Search for Charged Lepton Flavor Violation in Narrow Upsilon Decays


(The BABAR Collaboration)
this conservation law. However, SM processes involving charged lepton flavor violation (CLFV) remain unobservable because they are suppressed by the quantity $(\Delta m^2_{\nu}/M_W^2) \lesssim 10^{-48}$ [1,3]. Here $\Delta m^2_{\nu}$ is the difference between the squared masses of neutrinos of different flavors and $M_W$ is the charged weak vector boson mass. Hence CLFV represents an unambiguous signature of

Charged lepton flavor violating processes are unobservable in the standard model, but they are predicted to be enhanced in several extensions to the standard model, including supersymmetry and models with leptoquarks or compositeness. We present a search for such processes in a sample of $99 \times 10^6 \phi(2S)$ decays and $117 \times 10^6 \phi(3S)$ decays collected with the BABAR detector. We place upper limits on the branching fractions $\mathcal{B}(\tau \to e\nu\ell^-)$ and $\mathcal{B}(\tau \to \mu\nu\ell^-)$ ($\ell = e, \mu$) at the $10^{-6}$ level and use these results to place lower limits of order 1 TeV on the mass scale of charged lepton flavor violating effective operators.


In the original formulation of the standard model (SM) in which neutrinos are massless, lepton flavor is an accidentally conserved quantum number. The extension of the SM to include neutrino masses introduces oscillations between the neutrino flavors, which violate this conservation law. However, SM processes involving charged lepton flavor violation (CLFV) remain unobservable because they are suppressed by the quantity $(\Delta m^2_{\nu}/M_W^2) \lesssim 10^{-48}$ [1,3]. Here $\Delta m^2_{\nu}$ is the difference between the squared masses of neutrinos of different flavors and $M_W$ is the charged weak vector boson mass. Hence CLFV represents an unambiguous signature of

Charged lepton flavor violating processes are unobservable in the standard model, but they are predicted to be enhanced in several extensions to the standard model, including supersymmetry and models with leptoquarks or compositeness. We present a search for such processes in a sample of $99 \times 10^6 \phi(2S)$ decays and $117 \times 10^6 \phi(3S)$ decays collected with the BABAR detector. We place upper limits on the branching fractions $\mathcal{B}(\tau \to e\nu\ell^-)$ and $\mathcal{B}(\tau \to \mu\nu\ell^-)$ ($\ell = e, \mu$) at the $10^{-6}$ level and use these results to place lower limits of order 1 TeV on the mass scale of charged lepton flavor violating effective operators.


In the original formulation of the standard model (SM) in which neutrinos are massless, lepton flavor is an accidentally conserved quantum number. The extension of the SM to include neutrino masses introduces oscillations between the neutrino flavors, which violate this conservation law. However, SM processes involv
new physics (NP) [4–7]. Many extensions to the SM, including supersymmetry and models with leptoquarks or compositeness, predict an enhancement in the rates for these processes at levels close to experimental sensitivity. There have been considerable efforts in searches for CLFV in decays of particles such as $\mu$ and $\tau$ leptons and $B$ and $K$ mesons, but CLFV in the $\Upsilon$ sector remains relatively unexplored [8]. By using unitarity considerations, limits on CLFV $\tau$ branching fractions [9] have been used to place indirect limits on CLFV $\Upsilon$ branching fractions at the $\mathcal{O}(10^{-3})$ level [10]. In this letter we describe a search for CLFV $\Upsilon$ decays, which is a thousand times more sensitive than these indirect limits, using data collected with the BaBar detector at the PEP-II B factory at SLAC National Accelerator Laboratory. Since these decays are in general mediated by new particles produced off-shell in loops, their measurement probes mass scales up to the TeV scale, far exceeding the $e^+e^-$ center-of-mass (CM) collision energy $\sqrt{s} = M_{\Upsilon(nS)} = \mathcal{O}(10 \text{ GeV})$ [11]. Therefore, this analysis provides a NP probe which is complementary to direct searches ongoing at the Tevatron and to be performed at the Large Hadron Collider.

