Optimization of the cut selection of the $h^0/H^0/A^0 \to \mu^+\mu^-$ analysis

J. Philipp Grohs,

Prof. Michael Kobel, Dr. Anja Vest

With Support by Sebastian Stern
Max-Planck-Institut für Physik München

IKTP Institute's seminar
January 19, 2012

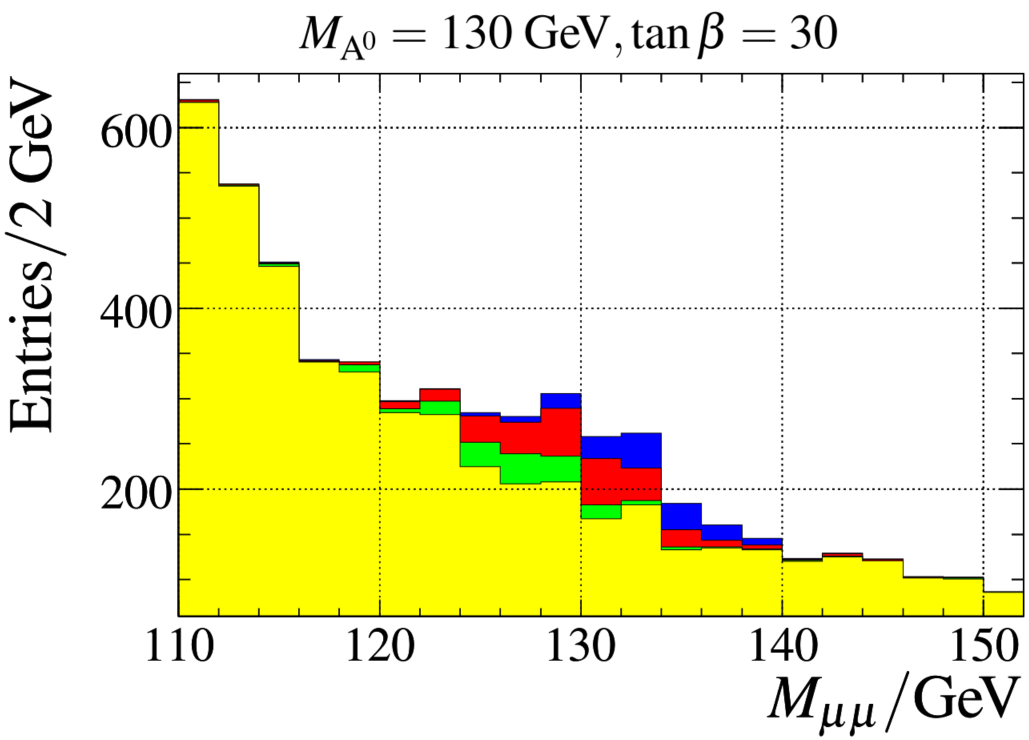
TECHNISCHE UNIVERSITÄT DRESDEN

Content

- Search for neutral MSSM Higgs bosons decaying into two muons
- Cut variables
- Distribution of di-muon invariant mass
- Optimization problem
    - Objective function
    - Optimization algorithm
- Results
- Summary

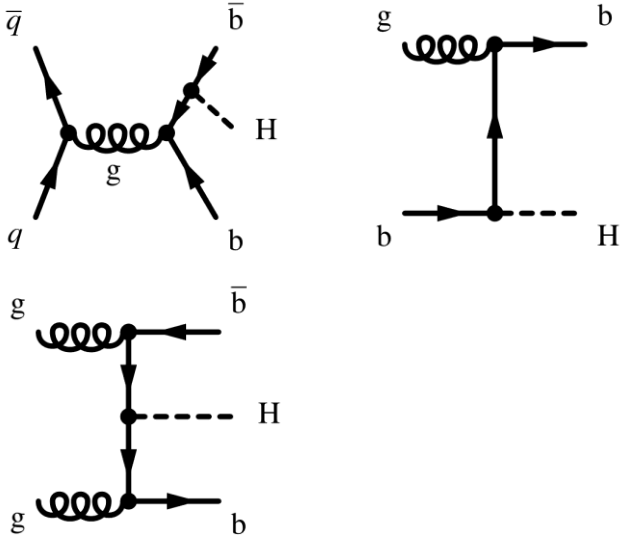
Introduction

- One of the three most interesting decay channels of neutral MSSM Higgs bosons for high $\tan\beta$ and low Higgs masses ($110\ GeV \leq M_{Higgs} \leq 250\ GeV$)
- Low branching ratio of $\sim 0.04\%$
- But: Very clean signal and high mass resolution

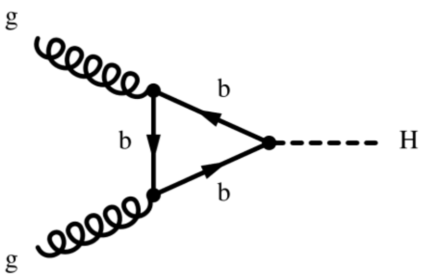


Signal processes

b-associated Production:



Gluon-Fusion:

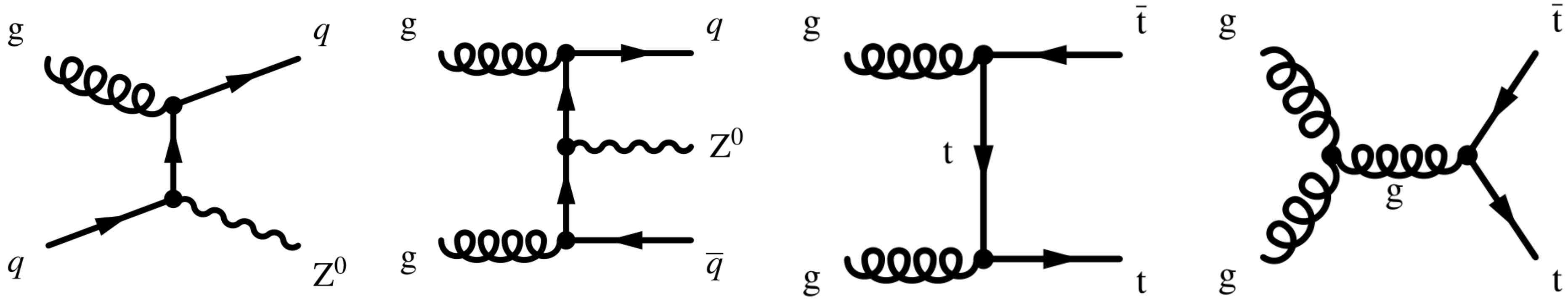


Decay Signature:

