FULLY BAYESIAN UNFOLDING IN HIGH-ENERGY PHYSICS*

Petr Baroň

Regional Centre of Advanced Technologies and Materials
Joint Laboratory of Optics of Palacký University, Czech Republic
and
Institute of Physics AS CR, Faculty of Science, Palacký University
17. listopadu 12, 771 46 Olomouc, Czech Republic

(Received April 14, 2020)

The process of unfolding is a crucial part of many particle physics analyses, representing the correction of measured spectra in data for the finite detector efficiency, acceptance and resolution from the detector to particle or from the particle to parton levels. Compared to other commonly used methods, the Fully Bayesian Unfolding (FBU) returns not only an unfolded value and its uncertainty, but provides the full binned posterior probability density. This study focuses on the dependence of unfolding results on the regularization parameter strength τ applied to different high-energy physics spectra.

DOI:10.5506/APhysPolB.51.1241

1. Introduction

Simplified schematic procedure for every unfolding method can be expressed for given data D, background B, and unfolded particle spectrum p in the case of migrations from the detector to particle level as

$$p = \frac{1}{\epsilon} M^{-1} \eta \left(D - B \right), \tag{1}$$

where ϵ and η are the efficiency and acceptance corrections respectively, and M^{-1} is the inverse migration matrix which maps migrations from the particle to detector levels of the studied spectrum. The symbol M^{-1} stands here also for different unfolding algorithms, *e.g.*, Invert, TUnfold, Svd, Ids, BinByBin, and IterativeBayes implemented as a part of the RooUnfold package [1].

* Presented at XXVI Cracow Epiphany Conference on LHC Physics: Standard Model and Beyond, Kraków, Poland, January 7–10, 2020.

All these approaches thus have the same input components D, B, M, ϵ and η, an example of which is shown in Fig. 1.

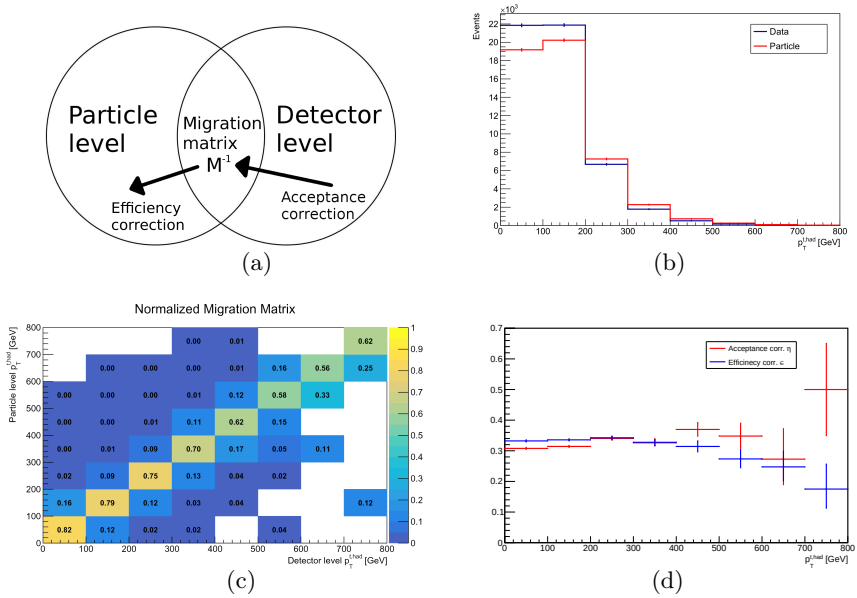


Fig. 1. (Color online) Unfolding ingredients. (a) Migration diagram; (b) Detector-level (gray/blue) and particle-level (light gray/red) spectra; (c) Migration matrix between particle and detector levels; (d) Efficiency (gray/blue) and acceptance (light gray/red) corrections as a function of the transverse momentum of the hadronically decaying top quark.

2. Fully Bayesian Unfolding

The Bayesian theorem is the main building block of the presented FBU [2] method and is based on the conditional probability of A given B

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)}\,. \tag{2}$$

In applications in high-energy physics, Eq. (2) can be rewritten to obtain the probability of the truth T (particle spectrum) given measured data D as

$$P(T|D) = \frac{P(D|T)\,\pi(T)}{\text{Norm.}}\,, \tag{3}$$

where $P(D|T)$ is the likelihood function $L(D|T)$ and $\pi(T)$ is the prior information on the truth spectrum. The prior is usually unknown, and if set to a constant, called *flat*.

