



Dealing with overlap of events

Tomas Dado (Dortmund)

(Re)interpretation of the LHC results for new physics, 14/12/2022

Overlaps

- Providing full likelihood = maximal information about a given distribution
 - But not enough for combinations
- Different types of correlations/overlaps
 - Overlap in data events data statistical correlations
 - Correlations in the nuisance parameters (systematics)
 - Correlations of MC statistical uncertainties
 - **Statistical correlations of the nuisance parameter measurements** (not correlations of the nuisance parameters!)
- **RED** ignoring for now too much effort for a little gain (usually)
- BLUE
 - Ideally: trivial (same name = same parameter, different name = different parameter)
 - In reality: correlations are probably not 0 or 100% even experiments struggle with this
- **GREEN** There is a sound technique to use!



Dealing with data statistical overlap

- <u>Bootstraps</u>
 - **Create pseudo-data** from real data following a precise recipe
 - N poisson random numbers for each data event
 - Use **predefined seed** (run)event numbers
 - Same event = same N random numbers!
- Documentation from ATLAS statistical forum
 - C++ tools (extension of ROOT::TH1) for bootstraps
- How to use it?
 - Can estimate correlations of POIs between measurements/searches
 - Check if the results are uncorrelated
 - Use the correlations in the statistical model
 - Eg. gaussian approximation





Storing bootstrap replicas on HEPData

- <u>A simple table</u>
 - **Column** = bin on the detector level
 - **Row** = bootstrap replica
- Allows to use only partial data, e.g. only SRs
- If full likelihood is provided allows to run the fit on

replicas

- Can get stat correlations
- This was already <u>done</u> in ATLAS at 7 TeV!
- Should this become a common practice?



