Generating the full SM at linear colliders

Mikael Berggren[1]

[1] DESY, Hamburg, On behalf of the generator group (LCGG) of the Linear Collider Collaboration (LCC)

40th International Conference on High Energy Physics, Prague, Czech republic , July 27 - August 6, 2020

Outline

1. Linear colliders
2. Generating the full SM
   - Process classification
   - Physics generator
   - Generating beam properties
   - Setup, integration, event generation, documentation
3. Outlook
4. Conclusions

Outline

Disclaimer

This is not a talk about physics in event generators, nor about detector simulation.

Linear colliders

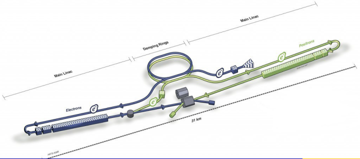
- polarised e^+e^- colliders with E_{CMS} = (90)-250 - 3000 GeV.
- e^+e^- means EW-production $\Rightarrow$ Low background.
  - Detectors w/ $\sim 4\pi$ coverage.
  - Rad. hardness not needed: only few % X_0 in front of calorimters.
  - No trigger
- e^+e^- means colliding point-like objects $\Rightarrow$ initial state known

Linear colliders

- polarised e^+e^- colliders with E_{CMS} = (90)-250 - 3000 GeV.
- e^+e^- means EW-production $\Rightarrow$ Low background.
  - Detectors w/ $\sim 4\pi$ coverage.
  - Rad. hardness not needed: only few % X_0 in front of calorimters.
  - No trigger
- e^+e^- means colliding point-like objects $\Rightarrow$ initial state known

ILC:

- 20 year running $\rightarrow$ 4 ab^{-1} @ 500 GeV, 2 ab^{-1} @ 250 GeV.
- Construction under political consideration in Japan.



CLIC:

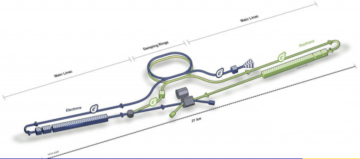
- 20 year running $\rightarrow$ 5 ab^{-1} @ 3 TeV, 2.5 ab^{-1} @ 1.5 TeV, 1 ab^{-1} @ 380 GeV.
- Possible future CERN project.

Linear colliders

- polarised e^+e^- colliders with E_{CMS} = (90)-250 - 3000 GeV.
- e^+e^- means EW-production $\Rightarrow$ Low background.
  - Detectors w/ $\sim 4\pi$ coverage.
  - Rad. hardness not needed: only few % X_0 in front of calorimters.
  - No trigger
- e^+e^- means colliding point-like objects $\Rightarrow$ initial state known

ILC:

- 20 year running $\rightarrow$ 4 ab^{-1} @ 500 GeV, 2 ab^{-1} @ 250 GeV.
- Construction under political consideration in Japan.



CLIC:

- 20 year running $\rightarrow$ 5 ab^{-1} @ 3 TeV, 2.5 ab^{-1} @ 1.5 TeV, 1 ab^{-1} @ 380 GeV.
- Possible future CERN project.

MC requirements for linear colliders

- Future LCs aim for extremely high precision measurements.
  - $\Rightarrow$ Need excellent detector, well controlled machine conditions
  - But also the best possible estimate of backgrounds.
- MC statistics or lacking channels must not be a major source of systematic errors $\Rightarrow$
  - All SM channels yielding at least a few events under the full lifetime of the projects need to be generated, with statistics largely exceeding that of the real data.
  - Also machine conditions need to be accurately taken into account.
- In addition: at an LC ALL events are interesting, and often fully reconstructed. More like a B-factory than LHC!
- This endeavour has been organised as a common effort between ILD and SiD at ILC and CLICdp at CLIC. The work is done within the generator group, LCGG.

MC requirements for linear colliders

- Future LCs aim for extremely high precision measurements.
  - ⇒ Need excellent detector, well controlled machine conditions
  - But also the best possible estimate of backgrounds.
- MC statistics or lacking channels must not be a major source of systematic errors ⇒
  - All SM channels yielding at least a few events under the full lifetime of the projects need to be generated, with statistics largely exceeding that of the real data.
  - Also machine conditions need to be accurately taken into account.
- In addition: at an LC ALL events are interesting, and often fully reconstructed. More like a B-factory than LHC!
- This endeavour has been organised as a common effort between ILD and SiD at ILC and CLICdp at CLIC. The work is done within the generator group, LCGG.

MC requirements for linear colliders

- Future LCs aim for extremely high precision measurements.
  - $\Rightarrow$ Need excellent detector, well controlled machine conditions
  - But also the best possible estimate of backgrounds.
- MC statistics or lacking channels must not be a major source of systematic errors $\Rightarrow$
  - All SM channels yielding at least a few events under the full lifetime of the projects need to be generated, with statistics largely exceeding that of the real data.
  - Also machine conditions need to be accurately taken into account.
- In addition: at an LC ALL events are interesting, and often fully reconstructed. More like a B-factory than LHC!
- This endeavour has been organised as a common effort between ILD and SiD at ILC and CLICdp at CLIC. The work is done within the generator group, LCGG.

