Unraveling Hadron Mass and Quark Structure with COMPASS & COMPASS++/AMBER

(Nuclear Physics - Frontera LRAC allocation)

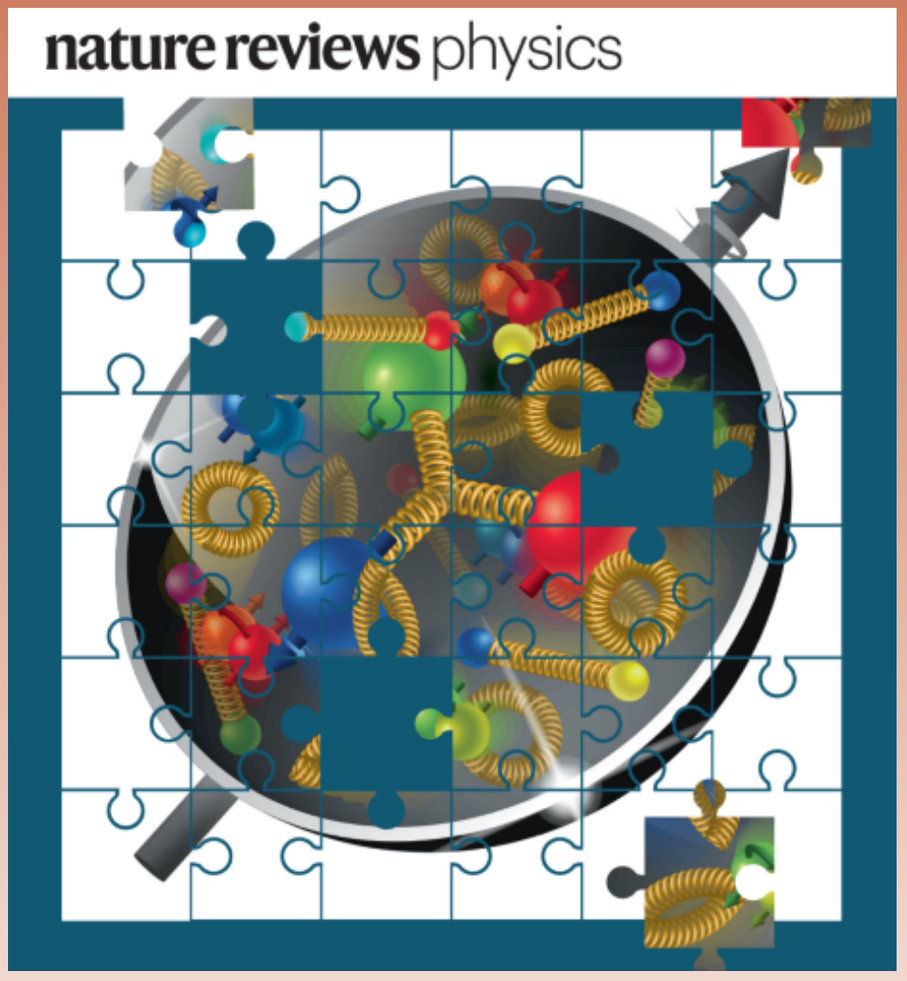


Image: Courtesy of Brookhaven National Laboratory. Cover design: Charlotte Gurr.

Vincent Andrieux[2], Riccardo Longo[2], Greg Mattson, Matthias Perdekamp[2], April Townsend, Caroline Riedl[1] (UIUC)

Frontera User Meeting 2021

January 28, 2021

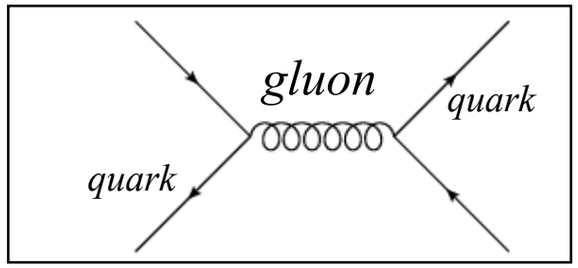
ILLINOIS

[1] PI and contact
[2] co-PI

The proton and the strong nuclear force

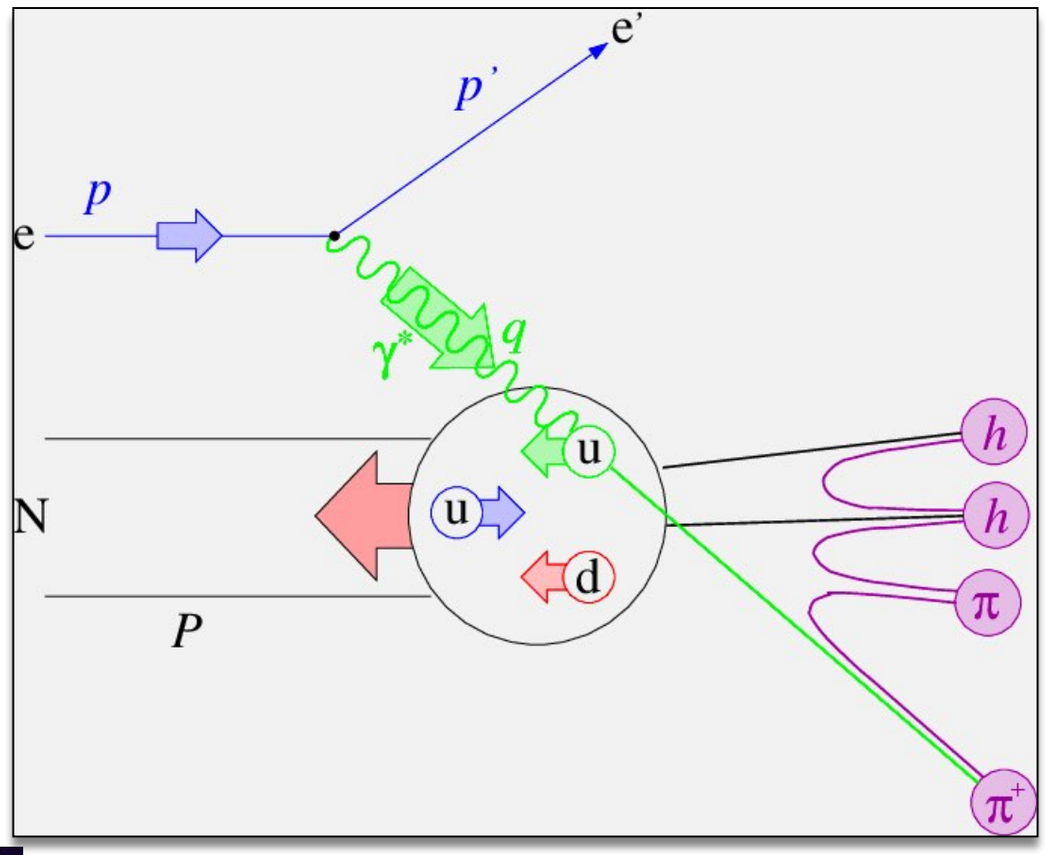
- Proton = nucleus of the hydrogen atom:
  ~ 1 femto meter (10^{-15}m), ~ 10^{-27} kg or 938 MeV

- Consists of
  - 3 valence quarks
  - quark-antiquark pairs = sea quarks
  - gluons mediating strong nuclear force



- Quantum Chromo Dynamic (QCD)
  = quantum field theory of the strong nuclear force

Besides electromagnetism, weak nuclear force & gravity one of the 4 fundamental forces in nature.



