GNN fit robustness II

06.09.24

"Mixing"

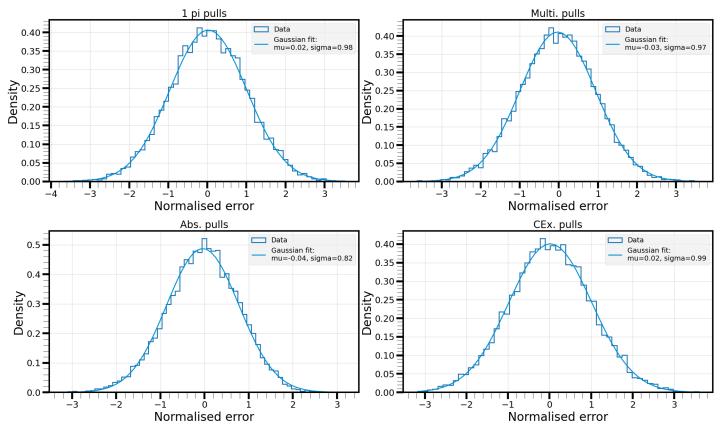
- Same source for all MC draw from "true" distribution
- Previously, this was divided into two samples (files).
- From these, sub-samples were created.
- Treating the "true" distribution as a multinomially distributed set of bins.
- When sampling from the subsample, we sample with a probability $\frac{X}{N_{sub}} = p_{i,sub} \neq p_{i,true}$, where $X \sim \text{Multi}(N_{full}, p_{i,true})$ from the true multinomial.

"Mixing"

- Before showing any results:
- How large a pull is acceptable?
 - I've already looked at a bunch of results, so don't trust myself to think about it sensibly

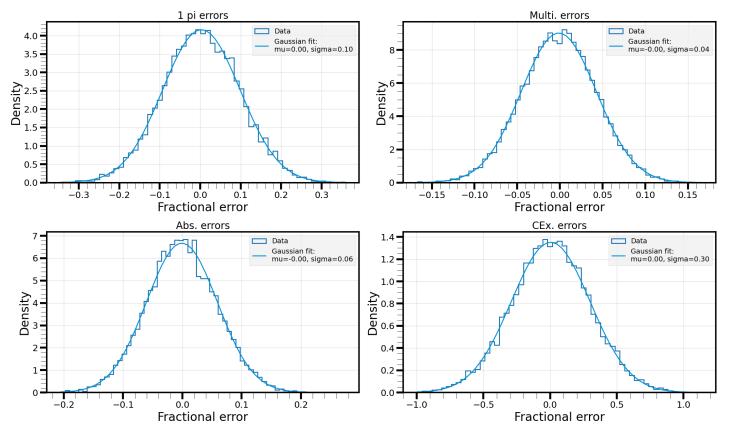
"Mixing" pulls

 Combining all MC events, and randomly splitting these removes pulls.



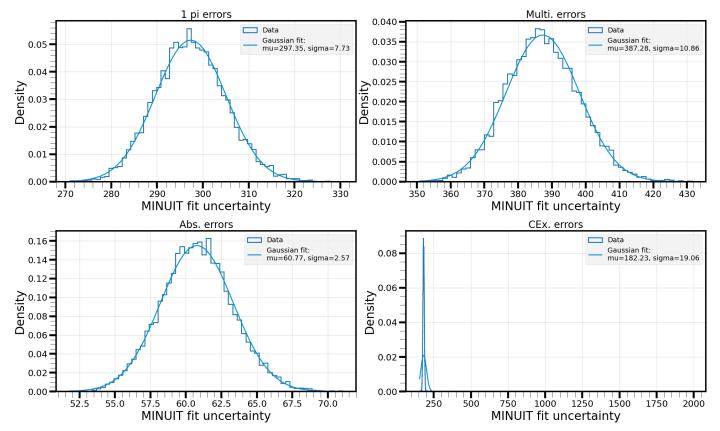
"Mixing" fractional errors $\frac{Prediction}{True} - 1$

• Combining all MC events, and randomly splitting these removes pulls.



"Mixing" fit uncertainties

 Combining all MC events, and randomly splitting these removes pulls.