MC requirements for linear colliders

- Future LCs aim for extremely high precision measurements.
  - $\Rightarrow$ Need excellent detector, well controlled machine conditions
  - But also the best possible estimate of backgrounds.
- MC statistics or lacking channels must not be a major source of systematic errors $\Rightarrow$
  - All SM channels yielding at least a few events under the full lifetime of the projects need to be generated, with statistics largely exceeding that of the real data.
  - Also machine conditions need to be accurately taken into account.
- In addition: at an LC ALL events are interesting, and often fully reconstructed. More like a B-factory than LHC!
- This endeavour has been organised as a common effort between ILD and SiD at ILC and CLICdp at CLIC. The work is done within the generator group, LCGG.

MC requirements for linear colliders

- Future LCs aim for extremely high precision measurements.
  - $\Rightarrow$ Need excellent detector, well controlled machine conditions
  - But also the best possible estimate of backgrounds.
- MC statistics or lacking channels must not be a major source of systematic errors $\Rightarrow$
  - All SM channels yielding at least a few events under the full lifetime of the projects need to be generated, with statistics largely exceeding that of the real data.
  - Also machine conditions need to be accurately taken into account.
- In addition: at an LC ALL events are interesting, and often fully reconstructed. More like a B-factory than LHC!
- This endeavour has been organised as a common effort between ILD and SiD at ILC and CLICdp at CLIC. The work is done within the generator group, LCGG.

Generating the full SM

Generate the full SM? What's the problem?

- Just select a generator, and press <RET>, right?
- Noooo..., not really. Lots of details:
  - What collides ($e^{+/-}$ or γ ?)
    - What energy do they have, and how are they polarised ?
  - Where do they collide ?
    - Beam-spot properties
  - What else happens?
    - Beam-strahlung gives pairs.
      Do they hit anything ? Maybe forward calorimetry, or the tracking system ?
    - Multiple interactions (pile-up) ?

Generating the full SM

Generate the full SM? What's the problem?

- Just select a generator, and press <RET>, right?
- Noooo..., not really. Lots of details:
  - What collides ($e^{+/-}$ or γ ?)
    - What energy do they have, and how are they polarised ?
  - Where do they collide ?
    - Beam-spot properties
  - What else happens?
    - Beam-strahlung gives pairs. Do they hit anything ? Maybe forward calorimetry, or the tracking system ?
    - Multiple interactions (pile-up) ?

Generating the full SM

Generate the full SM? What's the problem?

- Just select a generator, and press <RET>, right?
- Noooo..., not really. Lots of details:
  - What collides (e$^{+/-}$ or γ ?)
    - What energy do they have, and how are they polarised ?
  - Where do they collide ?
    - Beam-spot properties
  - What else happens?
    - Beam-strahlung gives pairs
      Do they hit anything ? Maybe forward calorimetry, or the tracking system ?
    - Multiple interactions (pile-up) ?

Generating the full SM

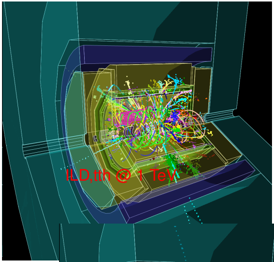
Generate the full SM? What's the problem?

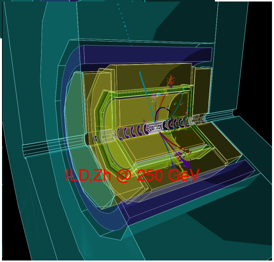
- Just select a generator, and press <RET>, right?
- Noooo..., not really. Lots of details:
  - What collides (e$^{+/-}$ or γ ?)
    - What energy do they have, and how are they polarised ?
  - Where do they collide ?
    - Beam-spot properties
  - What else happens?
    - Beam-strahlung gives pairs
      Do they hit anything ? Maybe forward calorimetry, or the tracking system ?
    - Multiple interactions (pile-up) ?

Generating the full SM

In addition, the physics channels themselves:

- At an LC, ALL events are interesting !
- Huge spread in cross-sections
- But for any given study, it might be a tiny cross-section one that dominates...
- ... or maybe a tiny fraction of a huge cross-section one.
- We want to make nice stacked histos of different backgrounds - different from analysis to analysis.
- ⇒ Good idea to separate channels in an intelligent and well documented way

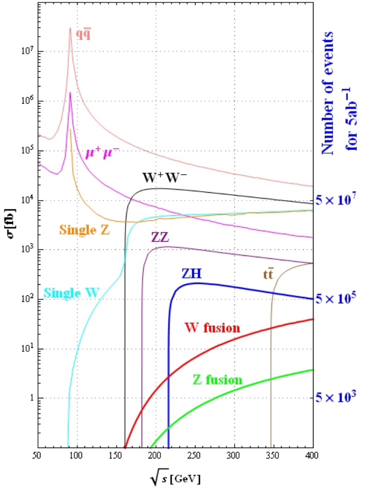




Generating the full SM

In addition, the physics channels themselves:

- At an LC, ALL events are interesting !
- Huge spread in cross-sections
- But for any given study, it might be a tiny cross-section one that dominates...
- ... or maybe a tiny fraction of a huge cross-section one.
- We want to make nice stacked histos of different backgrounds - different from analysis to analysis.
- $\Rightarrow$ Good idea to separate channels in an intelligent and well documented way

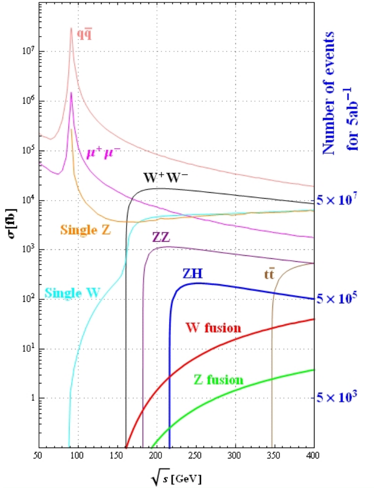


[from CepC, arXiv:1505.01008]

