ATLAS Muon Resolution Studies for the Higgs Mass Measurement

Alex Yuanming Wen[abc], Lukas Adamek[ac], Pierre Savard[ac]
[a] *The University of Toronto,*
[b] *the University of British Columbia, and*
[c] *the European Organization for Nuclear Research (CERN)*
(Dated: August 21, 2020)

The precise measurement of the Higgs mass at ATLAS is a primary goal of the current scientific programme and is an important parameter for many current and future theoretical models and phenomenology. Concerning the $H \to ZZ \to 4l$ decay channel, an original method to arrive at a four lepton invariant mass resolution is presented, which in turn relies on an determination of the muon momentum resolution. The accuracy of this invariant mass resolution is demonstrated. We then present the method to form a Higgs mass estimate and its uncertainty given a signal model and an invariant mass resolution estimate, and explore some avenues of improvement.

I. INTRODUCTION

The Higgs boson is one of the newest particles known since its discovery in 2012, and it is critical for its role in mass generation for the Standard Model [2]. As it is the subject of many theoretical and phenomenological studies, precise information on the Higgs, in particular its mass, is necessary. This is a current goal of ATLAS, especially with the availability of the Run 2 dataset, collected from 2015-2018 with a center-of-mass energy of $\sqrt{s} = 13\ TeV$ and an integrated luminosity of 139 fb^{-1} [1], the most statistically-rich ATLAS dataset collected to date.

The Higgs can be produced and decay in a variety of ways. Of particular interest is the $H \to ZZ \to 4l$ channel; while $H \to ZZ$ has a branching ratio of 2.6% at a Higgs mass $m_H = 125\ GeV$ [2], the production of 4 leptons means that it forms a relatively clean and easily detectable signal to the ATLAS detector. In the present analysis, we focus on $H \to 4\mu$ events.

In the detector, the transverse momentum p_T, pseudorapidity η, and azimuthal angle ϕ of each muon are measured; this information is collected for all 4 muons and is used to calculate an invariant mass m_{4l} that could then be fitted to estimate m_H. The directions (η, ϕ) and transverse momenta (p_T) of reconstructed muon candidates are measured by the inner detector and muon spectrometer of ATLAS. When calculating the reconstructed mass observable that is used to perform the mass measurement, the largest contribution to the observed resolution of the invariant mass spectrum comes overwhelmingly from the detector's ability to resolve the transverse momentum p_T, and not η and ϕ, of the muon candidates. Therefore, the entire measurement is limited largely by the p_T resolution, and it is important to make sure that this quantity is determined accurately. Contrary to previous approaches, most notably involving a neural network [1], we present a novel process to determine the muon p_T resolution $\sigma(p_T)$ and propagate this value for the invariant mass calculation.

In this work, the Monte Carlo simulation samples used were generated by the POWHEG QCD interfacing method [3] and the Pythia8 event generator [4], modelling the three data taking periods (2015-16, 2017, 2018) of Run 2, with gluon-gluon fusion (ggH) as the production mode for all mass points and additional vector boson fusion (VBF) mode events for the 125 GeV mass point. The mass points considered are 123, 124, 125, 126, and 127 GeV. While the MC samples include $4e$, $2\mu 2e$ and $2e2\mu$ events, we presently

only consider 4μ events. For the 125 GeV mass point 4μ samples, there are around 76 events given Standard Model predictions for the cross section; this is also after all event-level selections.

II. EXTRACTING THE MUON MOMENTUM RESOLUTION

Using a binning scheme presented in [1], we bin the muon population from all 4μ MC events into the $p_T - |\eta|$ parameter space.

In each bin, we plot the distribution of $(p_T^{true} - p_T)/p_T^{true}$ and obtain a distribution. We subsequently fit a Gaussian to this distribution in each bin and take the width of the Gaussian fit to be the $\sigma(p_T)/p_T$ value that we assign to every muon in that bin. Multiplying by the p_T of every muon, we then obtain a muon resolution $\sigma(p_T)$ value. Figure 1 is the heatmap which shows the value of $\sigma(p_T)/p_T$ in each bin.

Figure 2 shows the Gaussian fit in one particular bin. The distribution in each bin is not, however, perfectly Gaussian - hence the fit to only a limited range in the center of the distribution.

