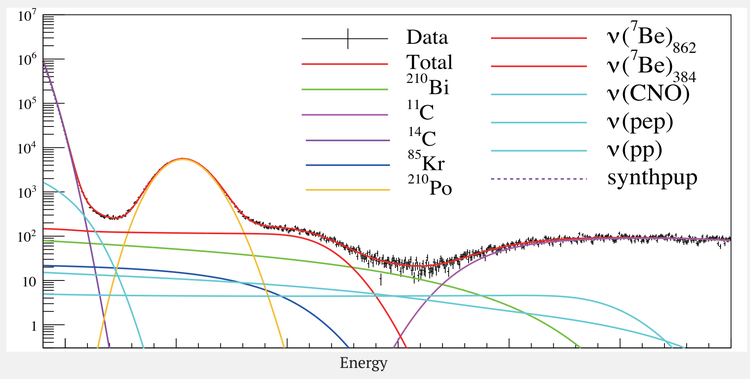
Speeding up complex multivariate data analysis in Borexino with parallel computing based on Graphics Processing Unit

Xuefeng Ding[1] Alina Vishneva[2] Ömer Penek[3] Simone Marcocci[1]

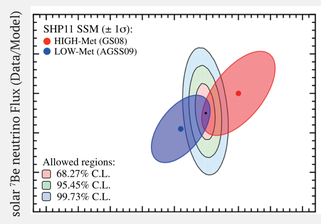
1. Gran Sasso Science Institute (INFN), 67100, L' Aquila, Italy
2. Joint Institute for Nuclear Research , 141980, Dubna, Russia
3. Forschungszentrum Jülich, 52428, Jülich, Germany

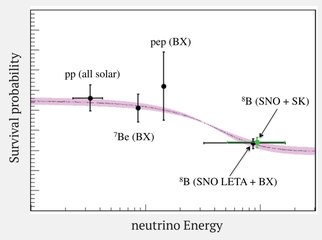
Solar neutrino with Borexino

- Disentangle signal-background with **spectral fitting**



- Study the **neutrino flux** or the **neutrino oscillations**.



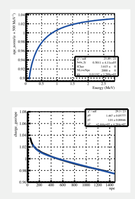


Challenges and Complexity in Borexino spectral analysis

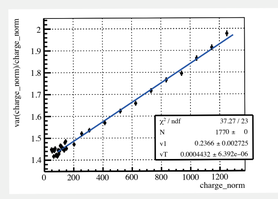
- Various **energy estimators** with advantage in different energy regions:
  - **Number of fired PMTs, number of hits**: free of extra PMT charge resolution
  - **Sum of Integrated PMT Charge**: more linear, better resolution at high energies
- Need of analytical approach:
  - Fit the **light yield**. Important for understanding the solar pp neutrino.
  - Fit the **resolution**. For charge variable it can fit the single p.e. charge resolution.
- Challenge to get the precise analytical detector response model
  - **The shape of mono-energetic uniform-in-fiducial-volume events**: scaled poisson, modified gaussian, generalized gamma. Validation of it on Monte Carlo. [1]
  - **Reliable and flexible non-linearity and resolution model**.
  - **Temporal dependence of detector response**.
  - Describe **the pile-up with ^{14}C effect**, critical effect for solar pp neutrinos.
  - Describe the spikes for the **normalized integer variable**.
- Various studies need to be done to get robust fitting procedure and fit result:
  - Validation with correlation plot: fit of toyMC spectra, reproduce the injected rate
  - Scan of parameters: Varying fit range, binning, all fixed parameter etc.
  - Systematics: Inject deformation of detector response and see how stable is the fit.

Common challenges for all experiments! Need to be **fast**! Need **GPU**!