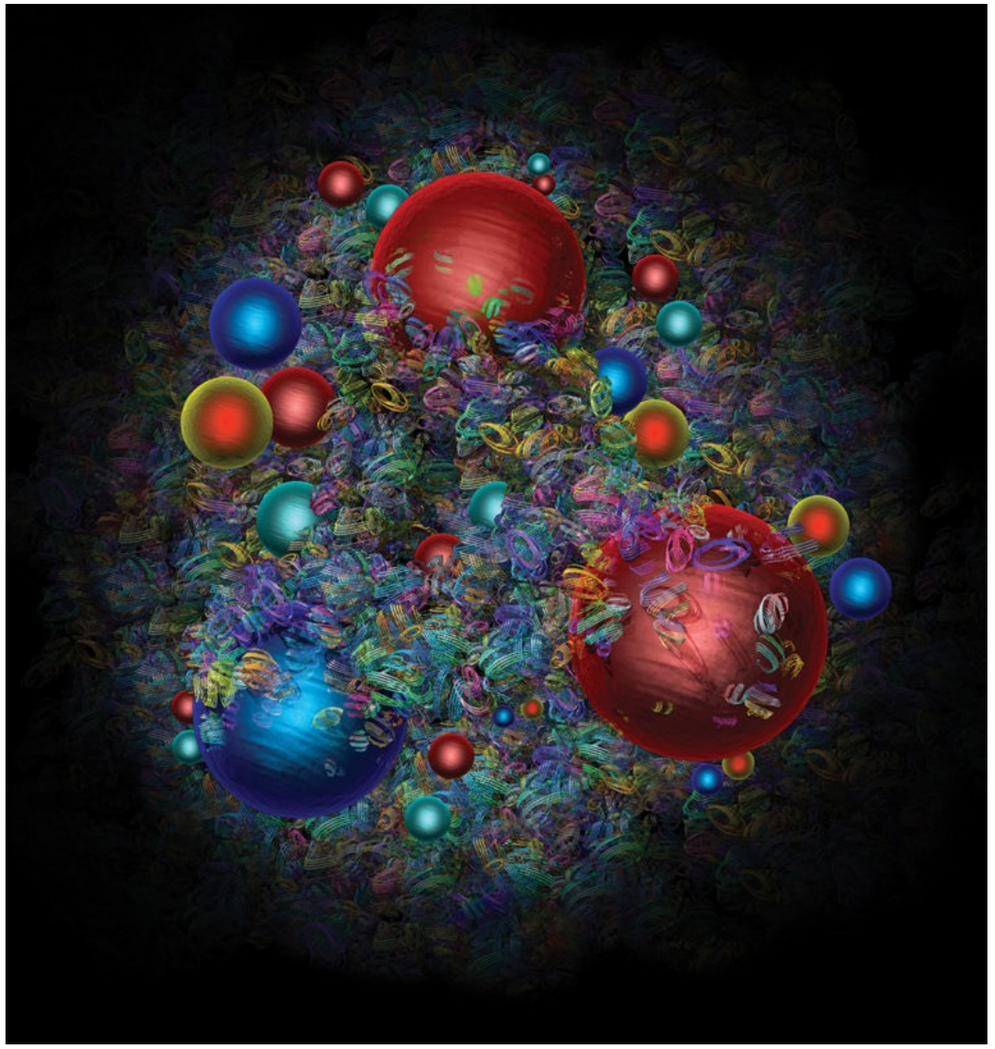


Image from D. Dominguez, CERN courier May/June 2019
"The proton laid bare"

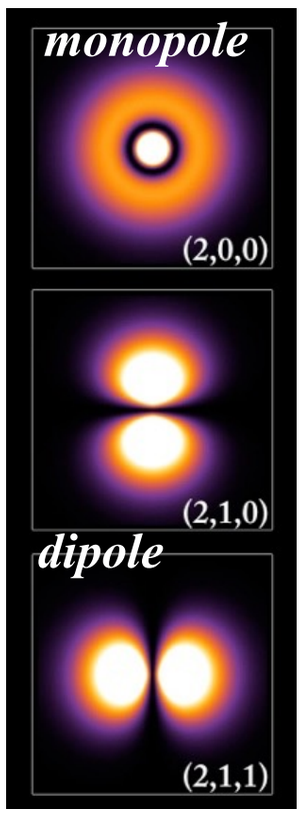
- Proton structure probed in scattering experiments, for example deep-inelastic proton-electron scattering

- Worldwide only few places: need accelerator laboratory like BNL, JLab, FNAL, SLAC, CERN, DESY, RIKEN

Proton “orbitals”: from QED to QCD

~ 1930’s: Quantum Electro Dynamics QED
(the theory of the electromagnetic force between charged particles)

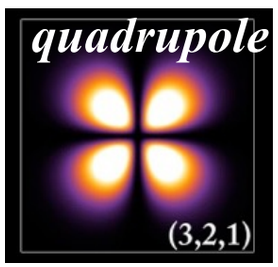
Hydrogen atom - “wave function”



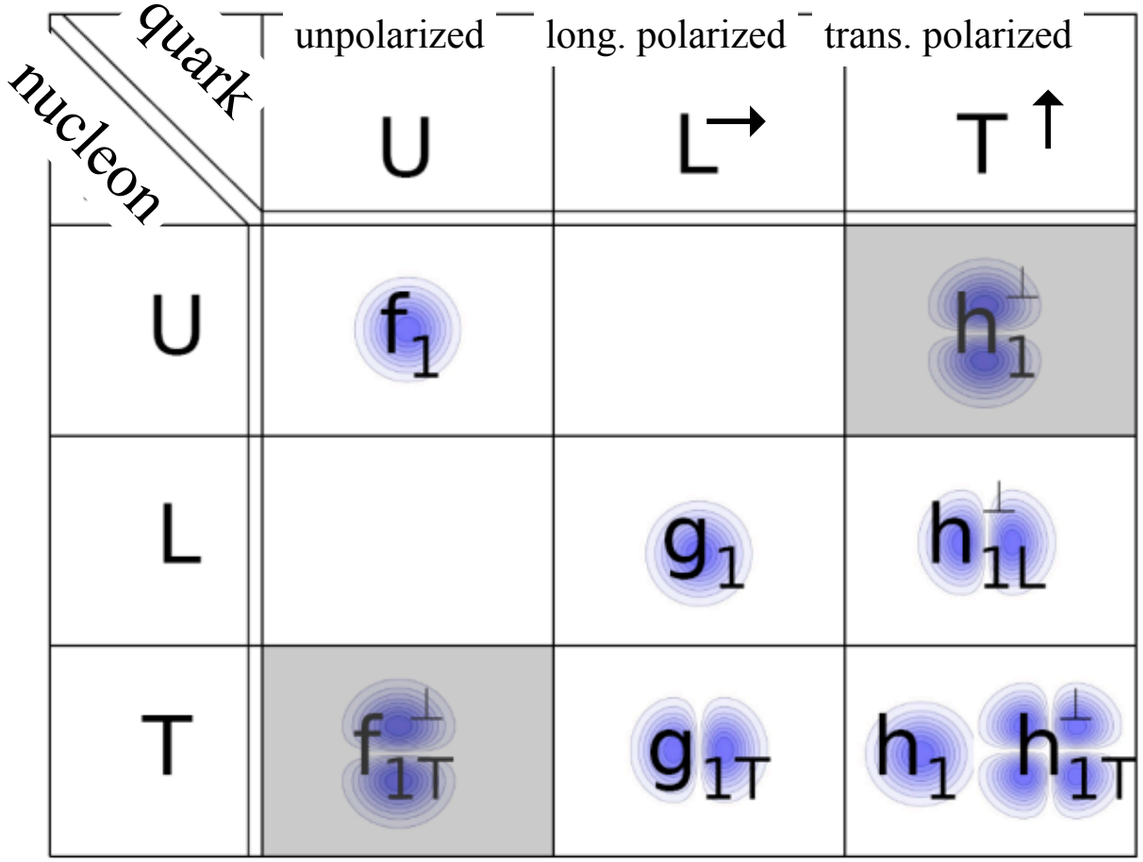
$$\Psi_{nlm}(r,\theta,\phi)$$

spin

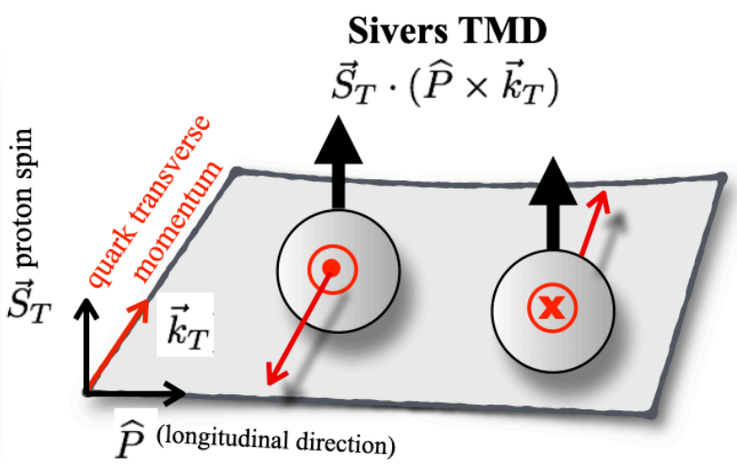
orbital angular momentum



~ 2010’s+: Quantum Chromo Dynamics QCD
(the theory of the **strong nuclear force between quarks**)
with proton polarization & transverse degrees of freedom



Correlation between transverse spin of proton and transverse momentum of quarks is indicative of orbital angular momentum of quarks in the proton

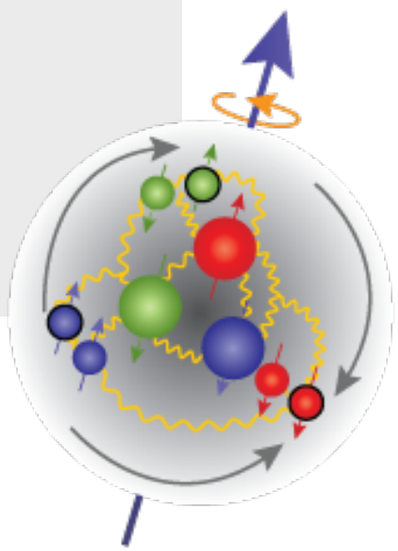


Proton: spin and mass… more complex than one thinks!