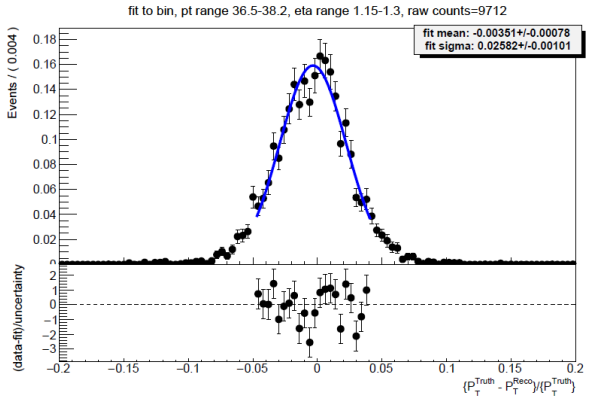


FIG. 2: the Gaussian fit in one particular bin of the heatmap pictured in figure 1.

III. ESTIMATING THE INVARIANT MASS RESOLUTION

After estimating the muon $\sigma(p_T)$, we must propagate this resolution on all four muons to form a four lepton invariant mass resolution $\sigma(m_{4l})$ estimate.

To estimate the m_{4l} mass and to easily propagate an uncertainty on p_T, we must express m_{4l} in terms of the p_T, η and ϕ of each muon. Starting with

$$m_{4l}^2 = \left(\sum_i E_i\right)^2 - \left|\sum_i \vec{p}_i\right|^2 \tag{1}$$

with i iterating over all 4 muons, E being energy and $\vec{p}$ being Cartesian momentum, we realize that in ATLAS the muon mass (order 100 MeV) is negligible compared to its momentum (order tens of GeV). Hence we assume $E \approx p = |\vec{p}|$ and therefore

$$m_{4l}^2 \approx \left(\sum_i p_i\right)^2 - \left|\sum_i \vec{p}_i\right|^2 . \tag{2}$$

We can expand and simplify equation 2 to obtain

$$m_{4l}^2 \approx \sum_{i \neq j} (p_i p_j - \vec{p}_i \cdot \vec{p}_j) \tag{3}$$

where i, j sum over the muons - in this case they would iterate from 1 to 4. This is not helpful, however, because we have uncertainties on p_T - so to introduce p_T into the expression, we must use the conversion from Cartesian momentum to accelerator coordinates: $p_x = p_T \cos(\phi)$, $p_y = p_T \sin(\phi)$, and $p_z = p_T \sinh(\eta)$. Furthermore, $p = |\vec{p}| = p_T\sqrt{\cos^2(\phi) + \sin^2(\phi) + \sinh^2(\eta)} = p_T\sqrt{1 + \sinh^2(\eta)} = p_T \cosh(\eta)$. Applying these identities and simplifying, equation 3 takes the form

$$m_{4l}^2 \approx \sum_{i \neq j} p_{T,i} p_{T,j} \left(\cosh(\eta_i - \eta_j) - \cos(\phi_i - \phi_j)\right) \tag{4}$$

which makes it easy to calculate a value for $\sigma(m_{4l}^2)$, using a familiar method of simple error

