Genetic Invariant Mass for Reconstruction of Resonance Decays

Won Sang Cho (IBS-CTPU)

Work in progress in collaboration with
Kiwoon Choi (IBS-CTPU),
Sung-Hak Lim (KAIST/IBS-CTPU),
Chan Beom Park (KIAS)

KIAS 4th Phenomenology Workshop
2014. 10. 27

1

Table of Contents

- Reconstruction of resonance decay
- Reconstruction via optimisation (event by event)
  - MT / MT2 / M2 by mass scale minimisation
  - MAOS
- What is Genetic Mass?
  - via collective stochastic optimisation
  - using Genetic Algorithm
  - utilising unused Pz of visible particle
  - evolution of solutions so that the evolved distribution best fit into a physical p.d.f, trying to inherit the boost invariance of a mass distribution maximally.
- Example of MW reconstruction

Reconstruction of Resonance Decay in HEP

- **Given a hypothetical RESONANCE decay topology, problems of unknown kinematic information to be reconstructed in HEP events :**

  - **Combinatorial problem**
    - Many many visible objects to be tagged at correct decay vertices, which cannot be easily determined by simple kinematical cuts.

  - **Missing momenta from invisible particles**
    - Neutrinos, Dark matter candidates
    - Basically resolved with kinematic constraints, given in an event / combined events.

Reconstruction of Resonance Decay in HEP

- **Given a hypothetical RESONANCE decay topology, problems of unknown kinematic information to be reconstructed in HEP events :**

  - **Combinatorial problem**
    - Ma_____which
      ca_____

  - **Missi**
    - Ne
    - Ba_____combined
      ev

Momentum Reconstruction via Optimisation

- **Some mass variables constructed by minimising the involved resonance mass scale :**

$$Resonances(P) \rightarrow Visibles(p) + Invisibles(q)$$

- **Transverse Mass (M_T) for single resonance decay**

$$M_T^2 = \min_{q_x, q_y, q_z} (p + q)^2 \quad s.t. \quad p_T + q_T = 0$$

- **Stransverse Mass (M_T2) for double resonance decays**

$$M_{T2}^2 = \min_{q1, q2} \left\{ \max \left[(p_1 + q_1)^2, (p_2 + q_2)^2 \right] \right\} \quad s.t. \quad p_{1T} + p_{2T} + q_{1T} + q_{2T} = 0$$

- **Constrained-M_2/M_N with general kinematic constraints :**
  - A.Barr, T.Khoo, P. Konar, K.Kong, C.Lester, K.Matchev, M.Park [Phys.Rev. D84 (2011) 095031]
  - W. Cho, J. Gainer, D. Kim, K. Matchev, F. Moortgat, L. Pape and M. Park [JHEP 1408 (2014) 070]

$$\bar{M}^2 = \min_{q \in \mathcal{R}^n} \tilde{M}^2(p, q) \quad s.t. \quad c_{i=1..m}(p, q) = 0$$

- via the optimisation(minimisation in mass scale) process, unknown momentum d.o.f are fixed by some values.

  ex) $Minimisation\ for\ M_T \rightarrow \eta_\nu = \eta_l$

- studied as 'MAOS (M_T2-Assisted On-Shell) momenta' for spin discrimination / mass peak reconstruction

  - W.S. Cho, K. Choi, Y.G.Kim, C.B.Park [Phys.Rev. D79 (2009) 031701]
  - C.B. Park et al.[JHEP 1401 (2014) 030], [Phys.Rev. D84 (2011) 096001] [Phys.Rev. D82 (2010) 113017]

Need New Approach for better mass variable ??

1. no more kinematic constraints (in single event)

2. scale-minimised (transverse) mass variables are unstable under the boost profile of mother resonance : Main source of systematic uncertainties in W-mass measurement

3. Its momentum solution is meaningful only for near-endpoint events.

4. While transversification, longitudinal information of visible particles are often ignored.

- Then, we try to optimise the unknown d.o.f from whole event sample collectively, in search for the solution set whose invariant distribution most likely resembles the shape of essential Breit-Wigner distribution.

- Optimisation technology = Genetic Algorithm

- Hoping to get an engineered solution/mass distribution, which evolves and approximately/ maximally inherit the boost invariant property :
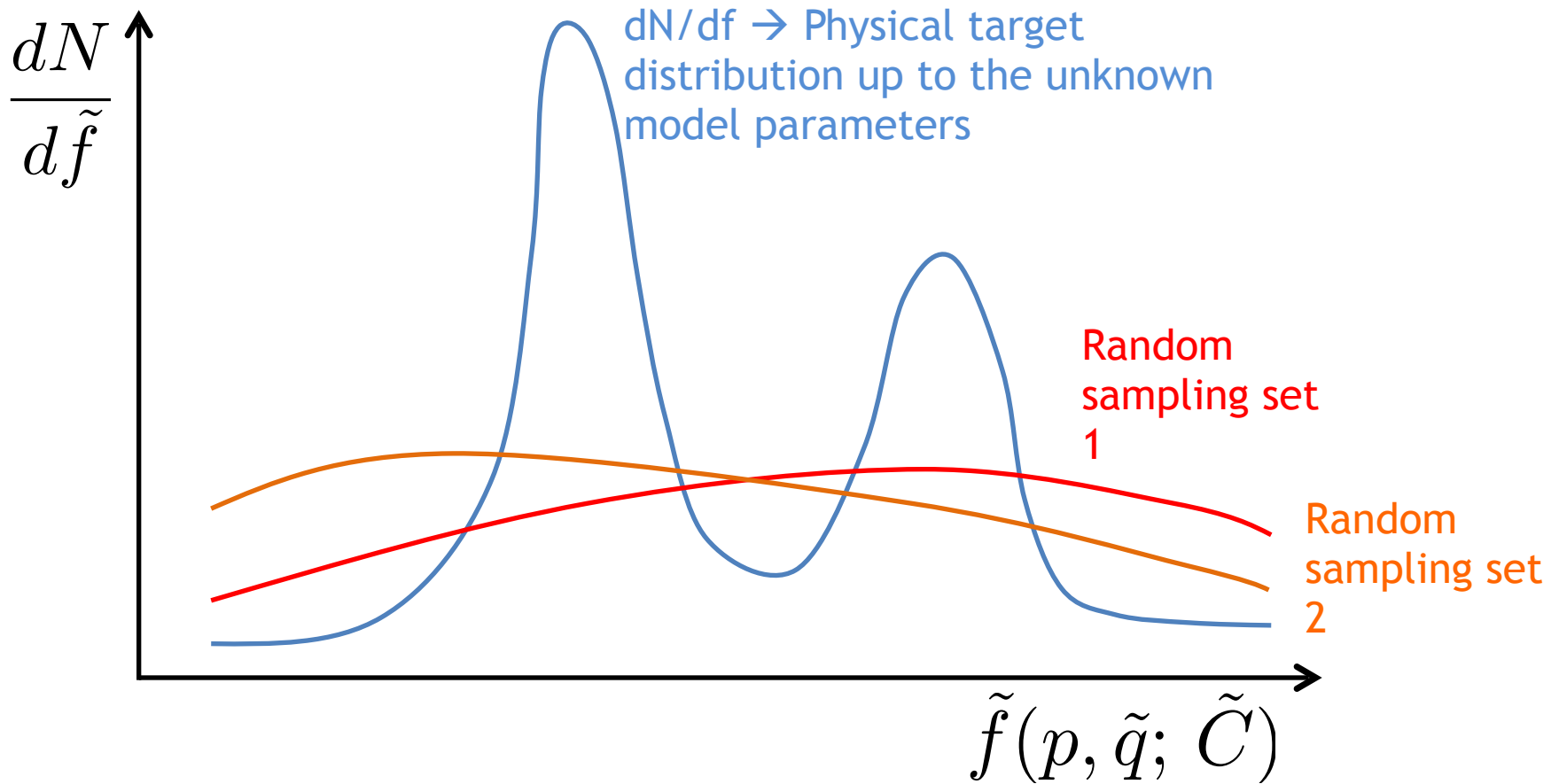
  => Genetic Mass