Assuming that the partial widths for CLFV $\Upsilon$ decays are comparable at the $\Upsilon(2S)$, $\Upsilon(3S)$ and $\Upsilon(4S)$ resonances, the branching fractions for rare decays of the narrow $\Upsilon(nS)$ resonances (henceforth $n = 2, 3$) are enhanced by approximately $\Gamma_{\Upsilon(4S)}/\Gamma_{\Upsilon(nS)} = \mathcal{O}(10^3)$ [12] with respect to those of the $\Upsilon(4S)$. We search for the CLFV decays $\Upsilon(nS) \rightarrow e^+\tau^-$ and $\Upsilon(nS) \rightarrow \mu^+\tau^-$, while the decay $\Upsilon(nS) \rightarrow e^+\mu^-$ is constrained by unitarity considerations to be less than $\mathcal{O}(10^{-8})$ [10]. No signal is expected in data collected at the $\Upsilon(4S)$ since the CLFV branching fractions are strongly suppressed, or in data collected away from the $\Upsilon$ resonances, since this data contains very few $\Upsilon$ decays. We search for CLFV in a sample of $(98.6 \pm 0.9) \times 10^6 \Upsilon(2S)$ decays and $(116.7 \pm 1.2) \times 10^6 \Upsilon(3S)$ decays corresponding to integrated luminosities of $13.6 \text{ fb}^{-1}$ and $26.8 \text{ fb}^{-1}$, respectively. Data collected at the $\Upsilon(4S)$ after the upgrade of the muon detector system (77.7 $\text{ fb}^{-1}$) and data collected 30 MeV below the $\Upsilon(2S)$ and $\Upsilon(3S)$ resonances (off-peak data corresponding to 2.6 $\text{ fb}^{-1}$ and 1.3 $\text{ fb}^{-1}$, respectively) constitute control samples that are used to validate the fit procedure. An additional data control sample collected at the $\Upsilon(3S)$ resonance (1.2 $\text{ fb}^{-1}$) is used in a preliminary unblinded analysis to validate the analysis procedure and to ensure agreement between data and events simulated using Monte Carlo (MC) techniques. Simulated background processes consisting of continuum QED events [14, 15] and generic $\Upsilon(nS)$ decays, as well as signal $\Upsilon(nS) \rightarrow e^+\tau^-$ ($\ell \equiv e, \mu$) decays [16], are produced and analyzed to optimize the fit procedure. The GEANT4 [17] software is used to simulate the interactions of particles traversing the BaBar detector, which is described in detail elsewhere [18].

The signature for $\Upsilon(nS) \rightarrow e^+\tau^-$ events consists of exactly two oppositely charged particles: a primary lepton, an electron (muon) for the $\Upsilon(nS) \rightarrow e^+\tau^-$ ($\Upsilon(nS) \rightarrow \mu^+\tau^-$) search, with momentum close to the beam energy $E_B = \sqrt{s}/2$, and a secondary charged lepton or charged pion from the $\tau$ decay. Here and in the following all quantities are defined in the CM frame unless otherwise specified. If the $\tau$ decays to leptons, we require that the primary and $\tau$-daughter leptons are of different flavor. If the $\tau$ decays to hadrons, we require one or two additional neutral pions from this decay. These requirements on the identified particle types are necessary to suppress Bhabha and $\mu$-pair backgrounds. Thus we define four signal channels, consisting of leptonic and hadronic $\tau$ decay modes for the $\Upsilon(nS) \rightarrow e^+\tau^-$ and $\Upsilon(nS) \rightarrow \mu^+\tau^-$ searches, hereafter referred to as the leptonic and hadronic $e\tau$ and $\mu\tau$ channels. The main source of background to our events comes from $\tau$-pair production, for which the final state particles are the same as for the signal. There is a background contribution to the $e\tau$ channels from Bhabha events in which one of the electrons is misidentified, and to the $\mu\tau$ channels from $\mu$-pair events in which one of the muons is misidentified or decays in flight, or an electron is generated in a material interaction. An additional background consisting of events with multiple pions and possible additional photons (‘$\pi$-hadron background’), in which a charged pion is misidentified as a lepton and the remaining particles pass the selection criteria for the $\tau$ decay products, contributes to the hadronic $e\tau$ and $\mu\tau$ channels.