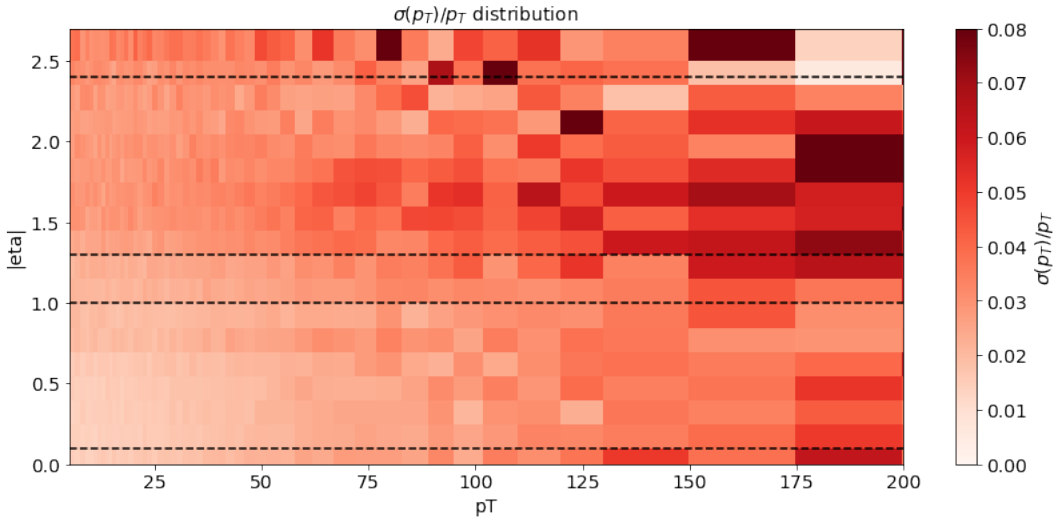


FIG. 1: the heatmap used for the binned approach in determining $\sigma(p_T)/p_T$. For the bins at the more extreme values of p_T and η, there are very few events and consequently the fits are poor, yielding less accurate values.

propagation, if we assume uncertainties only on p_T:

$$\sigma(m_{4l}^2) = \sqrt{\sum_i \left(\frac{\partial m_{4l}^2}{\partial p_{T,i}}\sigma(p_{T,i})\right)^2}. \qquad (5)$$

Figure 3 shows the distribution of $\sigma(m4l)$ and figure 4 shows the distribution of m_{4l} calculated using this method, from the 125 GeV MC sample. They broadly agree with first-order expectations of the range and peak of the distributions.

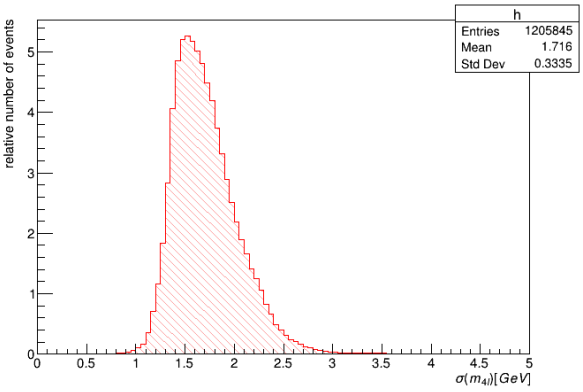


FIG. 3: Distribution of $\sigma(m_{4l})$ calculated from 125 GeV MC.

IV. VERIFYING THE INVARIANT MASS RESOLUTION

A. Pull Plot

To check that the estimated resolution is useful, and describes the variation in measured masses from the detector resolution, a pull distribution is studied. It is ideal, but not necessary that the pull distribution has a mean of 0.0 and width of exactly 1.0. Things are complicated by slightly non-Gaussian momentum mismeasurement distributions (see figure 2), and by other effects such as final state radiation. A pull plot in this case is a plot of $(m_{4l} - m_{4l}^{truth})/\sigma(m4l)$, where m_{4l}^{truth} is available from MC. Figure 5 shows the pull plot.

Both the mean (~ -0.09) and the standard deviation (~ 1.06) do not agree with expected values, this disagreement surpassing the magnitude of their statistical uncertainties. Due to the non-Gaussian appearance of fitted distributions as exemplified in figure 2, we expect some sort of

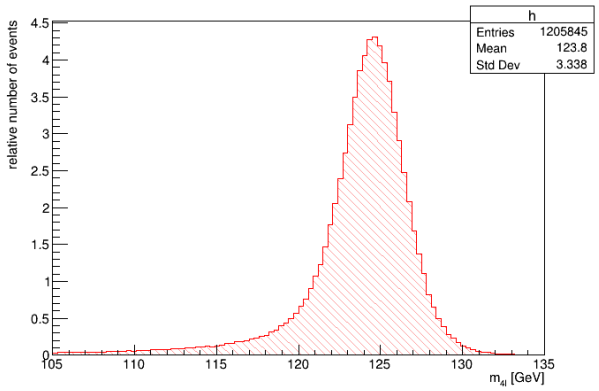


FIG. 4: Distribution of m_{4l} calculated from 125 GeV MC.

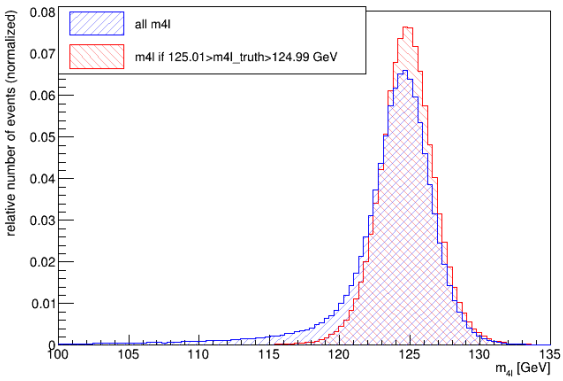


FIG. 6: the m_{4l} with and without a cut on truth m_{4l} that removes most FSR effects. Both histograms are normalized to unity.