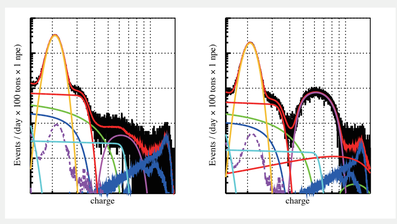
Non-linearity model



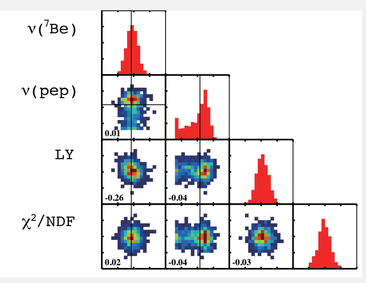
Resolution dependence



An example fit with analytical approach



Thousands of fit: check correlation and bias

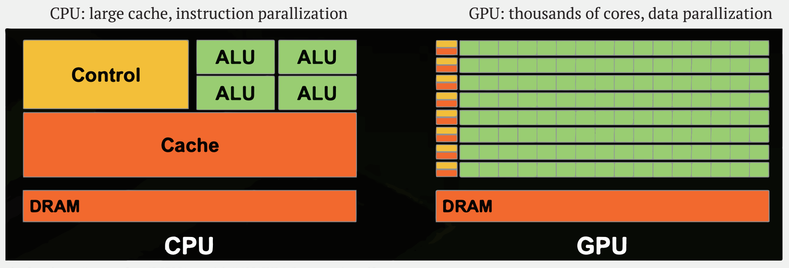


Parallel computing

This project is based on C++11, ROOT, cuda and GooFit

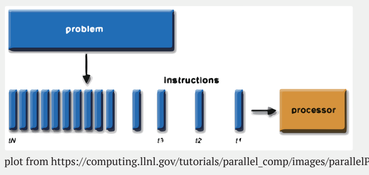
Graphic Processing Unit: thousands of cores, data parallelization

CPU: large cache, instruction parallization GPU: thousands of cores, data parallization

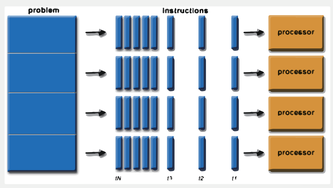


plot from https://www.ogf.org/OGF25/materials/1605/CUDA_Programming.pdf

Parallel computing: divide into small tasks and solved simultaneously



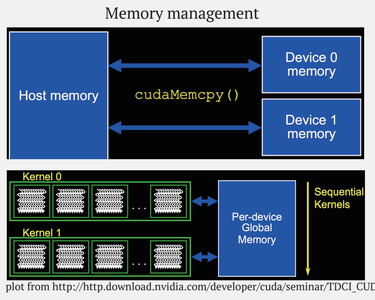
VS



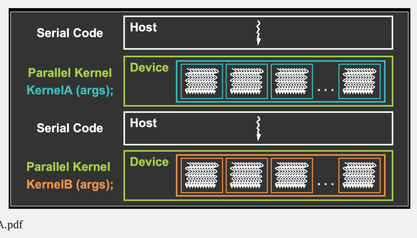
plot from https://computing.llnl.gov/tutorials/parallel_comp/images/parallelProblem2.gif

Scheme of Graphic Processing Unit based parallization

Memory management

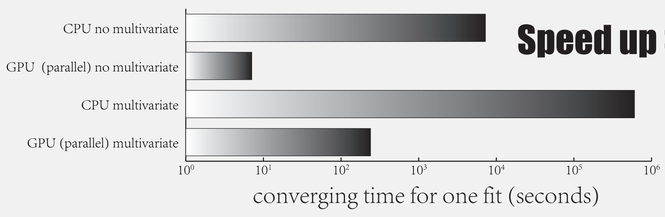


Program flow



plot from http://http.download.nvidia.com/developer/cuda/seminar/TDCI_CUDA.pdf

Performance and physics results



Speed up >x1000 !

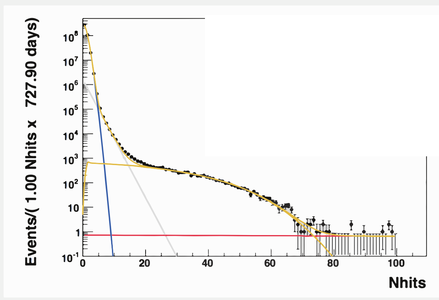
The physics result got with this tool

- The global spectral-fit of solar neutrino rate with analytical approach
- The systematic uncetrtainty, which is got with tens of thousands of fits
- The upper limit for the neutrino magnetic moment with phase-II data

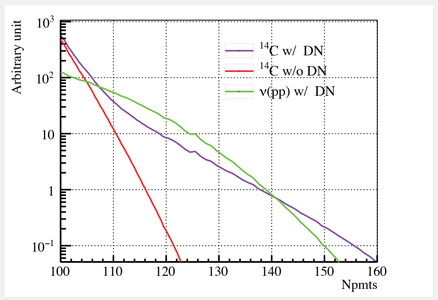
Physics motivated Statistical tool for Borexino analysis

How to fit pp neutrino? Account for Pile-up from ^{14}C

- The major background for pp neutrino is **^{14}C**. The **pile-up effect modifies its shape** in region of interest of pp neutrino drastically.
- ^{14}C will **pile-up with every-thing**: ^{14}C, ^{210}Po etc.
- Take **spectrum with random trigger**, which is the spectrum to be piled on ^{14}C
- **Convolve the analytical ^{14}C shape with such spectrum** to account for the pile-up effect