However, the prior might be chosen as an arbitrary function of T, driven by some reasonable arguments of the analyzer, in this way, the *regularization* of the FBU method is introduced and defined by $\pi(T)$.

The detailed formula for the unfolded spectrum with N bins is constructed using a product of Poisson distributions

$$\begin{aligned} P(T|D) &\propto L(D|T)\,\pi(T) \\ &= \left(\prod_{i=1}^{N} \frac{1}{\epsilon_i} \frac{\left(\sum\limits_{j=1}^{N} M_{ij} T_j \right)^{[\eta_i (D_i - B_i)]}}{[\eta_i (D_i - B_i)]!} \mathrm{e}^{-\left(\sum\limits_{j=1}^{N} M_{ij} T_j \right)} \right) \mathrm{e}^{-\tau_{\mathrm{abs}} S(T)} , \end{aligned} \tag{4}$$

where the prior is introduced as an exponential $\pi(T) = \mathrm{e}^{-\tau_{\mathrm{abs}} S(T)}$ with the parameter τ_{abs} describing the strength of the regularization. In other words, it implies in the case of $\tau_{\mathrm{abs}} = 0$ that no regularization is applied and the prior is flat (equal to one). In contrast, non-zero τ_{abs} with regularization function $S(T)$ defines how the resulting spectrum is regularized.

In this study, two approaches of regularization were taken into account. First, the curvature regularization

$$S(T) = \sum_{t=2}^{N-1} \left(\Delta_{t+1,t} - \Delta_{t,t-1} \right)^2 \tag{5}$$

using the sum of second derivatives (differences)

$$\Delta_{t_1,t_2} = T_{t_1} - T_{t_2} \tag{6}$$

and second, the entropy regularization

$$S(T) = -\left[-\sum_{t=1}^{N} \frac{T_t}{\sum T_{t'}} \log \left(\frac{T_t}{\sum T_{t'}} \right) \right] . \tag{7}$$

In order to get the posterior of the unfolded spectrum for each i^{th} bin, the marginalization of the full multi-dimensional posterior is performed

$$p_i(T_i|D) = \int \int P(T|D) \mathrm{d}T_1 \ldots \mathrm{d}T_{i-1} \mathrm{d}T_{i+1} \ldots \mathrm{d}T_N \, . \tag{8}$$

The unfolded result is taken as the fitted mean of the fit Gauss function and the uncertainty is taken as the posterior σ_{Gauss} standard deviation. Example of four bins of the spectrum of transverse momentum of hadronically decaying top quark derived in private simulation using MadGraph [3] and Delphes [4] is given in Fig. 2 and Fig. 3. See the details below.

