Science & Technology Facilities Council

UK Research and Innovation

Forward Folding Issues and Ideas

Lukas Koch

STFC Rutherford Appleton Laboratory

on behalf of the T2K collaboration

NuInt 18, GSSI Italy, October 2018

The cross-section cave

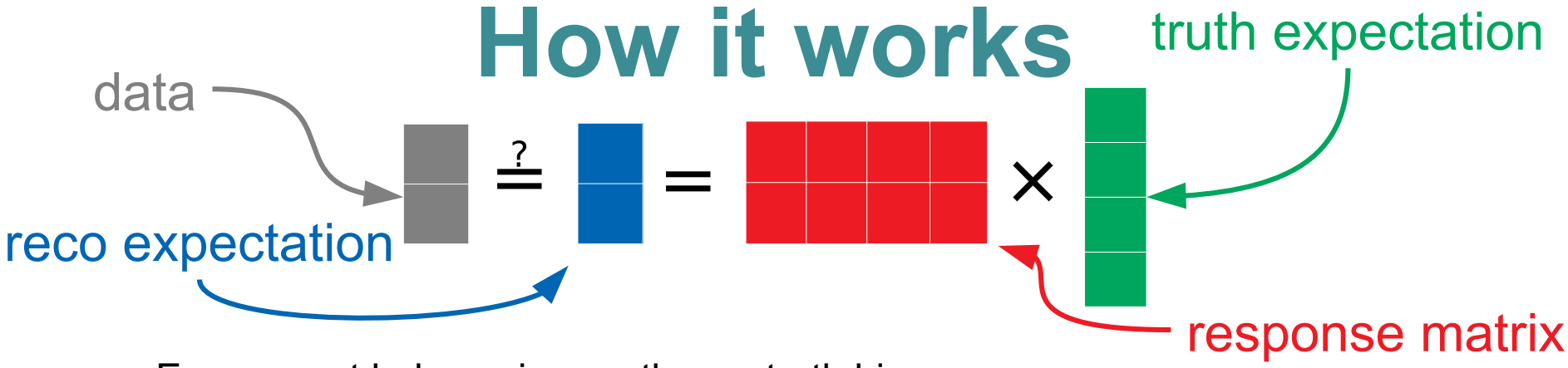
- What we see is not what we are interested in
  - Lost events due to efficiency
  - Added events due to background
  - Different event properties due to smearing

cross-section extraction

- The canonical way: Unfolding
  - "Undo" the detector and selection effects
  - Challenging to do right without introducing bias

- Another way: Forward-folding
  - Apply detector effects to theory
  - Brings its own sets of challenges