In order to reduce background, we first apply requirements common to all the decay modes and then a channel specific selection. All events are required to have exactly two tracks of opposite charge, both consistent with originating from the primary interaction point and with opening angle greater than 90°. To suppress Bhabha and $\mu$-pair backgrounds, we require that $M_{\text{miss}}/\sqrt{s} < 0.95$, where $M_{\text{miss}}$ is the invariant mass of the sum of the 4-vectors of the two charged particles and of all photon candidates in the event. To ensure that the missing momentum is not pointing toward the holes in the detector near the beamline, we require that $\cos(\theta^\text{lab}_{\text{miss}}) < 0.9$ and $\cos(\theta^\text{CM}_{\text{miss}}) > -0.9$, where $\theta^\text{lab}_{\text{miss}}$ ($\theta^\text{CM}_{\text{miss}}$) is the polar angle of the missing momentum in the lab (CM) frame. To suppress two-photon processes, we require that $(p_1 + p_2)_L/|\sqrt{s} - |p_1| - |p_2|| > 0.2$, where $p_1$ and $p_2$ are the momenta of the two charged particles and $L$ indicates the transverse component with respect to the beam axis.

Particle identification is performed using a multivariate analysis [19] which uses measurements from all of the detector subsystems. An electron selector and a muon veto, combined with the requirement that the particle falls within the angular acceptance of the electromagnetic calorimeter (EMC), are used to identify electrons. A muon selector and an electron veto are used to identify muons, while a charged pion selector, an electron
veto and a muon veto are used to identify charged pions. The particle misidentification efficiencies are $O(10^{-6})$ ($\mu \to e$), $O(10^{-5})$ ($e \to \mu$) and $O(10^{-1})$ ($\ell \to \pi$), where $x \to y$ indicates that a particle of type $x$ is misidentified as a particle of type $y$. A photon candidate must deposit at least 50 MeV in the EMC and have a shower profile consistent with that expected from an electromagnetic shower. All pairs of photons with an invariant mass between 0.11 GeV and 0.16 GeV are selected as neutral pion candidates.

The channel specific selection classifies events into one of the four signal channels. The momentum of the primary lepton normalized to the beam energy is required to satisfy $x \equiv |p_1|/E_B > 0.75$. For the hadronic $\tau$-decay channels, the momentum of the $\tau$-daughter charged pion is required to satisfy $|p_2|/E_B < 0.8$. Since these $\tau$ decays to hadronic final states are dominated by the decays $\tau^\pm \to \rho^\pm \nu_\tau$ and $\tau^\pm \to a_1^\pm \nu_\tau$, the masses of the $\pi^\pm \pi^0$ and $\pi^\pm \pi^0 \pi^0$ systems are required to be consistent with the masses $m_\rho = 0.77$ GeV and $m_{a_1} = 1.26$ GeV, respectively, where the requirement on the $\pi^\pm \pi^0 \pi^0$ system mass is included only if there are two neutral pions in the event. In order to suppress Bhabha events in which an electron is misidentified as a muon, for the leptonic $e\tau$ channel the $\tau$-daughter muon is required to penetrate deeply into the muon detector. In order to suppress $\mu$-pair events in which the tracks are back-to-back, for the leptonic $\mu\tau$ channel the tracks are required to satisfy $\Delta \phi < 172^\circ$, where $\Delta \phi$ is the difference between the track azimuthal angles. After including all selection requirements, typical signal efficiencies determined from MC are $(4 - 6)\%$ [20], including the $\tau$ decay branching fractions. The typical number of events passing the selection criteria is $(10 - 15) \times 10^3$ for $\Upsilon(2S)$ data and $(20 - 30) \times 10^3$ for $\Upsilon(3S)$ data, depending on the signal channel. These yields are consistent with background expectations from MC simulations.

After selection an unbinned extended maximum likelihood fit is performed to the distribution of the discriminant variable $x$. The signal peaks at $x \approx 0.97$, while the $\tau$-pair background $x$ distribution is smooth and approaches zero as $x \to x_{MAX}$, where $x_{MAX} \approx 0.97$ is the effective kinematic endpoint for the lepton momentum in the decay $\tau^\pm \to \ell^\pm \nu_\ell \nu_\tau$, boosted into the $\Upsilon(nS)$ rest-frame. The $x$ distributions for the Bhabha/$\mu$-pair backgrounds have a peaking component near $x = 1$, about $(2.5-3)\sigma_x$ above the signal peak, where $\sigma_x \approx 0.01$ denotes the detector $x$ resolution. The $x$ distribution for the $\pi$-hadron background is smooth and falls off sharply near $x = x_{MAX}$. Probability density functions (PDFs) for signal, $\tau$-pair, Bhabha/$\mu$-pair and $\pi$-hadron backgrounds are determined as discussed below, and a PDF consisting of the sum of these components weighted by their yields is fitted to the data for each signal channel, with the yields of the components allowed to vary in the fit.