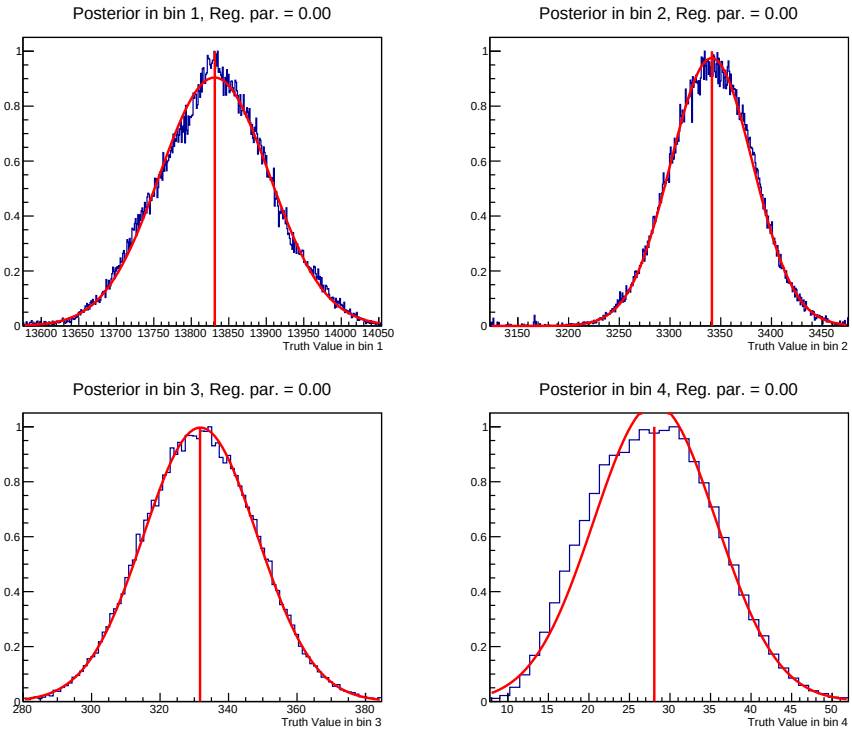


Fig. 2. (Color online) Marginalized posterior probability for four bins of the transverse momentum of hadronically decaying top quark in private simulation. The vertical (red) line represents the mean of the fitted Gauss function.

The dimension of the likelihood function $L(T)$ is given by the number of bins N of the studied spectrum. While running the FBU method, many pseudo-experiments (truth spectrum) need to be generated and the calculated likelihood function has to be efficiently sampled in the truth space.

One of the efficient ways is to use a Monte Carlo Markov Chain algorithm, and especially the Hamiltonian Monte Carlo Markov Chain with the No-U-Turn sampler [5]. The idea is to transfer the sampling of $L(T)$ to classical motion of a virtual particle in an N-dimension hyperspace with the potential $L(T)$. The name *chain* refers to the fact that the motion of the particle is derived step-by-step and creates *chains*, see Fig. 4. The feature of No-U-Turn samples avoids the particle coming back to the same point in the hyper-space so the computation converges faster.

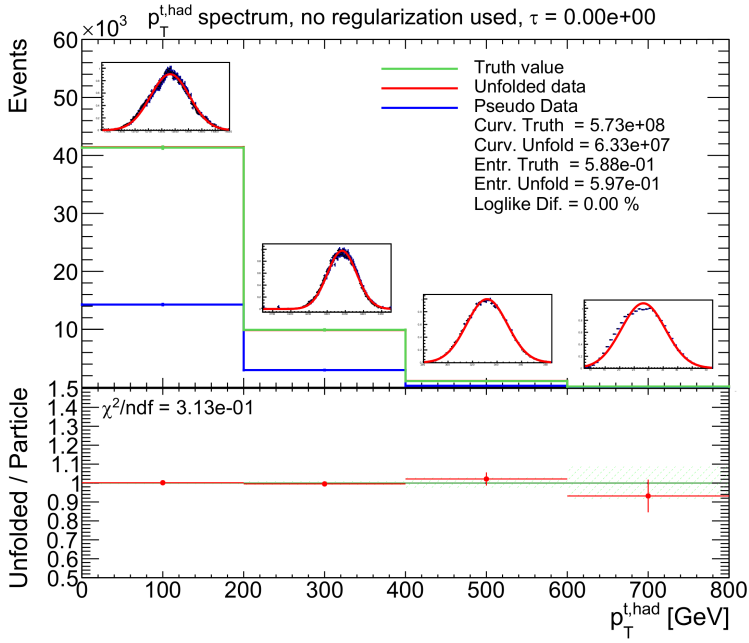


Fig. 3. Unfolded spectrum of the transverse momentum of the hadronically decaying top quark produced in the process of top-quark pair production in proton–proton collisions, private simulation. Insets show the FBU posteriors for each bin.

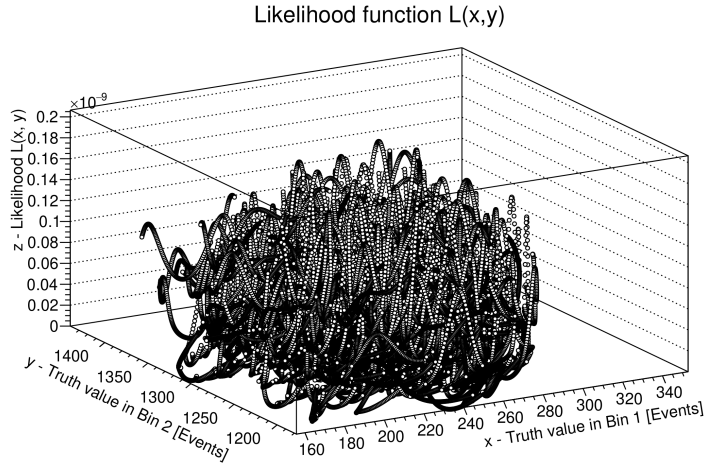


Fig. 4. Example of the likelihood function sampling using HMCMC and No-U-Turn sampler.

