Continuous simulation of hypothetical physics processes with multiple free parameters

J Zhong[1,2,3] and S-C Lee[2]

[1] School of Physics, Nanjing University, Nanjing, CN - Jiangsu 210093, China
[2] Institute of Physics, Academia Sinica, TW - Taipei 11529, Taiwan
[3] Physics Department, University of Oxford, UK - Oxford OX13RH, United Kingdom

E-mail: Jiahang.Zhong@cern.ch, phsclee@phys.sinica.edu.tw

Abstract. We present a new approach to simulate Beyond-Standard-Model (BSM) processes which are defined by multiple parameters. In contrast to the traditional grid-scan method where a large number of events are simulated at each point of a sparse grid in the parameter space, this new approach simulates only a few events at each of a selected number of points distributed randomly over the whole parameter space. In subsequent analysis, we rely on the fitting by the Bayesian Neural Network (BNN) technique to obtain accurate estimation of the acceptance distribution. With this new approach, the signal yield can be estimated continuously, while the required number of simulation events is greatly reduced.

1. Introduction

Many Beyond Standard Model (BSM) physics are extensively searched in collider experiments, such as Large Hadron Collider (LHC). As the properties of such new physics are unknown, they are often defined by multiple free parameters, including the masses of hypothetical particles and the related coupling constants. These free parameters can largely affect the kinematics and event topology, as well as cross-sections and branching ratios of these processes. It's important to understand such dependency, in order to constrain the phase space of possible new physics.

The conventional approach to study such dependency is the so-called "grid-scan" approach. A discrete grid with regular space is defined within the concerned parameter space, and a Monte-Carlo (MC) sample is generated at each point of the grid. There are several disadvantage of this approach. Firstly, it suffers the "curse of dimensionality", i.e. the number of points on the grid increases exponentially with the number of parameters. The grid often has to be reduced to two-dimensions, with all the other parameters fixed to certain values. Secondly, as the MC samples at different points are treated independently, a decent statistics is need for each sample in order to evaluate the acceptance precisely. Thirdly, the discreteness of the grid requires interpolation between points, which may leads to non-negligible systematic uncertainties when the spacing is large.

In this presentation, we demonstrated a new method to provide prediction of signal yield continuously over the multivariate parameter space, with much less MC simulation required.

2. Continuous simulation

Given a set of free parameters $\mathbf{x}$, the signal yield is defined as

$$N = \mathcal{L}\sigma(\mathbf{x})\epsilon(\mathbf{x}) \tag{1}$$

$\mathcal{L}$ is the integrated luminosity, which is irrelevant to the BSM parameters. σ is the cross-section times branching ratio of the concerned process. Although it is affected by the parameters, it can be calculated easily at generator level. ϵ represents the detector acceptance and reconstruction efficiency. It needs to be estimated with many events of detector simulation, which is generally the bottleneck of computing power. In this new method, we focused on facilitate an easier estimation of ϵ as a function of the parameters.

We first pick a large number of space points stochastically from the concerned parameter space. Then only one or a few events are simulated at each point $\mathbf{x}$. After that, we process all the events through the reconstruction and analysis procedures. A target value t of 1 or 0 is assigned to each event, depending on whether it passed the selection criteria of the analysis. A Bayesian Neural Network (BNN) [1, 2] is then used to perform an unbinned fitting with all these events, in order to approximate the probability distribution $\epsilon(\mathbf{x})$ by its output distribution $y(\mathbf{x})$. This is achieved by minimizing the cross-entropy cost function

$$\mathrm{CE} = \sum_{k}(-t_k \log y(\mathbf{x_k}) - (1 - t_k)\log(1 - y(\mathbf{x_k}))) \tag{2}$$

k loops over all simulated events. Minimizing this cost function is equivalent to maximizing the Bernoulli likelihood, which naturally gives $y(\mathbf{x})$ the meaning of success probability.

The Neural Network (NN) has been extensively used in Partticle Physics in the past decade, mostly as a multivariate discrimnant for signal/background separation [3, 4, 5]. In recent years, its usage as a universal function approximator [6, 7] has also been exploited by physicists [8, 9, 10].

