Abstract

A good knowledge of the luminosity spectrum is mandatory for many measurements at future $e^+e^-$ colliders. As the beam-parameters determining the luminosity spectrum cannot be measured precisely, the luminosity spectrum has to be measured through a gauge process with the detector. The measured distributions, used to reconstruct the spectrum, depend on Initial State Radiation, cross-section, and Final State Radiation. To extract the basic luminosity spectrum, a parametric model of the luminosity spectrum is created, in this case the spectrum at the 3 TeV CLIC. The model is used in a reweighting technique to extract the luminosity spectrum from measured Bhabha event observables, taking all relevant effects into account. The centre-of-mass energy spectrum is reconstructed within 5% over the full validity range of the model. The reconstructed spectrum does not result in a significant bias or systematic uncertainty in the exemplary physics benchmark process of smuon pair production.
1. Introduction

Small, nanometre-sized beams are necessary to reach the required luminosity at future linear colliders. Together with the high energy, the small beams cause large electromagnetic fields during the bunch crossing. These intense fields at the interaction point squeeze the beams. This so-called pinch effect increases the instantaneous luminosity. However, the deflection of the particles also leads to the emission of Beamstrahlung photons – which reduce the nominal energy of colliding particles – and produces collisions below the nominal centre-of-mass energy [1, 2, 3, 4]. The resulting distribution of centre-of-mass energies is the luminosity spectrum.

The knowledge of the shape of this luminosity spectrum is mandatory for the precision measurements in which a cross-section has to be known. While the cross-section depends on the centre-of-mass energy, the observables measured in the lab-frame also depend on the difference in energy of the colliding electrons\(^1\), which determines the Lorentz-boost of the system.

Unlike the electron structure functions (i.e, Initial State Radiation (ISR)) – which can be calculated precisely – the beam-beam forces, and therefore the Beamstrahlung, strongly depend on the geometry of the colliding bunches. The actual beam-beam interaction taking place at the interaction point cannot be precisely simulated, because the geometry of the bunches cannot be measured. Therefore, the luminosity spectrum at the interaction point has to be measured using a physics channel with well known properties, e.g., Bhabha scattering.

The observables measured in the events are affected by detector resolutions. The distributions used for the reconstruction of the luminosity spectrum also depend on the cross-section of the process, and Initial and Final State Radiation (FSR). All effects have to be taken into account for the reconstruction of the luminosity spectrum.

It was pointed out by Frary and Miller [5] that a precise reconstruction of the peak of the luminosity spectrum, necessary for a top-quark threshold scan, can only be achieved with a measurement of the angles of the outgoing electrons from Bhabha scattering. The angles of the two particles are the most precisely measurable observable [5]. The angles of the outgoing electrons – or rather the acollinearity between the two particles – are sufficient to extract a relative centre-of-mass energy, which gives access to the luminosity spectrum.

Toomi et al. [6] showed that the reconstruction of a parametrised luminosity spectrum is possible using a template fit. In their parametrisation only three parameters were used to describe the effective centre-of-mass energy spectrum. However, as the boost of the initial system and correlation between the energies of the two particles cannot be neglected [7], a description of the energies of the pairs of colliding particles is necessary.

The reconstructed relative centre-of-mass energy from the acollinearity is equal to unity for back-to-back particles, and always smaller than unity for larger acollinearity, regardless whether one of the particles has a higher or lower energy than nominal. Therefore, Shibata et al. [8] proposed to calculate the distribution of the four-vectors of the Bhabha electrons and extract the luminosity spectrum with the iterative Expectation–Maximisation algorithm. However, they have considered neither detector resolutions, nor Initial and Final State Radiation. For a full description of the outgoing Bhabha electrons, the luminosity spectrum would have to be weighted with the Bhabha cross-section and convoluted with the detector resolutions, which would require

\(^1\)Unless explicitly stated, electron always refers to both electrons and positrons.
a huge computational effort, when using their method. A reconstruction of the energy of the particle pairs was done for the 500 GeV ILC [9]. The acollinearity and the energies of the electrons measured in the calorimeter were used in a reweighting fit to reconstruct the luminosity spectrum. The parametrisation – necessary for the reweighting fit – accounted for the correlation between the two beams and the beam-energy spread.

This note follows the approach of Reference [9], extends it, and applies it to the luminosity spectrum of the 3 TeV CLIC [10], which is the most challenging luminosity spectrum. This note is structured as follows: in Section 2 the basic and cross-section scaled luminosity spectrum are defined. The Bhabha scattering and observables used for the reconstruction are also introduced. In Section 3 the model of the luminosity spectrum, which is required to perform a reweighting fit, is derived. The reweighting technique is explained in Section 4, and in Section 5 it is applied to first validate the model against the luminosity spectrum at the 3 TeV CLIC; then all the relevant effects leading to the measured observables are included, and the luminosity spectrum is reconstructed from these distributions. In Section 6 the impact of the reconstructed luminosity spectrum on the measurement of the masses of supersymmetric particles in a CLIC benchmark process is estimated. The note closes with a summary, conclusions, and outlook in Section 7.

2. Luminosity Spectrum, Bhabha Scattering, and the Measurement

The nominal centre-of-mass energy \( \sqrt{s_{\text{nom}}} \) of a collider with two beams with the nominal beam energy \( E_{\text{Beam}} \) is \( \sqrt{s_{\text{nom}}} = 2E_{\text{Beam}} \). If the two interacting particles carry only a fraction of the nominal beam energy \( x_{1,2} = E_{1,2}/E_{\text{Beam}} \), the effective centre-of-mass energy becomes

\[
\sqrt{s} = E_{\text{Beam}} \sqrt{x_1x_2}.
\]

The basic luminosity spectrum describes either the distribution of the centre-of-mass energies \( \mathcal{L}(\sqrt{s}) \) or the distribution of energies of colliding particle pairs \( \mathcal{L}(E_1, E_2) \) prior to hard collisions and prior to Initial State Radiation. The two functions are connected via the integral along the lines of constant centre-of-mass energies, described by Equation 1. Therefore,

\[
\mathcal{L}(\sqrt{s}) = \int_0^{E_{\text{max}}} dE_1 \int_0^{E_{\text{max}}} dE_2 \delta(\sqrt{s} - 2\sqrt{E_1E_2}) \mathcal{L}(E_1, E_2).
\]

The luminosity spectrum affects all centre-of-mass energy dependent observables, for example the effective cross-section of any process is given by either

\[
\sigma_{\text{effective}}(\sqrt{s_{\text{nom}}}) = \int_0^{\sqrt{s_{\text{max}}}} d\sqrt{s} \mathcal{L}(\sqrt{s}) \sigma(\sqrt{s}),
\]

or for the two-dimensional luminosity spectrum

\[
\sigma_{\text{effective}}(\sqrt{s_{\text{nom}}}) = \int_0^{E_{\text{max}}} dE_1 \int_0^{E_{\text{max}}} dE_2 \mathcal{L}(E_1, E_2) \sigma(2\sqrt{E_1E_2}).
\]
Bhabha scattering is the process of choice for luminosity measurements. It can be calculated with high precision and has a large cross-section. To first order, the differential Bhabha cross-section is [11]

\[
\frac{d\sigma_{\text{Bhabha}}}{d\theta} = \frac{2\pi\alpha^2}{s} \frac{\sin \theta}{\sin^4(\theta/2)},
\]

where \(\alpha\) is the fine-structure constant and \(\theta\) the polar scattering angle.

Because cross-sections \(\sigma_i(\sqrt{s})\) depend on the centre-of-mass energy, any process used to reconstruct the basic luminosity spectrum will inherently contain a scaled luminosity spectrum

\[
L_{\text{scaled}}(\sqrt{s}) = \frac{\sigma_i(\sqrt{s})}{\int_{\sqrt{s}_{\text{min}}}^{\sqrt{s}_{\text{max}}} d\sqrt{s}' \sigma_i(\sqrt{s}') L(\sqrt{s}')},
\]

The observed centre-of-mass energy is further affected by Initial State Radiation. It is impossible to distinguish between Initial State Radiation and Beamstrahlung on an event-by-event basis. Initial State Radiation and Beamstrahlung have to be disentangled statistically.

Finally, the scattered particles are recorded in the detector, where their properties are reconstructed within the limits of the resolution of the respective sub-detector.

2.1. The Basic Luminosity Spectrum

The luminosity spectrum is a convolution of the beam-energy spread, which is inherent to the accelerator, and the Beamstrahlung due to the Beam-Beam effects. Figure 1 shows the beam-energy spread of the 3 TeV CLIC machine. It is obtained from a simulation of the main linear accelerator and the beam delivery system [12].

The energy of a particle depends on its longitudinal position in the bunch (Figure 1a). Due to intra-bunch wakefields, particles in the front of the bunch gain more energy from the RF cavities than particles in the back of the bunch [13]. This leads to the two distinct peaks near the minimal and the maximal value of the beam-energy spread (Figure 1b). The energy spread is not following a Gaussian distribution.

During the bunch crossing the intense electromagnetic fields – due to the opposing bunches – deflect the beam particles and cause Beamstrahlung.

The distribution of particles is used as the input to the beam-beam simulation. The simulation of the beam-beam effects is done with GUINEAPig [4]. GUINEAPig provides the luminosity spectrum in the form of a list of particle energy pairs. Figure 2 shows the full range and the region around the maximal energy of the two-dimensional luminosity spectrum. The square region in the distribution of the two energies is due to the beam-energy spread (see Figure 1b). Events with \(x_1 < 0.995\) or \(x_2 < 0.995\) were affected by the Beamstrahlung.

Figure 3a shows the basic luminosity spectrum with respect to the effective centre-of-mass energy \(\sqrt{s}\). The spectrum possesses a peak around the nominal centre-of-mass energy and a long tail down to less than 0.05\(\sqrt{s_{\text{nom}}}\). Figure 3b shows the peak of the luminosity spectrum produced by GUINEAPig. Because the beam-energy spread is not Gaussian, also the centre-of-mass energy peak is not Gaussian. Figure 3b also shows a spectrum obtained by randomly pairing the energies of two particles, i.e., removing the correlation between the energies of the two beams. There is a clear difference between the two cases. If the correlation between the
Figure 1: Energy distribution of the CLIC beams. (a) Dependence of the particle energy on the longitudinal position of the particles along the length of the bunch, where the beam travels towards the left. (b) The energy distribution of all particles.

Figure 2: (a) Energy spectrum of colliding particles as simulated with GUINEAPIG for 3 TeV CLIC. (b) Zoom of the luminosity spectrum around the nominal beam energies.

particle energies is not taken into account, the luminosity spectrum cannot be reconstructed properly.