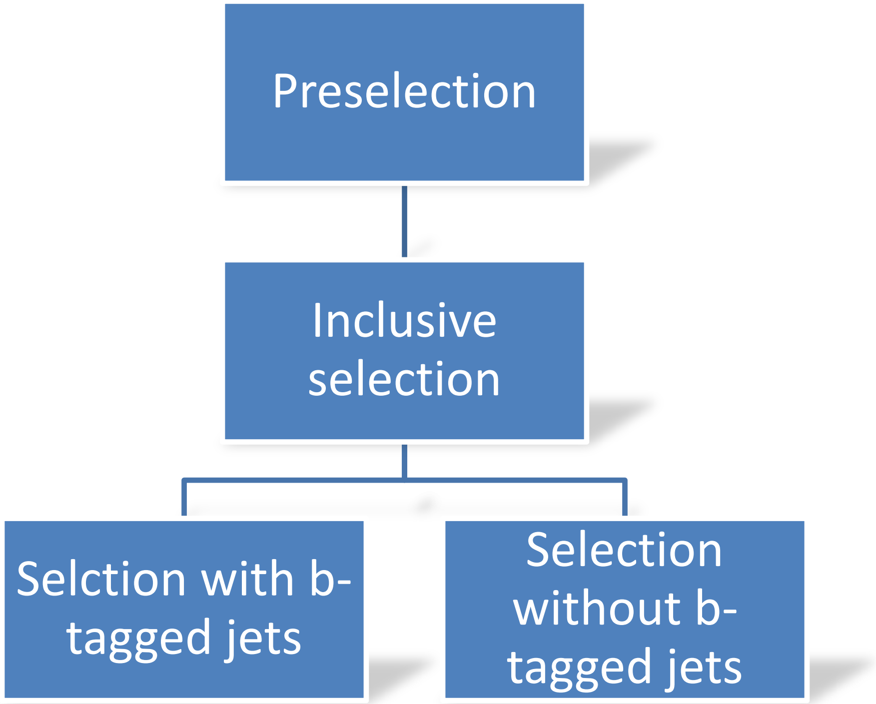
- Two muons oppositely charged
- Breit-Wigner peak of $M_{\mu\mu}$ around Higgs mass
- No missing transverse energy
- (Jets containing b-quarks)
- Higgs with low p_T, boosted along the beam pipe
  →large opening angle of muon pair

Background processes

- $Z \rightarrow \mu\mu$ process irreducible if produced with b-association
- $t\bar{t} \rightarrow \mu\mu X$ process:
  - Two b-quarks in final state
  - High jet activity (sum of p_T of all jets)
  - No angular correlation between muon pair due to different decays



Analysis strategies



Preselection:

- Object selection (muons, jets, missing E_T)
- Trigger
- Cleaning cuts

Inclusive selection:

- Mass window $M_{\mu\mu} > 70\ GeV$
- **Missing transverse energy** $E_T^{miss} < 40\ GeV$

Selection with b-tagged jets:

- $N_{b-tagged\ Jets} \geq 1$
- **Jet activity** $\sum p_T^{Jets} < 90\ GeV$
- **Muon opening angle** $|sin\Delta\phi_{\mu\mu}| < 0.75$

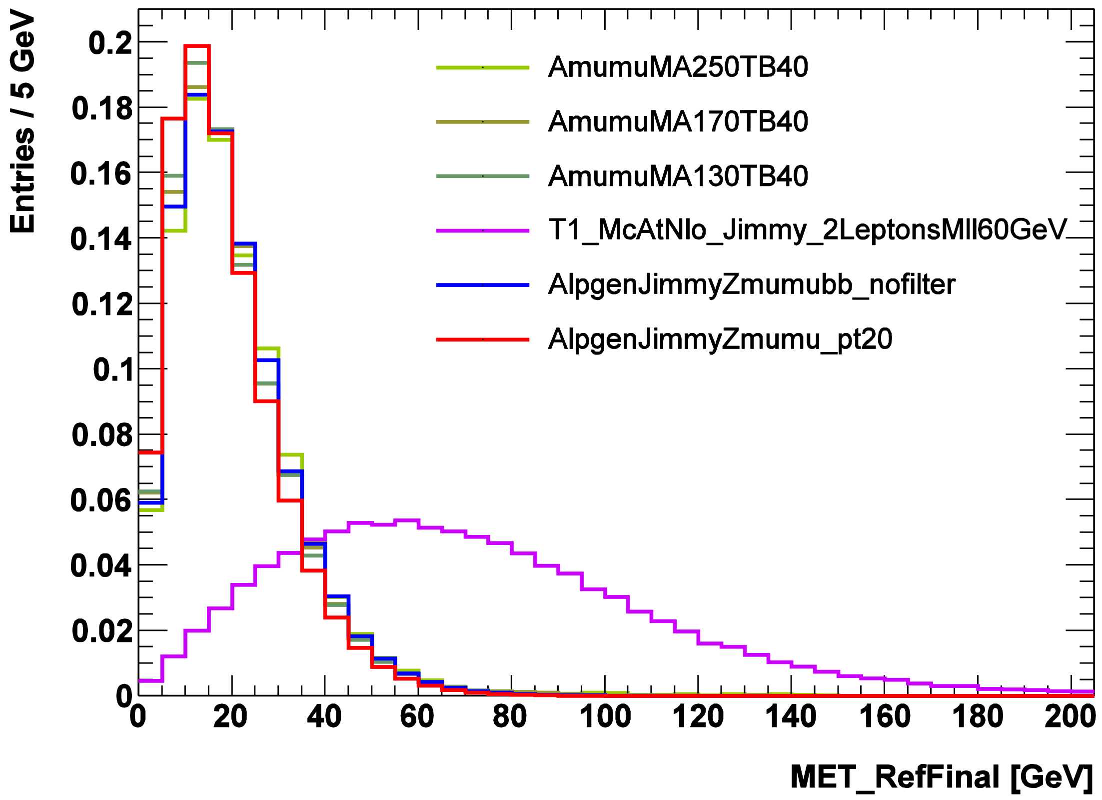
Selection without b-tagged jets:

- $N_{b-tagged\ Jets} = 0$

Cut variables

Missing transverse energy

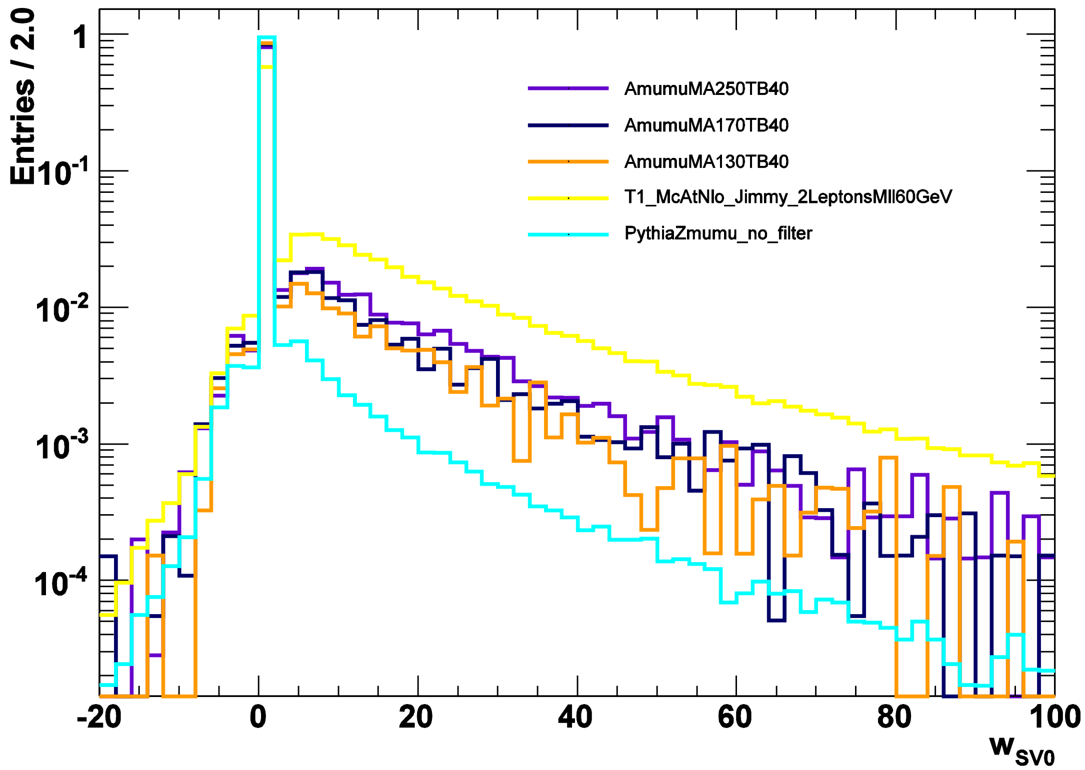
Normalized
to unity



Cut variables

b-tagging (SV0-weight)