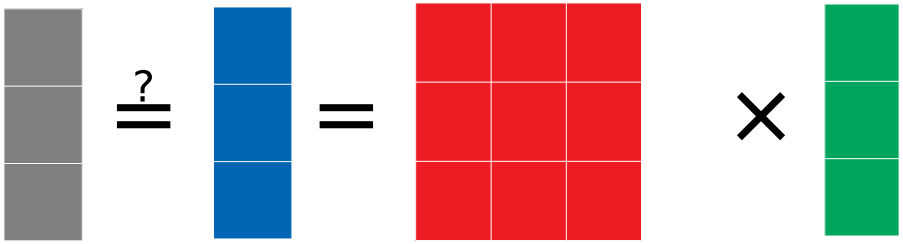
How it works

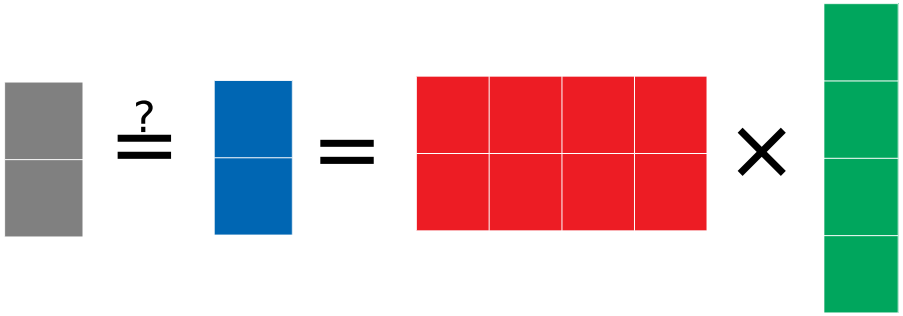


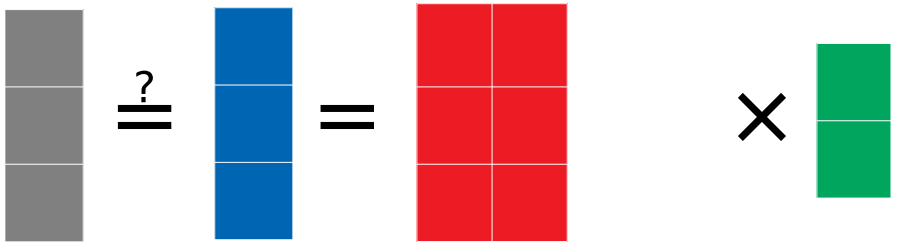
- Every event belongs in exactly one truth bin
  and up to one reconstructed bin (if it gets reconstructed)
- P(reco bin = i | truth bin = j) = R_{ij} = efficiency × smearing
  - Response matrix describes average detector response to true events
- reco expectation = response matrix × truth expectation
  - Can (and truth usually must) be binned in multiple variables
- The data is the data is the data
  - No uncertainty on the data points, 4 is exactly 4!
  - All systematics in response matrix or physics model
- All comparisons between data and theory (likelihoods, chi-squares, chi-by-eye) are done in reco space.

Reasons to do it: Flexible binning

- Flexible number of bins
- reco bins ≠ truth bins

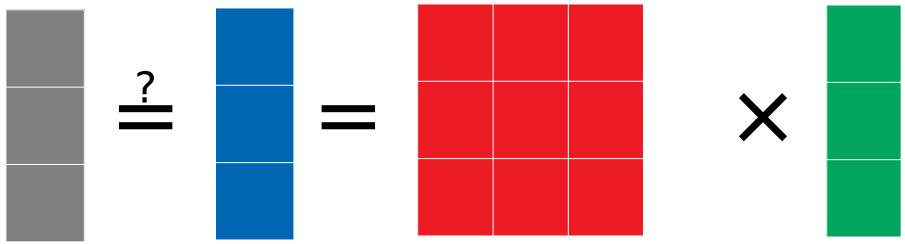


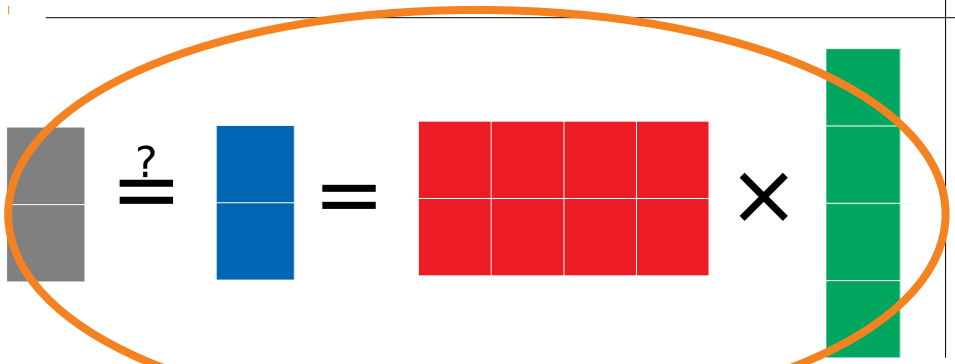


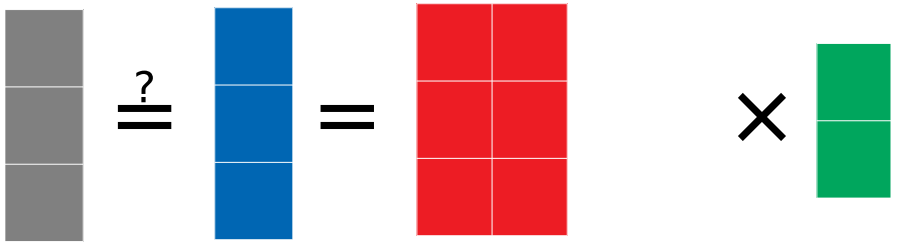


Reasons to do it: Flexible binning

- Flexible number of bins
- reco bins ≠ truth bins







- Combine coarse reco binning with fine truth binning
  - Great for analyses with low statistics
  - Admit we are not able to constrain truth completely