Collective Stochastic Optimisation for Genetic Mass

Event ID	1	2	...	N
Known	$\{p^1_{vis}\}$	$\{p^2_{vis}\}$	...	$\{p^N_{vis}\}$
Unknown	$\{q^1_{inv}\}, C^1$	$\{q^2_{inv}\}, C^2$	...	$\{q^N_{inv}\}, C^N$
Observable	$f(\{p^1_{vis}\}, \{q^1_{inv}\}; C^1)$	$f(\{p^2_{vis}\}, \{q^2_{inv}\}; C^2)$	...	$f(\{p^N_{vis}\}, \{q^N_{inv}\}; C^N)$

$\{p^i_{vis}\}$: set of visible momenta of the event-(i)

$\{q^i_{inv}\}$: set of unknown momenta of the event-(i)

$C^i \in \{1, ..., N_{combi}\}$: Particle assignment ID

$$\frac{dN}{d\tilde{f}}$$

dN/df → Physical target distribution up to the unknown model parameters

Random sampling set 1

Random sampling set 2

$$\tilde{f}(p, \tilde{q};\ \tilde{C})$$

Dimension of the random search space :

$$\{(\{\tilde{q}^i\},\ \tilde{C}^i)\,|\,i = 1..N_{event}\}$$

$$N_{par} = N_{event} \times (N_q + N_{combi})$$

Basic procedures for one random solution of the unknowns

1. Consider a random sampling of the unknowns for all event blocks.

$$\{(\{\tilde{q}^i\},\, \tilde{C}^i) \,|\, i = 1..N_{event}\}$$

2. Select a good collider observable (non-reconstructable as q & C are unknowns) :

$$f(p, q;\, C)$$

3. Calculate a sampled functional value using the sampled unknown values :

$$\tilde{f}(p, \tilde{q};\, \tilde{C})$$

4. Estimate the fitness as the distance between the sampled distribution and the physical target distribution where unknown model parameters are optimised to minimise distance between two distributions

$$Fitness = Distance\,between\,\frac{dN}{df}(\vec{\alpha}_0 = (m_0, \Gamma_0, ...))\,and\,\frac{dN}{d\tilde{f}}$$

The Genetic Algorithm does operations, generation by generation..

1. (Initially in the 1st generation) Consider a population of solutions (= a set of individual random solutions)

2. Estimate the fitness of all solutions

3. Sort the individuals by their ranking of the fitness

4. Select a portion of high ranked solutions, and discard others

5. Using reproduction algorithm(crossover/mutation of genes) of GA, it produces children solutions, filling the missing population.