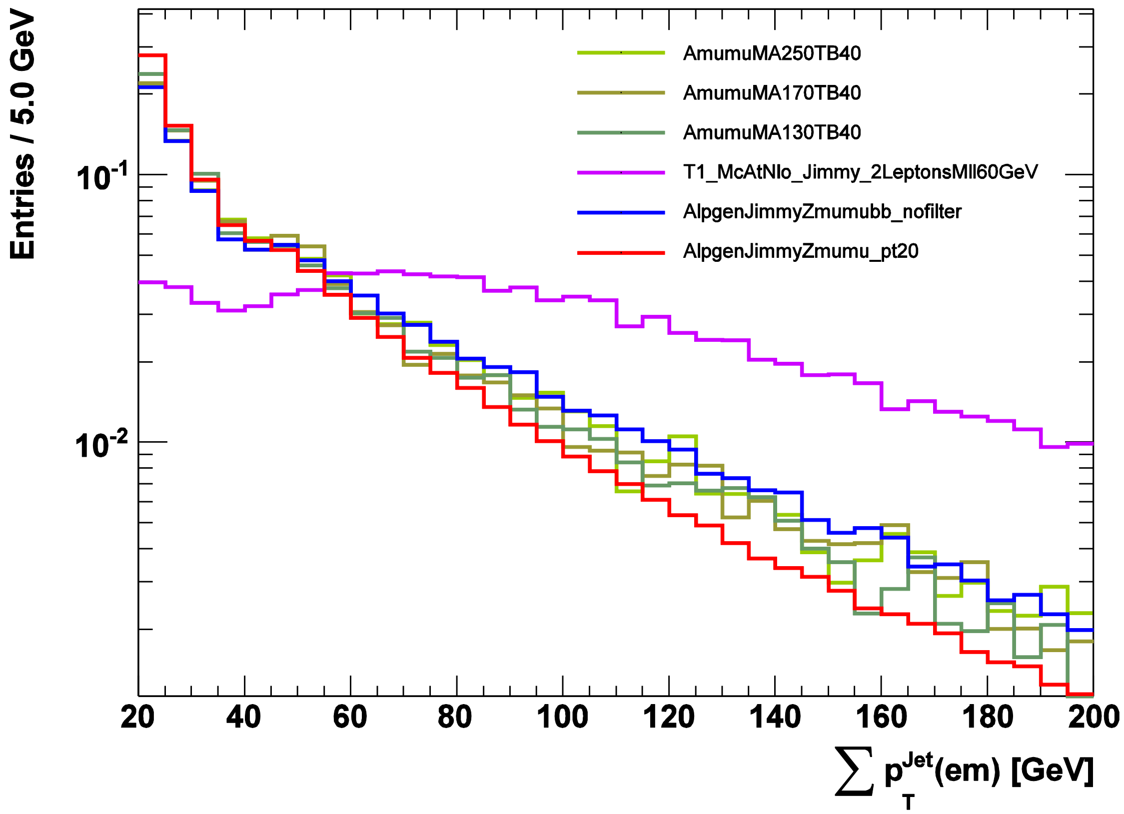
Normalized to unity



Cut variables

Jet activity

Normalized to unity

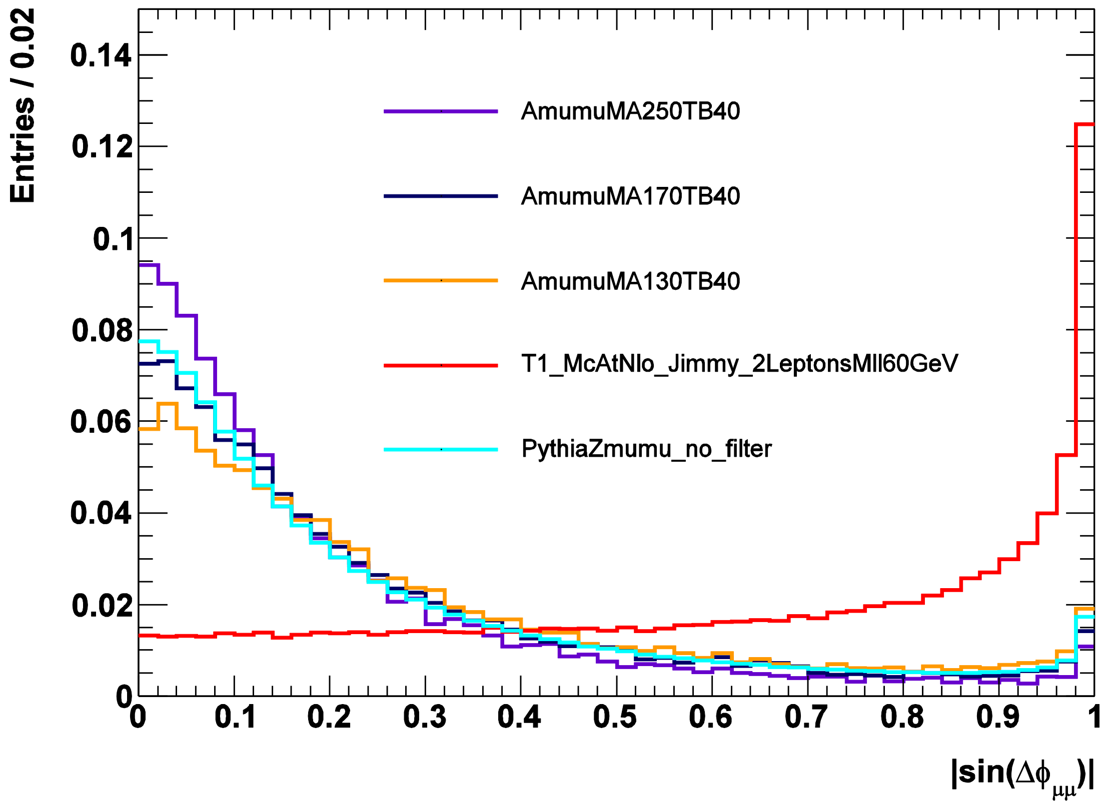


TECHNISCHE UNIVERSITÄT DRESDEN

Cut variables

Muon opening angle

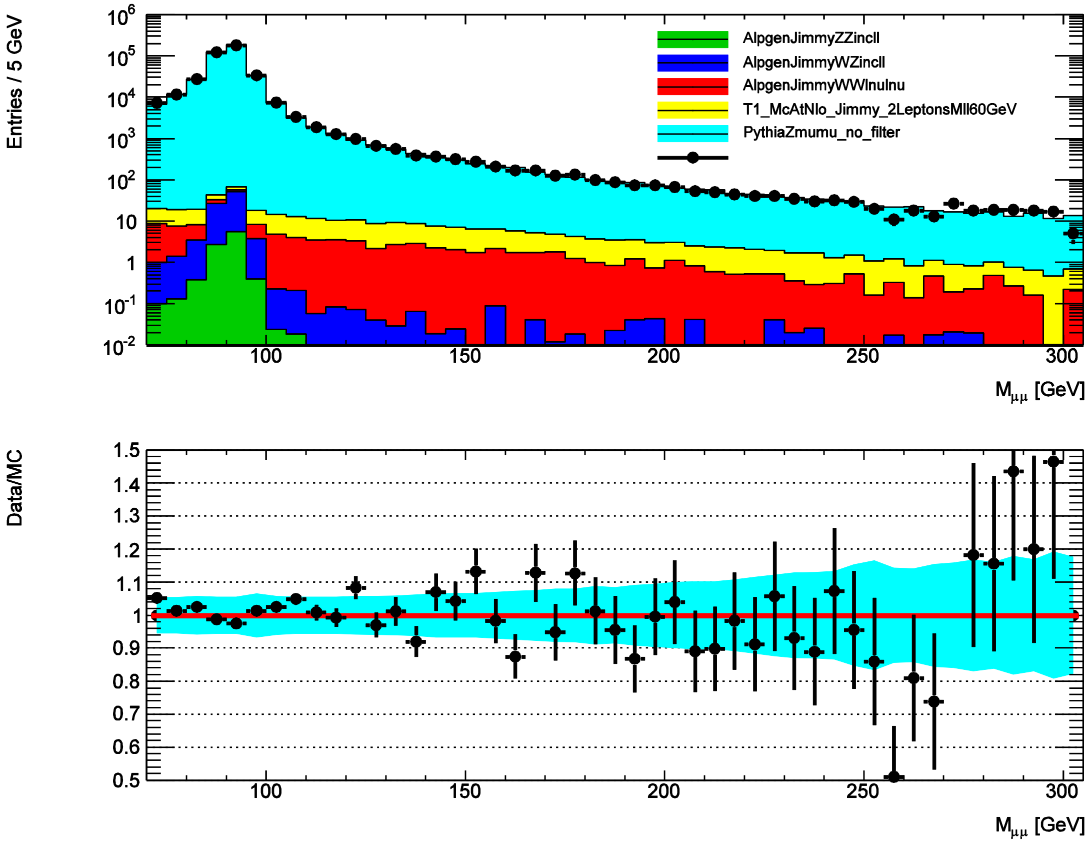
Normalized
to unity



Distribution of invariant mass