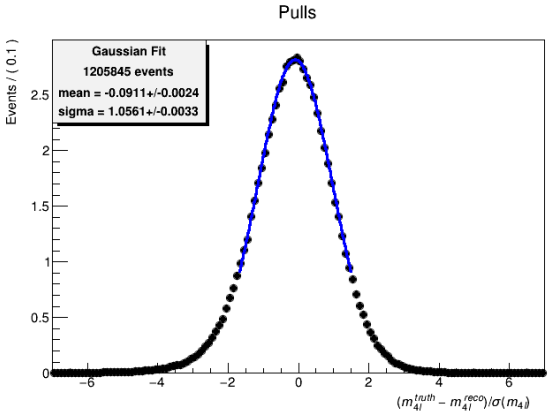


FIG. 5: the pull plot with respect to m_{4l}. The fitted mean is 0.091 ± 0.002 and the standard deviation is 1.056 ± 0.003.

discrepancy in the estimation of $\sigma(m_{4l})$. Strictly speaking, due to this non-Gaussian property, we are not estimating the resolution (width) of m_{4l} distributions - rather, we are creating a variable (we call "uncertainty" - $\sigma(m_{4l})$) that is highly correlated with the resolution, and we use this variable to categorize events based on the m_{4l} distribution resolution and to aid in our mass measurement. While it would be preferable, the pull plot need not be perfect; slight biases in the pull plot parameters do not necessarily render our estimate inaccurate. This process then also relies on the nature of the signal model used in the mass meausrement (see sections V and VI) to correctly model the relationship between $\sigma(m_{4l})$ and the real resolution.

Physically, the non-Gaussian nature of m_{4l} distributions can be attributed to final state radiation (FSR) effects, when emitted radiation decreases the momentum of a lepton before its detection, creating a low-mass tail in the final mass distribution. This can be demonstrated via the 125 GeV mass point MC samples, where if we cut away the truth m_{4l} values generated with FSR, the measured m_{4l} distribution becomes far more Gaussian. Figure 6 shows the difference between the entire measured m_{4l} distribution, and one that has added FSR effects removed via the cut.

B. Toy Analysis

Another way to check the error propagation is to assume that the p_T values are Gaussian-distributed on each muon (which they are not, but this method still verifies the process of error propagation), and sample many events from these four Gaussian distributions on each event to create a sample of toy events to accompany each event.

We then create a m_{4l} distribution of these toy events (which should be Gaussian) and com-

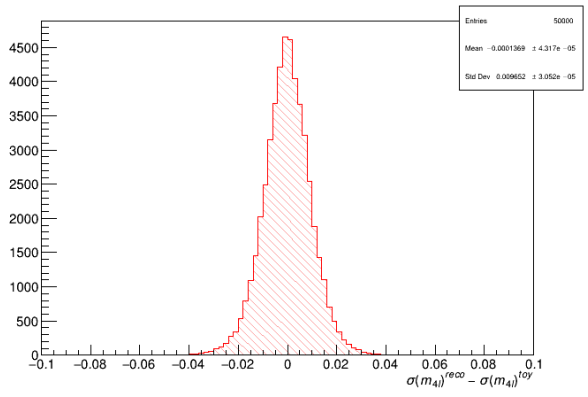


FIG. 7: a comparison between the toy m_{4l} width and the estimated $\sigma(m4l)$ from a random selection of 50,000 events. In agreement with expectation, there is minimal bias.

pare its width, $\sigma(m4l)^{toy}$, to the $\sigma(m_{4l})$ estimate for that event. These two values should agree, since the width of the m_{4l} distribution is exactly what we are estimating when we try to propagate uncertainty to m_{4l}. Figure 7 shows a plot of $\sigma(m_{4l})-\sigma(m_{4l})^{toy}$ for 50,000 random events. The distribution is almost unbiased, and we have verified that it narrows if we increase the number of toy samples we use to build our toy distribution. This agrees with expectations and shows that our error propagation process is valid (though it does not indicate the how good the muon $\sigma(p_T)$ estimate is, since we start with that value in building the toy distribution).