Generating the full SM

In addition, the physics channels themselves:

- At an LC, ALL events are interesting !
- Huge spread in cross-sections
- But for any given study, it might be a tiny cross-section one that dominates...
- ... or maybe a tiny fraction of a huge cross-section one.
- We want to make nice stacked histos of different backgrounds - different from analysis to analysis.
- ⇒ Good idea to separate channels in an intelligent and well documented way

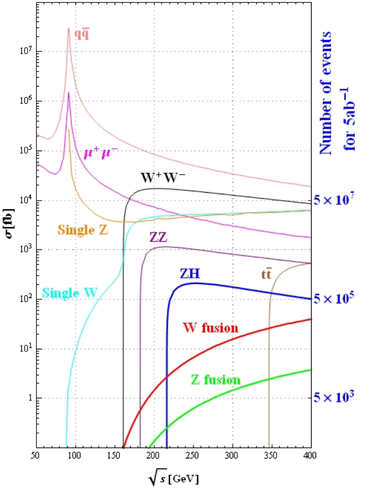


[from CepC, arXiv:1505.01008]

Generating the full SM

In addition, the physics channels themselves:

- At an LC, ALL events are interesting !
- Huge spread in cross-sections
- But for any given study, it might be a tiny cross-section one that dominates...
- ... or maybe a tiny fraction of a huge cross-section one.
- We want to make nice stacked histos of different backgrounds - different from analysis to analysis.
- ⇒ Good idea to separate channels in an intelligent and well documented way

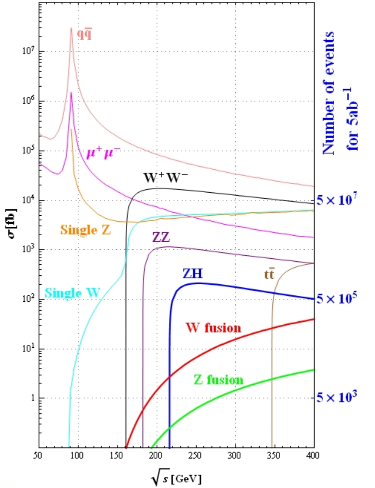


[from CepC, arXiv:1505.01008]

Generating the full SM

In addition, the physics channels themselves:

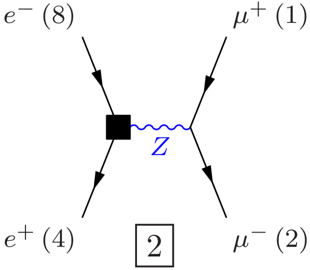
- At an LC, ALL events are interesting !
- Huge spread in cross-sections
- But for any given study, it might be a tiny cross-section one that dominates...
- ... or maybe a tiny fraction of a huge cross-section one.
- We want to make nice stacked histos of different backgrounds - different from analysis to analysis.
- $\Rightarrow$ Good idea to separate channels in an intelligent and well documented way



[from CepC, arXiv:1505.01008]

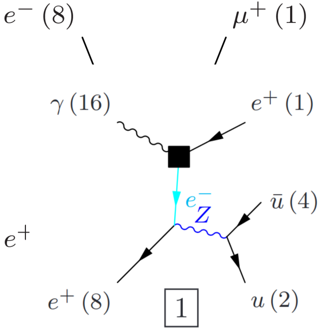
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \rightarrow f\bar{f}$ vs. $e^+e^- \rightarrow e^+e^- f\bar{f}$.



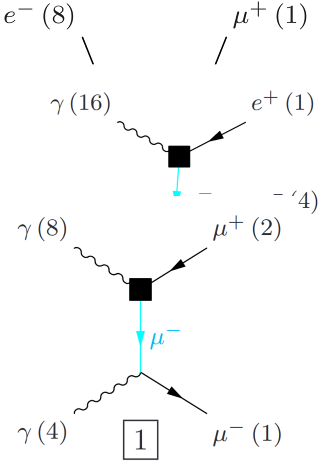
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $e^+e^- \to e^+e^-f\bar{f}$.



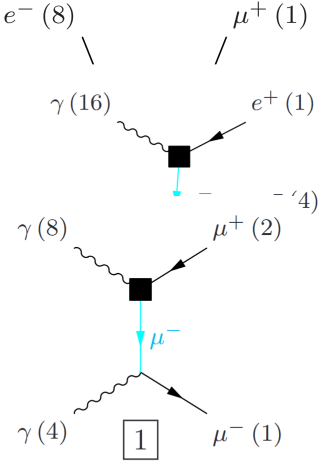
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \rightarrow f\bar{f}$ vs. $e^+e^- \rightarrow e^+e^-f\bar{f}$.



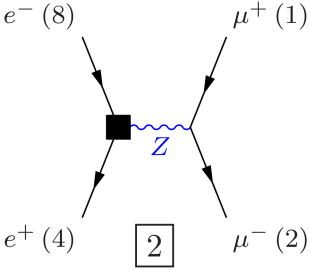
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $e^+e^- \to e^+e^- f\bar{f}$.