Reasons to do it: reco level data

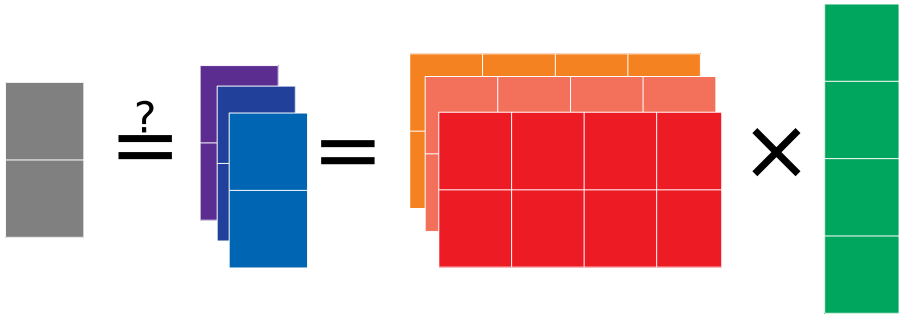
- No data point correlation
  - Theory predictions will be correlated, but probably much less than what unregularised unfolding might do
  - Chi-by-eye
- Robert D. Cousins, Samuel J. May, Yipeng Sun, [arXiv:1607.07038]
  Should unfolded histograms be used to test hypotheses?:

  "It seems remarkable that, even though unfolding by matrix inversion would appear not to lose information, in practice the way the information is used (linearizing the problem via expressing the result via a covariance matrix) already results in some failures of the bottom-line test of GOF. This is *without any regularization or approximate EM inversion*."

"D'Agostini"

Detector uncertainties

- Matrix only describes single possible detector
  - True detector probably behaves slightly differently
- Cover detector uncertainties with “toy simulations”
  - Variations and weights of same events
- Each toy yields own response matrix
- Each response matrix yields own reco prediction
- Compare to data w/ marginal, i.e. average, likelihood

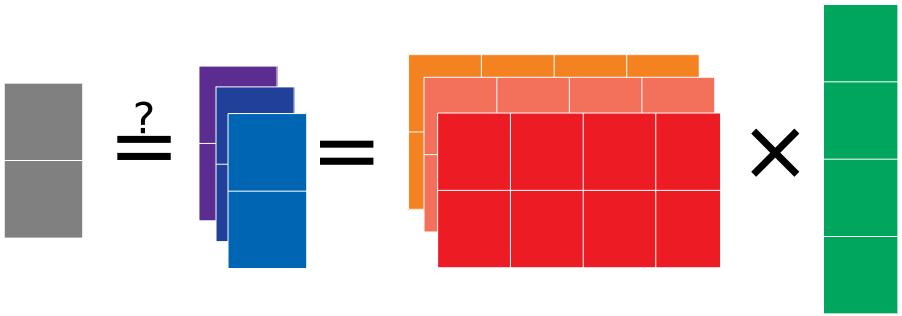


The data usage cycle

Experiment takes data, builds response matrix

Use for likelihood tests to reject models / fit parameters

Inside experiment collaboration



Publish parameter fits as well as raw data and matrix

Fits inform building of new models

Use published data and matrix for likelihood tests to reject models / fit parameters