The PDFs for signal and Bhabha/$\mu$-pair backgrounds are extracted from fits to the $x$ distributions of MC events. The signal is modeled by a modified Gaussian with low- and high-energy tails, hereafter referred to as a double Crystal Ball [21] function, which peaks near $x = 0.97$. The Bhabha and $\mu$-pair backgrounds have a threshold component truncating near $x = 1$, which is modeled by an ARGUS distribution [22], and a peaking component near $x = 1$, which is modeled by a Gaussian function. The $\pi$-hadron PDF is determined from data by modifying the selection to require that the primary lepton is instead identified as a charged pion. The resulting binned $x$ distribution is scaled by the probability for pions to be misreconstructed as charged leptons, as measured in data, to yield a binned PDF for the $\pi$-hadron background. The yield of this component is fixed in the maximum likelihood fit and an uncertainty of 10% is assessed. The $\tau$-pair background is modeled by the convolution of a polynomial, which vanishes above the kinematic endpoint $x_{MAX}$, and a detector resolution function. The detector resolution function is modeled by a double Crystal Ball function whose shape is extracted from $\tau$-pair MC events. Since the signal peaks in the region near the kinematic endpoint of the $\tau$-pair background $x$ distribution, the signal yield depends strongly on $x_{MAX}$, which must therefore be extracted from data. The value of this parameter is extracted from fits to the $\Upsilon(4S)$ data control sample and corrected for differences in the decay kinematics at the $\Upsilon(4S)$ vs. $\Upsilon(nS)$ resonances. The polynomial shape parameters, which are not strongly correlated with the signal yield, are allowed to vary in the fits to $\Upsilon(nS)$ data.

To validate the fit procedure, we perform fits to data control samples in order to verify that signal yields con-
sistent with zero are obtained. The \( \Upsilon(4S) \) data is divided into samples that are chosen to be comparable in size to the \( \Upsilon(2S) \) and \( \Upsilon(3S) \) data samples. The off-peak data and the 1.2 fb\(^{-1} \) of \( \Upsilon(3S) \) data constitute additional data control samples. Results consistent with zero signal yield are obtained for all signal channels in these data control samples.

The branching fraction \( B \) is calculated from the extracted signal yield \( N_{SIG} \) according to \( B = N_{SIG}/(\epsilon_{SIG} \times N_{\Upsilon(nS)}) \), where \( \epsilon_{SIG} \) is the signal selection efficiency and \( N_{\Upsilon(nS)} \) is the number of collected \( \Upsilon(nS) \) decays. The dominant systematic uncertainties in the signal yields, which arise from uncertainties in the PDF shapes, are determined by varying the shape parameters while taking into account the correlations between them. This uncertainty is 3-10 events depending on the signal channel, and the largest contribution is due to the uncertainty in the kinematic endpoint parameter \( x_{MAX} \). To assess the uncertainty in the signal efficiency, we take the relative difference between the yields for data and MC events from a portion of the sideband of the \( x \) distribution defined by \( 0.8 < x < 0.9 \), which is dominated by \( \tau \)-pair events. This difference is due to particle identification, tracking, trigger, and kinematic selection efficiency uncertainties. There is an additional statistical uncertainty in the signal efficiency, as well as an uncertainty arising from the relative difference between the yields for data and MC events. The total signal efficiency uncertainties are \( (2 - 4)\% \), depending on the signal channel. The uncertainty on the number of collected \( \Upsilon(nS) \) decays is approximately 1%.

To assess the possible bias in the fit procedure, several hundred simulated experiments are produced with the generated signal yield fixed to the larger of the value extracted by the fit to \( \Upsilon(nS) \) data, or zero. The bias is consistent with zero within the uncertainty of 0.2-0.7 events, depending on the signal channel. There is also an uncertainty resulting from a correction in the signal yield which is performed to compensate for primary leptons whose momentum is poorly measured. These particles populate a broad momentum range and some fall in the signal region defined as the interval within \( \pm 1.5\sigma \) of the signal peak. The number of these events is estimated using \( \tau \)-pair MC simulation and corrected using data and MC Bhabha and \( \mu \)-pair control samples. The expected contributions are subtracted from the signal yields extracted by the fit and an uncertainty of 100% times the correction is assessed. The corrections are approximately 3 events (5 events) for the \( \Upsilon(2S) \rightarrow \mu^{\pm}\tau^{\mp} \) (\( \Upsilon(3S) \rightarrow \mu^{\pm}\tau^{\mp} \)) channels and less than 1 event for the \( \Upsilon(nS) \rightarrow e^{\pm}\tau^{\mp} \) channels.

