

# A Study of Systematic Uncertainties

Matt Reader

University College London

3<sup>rd</sup> December 2024

PDF4LHC 2024 Meeting

# Objectives

- Introduce a Model that can incorporate Error on Errors
- Examining how this model behaves for both uncorrelated and correlated systematic errors
- Investigate what this model tells us about the ATLAS W,Z Data set
- A brief investigation into de-correlation

# Introduction

- Experimental data is becoming increasingly precise
- Experimental Errors are now dominated by systematic uncertainties
- Significant Errors on these systematic uncertainties
- So we need to include these "errors on errors" to correctly determine the errors in PDF fits

# Introduction

- Experimental data is becoming increasingly precise
- Experimental Errors are now dominated by systematic uncertainties
- Significant Errors on these systematic uncertainties
- So we need to include these "errors on errors" to correctly determine the errors in PDF fits
- We need to depart from the simple Gaussian treatment of errors if we want to include these "Error on Errors"

## Including Error on Errors 1 - The Gaussian Case<sup>1</sup>

- Consider a set of data,  $\mathbf{y}$ . The probability of  $\mathbf{y}$  can be written  $P(\mathbf{y}|\mu, \theta)$ , where  $\mu$  are parameters of interest and  $\theta$  are nuisance parameters that are required for the correctness of the model.

---

<sup>1</sup>Cowan arXiv:1809.05778v3

## Including Error on Errors 1 - The Gaussian Case<sup>1</sup>

- Consider a set of data,  $\mathbf{y}$ . The probability of  $\mathbf{y}$  can be written  $P(\mathbf{y}|\mu, \theta)$ , where  $\mu$  are parameters of interest and  $\theta$  are nuisance parameters that are required for the correctness of the model.
- Let  $\theta = (\theta_1, \dots, \theta_N)$  be independent Gaussian distributed values  $u = (u_1, \dots, u_N)$ , with standard deviations  $\sigma_u = (\sigma_{u_1}, \dots, \sigma_{u_N})$ :

---

<sup>1</sup>Cowan arXiv:1809.05778v3

# Including Error on Errors 1 - The Gaussian Case<sup>1</sup>

- Consider a set of data,  $\mathbf{y}$ . The probability of  $\mathbf{y}$  can be written  $P(\mathbf{y}|\mu, \theta)$ , where  $\mu$  are parameters of interest and  $\theta$  are nuisance parameters that are required for the correctness of the model.
- Let  $\theta = (\theta_1, \dots, \theta_N)$  be independent Gaussian distributed values  $u = (u_1, \dots, u_N)$ , with standard deviations  $\sigma_u = (\sigma_{u_1}, \dots, \sigma_{u_N})$ :

$$\begin{aligned} L(\mu, \theta) &= P(\mathbf{y}, \mathbf{u}|\mu, \theta) = P(\mathbf{y}|\mu, \theta)P(\mathbf{u}|\theta) \\ &= P(\mathbf{y}|\mu, \theta) \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i - \theta_i)^2 / 2\sigma_{u_i}^2} \end{aligned} \quad (1)$$

---

<sup>1</sup>Cowan arXiv:1809.05778v3

## Including Error on Errors 2 - Gaussian plus Gamma<sup>2</sup>

- Model the estimated variances,  $v_i$ , of  $\sigma_{u_i}^2$ , as Gamma distributed gives:

$$L(\mu, \theta, \sigma_{u_i}^2) = P(y|\mu, \theta) \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma_{u_i}} e^{-(u_i - \theta_i)^2 / 2\sigma_{u_i}^2} \frac{\beta_i^{\alpha_i}}{\Gamma(\alpha_i)} v_i^{\alpha_i - 1} e^{-\beta_i v_i}$$

$$\alpha_i = \frac{1}{4r_i^2} \quad \beta_i = \frac{1}{4r_i^2 \sigma_{u_i}^2}$$

- $r_i$  is defined as the relative uncertainty in the estimate of the systematic error. The parameters  $r_i$  can therefore be referred to as the "error on errors".

---

<sup>2</sup>Cowan arXiv:1809.05778v3



## Including Error on Errors 3 - t-distribution interpretation

- This model can be identically reinterpreted as a Student's t-distribution

$$L(\mu, \theta, \sigma_{u_i}^2) = P(y|\mu, \theta) \prod_{i=1}^N \frac{\Gamma(\frac{\nu_i+1}{2})}{\sqrt{\nu_i\pi}\Gamma(\nu_i/2)} \left(1 + \frac{t_i^2}{\nu_i}\right)^{-\frac{\nu_i+1}{2}} \quad (2)$$

where  $t_i = \frac{u_i - \theta_i}{\sqrt{v_i}}$  and  $\nu_i = \frac{1}{2r_i^2}$ .

- So we can treat our nuisance parameters as t-distributed!

- Consider the case:

$$y_i = d_i + \text{errors} = d_i + \sigma_i z_i + \sigma_{u_i} t_{u_i} + \sum_{j=1}^M \beta_{ij} t_j' \quad (3)$$

where for each observable  $y_i$  we have

- One statistical error  $\sigma_i$ , with a  $z_i$  that is a Normally distributed fluctuating variable.
- One uncorrelated systematic error  $\sigma_{u_i}$  with a  $t_{u_i}$  that is a t-distributed fluctuating variable with dof of  $\nu = 1/2r^2$
- M correlated systematic errors,  $\beta_{ij}$ , each with a fluctuation  $t_j'$  that are t-distributed with dof of  $\nu = 1/2r^2$ . These fluctuations are the same for all i.

