We present a measurement of the time-dependent CP-violating asymmetries in $B^0 \to K^{*0} \gamma (K^{*0} \to K_S^0 \pi^0)$ decays based on $124 \times 10^6 \ U(4S) \to BB$ decays collected with the BABAR detector at the PEP-II asymmetric-energy $B$ Factory at the Stanford Linear Accelerator Center. In a sample containing $105 \pm 14$ signal decays, we measure $S_{K^{*0}} = 0.25 \pm 0.63 \pm 0.14$ and $C_{K^{*0}} = -0.57 \pm 0.32 \pm 0.09$, where the first error is statistical and the second, systematic.

We identify $B^0 \to K^{*0} \gamma$ decays to $K^{*0} \gamma$ combinations using two nearly independent kinematic variables: the energy-substituted mass $m_{ES} = \sqrt{(s/2 + p_i \cdot p_B)^2 / E_i^2 - p_B^2}$ and the energy difference $\Delta E = E_{B^*} - \sqrt{s}/2$. Here $(E_i, p_i)$ and $(E_B, p_B)$ are the four-vectors of the initial $e^+e^-$ system and the $B$ candidate, respectively, $\sqrt{s}$ is the center-of-mass energy, and the asterisk denotes the center-of-mass (c.m.) frame. For signal decays, the $m_{ES}$ distribution peaks near the $B$ mass with a resolution of $3.5$ MeV/$c^2$, and $\Delta E$ peaks near $0$ MeV with a resolution of $50$ MeV. Both $m_{ES}$ and $\Delta E$ exhibit a low-side tail from energy leakage in the EMC. For the study of CPV asymmetries, we consider candidates within $5.2 < m_{ES} < 5.3$ GeV/$c^2$ and $|\Delta E| < 300$ MeV, which includes the signal as well as a large “sideband” region for background estimation. When more than one candidate is found in an event, we select the combination with the $\pi^0$ mass closest to the nominal $\pi^0$ value, and if ambiguity persists, we select the combination with the $K^0_S$ mass closest to the nominal $K^0_S$ value.

The sample of candidate events selected by the above requirements contains significant background contributions from continuum $e^+e^- \to q\bar{q} \ [q = (u, d, s, c)]$, as well as random combinations from other $B$ meson decays (mostly from other $b \to s \gamma$ decays [6]). We suppress both of these backgrounds by taking advantage of the expected angular distribution of the decay products of these processes. Angular momentum conservation restricts the $K^{*0}$ meson in the $B^0 \to K^{*0} \gamma$ decay to transversely polarized states, which leads to an angular distribution of $\sin^2 \theta_H$ for the decay products, where $\theta_H$ is the angle between the $K^0_S$ and the $B$ meson directions in the $K^{*0}$ rest frame.
Monte Carlo studies show that the background candidates peak near \( \cos \theta_{\mu} = -1 \). We require \( \cos \theta_{\mu} > -0.6 \), resulting in rejection of 68% of \( BB \) and 48% of continuum background candidates, while retaining 91% of the signal.

We exploit topological variables to further suppress the continuum backgrounds, which in the c.m. frame tend to retain the jet-like features of the \( q \bar{q} \) fragmentation process, as opposed to spherical \( BB \) decays. In the c.m. system we calculate the angle \( \theta_{s} \) between the sphericity axis of the \( B \) candidate and that of the remaining particles in the rest of the event. While \( |\cos \theta_{s}| \) is highly peaked near 1 for continuum background, it is nearly uniformly distributed for \( BB \) events. We require \( |\cos \theta_{s}| < 0.9 \), eliminating 58% of the continuum events.

We also employ an event-shape Fisher discriminant in the maximum-likelihood fit (described below) from which we extract the CPV measurements. This variable is defined as \( \mathcal{F} = 0.53 - 0.60L_{0} + 1.27L_{y} \), where \( L_{j} = \sum_{i \in \text{ROE}}|p_{i}^{j}| \cos \theta_{s}^{j} \), \( p_{i}^{j} \) is the momentum of particle \( i \) in the c.m. system, and \( \theta_{s}^{j} \) is the angle between \( p_{i}^{j} \) and the sphericity axis of the \( B \) candidate.