The CLIC luminosity spectrum covers almost the full energy range (see Figure 3a), but the beyond the Standard Model benchmark processes considered for the conceptual design report [14] usually have a threshold close to a centre-of-mass energy of 1.5 TeV. For these processes the luminosity spectrum below the threshold is irrelevant. Whereas a parametrisation of the full luminosity spectrum down to the smallest energies requires a number of additional parameters on top of those describing the spectrum down to about half the nominal centre-of-mass energy.
Therefore, to limit the number of parameters, a cut on the centre-of-mass energy is applied. This rejects events where the Model is not valid (see Section 3.2), it decreases the run-time of the fits, and makes the generation of the scaled luminosity spectrum more efficient (see Section 2.2).

2.2. Cross-Section-Scaled Luminosity Spectrum

Because the observed events are distributed according to the scaled luminosity spectrum (Equation 6), the events provided by GUINEAPig have to be either weighted with very large weights, or sampled through an accept–reject method [15] to obtain events with equal weights. Because large weights are undesirable, the accept–reject method is chosen.

More than three orders of magnitude of the Bhabha cross-section are covered by the 3 TeV luminosity spectrum, which means the accept–reject method is very inefficient. If a very large number of events for the basic luminosity spectrum were available, the scaled luminosity spectrum could be directly sampled from them. To avoid storing the large number of basic events the accept–reject method is directly added in GUINEAPig.

For the accept–reject method the differential cross-section of the Bhabha scattering has to be known. Instead of using Equation 5 to calculate the Bhabha cross-section, it is estimated with WHIZARD [16, 17] and BHWIDE [18]. BHWIDE includes higher-order effects and Initial State Radiation. For both generators only events with the electron and positron polar angle inside the tracking acceptance ($7^\circ < \theta < 173^\circ$) are accepted. The cross-section is estimated at precise centre-of-mass energies from 10 GeV to 3000 GeV without any luminosity spectrum. Figure 4a shows the cross-section as given by WHIZARD and BHWIDE. The cross-section is enhanced by about 10% around the Z-mass peak, which is barely visible. The two cross-sections agree within a few percent with one another.

To efficiently generate events with interesting (i.e., sufficiently large) centre-of-mass energies in GUINEAPig, every event is checked repeatedly. This results in the same lower energy events
being accepted multiple times. However, a centre-of-mass energy cut of $\sqrt{s} > 1.5$ TeV rejects all of these events. To study the luminosity spectrum towards its lower end, the repeated check of events can be replaced by independent GUINEAPig runs, given that a sufficient number of input files is available, which was not the case at the moment this note was written.

Figure 4b shows the basic luminosity spectrum obtained with GUINEAPig, the bin-wise multiplication of the luminosity spectrum with the cross-section, and the scaled luminosity spectrum from GUINEAPig with the cross-section used in the accept–reject method. The last two curves are nearly identical showing that the modified GUINEAPig produces a properly scaled luminosity spectrum with equally weighted events.

2.3. Generation of Bhabha Events

The different luminosity spectra are used in the Bhabha generator to create the events which are observed in the detector. The Bhabha events are generated with BHWIDE, where the energies of the initial electron and positron can be defined on an event-by-event basis, as implemented by Rimbault et al. [19]. The polar angle $\theta$ of the final state electrons must be $7^\circ < \theta < 173^\circ$ to ensure they will be observable in the tracker. BHWIDE produces also Initial and Final State Radiation photons and accounts for their effects during the Bhabha scattering.

The cross-section – including the luminosity spectrum – for events with a centre-of-mass energy above 1.5 TeV is 11 pb, which results in more than 1 million events for an integrated luminosity of 100 fb$^{-1}$. The expectation for the 3 TeV CLIC is an integrated luminosity of 500 fb$^{-1}$ per year [14].
2.4. Observables and Detector Resolutions

Three observables, which can be extracted from the final state electrons, are used for the reconstruction of the luminosity spectrum: the relative centre-of-mass energy calculated from the polar angles of the outgoing electrons \( \sqrt{s_{\text{acol}}} \), the energy of the electron \( E_1 \), and the energy of the positron \( E_2 \). The relative centre-of-mass energy reconstructed from the acollinearity of the final state electrons is \([7, 9]\)

\[
\sqrt{s_{\text{acol}}} = \sqrt{\sin(\theta_1) + \sin(\theta_2) + \sin(\theta_1 + \theta_2) - \sin(\theta_1 + \theta_2)},
\]

(7)

where \( \theta_1 \) is the polar angle of the electron and \( \theta_2 \) that of the positron. \( \sqrt{s_{\text{acol}}} \) is equal to the effective centre-of-mass energy \( \sqrt{s'} \), if only one of the particles radiated photons – Beamstrahlung or Initial State Radiation. If both the electron and the positron radiated photons, the reconstructed centre-of-mass energy \( \sqrt{s_{\text{acol}}} \) will be larger than the effective centre-of-mass energy \( \sqrt{s'} \).

The GeANT4 simulation of tens of millions of electrons is too time-consuming. To include the detector effects, resolution functions of the energy and angles have been obtained from fully simulated and reconstructed Bhabha events using the CLIC ILD CDR detector model [20]. The dominating beam-induced background, the \( \gamma \gamma \rightarrow \text{hadron} \) events, was taken into account [21].

The rate of electrons produced in Bhabha scattering falls with an increasing polar angle \( \theta \) (cf. Equation 5) and the events will be predominantly at small polar angles. Because the magnetic field is nearly collinear to those tracks, their curvature does not allow for an accurate measurement of the momentum. Therefore, the energy is measured using only the calorimeter information. The tracking information is used to measure the angles. The energy resolution is shown in Figure 5a. It is modelled in the analysis with

\[
\frac{\sigma_E}{E} = \frac{24.3%}{\sqrt{E/\text{GeV}}} \oplus 1.23%,
\]

(8)

obtained from reconstruction in which the \( \gamma \gamma \rightarrow \text{hadron} \) background for the 3 TeV CLIC was overlaid. The angular resolution needed for the computation of the relative centre-of-mass energy depends on the energy, shown in Figure 5b. The angular resolution is better than 20 \( \mu \)rad for particle energies above 200 GeV.

The distributions of the particle energies and the relative centre-of-mass energy are shown in Figure 6. Figures 6a and 6b show the distributions before and Figures 6c and 6d after the application of the resolution effects via four-vector smearing. The relative centre-of-mass energy is hardly affected by the resolution, due to the high angular resolution of the tracking detectors. The energy of the particles is much more affected by the detector resolution.

3. Modelling the Luminosity Spectrum

For the reconstruction of the basic luminosity spectrum with the reweighting fit, a model or parametrisation of the luminosity spectrum is needed. If the beam energy were not affected by

\footnotesize{\textsuperscript{2}}Strictly speaking, \( \theta_1 \) and \( \theta_2 \) are the angles with respect the positive or negative z-axis. The angles have to fulfill \( \theta_1 + \theta_2 > \pi \) by construction.}
beam-energy spread or Beamstrahlung, it could simply be described by a delta-distribution

\[ E_{\text{Beam}}(x) = \delta(x - 1), \tag{9} \]

and the random variates for this function would be \( x_{E_{\text{Beam}}} = 1 \). The contributions from beam-energy spread \( x_{\text{Spread}} \) and Beamstrahlung \( x_{\text{Strahlung}} \) change the beam energy away from its initial value. The functions describing the two contributions will therefore describe the difference to the nominal value, and the final random variate will be

\[ x_{\text{Final}} = x_{E_{\text{Beam}}} + x_{\text{Spread}} + x_{\text{Strahlung}}. \tag{10} \]

Thus, the functions for the beam-energy spread and Beamstrahlung should describe the energy change of particles due to the respective effect.

The luminosity spectrum model will be built by describing the two particle energies. Using \( x_1 \) and \( x_2 \) as defined in Section 2, the simplest description of a two-dimensional model is \( L(x_1, x_2) = f(x_1)f(x_2) \). However, a purely factorising Ansatz is insufficient to describe the correlation between the particle energies. Therefore, the two-dimensional energy distribution is divided into four regions (see Figure 7): one region where neither particle radiated Beamstrahlung (called the ‘Peak’); two regions where one or the other particle radiated Beamstrahlung (called the ‘Arms’); and one region where both particles radiated Beamstrahlung (called the ‘Body’). This separation is only qualified by whether a particle produced Beamstrahlung or not and ignores the beam-energy spread for the moment. The result of this division is a piecewise function

\[
L(x_1, x_2) = \begin{cases} 
    f_{\text{Peak}}, & \text{for } x_1 = 1 \text{ and } x_2 = 1 \\
    f_{\text{Arm}1}, & \text{for } x_1 = 1 \text{ and } x_2 < 1 \\
    f_{\text{Arm}2}, & \text{for } x_1 < 1 \text{ and } x_2 = 1 \\
    f_{\text{Body}}, & \text{for } x_1 < 1 \text{ and } x_2 < 1.
\end{cases} \tag{11}
\]
For each region, the resulting particle energies are described by a product of the functions for the two particles $f_{\text{Region}} = f_1 \cdot f_2$, and the individual functions are convolutions of the beam-energy spread and Beamstrahlung functions, depending on the region.

### 3.1. Parametrisation of the Beam-Energy Spread

The function to describe the shape of the beam-energy spread (Figure 8a) has to rise very steeply at the two extremities. A hyperbolic cosine, parabola, or higher order polynomials (with a reasonable number of parameters, see Section 3.1.2) do not describe this energy distribution well. A beta-distribution $b(x)$ (see Equation 42 and Figure 28 in Appendix A) is used to describe the beam-energy spread. The parameters describing the form of a beta-distribution for the beam-energy spread are given by $\omega_1$ and $\omega_2$. As the beta-distribution is limited between 0
and 1, a variable transform

\[ t = \frac{x - x_{\text{min}}}{x_{\text{max}} - x_{\text{min}}} \]  

(12)

is used to describe the beam-energy spread between the two endpoints \(x_{\text{min}}\) and \(x_{\text{max}}\) near the maximal values at the beginning and end of the distribution. Here, the variable \(x\) is the relative difference between a particle’s energy \(E_{\text{Particle}}\) and the nominal beam energy \(E_{\text{Beam}}\),

\[ x = \frac{E_{\text{Particle}} - E_{\text{Beam}}}{E_{\text{Beam}}} = \frac{\Delta E}{E_{\text{Beam}}}, \]  

(13)

which corresponds to \(x_{\text{Spread}}\) from Equation 10. To also describe the particles with energy below the minimal peak and above the maximal peak, the beta-distribution is convoluted with a Gaussian distribution \(g(x)\) with a mean \(\mu = 0\) and a variable width \(\sigma\). The Beam-Energy Spread (BES) function is

\[ \text{BES}(x; \omega_1, \omega_2, \sigma) = b(x; \omega_1, \omega_2) \otimes g(x; \sigma). \]  

(14)

Due to Fubini’s theorem ([22], Appendix A) the convolution of two probability density functions always results in another probability density function.