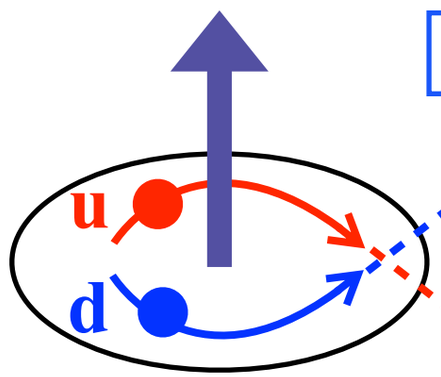
Proton spin budget & internal dynamics:
1/2 $\hbar$ = "spin quarks + spin gluons +

orbital angular momentum"

plus add transverse spin &
transverse momentum



Scattering off the spin-polarized proton - **COMPASS**



$\pi^-=\bar{u}d$ to the right

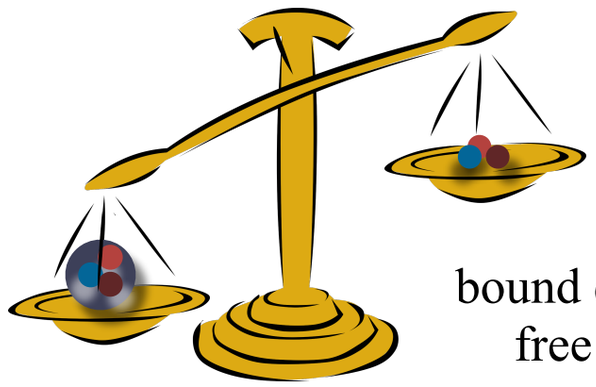
$A_N < 0$ for π^-

$\pi^+=u\bar{d}$ to the left

$A_N > 0$ for π^+

quarks' orbital movement inside the proton creates left-right asymmetry A_N

Emergence of hadron mass: only a small portion of the mass of visible matter comes from the Higgs mechanism (bare quark masses). The bulk of proton mass is emergent = created dynamically through the strong nuclear interaction.



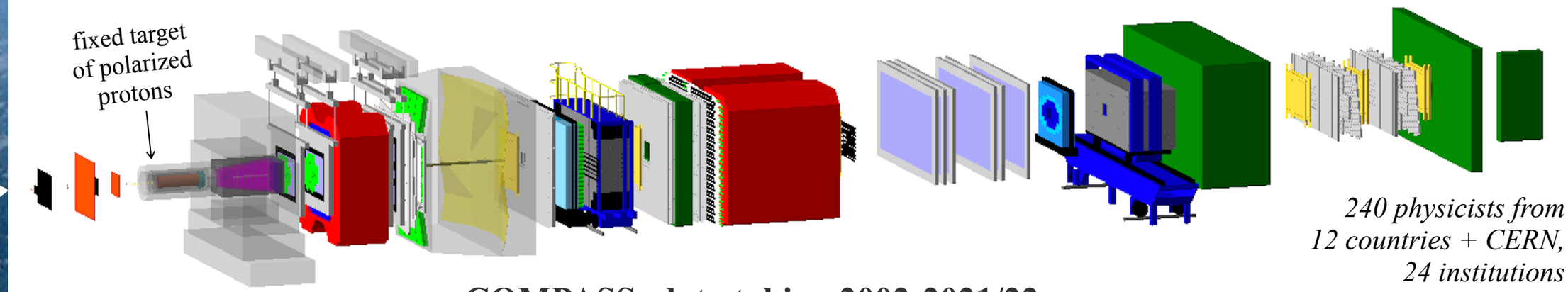
bound quarks vs. free quarks

Use different beam types (pions, kaons, anti-protons) - **COMPASS++/AMBER**

3-quark bound states (baryons) vs. 2-quark bound states (mesons)

COMPASS @ CERN

CERN = European Center for Nuclear Research
COMPASS = COmmon Muon Proton Apparatus for Structure and Spectroscopy



240 physicists from
12 countries + CERN,
24 institutions

COMPASS; data taking 2002-2021/22
COMPASS++/AMBER: 2022++

particle beam

CERN

COMPASS

Lake Geneva

France

CERN North Area

Switzerland

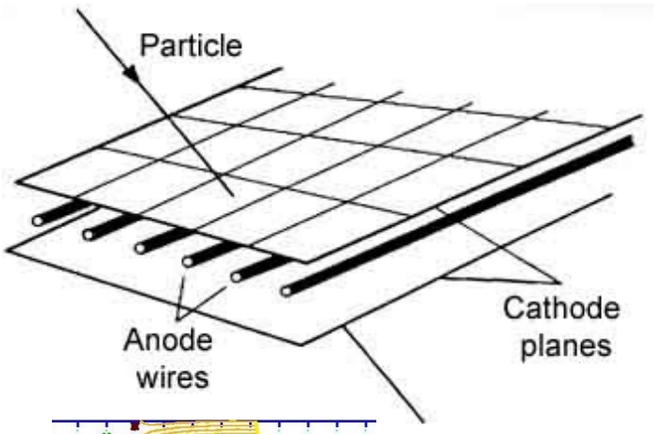
Geneva airport

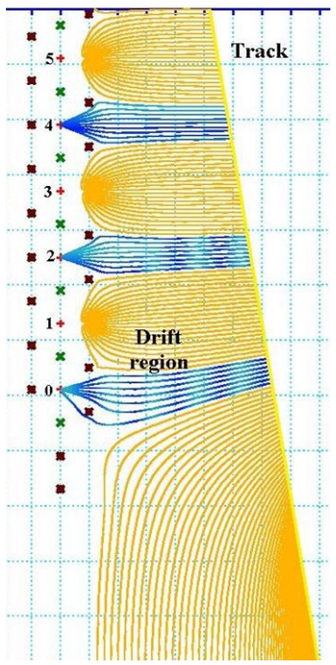
LHC

SPS

Future COMPASS++ / AMBER
- https://nqf-m2.web.cern.ch
- Letter of Intent 2018 arxiv:1808.00848[hep-ex]
- Proposal 2019 phase 1 (approved Oct. 2020) CERN-SPSC-2019-022

Tracking of high-energetic charged particles



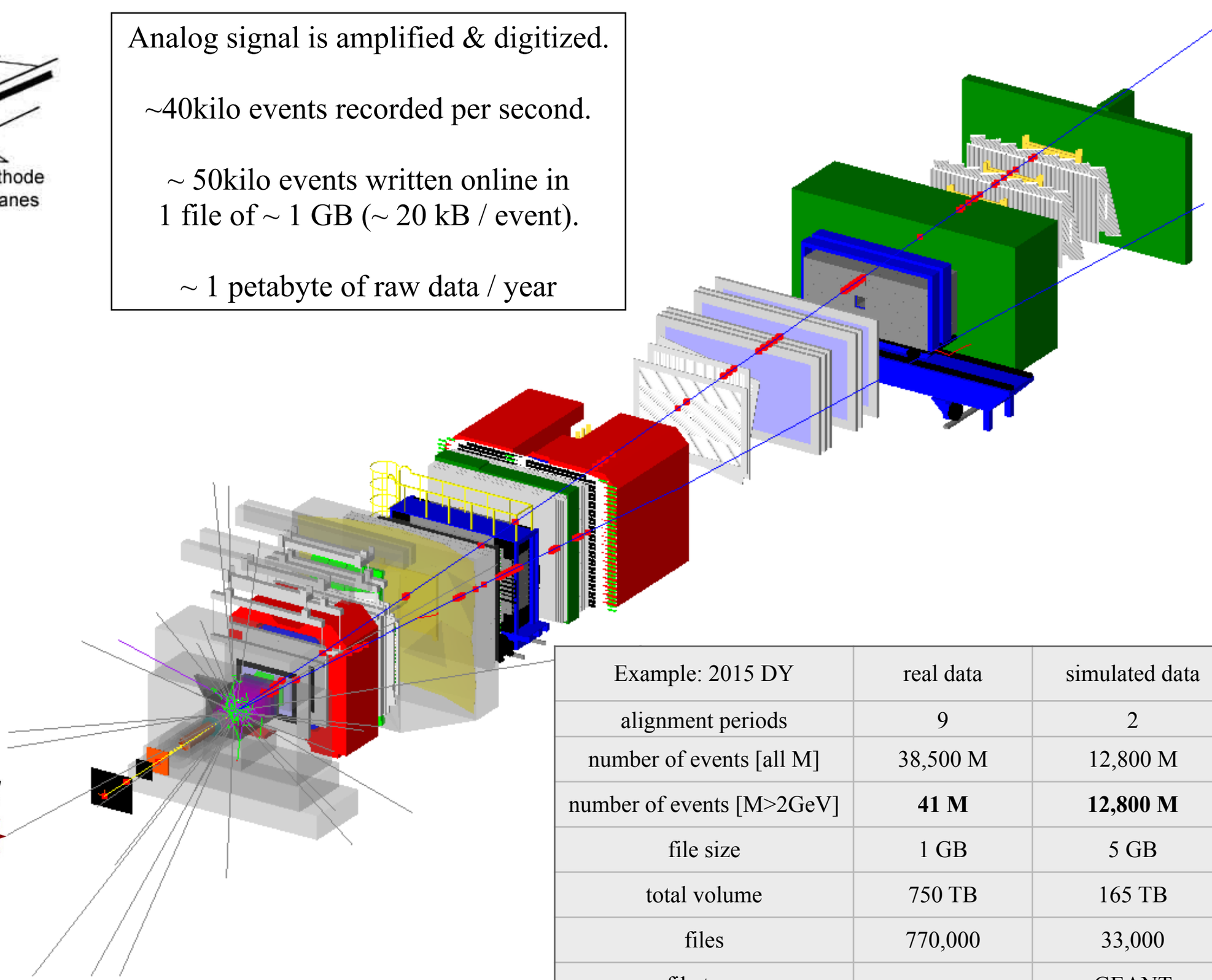


Analog signal is amplified & digitized.