## Treating our $z_i$ , $t_{u_i}$ and $t'$ as independent

- The likelihood function can be written up to some constants as:

$$L \propto \prod_{i=1}^N \exp \left[ -\frac{1}{2} \frac{(y_i - d_i - t_{u_i} \sigma_{u_i} - \sum_{j=1}^M \beta_{ij} t'_j)^2}{\sigma_i^2} \right] \left( 1 + \frac{t_{u_i}^2}{\nu} \right)^{-\frac{\nu+1}{2}}$$

$$\times \prod_{j=1}^M \left( 1 + \frac{t_j'^2}{\nu} \right)^{-\frac{\nu+1}{2}} \quad (4)$$

- Naturally leading to the Loglikelihood equation:

$$\chi^2 \equiv -2 \ln L = \sum_{i=1}^N \left( \frac{m_i - d_i - \sigma_{u_i} t_{u_i} - \sum_j \beta_{ij} t'_j}{\sigma_i} \right)^2$$

$$+ (\nu + 1) \sum_{i=1}^N \ln \left( 1 + \frac{t_{u_i}^2}{\nu} \right) + (\nu + 1) \sum_{j=1}^M \ln \left( 1 + \frac{t_j'^2}{\nu} \right) \quad (5)$$

## Treating our $z_i$ , $t_{u_i}$ and $t'$ as independent (Cont.)

- If we minimize with respect to  $t_{u_i}$  we obtain:

$$t_{u_i}^3 \frac{\sigma_{u_i}^2}{\sigma_i^2} \frac{1}{\nu} - \frac{t_{u_i}^2}{\nu} (y_i - d_i - \sum_j \beta_j t'_j) \frac{\sigma_{u_i}}{\sigma_i^2} + t_{u_i} \left( \frac{\sigma_{u_i}^2}{\sigma_i^2} + \frac{\nu + 1}{\nu} \right) - \frac{(y_i - d_i - \sum_j \beta_j t'_j) \sigma_{u_i}}{\sigma_i^2} = 0$$

- Minimizing with respect to  $t'_j$  yields

$$\sum_{i=1}^N \frac{\beta_{ij}^2}{\nu \sigma_i^2} (t'_j)^3 + \frac{D}{\nu} (t'_j)^2 + \left( \left( \sum_{i=1}^N \frac{\beta_{ij}^2}{\sigma_i^2} \right) + \frac{\nu + 1}{\nu} \right) t'_j + D = 0$$

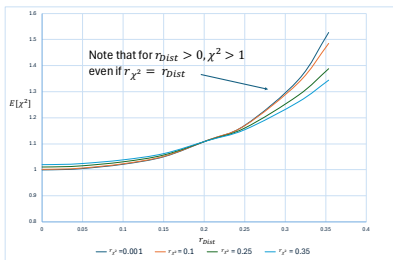
$$D = - \sum_{i=1}^N \frac{(y_i - d_i - \sigma_{u_i} t_{u_i} - \sum_{j \neq j'} \beta_{ij} t'_j) \beta_{ij'}}{\sigma_i^2}$$

- Solved Simultaneously.  $t_{u_i}$  solved analytically,  $t'_j$  has to be fitted.

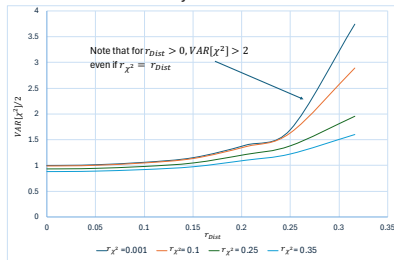
# Expectation and Variance of $\chi^2$ as a Function of $r$

- Let's consider the case:  $y_i = d_i + \sigma_i z_i + \sigma_{u_i} t_{u_i}$ ,  
with  $z_i \sim N(0, 1)$ ,  $t_{u_i} \sim t(0, \nu = 1/2r_{Dist}^2)$ .

i.e. focus on the uncorrelated systematic behaviour ( $\beta_{ij} = 0$ ).



Graph of  $E[\chi^2]$  as a Function of  $r_{Dist}$  for 4 different  $r_{\chi^2}$  ( $\sigma_i = \sigma_{u_i} = 1$ )



Graph of  $Var[\chi^2]$  as a Function of  $r_{Dist}$  for 4 different  $r_{\chi^2}$  ( $\sigma_i = \sigma_{u_i} = 1$ )

# Standard Deviation of the Simple Mean as a Function of $r$

- Consider the case:  $y_i = d_i + \sigma_i z_i + \sigma_{u_i} t_{u_i}$ ,  
with  $z_i \sim N(0, 1)$ ,  $t_{u_i} \sim t(0, \nu = 1/2r_{Dist}^2)$ . Also let  $E[d_i] = d$
- The standard deviation of the simple mean,  $y_{mean} = \sum_{i=1}^N y_i / N$ , is:

$$\sigma_{Mean} \approx \frac{\sqrt{\sum_{i=1}^N \sigma_i^2 + \sigma_{u_i}^2 \nu / (\nu - 2)}}{N} = \frac{\sqrt{\sum_{i=1}^N E[\chi_i^2(r_{\chi^2} \rightarrow 0)](\sigma_i^2 + \sigma_{u_i}^2)}}{N}$$

- Table showing  $\sigma_{mean}$  as a function of  $r$  and  $N$  (with  $\sigma_i = \sigma_{u_i} = 1$ )