The beam-energy spread distribution is fitted by the function \(\text{BES}(x)\) with a binned log-likelihood fit in ROOT version 5.34.01 [23]. Figure 8a shows the best fit to the beam-energy spread with this model, and the resulting parameters are given in Table 1. The distribution contains 300,000 entries.

The width of the Gaussian \(\sigma\) and the boundaries of the beam-energy spread beta-distributions \((x_{\text{min}}, x_{\text{max}})\) are fixed for all following fits. This assumes an existing knowledge of the beam-energy spread coming from the accelerator. Fixing these parameters can introduce a large systematic error, if they are not measured correctly.
Figure 8: (a) The beam-energy spread distribution from the accelerator simulation [12] and the best fit to the beam-energy spread with Equation 14. (b) Energy spread after the simulation in GUINEAPig. The distribution of the Peak requires that both particles have $E > 0.995E_{\text{Beam}}$, and the distribution in the Arms requires that one particle has $E < 0.995E_{\text{Beam}}$. The colour bands in both plots indicate the confidence interval at 99%.

Table 1: Parameters found by the fit of Equation 14 to the beam-energy spread from the accelerator simulation, and to the beam-energy spread for two different regions of the luminosity spectrum.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Factor</th>
<th>Value</th>
<th>Uncertainty</th>
<th>Value</th>
<th>Uncertainty</th>
<th>Value</th>
<th>Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\omega_1$</td>
<td>$-0.522$</td>
<td>0.001</td>
<td>$-0.333$</td>
<td>0.002</td>
<td>$-0.470$</td>
<td>0.001</td>
<td></td>
</tr>
<tr>
<td>$\omega_2$</td>
<td>$-0.409$</td>
<td>0.002</td>
<td>$-0.298$</td>
<td>0.002</td>
<td>$0.405$</td>
<td>0.004</td>
<td></td>
</tr>
<tr>
<td>$x_{\text{min}}$</td>
<td>$10^{-3}$</td>
<td>4.679</td>
<td>0.001</td>
<td>Fixed</td>
<td>Fixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$x_{\text{max}}$</td>
<td>$10^{-3}$</td>
<td>5.495</td>
<td>0.002</td>
<td>Fixed</td>
<td>Fixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$10^{-4}$</td>
<td>1.367</td>
<td>0.010</td>
<td>Fixed</td>
<td>Fixed</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$/ndf</td>
<td>764/195</td>
<td>6032/198</td>
<td>3803/198</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

3.1.1. Luminosity Weighted Beam-Energy Spread

The correlation between the particle energy and its position in the bunch causes a change in the effective beam-energy spread. The probability to radiate Beamstrahlung photons, and therefore the fraction of energy lost by particles, increases with the distance travelled in the electromagnetic field of the oncoming bunch.

Figure 8b shows two energy distributions, one of the Peak-region, where both particles possess an energy of more than 99.5% of the nominal beam energy $E_{\text{Beam}}$, and one of the Arms-region,
where only one of the particle contains more than 99.5% of the nominal beam energy. Both histograms contain 300,000 entries.

The energy spread of the Peak-region is clearly flatter than the energy spread coming from the accelerator (Figure 8a). In the column Peak, Table 1 lists also the parameters $\omega_1$ and $\omega_2$ found by fitting Equation 14 to the distribution. For this fit, the limits and the Gaussian width are fixed. For the best fit both $\omega_1$ and $\omega_2$ are closer to zero, but still negative, i.e., there are still two maxima at the bottom and top end of the distribution. The peak at the lower end of the spectrum is reduced, because the particles in the tail are less likely to interact with particles that did not radiate Beamstrahlung.

The energy spread of the Arms-region, where one of the particles radiated Beamstrahlung, shows a large peak at the lower energy, and almost no peak at the upper end of the spectrum. This is also caused by the correlation between the energy and the position in the bunch. Particles in the tail are more likely to collide with a particle that already radiated Beamstrahlung, and therefore the peak at the lower edge of the beam-energy spread is enhanced. Likewise only very few particles with the highest energy – located near the front of the bunch – interact with particles from the tail of the bunch, which leads to the disappearance of the peak at the highest energies. The beam-energy spread for the Arms-region is described by a beta-distribution for which $\omega_2 > 0$ (see column Arms in Table 1).

The $\chi^2$/ndf is larger for the fits to the luminosity weighted beam-energy spreads than for the fit to the initial beam-energy spread. The chosen function cannot perfectly model the distributions, however, the fits are only used to check qualitatively if the model can represent the luminosity spectrum at this stage. In addition, as there was only a single input file available to run GUINEAPIG, the macro-particles are re-used for luminosity events. This re-use means that the fluctuations in the number of entries are larger than what can be expected from the statistical uncertainties, which also increases the $\chi^2$/ndf.

### 3.1.2. Fitting with Chebyshev Polynomials

The convolution of the beta-distribution with a Gaussian function can be avoided by the use of Chebyshev polynomials to fit the beam-energy spread. The beam-energy spread is fitted by the function

$$f(x) = \sum_{i=0}^{N_{\text{Degrees}}-1} p_i T_i(x),$$

where $N_{\text{Degrees}}$ is the number of Chebyshev polynomial used in the fit, $p_i$ are the free parameters, and $T_i$ the $i$-th Chebyshev polynomial of the first kind. The range for $x$ is transformed into the range $-1 < x < 1$, as required by the Chebyshev polynomials. Figure 9 shows the curves for the fit with $N_{\text{Degrees}}$ equal to 5, 10, 26, and 35, and Table 2 gives the $\chi^2$/ndf for the different fits. With 5 or 10 free parameters the Chebyshev polynomials cannot reproduce the beam-energy spread. With 26 free parameters the $\chi^2$/ndf is similar to the $\chi^2$/ndf value of the fit by Equation 14, and with an even larger number of parameters the $\chi^2$/ndf approaches unity. It is also possible to limit the fit range to be between the two peaks, and convolute the function with a Gaussian function. However, this also requires a larger number of parameters than Equation 14. In summary, the possible gain by using Chebyshev polynomials is offset by the large increase in the number of parameters.
Table 2: The $\chi^2$/ndf for the different fits, one with the beta-distribution convoluted with a Gaussian function (BES, Equation 14) and for the different degrees of Chebyshev polynomials (Equation 15).

<table>
<thead>
<tr>
<th>Fit</th>
<th>$\chi^2$/ndf</th>
</tr>
</thead>
<tbody>
<tr>
<td>BES</td>
<td>764/195</td>
</tr>
<tr>
<td>Chebyshev $N_{Degrees} = 5$</td>
<td>46990/195</td>
</tr>
<tr>
<td>Chebyshev $N_{Degrees} = 10$</td>
<td>16699/190</td>
</tr>
<tr>
<td>Chebyshev $N_{Degrees} = 26$</td>
<td>671/174</td>
</tr>
<tr>
<td>Chebyshev $N_{Degrees} = 35$</td>
<td>228/165</td>
</tr>
</tbody>
</table>

Figure 9: Fit of the beam-energy spread with several Chebyshev polynomials of different degrees. The colour bands indicate the confidence interval at 99%.

Figure 10: Fit of the linear combination of one, two, and three beta-distributions (Equation 16) to the particle energy spectrum after Beamstrahlung. (a) Fit for $0.0 < x < 0.995$, (b) fit for $0.5 < x < 0.995$.  parameters.
3.2. Beamstrahlung

Following T. Ohl’s CIRCE model [24], the energy distribution of the particles after the emission of Beamstrahlung photons is modelled with a beta-distribution. Beamstrahlung will always reduce the energy of a particle, so that the variate would be in the range \(-1 < x_{\text{Strahlung}} < 0\) (cf. Equation 10). Beta-distributions are limited between 0 and 1, so that the function describing the Beamstrahlung effect will be convoluted with the \(\delta\)-distribution from Equation 9, which moves the range to \(0 < x_{\text{Strahlung}} < 1\) and no further variable-transform is necessary for the probability density function.

The parameters of the beta-distributions used to describe the energy distribution due to Beamstrahlung are called \(\eta_1\) and \(\eta_2\). The beta-distribution parameters must fulfil the conditions \(0 < \eta_1\) and \(-1 < \eta_2 < 0\) for the distribution to fall towards \(x = 0\) and rise towards \(x = 1\).

Previous studies have shown that the tail of the CLIC centre-of-mass energy distribution is better modelled by a sum of three beta-distributions [25]. Therefore, the energy distributions from Beamstrahlung are fitted by linear combinations of \(N_{\text{Beta-distributions}}\) beta-distributions

\[
b_{\text{linear}}(x) = \sum_{i=1}^{N_{\text{Beta-distributions}}} p_i b(x; [p]^i),
\]

and the constraint

\[
1 \equiv \sum_{i=1}^{N_{\text{Beta-distributions}}} p_i,
\]

where \(p_i\) are the respective fractions of the beta-distribution contribution and \([p]^i = \{\eta_1^i, \eta_2^i\}\) the parameter-set for each beta-distribution.

Figure 10 shows the fits with \(N_{\text{Beta-distributions}} = 1, 2, 3\) to the distribution of the particle energy. In Figure 10a the fit to the histogram is performed in the range of \(0.0 < x < 0.995\); above 0.995 the beam-energy spread is dominant and would have to be included for the fit. It is clearly visible that the function with three beta-distributions – with eight free parameters – shows a better agreement with the distribution than the other functions.

Figure 10b shows the same fit of linear combinations with a range limited to \(0.5 < x < 0.995\); all three fit-functions overlap. Therefore, a single beta-distribution is enough to describe the particle energy between \(0.5E_{\text{Beam}}\) and \(0.995E_{\text{Beam}}\). In this document, the Beamstrahlung is described by a single beta-distribution to reduce the number of free parameters. However, this will also limit the energy range in which our Model can be considered as valid.

3.3. The Model for the Full Luminosity Spectrum

The individual contributions discussed in the previous sections are now used to create the Model of the basic two-dimensional luminosity spectrum. As discussed in Section 3.1 the beam-energy spread \(\text{BES}(x)\) is described by a convolution of a Gaussian function \(g(x)\) and a beta-distribution \(b(x)\)

\[
\text{BES}(x) = (b \otimes g)(x).
\]
beam-energy spread with an incomplete beta-distribution with an upper limit of $\beta_{\text{limit}}^1 = 0.9999$ (see Appendix A),
\[ \text{BB}(x) = (b \otimes \text{BES})(x). \] (19)
The upper limit is chosen to be close to 1, so that the convolution with beam-energy spread causes an overlap with the Peak-region (cf. Figure 11).

To describe particle energy distributions that are only negligibly affected by the beam-energy spread, a beta-distribution with an upper limit of $\beta_{\text{limit}}^2 = 0.995$ convoluted with a Gaussian function is used
\[ \text{BG}(x) = (b \otimes g)(x). \] (20)
This upper limit separates the distribution from those strongly affected by the beam-energy spread.