V. MASS MEASUREMENT

The ultimate goal, which motivates the trouble of estimating $\sigma(m_{4l})$ by our method or another, is to improve the statistical uncertainty $\sigma(m_H)$ on the final mass measurement.

This can be done in many ways, and produces a mass estimate and uncertainty that is dependent on the signal model that we choose to employ. No matter the model, however, the basic process is the same. We first conduct a global m_{4l} fit on the MC samples from all true Higgs mass points m_H (in our case, integer masses from 123-127 GeV) and establish the fit parameters of our signal model. Fixing these parameters, on the MC or data sample we want to measure, we allow only m_H to vary and conduct a profile likelihood ratio scan, which establishes the optimal fit value of m_H and gives us a statistical uncertainty.

The default signal model is a double crystal ball function fit, which consists of a Gaussian core with power-law tails; it is often used in particle physics analysis to model lossy processes that would feature prominent tails unsuitable for a pure Gaussian. For this fit, the mean μ_{DCB} of the Gaussian core is parameterized as

$$\mu_{DCB} = a(m_H - 125\ GeV) + b \qquad (6)$$

where a and b are parameters to be fitted. The reasoning for this parameterization, as also described in [1], is to model the relationship between the true mass and the mean of the corresponding m_{4l} distribution, while at the same time trying to minimize the effect of the uncertainty of a (hence the subtracting of 125 GeV).

Otherwise, other parameters of the fit are simply fitted in the initial global fit and then fixed - this includes the σ_{DCB}. Note that for this signal model, we do not yet take into account the distinct value $\sigma(m_{4l})$ on each event. For testing purposes, we then run this fit on the 125 GeV mass point MC samples.

The profile likelihood ratio scan in m_H involves plotting the negative logarithm of the profile likelihood ratio (see [1] for more on this quantity) $\lambda(m_H)$ against m_H. Figure 8 shows this plot for the model, as well as the estimation of the (1 standard deviation) statistical uncertainty. From this measurement, we obtain

$$m_H = 125.000^{+0.262}_{-0.267}\ GeV$$

for the 125 GeV mass point MC. This statistical uncertainty is dependent on our signal model, and there are several avenues of improvement to the signal model which may improve the statistical uncertainty. We shall call the model that

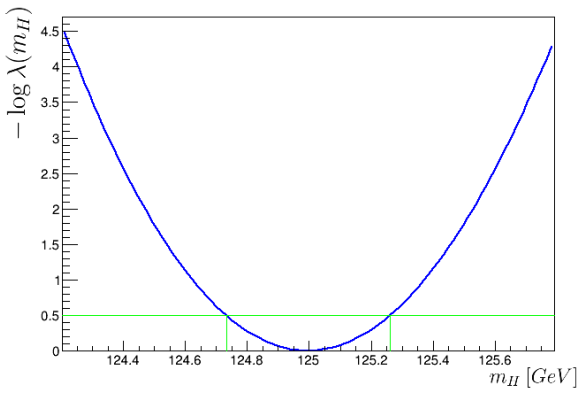


FIG. 8: a plot of $-\log\lambda$ against m_H. The minimum indicates the optimal value of m_H, and the locations of intersection with $-\log\lambda = 0.5$ mark 1 standard deviation.

produced this result, with a single fitted σ_{DCB} value, model A.

VI. MASS MEASUREMENT WITH PER-EVENT INVARIANT MASS UNCERTAINTIES

We explore two more models, where the process of mass measurement is exactly the same as before, and where the mass is parameterized as in equation 8. The only difference is in how we choose to model σ_{DCB}.

The first modification, which we shall call model B, simply assigns the estimated $\sigma(m_{4l})$ for every event to be the sigma of the double crystal ball fit for each event during the fits:

$$\sigma_{DCB} = \sigma(m_{4l}). \tag{7}$$

This is the most naive way of introducing per-event uncertainties into the mass calculation: σ_{DCB} is now no longer treated as a fixed value to drop out from the fit.

The other model tested, model C, also uses the estimated $\sigma(m_{4l})$, but instead relates it to σ_{DCB} in a more complicated way:

$$\sigma_{DCB} = \sqrt{\sigma(m_{4l})^2 + c^2}, \tag{8}$$

where c is a parameter to be determined from the fit. It is possible that the true width of the mass distribution is described by a sum in quadrature between the detector resolution and some true underlying width; the parameter c is designed to absorb any such effects.