$r_{Dist}$	N=2	N=3	N=5	N=10	N=100	N=500	$\frac{\sigma_{r_{Dist}}}{\sigma(r_{Dist}=0.001)}$ N=500
0.001	0.995	0.819	0.630	0.449	0.142	0.064	1.000
0.100	0.991	0.814	0.641	0.452	0.143	0.064	1.005
0.250	1.092	0.884	0.679	0.481	0.152	0.069	1.077
0.300	1.122	0.926	0.705	0.504	0.161	0.071	1.108
0.408	1.417	1.148	0.901	0.637	0.197	0.089	1.393

- This table gives the standard deviation for the fitted mean if  $r_{\chi^2} = 0.0001$

## Standard Deviation of the Fitted Mean as a Function of $r$

- What happens if we minimize the  $\chi^2$ , calculated with  $r_{\chi^2} = r_{Dist}$ , with respect to our mean?
- Table showing  $\sigma_{FIT}$  as a function of  $r$  and  $N$  (with  $\sigma_i = \sigma_{u_i} = 1$ )

$r$	N=2	N=3	N=5	N=10	N=100	N=500	$\frac{\sigma_{FIT}}{\sigma_{r_{Dist} = r_{\chi^2} = 0.001}}$ N=500
0.001	0.995	0.819	0.630	0.449	0.142	0.064	1.000
0.100	0.991	0.814	0.641	0.452	0.143	0.064	1.004
0.250	1.092	0.883	0.675	0.479	0.150	0.068	1.069
0.300	1.122	0.916	0.697	0.493	0.157	0.069	1.087
0.408	1.417	1.162	0.809	0.547	0.169	0.076	1.186

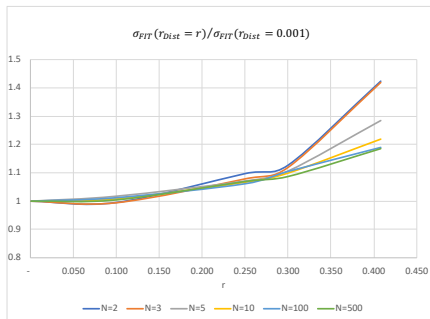
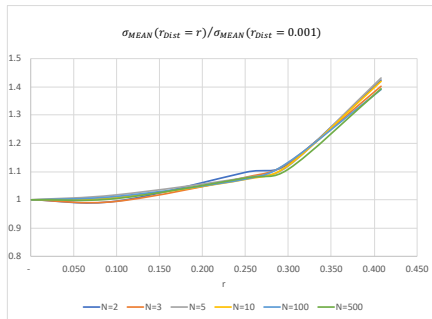
Data obtained using MC

- For Gaussian statistical errors and t-distributed uncorrelated systematic errors:

$$\sigma_{FIT} = \sqrt{\sum_{i=1}^N E[\chi_i^2(r_{\chi^2} = r_{Dist})] (\sigma_i^2 + \sigma_{u_i}^2)} / N$$

# Standard Deviation of the Mean Fitted Mean

- Graph on left shows ratio of  $\frac{\sigma_{Mean}}{\sigma_{Mean}(r_{Dist}=0.001)}$  as a function of  $r$  and  $N$ .
- Graph on right shows ratio of  $\frac{\sigma_{Fit}(r_{Dist}=r_{\chi^2=r})}{\sigma_{Fit}(r_{Dist}=r_{\chi^2=0.001})}$  as a function of  $r$  and  $N$ .





## Expectation and Variance of $\chi^2$ as a Function of $r$

- Consider the case of  $N$  observables each with a Gaussian statistical and  $M$  t-distributed correlated systematic errors :

$$y_i = d_i + \sigma_i z_i + \sum_{j=1}^M \beta_{ij} t_j' \quad z_i \sim N(0, 1), t_j' \sim t(0, \nu = 1/2r_{Dist}^2)$$

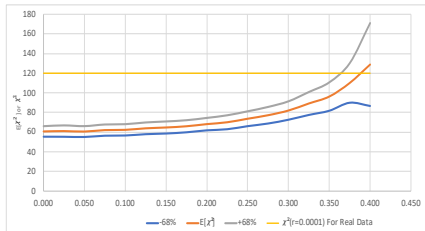
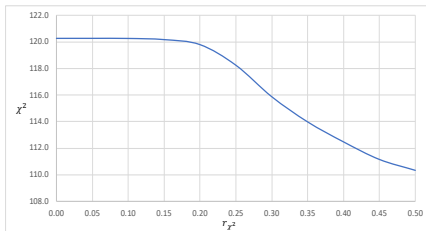
- In the case where  $r_{Dist} = r_{\chi^2}$ , and  $\sigma_i = \beta_{ij} = 1$ :