As described in Equation 11, the distributions in the four different regions are described by the product of two functions, one for each particle. The explicit piecewise description shown in Equation 11, however, is replaced by the use of delta-distribution and implicit ranges of the individual functions. The Peak region is described by two pure beam-energy spread functions (Equations 14 and 18) and delta-distributions to signify the absence of Beamstrahlung; the Arms are modelled by one beam-energy spread function and a delta distribution, and one function describing Beamstrahlung convoluted with the beam-energy spread (Equation 19); the Body is described by two functions describing only the Beamstrahlung (Equation 20).

\[
\mathcal{L}(x_1, x_2) = p_{\text{Peak}} \delta(1-x_1) \otimes \text{BES}(x_1; [p]_{\text{Peak}}) + p_{\text{Arm1}} \delta(1-x_1) \otimes \text{BES}(x_1; [p]_{\text{Arm1}}) \text{BB}(x_2; [p]_{\text{Arm1}}; \beta_{\text{limit}}^1) \\
+ p_{\text{Arm2}} \text{BB}(x_1; [p]_{\text{Arm2}}; \beta_{\text{limit}}^1) \delta(1-x_2) \otimes \text{BES}(x_2; [p]_{\text{Arm2}}) \\
+ p_{\text{Body}} \text{BG}(x_1; [p]_{\text{Body}}; \beta_{\text{limit}}^2) \text{BG}(x_2; [p]_{\text{Body}}; \beta_{\text{limit}}^2),
\] (21)
with $\beta_{\text{limit}}^1 = 0.9999, \beta_{\text{limit}}^2 = 0.995$. In addition, the constraint
\[ p_{\text{Body}} = 1 - p_{\text{Peak}} - p_{\text{Arm1}} - p_{\text{Arm2}} \] (22)
has to be fulfilled, which results in
\[ \int \mathcal{L}(x_1, x_2) \, dx_1 \, dx_2 = 1, \] (23)
as required for a probability density function. The function given in Equation 21 will be used to describe the luminosity spectrum. The random variates according to the individual parts of Equation 21 are shown in Figure 11. Each line of Equation 21 corresponds to one of the distributions. Due to the convolution of beam-energy spread and Beamstrahlung functions, the region around the nominal beam energies ($x_1 = 1, x_2 = 1$) is described by a superposition of individual contributions.

### 4. Reweighting Fit

It is possible to fit Equation 21 to the basic luminosity spectrum. The convolutions with the $\delta$-distribution can be performed explicitly. The other convolutions have to be performed numer-
Figure 11: Different parts of the Model, described by Equation 21.

ica, because the convolution between the beta-distribution and the Gaussian function cannot be expressed in a closed form\(^3\).

For the implementation of the function the numerical convolutions are evaluated with the QAG\(^4\) integration algorithm [26] interfaced via the GSLIntegrator from ROOT::MathMore. The evaluation of the function takes about 160 seconds for the full range. A direct fit with the function, requiring multiple iterations, would be slow. The fitting procedure can be sped up by using a reweighting fit and by exploiting the fact that the random variates according to Equation 21 can also be described by the sum of the random variates of the individual functions (see Appendix A).

The principle of the reweighting technique is shown in Figure 12. A \(\chi^2\) minimization, via

\(^3\)We have no formal proof of this statement. However, neither the integral of the Gaussian function (resulting in the error-function) nor the integral of the beta-distribution (yielding Gamma-functions (Equation 43)) can be expressed in a closed form with a finite number of elementary functions.

\(^4\)Quadrature Adaptive General integrand
MINUIT [27] implemented in ROOT, is used to fit a sample of Model events to the GUINEAPIG sample. The procedure starts with the generation of a large number of events according to the Model. This produces events consisting of pairs of beam energies \( (x_{i1}, x_{i2}) \) and the corresponding probability \( \mathcal{L}(x_{i1}, x_{i2}; [p]^0) \) to obtain a given event. The probability depends on the initial set of parameter values \( [p]^0 \).

The minimizer is used to obtain a new set of parameter values \( [p]^N \) that results in new probabilities \( \mathcal{L}(x_{i1}, x_{i2}; [p]^N) \) for each event. The event weight

\[
  w^i = \frac{\mathcal{L}(x_{i1}, x_{i2}; [p]^N)}{\mathcal{L}(x_{i1}, x_{i2}; [p]^0)}
\]  

(24)

is used to weight each event of the Monte Carlo distribution. For every set of parameter values \( [p]^N \) a reweighted Monte Carlo distribution is obtained. The minimum \( \chi^2 \) between the distribution of the Model and the distribution from GUINEAPIG corresponds to the optimal parameter values matching the GUINEAPIG sample.

The \( \chi^2 \) between the two histograms is calculated from the number of entries in bin \( j \) of the GUINEAPIG sample \( N^j_{GP} \) and its uncertainty \( \sigma^j_{GP} \), the sum of the weights in bin \( j \) of the Model sample \( N^j_{Model} \), and the uncertainty \( \sigma^j_{Model} \), which are calculated from the event samples
Thus Equation 24 becomes

\[ N_{GP}^j = \sum_{\text{GP Events } i \text{ in Bin } j} 1, \quad \sigma_{GP}^j = \sqrt{N_{GP}^j}, \]

\[ N_{Model}^j = \sum_{\text{Model Events } i \text{ in Bin } j} w^j, \quad \sigma_{Model}^j = \sqrt{\sum_{\text{Model Events } i \text{ in Bin } j} (w^j)^2}. \]

The \( \chi^2 \) to be minimized is then calculated with

\[ \chi^2 = \sum_{\text{Bins } j} \frac{(N_{GP}^j - f_S \cdot N_{Model}^j)^2}{(\sigma_{GP}^j)^2 + (f_S \cdot \sigma_{Model}^j)^2}, \]

where

\[ f_S = \frac{\sum_{\text{GP Events } i} 1}{\sum_{\text{Model Events } i} w^j} \]

is a scaling factor that takes into account the difference in the sample sizes and the normalisation of the event weights due to the limited number of Model-events. The entire procedure has the advantage that only one sample of Model-events is needed, contrary to traditional template-fit procedures that require generating new Monte Carlo samples for every parameter-set.

However, by itself this would require even more evaluations of Equation 21 – one for every Monte Carlo event used in the generated sample – but the random variates according to this function can also be described by the sum of the random variates of the individual functions (see Appendix A and [28, 29]). The particle energy can be built up from the individual contributions

\[ x_{\text{Particle}} = x_{\text{Strahlung}} + x_{\text{Spread}} + x_G, \]

neglecting the separate regions of the Model. Each \( x_{\text{Contribution}} \) can be generated according to its probability density function. During the generation of events, the combination of functions is chosen according to the probability given by the parameters for each part \( P_{\text{Peak/Arm1/Arm2/Body}} \).

The probability for a particle’s energy in an event is given by the product of all individual probabilities

\[ P(x_{\text{Strahlung}}, x_{\text{Spread}}, x_G) = b(x_{\text{Strahlung}}) \cdot b(x_{\text{Spread}}) \cdot g(x_G), \]

and the product of the probabilities for the individual particles multiplied by the probability for the region gives the probability for the event

\[ \mathcal{L}(x_{\text{Strahlung}}^1, x_{\text{Spread}}^1, x_G^1, x_{\text{Strahlung}}^2, x_{\text{Spread}}^2, x_G^2) \equiv P_{\text{Region}} \cdot P(x_{\text{Strahlung}}^1, x_{\text{Spread}}^1, x_G^1) \cdot P(x_{\text{Strahlung}}^2, x_{\text{Spread}}^2, x_G^2). \]

Thus Equation 24 becomes

\[ w^j = \frac{p_{\text{region}}^N b(x_{\text{Strahlung}}^1, [p]^N) b(x_{\text{Spread}}^1, [p]^N) g(x_G^1, [p]^N) b(x_{\text{Strahlung}}^2, [p]^N) b(x_{\text{Spread}}^2, [p]^N) g(x_G^2, [p]^N)}{p_{\text{region}}^0 b(x_{\text{Strahlung}}^1, [p]^0) b(x_{\text{Spread}}^1, [p]^0) g(x_G^1, [p]^0) b(x_{\text{Strahlung}}^2, [p]^0) b(x_{\text{Spread}}^2, [p]^0) g(x_G^2, [p]^0)}. \]
and no numerical convolutions have to be calculated.

The probability for obtaining $x_1$ and $x_2$ is not the same as the probability to obtain a specific group of variates, even if $x_1^{\text{Strahlung}} + x_1^{\text{Spread}} + x_1^{\text{G}} = x_1$ and $x_2^{\text{Strahlung}} + x_2^{\text{Spread}} + x_2^{\text{G}} = x_2$. There are many combinations of $x_1^{\text{Spread}}$, $x_1^{\text{Strahlung}}$, and $x_1^{\text{G}}$, which can lead to the same $x_1$ or $x_2$. To estimate the probability for any given pair of energies $L(x_1, x_2)$ the convolutions have to be performed either numerically or via Monte Carlo generation.

### 4.1. Application of the Reweighting Fit to other distributions

The reweighting fit is also used to fit the distributions after the inclusion of the Bhabha scattering, Initial and Final State Radiation, and detector resolutions. The individual events are passed through the Bhabha Monte Carlo generator and detector simulation, which can be understood as additional convolutions of the existing distribution. As can be seen in Equation 31, if the parameter governing one of the contributions does not change, the contribution does not affect the new weight. This enables the use of the reweighting fit also for the reconstruction of the spectrum from the Bhabha events, because the Bhabha scattering and detector resolutions $D(E_{1,2})$ are not varied during the fit.

A measured distribution $f$ can be approximately written as

$$f(E_1, E_2) \approx \sigma(E_1, E_2) \times L(E_1, E_2) \otimes \text{ISR}(E_1, E_2) \otimes \text{FSR}(E_1, E_2) \otimes D(E_1)D(E_2),$$

(32)

where $\sigma$ represents the centre-of-mass energy dependence of the Bhabha scattering, $\text{ISR}(E_1, E_2)$ the Initial State Radiation, and $\text{FSR}(E_1, E_2)$ the Final State Radiation distribution. If the cross-section and detector resolutions are well enough known, the only difference between the measured and generated distributions is the luminosity spectrum. For this study the same Bhabha generator and detector simulations are used for both samples, so the additional effects are statistically the same. Any difference for the contributions can lead to a systematic error in the reconstruction of the luminosity spectrum.

### 4.2. Equiprobability Binning

To obtain an unbiased estimator of the compatibility in a $\chi^2$-fit, all bins should contain at least seven entries, and the number of events in all bins should be similar [30, p.304]. These requirements can be fulfilled if an equiprobability binning is generated based on the respective GUINEAPIG sample used in the fits. With equal-size bins either a large number of bins could be used – where most would contain very few or no entries and would have to be rejected for the $\chi^2$ calculation – and the peak substructure could be resolved, or fewer bins with larger dimensions could be used, but then the peak could not be resolved. Therefore, the equiprobability binning can make better use of the available events.