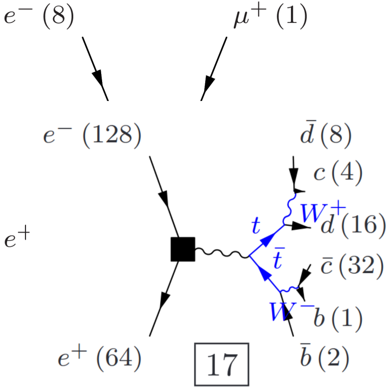
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $e^+e^- \to e^+e^- f\bar{f}$.



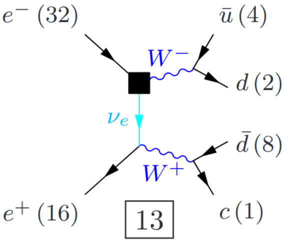
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptoni (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \rightarrow f\bar{f}$ vs. $e^+e^- \rightarrow e^+e^- f\bar{f}$.



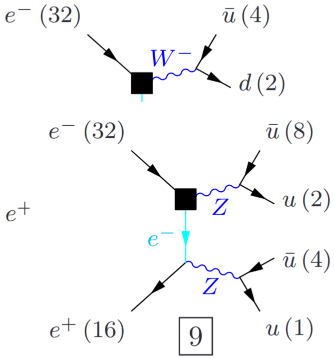
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptonic (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \rightarrow f\bar{f}$ vs. $e^+e^- \rightarrow e^+e^- f\bar{f}$.



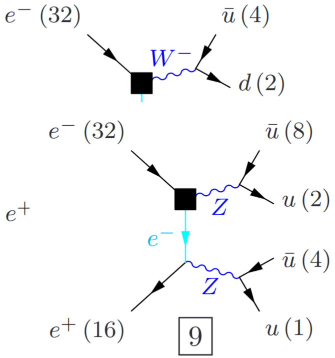
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptonic (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $e^+e^- \to e^+e^- f\bar{f}$.



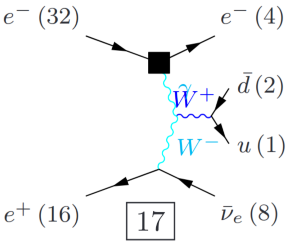
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptonic (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $e^+e^- \to e^+e^-f\bar{f}$.



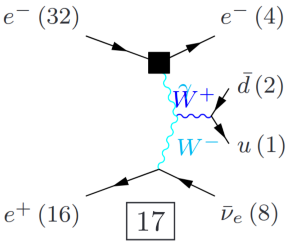
Process classification

- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptonic (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \to f\bar{f}$ vs. $\mathrm{e}^+\mathrm{e}^- \to e^+e^-f\bar{f}$.



Process classification

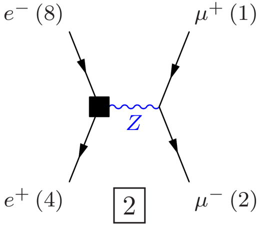
- Initial state
  - ee, eγ or $\gamma\gamma$
  - e - polarisation and γ type (real or virtual)
- Final state
  - Number of fermions (1 to 8)
  - Flavour-grouping: W or Z, or ambiguous
  - leptonic, hadronic, semi-leptonic (+ neutrino only, for Z-leptonic)
- Special considerations
  - Eg. 4f with $|L_e|$=2 $\Rightarrow$ dominated by single W or single Z (t-channel !)
  - Avoid double-counting: Eg. $\gamma^*\gamma^* \rightarrow f\bar{f}$ vs. $\mathrm{e^+e^-} \rightarrow e^+e^-f\bar{f}$.



Main generator: Whizard

Whizard is the generator of choice for e^+e^-

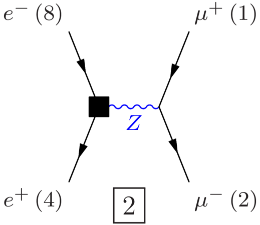
- Full matrix-element evaluation.
  Only at tree-level but:
  - Polarised beams.
  - Full helicity treatment.
  - Full colour flow, passed from the hard interaction to the P.S. code.
  - Can handle beam-spectrum, using Circe2.
  - Can handle polarised τ-decays, using TAUOLA [Comp Phys Comm,64,p275].
  - And: Can do $2 \rightarrow 8$ processes.
- ... which is more important than NLO for e^+e^- ! That was just one 5f tree-level diagram of ~ 500...



Main generator: Whizard

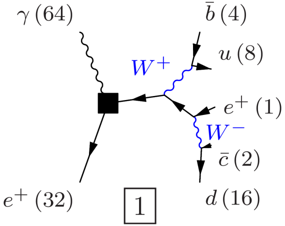
Whizard is the generator of choice for e^+e^-

- Full matrix-element evaluation.
  Only at tree-level but:
  - Polarised beams.
  - Full helicity treatment.
  - Full colour flow, passed from the hard interaction to the P.S. code.
  - Can handle beam-spectrum, using Circe2.
  - Can handle polarised τ-decays, using TAUOLA [Comp Phys Comm,64,p275].
  - And: Can do $2 \rightarrow 8$ processes.
- ... which is more important than NLO for e^+e^- ! That was just one 5f tree-level diagram of ~ 500...