N	M	r	$\nu$	$E[\chi^2(d_i)]$	$\sigma_{\chi^2}$	$\sigma_{\varphi_{FIT}} = \sigma_{FIT}(r_{\chi^2}=0.001)$	$\sigma_{\varphi_{FIT}}$	$\sigma_{\varphi_{FIT}}/\sigma_{\varphi_{MEAN}}$	$\frac{\sigma_{\varphi_{FIT}}}{\sigma_{\varphi_{FIT}}(r_{Dist}=r_{\chi^2}=0.001)}$
2	2	0.001	500000	1.99949	2.05734	1.58114	1.58114	1.000	1.000
2	2	0.25	8	2.28779	2.31064	1.78103	1.78970	1.005	1.132
2	2	0.40824829	3	2.87634	2.98990	2.51644	2.32738	0.925	1.472
5	5	0.001	500000	4.99873	3.20452	2.28036	2.28036	1.000	1.000
5	5	0.25	8	5.42717	3.44951	2.62217	2.62208	1.000	1.150
5	5	0.40824829	3	6.53625	4.49232	3.81179	3.51314	0.922	1.541
10	5	0.001	500000	9.99746	4.58094	2.25832	2.25833	1.000	1.000
10	5	0.25	8	10.47021	4.68864	2.64291	2.63296	0.996	1.166
10	5	0.40824829	3	11.61824	5.53776	4.08220	3.48161	0.853	1.542
10	10	0.001	500000	9.99746	4.53088	3.17806	3.17806	1.000	1.000
10	10	0.25	8	10.37782	4.69332	3.67337	3.64109	0.991	1.146
10	10	0.40824829	3	11.91221	5.81006	5.40928	4.72917	0.874	1.488

where  $E[\chi^2(d_j)]$  means that  $E[\chi^2]$  calculated with a mean equal to  $d_j$

# ATLAS W,Z Data analysis<sup>3</sup>

This very precise data gives strong constraint on the strange quark.  
 However fit is poor  $\chi^2/N_{pt} \sim 1.9$  for MSHT20 (NNLO).  $N_{pt} = 61$



- Graph on the left shows the  $\chi^2$  as a function of relative error,  $r_{\chi^2}$ .
- Graph on the right shows the  $E[\chi^2]$  in the Gaussian limit as a function of relative error,  $r_{Dist}$  of the simulated underlying systematic errors.
- For a  $\chi^2$  of 120  $r_{Dist} \approx 0.4$  (Ignoring all other effects, e.g, theoretical corrections).

<sup>3</sup><https://www.hepdata.net/record/ins1502620> Tables 9 - 15

## De-correlation

- Sometimes to improve our fits we use de-correlation techniques for certain systematic errors.
- E.g. when the systematic error is calculated as the difference between two different Monte Carlo runs with different input parameters - 2 point error.
- In the nuisance parameter approach, we would break down the 1 nuisance parameter into 2 or 4 new nuisance parameters.
- For the ATLAS 8 TeV Jet Data <sup>4</sup> we consider 4<sup>5</sup> such systematic errors.
- Different methods of de-correlating these 2 point systematic errors are proposed in the ATLAS paper, the MSHT20 <sup>6</sup> fit and others <sup>7</sup>.
- The 171 data points with 330 systematic errors have a no-decorrelation  $\chi^2 \sim 400$

---

<sup>4</sup> ATLAS arXiv:1706.03192

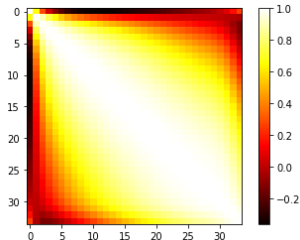
<sup>5</sup> 3 2-point errors in the data + hadronic correction which is treated in the same manner

<sup>6</sup> MSHT20 arXiv:2012.04684

<sup>7</sup> e.g. arXiv:2207.00690

## MSHT20 Correlation Interpretation

- The MSHT20 de-correlation procedure splits each 2-point error into 4 correlated systematic errors.
- Resulting correlation matrix for  $P_T$  space for the first  $y$  bin is:

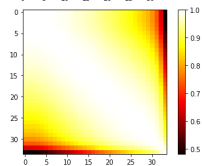
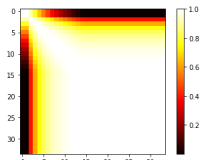


- MSHT20  $\chi^2 = 345$
- Some features to highlight:
  - Very flat correlation around the diagonal.
  - Dramatic drop in correlation towards the edges of the matrix. In fact some of the correlations become negative.

# Correlation Interpretation For The ATLAS 8 TeV Jet Data<sup>9</sup>

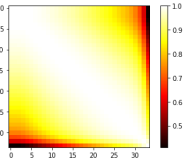
- Below are correlation matrices described by 4 different de-correlation methods described in the ATLAS paper.<sup>8</sup>

Splitting Option 5  $\chi^2 = 383$  (1  $\rightarrow$  2)



Splitting Option 17  $\chi^2 = 346$  (1  $\rightarrow$  3)

Splitting Option 16  $\chi^2 = 349$  (1  $\rightarrow$  3)



Splitting Option 18  $\chi^2 = 358$  (1  $\rightarrow$  3)

<sup>8</sup>For ease of presentation, heat maps only show the correlation structure due to the de-correlation of  $P_T$  in the lowest  $y$  bin.  $\chi^2$  numbers refer to full de-correlation including  $y$  space. 1  $\rightarrow$   $n$  refers to 1 correlated systematic error converted to  $n$  correlated errors.