The above selections yield 1916 \( B^{0} \rightarrow K^{*0} \gamma(K^{*0} \rightarrow K_{S}^{0}\pi^{0}) \) candidates. We extract our measurements from this sample using an unbinned maximum-likelihood fit to kinematic (\( \Delta E \) and \( K^{*0} \) mass), event shape (\( \mathcal{F} \)), flavor tag, and time-structure variables (described below).

As input to the fit, we parameterize the probability distribution functions (PDF) describing the observables of signal and \( BB \) background events using either more copious fully-reconstructed \( B \) decays in data or simulated samples. For the continuum background, we select the functional form of the PDFs describing each fit variable in data using the sideband regions of the other observables where the \( q \bar{q} \) background dominates. We include these regions in the fitted sample and simultaneously extract the parameters of the background PDFs along with the CPV measurements. We fit \( 105 \pm 14 \) signal and \( 19 \pm 15 \) other \( B \) decays in the selected sample. This signal yield is consistent with expectations from the previous measurements of the branching fractions [5–7]. Figure 1 displays the \( m_{ES} \) and \( M_{K^{*0}} \) distributions for signal-enhanced sub-samples of these events, selected using the PDFs employed in the fit (see below).

For each \( B^{0} \rightarrow K^{*0} \gamma \) candidate, we examine the remaining tracks and neutral particles in the event to determine if the other \( B \) in the event \( B_{\text{tag}} \) decayed as a \( B^{0} \) or a \( B^{0} \) (flavor tag). Time-dependent CPV asymmetries are determined by reconstructing the distribution of the proper decay time difference \( \Delta t = \tau_{CP} - \tau_{tag} \). At the \( \Upsilon(4S) \) resonance, the distribution of \( \Delta t \) follows

\[
P_{B}^{\Delta t}(\Delta t) = \frac{e^{-|\Delta t/\tau}|}{4\tau} \left\{ 1 \pm \left[ S_{f} \sin(\Delta t\Delta m_{d}) - C_{f} \cos(\Delta t\Delta m_{d}) \right] \right\},
\]

where the upper (lower) sign corresponds to \( B_{tag} \) decaying as \( B^{0} \) (\( B^{0} \)), \( \tau \) is the \( B^{0} \) lifetime, \( \Delta m_{d} \) is the mixing frequency, and \( S_{f} \) and \( C_{f} \) are the magnitude of the mixing-induced and direct CPV asymmetries, respectively. As stated above, in the SM we expect \( S_{K^{*}} = 2(m_{f}/m_{b}) \sin \beta = 0.05 \). We expect \( C_{K^{*}} = -A_{K^{*0}0}^{m} \), the direct CPV asymmetry measured in the self-tagging and more copious \( B^{0} \rightarrow K^{*0} \gamma(K^{*0} \rightarrow K^{+}\pi^{-}) \) decay.

We use a neural network to determine the flavor \( T \) of the \( B_{\text{tag}} \) meson from kinematic and particle identification information [10]. Each event is assigned to one of five mutually exclusive tagging categories, designed to combine flavor tags with similar performance and \( \Delta t \) resolution. We parameterize the performance of this algorithm in a data sample (\( B_{\text{flav}} \)) of fully-reconstructed \( B^{0} \rightarrow D^{(*)-}\pi^{+}/p^{+}/a_{1}^{+} \) decays. The average effective tagging efficiency obtained from this sample is \( Q = \sum (1 - 2w_{e})^{2} = 0.288 \pm 0.005 \), where \( w_{e} \) and \( e_{e} \) are the efficiency and mistag probabilities, respectively, for events tagged in category \( c \). In each tagging category, we extract the fraction of events (\( e_{e}^{c} \)) and the asymmetry in the rate of \( B^{0} \) and \( B^{0} \) tags in the continuum background events in the fit to the data.