3. Regularization strength parameter study

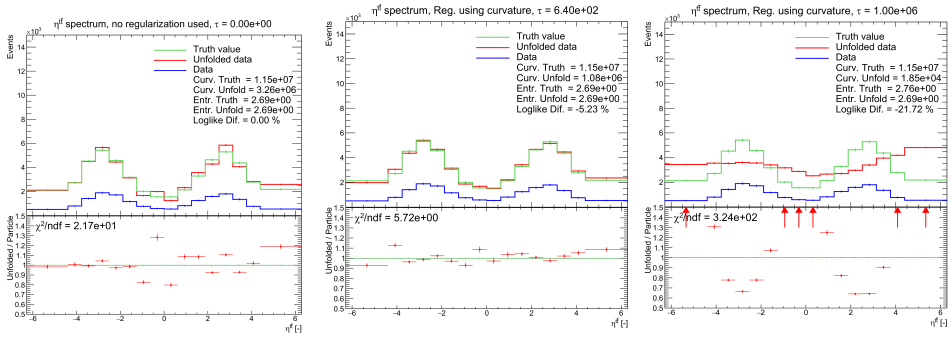
In this section, commonly used spectra were unfolded with both curvature and entropy based regularization.

The spectra of top quarks in proton–proton collisions at $\sqrt{s} = 14$ TeV were generated by the MadGraph generator [3], showered by PYTHIA 8 [6] to provide the particle level, and finally with the detector level simulated by Delphes [4].

Throughout this section, every τ parameter is normalized to curvature C_{truth} resp. entropy E_{truth} of the particle spectrum and number of bins N

$$\tau = \frac{\tau_{\text{abs}}}{C_{\text{truth}}\, N} \qquad \text{resp.} \qquad \tau = \frac{\tau_{\text{abs}}}{E_{\text{truth}}\, N} \tag{9}$$

so that the normalized τ is roughly comparable between spectra. As an example of the regularization effect, spectra of the top-quark pairs $\eta^{t\bar{t}}$, $m^{t\bar{t}}$ and $p_{\text{T}}^{t\bar{t}}$ were chosen with the gradual impact of regularization shown from left to right in figures 5, 6 and 7 using the curvature method.



(a) $\tau = 0$ (b) $\tau = 640$ (c) $\tau = 1\,000\,000$

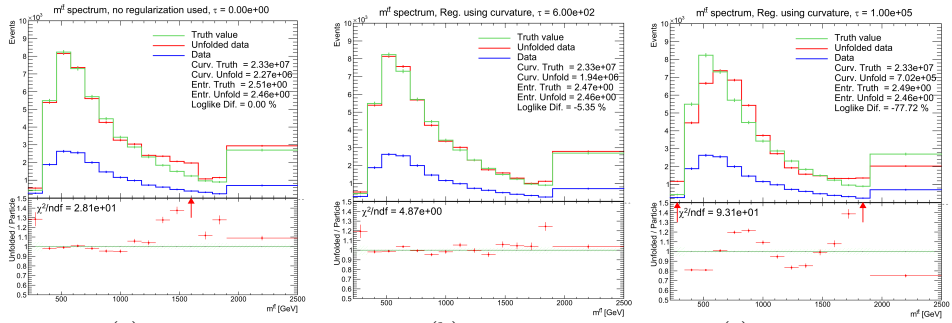
Fig. 5. Unfolding closure test of the $\eta^{t\bar{t}}$ over-binned spectrum with the average diagonal value of the normalized migration matrix $\bar{f}_{\text{diag}} = 0.4$ for different values of the regularization strength parameter τ.

Plots on the left are obtained without applying the regularization, in the middle regularization is applied with an optimal parameter strength τ, and on the right there are plots with extremely high τ.