Inclusive selection

Normalized to
$\int \mathcal{L}\, dt \cong 1\, fb^{-1}$

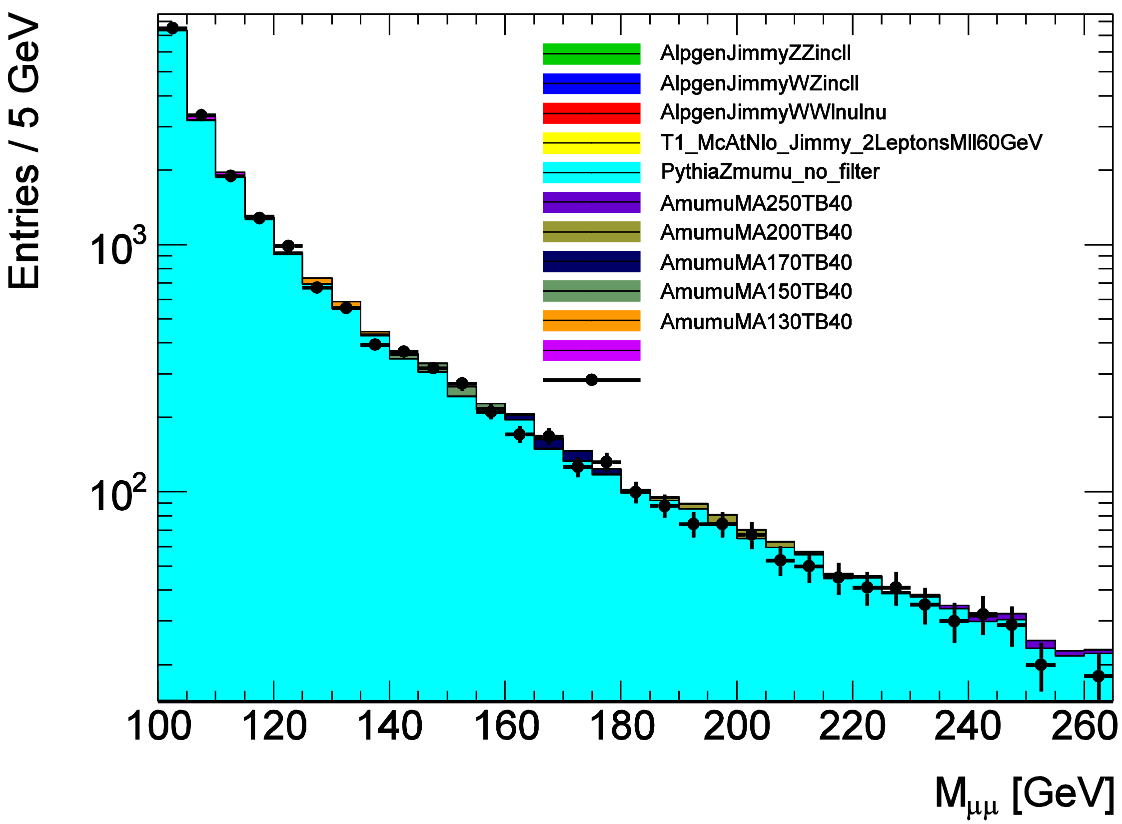


TECHNISCHE UNIVERSITÄT DRESDEN

Distribution of invariant mass

Inclusive selection

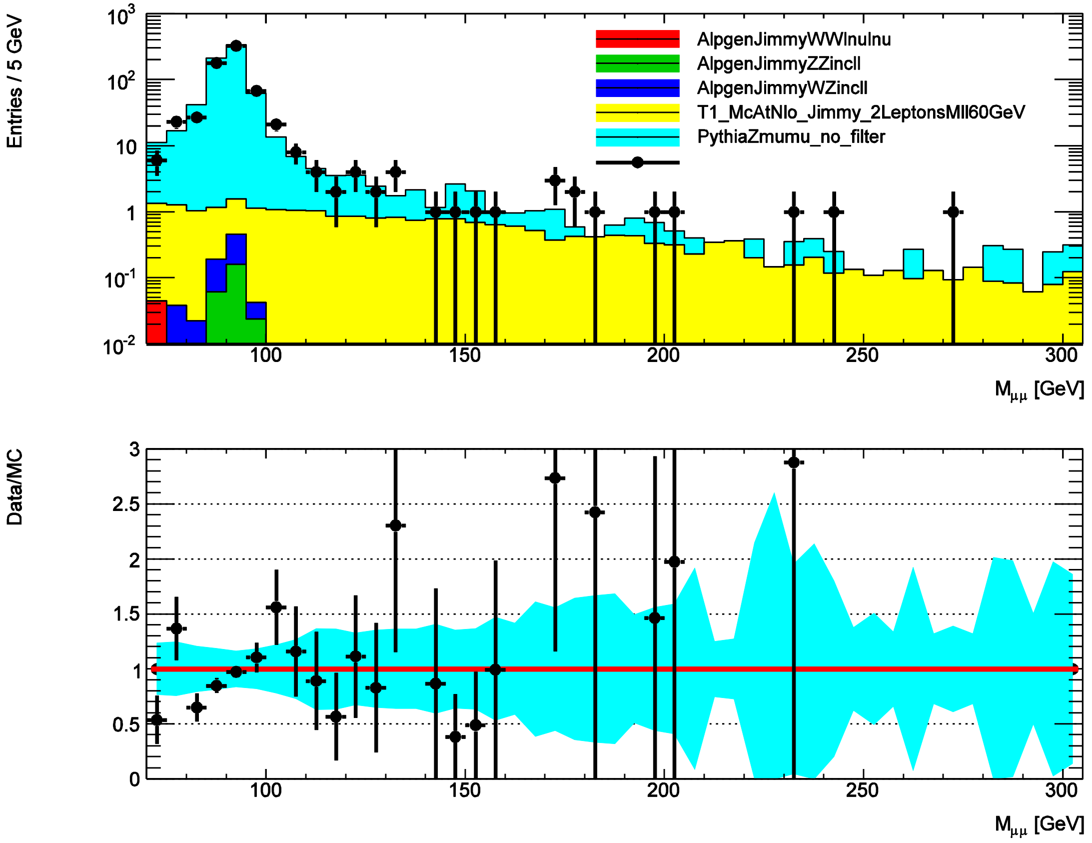
Normalized to
$\int \mathcal{L}\, dt \cong 1\, fb^{-1}$