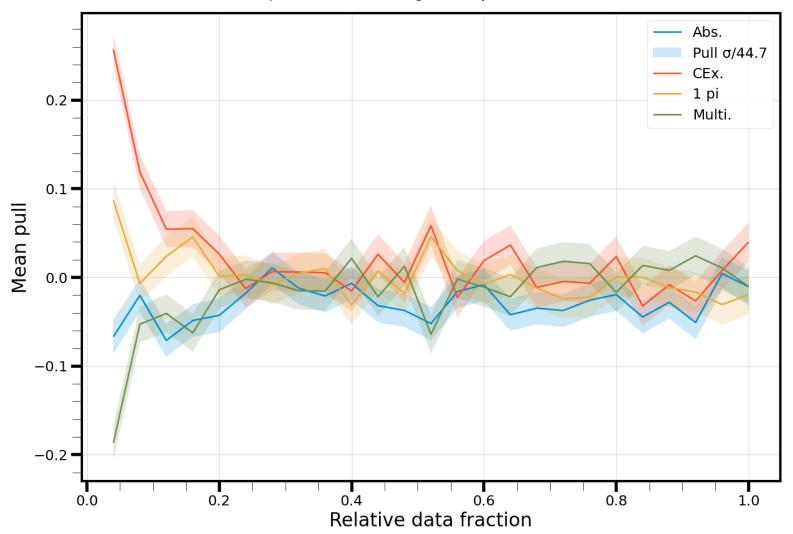


Tests

- Random fluctuation
- Data statistics
- Template statistics
- Re-weighted process fractions
- Initial fit predictions
- GNN score drift some form of smearing the underlying distributions
 - Note: need to confirm Minuit works with non integer templates (nuisances **Poisson** distribute template bins)
- Outliers

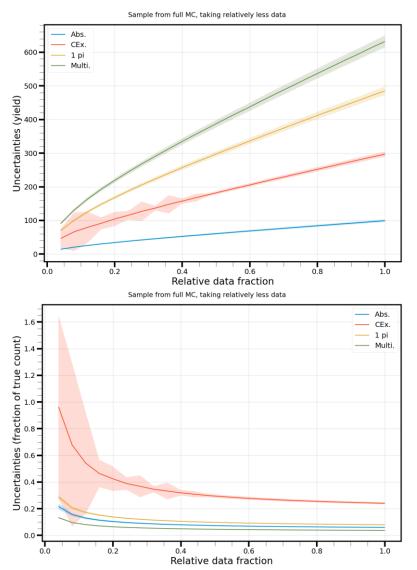
Data statistics

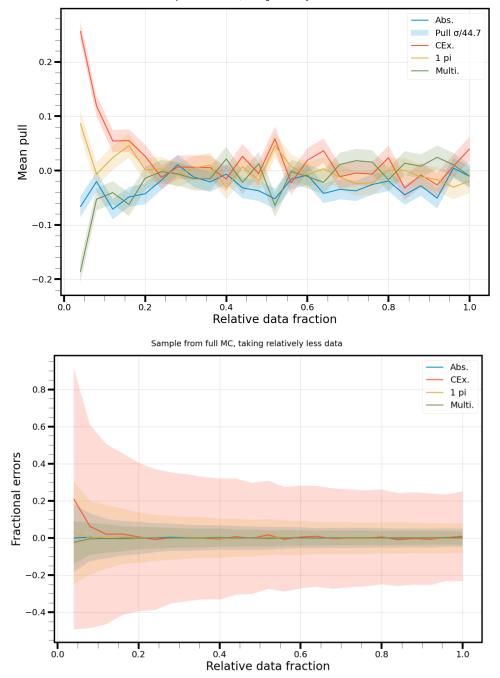
Sample from full MC, taking relatively less data