Finer binning was chosen on purpose to construct example spectra which are more difficult to unfold, because of statistical fluctuations. The variable describing this is the average fraction of events staying on the diagonal

$$\bar{f}_{\text{diag}} = \frac{\sum_{i=1}^{N} M_{ii}}{N}\,, \tag{10}$$

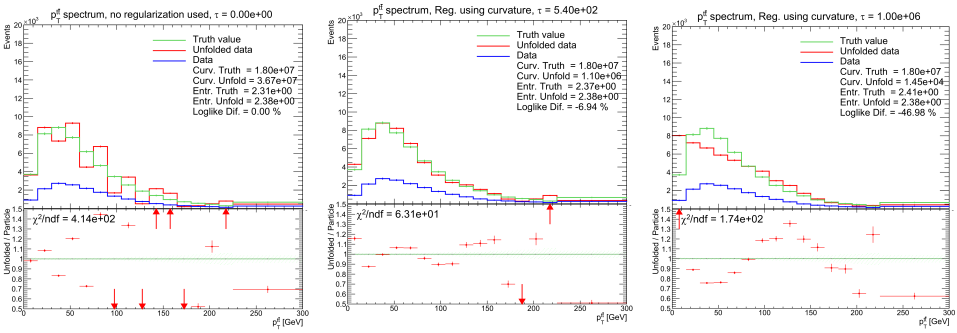
where N is the number of bins and M is the normalized unfolding matrix. The closer the $\bar{f}_{\text{diag}}$ is to one, the more stable unfolding process is. Usually, it is required at least 50% of migration in the diagonal bins, $\bar{f}_{\text{diag}} > 0.5$. The spectra in figures 5, 6 and 7 correspond to $\bar{f}_{\text{diag}} = 0.4$, 0.5 and even 0.22.



(a) $\tau = 0$ (b) $\tau = 600$ (c) $\tau = 100\,000$

Fig. 6. Unfolding closure test of the $m^{t\bar{t}}$ over-binned spectrum with the average diagonal value of the normalized migration matrix $\bar{f}_{\text{diag}} = 0.5$ for different values of the regularization strength parameter τ.

The unfolding of the $m^{t\bar{t}}$ spectrum with $\bar{f}_{\text{diag}} = 0.22$ oscillates without using regularization (figure 7 (a)). In this case, regularization is useful, as can be seen in figure 7 (b).



(a) $\tau = 0$ (b) $\tau = 540$ (c) $\tau = 1\,000\,000$

Fig. 7. Unfolding closure test of the $p_{\text{T}}^{t\bar{t}}$ over-binned spectrum with the average diagonal value of the normalized migration matrix $\bar{f}_{\text{diag}} = 0.22$ for different values of regularization strength parameter τ.

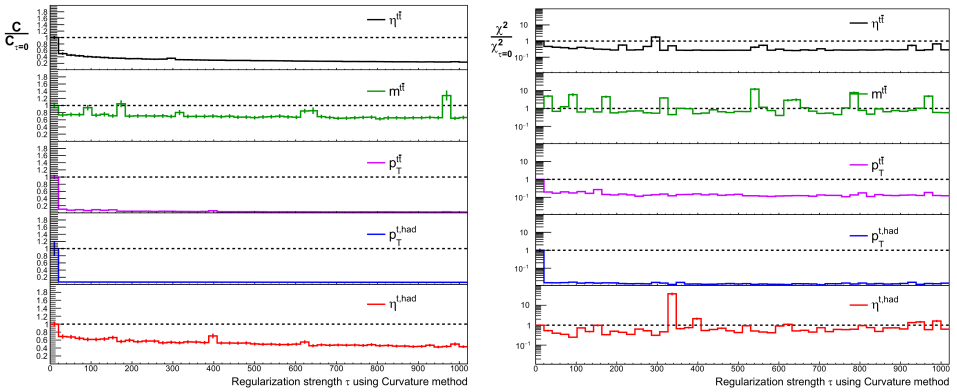
The problem which emerges is the selection of the optimal strength parameter τ to obtain the best result. As a metric of a successful unfolding, the relative $\chi^2_{\text{rel.}}/\text{n.d.f.}$ as a function of τ was chosen. The relative χ^2 is computed between the unfolded and the particle spectrum in simulation as

$$\chi^2_{\text{rel}}(\tau)/\text{n.d.f.} = \frac{\chi^2_{\text{reg}}(\tau)/\text{n.d.f.}}{\chi^2_{\tau=0}/\text{n.d.f.}} = \frac{\chi^2_{\text{reg}}(\tau)}{\chi^2_{\text{no-reg}\,\tau=0}}\,. \tag{11}$$

This study provides unfolding with the regularization parameter τ in the range $[0, 1000]$ in equidistant binning of width 20. Uncertainties are obtained using 20 unfolding processes for each bin with different starting random seed in the MCMC. The migration matrix and corrections are statistically-independent of input spectra.