6. Estimate the fitness of new individual solutions, and sort them again by their ranking.

7. If not converged, goto #5 again as the next generation.

- Then, the best solution in each generation-(k), fitted into the physical target distribution which is optimized w.r.t unknown model parameters, evolves toward the true distribution,
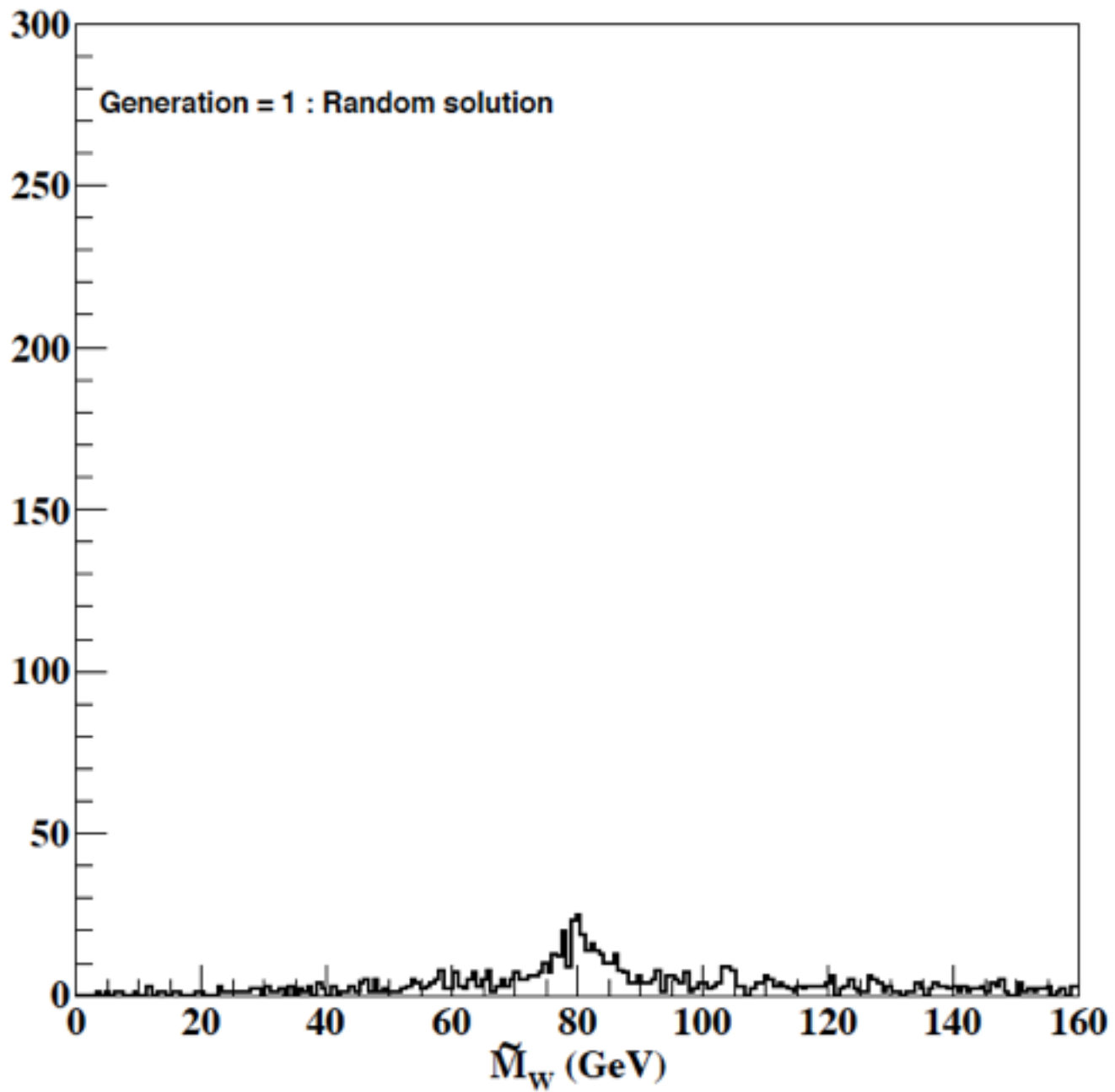
$$\lim_{k \to \infty} \frac{dN}{df}(\vec{\alpha}_k) = \frac{dN}{df}(\alpha_{true}) \sim \frac{dN}{d\tilde{f}_k}$$
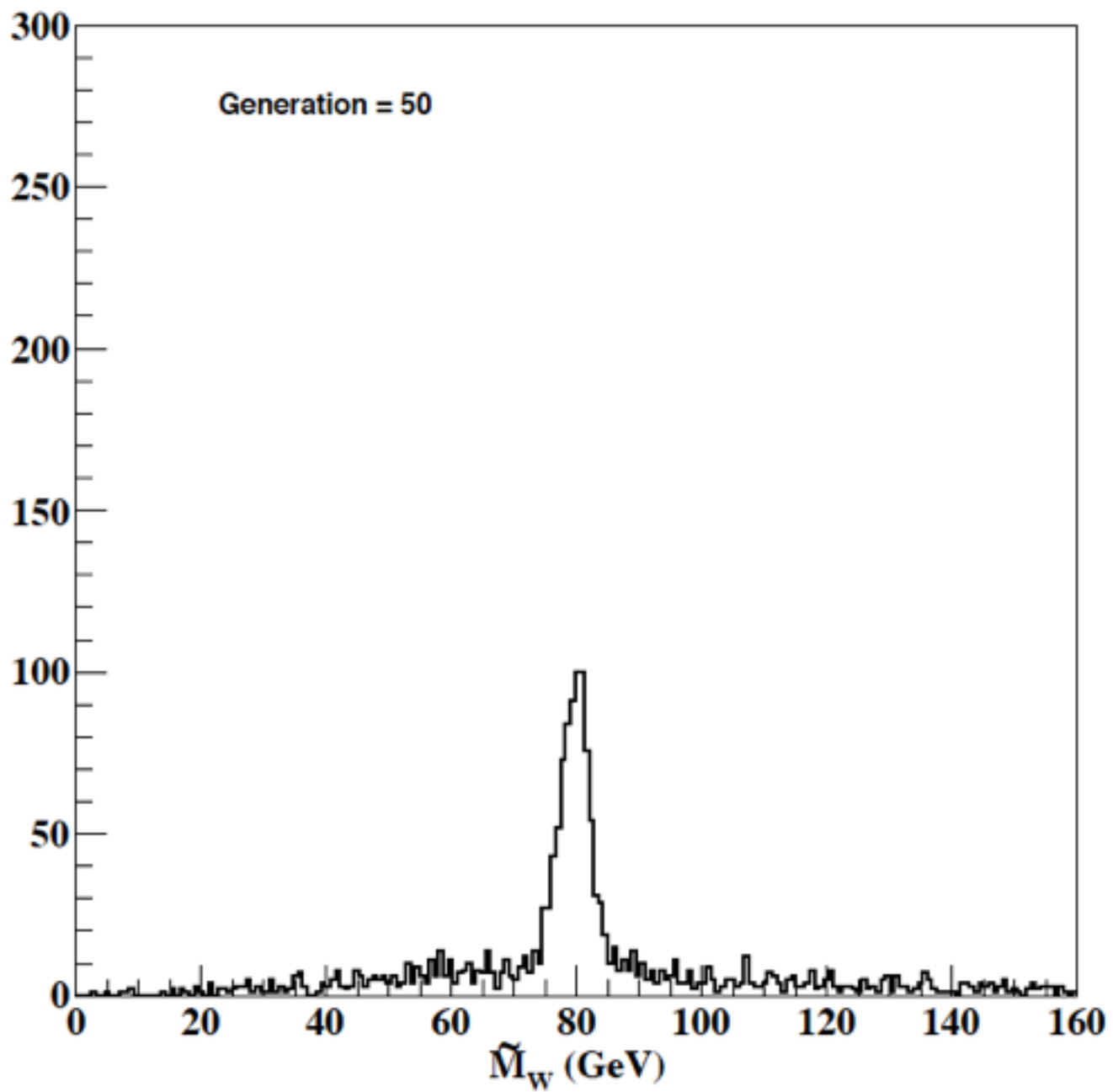
if the shape of our physical distribution possesses enough non-trivial physical information.
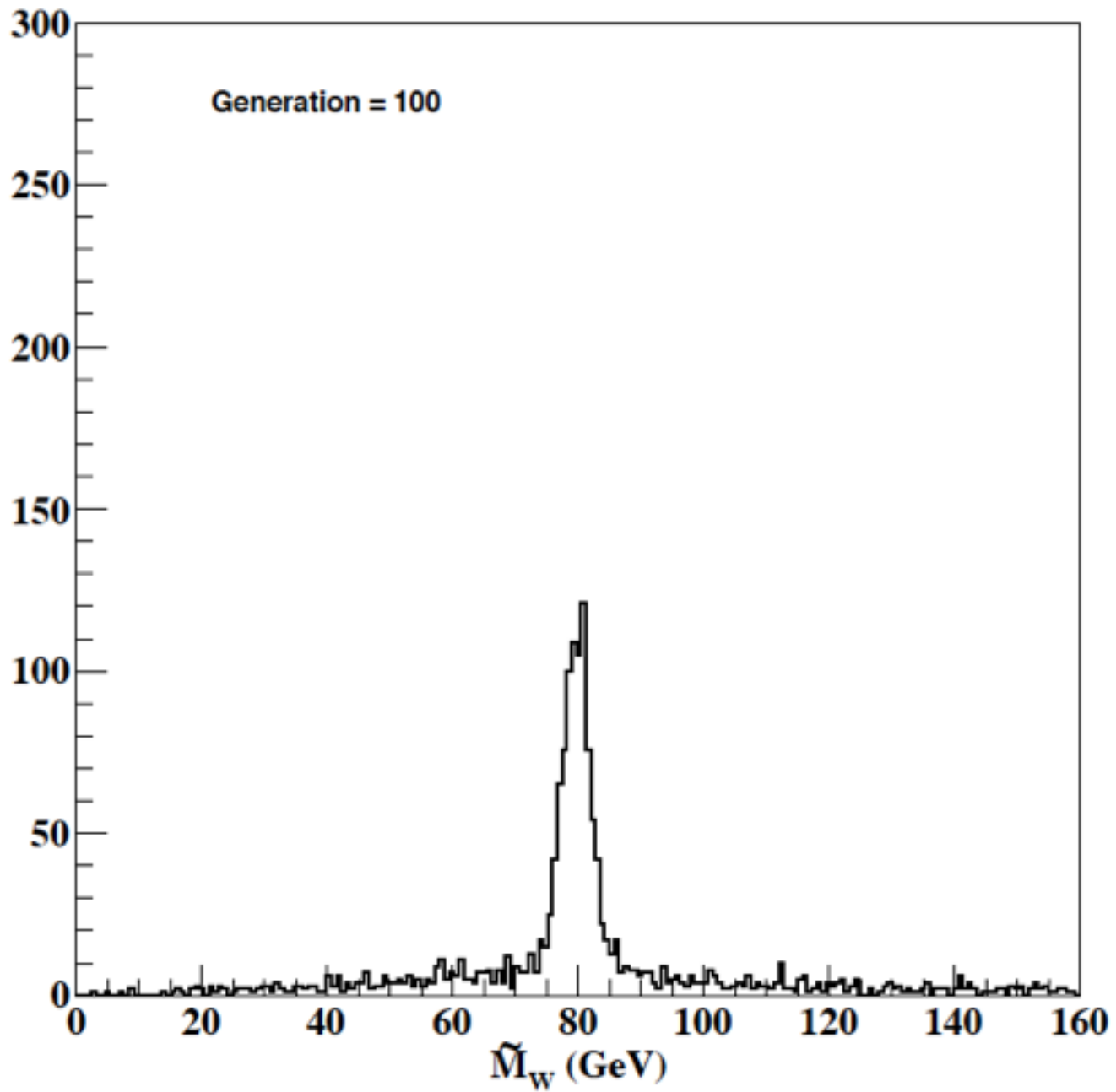
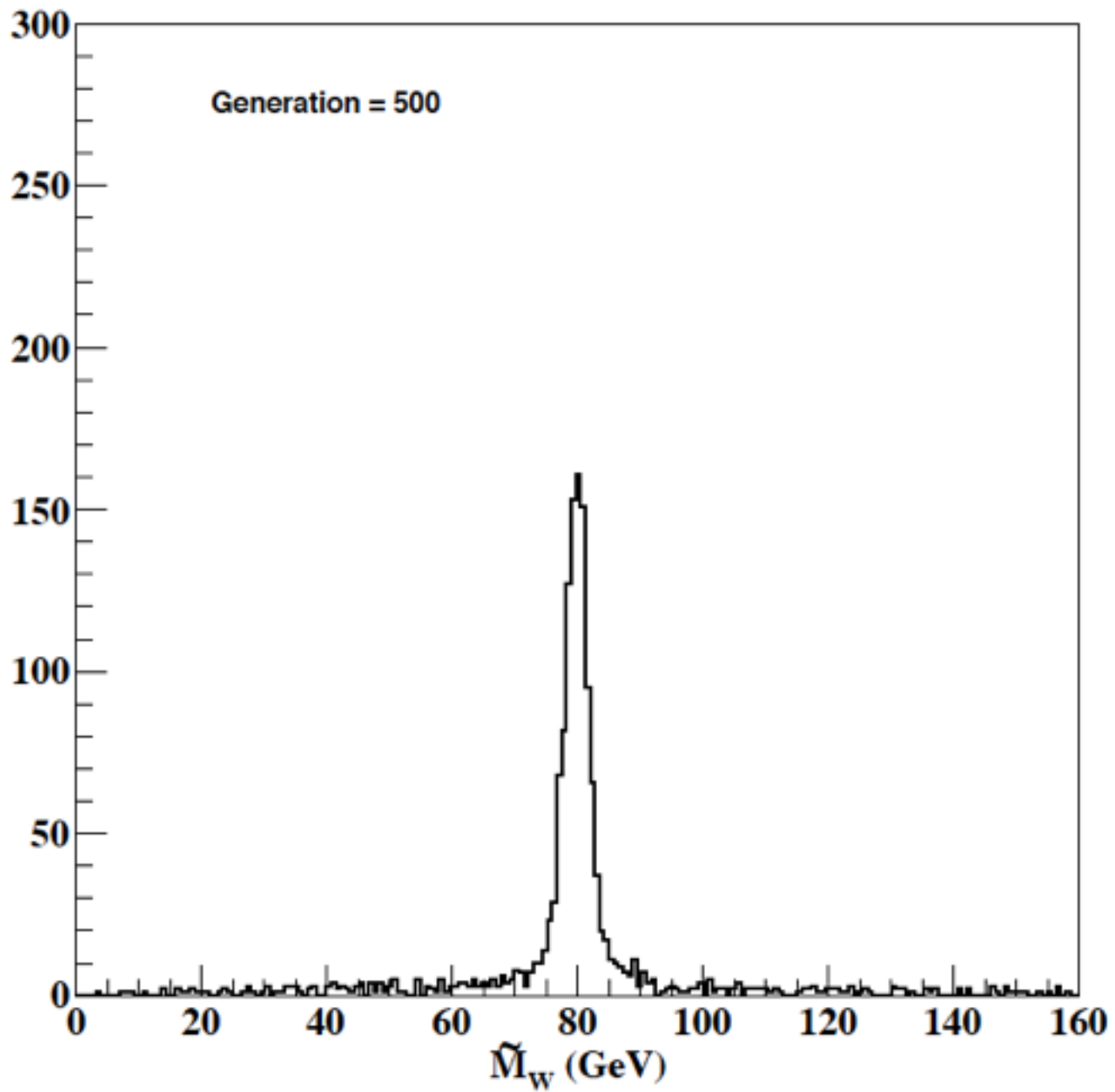- We apply this idea on the measurement of W boson mass