We compute the proper time difference \( \Delta t \) from the known boost of the \( e^{+}e^{-} \) system and the measured \( \Delta z = z_{CP} - z_{tag} \), the difference between the reconstructed decay vertex positions of the \( B^{0} \rightarrow K^{*0} \gamma \) and \( B_{\text{tag}} \) candidate.
along the boost direction (z). A description of the inclusive reconstruction of the $B_{\text{tag}}$ vertex using tracks in the rest of the event (ROE) is given in Ref. [10]. Replicating the vertexing technique developed for $B^0 \rightarrow K_S^0 \pi^0$ decays [9], we determine the decay point $z_{CP}$ for $B^0 \rightarrow K^{*0} \gamma(K^{*0} \rightarrow K_S^0 \pi^0)$ candidates from the intersection of the $K_S^0$ trajectory with the interaction region. This is accomplished by constraining the $B$ vertex to the interaction point (IP) in the plane transverse to the beam, which is determined in each run from the spatial distribution of vertices from two-track events. We combine the uncertainty in the IP position, which follows from the size of the interaction region (about 200 $\mu m$ horizontal and 4 $\mu m$ vertical), with the root mean square (RMS) of the transverse $B$ flight length distribution (about 30 $\mu m$) to assign an uncertainty to the IP constraint.

Simulation studies indicate that $B^0 \rightarrow K^{*0} \gamma(K^{*0} \rightarrow K_S^0 \pi^0)$ decays exhibit properties which are characteristic of the IP vertexing technique, namely, that the per-event estimate of the error on $\Delta$t, $\sigma_{\Delta t}$, reflects the expected dependence of the $z_{CP}$ resolution on the $K_S^0$ flight direction and the number of SVT layers traversed by its decay daughters. Though the fit extracts $C_{K^- \gamma}$ from all flavor tagged signal decays, we only allow 68% of these events to contribute to the measurement of $S_{K^- \gamma}$. This subset consists of candidates which are composed of $K_S^0$ decays with at least one hit in the SVT on both tracks and pass the quality requirements of $\sigma_{\Delta t} < 2.5$ ps and $|\Delta t| < 20$ ps. For 66% of this subset, both tracks have hits in the inner three SVT layers, which results in a mean $\Delta t$ resolution that is comparable to decays with the vertex directly reconstructed from charged particles originating at the $B$ decay point [10]. In the remainder of the subset, the resolution is nearly 2 times worse.

We obtain the PDF for the time-dependence of signal decays from the convolution of Eq. (1) with a resolution function $R(\delta t = \Delta t - \Delta t_{\text{true}}; \sigma_{\Delta t})$. The resolution function is parameterized as the sum of a “core” and a “tail” Gaussian function, each with a width and mean proportional to the reconstructed $\sigma_{\Delta t}$, and a third Gaussian centered at zero with a fixed width of 8 ps [10]. Using simulated data, we have verified that the parameters of $R(\delta t, \sigma_{\Delta t})$ for $B^0 \rightarrow K^{*0} \gamma$ decays and the $B\bar{B}$ backgrounds are similar to those obtained from the $B_{\text{tag}}$ sample, even though the distributions of $\sigma_{\Delta t}$ differ considerably. Therefore, we extract these parameters from a fit to the $B_{\text{tag}}$ sample. We find that the $\Delta t$ distribution of continuum background candidates is well described by a delta function convoluted with a resolution function with the same functional form as used for signal events. We determine the parameters of the background function in the fit to the $B^0 \rightarrow K^{*0} \gamma(K^{*0} \rightarrow K_S^0 \pi^0)$ data set.

To extract the CPV asymmetries we maximize the logarithm of the likelihood function

$$L(S_J, C_J, N_h, f_B, e_{q\bar{q}}, \bar{\alpha}) = e^{-(N_S+N_B+N_{q\bar{q}})} \prod_{\epsilon \in \epsilon/\Delta \epsilon} \left[ N_S f_S e^{S_i} \mathcal{P}_S(\bar{x}_i, \bar{y}_i; S_J, C_J) + N_B f_B e^{B_i} \mathcal{P}_B(\bar{x}_i, \bar{y}_i) + N_{q\bar{q}}(1-f_{q\bar{q}}) e_{q\bar{q}} \mathcal{P}_{q\bar{q}}(\bar{x}_i, \bar{y}_i; \bar{\alpha}) \right]$$

where the second (third) factor on the right hand side is the contribution from events with (without) $\Delta t$ information. The vectors $\bar{x}_i$ and $\bar{y}_i$ represent the time-structure and remaining observables, respectively, for event $i$. The PDFs

$$\mathcal{P}_h(\bar{x}_i, \bar{y}_i) = P_h(m_{h_{\epsilon/\Delta \epsilon}}) P_h(\Delta E_i) P_h(\epsilon_i) P_h(M_{K^{*0}_i}) \times P^c_\epsilon(\epsilon_i; \epsilon, \Delta \epsilon, T_i)$$

