Recent code updates

xFitter developer's meeting,
14 June 2023, S. Glazov

Recent updates

- Use the fork() for multiprocess computation of
  - PDF set variations in Profiler
  - Jacobian in CERES minimization
- General improvements in CERES minimizer
  - Add treatment of “bounds” and “priors” for fitted parameters
  - Add “ForwardDerivative” option
- Update in tools/test.sh: treat “,” as a delimiter
- Merged requests finalized:
  - merge_requests/292: chebyshev param (but example is failing, disabled)
  - merge_requests/277: treat matrix “forced positive” as “error”
  - merge_requests/300: Reaction RT NC: change correction to “additive”
  - merge request/288: Interface to pineapple grid
- Synchronize the CKM matrix in xFitter and APPLgrid. Special treatment of the V_{tx} elements (keep at 0 if they are at 0).
- Several minimizers can be executed one after another

Test of parallel profiling

profilerLHAPDF	112 sec
profilerLHAPDF threads=20	26 sec

Tests using “threads: 20” on i7-12800H (6x2 performance + 8 efficiency cores)

No large increase in consumed memory due to “copy-on-write” mechanism of the fork() command

Tests of parallel derivative

CERES-fit	102 sec
CERES-fit Forward derivative	46 sec
CERES-fit parallel	16 sec
CERES-parallel (RT)	39 sec
CERES-parallel (RT) Forward derivative	23 sec

Tests using “threads: 20” on i7-12800H (6x2 performance + 8 efficiency cores)

Test of generic minimization

```
real    3m21,865s
user    22m15,799s
sys     0m12,223s
```

```
Time (in seconds):
Preprocessor                            0.000082

  Residual only evaluation            60.748138 (21)
  Jacobian & residual evaluation     119.145190 (13)
  Linear solver                        0.032123 (21)
Minimizer                            179.929595

Postprocessor                           0.000005
Total                                179.929682
```

Fit to **HERA** (Q2>=3.5 GeV2) and **ATLAS W+Z 2016** data using **FONLL** for DIS.

Default HERAPDF 14 parameters plus **alpha$_s$** and **V$_{cs}$** (**16 parameters**)

```
 First iteration    14085.600565062623                 1190    11.836639130304725
After minimisation    1506.42  1190      1.266
```

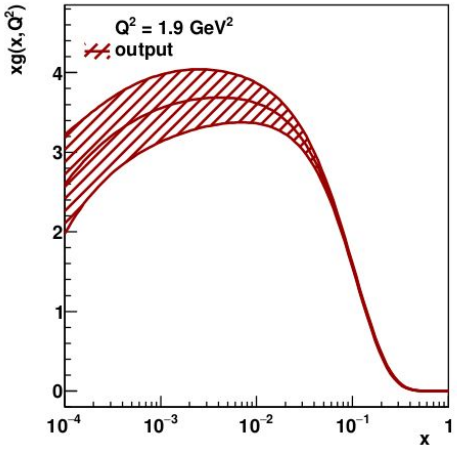
Multiple minimizers

```
Minimizers : [CERES, MINUIT]

CERES:
  offset: 2
  tolerance: 1e-5
  strategy: 0
  covariance: 0
  threads: 20

MINUIT:
  Commands: |
    set str 2
    call fcn 3
  doErrors :  Pumplin
```

Test on 8(16) core i7-10700F:

```
real    95m24,275s
user    663m39,117s
sys     0m43,557s
```

Parameter	output
'Adbar'	0.1522 ± 0.0015
'Agp'	1.185 ± 0.011
'Bdbar'	−0.1354 ± 0.0013
'Bdv'	0.6931 ± 0.0071
'Bg'	−0.1251 ± 0.0012
'Bgp'	−0.2298 ± 0.0023
'Buv'	0.7009 ± 0.0071
'Cdbar'	5.974 ± 0.062
'Cdv'	3.987 ± 0.041
'Cg'	10.75 ± 0.10
'Cubar'	1.251 ± 0.014
'Cuv'	3.914 ± 0.041
'Dubar'	−2.081 ± 0.021
'Euv'	4.504 ± 0.055
'Vcs'	0.9079 ± 0.0091
'alphas'	0.1097 ± 0.0011



- Several minimizers can be called in a single job
- Useful for CERES which has limited evaluation of uncertainties.
- Preliminary work to make Jon Pumplin's "Iterate" parallel rewriting in C++.
  - Done for the slowest Hessian computation part.
  - And for the first (out of two) method scanning around the minimum

Switch to LFS for xfitter-datafiles

- Branch containing recent ATLAS data pushes repository beyond the limit such that some of the files are not stored directly:
  - git clone does not copy some of the grid root files directly
  - to get the files, it is required to install git-lfs and use `git lfs checkout` command

```
- git lfs install
- git lfs fetch
- git lfs checkout
```

← One-time

- Tested to work with CI, plan to switch to it soon.
- Restore script to copy individual datasets?