Following the recipe of James [30, p.305], the events are first evenly separated along one axis, and then all events falling in the range on this first axis are again evenly separated in the second axis. If additional dimensions are used, the separation is repeated. In this way each bin has different dimensions along each axis, but the number of events per bin is constant. Two classes were developed to generate and store the equiprobability structure, and use it for the fits.
Figure 13: Example of the equiprobability binning in (a) two and (b) three dimensions. The colours are arbitrary. By construction every cell contains a similar number of events.

For the generator level fit, as discussed in Section 5.1, the distribution of the two particle energies is stored in a two-dimensional histogram. For the reconstruction of the spectrum from the Bhabha events (Sections 5.3.1 to 5.3.3) the energy of the scattered electron and positron, and the relative centre-of-mass energy reconstructed from the acollinearity is filled into a three-dimensional histogram. Figure 13 shows examples for a two- and three-dimensional bin structure. It can be seen that around the nominal beam energies the size of the bins becomes smaller. Because the separation of events is done individually along each axis, the bin structures are not symmetric.

4.3. Statistical Validation of the Model

In order to ensure that the chosen Model is unbiased and consistent, a large number of Model vs. Model fits were performed, with a varying number of bins. In each case, the procedure is as follows: two sets of events are created according to the Model of the basic luminosity spectrum. The samples are then used in the fit procedure described before. In each fit a cut on the centre-of-mass energy of \( \sqrt{s} > 1.5 \) TeV is applied.

The pull distribution of every free parameter is obtained and fitted with a Gaussian function. The Model is unbiased, if the mean of every pull distribution is close to zero. The uncertainty is correctly estimated, if the pull width is compatible with unity. This is called the Normality condition [30, p.310].

For all parameters the pull distributions are independent of the binning. Most parameters are unbiased (i.e., the mean is zero). Figure 14a shows two examples for normally distributed pulls. Exceptions are the \( \eta^{1}_{\text{Body1}} \) and \( \eta^{1}_{\text{Body2}} \) parameters, whose pulls are not normally distributed, as
shown in Figure 14b. These parameters describe the behaviour of the beta-distribution at the lower edge of the respective particle energy distribution. Therefore, the bias is caused by the cut on the centre-of-mass energy, which reduces the sensitivity to the lower energy part of the Body. The lower limit of these parameters is zero, which is often found by the minimizer instead of the nominal value. When the cut on the centre-of-mass energy is removed, Figure 14c is obtained. The pulls are then symmetric, and the parameters are correctly estimated.

The Model is unbiased and consistent for most parameters, and the results do not depend on the chosen binning. The parameters for which the results are biased can be recovered once the cut of $\sqrt{s} > 1.5$ TeV is dropped when the Model includes a better description of the Body.

5. Luminosity Spectrum Reconstruction

All ingredients for the reconstruction of the luminosity spectrum – the Model and the reweighting procedure – are now available. For completeness and the separation of their impact, the effects leading to the final observables are included one after the other. For each stage a reweighting fit is done to extract the basic luminosity spectrum.

The fits based on the basic luminosity spectrum (Section 5.1) are used to assess the similarity between the Model and the GUINEAPig spectrum. The energies of the electron and positron pairs are filled into the two-dimensional equiprobability structure used in the reweighting fit. The energies of the initial electron and positron are also used for the fit to the scaled luminosity spectrum (Section 5.2).

The results of the reconstruction of the luminosity spectrum from the observables (as defined in Section 2.4) given by the generated Bhabha events are shown in Sections 5.3.1 to 5.3.3. The relative centre-of-mass energy and the two measured energies – from the electron and positron – are filled into the three-dimensional equiprobability structure.

The initial values of the parameters, used to generate the Model events, are given in Table 3. All the regions are chosen to start with a similar number of events (0.25). The starting $\omega$ parameters are taken from the fit to the beam-energy spread before the collisions (Table 1). The parameters $\omega_{\text{Arm}1}$ and $\omega_{\text{Arm}2}$ are set to 0.35 to cause the distribution to fall, as is expected. Similarly, the other parameters are chosen arbitrarily in a way to cause a behaviour similar to the GUINEAPig luminosity spectrum. The position of the two boundaries for the beam-energy spread ($x_{\text{min}}$ and $x_{\text{max}}$) are also taken from Table 1. Table 3 also lists the lower and upper bounds limiting the values for the minimizer.

In Sections 5.1, 5.2 and 5.3.1 to 5.3.3 only the special features of the different fits are described, the discussion of the results is presented in Section 5.4.

5.1. Fit to the Basic Luminosity Spectrum

To verify that the Model can represent the basic luminosity spectrum from GUINEAPig, the distribution of the initial particle energies are used in the $\chi^2$-fit. The data histogram is shown in Figure 2. The Monte Carlo sample is shown in Figure 11. The GUINEAPig sample consists of 3 million events and the Model provides 10 million events. Fits are done with a binning varied from $50 \times 50$ bins to $300 \times 300$ bins in steps of 10 bins. Only events with $\sqrt{s} > 1.5$ TeV are
Figure 14: Pull distribution for selected parameters from the fits with 100 × 100 bins. The distributions are fitted to Gaussian functions. (a) $p_{\text{Arm2}}$ (left) and $\omega_{\text{Peak1}}$ (right). The $\eta_{\text{Body1}}$ and $\eta_{\text{Body2}}$ parameters (b) with the cut $\sqrt{s} > 1.5\text{ TeV}$ and (c) with a looser cut.
Table 3: Initial parameter values used for the generation of the events. Also listed are the lower and upper bounds used in the reweighting fits.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Lower Bound</th>
<th>Nominal Value</th>
<th>Upper Bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>$p_{\text{Peak}}$</td>
<td>0.0</td>
<td>0.25</td>
<td>0.4</td>
</tr>
<tr>
<td>$p_{\text{Arm1}}$</td>
<td>0.0</td>
<td>0.25</td>
<td>0.3</td>
</tr>
<tr>
<td>$p_{\text{Arm2}}$</td>
<td>0.0</td>
<td>0.25</td>
<td>0.3</td>
</tr>
<tr>
<td>$\omega_{\text{Peak1}}$</td>
<td>$-1.0$</td>
<td>$-0.522336$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Peak1}}^2$</td>
<td>$-1.0$</td>
<td>$-0.409289$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Peak2}}$</td>
<td>$-1.0$</td>
<td>$-0.522336$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Peak2}}^2$</td>
<td>$-1.0$</td>
<td>$-0.409289$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Arm1}}$</td>
<td>$-1.0$</td>
<td>$-0.522336$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Arm1}}^2$</td>
<td>$-1.0$</td>
<td>0.35</td>
<td>10.0</td>
</tr>
<tr>
<td>$\omega_{\text{Arm2}}^1$</td>
<td>$-1.0$</td>
<td>$-0.522336$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\omega_{\text{Arm2}}^2$</td>
<td>$-1.0$</td>
<td>0.35</td>
<td>10.0</td>
</tr>
<tr>
<td>$\eta_{\text{Arm1}}$</td>
<td>0.0</td>
<td>2.5</td>
<td>10.0</td>
</tr>
<tr>
<td>$\eta_{\text{Arm1}}^2$</td>
<td>$-1.0$</td>
<td>$-0.75$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\eta_{\text{Arm2}}$</td>
<td>0.0</td>
<td>2.5</td>
<td>10.0</td>
</tr>
<tr>
<td>$\eta_{\text{Arm2}}^2$</td>
<td>$-1.0$</td>
<td>$-0.75$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\eta_{\text{Body1}}$</td>
<td>0.0</td>
<td>0.15</td>
<td>10.0</td>
</tr>
<tr>
<td>$\eta_{\text{Body1}}^2$</td>
<td>$-1.0$</td>
<td>$-0.55$</td>
<td>0.0</td>
</tr>
<tr>
<td>$\eta_{\text{Body2}}$</td>
<td>0.0</td>
<td>0.15</td>
<td>10.0</td>
</tr>
<tr>
<td>$\eta_{\text{Body2}}^2$</td>
<td>$-1.0$</td>
<td>$-0.55$</td>
<td>0.0</td>
</tr>
</tbody>
</table>

used in the fit. The cut is applied because the Model has limited validity range, and the events below half the nominal centre-of-mass energy would have a negative impact on the fit result.

As an example for the result of the reweighting fit, Figure 15 shows a small section of the histogram mapped onto one dimension and the pull distribution for all the bins before and after the fit. The data histogram has a constant number of events per bin, as designed by the equiprobability binning. The pull distribution after the fit converged is well centred around 0 with a width of 2.3. The width of the pull distribution is not equal to 1, because the $\chi^2$/ndf is larger than 1. This means that the Model is not completely identical to the GUINEAPig distribution. Two of the differences are the limited number of beta-distributions used to model the tail of the spectrum (see Section 3.2), where deviations appear, and the differences in the peak of the spectrum (see Section 3.1), where a much larger number of parameters would be needed. As the $\chi^2$/ndf is not equal to unity, additional parameters would enable a better description of the spectrum.

At this point, the Model’s consistency can be validated using independent Monte Carlo samples fitted against the same GUINEAPig sample. If the fitting procedure and the Model are consistent, the fit results are always compatible with each other. The fit was done 198 times with the same GUINEAPig sample and statistically independent Model samples. All the parameters vary within their uncertainties. Therefore, the Model and the fit are consistent.
Figure 15: (a) Blow-up of a small section of the bins used in the re-weighting fit of the initial electron energies. The histogram for GuineaPig (black) has, by construction, a constant number of events per bin. Also shown are the histograms for the Model with the initial parameter values before the fit and after the fit. (b) Distribution of the pulls \( \Omega = (N_{\text{GP}} - f_S \cdot N_{\text{Model}}) / ((\sigma_{\text{GP}})^2 + (f_S \cdot \sigma_{\text{Model}})^2)^{1/2} \) (symbols are defined in Equations 25 and 27) between the GuineaPig and Model samples before and after the re-weighting fit. A Gaussian function is fitted to the distribution of pulls after the fit.

5.2. Fit to the Scaled Luminosity Spectrum

For this fit the samples including the cross-section dependence are used. In the generation of the events according to the Model the accept–reject method is used to scale the model to the Bhabha cross-section. The method is also used for the GuineaPig events (see Section 2.2). In the generation of events a centre-of-mass energy cut of several hundred GeV is applied for efficiency reasons. The samples sizes are again 3 million from GuineaPig and 10 million for the Model. The same cut on the centre-of-mass energy of \( \sqrt{s} > 1.5 \) TeV is used. Fits are done with a binning from 50 × 50 bins to 300 × 300 bins in steps of 10 bins.