Response Matrix Utilities

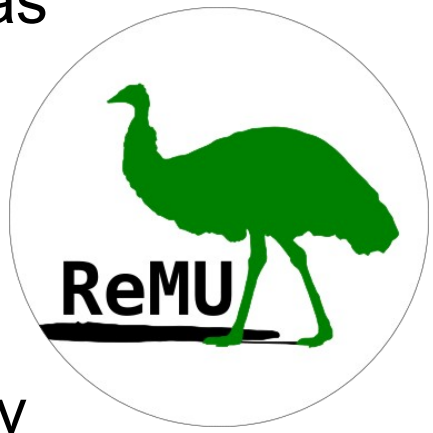
- Implements all of this (and then some)
- Input:
  - Toy variations of selection (detector systematics)
  - Truth and reco binning
- Provides methods to:
  - Build matrix
  - Evaluate statistical detector uncertainty
  - Forward-fold truth (i.e. model)
  - Compare to data (e.g. compute likelihoods, p-values, MCMC)
- Pure python (+ standard scientific packages numpy, etc)
  - Easy to install and use
  - $ pip install remu
- Tell me what you expect/want/need!

The hard part for the analyser

- Make the response matrix model-independent!
- What:
  - Matrix elements depend *only* on detector properties
- Why:
  - A model-dependent matrix defeats the purpose of being able to test arbitrary models with it
- How:
  - Understand your detector and analysis
  - Choose an appropriate truth binning (variables to bin in, granularity of binning)
  - See backup slides

To conclude

- Forward folding is every bit as challenging as unfolding
  - Need to really understand the detector to decide which variables to bin in and how
  - High MC statistics requirements
- ReMU implements the necessary machinery
- Method promises some advantages over unfolding
  - Works with low real data statistics
  - Best model separation power in reco space [arXiv:1607.07038]
- Method paper in preparation



A few things to think about

- How to best handle backgrounds
  - Backgrounds are just another set of truth bins
  - Can be handled organically (simultaneous fits)
- Provide background templates
- Provide experiment/analysis specific convenience functions
- Plot release → data release → algorithm release
  - $ pip install t2k-results

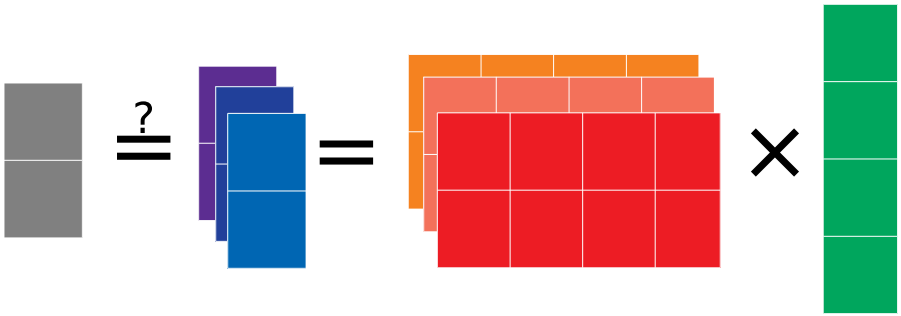
    >>> t2k_results.thisorthat_xsec_result.fit(my_model)
  - No more manual overlaying plots copied from papers

Thank you!

https://remu.readthedocs.io https://github.com/ast0815/remu

Statistical uncertainty

- Generating MC costs time and money
  - In theory we could have arbitrarily precise matrices
  - In practice we don't
- Quantify statistical uncertainty of matrix elements
- Generate random matrices according to stats
- Handle just like systematic uncertainties
  - In a way the statistical uncertainty is just another detector systematic



Three step matrix building model

- $R_{ij} = eff_j \times smear_{ij} \times weight_{ij}$
- Efficiency
  - Binomial process
  - Parameters ~ conjugate distribution: Beta
- Smearing
  - Multinomial process
  - Parameter ~ conjugate distribution: Dirichlet
- Weighting
  - What matters are the average weights
  - Use standard error of the mean: Normal

What to bin in

- Ideal:
  - Bin in *all* truth variables that affect reconstruction
- This goes beyond the variables of physical interest, i.e reco variables!
  - Measuring muon momentum distribution, but true cos(theta) affects efficiency? You *must* bin in true cos(theta)!
  - Might lead down some weird rabbit holes (angular separation of tracks, total particle multiplicity, ...)
- Realistic:
  - Bin in *most important* variables that affect reconstruction