Reconstruction of the Breit-Wigner resonance peak of the W boson using genetic algorithm
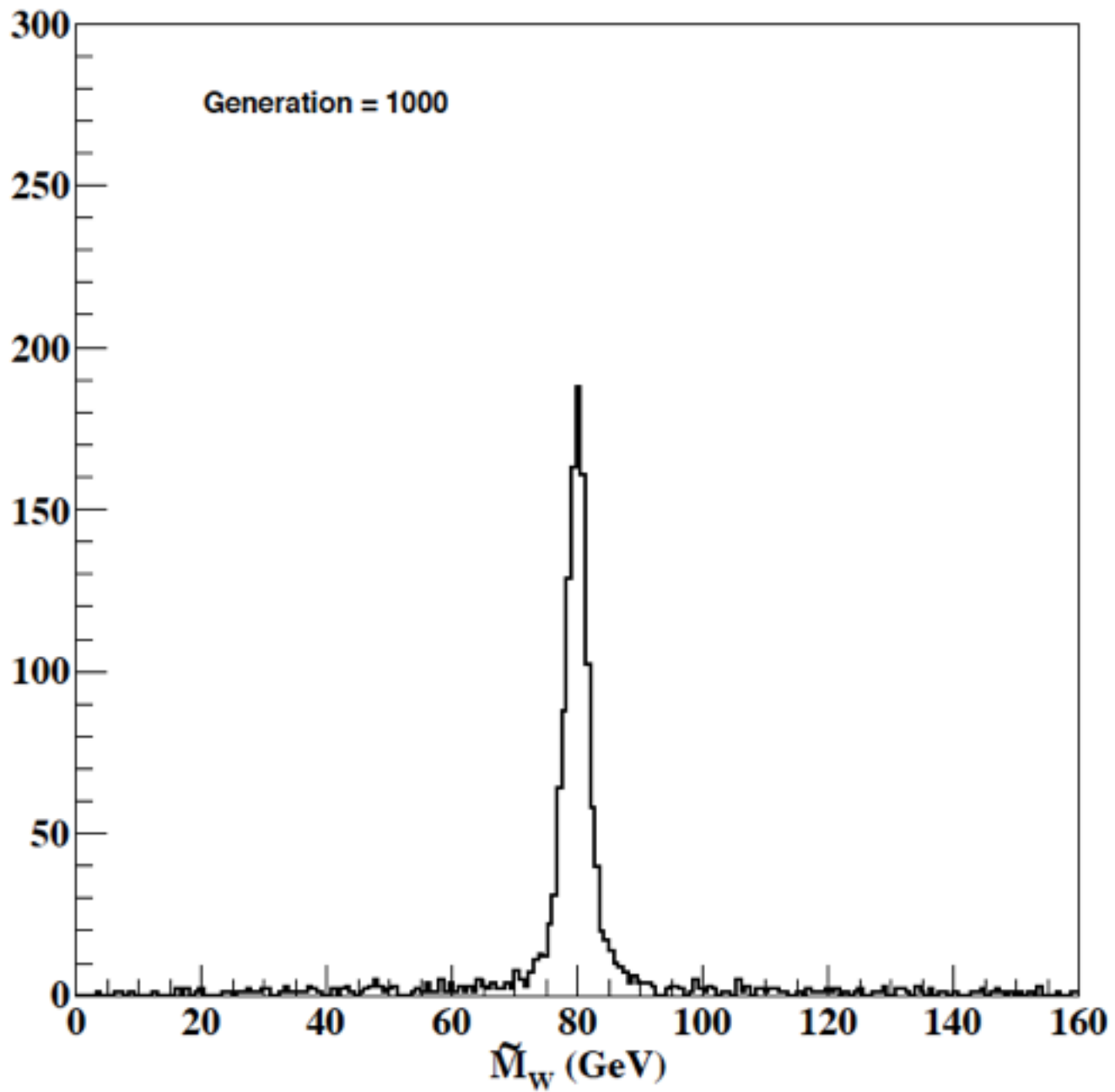
- Unknown : Pz or rapidity of neutrino in each event