<sup>9</sup>arXiv:1706.03192

## Is this what we might expect?

- All the correlation matrices exhibit a similar shape - very flat for the majority of the matrix and then fall off rapidly at the wings.
- The off diagonals are very close to 1
- It can be shown that this is a feature of a nuisance parameter approach to de-correlation, where the number of nuisance parameters is much smaller than the number of observables.
- Can we instead create a correlation matrix and use this to generate our nuisance parameters (via diagonalising)?
- The properties that we would like to see:
  - As the kinematic variables become more dis-similar we would like to see the correlation decrease in a smooth and predictable fashion.
  - We would like the correlations for neighbouring bins to slowly increase as the kinematic variables increase. E.g. we would expect the correlation for neighbouring bins at high  $P_T$  to be higher than neighbouring bins at low  $P_T$

## Alternative Approach

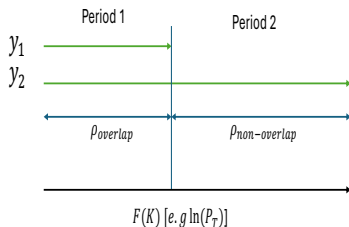
- For a given kinematic variable,  $K$ , we may want the following behaviour for a 2-point error [ $z_i \sim N(0, 1)$ ]:

$$y_1 \sim \sqrt{F(K_1)} z_1$$

$$y_2 \sim \sqrt{F(K_1)} z_2 + \sqrt{F(K_2) - F(K_1)} z_3$$

with correlations over the 2 periods:

$$\rho_{z_1, z_2} = \rho_{\text{overlap}} \quad \rho_{z_2, z_3} = \rho_{\text{non-overlap}}$$



- We can simply write down the correlation between  $y_1$  and  $y_2$  as:

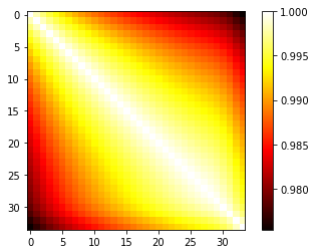
$$\rho_{1,2}(F(K)) = \frac{\rho_{\text{overlap}}(\sqrt{F(K_1)} + \rho_{\text{non-overlap}}\sqrt{F(K_2) - F(K_1)})}{\sqrt{F(K_2) + 2\rho_{\text{non-overlap}}\sqrt{F(K_1)}\sqrt{F(K_2) - F(K_1)}}}$$

## Alternative Approach (Continued)

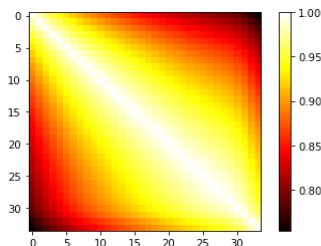
- If we now assume for simplicity that  $\rho_{1,2} = \rho_{1,2}(F(P_T)) \times \rho_{1,2}(F(y))$
- and using the simple functions:

$$F(P_T) = \text{Ln}(P_T) \quad F(y) = y$$

- We obtain the following correlation matrices in the case  $\rho_{\text{overlap}} = 1$   
 $\rho_{\text{non-overlap}} = 0.90$                        $\rho_{\text{non-overlap}} = 0$



$$\chi^2 = 329$$



$$\chi^2 = 233$$

- To match the MSHT20 de-correlated  $\chi^2 = 345$ ,  $\rho_{\text{non-overlap}} = 0.94$



## Conclusions

- We have shown how we can incorporate Errors on Errors into the calculation of a  $\chi^2$
- Expected  $\chi^2$  and Variance of  $\chi^2$  increase as the relative errors of the systematic errors increase
- We have noted that for ATLAS W,Z data set  $r \approx 0.4$ . If we compare the expected standard deviation of the mean, calculated with  $r_{\chi^2} = 0.001$ , where the systematic errors have an  $r_{Dist} = 0.4$  to that where  $r_{Dist} = 0.001$  we obtain a ratio of approx. 1.2 – 1.5. This is suggestive of using a tolerance,  $T^2$  in the region of 1.5 – 2 in this example.
- When de-correlating systematic errors we need to be incredibly careful about the correlation matrix that we are producing

## Back Up Slides

## Behaviour as $r \rightarrow 0$

- As  $r \rightarrow 0$ ,  $v \rightarrow \infty$  :

$$-2\text{Ln}L = \sum_{i=1}^N \left( \frac{y_i - d_i - \sigma_{u_i} t_{u_i} - \sum_j \beta_{ij} t'_j}{\sigma_i} \right)^2 + \sum_{i=1}^N t_{u_i}^2 + \sum_j^M t'^2_{j=1} \quad (6)$$

- Taking derivatives with respect to  $t_{u_i}$  and setting to zero gives:

$$\frac{\partial(-2\text{Ln}L)}{\partial t_{u_i}} = 0 \implies t_{u_i} = \frac{\sigma_{u_i} (y_i - d_i - \sum_j \beta_{ij} t'_j)}{\sigma_i^2 + \sigma_{u_i}^2} \quad (7)$$

- Substituting equation (7) into (6) gives our expected Gaussian formulation:

$$-2\text{Ln}L = \sum_{i=1}^N \frac{(y_i - d_i - \sum_j \beta_{ij} t'_j)^2}{\sigma_i^2 + \sigma_{u_i}^2} + \sum_{j=1}^M t'^2 \quad (8)$$

## Treating our $r_i$ , $r_{u_i}$ and $r'$ as 0 Correlated - Model 2

- In this scenario we obtain the Loglikelihood equation:

$$\begin{aligned}
 -2\text{Ln}L &= \sum_{i=1}^N \left( \frac{y_i - t_i - \sigma_{u_i} r_{u_i} - \sum_j \beta_{ij} r'_j}{\sigma_i} \right)^2 \\
 &+ \sum_{i=1}^N (\nu + 1) \text{Ln} \left( 1 + \frac{r_{u_i}^2}{\nu} \right) + (\nu + M) \text{Ln} \left( 1 + \frac{\sum_j^M r_j'^2}{\nu} \right) \quad (9)
 \end{aligned}$$

- Again we obtain cubic equations for  $r_{u_i}$  and  $r'_j$ .
- The equation for  $r_{u_i}$  can be solved analytically whilst solving for  $r'_j$  numerically.
- Model 1 and Model 2 have a different dependence on  $\nu$ . In order to make Model 1 and Model 2 closer we shall let  $\nu \rightarrow M\nu$ .