Distribution of invariant mass

Selection with b-tagged jets

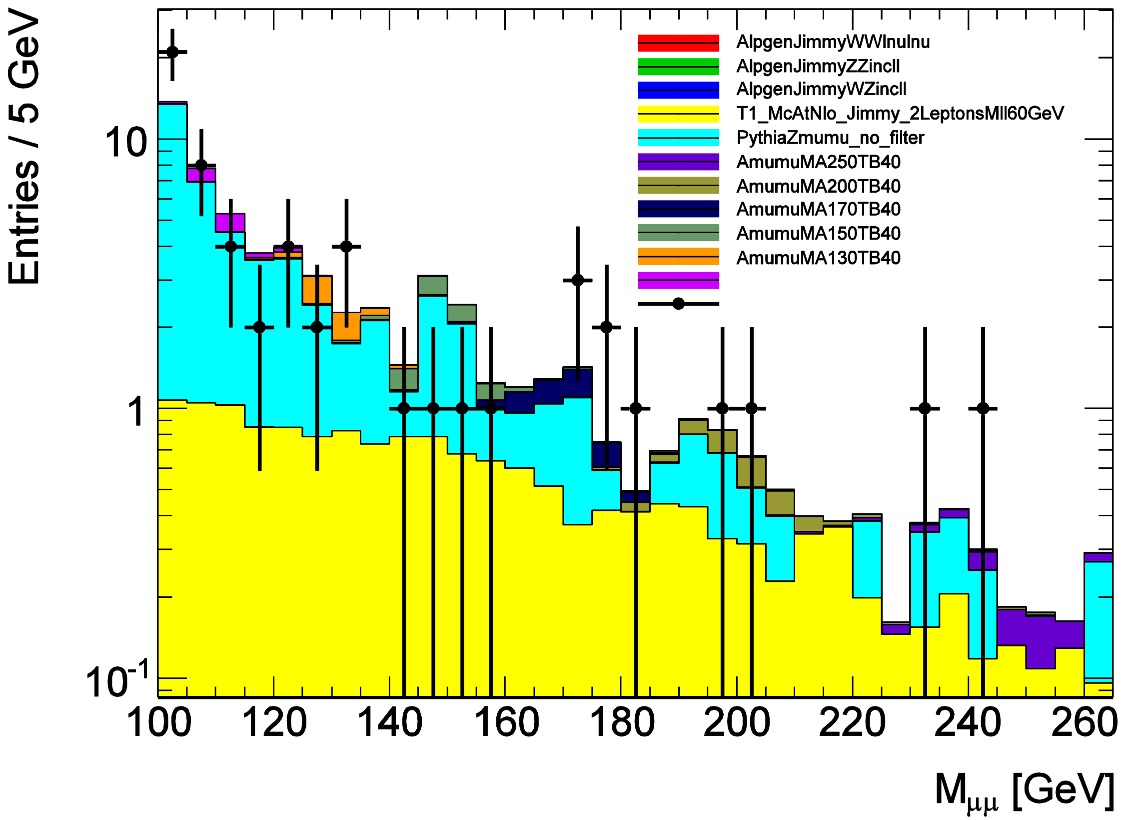
Normalized to
$\int \mathcal{L}\, dt \cong 1\, fb^{-1}$



Distribution of invariant mass

Selection with b-tagged jets

Normalized to
$\int \mathcal{L}\, dt \cong 1\, fb^{-1}$



Optimization of cut values

Optimization always means to find a compromise between two or more aspects:

- Programming: Speed versus size
- Smartphone: Screen size versus device size

…

- HEP Analysis: Signal efficiency versus signal purity

→ Criteria for a good compromise described by objective function

Optimization problem:

$\text{maximize}_x\, f(x)$ where $f(x): \mathbb{R}^n \rightarrow \mathbb{R}$ is the objective function (fitness, figure of merit)

In our Case: $f(x): \mathbb{R}^3 \rightarrow \mathbb{R}$ due to three cut variables

⇒Two decisions to be made:

- Selection of objective function
- Selection of optimization algorithm

Selection of objective function

Example: Measurement of signal cross section σ_S with well-known background B

Expected number of total observed events:

$$\langle N \rangle = \langle S \rangle + \langle B \rangle = \epsilon \cdot \int \mathcal{L} dt \cdot \sigma_S + \langle B \rangle,$$

with signal efficiency ϵ.

Measurement of cross section:

$$\sigma_{meas} = \frac{N - \langle B \rangle}{\epsilon \cdot \int \mathcal{L} dt}, \qquad \delta\sigma_{meas} = \frac{\sqrt{N}}{\epsilon \cdot \int \mathcal{L} dt}$$

Selection of objective function

Example: Measurement of signal cross section σ_S with well-known background B

Definition of purity:

$$\pi = \frac{\langle S \rangle}{\langle S \rangle + \langle B \rangle} = \frac{\langle S \rangle}{\langle N \rangle}$$

Relative error of measurement:

$$\frac{\langle \delta\sigma_{meas} \rangle}{\langle \sigma_{meas} \rangle} = \frac{1}{\sqrt{\epsilon \cdot \pi \cdot \int \mathcal{L}dt \cdot \sigma_S}}$$

Objective function:

$$\epsilon \cdot \pi, \qquad \frac{S}{\sqrt{S+B}}$$

Assumption: $\langle \delta B \rangle \approx 0$

Selection of objective function

Example: Search for a new signal

Objective function $\epsilon \cdot \pi$ not best choice:

- MC generator for bkg. does not describe tails of distributions perfect $\longrightarrow \langle \delta B \rangle \approx 0$ not valid
- π depends on signal cross section

New ansatz: Experiment should be sensitive to the smallest possible cross section: $\sigma_s \approx 0$

$\longrightarrow$ New purity:

$$\pi \approx \frac{\langle S \rangle}{\langle B \rangle}$$

$\longrightarrow$ New objective function: (all equivalent and independent from σ_s)

$$\epsilon \cdot \pi \approx \frac{\langle S \rangle^2}{\langle B \rangle}, \qquad \frac{\epsilon}{\sqrt{B}}, \qquad \frac{\epsilon}{\sqrt{\epsilon_{Bkg}}}, \qquad \frac{S}{\sqrt{B}}$$

Selection of objective function

Example: Different approach by G. Punzi

Ansatz (Hypothesis testing): Power function describes probability of discovery

$$1 - \beta_\alpha(m) > CL$$

Definition of region in parameter space (m) where experiment is sensitive to searched signal.