- Encoding representation for a gene : real continuos value

- Chromosome : a set of qz, {qz} for whole event set

- Population : randomly generated chromosomes

- Fitness (of a chromosome = a solution set) : calculated by various probability measures like as chi2, log-likelihoods, relative entropy ..

- Using the (Gaussian convoluted) BW shape as a physical target distribution.
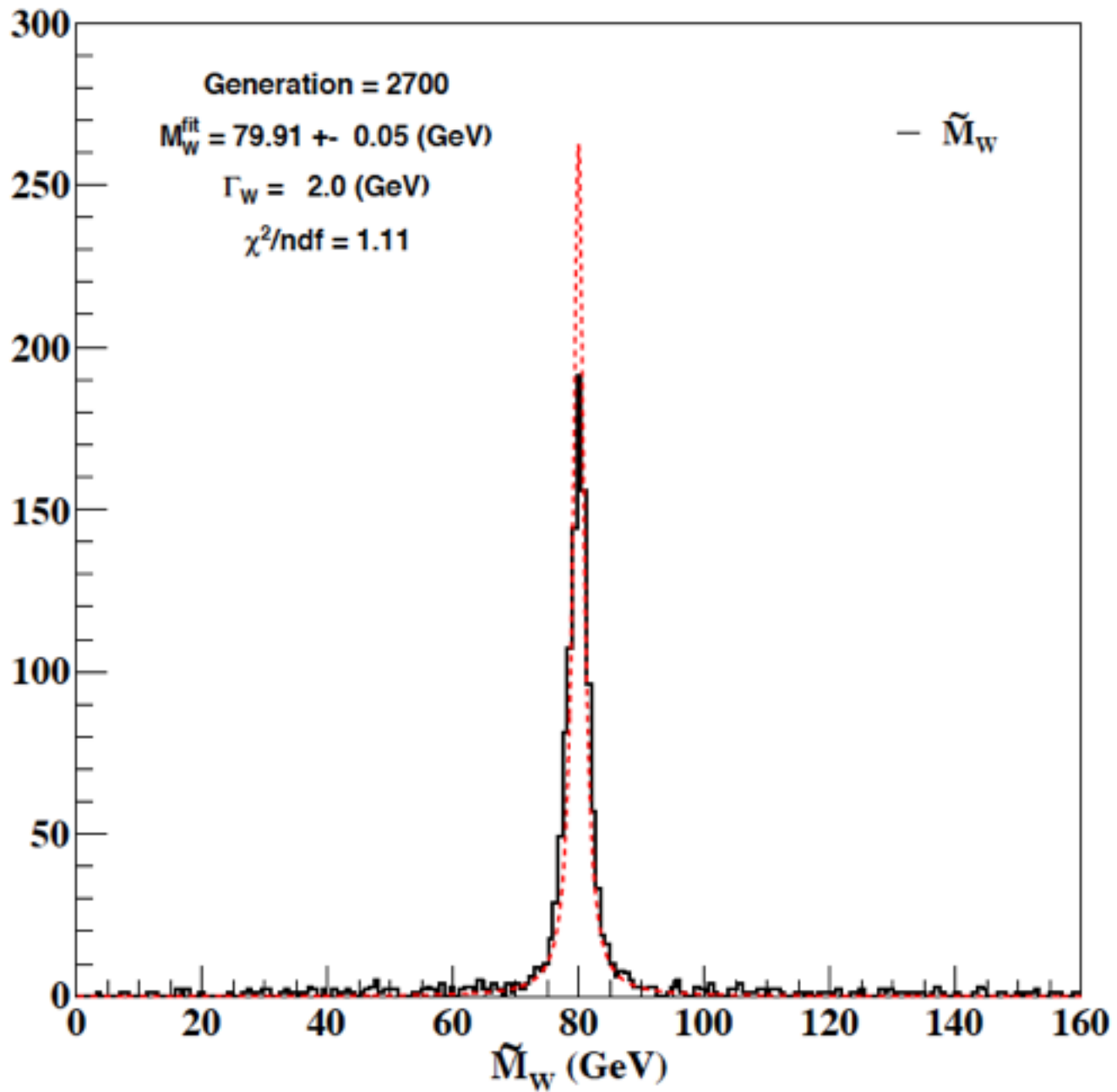
Generation = 500

$\tilde{M}_W$ (GeV)

Generation = 1000

$\tilde{M}_W$ (GeV)

Generation = 2700

$M_W^{fit} = 79.91 +- 0.05$ (GeV)

$\Gamma_W = 2.0$ (GeV)

$\chi^2/ndf = 1.11$

$- \quad \tilde{M}_W$

$\tilde{M}_W$ (GeV)

Conclusions

- We are engineering to produce the Genetic Mass which resembles the properties of BW resonance.