There are several benefits to use the NN algorithm for this fitting. Firstly, it's a non-parametric model which does not require *a prior* knowledge of the distribution. Secondly, this algorithm assumes smoothness in the target distribution, which generally is true over the BSM parameter space. Thirdly, as an unbinned fitting, the information of input events can be fully exploited. Fourthly, NN is well-known as a multivariate tool. It can handle higher dimensional distributions easier, even if the dependency of $\epsilon(\mathbf{x})$ on each BSM parameters are correlated.

In addition, the Bayesian implementation of the BNN algorithm also allows us to estimate the uncertainty of each predicted $\epsilon(\mathbf{x})$ based on Bayesian inference framework. It provides us a handle on both the statistical fluctuation of the generated events, as well as the goodness of BNN fitting.

3. LRSM acceptance study

To demonstrate this new technique, we studied the acceptance $\epsilon(\mathbf{x})$ of di-lepton final state, for a hypothetical Left-Right Symmetrical Models (LRSM) processes [11]. In this process, a right-handed W boson (W_R) is produced, which subsequently decays into a heavy Majorana neutrino (N_R), as shown in figure 1.

The masses of the hypothetical W_R and N_R are the free parameters of this model. We considered the parameter space region with W_R mass between 0.5 TeV and 1.5 TeV, and N_R mass between 0 GeV and $M(W_R)$. The Pythia generator [12] is used to simulate this process. We randomly pick 100,000 points form the concerned parameter space, and simulate one event at each point. For simplicity, we only studied the acceptance at generator level, with 100% fiducial efficiency assumed for particle identification. Two leptons (e,μ) are required in each event, with transverse momentum greater than 20 GeV, and pseudo-rapidity η between -2.5 and 2.5. A

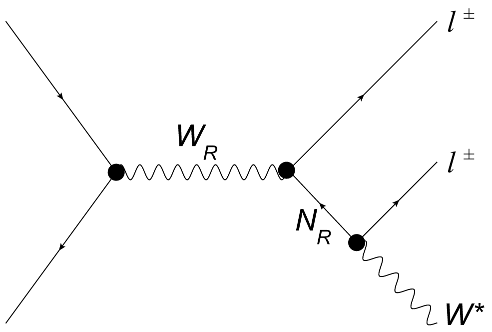
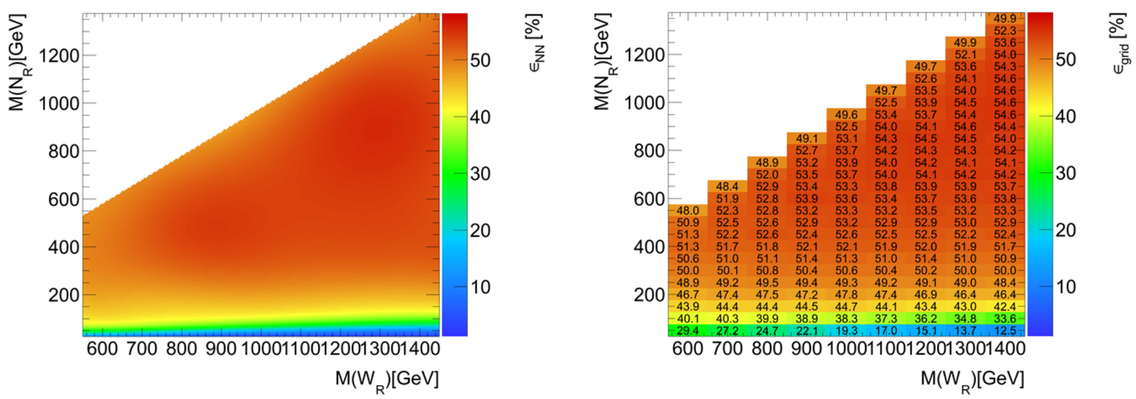


Figure 1. The Feynman diagram of W_R production in the LRSM model. The right-handed W boson will couple with a lepton and a hypothetical Majorana neutrino, which subsequently decays into another lepton and two jets.

typical isolation cut for QCD suppression is used, requiring the transverse energy flow of all interactive (non-neutrino) particles around the lepton, within a $\Delta R = \sqrt{(\Delta\eta)^2 + (\Delta\phi)^2} \leq 0.2$ cone, to be less than 10% of the lepton's transverse momentum.