and

$$\mathcal{P}_h^c(\bar{y}_i) = P_h(m_{h_{\epsilon/\Delta \epsilon}}) P_h(\Delta E_i) P_h(\epsilon_i) P_h(M_{K^{*0}_i}) P^c_\epsilon(T_i)$$

are the products of the PDFs described above for hypothesis $h$ of signal ($S$), $B\bar{B}$ background ($B\bar{B}$), and continuum background ($q\bar{q}$). With the CPV asymmetries $S_J$ and $C_J$, the fit extracts the yields $N_S$, $N_{B\bar{B}}$, and $N_{q\bar{q}}$, the fractions of events with $\Delta t$ information $f_S$ and $f_{q\bar{q}}$, and the parameters $\bar{\alpha}$ which describe the background PDFs. We determine $e_B$ and $f_{B\bar{B}}$ in simulated $B\bar{B}$ decays to all final states.

The fit to the data sample yields $S_{K^- \gamma} = 0.25 \pm 0.63 \pm 0.14$ and $C_{K^- \gamma} = -0.57 \pm 0.32 \pm 0.09$, where the uncertainties are statistical and systematic, respectively. The fit reports a correlation of 1% between these parameters. The systematic uncertainties are described below. The result for $C_{K^- \gamma}$ is consistent with a fit that does not employ $\Delta t$ information. Since the present measurements of $A_{K^{*0} \gamma}$ [6,7] are consistent with zero, we also fit the data sample with $C_{K^- \gamma}$ fixed to zero and obtain $S_{K^- \gamma} = 0.25 \pm 0.65 \pm 0.14$.

The event selection criteria employed to isolate signal-enhanced samples displayed in Fig. 1 are based on a cut on the likelihood ratio $R = \mathcal{P}_S/(\mathcal{P}_S + \mathcal{P}_{B\bar{B}} + \mathcal{P}_{q\bar{q}})$ calculated without the displayed observable. The dashed and solid curves indicate background and signal-plus-background contributions, respectively, as obtained from the fit but corrected for the selection efficiency of $R$. Figure 2 shows distributions of $\Delta t$ for $B^0$- and $\overline{B}^0$-tagged events, and the asymmetry...
We estimate the impact of potential biases in the systematic uncertainty of CPV asymmetries in the B estimate an uncertainty of 0.12 (0.03) due to potential tagging asymmetries in the signal and 0.02 (0.02) due to possible asymmetries in the rate of B decays with K\(\rightarrow K^{\pm} \gamma(K^{\pm} \rightarrow K^{\pm}_S \pi^0)\) decays. Our measurement is consistent with the SM expectation of very small CPV asymmetries.

We are grateful for the excellent luminosity and machine conditions provided by our PEP-II colleagues and for the substantial dedicated effort from the computing organizations that support BABAR. The collaborating institutions wish to thank SLAC for its support and kind hospitality. This work is supported by DOE and NSF (U.S.A), NSERC (Canada), IHEP (China), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MIST (Russia), and PPARC (United Kingdom). Individuals have received support from CONACyT (Mexico), A.P. Sloan Foundation, Research Corporation, and Alexander von Humboldt Foundation.

*Present address: Department of Physics, University of Warwick, Coventry, United Kingdom.
†Also at Università della Basilicata, Potenza, Italy.
‡Also at IFIC, Instituto de Física Corpuscular, CSIC-Universidad de Valencia, Valencia, Spain.
§Deceased.

3] Unless explicitly stated, charge conjugate decay modes are assumed throughout this Letter.

FIG. 2. Distributions of \(\Delta t\) for events enhanced in signal decays with \(B_{\text{tag}}\) tagged as (a) \(B^0\) or (b) \(\overline{B}^0\), and (c) the resulting asymmetry \(A_{K^{\pm} \gamma}(\Delta t)\). The dashed and solid curves represent the fitted background and signal-plus-background contributions, respectively, as obtained from the maximum-likelihood fit. The raw asymmetry projection corresponds to approximately 38 signal and 19 background events.

\[ A_{K^{\pm} \gamma}(\Delta t) = \frac{N_{B^0} - N_{\overline{B}^0}}{N_{B^0} + N_{\overline{B}^0}} \] as a function of \(\Delta t\), also for a signal-enhanced sample.