Sample from full MC, taking relatively less data

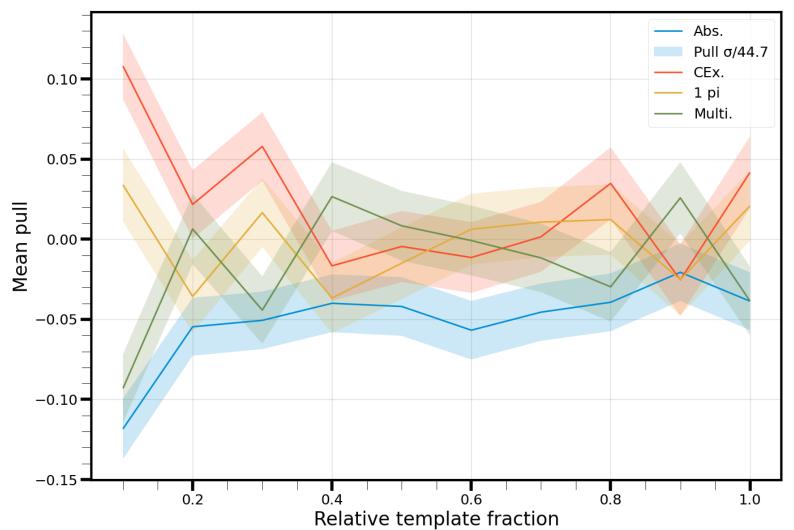






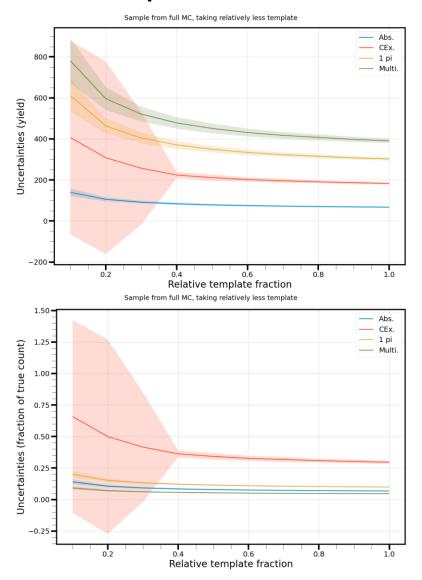
Template statistics

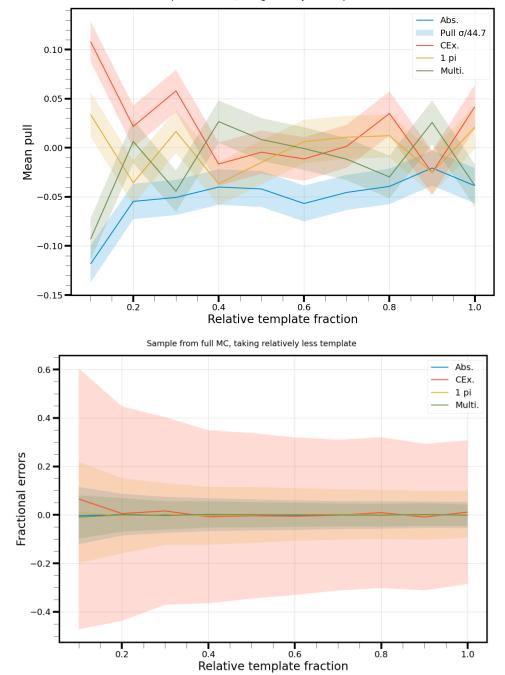
Sample from full MC, taking relatively less template



Sample from full MC, taking relatively less template

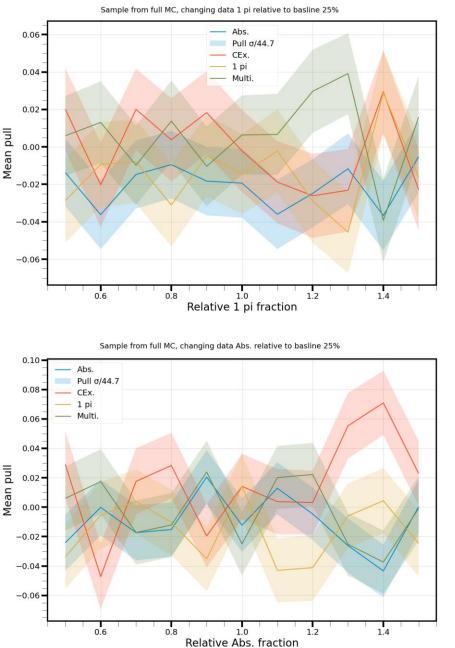
Temp. statistics

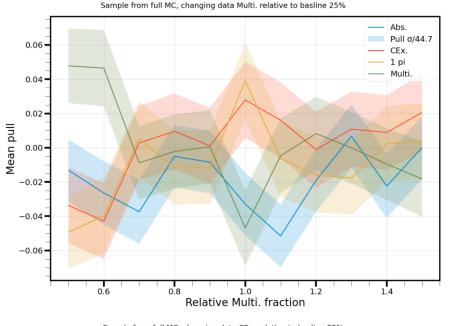


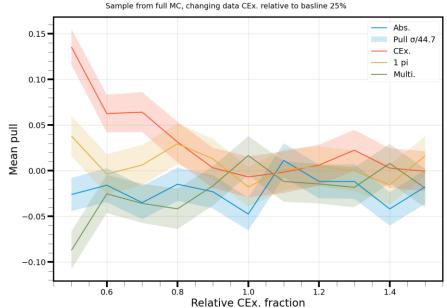


Weighting data

A bad method should show a negative gradient (over-estimate when data has a deficit)