Main generator: Whizard

Whizard is the generator of choice for e^+e^-

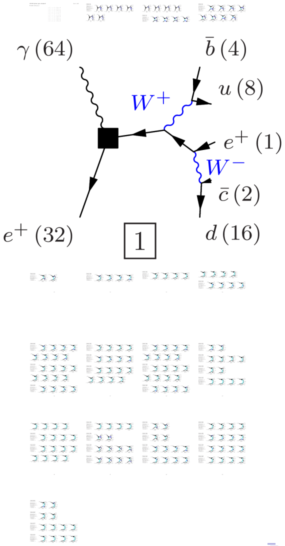


- Full matrix-element evaluation. Only at tree-level but:
  - Polarised beams.
  - Full helicity treatment.
  - Full colour flow, passed from the hard interaction to the P.S. code.
  - Can handle beam-spectrum, using Circe2.
  - Can handle polarised τ-decays, using TAUOLA [Comp Phys Comm,64,p275].
  - And: Can do $2 \rightarrow 8$ processes.
- ... which is more important than NLO for e^+e^- ! That was just one 5f tree-level diagram of ~ 500...

Main generator: Whizard

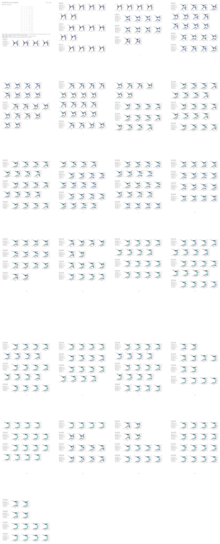
Whizard is the generator of choice for e^+e^-

- Full matrix-element evaluation. Only at tree-level but:
  - Polarised beams.
  - Full helicity treatment.
  - Full colour flow, passed from the hard interaction to the P.S. code.
  - Can handle beam-spectrum, using Circe2.
  - Can handle polarised τ-decays, using TAUOLA [Comp Phys Comm,64,p275].
  - And: Can do $2 \to 8$ processes.
- ... which is more important than NLO for e^+e^- ! That was just one 5f tree-level diagram of ~ 500...



Main generator: Whizard

Whizard is the generator of choice for e^+e^-.



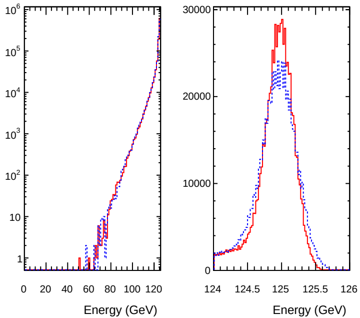
- Full matrix-element evaluation. Only at tree-level but:
  - Polarised beams.
  - Full helicity treatment.
  - Full colour flow, passed from the hard interaction to the P.S. code.
  - Can handle beam-spectrum, using Circe2.
  - Can handle polarised τ-decays, using TAUOLA [Comp Phys Comm,64,p275].
  - And: Can do $2 \rightarrow 8$ processes.
- ... which is more important than NLO for e^+e^- ! That was just one 5f tree-level diagram of ~ 500...

Main generator: Whizard

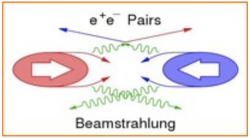
- The subsequent parton-shower and hadronisation is done by PYTHIA6.4.
  - LCGG has tuned hadronisation using input from OPAL at LEPII [Phys.Rept. 291 (1997) 107-217, D. Ward, private communication.].
- The process-definition given in the Whizard steering file (aka the *sindarin*) is also the driver for the scripts that organises the production: One ring to rule them all.
- Use powerful grouping and aliasing capabilities of sindarin to assure that no processes are over-looked.

Beam effects

- Beam-spectrum.
  1. Incoming beam-spread
  2. But also: very strongly focused beams ⇒ Beam-beam interactions

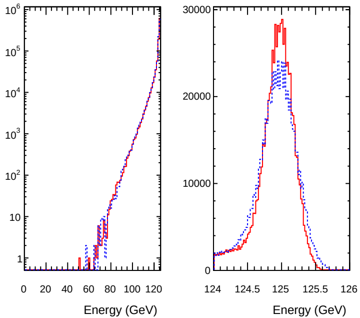


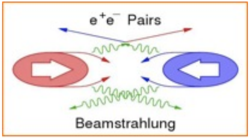
- Photons
  1. How many photons?
  2. Are they virtual or real?
- Need beam-beam interaction simulation input.
- Simulate interaction region: GuineaPig [CERN-PS-99-014-LP]. Gives:
  - Beam-spectrum for electrons and positrons independently
  - Amount and spectrum of real photons
  - Distribution of interaction point



Beam effects

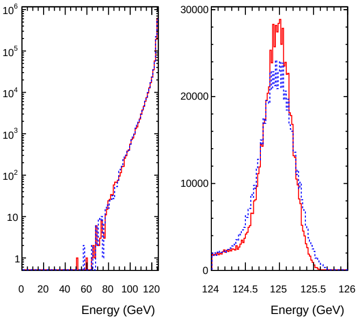
- Beam-spectrum.
  1. Incoming beam-spread
  2. But also: very strongly focused beams ⇒ Beam-beam interactions
- Photons
  1. How many photons?
  2. Are they virtual or real?
- Need beam-beam interaction simulation input.
- Simulate interaction region: GuineaPig [CERN-PS-99-014-LP]. Gives:
  - Beam-spectrum for electrons and positrons independently
  - Amount and spectrum of real photons
  - Distribution of interaction point