On the 125 GeV mass point MC samples, model B yields a mass value of

$$m_H = 125.000^{+0.242}_{-0.253}\ GeV$$

and model C yields a mass value of

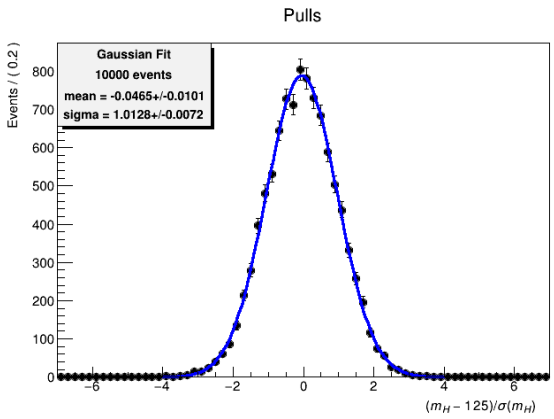
$$m_H = 125.000^{+0.259}_{-0.256}\ GeV.$$

VII. COMPARING MODEL PERFORMANCES

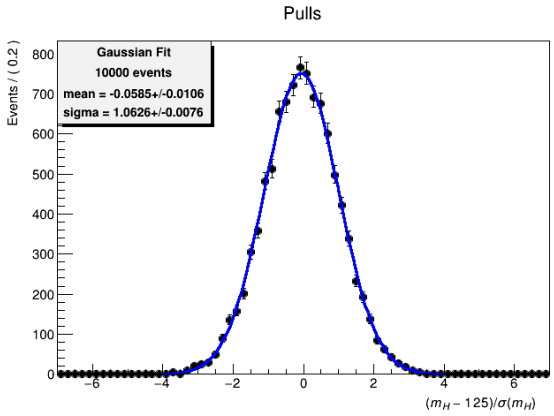
A way to test the mass measurement and its statistical uncertainty $\sigma(m_H)$ is to again generate pull plots. Given the 125 GeV mass point MC sample we did our measurement on, we randomly select a subset of events and perform the mass measurement as described in sections V and VI, yielding a mass and statistical uncertainty. Repeating this many times, we can build a pull plot which plots $(m_H - 125\ GeV)/\sigma(m_H)$. If m_H overestimates 125 GeV, we use the high-side $\sigma(m_H)$, and if m_H underestimates 125 GeV, we use the low-side $\sigma(m_H)$.

We can reasonably expect these pull plots to have a mean of 0 and a standard deviation of 1 if the uncertainty were estimated well. Nonetheless, because the quality of the pull plots is also dependent on factors like the number of events we select in our subset to generate one mass measurement, the key goal is to compare the quality of the pull plots with respect to other signal models.

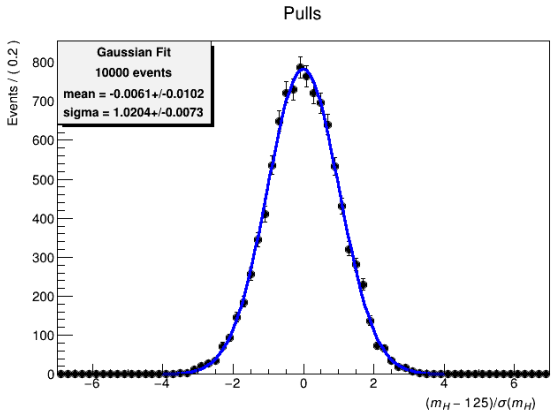
Figure 9 features the pull plots for models A, B, and C. Table I summarizes the statistical uncertainty for each model, and the pull plot parameters.



(a)



(b)



(c)

FIG. 9: pull plots for models A (a), B (b) and C (c).

We make a couple of key observations:

- model B has the best statistical uncertainty, but it also has the most bias in the pull plot mean and standard deviation - in particular, the large standard deviation suggests that the statistical uncertainty is underestimated;
- model C is a definite, if slight, improvement over model A, as it offers an improved statistical uncertainty, has a much less biased pull plot mean (within statistical uncertainty of 0), and only a slightly more biased pull standard deviation.

We may try to partially rectify any bias in uncertainty estimate. If a pull plot standard deviation is too large, it indicates that the statistical uncertainty tends to be underestimated. In particular, we may provide a correction for the statistical uncertainty given a specific bias in the standard deviation: if the standard deviation is greater than 1 by $x\%$, then we scale up the statistical uncertainty by $x\%$. These values are shown in the "corrected" statistical uncertainty column in table I. We should not, however, forget about the bias in the pull plot mean that we cannot easily correct for.

