible interpretations of the Bell inequality for this experiment. The initial proposal named it a test of a "position and time" inequality. The related experiment by Rarity and Tapster [8] describes a "momentum-phase" inequality. In a private communication, Caves and Braunstein have maintained that the variables in question are simply *which port* of the final beam splitter each photon exits. We describe the present experiment in terms of a Bell inequality concerning energy and time. As discussed earlier, the underlying mechanism for the observed interference is the strong correlations of both of these variables in the conjugate photons. Although Bell's positive-definite Wigner function is equivalent to a classical probability distribution for all measurements made *directly* on these two incompatible observables, it is not truly an LHV model, since it includes all the QM correlations. When measurements are made involving the coherent superposition of the field at different times, as in our MZ interferometers, it is possible to bring out the irreducible nonlocality. An example of an LHV model inconsistent with our results is one that ascribes to each photon some definite but unknown energy, but allows this energy to differ among the members of the ensemble, thus washing out any fringes in single-event detection. It is straightforward to show that such a theory will lead to no more than 50% visibility in this experiment, even for a short coincidence window.

In conclusion, we report here the observation of high-visibility fringes in the "Franson experiment." This extends the work of previous research with Michelson interferometers to spatially-separated Mach-Zehnder interferometers. By employing fast detectors and larger path-length imbalances, we were able to remove the unwanted background of long-short and short-long coincidences, which had limited previous results with separated beams to $V \leq 50\%$. Our visibility of $80.4 \pm 0.6\%$ allows us to exclude any classical field models, which necessarily have $V \leq 50\%$, without reference to separate measurements. Furthermore, contingent on reasonable extra assumptions, we can infer a violation of a Bell inequality by more than 16 standard deviations. We interpret these results to rule out the possibility of any local realistic theory underlying the simultaneous energy and time correlations of down-converted photon pairs.

This work was supported by the Office of Naval Research under Grant No. N00014-90-J-1259.

---

[1] A. Einstein, B. Podolsky, and N. Rosen, Phys. Rev. **47**, 777 (1935).

[2] J. S. Bell, Physics **1**, 195 (1964).

[3] D. Bohm, in *Quantum Theory and Measurement,* edited by J. A. H. Wheeler and W. H. Zurek (Princeton University Press, Princeton, NJ, 1983), pp. 356–368.

[4] J. F. Clauser and A. Shimony, Rep. Prog. Phys. **41**, 1881 (1978).

[5] A. Aspect, J. Dalibard, and G. Roger, Phys. Rev. Lett. **49**, 1804 (1982).

[6] Y. H. Shih and C. O. Alley, Phys. Rev. Lett. **61**, 2921 (1988).

[7] Z. Y. Ou and L. Mandel, Phys. Rev. Lett. **61**, 50 (1988).

[8] J. G. Rarity and P. R. Tapster, Phys. Rev. Lett. **64**, 2495 (1990).

[9] J. D. Franson, Phys. Rev. Lett. **62**, 2205 (1989).

[10] P. G. Kwiat, W. A. Vareka, C. K. Hong, H. Nathel, and R. Y. Chiao, Phys. Rev. A **41**, 2910 (1990).

[11] Z. Y. Ou, X. Y. Zou, L. J. Wang, and L. Mandel, Phys. Rev. Lett. **65**, 321 (1990).

[12] J. G. Rarity, P. R. Tapster, E. Jakeman, T. Larchuk, R. A. Campos, M. C. Teich, and B. E. A. Saleh, Phys. Rev. Lett. **65**, 1348 (1990).

[13] A. K. Ekert, Phys. Rev. Lett. **67**, 661 (1991).

[14] J. Brendel, E. Mohler, and W. Martienssen, Phys. Rev. Lett. **66**, 1142 (1991).

[15] C. H. Bennett, Phys. Rev. Lett. **68**, 3121 (1992).

[16] J. G. Rarity and P. R. Tapster, Phys. Rev. A **45**, 2052 (1992).

[17] M. H. Rubin and Y. H. Shih, Phys. Rev. A **45**, 8138 (1992).

[18] A. K. Ekert, J. G. Rarity, P. R. Tapster, and G. M. Palma, Rev. Lett. **69**, 1293 (1992).

[19] A. M. Steinberg, P. G. Kwiat, and R. Y. Chiao, Phys. Rev. Lett. **68**, 2421 (1992).

[20] C. K. Hong, Z. Y. Ou, and L. Mandel, Phys. Rev. Lett. **59**, 2044 (1987).

[21] P. G. Kwiat, A. M. Steinberg, and R. Y. Chiao, Phys. Rev. A **45**, 7729 (1992).

[22] J. S. Bell, Ann. N.Y. Acad. Sci. **480**, 263 (1986).

[23] By "coincident detection," we mean that the time between detection events falls within a specified "gate" window, after electronic and optical delays have been accounted for.

[24] M. Horne, A. Shimony, and A. Zeilinger, in *Quantum Coherence,* edited by J. S. Anandan (World Scientific, Singapore, 1990), pp. 356–372.

[25] We actually observe an average step size of 6 nm, and individual steps that vary over a 3-μm cycle from about 2 nm to 8 nm.

[26] The optimal visibility of interference between two processes if the probability of one process is reduced by a factor of α relative to that of the other is $2\alpha^{1/2}/(1+\alpha)$.

[27] C. Su and K. Wódkiewicz, Phys. Rev. A **44**, 6097 (1990).

[28] J. D. Franson [Phys. Rev. Lett. **67**, 290 (1991)] has pointed out that no classical field model can account for *both* the 50% visibility and the tight time correlations seen in other experiments.

[29] In particular, for detection efficiencies below about 83%, the assumption must be made that the observed counts are "representative" of the entire ensemble, including those particles that escape detection. It has been pointed out that this "detection loophole" can never be evaded in the Franson experiment due to the necessity of throwing away half the counts, i.e., the *s-l* and *l-s* processes. The other principal assumption is that the state of the particles emitted by the source is independent of the parameter settings at the analyzers.