~40kilo events recorded per second.

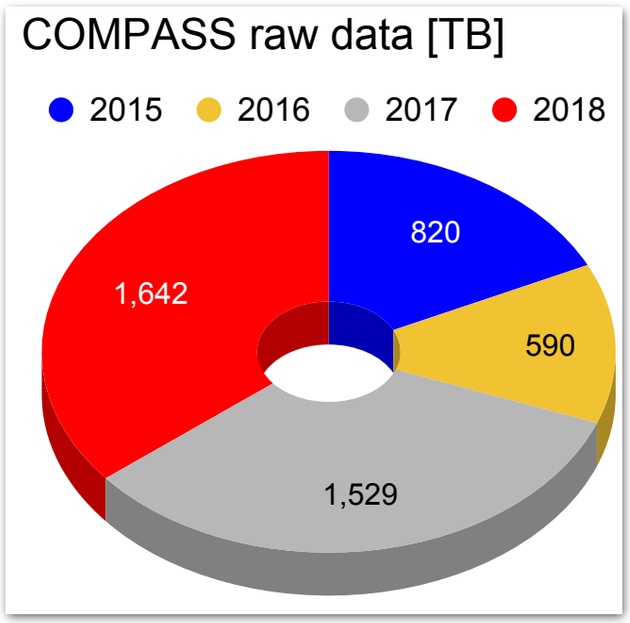
~ 50kilo events written online in
1 file of ~ 1 GB (~ 20 kB / event).

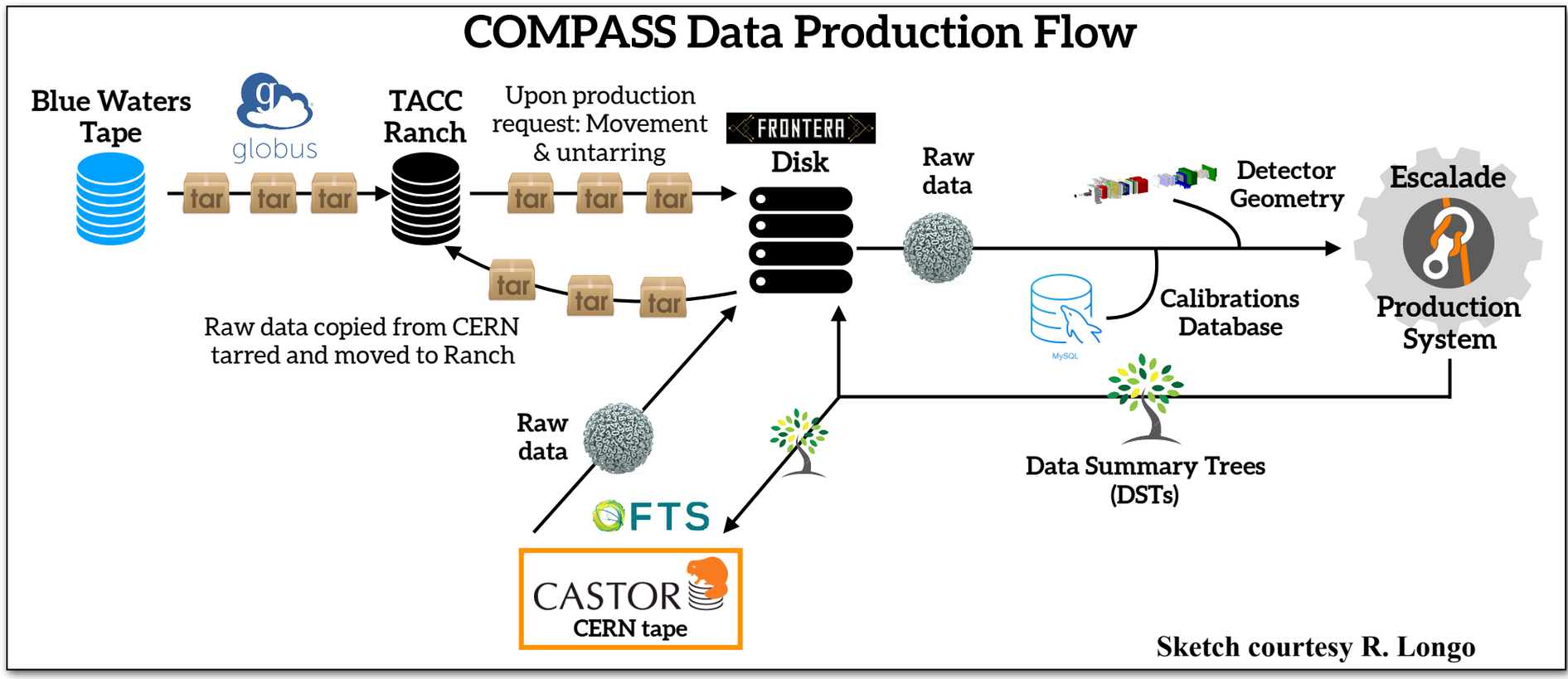
~ 1 petabyte of raw data / year



Example: 2015 DY	real data	simulated data
alignment periods	9	2
number of events [all M]	38,500 M	12,800 M
number of events [M>2GeV]	**41 M**	**12,800 M**
file size	1 GB	5 GB
total volume	750 TB	165 TB
files	770,000	33,000
file type	raw	GEANT

Submissions to Frontera grid





1 node = 55 raw files

1	2	3	4	5
6	7	8	9	10
11	12	13	14	15
16	17	18	19	20
21	22	23	24	25
26	27	28	29	30
31	32	33	34	35
36	37	38	39	40
41	42	43	44	45
46	47	48	49	50
51	52	53	54	55

56 CPUs on 1 node

56 calibration data base

1 job ~ 8k files

1	2	3	…
…			
147	148	149	150

1 submission ~ 80k files

- job 1
- job 2
- job 3
- job 4
- job 5
- job 6
- job 7
- job 8
- job 9
- job 10

- On a given node, tasks are dispatched and terminated using an MPI-based code (pcp, https://github.com/wtsi-ssg/pcp)
- 10-job-limitation due to I/O restraints.
- Copy of output data from /tmp to /scratch at the end of a job has issues since late 2020 - higher occupancy of Frontera?

Frontera usage

Our group ran COMPASS data productions on NCSA's Blue Waters 2016-2019. We moved to Frontera in 2019:

- Experimental data productions for 2018 COMPASS data
- Data analysis of 2018 and 2015 data
- Detector efficiency maps for 2016 data (CPU intensive)
- Monte-Carlo studies for 2018 data

- Planned:
  - close-to-final productions for 2018 data in nearest future
  - more detector efficiency maps for 2018 & 2016 data
  - Monte-Carlo mass productions for 2018 data
  - Detailed simulations for COMPASS++/AMBER

Frontera LRAC allocation - usage in kilo node hours

	unused	RD productions	detector maps	COMPASS simulations	AMBER simulations	physics analysis
actual	972	389	23	91	0	20
original plan		400	600	300	100	100
fulfilled	35%	97%	4%	30%	0%	20%

➤ Much faster data productions than would be possible on other computing clusters

➤ Minimization of systematic uncertainties due to the possibility of generating larger simulated samples and perform CPU-intensive simulation studies (*)

➤ Unprecedented precision determination of detector efficiencies

Team members:

UIUC: Vincent Andrieux (2), Riccardo Longo (2), Greg Mattson (3), Marco Meyer-Condo (2), Matthias Perdekamp (1), April Townsend (3), Caroline Riedl (1)
JINR, Russia: Artem Petrosyan (3)
Academia Sinica, Taipei: Yu-Shiang Lian (3), Chia-Yu Hsieh (3)
Warsaw University, Poland: Anatolii Koval (3)
LIP, Portugal: Catarina Quintans (1)
Bonn University, Germany: Henri Pekeler (3)

(1) senior, (2) postdoc, (3) grad student

(*) multi-dimensional acceptance studies; pile-up; minimum bias trigger

Summary: Unraveling Hadron Mass and Quark Structure with COMPASS & COMPASS++/AMBER

- Frontera allows the COMPASS data to be analyzed in a fast way, at high precision and using novel approaches
- Frontera allows to simulate COMPASS++/AMBER data to prepare the best possible instrumentation
- Outlook: Frontera is essential to prepare the 2018 data for public presentation at 2021 spring conferences
- Education of students and young postdocs in petascale computing.
  Creation of reference data productions shared with dozens of researchers.

Project abstract: Hadrons are the bound states of quarks held together by gluons, particles mediating the strong nuclear force. The mass of the quark compounds cannot be understood by only summing its constituent masses. In addition a dynamic contribution introduced by the strong force has to be considered. The observed hadron mass hierarchy remains a mystery to date. Simulating detector setups on Frontera, we intend to determine an optimized set of instrumentation to experimentally address the hadron mass puzzle with the future COMPASS++/AMBER experiment to run at CERN after 2021. Using Frontera, we will also complete the 2015-2018 COMPASS measurement campaigns to unravel the transverse hadron structure in momentum and position space, and the origin of proton spin.