5.3. Fits to the Observables

The observables are defined in Section 2.4. Binnings\(^5\) from 10 × 10 × 10 bins to 80 × 50 × 50 bins were used in the fits. The binning step is 5 bins for the relative centre-of-mass energy and 10 bins for the particle energies.

5.3.1. Fit to Observables with the Basic Luminosity Spectrum

Events with a lower centre-of-mass energy are rejected by a cut on the final state particles. A cut on the centre-of-mass energy calculated from the momentum four-vectors of the electrons

\(^5\)For the number of bins given, the first number represents the number of bins for the relative centre-of-mass energy, and the second and third number represents the number of bins for the two particle energies.
of $\sqrt{s_{\text{4-vec}}} > 1.5$ TeV is applied. For the fit the GUINEAPig sample consists of 3 million events, and for the Model 10 million events are used.

### 5.3.2. Fit to Observables with the Scaled Luminosity Spectrum

The final state electrons are generated with BHWide, but this time the scaled luminosity spectra are used as the input to BHWide. The same selection cut on the centre-of-mass energy calculated from the momentum four-vectors of $\sqrt{s_{\text{4-vec}}} > 1.5$ TeV is applied on the sample. Due to the larger rate of lower energy events which are cut, the GUINEAPig sample has become smaller, only about 2 million events are available. For the Model 10 million events are used.

### 5.3.3. Fit to Observables with the Scaled Luminosity Spectrum and Detector Resolutions

For the last step, the Bhabha events generated with the scaled luminosity spectrum are smeared with the detector resolutions as described in Section 2.4. The selection cuts had to be modified, and the cut is applied on the centre-of-mass energy calculated from the (smeared) four-vectors $\sqrt{s_{\text{4-vec}}} > 1.5$ TeV and in addition on the individual particle energies $E_1 > 150$ GeV and $E_2 > 150$ GeV. To recover Final State Radiation, the energy of all photons in a $3^\circ$ cone around an electron is summed up.

### 5.4. Discussion of the Results

For every stage of the reconstruction multiple fits with different binnings were done. However, as the reconstructed spectra are fairly similar, only one reconstructed luminosity spectrum per stage is shown in detail. In addition, the parameter dependence on the number of bins is shown. For the fits to the basic and scaled luminosity spectrum the results with $100 \times 100$ bins are taken. For the reconstruction from the observables the fits with $40 \times 50 \times 50$ bins are shown.

In Table 4 the $\chi^2$/ndf and parameters extracted by in selected fit stages are listed. The reconstructed parameters are far away from the initial values of the parameters (cf. Table 3). The fit results are not artificially improved by using a good starting point. The final values of the beam-energy spread parameters $\omega$ are close to the values found by the one-dimensional fit to the beam-energy spread distributions detailed in Section 3.1. Due to the cut on the minimum centre-of-mass energy, the sensitivity on the lower Beamstrahlung parameter $\eta_1$ is lost, and all fit stages give a result of $\approx 0$ for these parameters. The reconstruction of the upper Beamstrahlung parameter $\eta_2$ is quite consistent. While there are significant differences from stage to stage, the spread is small.

The largest variation in the parameters is observed for the beam-spread parameters $\omega$. Especially, once the detector effects are added to the samples, the uncertainties increase by about a factor four and the values are significantly changed. The larger uncertainties can also be due to the smaller GUINEAPig sample (cf. Section 5.3.2).

There are significant correlations between the parameters. Figure 16a gives the correlation matrix for the fit to the basic luminosity spectrum and Figure 16b for the fit to the observables with the scaled luminosity spectrum and detector resolutions. The order of parameters is
Table 4: The parameter values found in selected fits. The order of the results is the same as the Sections 5.1, 5.2 and 5.3.1 to 5.3.3. The details of the fits are given in the text.

<table>
<thead>
<tr>
<th>$\chi^2$/ndf</th>
<th>$P_{\text{Peak}}$</th>
<th>$P_{\text{Arm1}}$</th>
<th>$P_{\text{Arm2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>63832 / 10000</td>
<td>0.2387 ± 0.0004</td>
<td>0.2672 ± 0.0004</td>
<td>0.2659 ± 0.0004</td>
</tr>
<tr>
<td>62697 / 10000</td>
<td>0.2402 ± 0.0004</td>
<td>0.2666 ± 0.0004</td>
<td>0.2642 ± 0.0004</td>
</tr>
<tr>
<td>114039 / 100000</td>
<td>0.2439 ± 0.0007</td>
<td>0.2666 ± 0.0006</td>
<td>0.2652 ± 0.0006</td>
</tr>
<tr>
<td>109972 / 100000</td>
<td>0.2479 ± 0.0009</td>
<td>0.2652 ± 0.0007</td>
<td>0.2627 ± 0.0007</td>
</tr>
<tr>
<td>100593 / 100000</td>
<td>0.2483 ± 0.0010</td>
<td>0.2681 ± 0.0009</td>
<td>0.2632 ± 0.0009</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\omega^1_{\text{Peak1}}$</th>
<th>$\omega^2_{\text{Peak1}}$</th>
<th>$\omega^1_{\text{Peak2}}$</th>
<th>$\omega^2_{\text{Peak2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.2788 \pm 0.0016$</td>
<td>$-0.3425 \pm 0.0013$</td>
<td>$-0.2805 \pm 0.0016$</td>
<td>$-0.3417 \pm 0.0013$</td>
</tr>
<tr>
<td>$-0.2772 \pm 0.0019$</td>
<td>$-0.3370 \pm 0.0015$</td>
<td>$-0.2769 \pm 0.0019$</td>
<td>$-0.3403 \pm 0.0015$</td>
</tr>
<tr>
<td>$-0.2668 \pm 0.0028$</td>
<td>$-0.3257 \pm 0.0022$</td>
<td>$-0.2670 \pm 0.0028$</td>
<td>$-0.3232 \pm 0.0022$</td>
</tr>
<tr>
<td>$-0.2583 \pm 0.0037$</td>
<td>$-0.3107 \pm 0.0030$</td>
<td>$-0.2659 \pm 0.0037$</td>
<td>$-0.3225 \pm 0.0029$</td>
</tr>
<tr>
<td>$-0.3879 \pm 0.0149$</td>
<td>$-0.3882 \pm 0.0135$</td>
<td>$-0.3058 \pm 0.0175$</td>
<td>$-0.3283 \pm 0.0153$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\omega^1_{\text{Arm1}}$</th>
<th>$\omega^2_{\text{Arm1}}$</th>
<th>$\omega^1_{\text{Arm2}}$</th>
<th>$\omega^2_{\text{Arm2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$-0.4399 \pm 0.0012$</td>
<td>$0.3243 \pm 0.0037$</td>
<td>$-0.4399 \pm 0.0012$</td>
<td>$0.3364 \pm 0.0036$</td>
</tr>
<tr>
<td>$-0.4509 \pm 0.0012$</td>
<td>$0.3581 \pm 0.0039$</td>
<td>$-0.4473 \pm 0.0012$</td>
<td>$0.3648 \pm 0.0039$</td>
</tr>
<tr>
<td>$-0.4057 \pm 0.0034$</td>
<td>$0.4127 \pm 0.0090$</td>
<td>$-0.4078 \pm 0.0033$</td>
<td>$0.4180 \pm 0.0090$</td>
</tr>
<tr>
<td>$-0.4192 \pm 0.0040$</td>
<td>$0.4762 \pm 0.0112$</td>
<td>$-0.4162 \pm 0.0039$</td>
<td>$0.4688 \pm 0.0108$</td>
</tr>
<tr>
<td>$-0.4994 \pm 0.0107$</td>
<td>$0.3054 \pm 0.0305$</td>
<td>$-0.5501 \pm 0.0098$</td>
<td>$0.1842 \pm 0.0292$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\eta^1_{\text{Arm1}}$</th>
<th>$\eta^2_{\text{Arm1}}$</th>
<th>$\eta^1_{\text{Arm2}}$</th>
<th>$\eta^2_{\text{Arm2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000 ± 0.0008</td>
<td>$-0.6253 \pm 0.0011$</td>
<td>0.0000 ± 0.0007</td>
<td>$-0.6268 \pm 0.0011$</td>
</tr>
<tr>
<td>0.0000 ± 0.0004</td>
<td>$-0.6243 \pm 0.0011$</td>
<td>0.0000 ± 0.0005</td>
<td>$-0.6306 \pm 0.0011$</td>
</tr>
<tr>
<td>0.0000 ± 0.0005</td>
<td>$-0.6021 \pm 0.0020$</td>
<td>0.0000 ± 0.0007</td>
<td>$-0.6120 \pm 0.0020$</td>
</tr>
<tr>
<td>0.0000 ± 0.0003</td>
<td>$-0.5987 \pm 0.0023$</td>
<td>0.0000 ± 0.0004</td>
<td>$-0.6041 \pm 0.0023$</td>
</tr>
<tr>
<td>0.0000 ± 0.0003</td>
<td>$-0.6054 \pm 0.0027$</td>
<td>0.0000 ± 0.0004</td>
<td>$-0.6080 \pm 0.0028$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$\eta^1_{\text{Body1}}$</th>
<th>$\eta^2_{\text{Body1}}$</th>
<th>$\eta^1_{\text{Body2}}$</th>
<th>$\eta^2_{\text{Body2}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0000 ± 0.0002</td>
<td>$-0.6640 \pm 0.0012$</td>
<td>0.0000 ± 0.0002</td>
<td>$-0.6636 \pm 0.0012$</td>
</tr>
<tr>
<td>0.0000 ± 0.0001</td>
<td>$-0.6643 \pm 0.0010$</td>
<td>0.0000 ± 0.0001</td>
<td>$-0.6675 \pm 0.0010$</td>
</tr>
<tr>
<td>0.0000 ± 0.0005</td>
<td>$-0.6538 \pm 0.0027$</td>
<td>0.0000 ± 0.0006</td>
<td>$-0.6540 \pm 0.0027$</td>
</tr>
<tr>
<td>0.0000 ± 0.0003</td>
<td>$-0.6550 \pm 0.0025$</td>
<td>0.0000 ± 0.0004</td>
<td>$-0.6571 \pm 0.0025$</td>
</tr>
<tr>
<td>0.0000 ± 0.0004</td>
<td>$-0.6421 \pm 0.0029$</td>
<td>0.0000 ± 0.0005</td>
<td>$-0.6415 \pm 0.0029$</td>
</tr>
</tbody>
</table>

The largest correlations are between parameters from the same beta-distribution. Comparing Figure 16a with Figure 16b shows that the correlations increase when the additional effects are taken into account. Some of the changes of the parameters could be due to increased correlations.

Table 5 lists the fraction of events with a centre-of-mass energy larger than $0.99\sqrt{s_{\text{nom}}}$ from
Table 5: Summary of the fraction of events with $\sqrt{s} > 0.99\sqrt{s_{\text{nom}}}$ from GUINEAPig and the reconstructed luminosity spectra from the different fit stages.