- Genetic Algorithm, based on the principe of Natural Selection, is found to be very efficient and powerful for implementing the collective stochastic optimisation.

- Collective stochastic optimisations in light of stable physical target distribution can provide a good chance to solve complicated event reconstruction problems

Backup

Genetic Algorithm

- Genetic algorithm is a search and optimisation technique based on Darwin's Principle of Natural Selection.

- "Select the best individuals with good fitness, and Exchange good genes, and Select .."

- Optimises a large number of (continuous/discrete) parameters with extremely complex objective function. It can rather easily jump out of a local minimum, in compared to the usual gradient-based search algorithms.

- Does not require derivative information.

- May encode the parameters so that the optimisation is done with the encoded parameters

Simple GA works with the binary encoding for each variable.

$$gene^i \quad = \quad \text{an encoded unknown value } (q^i \text{ or } C^i)$$
$$\text{of an event block}$$

$$
\begin{array}{cccc}
Gene^1 & Gene^2 & .... & Gene^{N_{event}}
\end{array}
$$

$$chromosome \quad = \quad \left(\begin{array}{cccc} 010101 & 101010 & ..... & 110010 \end{array} \right)$$

$$(= individual) \quad = \quad \text{a set of genes } = \text{ a set of encoded unknowns}$$
$$\leftrightarrow \quad \{\tilde{q}^1/\tilde{C}^1, \ \tilde{q}^2/\tilde{C}^2, ..., \ \tilde{q}^{N_{event}}/\tilde{C}^{N_{event}}\}$$
$$\rightarrow \quad \text{A distribution of } \{\tilde{m}(p^i, \tilde{q}^i | \tilde{C}^i)\}$$

$$population \quad = \quad \text{set of (random)chromosomes}$$
$$\leftrightarrow \quad \text{set of (random)solutions}$$
$$\rightarrow \quad \text{larger population } \rightarrow \text{ better genetic diversity}$$
$$\rightarrow \quad \text{GA can survey larger solution space.}$$
$$fitness \quad = \quad \text{quality of a solution}$$
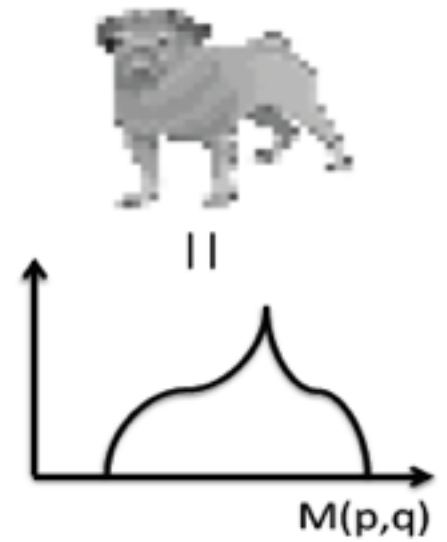
Flow-chart of Genetic Algorithm

1. Encode Chromosomes

   1.1 Binary encoding
   1.2 Continous encoding

2. Generating initial population

3. Evaluation of fitness values of all chromosomes

4. Convergence test (continue, if not converged)

5. Selection and Mating

   5.1 Roulette Wheel Selection in proportion to fitness
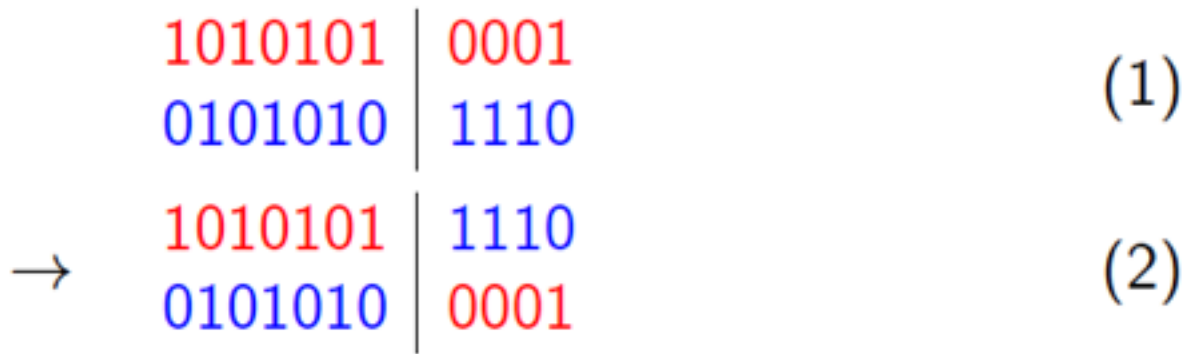
6. Reproduction (Crossover, Mutation)

a binary code represents each dog

dogs bark and receive rating

dogs selected for mating

offspring result & replace poor performers in population

introduce mutations

101101
010001
001011
011010

bowwow woof

yip 2
7
3 arf
5

010 001
011 010

010 010
011 001

011010

010010

[101101 ...]

= {q_1, q_2, ... q_N}

$M(p,q)$
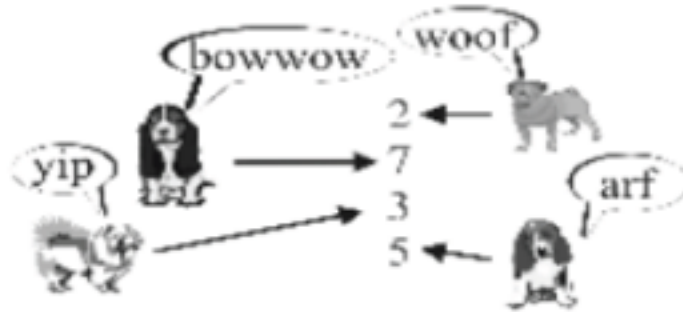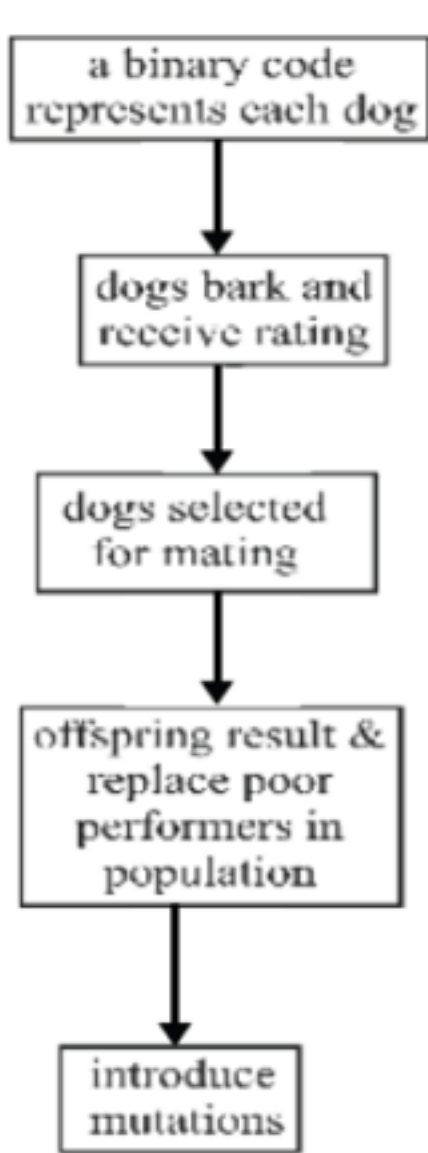
From 'Practical Genetic Algorithms' by Haupt.