Result for a counting experiment:

$$\frac{\epsilon}{\frac{a}{2} + \sqrt{B}},$$

where a is the number of sigmas corresponding to one-sided Gaussian test at significance α.

Selection of objective function

Conclusion

- Objective function strongly depends on type of measurement
- No general objective function available
- Selection for cut optimization: $\frac{S}{\sqrt{B}}$ (due to search for new signal) and $\frac{\epsilon}{\frac{a}{2}+\sqrt{B}}$ as cross check
- Reading:
  - Benno List, “Why and When to Optimize Efficiency Times Purity”, 2002, https://www.desy.de/~blist/notes/whyeffpur.ps.gz
  - Glen Cowan, Eilam Gross, “Discovery significance with statistical uncertainty in the background estimate”, 2008, www.pp.rhul.ac.uk/~cowan/stat/notes/SigCalcNote.pdf
  - Giovanni Punzi, “Sensitivity of searches for new signals and its optimization”, 2003, http://arXiv.org/abs/physics/0308063v2

Selection of optimization algorithm

- Most optimization problems cannot be solved analytically
- Lots of different algorithms for different optimization problems available
- Easiest solution: Brute-force search
  - Find every possible set of cut values and calculate it's objective function
  - Independent from actual optimization problem
  - Discretisation of solution space necessary
  - Applicable for one or two cut variables
  - Too many possible solutions for three or more dimensional solution space (Combinatorial explosion)

  →More sophisticated algorithms needed

- One important group of algorithms: Metaheuristics

Selection of optimization algorithm

Metaheuristics

- Metaheuristics weakly depend on actual optimization problem
- For combinatorial optimization (e.g. travelling salesman problem) and real-valued search-spaces
- Metaheuristics do not guarantee to find the best solution
- Often used: stochastic components
- Selected type of algorithm for cut optimization: Genetic algorithm
    - Most popular type of evolutionary algorithms
    - Put back evolution of populations
    - Genetic operations (e.g. reproduction) contains statistical elements

Selection of optimization algorithm

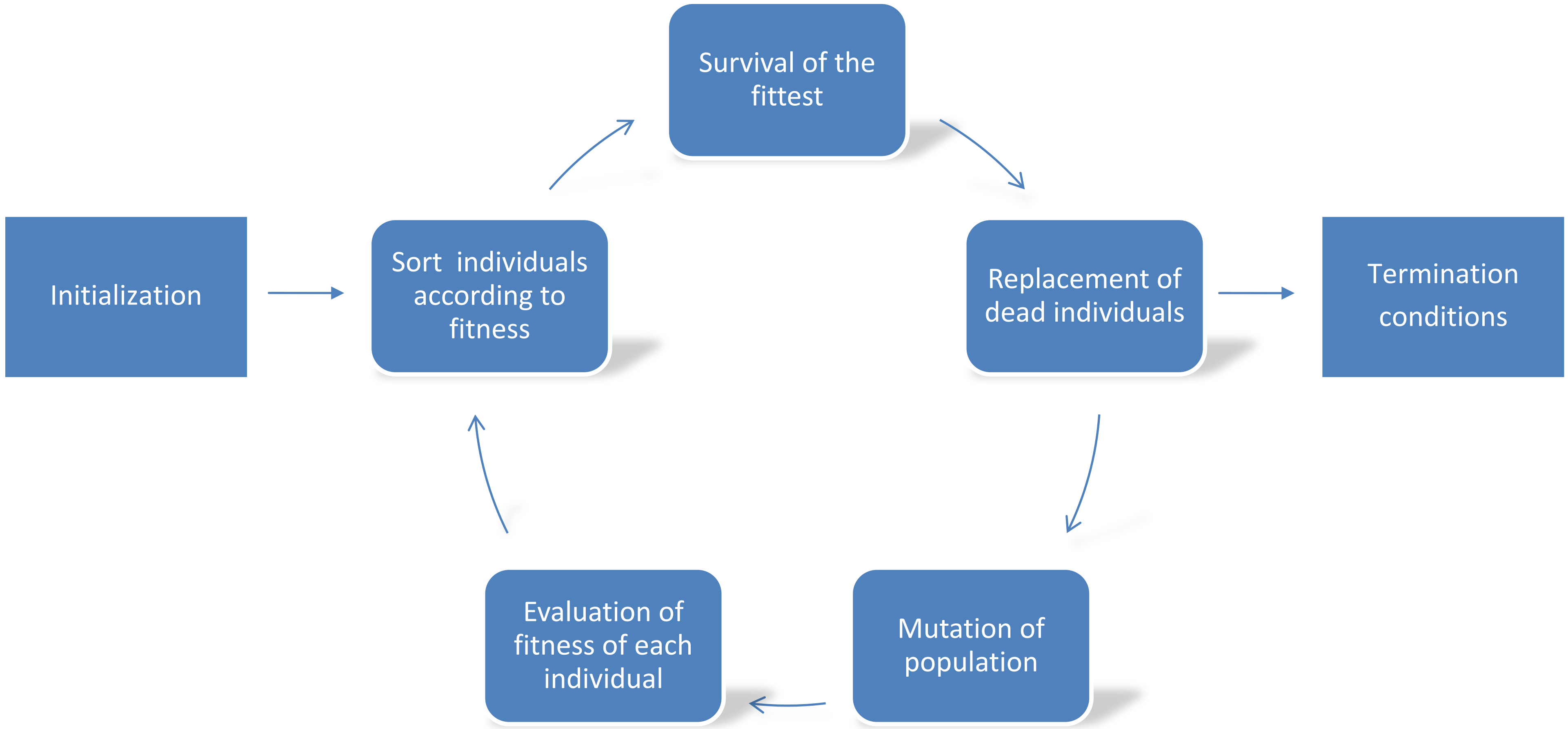
Genetic Algorithm

Terminology:

- Individual: Possible solution of the problem (one set of cut values)
- Genome: Complete description of a solution (actual cut values)
- Population: Set of individuals
- Fitness: Value of objective function for an individual,
  The higher the fitness of an individual the better the solution