Backup

COMPASS data productions on Frontera

RD (Real Data)

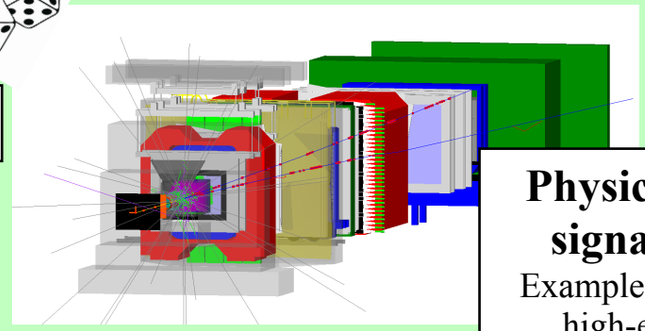
If coincident signal in trigger detectors: Read out all detector planes (>300) & write to disk at CERN: **Data AcQuisition = DAQ**

raw data

Calibration: physical energy deposits in detectors

mySQL DB

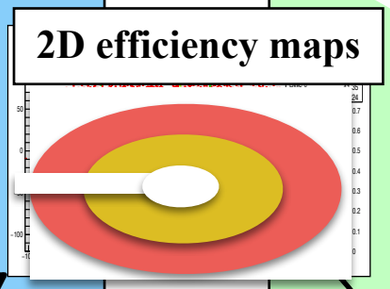
MC (Monte Carlo simulation)



Physics event generator signal & background
Example PYTHIA: generation of high-energy physics events
http://home.thep.lu.se/Pythia

GEANT4
Toolkit for the simulation of the passage of particles through matter
http://geant4.web.cern.ch

CPU intensive part



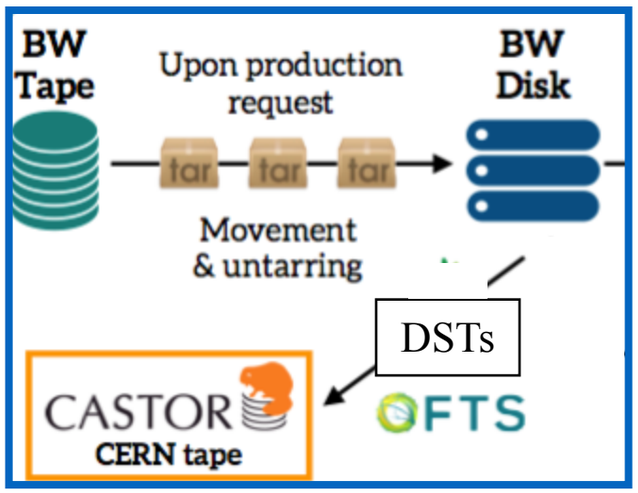
2D efficiency maps

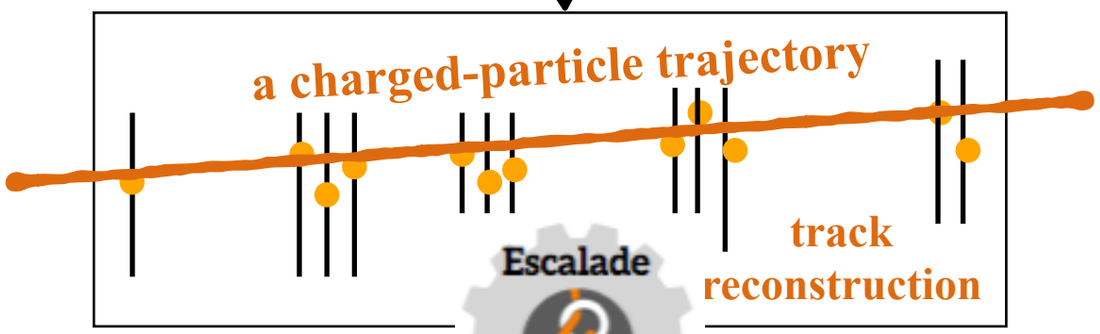
Alignment: spatial hits in detectors

CORAL = COmpass Reconstruction Analysis Library

if RD

if MC





Escalade

Production System

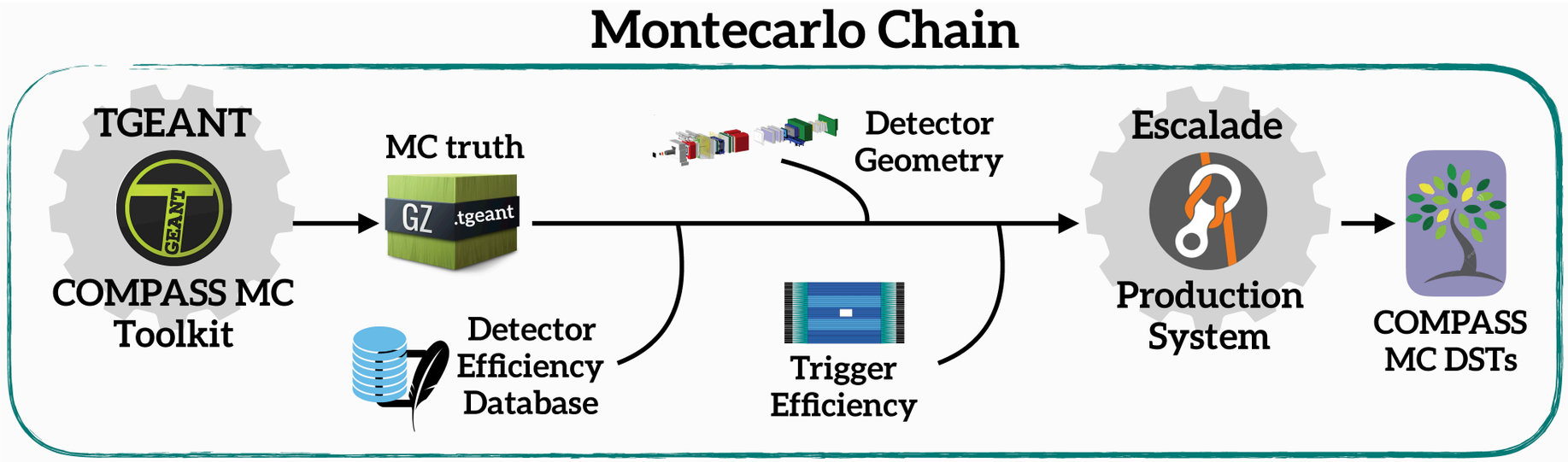
DSTs
(skimmed Data Summary Trees)
factor 5-10 reduction

User analysis

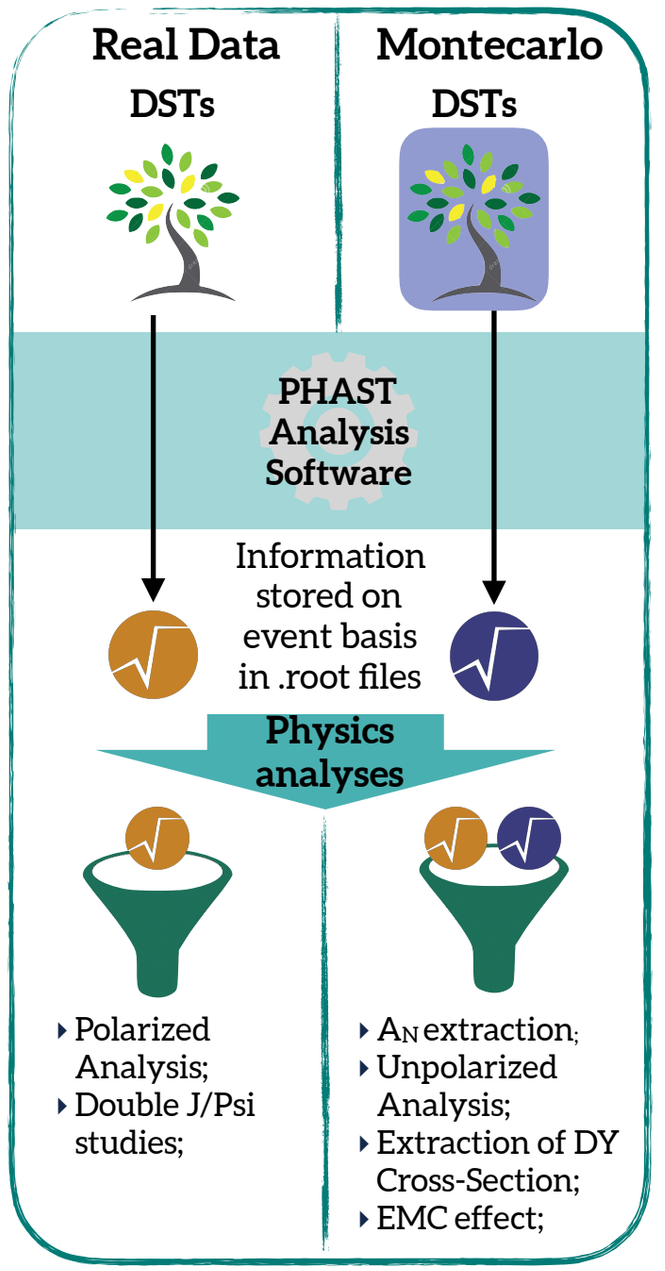
Simulations of the detectors play a central role in understanding subtle detector effects and removing background events from the data sample.