In light of these results, if we consider the nominal (not "corrected") uncertainty, we can conclude that assigning a per-event $\sigma(m_{4l})$ into the signal model appears to slightly affect the statistical uncertainty of the mass measurement. In particular, the poor performance of model B in the pull plot is more or less expected, as the naive $\sigma_{DCB} = \sigma(m_{4l})$ relationship defining this model is not expected to be accurate. Model C's performance over model A is existent but marginal, which may ask questions of the necessity of a per-event $\sigma(m_{4l})$.

If we assume that the "corrected" uncertainty is a better estimate, most of our conclusions still hold - model B still remains the best, but still suffers from the largest bias in the pull plot mean. Model C, again, offers an enticing compromise, as it both improves upon the uncertainty in model A and has (by far) the smallest bias, which is within statistical uncertainty of 0, on the pull plot mean.

Model	Mass & Statistical Uncertainty [GeV]	Mass & Corrected Statistical Uncertainty [GeV]	Pull Plot Mean	Pull Plot Std. Deviation
A	$125.000^{+0.262}_{-0.267}$	$125.000^{+0.265}_{-0.270}$	-0.047 ± 0.010	1.013 ± 0.007
B	$125.000^{+0.242}_{-0.253}$	$125.000^{+0.257}_{-0.269}$	-0.059 ± 0.011	1.063 ± 0.008
C	$125.000^{+0.259}_{-0.256}$	$125.000^{+0.264}_{-0.261}$	-0.006 ± 0.010	1.020 ± 0.007

TABLE I

Ultimately, we observe that the inclusion of per-event $\sigma(m_{4l})$ values in our signal model has a small but still improving effect on the statistical uncertainty of the mass measurement. We also see that varying the signal model has potential to improve the measurement, or to reduce bias in our pull plot analysis. Keep in mind, however, that our test of only three models is far from comprehensive, and these conclusions should be taken as indicative, not definitive.

VIII. CONCLUSIONS

In this work we constructed a way to estimate a per-event $\sigma(m_{4l})$ and demonstrated its validity; using this, we explored some signal models to help formulate a mass measurement and its uncertainty. We found that the uncertainty is likely improved by the consideration of the per-event invariant mass resolution estimate.

These results will hopefully be part of the larger ATLAS initiative to form the most precise measurement ever of the Higgs mass. While our analysis was centered on a particular set of Monte Carlo samples, the way we have tested results relative to each other and the robustness of the per-event invariant mass resolution estimation leads us to be confident that this work is applicable for other data.

Finally, there are several logical extensions to this work. They include

1. performing the same analysis, and verifying its efficacy, using the other $4l$ channels ($4e$, $2\mu 2e$, $2e2\mu$) and eventually combining them for a consolidated measurement,

2. probing further improvements to the signal model used to make the mass measurement, likely by exploring other ways in which σ_{DCB} may be related to $\sigma(m_{4l})$,

3. and improving the muon momentum resolution, by improving the heatmap (figure 1) or modelling the non-Gaussian-ness of the distribution in its bins, which may likely improve our mass estimate statistical uncertainty.

IX. ACKNOWLEDGEMENTS

Alex would like to thank Lukas Adamek and Pierre Savard for their close mentorship despite the hardships of working remotely at the height of the COVID-19 pandemic. Additionally, he would like to acknowledge everyone else in the Higgs mass measurement group at the University of Toronto: Pekka Sinervo, Bianca Ciungu, and Baria Khan. This work was done as part of the Canadian Institute of Particle Physics (IPP) CERN Summer Student Program, and supported by a NSERC USRA grant at the University of Toronto.

[1] ATLAS CONF Note *ATLAS-CONF-2020-005*, April 7, 2020.

[2] M. Tanabashi et al. (Particle Data Group), Status of Higgs Boson Physics, Phys. Rev. D 98, 030001 (2018) and 2019 update.

[3] C. Oleari, The POWHEG-BOX, July 22, 2010. arXiv:1007.3893 [hep-ph].

[4] T. Sjöstrand et al., An Introduction to PYTHIA 8.2, October 11, 2014. arXiv:1410.3012 [hep-ph].