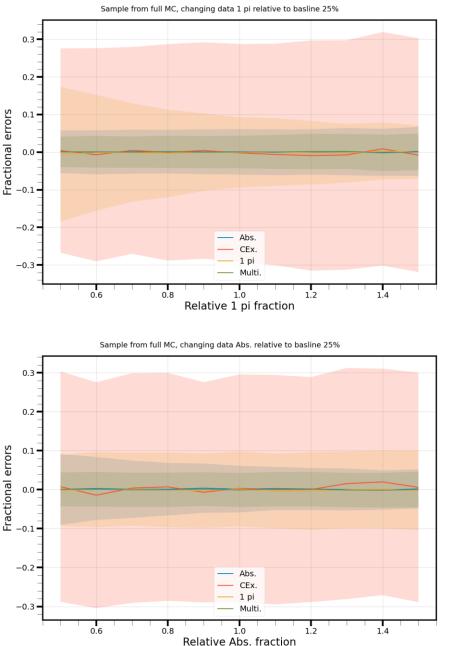


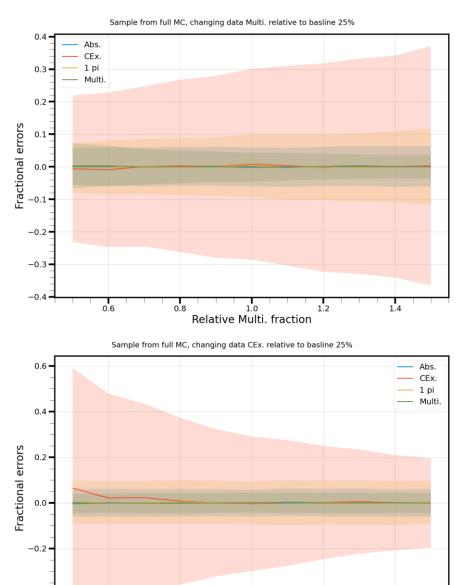




Weighting data

A bad method should show a negative gradient (over-estimate when data has a deficit)





1.4

-0.4

0.6

0.8

1.0

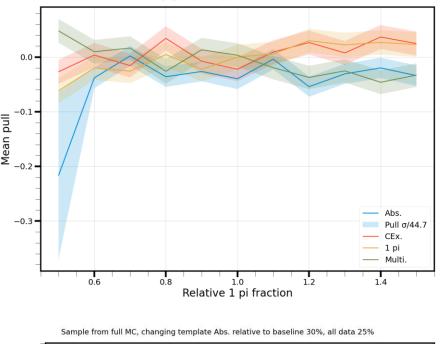
Relative CEx. fraction

1.2

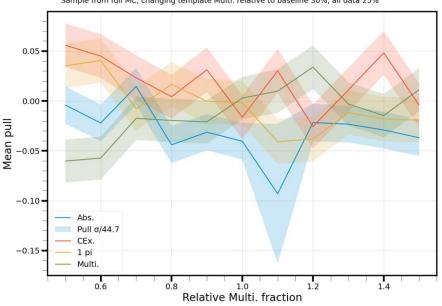
Weighting template

Sample from full MC, changing template 1 pi relative to baseline 30%, all data 25%

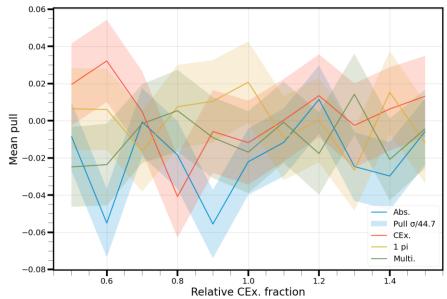
A bad method should show a positive gradient (under-estimate when template has a deficit)



0.08 Abs. Pull σ/44.7 CEx. 0.06 1 pi Multi 0.04 0.02 Mean pull 0.00 -0.02 -0.04-0.06-0.08 0.8 0.6 1.0 1.2 1.4 Relative Abs. fraction



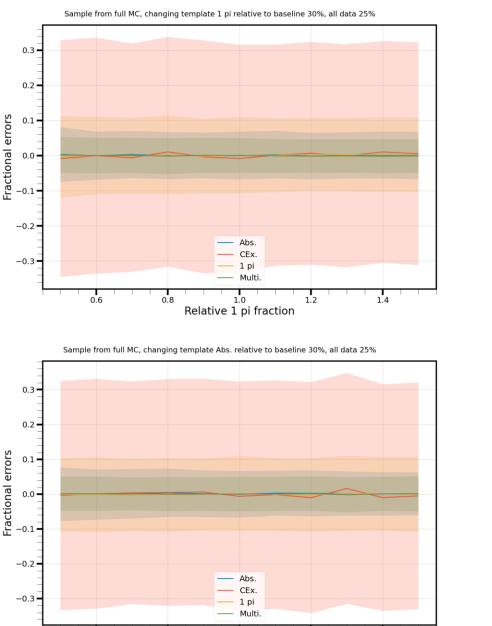
Sample from full MC, changing template CEx. relative to baseline 30%, all data 25%



Sample from full MC, changing template Multi, relative to baseline 30%, all data 25%

Weighting template

A bad method should show a positive gradient (under-estimate when template has a deficit)



1.0

Relative Abs. fraction

1.2

1.4

0.8

0.6