Backup workflow on Frontera

Montecarlo Chain



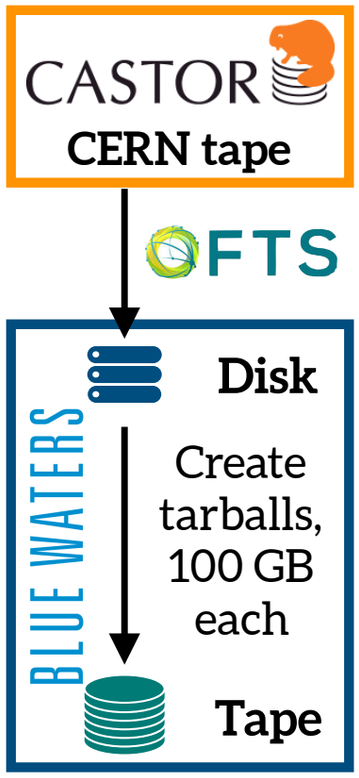
DY Analysis Flow

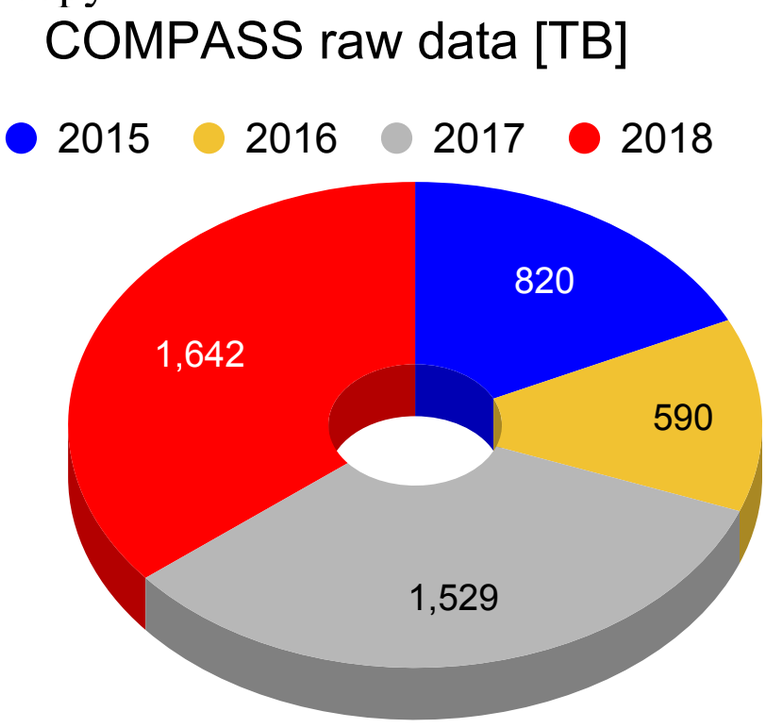


Sketches courtesy R. Longo

COMPASS data transfer

- Transferred ~ 3 petabytes of raw COMPASS data from BW or CERN to Frontera, ~ 3 million files of ~ 1GB
- Transfer of produced DSTs (~1/10 reduction factor) back to CERN
- Using FTS3 **(File Transfer System)**: bulk data mover created to globally distribute LHC data. FTS3 effectively uses globus-url-copy.



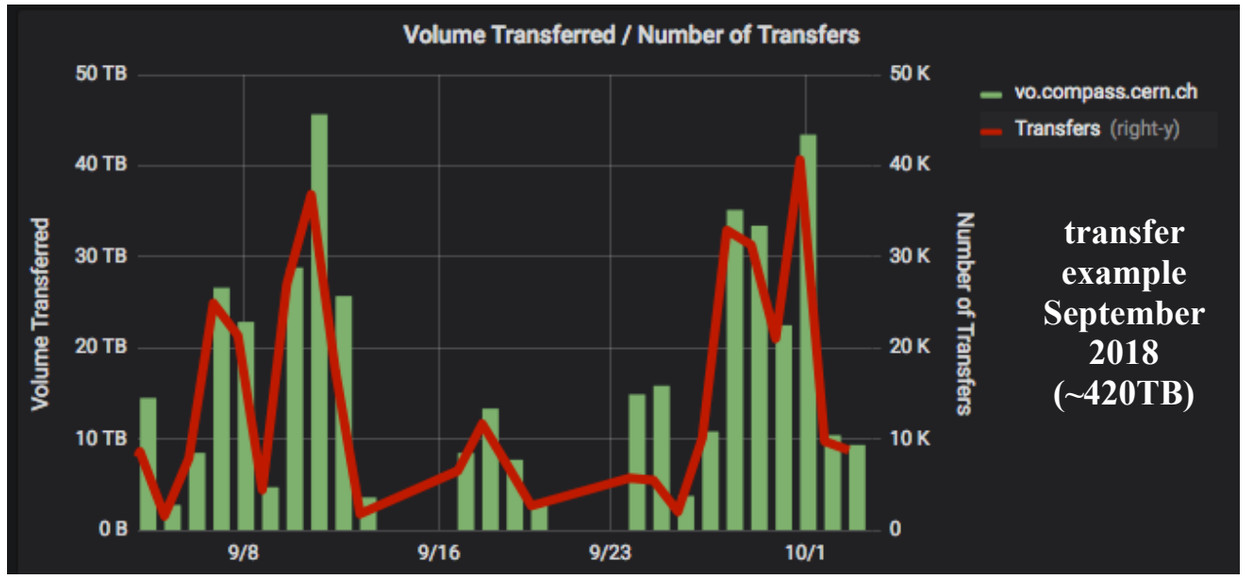


1 data period = 2 weeks of data taking

raw dy15	[TB]	chunks
W07	104	97,766
W08	104	98,067
W09	112	105,450
W10	85	80,188
W11	138	129,467
W12	96	89,995
W13	84	78,970
W14	57	53,978
W15	40	37,234

raw dvcs2016	[TB]	chunks
W06 (P01)	28	27,332
W07 (P02)	57	53,711
W08 (P03)	60	57,467
W09 (P04)	60	57,661
W10 (P05)	55	51,919
W11 (P06)	57	54,473
W12 (P07)	70	66,080
W13 (P08)	77	72,444
W14 (P09)	62	58,533
W15 (P10)	64	60,608
W16 (P11)	14	13,652

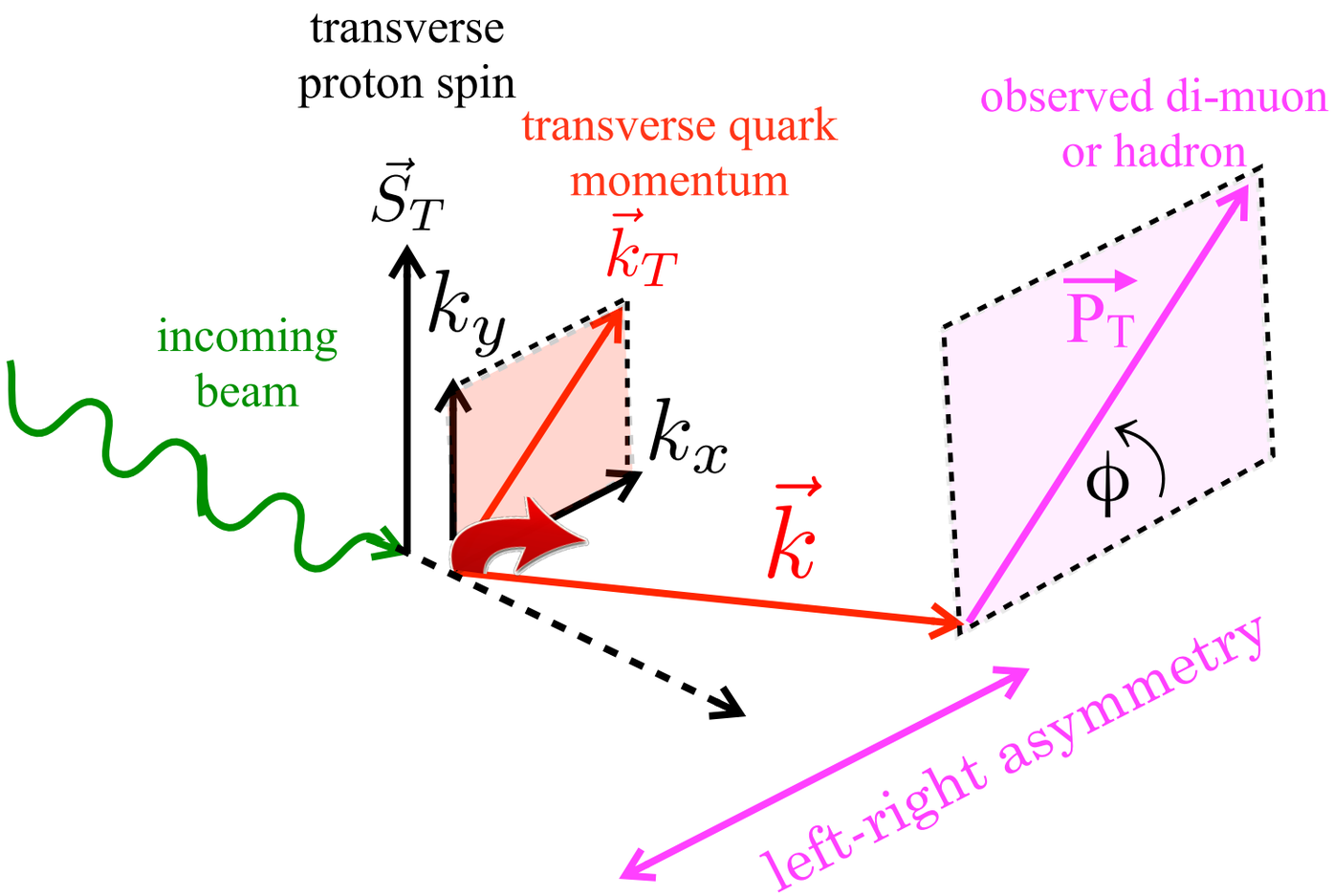
raw dy18		[TB]	chunks
dy18W01	P0	214	200,312
dy18W02	P0	116	108,693
dy18W03	P1	65	60,931
dy18W04	P1	62	58,422
dy18W05	P2	49	45,533
dy18W06	P2	71	66,655
dy18W07	P2	79	73,675
dy18W08	P2	55	51,807
dy18W09	P3	100	93,428
dy18W10	P3	78	72,869
dy18W11	P4	53	49,305
dy18W12	P4	48	45,398
dy18W13	P4	82	77,048
dy18W14	P5	64	60,146
dy18W15	P5	47	43,748
dy18W16	P6	59	55,418
dy18W17	P6	65	60,551
dy18W18	P7	71	67,077
dy18W19	P7	77	72,380
dy18W20	P7	52	48,401
dy18W21	P7	51	48,084
dy18W22	P8	55	50,951
dy18W23	P8	29	27,248