<table>
<thead>
<tr>
<th></th>
<th>Fraction [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>GUINEAPig sample</td>
<td>35.41±0.06</td>
</tr>
<tr>
<td>Fit to Basic Luminosity Spectrum</td>
<td>34.61±0.01</td>
</tr>
<tr>
<td>Fit to Scaled Luminosity Spectrum</td>
<td>34.74±0.01</td>
</tr>
<tr>
<td>Fit to Observables with Basic Luminosity Spectrum</td>
<td>34.42±0.01</td>
</tr>
<tr>
<td>Fit to Observables with Scaled Luminosity Spectrum</td>
<td>34.53±0.02</td>
</tr>
<tr>
<td>Fit to Observables with Scaled Luminosity Spectrum and Det. Res.</td>
<td>34.72±0.07</td>
</tr>
</tbody>
</table>

GUINEAPig and from selected fits of the different fit stages. The uncertainty of the GUINEAPig value is the statistical uncertainty from one million events. The uncertainty for the fits is calculated from the uncertainty of the individual parameters and accounts for the correlation between them. The inclusion of each individual effect increases the size of the uncertainty.

The difference of the fractions between GUINEAPig and the Model is less than one percentage point. Given the size of the uncertainties the difference is significant. However, processes with lower cross-section will effectively use smaller samples from the luminosity increasing the uncertainty to around one percentage point. The difference in the fraction of events in the top 1% might therefore be insignificant for other measurements at 3 TeV.

The basic luminosity spectrum from GUINEAPig compared with the basic luminosity spectrum reconstructed in the different fit stages for the selected fits are shown in Figures 17, 19, 21, 23 and 25. For the ratios the green error bars show the statistical uncertainty for one million GUINEAPig events and the barely visible red error bars show the uncertainty coming from the parametrisation. For all stages the luminosity spectrum is reconstructed within 5% at least for $0.55 < \sqrt{s}/\sqrt{s_{\text{nom}}} < 0.995$. Above $0.995\sqrt{s_{\text{nom}}}$, the beam-energy spread is the dominant effect and the difficulty of modelling this peak becomes visible. Still, this difference is seen only, when looking at small bin sizes (e.g., compare the bins around 1 in Figure 17d or Figure 17e with Figure 17f). As Table 5 shows, the average fraction around the peak is reconstructed within 1 percentage point. Improved parametrisations should be able to better describe and reconstruct the shape of the peak.

Below $0.5\sqrt{s_{\text{nom}}}$, the Model is much more inconsistent with GUINEAPig, but this is given by the design of this Model and the cut on the centre-of-mass energy applied for the fits.

Some of the reconstructed parameters depend on the number of bins used in the fit. Figure 18 shows the dependence of the reconstructed parameters on the number of bins used in the fit. Fits with a binning of $50 \times 50$ bins to $300 \times 300$ bins with the same number of events were done. The parameters $\eta^2_{\text{Arm}1}$ and $\eta^2_{\text{Arm}2}$, which represent the upper edge of the beam-energy spread of the Arms, show a significant dependence on the binning. For the other parameters the change is below one sigma. It is also visible that with more bins the parameter $p_{\text{Peak}}$ rises, while the two parameters $p_{\text{Arm}1}$ and $p_{\text{Arm}2}$ fall, which can also be seen in the correlation matrix and their correlation coefficient of about $-0.4$ (see Figure 16a).

Comparing Figure 18 and Figure 20, the inclusion of the cross-section scaling in the distri-
The parameter value dependence on the binning after the inclusion of the cross-section scaling in the fit to the observables is shown in Figure 24. As with the inclusion of the cross-section before the Bhabha scattering, there is little impact on the beam-energy spread parameters and the $\eta^2$-parameters for the Arms show the strongest change.

After including the detector resolutions, the parameter values depend much stronger on the number of bins (see Figure 26). Without a minimum number of bins the peak structure cannot be resolved, and the $\omega$-parameters are completely different from the previous results, and show large fluctuations in their values. If a large enough number of bins is used, the results are only a few sigma different from the previous fit results. The detector resolutions have a strong impact on resolving the structure of the luminosity peak.
Figure 16: Correlation matrices (a) for the fit to the basic luminosity spectrum and (b) for the fit to the Bhabha observables with the scaled luminosity spectrum and detector resolutions. Only the values above the diagonal and with an absolute value larger than 0.01 are shown.
Figure 17: Resulting spectra for the fit to the basic luminosity spectrum and a binning of 100 × 100. (a)–(c) Basic luminosity spectrum from GUINEA PIG compared to the Model after the fit. The integral over the full range is normalized. (d)–(f) Relative difference between GUINEA PIG and the Model after the fit. The error bars are the statistical uncertainty from GUINEA PIG (green), and the uncertainty from the parameters (red).
Figure 18: Parameter variation with respect to the Binning ID for the fits to the basic luminosity spectrum. The entries are sorted by falling $\chi^2$.
Figure 19: Resulting spectra for the fit to the scaled luminosity spectrum and a binning of $100 \times 100$. (a)–(c) Basic luminosity spectrum from GUINEAPIG compared to the Model after the fit. The integral over the full range is normalized. (d)–(f) Relative difference between GUINEAPIG and the Model after the fit. The error bars are the statistical uncertainty from GUINEAPIG (green), and the uncertainty from the parameters (red).
Figure 20: Parameter variation with respect to the Binning ID for the fits to the scaled luminosity spectrum. The entries are sorted by falling $\chi^2$.
Figure 21: Resulting spectra for the fit to the observables with the basic luminosity spectrum, without detector resolutions, and a binning of $40 \times 50 \times 50$. (a)–(c) Basic luminosity spectrum from GUINEAPig compared to the Model after the fit. The integral over the full range is normalized. (d)–(f) Relative difference between GUINEAPig and the Model after the fit. The error bars are the statistical uncertainty from GUINEAPig (green), and the uncertainty from the parameters (red).
<table>
<thead>
<tr>
<th>Binning ID</th>
<th>Peak 1</th>
<th>Arm 1</th>
<th>Peak 2</th>
<th>Arm 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$0.235$</td>
<td>$0.26$</td>
<td>$0.25$</td>
<td>$0.26$</td>
</tr>
<tr>
<td></td>
<td>$0.24$</td>
<td>$0.265$</td>
<td>$0.255$</td>
<td>$0.265$</td>
</tr>
<tr>
<td></td>
<td>$0.245$</td>
<td>$0.27$</td>
<td>$0.275$</td>
<td>$0.27$</td>
</tr>
</tbody>
</table>

Figure 22: Parameter variation with respect to the Binning ID for the fits to the observables with the basic luminosity spectrum. The entries are sorted by falling $\chi^2$ and the colours give the number of bins used for the energy observables: Black (20), Blue (30), Green (40), Red (50).
Figure 23: Resulting spectra for the fit to the observables with the scaled luminosity spectrum, without detector resolutions, and a binning of $40 \times 50 \times 50$. (a)–(c) Basic luminosity spectrum from GUINEAPig compared to the Model after the fit. The integral over the full range is normalized. (d)–(f) Relative difference between GUINEAPig and the Model after the fit. The error bars are the statistical uncertainty from GUINEAPig (green), and the uncertainty from the parameters (red).
Figure 24: Parameter variation with respect to the Binning ID for the fits to the observables with the scaled luminosity spectrum. The entries are sorted by falling $\chi^2$ and the colours give the number of bins used for the energy observables: Black (20), Blue (30), Green (40), Red (50).
Figure 25: Resulting spectra for the fit to the observables with the scaled luminosity spectrum, detector resolutions, and a binning of $40 \times 50 \times 50$. (a)–(c) Basic luminosity spectrum from GUINEAPIG compared to the Model after the fit. The integral over the full range is normalized. (d)–(f) Relative difference between GUINEAPIG and the Model after the fit. The error bars are the statistical uncertainty from GUINEAPIG (green), and the uncertainty from the parameters (red).
Figure 26: Parameter variation with respect to the Binning ID for the fits to the observables with the scaled luminosity spectrum and including detector resolutions. The entries are sorted by falling $\chi^2$ and the colours give the number of bins used for the energy observables: Black (20), Blue (30), Green(40), Red(50).

There are significant differences between the reconstructed luminosity spectrum and the one from GUINEAPIG when looking at large event samples. Typical cross-sections for New Physics phenomena will be much smaller than that of Bhabha scattering, and the luminosity spectrum sampled for a specific process will therefore have larger statistical fluctuations, so that the difference between the reconstructed and actual spectrum might not be significant. To estimate the impact of the difference between GUINEAPIG and the reconstructed spectrum, the measurement of the smuon mass $m_{\tilde{\mu}} \pm$ and neutralino mass $m_{\tilde{\chi}_0}$ from smuon pair production is used.

The smuon decays into a muon and a neutralino, so that the energy spectrum of the muons $f(E_\mu)$ can be used to extract the smuon and neutralino masses. The details of the analysis are described elsewhere [31], here only the parts directly concerning the systematic uncertainty from the luminosity spectrum are repeated. There are some differences in the treatment of the statistical uncertainty between the version of the fitting program used here, and the one used in the original paper [31].

In an ideal situation – with a single centre-of-mass energy $\sqrt{s_{\text{nom}}}$ – the muon energy spectrum is a uniform distribution $U(E_\mu)$ with the boundaries

$$E_{H,L} = \frac{\sqrt{s_{\text{nom}}}}{4} \left( 1 - \frac{m_{\tilde{\mu}}^2}{m_{\tilde{\chi}_0}^2} \right) \left( 1 \pm \sqrt{1 - \frac{4 m_{\tilde{\mu}}^2}{s_{\text{nom}}}} \right).$$ (33)

The uniform distribution therefore depends on the smuon and neutralino masses.

In reality, there is not a single centre-of-mass energy. For every centre-of-mass energy the uniform distribution has different limits. Therefore, the measured muon-energy spectrum is affected by the basic luminosity spectrum, the Initial State Radiation, the cross-section, and the detector resolution $D(E_\mu)$. The luminosity spectrum $\mathcal{L}'(\sqrt{s})$, Initial State Radiation $\text{ISR}(\sqrt{s})$, and cross-section $\sigma_x(\sqrt{s})$ can be convoluted into an effective luminosity spectrum

$$\mathcal{L}_{\text{Eff}}(\sqrt{s}) = \mathcal{L}'(\sqrt{s}) \otimes \text{ISR}(\sqrt{s}) \otimes \sigma_x(\sqrt{s}).$$ (34)

The function to fit the muon energy spectrum is then the convolution of the uniform energy spectrum with the detector resolution weighted by the effective luminosity spectrum

$$f(E_\mu) = \int_{\sqrt{s_{\text{min}}}}^{\sqrt{s_{\text{max}}}} \mathcal{L}_{\text{Eff}}(\sqrt{s}) \cdot \int_{E_{L_{\text{eff}}}(\sqrt{s})}^{E_{H_{\text{eff}}}(\sqrt{s})} U(\sigma_x, m_{\tilde{\mu}} \pm, m_{\tilde{\chi}_0}, \sqrt{s}, \tau) \cdot D(E_\mu - \tau) \, d\tau \, d\sqrt{s}. \quad (35)$$

The luminosity spectrum can either be taken directly from GUINEAPIG or from the reconstruction. Figure 27a shows the background-subtracted signal sample and an example fit with Equation 35. In the original study [31], the effective luminosity spectrum was taken directly from the signal events, where Initial State Radiation, cross-section and Beamstrahlung are convoluted by WHIZARD. Here, the GUINEAPIG spectrum is treated the same way as the reconstructed spectrum to avoid any differences from different implementations of the Initial State Radiation or cross-section.