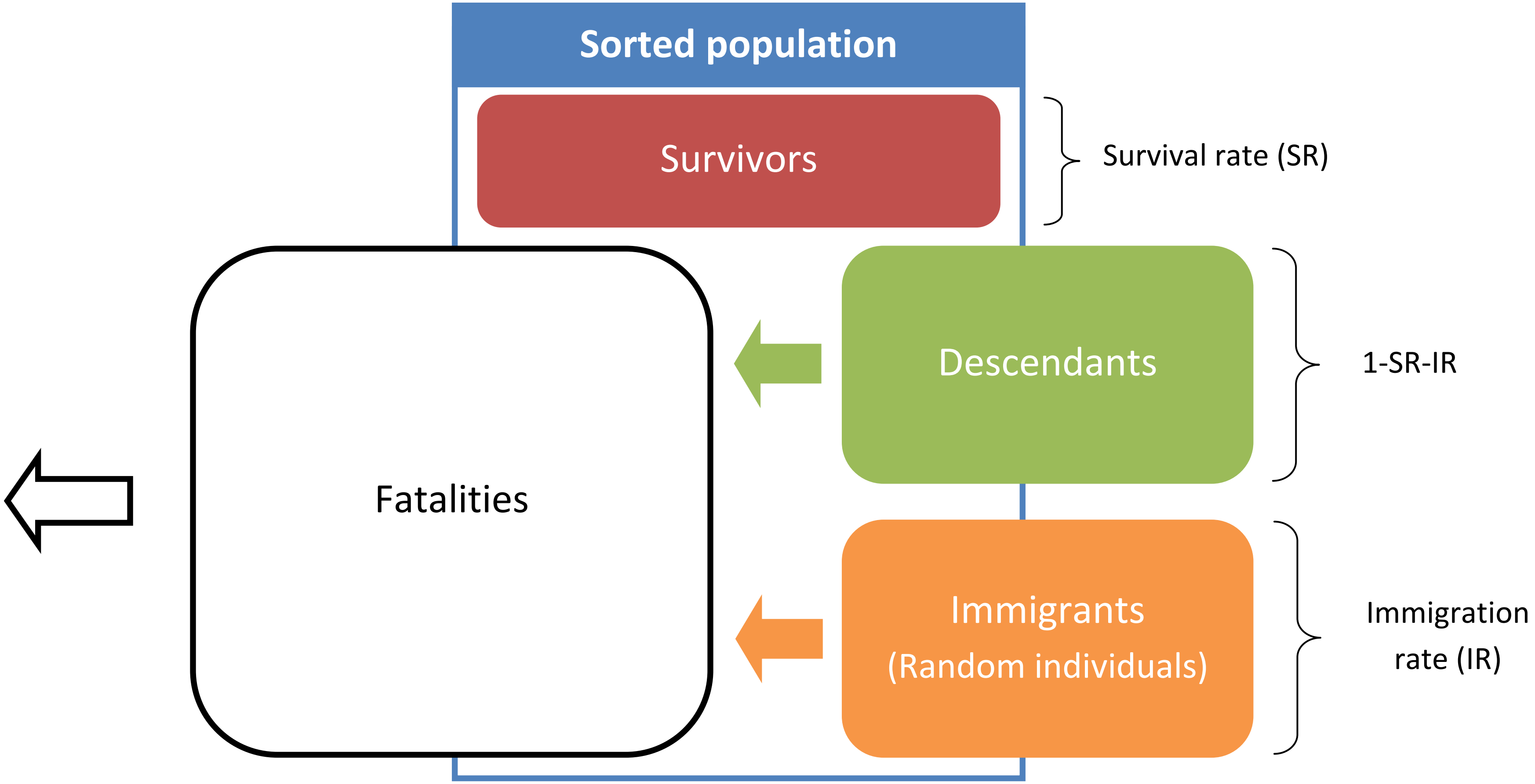
Selection of optimization algorithm

Genetic algorithm



Selection of optimization algorithm

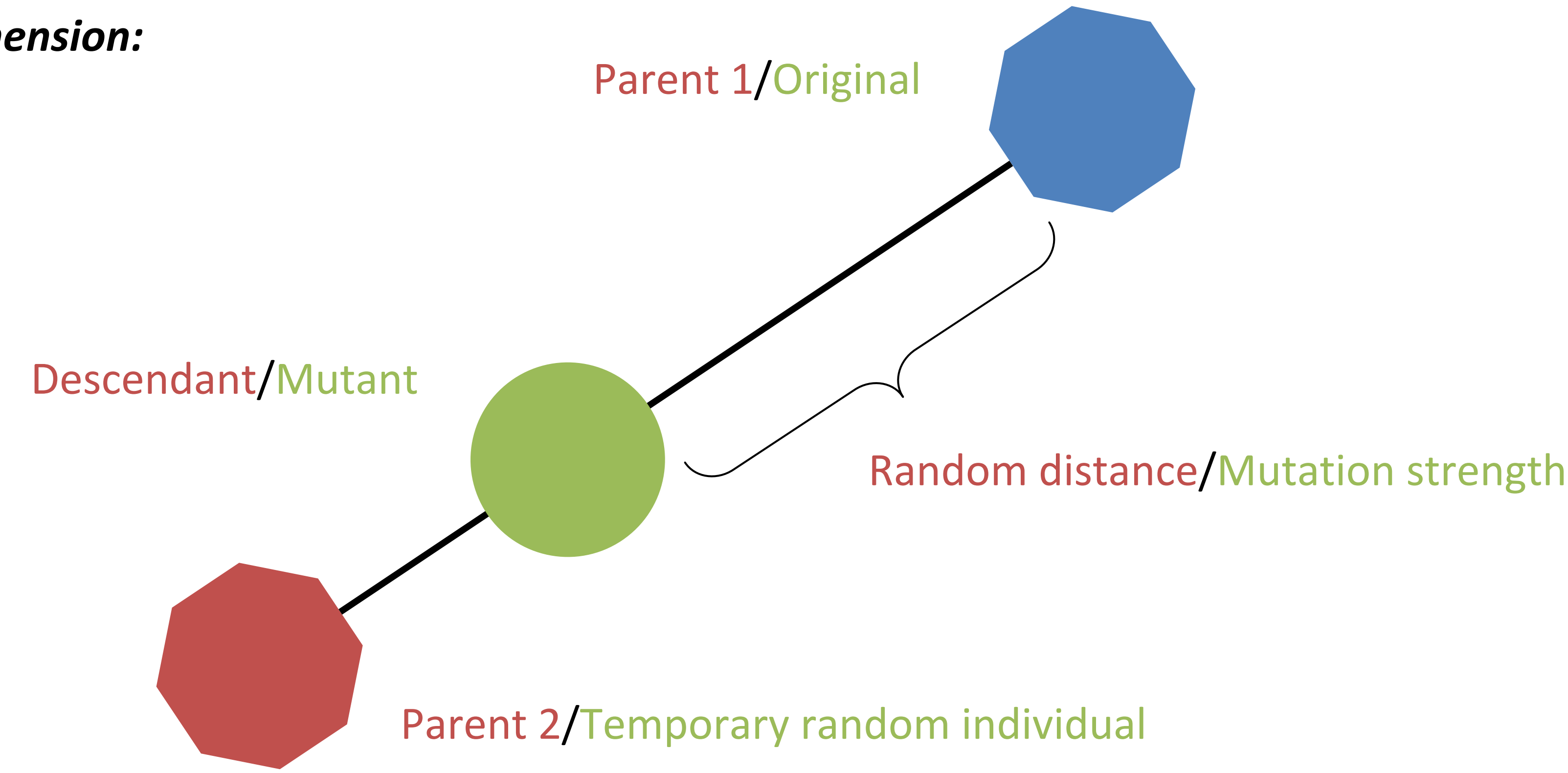
Genetic Algorithm: Replacement of dead Individuals



Selection of optimization algorithm

Genetic algorithm: Reproductions/Mutations

One Dimension:



Selection of optimization algorithm

Genetic Algorithm

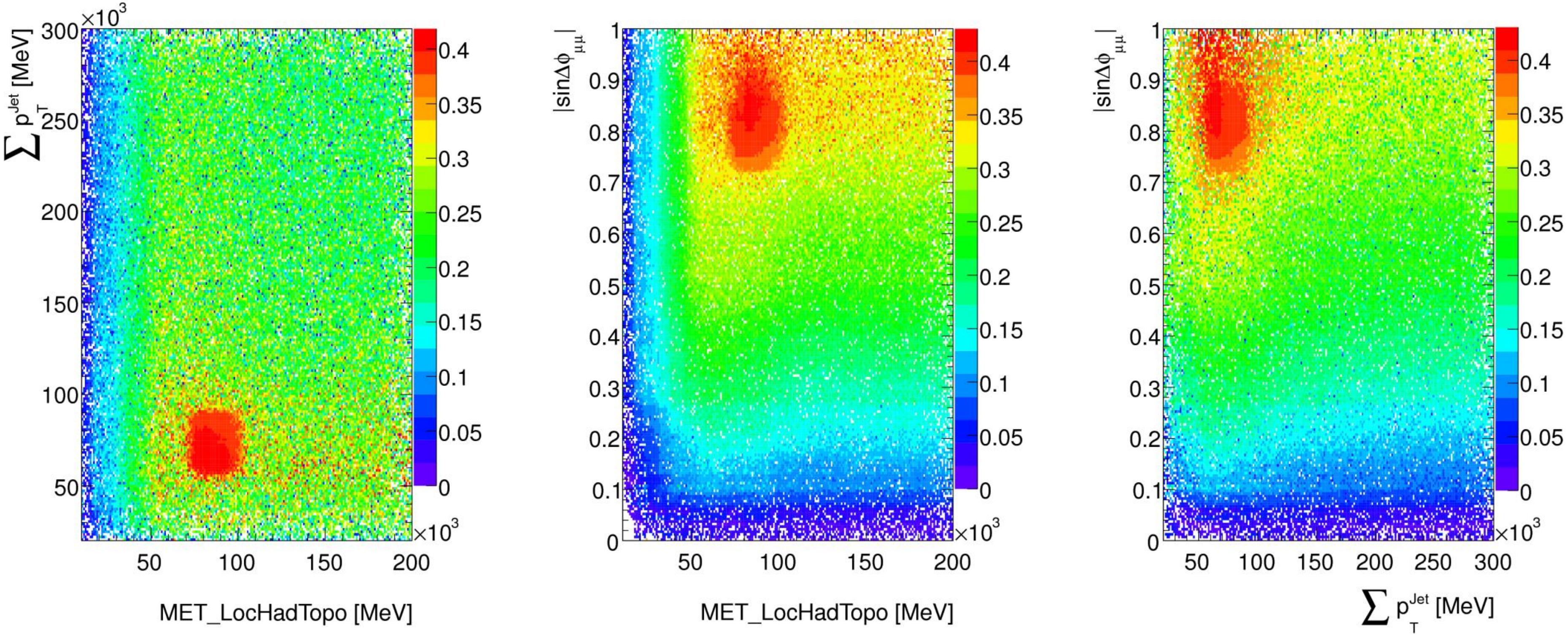
Genetic Algorithm requires several parameters:

• Population size:	500
• Number of generations (Termination condition)	500
• Survival rate:	20%
• Immigration rate:	40%
• Mutation rate:	70%
• Mutation strength:	10%

→Hard conditions

Optimization results

Projections of objective function $\frac{S}{\sqrt{B}}$ for selection with b-tagged jets, Higgs mass 150 GeV:

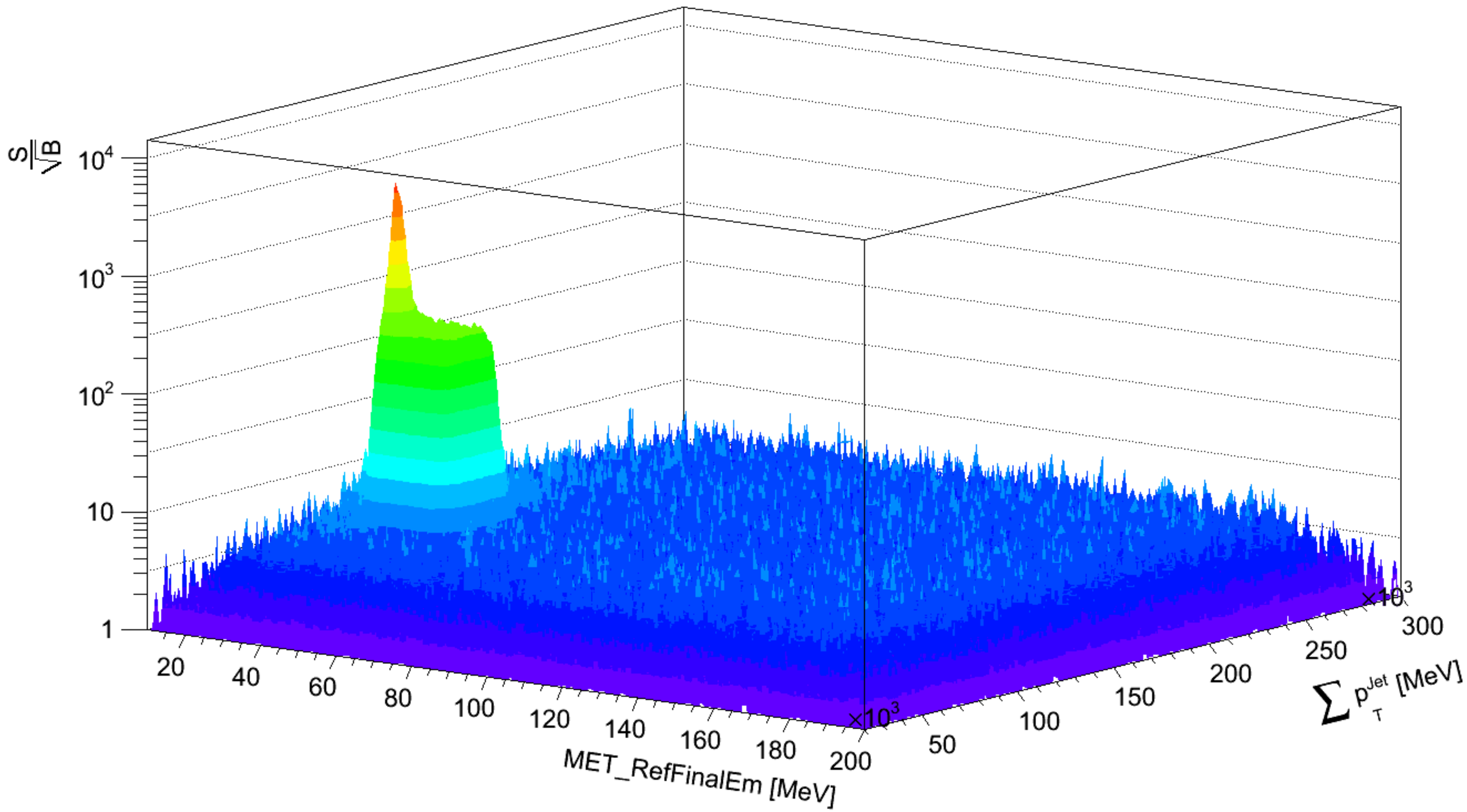


J. Philipp Grohs

TECHNISCHE UNIVERSITÄT DRESDEN

Optimization results

Sampling density:



Optimization results

New cut values

• Missing transverse energy	$< 40\ GeV$	$< 40\ GeV$
• Jet-activity	$< 70\ GeV$	$< 90\ GeV$
• Muon opening angle	< 0.80	< 0.75

→One third less $t\bar{t}$-background without reducing signal

Summary

- Search for neutral MSSM Higgs bosons decaying into two muons presented
- Optimization is a common problem

  $\longrightarrow$ Lots of solutions available
- Selection of objective function is crucial
- Working principles of genetic algorithm presented
- Resulting new cut values can reduce $t\bar{t}$-background by one third

Thank you for your attention!