Reproduction operators

- ▶ Crossover : random points are chosen on the individual chromosomes and the genes are exchanged at this point. This is the first operation where the GA explores another points in the variable space.

  e.g.) single point crossover

$$\begin{array}{c|c} \color{red}{1010101} & \color{red}{0001} \\ \color{blue}{0101010} & \color{blue}{1110} \end{array} \tag{1}$$

$$\rightarrow \quad \begin{array}{c|c} \color{red}{1010101} & \color{blue}{1110} \\ \color{blue}{0101010} & \color{red}{0001} \end{array} \tag{2}$$

► Mutation : this is the process by which a bit/string is deliberately changed so as to maintain diversity in the population set

$$\# \, of \, mutations \;\; = \;\; \mu \times (N_{pop} - N_{elite}) \times N_{bits},$$
$$\mu \;\; = \;\; mutation \, rate$$

- Clustering of the showered remnants by their partonic origins

- Easy example
1. Clustering 4 particles into two groups by their (resonant) decay origins

2. Clustering N-particles into M-groups by their origins

1. Grouping 4 visible momenta into 2 pairs - (a1,b1), (a2,b2) by their origins - 2 on-shell mother particles, each decays to (a_i, b_i), without the knowledge of M_{mother}.
   ($N_{event} = 100$, $N_{pop} = 1000$, Elite rate $= 0.2$, $\mu = 0.25$)
   (Fitness $=$ Height of the peak)
   1st generation :
   Best chromosome : 0001110........2112020
   fitness : 98 , Nfalse : 51 , NEvent : 100

   ....

   42nd generation :
   Best chromosome : 0000..1..00000000000
   fitness : 198 , Nfalse : 1 , NEvent : 100
   Running time - Real Time : 3.846, Cpu : 3.610 s

Gain :

1. Naive search : # of fitness estimation $\sim 3^{100} = 59049^{10}$
   $\rightarrow$ may not be possible before the end of the universe.

2. Stochastic search using GA : # of fitness estimation $\sim$
   $N_{pop}/2 \times N_{gen} = 500 \text{x} 42 = 21000$!!!
   $\rightarrow$ just 3-sec using an old single core.

- In the Electroweak theory of the SM, W mass is predictable

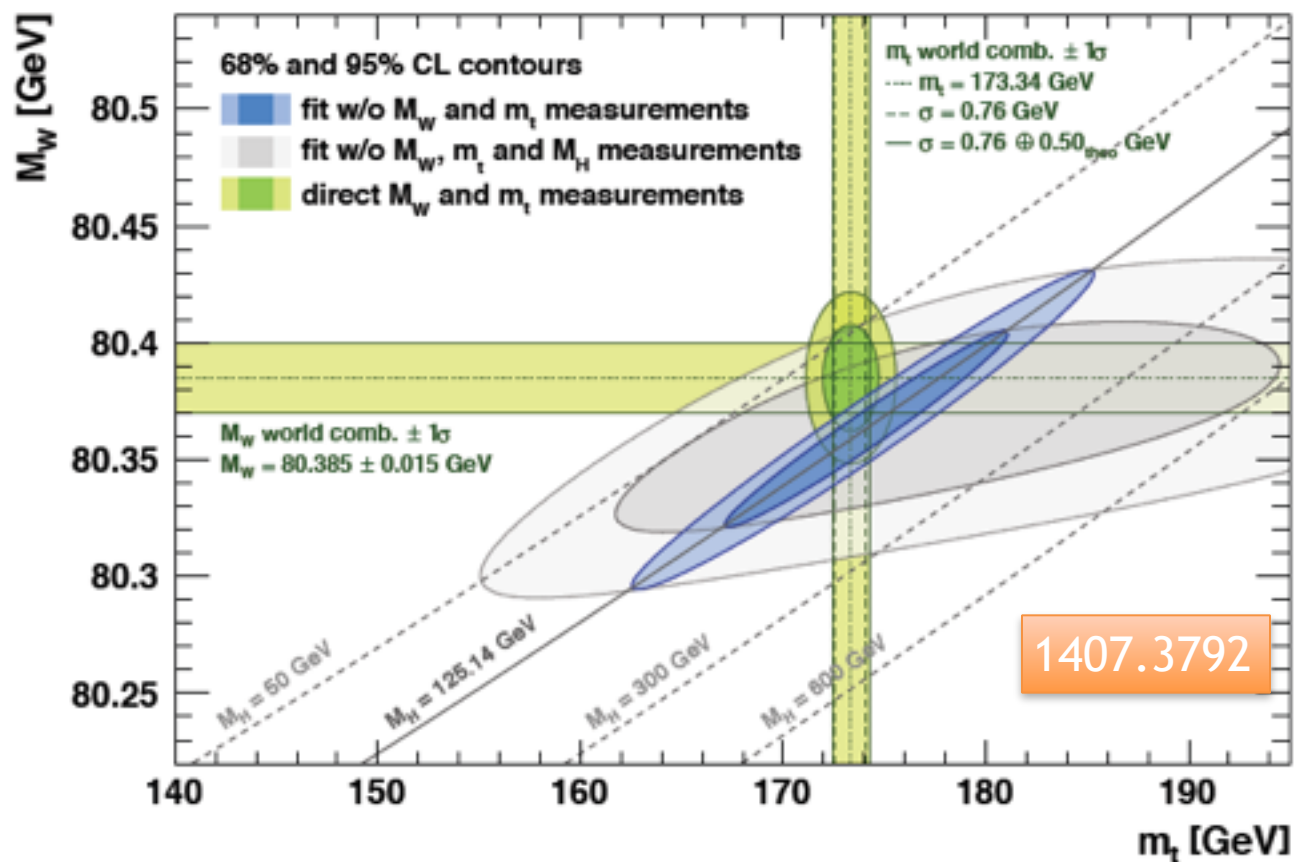$$M_W = \sqrt{\frac{\pi\alpha}{\sqrt{2}G_F}\frac{1}{\sin\theta_W\sqrt{1-\Delta r}}}$$