The maximum likelihood fit results for a sample channel are displayed in Fig. \( \text{[1]} \) and the fit results for all channels are available in \( \text{[20]} \). After including statistical and systematic uncertainties, the extracted signal yields for all channels are consistent with zero within \( \pm 1.8\sigma \). We conclude that no statistically significant signal is observed and determine 90% confidence level (CL) upper limits (UL) using a Bayesian technique, in which the prior likelihood is uniform in \( B \) and assumes that \( B > 0 \). The resulting ULs, summarized in Table \( \text{[1]} \), are \( \mathcal{O}(10^{-6}) \) and represent the first constraints on \( B(\Upsilon(nS) \rightarrow \ell^{\pm}\tau^{\mp}) \). These results improve the sensitivity by factors of 3.7 and 5.5, respectively, with respect to the previous ULs on \( B(\Upsilon(2S) \rightarrow \mu^{\pm}\tau^{\mp}) \) and \( B(\Upsilon(3S) \rightarrow \mu^{\pm}\tau^{\mp}) \) \( \text{[8]} \).

Our results can be used to constrain NP using effective field theory. The CLFV \( \Upsilon(nS) \) decays may be parameterized as an effective \( b \bar{b} \ell^{\pm}\tau^{\mp} \) 4-fermion interaction given by \( \text{[8]} \):

\[
\Delta L = \frac{4\pi \alpha_{\ell\tau}}{A_{\ell\tau}^2} \langle \bar{b} \Gamma_{\mu} (\tau) \rangle \langle \bar{b} \gamma^{\mu} b \rangle,
\]

where \( \Gamma_{\mu} \) is a vector or an axial current or their combination, \( \alpha_{\ell\tau} \) and \( A_{\ell\tau} \) are the NP coupling constant and mass scale, respectively. This allows the following relation to be derived \( \text{[23]} \):

\[
\frac{\alpha_{\ell\tau}^2}{A_{\ell\tau}^2} = \frac{B(\Upsilon(nS) \rightarrow \ell^{\pm}\tau^{\mp})}{B(\Upsilon(nS) \rightarrow \ell^{\pm}\ell^{-})} \frac{2q_b^2 \alpha^2}{(M_{\Upsilon(nS)})^4}.
\]

Here \( q_b = -1/3 \) is the charge of the \( b \) quark, \( \alpha = \alpha(M_{\Upsilon(nS)}) \) is the fine structure constant evaluated at the \( \Upsilon(nS) \) mass, and we take the dilepton branching fraction \( B(\Upsilon(nS) \rightarrow \ell^{\pm}\ell^{-}) \) from the average of the \( \Upsilon(nS) \) dielectron and dimuon branching fractions \( \text{[12]} \). Using these values and taking into account the 7% uncertainty in \( B(\Upsilon(nS) \rightarrow \ell^{\pm}\ell^{-}) \), we determine the likelihood as a function of the quantity \( \alpha_{\ell\tau}^2/A_{\ell\tau}^2 \) and extract the 90% CL UL using the same Bayesian method as above. We use these results to exclude regions of the \( \alpha_{\ell\tau} \) vs. \( \alpha_{\ell\tau} \) parameter spaces as shown in Fig. \( \text{[2]} \). Assuming \( \alpha_{\ell\tau} = \alpha_{\mu\tau} = 1 \), these results translate to the 90% CL lower limits \( \alpha_{\tau\tau} > 1.6 \) TeV and \( \alpha_{\mu\tau} > 1.7 \) TeV on the mass scale of NP contributing to CLFV \( \Upsilon(nS) \) decays, which improve upon the previous lower limit on \( \alpha_{\mu\tau} \) \( \text{[8]} \).