# Decorrelation Procedure For MSHT20<sup>10</sup> for the 8 TeV Data

$$x_{p_{\perp}} = \frac{\log(p_{\perp}^j) - \log(p_{\perp,min}^j)}{\log(p_{\perp,max}^j) - \log(p_{\perp,min}^j)}$$

$$x_y = \frac{y_j - y_{j,min}}{y_{j,min} - y_{j,max}}$$

$$r = \frac{1}{\sqrt{2}} \sqrt{x_{p_{\perp}}^2 + x_y^2}, \quad \phi = \arctan\left(\frac{x_y}{x_{p_{\perp}}}\right)$$

$$L_{trig}(z, z_{min}, z_{max}) = \cos\left[\pi\left(\frac{z - z_{min}}{z_{max} - z_{min}}\right)\right]$$

$$\beta_i^{(1)} = L_{trig}(r, 0, 1) \cdot L_{trig}\left(\phi, 0, \frac{\pi}{2}\right) \beta_i^{tot}$$

$$\beta_i^{(2)} = \sqrt{1 - L_{trig}(r, 0, 1)^2} \cdot L_{trig}\left(\phi, 0, \frac{\pi}{2}\right) \beta_i^{tot}$$

$$\beta_i^{(3)} = L_{trig}(r, 0, 1) \cdot \sqrt{1 - L_{trig}\left(\phi, 0, \frac{\pi}{2}\right)^2} \beta_i^{tot}$$

$$\beta_i^{(4)} = \sqrt{1 - L_{trig}(r, 0, 1)^2} \cdot \sqrt{L_{trig}\left(\phi, 0, \frac{\pi}{2}\right)^2} \beta_i^{tot}$$

<sup>10</sup>MSHT20 arXiv:2012.04684

# Decorrelation Procedure proposed by ATLAS for the 8 TeV Data

- $L(x, min, max) = (x - min) / (max - min)$ , for  $x$  in the range  $[min, max]$ ,  $L(x, min, max) = 0$  for  $x < min$ ,  
 $L(x, min, max) = 1$  for  $x > max$
- Splitting Option 5  
 $L((\ln(p_T[TeV]))^2, (\ln(0.1))^2, (\ln(2.5))^2) \times \text{uncertainty}$
- Splitting Option 16  
 $\frac{\sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} \sqrt{1 - L(|y|, 0, 1.5)^2}}{\text{uncertainty} \sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} L(|y|, 1.5, 3)} \times \text{uncertainty}$
- Splitting Option 17  
 $\frac{\sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} \sqrt{1 - L(|y|, 0, 1)^2}}{\text{uncertainty} \sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} L(|y|, 1, 3)} \times \text{uncertainty}$
- Splitting Option 18  
 $\frac{\sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} \sqrt{1 - L(|y|, 0, 2)^2}}{\text{uncertainty} \sqrt{1 - L(\ln(p_T[TeV]), \ln(0.1), \ln(2.5))^2} L(|y|, 2, 3)} \times \text{uncertainty}$
- An extra (complementary) sub-component completes them, such that the sum in quadrature of all the sub-components in each splitting option equals the original uncertainty

# $\chi^2$ as Function of the relative $r$ for the Different ATLAS Decorrelation proposals <sup>11</sup>

$r$	1	2	3	4	5	6	7	8	9	10
0.00001	391.24	397.66	398.94	400.21	383.21	392.39	390.45	398.23	395.46	392.49
0.1	389.18	395.99	397.31	398.50	380.86	390.38	388.56	396.52	393.04	390.92
0.2	381.42	392.21	393.76	394.76	375.55	385.45	385.02	393.18	387.17	385.45
0.3	357.35	370.75	389.15	389.98	369.32	376.32	381.49	389.35	376.54	361.19
0.4	336.21	345.11	373.27	373.32	363.50	364.05	374.03	375.72	361.46	333.28
0.5	312.76	327.64	362.44	362.39	344.40	347.49	369.40	363.66	346.35	293.64
0.6	295.49	314.50	349.10	348.36	332.12	333.27	362.11	359.21	329.03	281.08
0.7	288.32	309.35	340.53	339.60	325.33	321.76	359.35	351.07	324.11	273.56

$r$	11	12	13	14	15	16	17	18	MSHT20
0.00001	398.86	397.99	387.73	384.07	391.90	349.33	346.33	358.35	344.66
0.1	396.85	396.44	387.38	383.67	390.97	348.19	344.02	357.11	342.45
0.2	392.43	392.92	386.27	382.01	388.54	345.49	337.91	355.03	335.61
0.3	386.57	385.46	383.39	376.68	382.83	343.03	330.35	353.95	321.29
0.4	369.81	347.67	369.02	353.47	364.49	339.29	312.46	349.59	306.77
0.5	358.33	330.31	352.00	341.91	346.48	327.21	303.16	344.61	298.41
0.6	341.59	308.26	345.22	332.68	324.46	311.64	298.76	340.36	291.65
0.7	331.59	297.90	341.38	324.29	313.11	303.87	296.59	332.46	284.49