Then all the simulated events are passed as input into a BNN. Target values t were assigned based on whether the event passed the above-mentioned selections. The BNN has a topology of 20 neurons in the first hidden layer and 10 neurons in the second hidden layer. The probability function of fitted $\epsilon(\mathbf{x})$ is shown in figure 2(a).

To demonstrate the performance, we also simulated another set of grid-scan samples. In this typical 2D grid, the values of $M(W_R)$ are from 600 GeV to 1400 GeV with 100 GeV steps, and the values of $M(N_R)$ are from 50 GeV to $(M(W_R) - 50)$ GeV with 50 GeV steps. The $\epsilon(\mathbf{x})$ evaluated from these samples can be seen in figure 2(b).



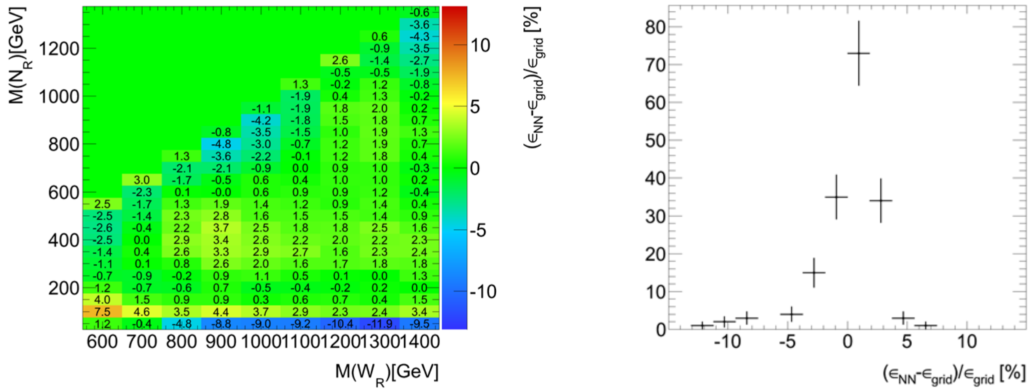
(a)

(b)

Figure 2. The acceptance distribution fitted by BNN (a) and estimated by the grid-scan samples (b).

The relative difference between the BNN fitted distribution and the reference values estimated from the grid-scan samples are shown in figure 3. For most points, the deviation is quite small, except the region where $M(N_R)$ approaches 0. In this edge region, the accuracy is worse because that $\epsilon(\mathbf{x})$ is relatively smaller, also the statistics of training sample is poorer due to the lack of neighbouring points.

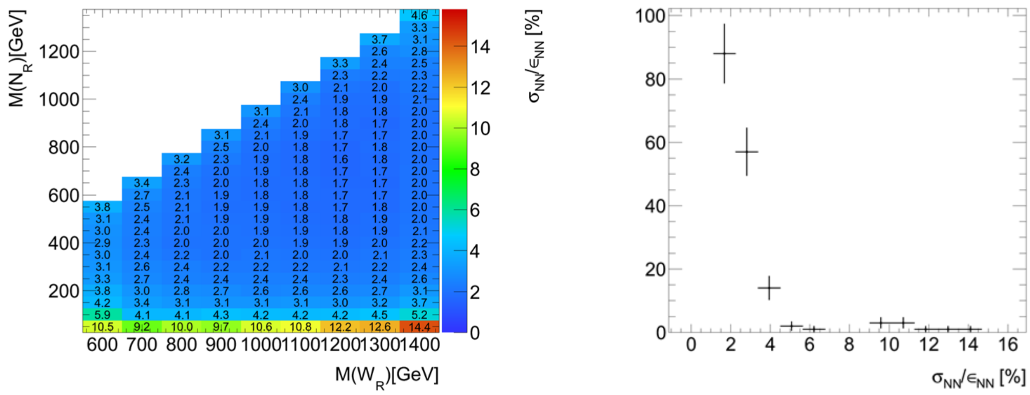
With the Bayesian uncertainty estimation, the above mentioned poorer accuracy is actually predictable. Figure 3 shows the uncertainty estimated by the BNN, based on the input statistics and goodness of fitting. We denote this uncertainty as $\sigma_{NN}(\mathbf{x})$. Further comparing this



(a) (b)

Figure 3. The relative deviation of the BNN predictions from the estimations by grid-scan, at the selected parameter space points (a), and the distribution of these deviations (b).

uncertainty to the actual deviation between the fitted values and the reference values (figure 3), we can see good consistency over the whole parameter space. This gives us the confidence on the BNN's uncertainty, allowing future application to evaluate the precision of BNN estimation without generating grid-scan samples for reference.