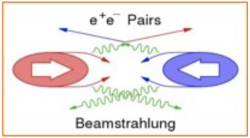




Beam effects

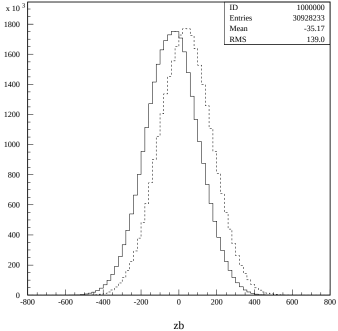
- Beam-spectrum.
  1. Incoming beam-spread
  2. But also: very strongly focused beams $\Rightarrow$ Beam-beam interactions
- Photons
  1. How many photons?
  2. Are they virtual or real?
- Need beam-beam interaction simulation input.
- Simulate interaction region: GuineaPig [CERN-PS-99-014-LP]. Gives:
  - Beam-spectrum for electrons and positrons independently
  - Amount and spectrum of real photons
  - Distribution of interaction point

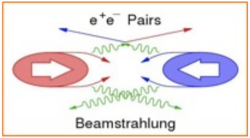




Beam effects

- Beam-spectrum.
  1. Incoming beam-spread
  2. But also: very strongly focused beams $\Rightarrow$ Beam-beam interactions
- Photons
  1. How many photons?
  2. Are they virtual or real?
- Need beam-beam interaction simulation input.
- Simulate interaction region: GuineaPig [CERN-PS-99-014-LP]. Gives:
  - Beam-spectrum for electrons and positrons independently
  - Amount and spectrum of real photons
  - Distribution of interaction point

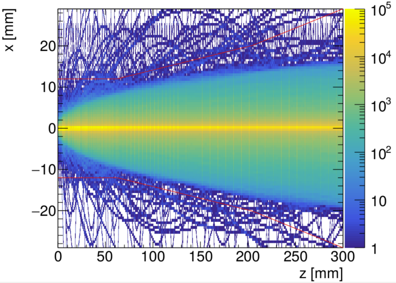




Spurious interactions ("pile-up")

Two types:

- Pair-background: Pair-creation of photons in the beam by the strong fields. GuineaPig can generate the full activity during a beam-crossing (a "BX").
- low-p$_\perp$ hadrons, ie. $\gamma^{(*)}\gamma^{(*)}$ interaction with small $M_{\gamma\gamma}$ and multiplicity. NB: only $\mathcal{O}(1)$/BX !
  - ME can't do this, and PYTHIA is good down to $M_{\gamma\gamma} \sim 2$ GeV.
  - Below: fit to data - Custom generator developed by LCGG.
- Both types: Pre-generate pool of events, pick at random and overlay on main event.

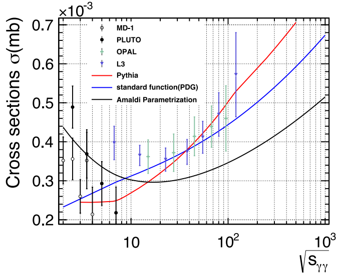


- For pairs: Do $\sim 10^5$ BXes, use fast detector simulation SGV to filter out track hitting tracking ($\sim$10/BX), rest used to build a map of background on BeamCal.

Spurious interactions (“pile-up”)

Two types:

- Pair-background: Pair-creation of photons in the beam by the strong fields. GuineaPig can generate the full activity during a beam-crossing (a “BX”).
- low-p$_\perp$ hadrons, ie. $\gamma^{(*)}\gamma^{(*)}$ interaction with small $M_{\gamma\gamma}$ and multiplicity. NB: only $\mathcal{O}(1)$/BX !
  - ME can't do this, and PYTHIA is good down to $M_{\gamma\gamma} \sim 2$ GeV.
  - Below: fit to data - Custom generator developed by LCGG.
- Both types: Pre-generate pool of events, pick at random and overlay on main event.

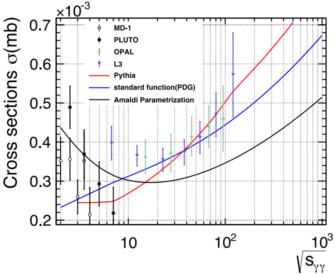


- For pairs: Do $\sim 10^5$ BXes, use fast detector simulation SGV to filter out track hitting tracking ($\sim$10/BX), rest used to build a map of background on BeamCal.

Spurious interactions (“pile-up”)

Two types:

- Pair-background: Pair-creation of photons in the beam by the strong fields. GuineaPig can generate the full activity during a beam-crossing (a “BX”).
- low-p$_\perp$ hadrons, ie. $\gamma^{(*)}\gamma^{(*)}$ interaction with small $M_{\gamma\gamma}$ and multiplicity. NB: only $\mathcal{O}(1)$/BX !
  - ME can't do this, and PYTHIA is good down to $M_{\gamma\gamma} \sim 2$ GeV.
  - Below: fit to data - Custom generator developed by LCGG.
- Both types: Pre-generate pool of events, pick at random and overlay on main event.

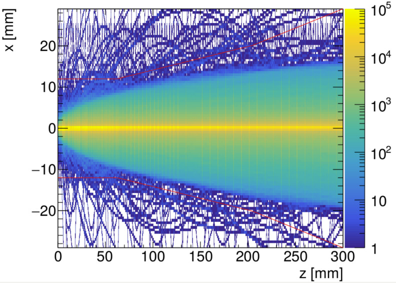


- For pairs: Do $\sim 10^5$ BXes, use fast detector simulation SGV to filter out track hitting tracking ($\sim$10/BX), rest used to build a map of background on BeamCal.