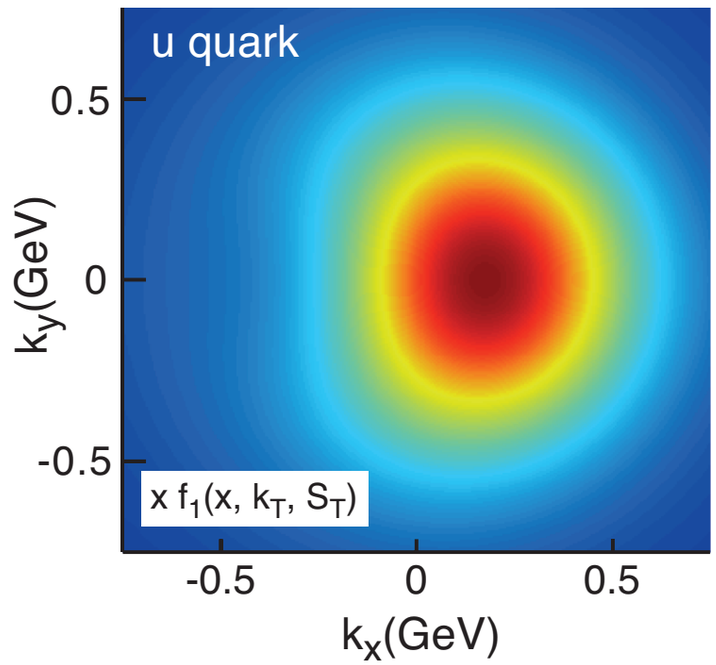


transfer example September 2018 (~420TB)

Sketches courtesy R. Longo

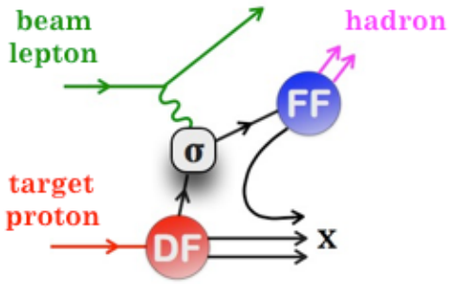
The Sivers effect





EIC "White Paper" arXiv:1212.1701, based on M. Anselmino et al., J. Phys. Conf. Ser. 295, 012062 (2011), arXiv:1012.3565

Semi-Inclusive Deep-Inelastic Scattering

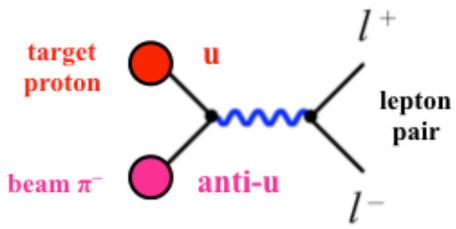


Sivers (proton) ⊗ unpolarized fragmentation function

The Sivers function is expected to have the **same magnitude** but **opposite sign** in DY. Crucial test of QCD-TMD framework.

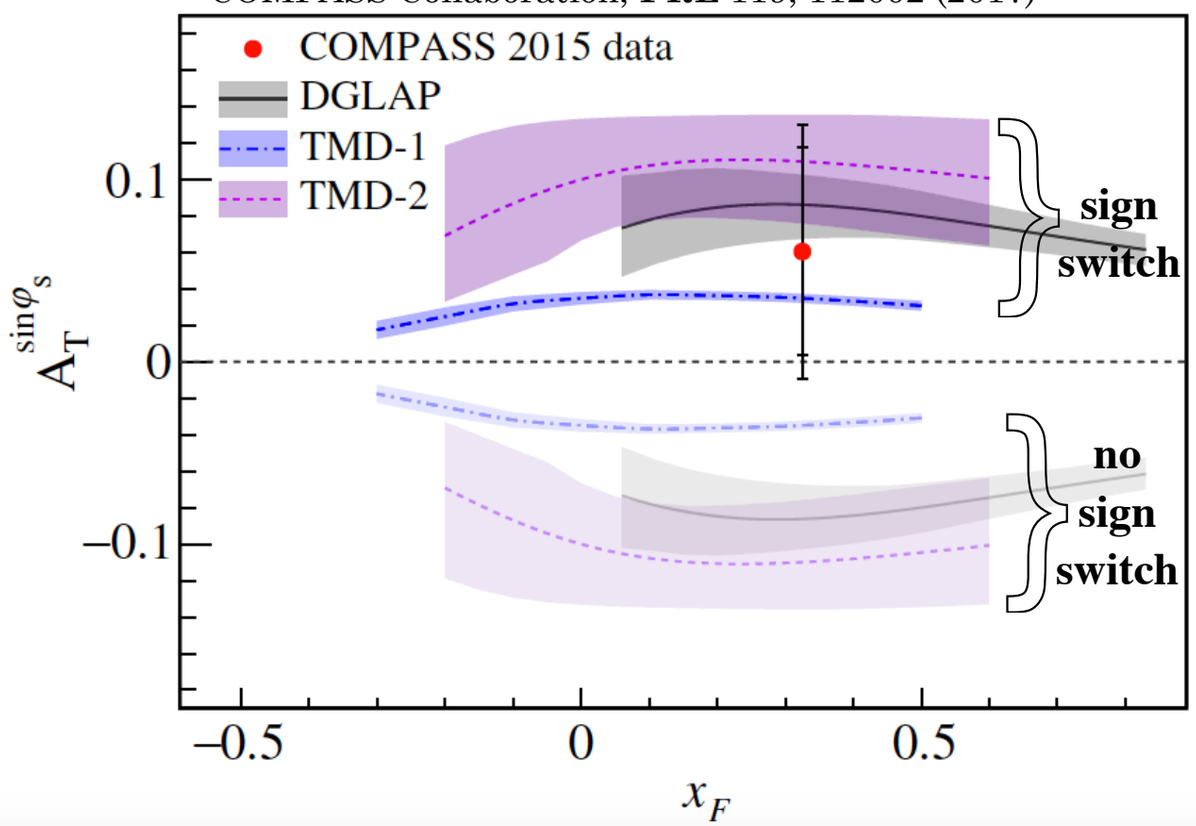
Sivers (SIDIS) = −(1)·Sivers (DY)

Drell-Yan



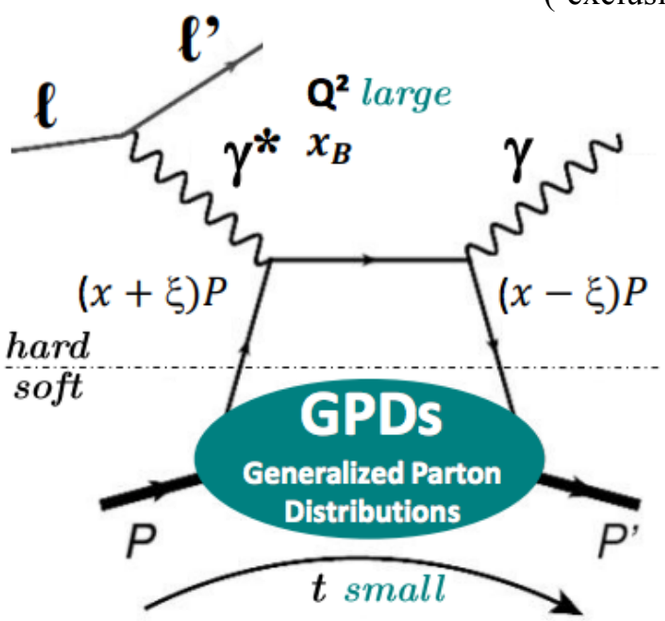
Sivers (proton) ⊗ unpolarized distribution function (pion)