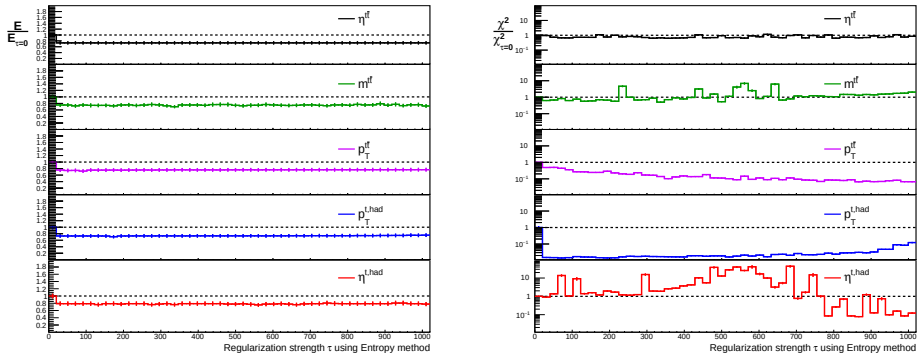
3.1. Results

Figure 8 represents results of many unfolding rounds using regularization based on minimizing the curvature or entropy (figure 9). On the left, there are drawn the curvature or entropy, relatively to values without applying regularization. Thus values in the first bin are equal to one by definition. On the right, relative χ^2 is plotted as a function of τ. Expected decreasing behavior of the relative curvature and entropy with respect to τ was proven (figures 8 (a), 9 (a)).



(a) $\frac{C(\tau)}{C_{\tau=0}}$ (b) $\frac{\chi^2(\tau)}{\chi^2_{\tau=0}}$

Fig. 8. Relative curvatures (left) and χ^2 (right) of five spectra as a function of the regularization parameter τ.



(a) $\frac{E(\tau)}{E_{\tau=0}}$ (b) $\frac{\chi^2(\tau)}{\chi^2_{\tau=0}}$

Fig. 9. Relative entropies (left) and χ^2 (right) of five spectra as a function of the regularization parameter τ.

4. Conclusion

Results show the ability to use regularization in the FBU unfolding method. The expected behavior is the improvement in terms of χ^2 with an optimal strength parameter τ, but for larger τ, the χ^2 rises to the point where the unfolded spectrum is close to a constant with the lowest curvature and entropy.

This is demonstrated in the case of different spectra and binning, to judge a variety of results. On the other hand, the normalization of τ to the truth curvature and entropy should make the τ comparable between spectra.

Despite this fact, no common minimum for all five spectra of the relative χ^2 was found. To be able to derive a general formula for the optimal strength parameter τ choice, the region of τ on the x-axis has to be extended with finer binning and the process needs to be more deeply understood, *e.g.*, different normalization of τ could be one of the possible ways.

The author gratefully acknowledges the support from the projects IGA PrF 2019 008 and IGA PrF 2020 007 of the Palacký University as well as the grant of the Ministry of Education, Youth and Sports, Czech Republic, LTT-17018, and of GACR, 19-21484S.

REFERENCES

[1] T. Adye, H.B. Prosper, L. Lyons (Eds.) «Proceedings of the PHYSTAT 2011 Workshop on Statistical Issues Related to Discovery Claims in Search Experiments and Unfolding», *CERN*, Geneva, Switzerland, 17–20 January 2011, pp. 313–318.

[2] G. Choudalakis, «Fully Bayesian Unfolding», arXiv:1201.4612 [physics.data-an].

[3] J. Alwall *et al.*, *J. High Energy Phys.* **1407**, 079 (2014), arXiv:1405.0301 [hep-ph].

[4] DELPHES 3 Collaboration (J. de Favereau *et al.*), *J. High Energy Phys.* **1402**, 057 (2014), arXiv:1307.6346 [hep-ex].

[5] M.D. Hoffman, A. Gelman, *J. Machine Learning Res.* **15**, 1593 (2014), http://jmlr.org/papers/v15/hoffman14a.html

[6] T. Sjöstrand *et al.*, *Comput. Phys. Commun.* **191**, 159 (2015), arXiv:1410.3012 [hep-ph].