Spurious interactions ("pile-up")

Two types:

- Pair-background: Pair-creation of photons in the beam by the strong fields. GuineaPig can generate the full activity during a beam-crossing (a "BX").
- low-p$_\perp$ hadrons, ie. $\gamma^{(*)}\gamma^{(*)}$ interaction with small $M_{\gamma\gamma}$ and multiplicity. NB: only $\mathcal{O}(1)$/BX !
  - ME can't do this, and PYTHIA is good down to $M_{\gamma\gamma} \sim 2$ GeV.
  - Below: fit to data - Custom generator developed by LCGG.
- Both types: Pre-generate pool of events, pick at random and overlay on main event.



- For pairs: Do $\sim 10^5$ BXes, use fast detector simulation SGV to filter out track hitting tracking ($\sim$10/BX), rest used to build a map of background on BeamCal.

Setup & integration

- Whizard process definition is parsed to build a directory-tree structure - one unique directory per process.
- Process-specific code is generated and compiled (interactive)
- The tree is traversed to do a "pre-integration" of all channels, to flag zero cross-section ones (interactive).
- Full integration of all channels, with error goal 0.1 % submitted (local batch-farm under Condor). Over-night for $\leq$ 5f.
- Pilot generation of all channels, (1k events/channel), to evaluate CPU time and storage needed.

Setup & integration

- Whizard process definition is parsed to build a directory-tree structure - one unique directory per process.
- Process-specific code is generated and compiled (interactive)
- The tree is tra[...] all channels, to flag zero cross[...]
- Full integration of all channels, with error goal 0.1 % submitted (local batch-farm under Condor). Over-night for $\leq$ 5f.
- Pilot generation of all channels, (1k events/channel), to evaluate CPU time and storage needed.

At this point, the system is ready to generate the full SM

Generation

- Done on a batch-farm.
- In most cases: one process per generation job.
- Some processes alone represent $\sim$ billion events. These must be split in several jobs.
- At the end of each job, the events (in LCIO format), metadata, and input+logfile tarballs are uploaded to the grid, where the simulation and reconstruction system under DIRAC kicks in - details in talk of R. Ete in this track on Friday.
- At this point, disk can be freed on the batch cluster (otherwise disk-space might become an issue).

Generation

- Done on a batch-farm.
- In most cases: one process per generation job.
- Some processes alone represent $\sim$ billion events. These must be split in several jobs.
- At the end of each job, the events (in LCIO format), metadata, and input+logfile tarballs are uploaded to the grid, where the simulation and reconstruction system under DIRAC kicks in - details in talk of R. Ete in this track on Friday.
- At this point, disk can be freed on the batch cluster (otherwise disk-space might become an issue).

Documentation

Created by generation job, driven by the contents of the process-definition Sindarin script and common conditions. It goes into:

- The event header:
  - Process-id, beam-polarisation, and cross-section of each event.
- The generator meta-data files:
  - Condenses job-specific information from Whizard logs.
  - Contains: process, cross section, polarisation, file-names, ...
  - Browasble on the Web and uploaded to the Grid.
  - Once on the grid, the metadata is read by DIRAC to orchestrate the detector simulation.
- Steering-files, logs, pdf:s with diagrams, integration grids, output other than the events,...:
  - On the Web, in full.
  - In tar files - parallel to generated files, also stored on the Grid.

Documentation

Created by generation job, driven by the contents of the process-definition Sindarin script and common conditions. It goes into:

- The event header:
  - Process-id, beam-polarisation, and cross-section of each event.
- The generator meta-data files:
  - Condenses job-specific information from Whizard logs.
  - Contains: process, cross section, polarisation, file-names, ...
  - Browasble on the Web and uploaded to the Grid.
  - Once on the grid, the metadata is read by DIRAC to orchestrate the detector simulation.
- Steering-files, logs, pdf:s with diagrams, integration grids, output other than the events,...:
  - On the Web, in full.
  - In tar files - parallel to generated files, also stored on the Grid.

Documentation

Created by generation job, driven by the contents of the process-definition Sindarin script and common conditions. It goes into:

- The event header:
  - Process-id, beam-polarisation, and cross-section of each event.
- The generator meta-data files:
  - Condenses job-specific information from Whizard logs.
  - Contains: process, cross section, polarisation, file-names, ...
  - Browsable on the Web and uploaded to the Grid.
  - Once on the grid, the metadata is read by DIRAC to orchestrate the detector simulation.
- Steering-files, logs, pdf:s with diagrams, integration grids, output other than the events,...:
  - On the Web, in full.
  - In tar files - parallel to generated files, also stored on the Grid.

Documentation

Created by generation job, driven by the contents of the process-definition Sindarin script and common conditions. It goes into:

- The event header:
  - Process-id, beam-polarisation, and cross-section of each event.
- The generator meta-data files:
  - Condenses job-specific information from Whizard logs.
  - Contains: process, cross section, polarisation, file-names, ...
  - Browsable on the Web and uploaded to the Grid.
  - Once on the grid, the metadata is read by DIRAC to orchestrate the detector simulation.
- Steering-files, logs, pdf:s with diagrams, integration grids, output other than the events,...:
  - On the Web, in full.
  - In tar files - parallel to generated files, also stored on the Grid.

Future tasks

Soon: massive generation for ILC@250, with latest ILC parameters

- $\mathcal{O}(10)$ times more than the DBD, more than expected real data.
  - For $\leq$ 5f: CPU $\sim$20000 h - Disk $\sim$10 TB : Much, but not too much...
  - Using $\mathcal{O}(100)$ cores, this generation is done in 1 to 2 weeks.