When the effective luminosity spectrum is created with the luminosity spectrum reconstructed from the Bhabha events, the masses extracted from the fit with Equation 35 become a function...
Figure 27: (a) Background subtracted signal sample and the best fit to extract the smuon and neutralino mass. (b) Reconstructed masses with the luminosity spectra taken from the fits to the Bhabha observables with different binnings.

of the parameters $\vec{p}$ from the spectrum reconstruction $m = m(\vec{p})$. To estimate the systematic uncertainty due to the reconstruction of the spectrum, the fit is performed with the nominal set of parameters $\vec{p}$ and with each parameter $p_i$ increased or decreased by half of a standard deviation $\sigma_{p_i}$,

$$m_j^+ = m\left(\vec{p} + \vec{e}_i\frac{\sigma_{p_i}}{2}\right) \quad m_j^- = m\left(\vec{p} - \vec{e}_i\frac{\sigma_{p_i}}{2}\right).$$

The systematic uncertainty on the fitted value is then given by

$$\sigma_m^2 = \left(\sum_{i,j} \delta C_{ij} \delta_j\right)^{1/2}$$

with $\delta = m_j^+ - m_j^-$, and the correlation matrix $C$ (i.e., Figure 16b).

Table 6 lists the smuon and neutralino masses from the fit when the luminosity spectrum is taken from GUINEAP1G and from the fit to the observables with the scaled luminosity spectrum and detector resolutions with a binning of $50 \times 40 \times 40$ bins. The difference in the reconstructed masses for these two luminosity spectra is smaller than the statistical uncertainty. However, as the reconstructed luminosity spectrum shows a dependence on the binning, so do the reconstructed masses. Figure 27b shows the reconstructed masses for the spectra reconstructed with different binnings. There is a dependence of the reconstructed masses on the number of bins, but the spread of the reconstructed masses is smaller than the statistical uncertainty (cf. Table 6).

As the difference between the obtained masses and the spread of masses is smaller than the statistical uncertainty, the reconstruction of the luminosity spectrum does not introduce a significant bias. The systematic uncertainty due to the luminosity spectrum reconstruction is also much smaller than the statistical uncertainty, so that the total uncertainty on the reconstructed mass is not increased significantly.
Table 6: Extracted smuon and neutralino masses from the fits to the signal sample using different (effective) luminosity spectra.

<table>
<thead>
<tr>
<th>Spectrum</th>
<th>Smuon Results [GeV]</th>
<th>Neutralino Results [GeV]</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mass</td>
<td>σ_{Stat}</td>
</tr>
<tr>
<td>GUINEAPig-spectrum</td>
<td>1011.77 ± 3.05</td>
<td></td>
</tr>
<tr>
<td>Fit 50 × 40 × 40</td>
<td>1011.56 ± 3.05</td>
<td>±0.04</td>
</tr>
</tbody>
</table>

7. Summary, Conclusions and Outlook

A framework has been developed for the reconstruction of the basic luminosity spectrum at future linear colliders. The spectrum can be reconstructed from Bhabha events measured with the tracking detectors and calorimeters. To test the reconstruction method all relevant effects were included step-by-step: the full luminosity spectrum from beam-beam simulations – including the CLIC beam-energy spread – the Bhabha cross-section dependence, Initial and Final State Radiation, and the detector resolutions.

The Model, needed for the reweighting fit, has some limitations. For technical reasons the energy range to describe the Beamstrahlung tail is limited to $\sqrt{s} > 1500$ TeV, and the peculiar beam-energy spread cannot be modelled precisely with few parameters. The reweighting fit itself does not impair the reconstructed spectrum. The differences between GUINEAPig and the reconstructed spectrum do not significantly change between the fit to the basic luminosity spectrum and the fit to the observables with the scaled luminosity spectrum and including detector resolutions. With an improved model, and larger processing power, an improved reconstruction of the CLIC 3 TeV spectrum should be possible.

The fraction of events above 99% of the nominal centre-of-mass energy is reconstructed within 1 percentage point. The centre-of-mass energy distribution is reconstructed within 5% between the nominal and about half the nominal centre-of-mass energy, the validity limit of our Model. These results are obtained regardless of the included level of details, so that one can conclude that the limitations of the Model cause most of the discrepancies to the simulated spectrum, and if a better model is used, the discrepancies should be reduced.

To estimate the systematic impact on other physics measurements, the reconstructed spectrum was used in the study of smuon decays, one of the CLIC 3 TeV benchmark processes. The reconstructed spectrum does not induce a significant bias on the measured mass, nor does it cause a significant systematic uncertainty. The systematic uncertainty from the spectrum reconstruction is an order of magnitude smaller than the statistical uncertainty.

7.1. Outlook

The framework can also be applied for the reconstruction of the luminosity spectrum at other centre-of-mass energies and linear electron–positron colliders than CLIC. Depending on the beam-energy spread and the demanded range of the reconstruction, the Model has to be adapted, but this will not increase the computational complexity of the reconstruction.
Increasing the energy range of the current Model is straightforward: the single Beamstrahlung beta-distributions have to be replaced by linear combinations of beta-distribution. An improved description of the beam-energy spread is less obvious without a large increase in the number of parameters.

The boundaries of the beam-energy spread – the parameters $x_{\text{min}}$ and $x_{\text{max}}$ – were fixed during the fit. It should be evaluated how much the measurement is affected, if these parameter values differ from those of the beam-energy spread. It should also be tried to vary the boundaries of the beam-energy spread during the re-weighting fit.

The observables from the Bhabha events can also be exchanged for other suitable choices, always keeping the detector resolutions in mind. The impact of the detector resolutions on the reconstructed spectrum can be easily studied by changing the resolutions used in the four-vector smearing. The same detector resolutions and Bhabha generator were used for the GUINEAPIG and Model events. Differences in the predicted detector resolution and Bhabha scattering to the actual events can introduce systematic errors into the reconstruction. These effects could be studied by varying the detector resolutions or the Bhabha cross-section independently for the two samples used in the fit.

Only Bhabha events – and no other physics processes – were considered. It should be checked if multi-peripheral two-photon events, in which the spectator electrons scatter at large angles are, a background.

As the luminosity spectrum depends on the accelerator, the impact of possible variations of the beam parameter on the reconstruction of the luminosity spectrum should be studied with realistic variations of the beam parameters.

**Acknowledgements**

We are grateful to Klaus Mönig for proposing the reweighting fit to reconstruct the luminosity spectrum at future linear colliders; Barbara Dalena for providing the input files for GUINEAPIG and luminosity spectra; Daniel Schulte for useful discussions about the peculiarities of the CLIC luminosity spectrum; Jean-Jacques Blaising for the estimate of the detector resolutions and for providing the fitting code for the smuon studies; and Konrad Elsener for careful reading of this document. Our thanks also go to the members of the FCal-collaboration and the LCD analysis working group for useful discussions, inputs, and comments.
A. Probability Density Functions, Convolutions, and Random Variates

A probability density function (p.d.f.) $f(x)$ describes the distribution of a variable $x$. It is an integrable function with the properties

$$f(x) \geq 0 \quad \forall x$$

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$  \(36\)

A random variate $x_f$ is a randomly chosen value based on a probability density function $f$. The probability for an event $x$ to occur is $P(x) = f(x)dx$.

The distribution of the sum of two independent random variates can be described by the convolution

$$h(x) \equiv (f \otimes g)(x) \equiv \int_{-\infty}^{\infty} f(\tau)g(x-\tau)d\tau,$$

of the two probability density functions $f(x)$ and $g(x)$. This convolution always gives a new p.d.f. $h(x)$, since, according to Fubini’s theorem [22],

$$\left(\int_{-\infty}^{\infty} f(x)dx\right) \left(\int_{-\infty}^{\infty} g(x)dx\right) = \int_{-\infty}^{\infty} (f \otimes g)(x)dx.$$  \(39\)

A random variate $x_h$ can now either be sampled directly from the convolution $h$, or as the sum of two random variates sampled from the functions $f$ and $g$ [28, 29]

$$x_h = x_f + x_g.$$  \(40\)

Two probability density functions will be used for our study. The Gaussian function

$$g(x; \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right),$$

with the parameter $\sigma$; in this note, the Gaussian function will always have a mean of 0. The second p.d.f. is the beta-distribution

$$b(x; a_1, a_2) = \frac{1}{N} x^{a_1} (1-x)^{a_2},$$

with the parameters $a_1$ and $a_2$. The beta-distribution is only defined in the region $0 < x < 1$; and is only integrable when $-1 < a_1$ and $-1 < a_2$. With the values $a_1 < 0$ the beta-distribution rises towards 0, for $a_2 < 0$ it rises towards 1, see Figure 28. The normalisation $N$ in Equation 42 is given by the beta-function

$$N = B(a_1, a_2) = \int_{0}^{1} x^{a_1} (1-x)^{a_2}dx = \frac{\Gamma(a_1+1)\Gamma(a_2+1)}{\Gamma(2+a_1+a_2)},$$

\(43\)
where $\Gamma$ is the Gamma function. The beta-function is implemented in the ROOT::Math framework [23] as TMath::Beta($a_1 + 1, a_2 + 1$). In case of a limited range of the beta-distribution

$$ b(x; a_1, a_2, \beta_{\text{limit}}) = \begin{cases} \frac{1}{N} x^{a_1} (1-x)^{a_2} & x < \beta_{\text{limit}} \\ 0 & \beta_{\text{limit}} \leq x \end{cases} $$

the normalisation $N$ is given by the unregularized incomplete beta-function

$$ B(a_1, a_2; \beta_{\text{limit}}) = \int_0^{\beta_{\text{limit}}} x^{a_1} (1-x)^{a_2} dx. $$

In ROOT::Math only the regularized incomplete beta-function, that is the limited integral divided by the complete integral, is implemented as TMath::BetaIncomplete($\beta_{\text{limit}}, a_1 + 1, a_2 + 1$), which requires that the function return value is multiplied by the beta-function.

References