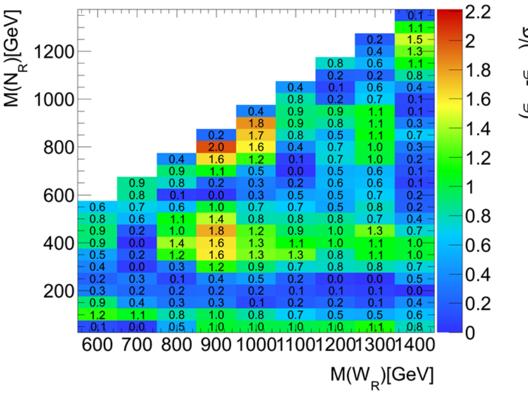


(a) (b)

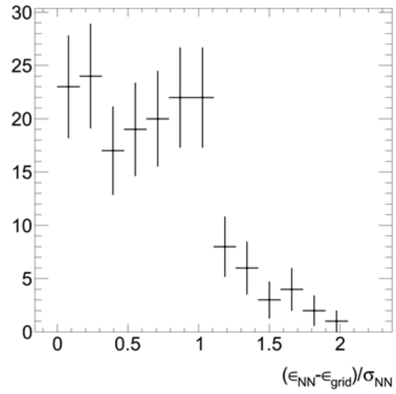
Figure 4. The uncertainty estimated by the BNN, at the selected parameter space points (a), and the distribution of these values (b).

4. Summary

In this work, we introduced a novel technique to simulate and analyse BSM processes defined by multiple parameters. A BNN algorithm is used for the unbinned fitting of the acceptance, which could help to suppress the required computing load greatly, while providing continuous estimation over the parameter space. The exemplar demonstration with LRSM model indicated that the technique can provide good approximation of the acceptance distribution, as well as precise uncertainties associated to the fitted distributions.



(a)



(b)

Figure 5. The deviation between the BNN predictions and the estimations by grid-scan, divided by the BNN uncertainties, at the selected parameter space points (a), and the distribution of these ratios (b).

Acknowledgments
We thank Song-Ming Wang, Rachid Mazini for helpful discussions. Jiahang Zhong and Shih-Chang Lee are partially supported by the National Science Council, Taiwan under the contract number NSC99-2119-M-001-015.

References

[1] MacKay D J C 1995 *Network: Computation in Neural Systems* **6** 469–505
[2] Zhong J, Huang R S and Lee S C 2011 *Computer Physics Communications* **182** 2655 – 2660 ISSN 0010-4655
[3] Abazov V M *et al.* (D0) 2001 *Phys. Lett.* **B517** 282–294
[4] Chiappetta P, Colangelo P, De Felice P, Nardulli G and Pasquariello G 1994 *Phys. Lett.* **B322** 219–223
[5] Abramowicz H, Caldwell A and Sinkus R 1995 *Nucl. Instrum. Meth.* **A365** 508–517
[6] Cybenko G 1989 *Mathematics of Control, Signals and Systems* **2** 303–314
[7] Hornik K, Stinchcombe M and White H 1989 *Neural Networks* **2** 359 – 366 ISSN 0893-6080
[8] Forte S, Garrido L, Latorre J I and Piccione A 2002 *Journal of High Energy Physics* **2002** 062
[9] Rojo J 2006 *Journal of High Energy Physics* **2006** 040
[10] Graczyk K M, Plonski P and Sulej R 2010 *JHEP* **09** 053 (*Preprint* `1006.0342`)
[11] Ferrari A, Collot J, Andrieux M L, Belhorma B, de Saintignon P, Hostachy J Y, Martin P and Wielers M 2000 *Phys. Rev. D* **62** 013001
[12] Sjostrand T, Mrenna S and Skands P 2006 *Journal of High Energy Physics* **2006** 026