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

<table>
<thead>
<tr>
<th>( B(\Upsilon(nS) \rightarrow \ell^{\pm}\tau^{\mp}) )</th>
<th>UL (10^{-6})</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.6^{+1.5+0.5}_{-1.4-0.6} \times 3</td>
<td>&lt; 3.2</td>
</tr>
<tr>
<td>0.2^{+1.5+1.0}_{-1.3-1.2} \times 3</td>
<td>&lt; 3.3</td>
</tr>
<tr>
<td>1.8^{+1.7+0.8}_{-1.4-0.7} \times 3</td>
<td>&lt; 4.2</td>
</tr>
<tr>
<td>-0.8^{+1.5+1.4}_{-1.5-1.3} \times 3</td>
<td>&lt; 3.1</td>
</tr>
</tbody>
</table>

TABLE I: Branching fractions and 90% CL ULs for signal decays. The first error is statistical and the second is systematic.
FIG. 2: Excluded regions of effective field theory parameter spaces of mass scale $\Lambda$ versus coupling constant $\alpha$. The dotted blue line is derived from $T(2S)$ results only, the dashed red line is derived from $T(3S)$ results only, and the solid black line indicates the combined results. The yellow shaded regions are excluded at 90% CL.

NSF (USA), NSERC (Canada), CEA and CNRS-IN2P3 (France), BMBF and DFG (Germany), INFN (Italy), FOM (The Netherlands), NFR (Norway), MES (Russia), MEC (Spain), and STFC (United Kingdom). Individuals have received support from the Marie Curie EIF (European Union) and the A. P. Sloan Foundation.

** Now at Temple University, Philadelphia, Pennsylvania 19122, USA
† Also with Università di Perugia, Dipartimento di Fisica, Perugia, Italy
‡ Also with Università di Roma La Sapienza, I-00185 Roma, Italy
§ Now at University of South Alabama, Mobile, Alabama 36688, USA
∗ Also with Laboratoire de Physique Nucléaire et de Hautes Energies, IN2P3/CNRS, Université Pierre et Marie Curie-Paris6, Université Denis Diderot-Paris7, F-75252 Paris, France
∗∗ Also with Università di Sassari, Sassari, Italy
†† Deceased

[20] Additional plots are available in the Appendix.
APPENDIX: EPAPS MATERIAL

The following includes supplementary material for the Electronic Physics Auxiliary Publication Service.

TABLE II: Signal efficiencies $\epsilon_{SIG}$ and signal yields $N_{SIG}$ extracted by the maximum likelihood fit for the four signal channels in $\Upsilon(2S)$ and $\Upsilon(3S)$ data. The first error is statistical and the second is systematic.

<table>
<thead>
<tr>
<th></th>
<th>lepton $e\tau$</th>
<th>hadron $e\tau$</th>
<th>lepton $\mu\tau$</th>
<th>hadron $\mu\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\Upsilon(2S)$ $\epsilon_{SIG}$ (%)</td>
<td>5.10 ± 0.03 ± 0.10</td>
<td>5.53 ± 0.03 ± 0.16</td>
<td>4.22 ± 0.03 ± 0.11</td>
<td>6.08 ± 0.03 ± 0.12</td>
</tr>
<tr>
<td>$N_{SIG}$</td>
<td>-8 ± 8 ± 3</td>
<td>19 ± 12 ± 4</td>
<td>-11 ± 8 ± 5</td>
<td>17 ± 13 ± 7</td>
</tr>
<tr>
<td>$\Upsilon(3S)$ $\epsilon_{SIG}$ (%)</td>
<td>5.24 ± 0.05 ± 0.17</td>
<td>5.56 ± 0.05 ± 0.12</td>
<td>4.20 ± 0.05 ± 0.06</td>
<td>6.07 ± 0.06 ± 0.10</td>
</tr>
<tr>
<td>$N_{SIG}$</td>
<td>24 ± 13 ± 4</td>
<td>-6 ± 13 ± 6</td>
<td>-11 ± 10 ± 9</td>
<td>4 ± 16 ± 11</td>
</tr>
</tbody>
</table>

FIG. 3: Maximum likelihood fit results for the four signal channels in $\Upsilon(2S)$ data (left) and $\Upsilon(3S)$ data (right). The red dotted line indicates the signal PDF, the green dashed line indicates the sum of all background PDFs and the solid blue line indicates the sum of these components. The inset shows a close-up of the region $0.95 < x < 1.02$. The pull denotes the difference in each bin between the data and total PDF, normalized to the statistical uncertainty in data. The corrected signal yield $N_{SIG}$ is displayed with statistical and systematic errors.