Whizard developments

- gluon matching between ME and PS:
  - Now: no gluons included in ME (setting $\alpha_s = 0$) to avoid double-counting with (unmatched) parton shower in PYTHIA.
  - Whizard 2 does parton-shower with "MLM matching". Exploit this !
- Medium-term wishes for Whizard:
  - γ ISR/FSR matching
  - Work out priority processes for EW-NLO (!)

In general it would be nice to also have other generators

- BHWide for better Bhabhas. BDK/BDKRC for $\gamma\gamma \to \ell\ell$
- Pythia8, MadGraph, Sherpa for double-checks.
- Pythia8 instead of Pythia6 for hadronisation.

Future tasks

Soon: massive generation for ILC@250, with latest ILC parameters

- $\mathcal{O}(10)$ times more than the DBD, more than expected real data.
  - For $\leq$ 5f: CPU ∼20000 h - Disk ∼10 TB : Much, but not too much...
  - Using $\mathcal{O}(100)$ cores, this generation is done in 1 to 2 weeks.

Whizard developments

- gluon matching between ME and PS:
  - Now: no gluons included in ME (setting α_s = 0) to avoid double-counting with (unmatched) parton shower in PYTHIA.
  - Whizard 2 does parton-shower with "MLM matching". Exploit this !
- Medium-term wishes for Whizard:
  - γ ISR/FSR matching
  - Work out priority processes for EW-NLO (!)

In general it would be nice to also have other generators

- BHWide for better Bhabhas. BDK/BDKRC for $\gamma\gamma \to \ell\ell$
- Pythia8, MadGraph, Sherpa for double-checks.
- Pythia8 instead of Pythia6 for hadronisation.

Future tasks

Soon: massive generation for ILC@250, with latest ILC parameters

- $\mathcal{O}(10)$ times more than the DBD, more than expected real data.
  - For $\leq$ 5f: CPU ~20000 h - Disk ~10 TB : Much, but not too much...
  - Using $\mathcal{O}(100)$ cores, this generation is done in 1 to 2 weeks.

Whizard developments

- gluon matching between ME and PS:
  - Now: no gluons included in ME (setting α_s = 0) to avoid double-counting with (unmatched) parton shower in PYTHIA.
  - Whizard 2 does parton-shower with “MLM matching”. Exploit this !
- Medium-term wishes for Whizard:
  - γ ISR/FSR matching
  - Work out priority processes for EW-NLO (!)

In general it would be nice to also have other generators

- BHWide for better Bhabhas. BDK/BDKRC for $\gamma\gamma \rightarrow \ell\ell$
- Pythia8, MadGraph, Sherpa for double-checks.
- Pythia8 instead of Pythia6 for hadronisation.

Conclusions

- We have explained that the precision-goal of future e^+e^- colliders are such that it is not permissible that MC statistics or coverage could constitute an major systematic uncertainty.
- We showed how the generation of the full SM can be achieved.
  - It consists of bringing a large number of different codes together:
    - Whizard, Pythia, and Tauola for physics generation
    - GuineaPig and Circe2 for beam-properties
    - SGV+GuineaPig and the Peskin/Barklow generator for pile-up
    - In addition, input from machine-physics and data from LEPII was used.
- This full data is organised and documented in a physics-oriented fashion, for the benefit of the end-user.
- The entire system pivots around one data-source, the Whizard process definition file.

Conclusions

- We have explained that the precision-goal of future e^+e^- colliders are such that it is not permissible that MC statistics or coverage could constitute an major systematic uncertainty.
- We showed how the generation of the full SM can be achieved.
  - It consists of bringing a large number of different codes together:
    - Whizard, Pythia, and Tauola for physics generation
    - GuineaPig and Circe2 for beam-properties
    - SGV+GuineaPig and the Peskin/Barklow generator for pile-up
    - In addition, input from machine-physics and data from LEPII was used.
- This full data is organised and documented in a physics-oriented fashion, for the benefit of the end-user.
- The entire system pivots around one data-source, the Whizard process definition file.

Conclusions

- We have explained that the precision-goal of future e^+e^- colliders are such that it is not permissible that MC statistics or coverage could constitute an major systematic uncertainty.
- We showed how the generation of the full SM can be achieved.
  - It consists of bringing a large number of different codes together:
    - Whizard, Pythia, and Tauola for physics generation
    - GuineaPig and Circe2 for beam-properties
    - SGV+GuineaPig and the Peskin/Barklow generator for pile-up
    - In addition, input from machine-physics and data from LEPII was used.
- This full data is organised and documented in a physics-oriented fashion, for the benefit of the end-user.
- The entire system pivots around one data-source, the Whizard process definition file.

Conclusions

- We have explained that the precision-goal of future e^+e^- colliders are such that it is not permissible that MC statistics or coverage could constitute an major systematic uncertainty.
- We showed how the generation of the full SM can be achieved.
    - It consists of bringing a large number of different codes together:
        - Whizard, Pythia, and Tauola for physics generation
        - GuineaPig and Circe2 for beam-properties
        - SGV+GuineaPig and the Peskin/Barklow generator for pile-up
        - In addition, input from machine-physics and data from LEPII was used.
- This full data is organised and documented in a physics-oriented fashion, for the benefit of the end-user.
- The entire system pivots around one data-source, the Whizard process definition file.