COMPASS Collaboration, **PRL** 119, 112002 (2017)



Exclusive reactions, hadron multiplicities & spectroscopy

('exclusive' means all particles are detected. Particles made of quarks are called hadrons)

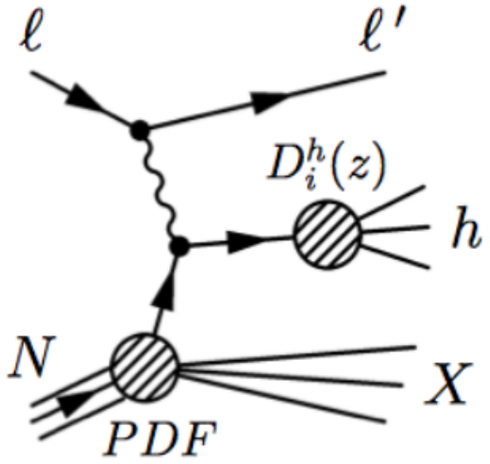


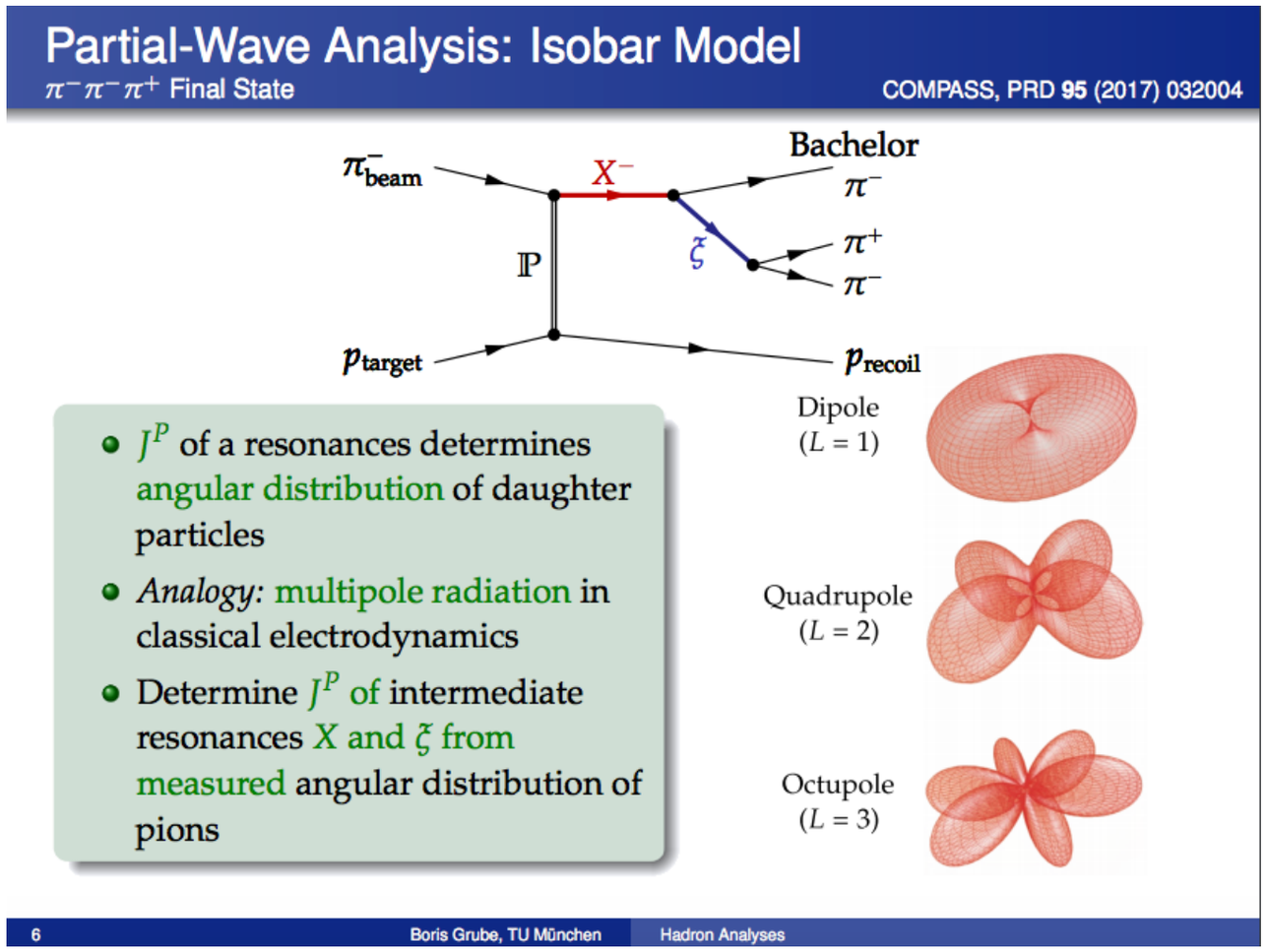
Fragmentation Functions (FFs) :

$$\frac{dM^h(x,Q^2,z)}{dz} = \frac{\sum_q e_q^2 q(x,Q^2) D_q^h(z,Q^2)}{\sum_q e_q^2 q(x,Q^2)}$$

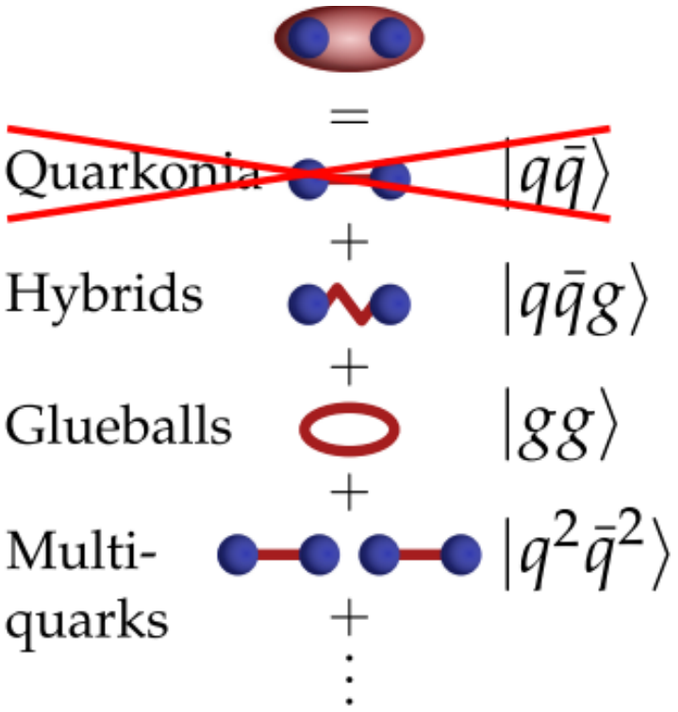
quark to hadron FFs

quark PDFs





QCD permits color-neutral configurations beyond $q\bar{q}$



"Exotic" mesons

Courtesy A. Vidon/N. D'Hose, N. Pierre, B. Grube

Blue Waters (NCSA) & Frontera (TACC): history

machine	name of allocation	requested node hours	allocated node hours	used node hours	proposal submission
Blue Waters	exploratory	40k	40k	55k	Mar 2016
Blue Waters	campus 17	960k	200k	200k	Sep 2016
Blue Waters	PRAC-A	9,440k	9,440k	3,860k	Nov 2016 *
Blue Waters	PRAC-B	9,440k	9,440k	9,440k	(together with PRAC-A)
Blue Waters	campus 19	1,000k	600k	588k	Oct 2018
Blue Waters	supplement (SP)	20,500k	0	n/a	Jan 2019 *
Blue Waters	campus 20	5,000k	0	n/a	Sep 2019
Frontera	early science (SP)	1,200k **	450k	140k	Jan 2019 *
Frontera	LRAC (DCL)	1,500k	1,500k	528k	Jan 2020
Frontera	travel proposal	- ($ only)	- ($ only)	n/a	April 2020 *
XSEDE	startup	- (for grid transfer only)	- (for grid transfer only)	n/a	June 2020
Frontera	renewal	???	tba	tba	deadline Jan 15, 2021

	Jan	Feb	Mar	Apr	May	June	July	Aug	Sep	Oct	Nov	Dec
2016												
2017												
2018												
2019												
2020												
2021												

SP = NSF supplemental proposal
DCL = NSF Dear Colleague Letter
* = proposal submitted to NSF via Fastlane, otherwise submitted directly to supercomputing institution
** requested COMPASS & AMBER, got only COMPASS
1 Frontera nh is about 5 BW nh