One loop radiative correction Δr to W mass depends quadratically on the top mass, and logarithmically on Higgs mass.
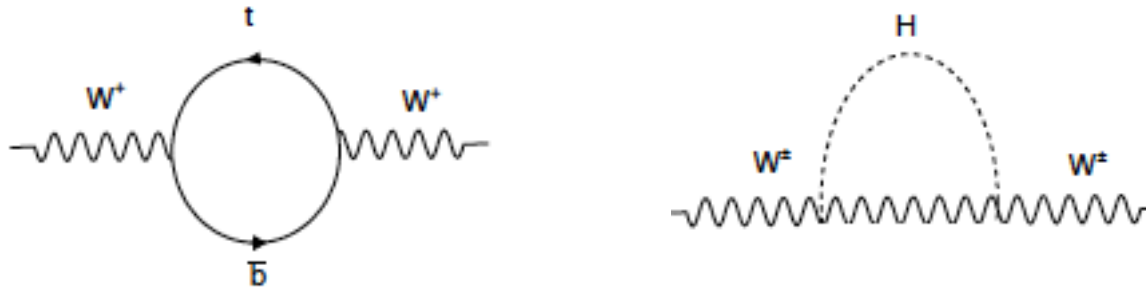
Measuring Mw and Mt, one can extract the information on MH, vice versa.

Some new physics can contribute on the Mw

With the discovery of the Higgs and direct Higgs mass measurement..

- Global EW fitting with MH(exp) indirectly estimated MW and Δ MW ~ 8 MeV
- Compared to the ΔMt(exp) ~ 0.9 GeV, precision on ΔMW(exp) ~ 5 MeV is required assuming equivalent contribution on Higgs mass precision.
- Both values motivate some improvements on current experimental precision, ΔMW ~ 15 MeV ➔ Precision MW measurement is one of the key EW precision tests for the SM consistency.

Precision W mass measurement – Status

- World average : 80385 ± 15 MeV (sys+stat)

- CDF result (80375 MeV, 2012)   D0 (80387 MeV 2012)

Source	Uncertainty (MeV)
Lepton energy scale and resolution	7
Recoil energy scale and resolution	6
Lepton removal from recoil	2
Backgrounds	3
Experimental subtotal	10
Parton distribution functions	10
QED radiation	4
$p_T(W)$ model	5
Production subtotal	12
Total systematic uncertainty	15
W-boson event yield	12
Total uncertainty	19

Source	Uncertainty (MeV)
Electron energy calibration	16
Electron resolution model	2
Electron shower modeling	4
Electron energy loss model	4
Recoil energy scale and resolution	5
Electron efficiencies	2
Backgrounds	2
Experimental subtotal	18
Parton distribution functions	11
QED radiation	7
$p_T(W)$ model	2
Production subtotal	13
Total systematic uncertainty	22
W-boson event yield	13
Total uncertainty	26

- Systematic uncertainty, (CDF:15, D0: 22) MeV > Statistical uncertainty, (CDF:22, D0 : 13) MeV

- How to reduce the systematic uncertainties?

Systematic uncertainties on MW measurement

1) **Experimental sources**

   1) **Lepton energy scale/energy resolution/shower model/energy loss/ efficiencies**
   2) **Backgrounds**
   3) **Recoil model**

   : limited by the size of the Z->ll sample => Enhanced event statistics can help.

2) **W and Z production and decay models**

   1) **PDF** : Boost of W-boson along longitudinal beam direction does not change a MT value. However uncertainty on the PDF affects on event acceptance rate, in particular with respect to the imposed pseudo-rapidity range, so MT distribution is also changed.
   2) **Transverse momentum of W boson** : PT(l) distribution is very sensitive on the non-zero PT(W). MT distribution is also affected by W boson's PT.
   3) **Photon radiation (QED)** : final state radiation from lepton

Measurement strategy and sources of systematic uncertainty

- W-boson production and decay
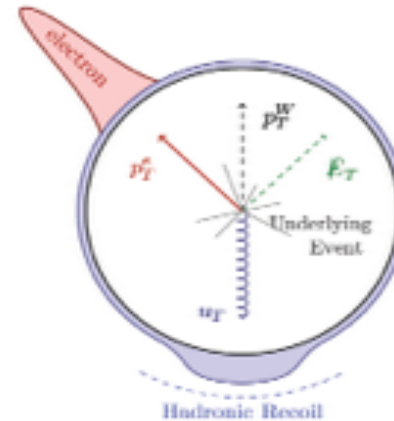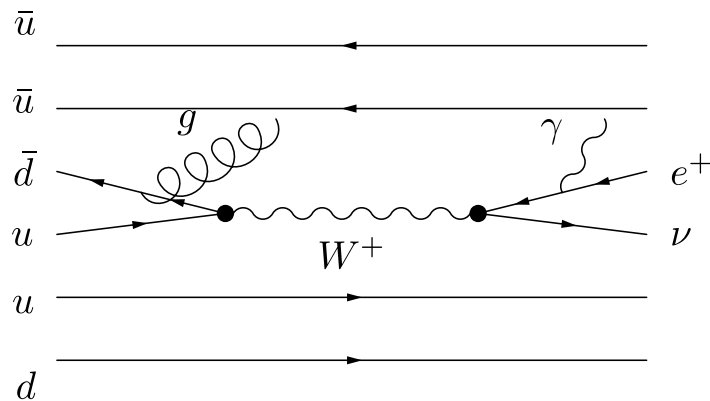
- Template method, using MT, PT(lepton) and PT(neutrino) variables

$$M_T = \sqrt{2P_T^l P_T^\nu (1 - \cos(\phi^l - \phi^\nu))}$$

- The distributions of the variables are unstable w.r.t the momentum profile of the mother particle, W production

  - Both MT and PT(l) distributions are not invariant under PT(W) and PDF.
  - One solution is to do precise calculation and simulation on the W production (and decay) to reduce the systematic uncertainties literally.

Consider a 'sampled invariant mass' of W, where the missing momentum along beam direction is assigned by random variable

$$\tilde{M}_W(p^l, p_T^\nu; \tilde{p}_z^\nu)^2 \equiv \left(p^l + \tilde{p}^\nu\right)^2 \quad \text{with} \quad \tilde{p}^\nu = (\sqrt{\mathbf{p}_T^{\nu\,2} + \tilde{p}_z^{\nu\,2}}, \, p_x^\nu, \, p_y^\nu, \, \tilde{p}_z^\nu),$$

$$= 2\left[|\mathbf{p}^l|\sqrt{\mathbf{p}_T^{\nu\,2} + \tilde{p}_z^{\nu\,2}} - \mathbf{p}_T^l \cdot \mathbf{p}_T^\nu - p_z^l\,\tilde{p}_z^\nu\right],$$

$$= 2|\mathbf{p}_T^l||\mathbf{p}_T^\nu|\left[\cosh(\eta^l - \tilde{\eta}^\nu) - \cos\phi_{l\nu}\right],$$

In this picture, M_T is also one sampled invariant mass of W, but which is especially optimised to minimise the mass scale.

$$M_T(W) \equiv \min_{\tilde{p}_Z} \tilde{M}_W \quad \rightarrow \quad \tilde{\eta}^\nu = \eta^l$$