<sup>11</sup> ATLAS arXiv:1706.03192

# First 9 points of the Correlation Matrix

■ Correlation For  $\rho_{overlap} = 1, \rho_{non-overlap} = 0.9$

	1	2	3	4	5	6	7	8	9
1	1	0.99748009	0.99569046	0.99418454	0.99300046	0.99193164	0.99096311	0.99014979	0.98939283
2	0.99748009	1	0.99773847	0.99601303	0.99469876	0.99353263	0.9924881	0.99161806	0.99081313
3	0.99569046	0.99773847	1	0.99785259	0.99635848	0.99507089	0.9939368	0.99300232	0.99214428
4	0.99418454	0.99601303	0.99785259	1	0.99813843	0.99667175	0.99541903	0.99440392	0.99348175
5	0.99300046	0.99469876	0.99635848	0.99813843	1	0.99820989	0.99680264	0.99569264	0.99469896
6	0.99193164	0.99353263	0.99507089	0.99667175	0.99820989	1	0.99828798	0.99704117	0.99595385
7	0.99096311	0.9924881	0.9939368	0.99541903	0.99680264	0.99828798	1	0.99848236	0.99725955
8	0.99014979	0.99161806	0.99300232	0.99440392	0.99569264	0.99704117	0.99848236	1	0.99853457
9	0.98939283	0.99081313	0.99214428	0.99348175	0.99469896	0.99595385	0.99725955	0.99853457	1

■ Correlation For  $\rho_{overlap} = 1, \rho_{non-overlap} = 0.0$

	1	2	3	4	5	6	7	8	9
1	1	0.98219624	0.96674123	0.95239876	0.94037583	0.92901594	0.91834121	0.90911878	0.90033787
2	0.98219624	1	0.98426484	0.9696624	0.95742153	0.94585573	0.9349875	0.9255979	0.91665782
3	0.96674123	0.98426484	1	0.98516411	0.97272755	0.96097685	0.94993488	0.94039517	0.93131216
4	0.95239876	0.9696624	0.98516411	1	0.98737616	0.9754485	0.96424024	0.95455687	0.94533708
5	0.94037583	0.95742153	0.97272755	0.98737616	1	0.98791984	0.97656828	0.96676111	0.95742344
6	0.92901594	0.94585573	0.96097685	0.9754485	0.98791984	1	0.98850964	0.97858254	0.9691307
7	0.91834121	0.9349875	0.94993488	0.96424024	0.97656828	0.98850964	1	0.98995751	0.9803958
8	0.90911878	0.9255979	0.94039517	0.95455687	0.96676111	0.97858254	0.98995751	1	0.99034129
9	0.90033787	0.91665782	0.93131216	0.94533708	0.95742344	0.9691307	0.9803958	0.99034129	1



# Using Off-Diagonal plus serial correlation to define the Correlation Matrix

- One can define the off diagonal and use serial correlation to define the rest of the correlation matrix
- This is exactly the same as building a variate space where one only correlates neighbouring points (i.e. those on the off-diagonal), and infers the other correlations.
- For the First 9 points you would obtain the following correlation matrix if you defined the off-diagonal to be  $\rho_{n,n+1} = 0.97 + 0.002(n - 1)$ :

	1	2	3	4	5	6	7	8	9
1	1	0.97	0.94284	0.91832616	0.89628633	0.87656803	0.85903667	0.84357401	0.83007683
2	0.97	1	0.972	0.946728	0.92400653	0.90367838	0.88560482	0.86966393	0.85574931
3	0.94284	0.972	1	0.974	0.950624	0.92971027	0.91111607	0.89471598	0.88040052
4	0.91832616	0.946728	0.974	1	0.976	0.954528	0.93543744	0.91859957	0.90390197
5	0.89628633	0.92400653	0.950624	0.976	1	0.978	0.95844	0.94118808	0.92612907
6	0.87656803	0.90367838	0.92971027	0.954528	0.978	1	0.98	0.96236	0.94696224
7	0.85903667	0.88560482	0.91111607	0.93543744	0.95844	0.98	1	0.982	0.966288
8	0.84357401	0.86966393	0.89471598	0.91859957	0.94118808	0.96236	0.982	1	0.984
9	0.83007683	0.85574931	0.88040052	0.90390197	0.92612907	0.94696224	0.966288	0.984	1

As can be seen this exhibits many of the features we were hoping for.

# De-correlation For 7 TeV Jet Data <sup>12</sup>

- Sometimes we use decorrelation techniques for systematic errors.
- E.g. would be when a systematic error is calculated as the difference between two different Monte Carlo runs with different input parameters.
- For the ATLAS 7 TeV Jet Data, as is done in the MSHT20 paper, we could use a parameterisation which allow data points that are distant in  $(y_j, p_{\perp}^j)$  space to have different systematic variations:

$$x_{p_{\perp}} = \frac{\log(p_{\perp}^j) - \log(p_{\perp, \min}^j)}{\log(p_{\perp, \max}^j) - \log(p_{\perp, \min}^j)}$$

$$x_y = \frac{y_j - y_{j, \min}}{y_{j, \min} - y_{j, \max}}$$

$$r = \frac{1}{\sqrt{2}} \sqrt{x_{p_{\perp}}^2 + x_y^2}, \quad \phi = \arctan\left(\frac{x_y}{x_{p_{\perp}}}\right)$$

$$L_{\text{trig}}(z, z_{\min}, z_{\max}) = \cos\left[\pi\left(\frac{z - z_{\min}}{z_{\max} - z_{\min}}\right)\right]$$

- If we implement this for the , then if  $r = 0.0001$ , then the  $\chi^2 = 243.43$ .