What *not* to bin in

- Never ever use truth variables that need a "physics" model to propagate to the reco level!
- Neutrino energy? Bad choice.
  - Measurable effect depends on interaction model, nuclear model, FSI…
- Muon momentum? Good choice.
  - Directly accessible by detector (track curvature)
- HMN momentum? Even better choice!
  - Do you assume the muon to be selected as HMN?
  - What about confusion with high-momentum pions?
- Rule of thumb:
  - Bin in variables as "close" to low-level reconstructed quantities
  - "Could you see it in an event display?"

The exponential #bins problem

- #bins = (#bins/variable) ^ (#variables)
- MC stats are cheap (compared to data) but not free
- Need to compromise
  - Bin coarsely (but beware in-bin variations!)
  - Concentrate on most important truth variables
  - Reduce #reco bins
  - #response matrix bins = #truth bins × #reco bins
- Aim: Reduce model-dependence to a negligible level
  - Will never remove it completely
- Constraint: Sufficient MC events in bins

The bitter truth

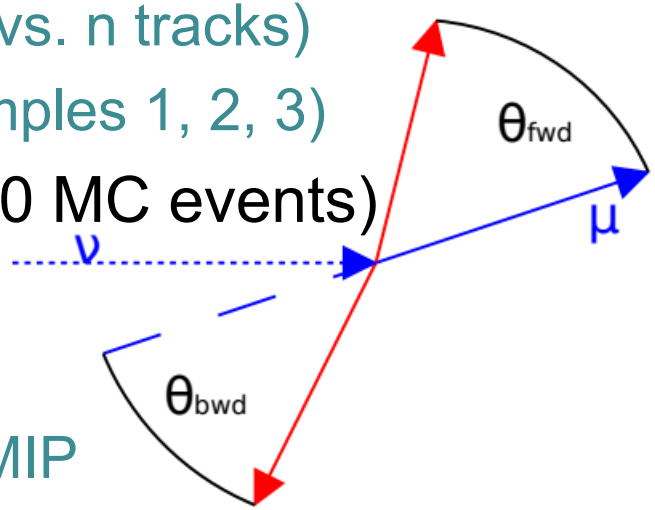
- There will be truth bins with not enough events
- Constrains the phase space of testable models
  - $n_{test} < (n_{generated} / \text{safety factor})$ in all truth bins
- Best way to avoid this:
  - Build response matrix with MC covering the full phase space (“particle parties”)
- Realistic way to mitigate this:
  - Build response matrix with MC from multiple generators, turn dials to widest possible phase space
- Response matrix depends only on detector properties
  - Mix and merge all the models!

Gas interaction example (WIP)

- Reco binning: 16 bins
  - 2 bins in main MIP (muon or pion) angle (forwards vs. backward)
  - 2 bins in particle multiplicity (1 track vs. n tracks)
  - 4 bins in selection (main, control samples 1, 2, 3)

Gas interaction example (WIP)

- Reco binning: 16 bins
  - 2 bins in main MIP (muon or pion) angle (forwards vs. backward)
  - 2 bins in particle multiplicity (1 track vs. n tracks)
  - 4 bins in selection (main, control samples 1, 2, 3)
- Truth binning: 11760 bins (5353 w/ >0 MC events)
  - 7 bins in true MIP momentum
  - 7 bins in true MIP cos(theta)
  - 5 bins in true forward separation of MIP
  - 6 bins in true backwards separation of MIP
  - 8 bins in event category (4 in FV + 4 out of FV)



The other hard part

- Getting everyone on board to use this
  - This will mean extra work for theorists/model-builders
- But it is worthwhile
  - Better model separation power
  - Works with low statistics
  - Endorsed by actual statisticians!
- This is not just dumping work on theorists
  - This is hard for experimentalists too!
  - Have to work together for better physics results
- Make this as painless as possible
  - There will be *some* pain…
  - Tell me what you want/need/expect!