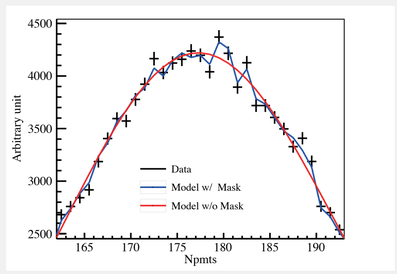
Spectrum acquired with random trigger



Primary bkg. for *ν*(pp): ^{14}C-^{14}C pile up

Spikes from integer

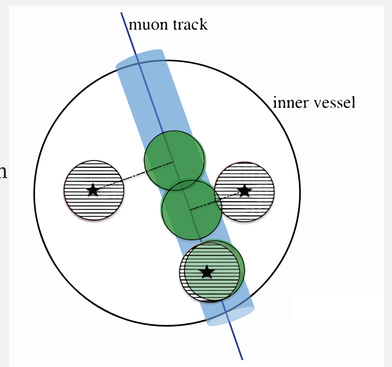
Use number of fired pmts as energy estimator, which is integer. Need correction to account for spikes due to bining.

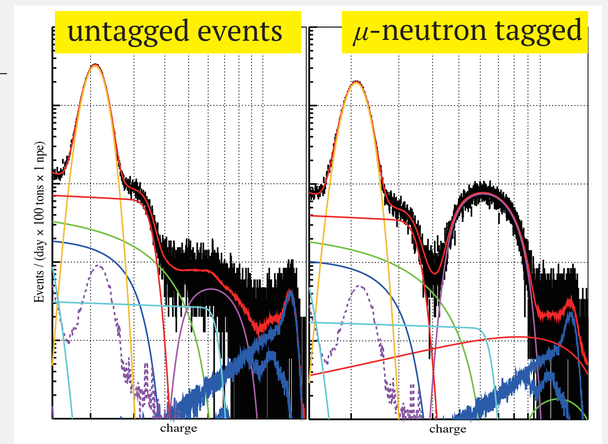


Red vs Blue: w/ vs w/o correction

reject ^{11}C with μ-neutron tagging: Complementary

- **cosmogenic ^{11}C** is the major background for **pep/CNO neutrinos**
- The production of **cosmogenic ^{11}C** is usually **associated with one or more spallation neutrons**.
- If a physics event is asscociated with a muon track and a tagged neutron, it is tagged to be more likely a cosmogenic ^{11}C.
- Rather than cut these events, we build another spectrum and sum up the likelihood of both. **Sensitivity on solar pp and ^{7}Be neutrino species is improved**.
- For analytical approach, the enhanced ^{11}C **improves the knowledge on the resolution model**.
- Massive tests have been performed and this method is validated.

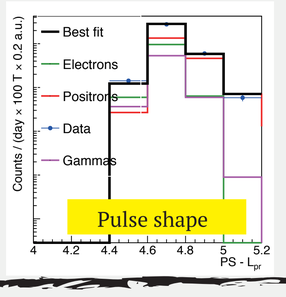


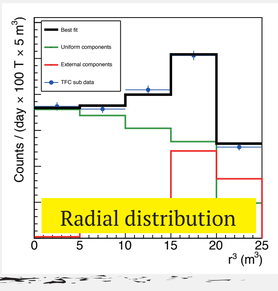


Further improvement of pep and CNO sensitivity

β^{+}/β^{-} **P**ulse **S**hape **D**iscrimination to suppres ^{11}C
Radial **D**istribution to suppress external γ

Integrate PSD/Radial distribution into the likelihood: **Multivariate fit**





And... **Year-by-year simutaneous fitting**, constrain neutrino to be constant while background different each year, North/South splitting, constrain light yield, resolution etc. follow **linear trend over time** etc.

[1] Oleg Smirnov. NIM, 595, 2008 (410-418) [2] Bellini, G. et al. PRD, 89, 112007, 2014 [3] New Borexino paper

GRAN SASSO SCIENCE INSTITUTE — SCHOOL OF ADVANCED STUDIES — Scuola Universitaria Superiore

TAUP 2017

XV International Conference on Topics in *Astroparticle and Undergroup Physics*

Sudbury, ON, Canada
24 - 28 July 2017

Contact: Xuefeng Ding <xuefeng.ding@gssi.infn.it>
Project page: https://github.com/DingXuefeng/GooStats