$$\beta_i^{(1)} = L_{\text{trig}}(r, 0, 1) \cdot L_{\text{trig}}\left(\phi, 0, \frac{\pi}{2}\right) \beta_i^{\text{tot}}$$

$$\beta_i^{(2)} = \sqrt{1 - L_{\text{trig}}(r, 0, 1)^2} \cdot L_{\text{trig}}\left(\phi, 0, \frac{\pi}{2}\right) \beta_i^{\text{tot}}$$

$$\beta_i^{(3)} = L_{\text{trig}}(r, 0, 1) \cdot \sqrt{1 - L_{\text{trig}}\left(\phi, 0, \frac{\pi}{2}\right)^2} \beta_i^{\text{tot}}$$

$$\beta_i^{(4)} = \sqrt{1 - L_{\text{trig}}(r, 0, 1)^2} \cdot \sqrt{L_{\text{trig}}\left(\phi, 0, \frac{\pi}{2}\right)^2} \beta_i^{\text{tot}}$$

<sup>12</sup><https://www.hepdata.net/record/ins1325553-tables-7-12-for-R=0.6>





## Closer Look at the Correlation Matrix - $\pi/2$ versus $\pi/4$ (ATLAS 7 TeV Data)

In some parts of the correlation matrix using  $\pi/2$  looks better, but in other parts it looks worse.

Worse		Better									
	x (P T)	0.02572464	0.23994183	0.51009257	0.74286895	0.95880057	0.02572464	0.23994183	0.51009257	0.74286895	0.93724088
x (P T)	x y	0.08333333	0.08333333	0.08333333	0.08333333	0.08333333	0.25	0.25	0.25	0.25	0.25
0.02572464	0.08333333	0.9994554	-0.2784397	-0.3488771	-0.0705975	0.26125823	0.86256755	0.50062405	-0.0291488	-0.0074044	0.16565448
0.23994183	0.08333333	-0.2784397	0.9994555	0.78504325	0.41223169	-0.0025533	-0.6418794	0.57461627	0.74366462	0.38358978	-0.0195558
0.51009257	0.08333333	-0.3488771	0.78504325	0.9997457	0.86694258	0.54252776	-0.7172612	0.25926028	0.82672441	0.78554601	0.52416333
0.74286895	0.08333333	-0.0705975	0.41223169	0.86694258	0.99974561	0.88687464	-0.4106197	0.144004	0.71176327	0.90750779	0.84133983
0.95880057	0.08333333	0.26125823	-0.0025533	0.54252776	0.88687464	0.99974552	0.00822859	0.02654636	0.47350771	0.81963561	0.93970072
0.02572464	0.25	0.86256755	-0.6418794	-0.7172612	-0.4106197	0.00822859	0.99974556	0.23816247	-0.3345382	-0.2483446	0.0192284
0.23994183	0.25	0.50062405	0.57461627	0.25926028	0.114004	0.02654636	0.23816247	0.99974544	0.67364362	0.32225615	0.08530261
0.51009257	0.25	-0.0291488	0.74366462	0.82672441	0.71176327	0.47350771	-0.3345382	0.67364362	0.9997456	0.85680031	0.57989632
0.74286895	0.25	-0.0074044	0.38358978	0.78554601	0.90750779	0.81963561	-0.2483446	0.32225615	0.85680031	0.99974561	0.90795241
0.93724088	0.25	0.16565448	-0.0195558	0.52416333	0.84133983	0.93970072	0.0192284	0.08530261	0.57989632	0.90795241	0.99974551

$\pi/2$

Better		Worse									
	x (P T)	0.02572464	0.23994183	0.51009257	0.74286895	0.95880057	0.02572464	0.23994183	0.51009257	0.74286895	0.93724088
x (P T)	x y	0.08333333	0.08333333	0.08333333	0.08333333	0.08333333	0.25	0.25	0.25	0.25	0.25
0.02572464	0.08333333	0.99974556	0.9228621	0.49272974	0.07584232	-0.240455	0.65907332	-0.2405758	0.39051862	0.02322744	-0.3757966
0.23994183	0.08333333	0.9228621	0.99974549	0.64364632	0.28738736	-0.0015913	0.60011925	-0.1478024	0.6775356	0.38337488	-0.0189196
0.51009257	0.08333333	0.49272974	0.64364632	0.99974564	0.85380853	0.52406367	0.80746022	-0.6913689	0.38290406	0.65961034	0.4934031
0.74286895	0.08333333	0.07584232	0.28738736	0.85380853	0.99974562	0.88354433	0.45112934	-0.5417014	0.17837245	0.65642429	0.72844584
0.95880057	0.08333333	-0.240455	-0.0015913	0.52406367	0.88354433	0.99974552	-0.0088366	-0.1829543	0.06056808	0.535064	0.76717266
0.02572464	0.25	0.65907332	0.60011925	0.80746022	0.45112934	-0.0088366	0.99974556	-0.8614181	0.12004887	0.23849793	-0.0207453
0.23994183	0.25	-0.2405758	-0.1478024	-0.6923689	-0.5427014	-0.1829543	-0.8614181	0.99974559	0.28894148	-0.1066135	-0.0796406
0.51009257	0.25	0.39051862	0.6775356	0.38290406	0.17837245	0.06056808	0.12004887	0.28894148	0.99974547	0.76794774	0.44568291
0.74286895	0.25	0.02322744	0.38337488	0.65961034	0.65642429	0.535064	0.23849793	-0.1066135	0.76794774	0.99974561	0.88563822
0.93724088	0.25	-0.3757966	-0.0189196	0.4934031	0.72844584	0.76717266	-0.0207453	-0.0796406	0.44568291	0.88563822	0.99974545

$\pi/4$