André Walker-Loud
Staff Scientist
Lawrence Berkeley National Laboratory
S91010 - Accelerating our Understanding of the Nuclear Physics and the Early Universe

Accelerating our Understanding of the Nuclear Physics and the Early Universe

André Walker-Loud

Accelerating our Understanding of ~~the~~ Nuclear Physics and the Early Universe

André Walker-Loud

BERKELEY LAB

This Talk

- What are the driving science questions we are trying to understand?
- Why do we need High-Performance-Computing?
- How do we make optimal use of the Leadership Class Supercomputers (like Summit at Oak Ridge National Laboratory)?
- Preliminary new results from new machines

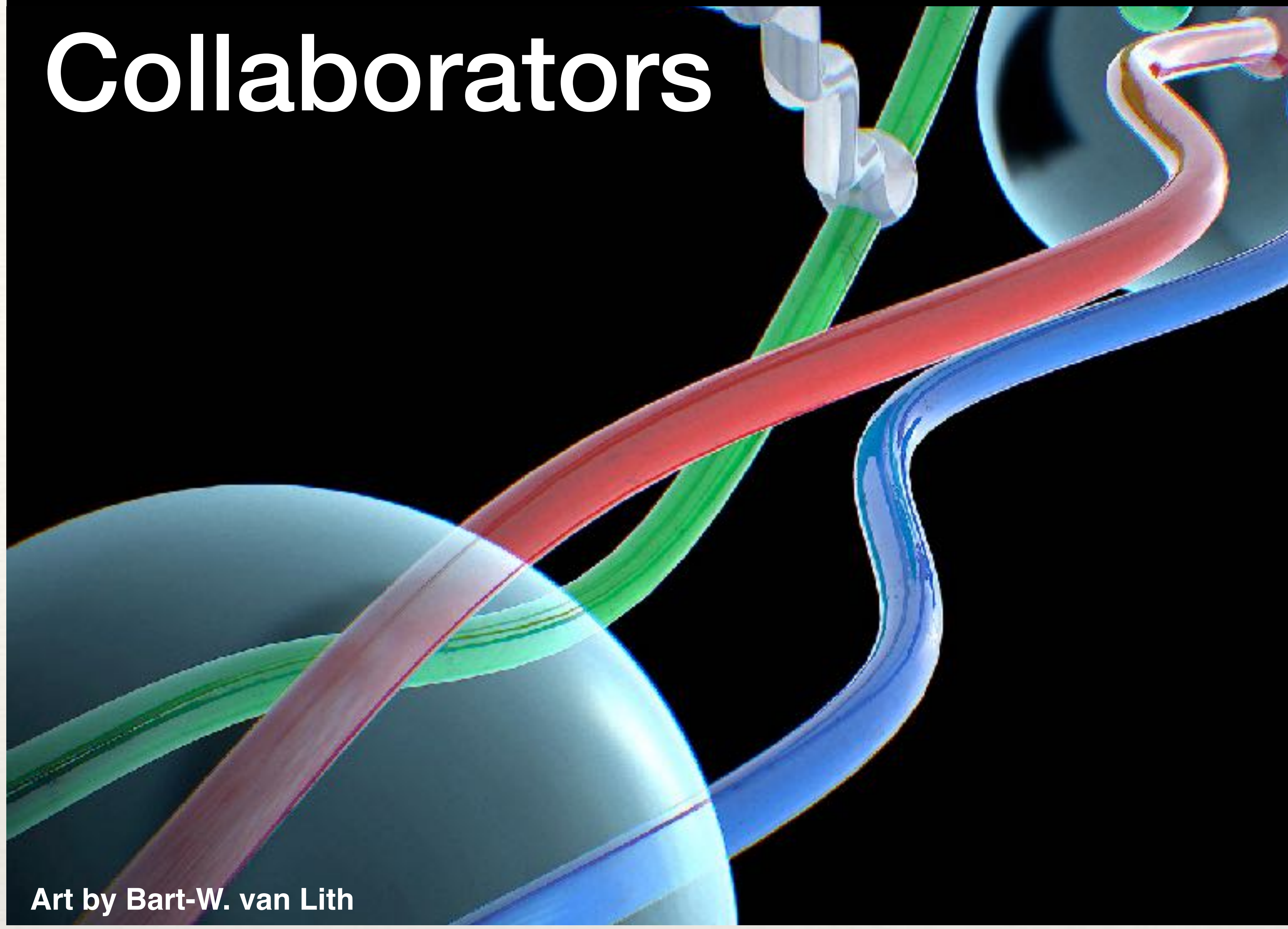
Acknowledgements

- I would like to thank Jack Wells for suggesting us for this talk

- We would like to thank the Lawrence Livermore CORAL team
  Scott Futral, Greg Tomaschke, Adam Bertsch, John Gyllenhal, Py Watson
  for providing us early access to Sierra and help understanding how to use the machine

- We would like to thank the Oak Ridge Leadership Computing Facility team
  Jack Wells, Tjerk Straatsma, Chris Fuson, Gustav Jansen, …
  for providing us early access (and an Early Science Award) on Summit and help understanding how to use the machine

Collaborators



Art by Bart-W. van Lith

Gordon Bell
Evan Berkowitz
Kate Clark
Arjun Gambhir
Ken McElvain
Amy Nicholson
Enrico Rinaldi
Pavlos Vranas
André Walker-Loud
Chia Cheng (Jason) Chang
Balint Joó
Thorsten Kurth
Kostas Orginos

Nucleon Axial Coupling
Chia Cheng (Jason) Chang
Amy Nicholson
Enrico Rinaldi
Evan Berkowitz
Nicolas Garron
David Brantley
Henry Monge-Camacho
Chris Monahan
Chris Bouchard
Kate Clark
Balint Joó
Thorsten Kurth
Kostas Orginos
Pavlos Vranas
André Walker-Loud

INSTITUTE for NUCLEAR THEORY
Berkeley UNIVERSITY OF CALIFORNIA
NVIDIA
RIKEN BNL Research Center
BERKELEY LAB
Lawrence Livermore National Laboratory
California Lattice Collaboration
WILLIAM & MARY
University of Glasgow
Jefferson Lab Thomas Jefferson National Accelerator Facility
UNIVERSITY OF LIVERPOOL
NERSC
iTHEMS
THE UNIVERSITY of NORTH CAROLINA at CHAPEL HILL
JÜLICH FORSCHUNGSZENTRUM

*not all in California

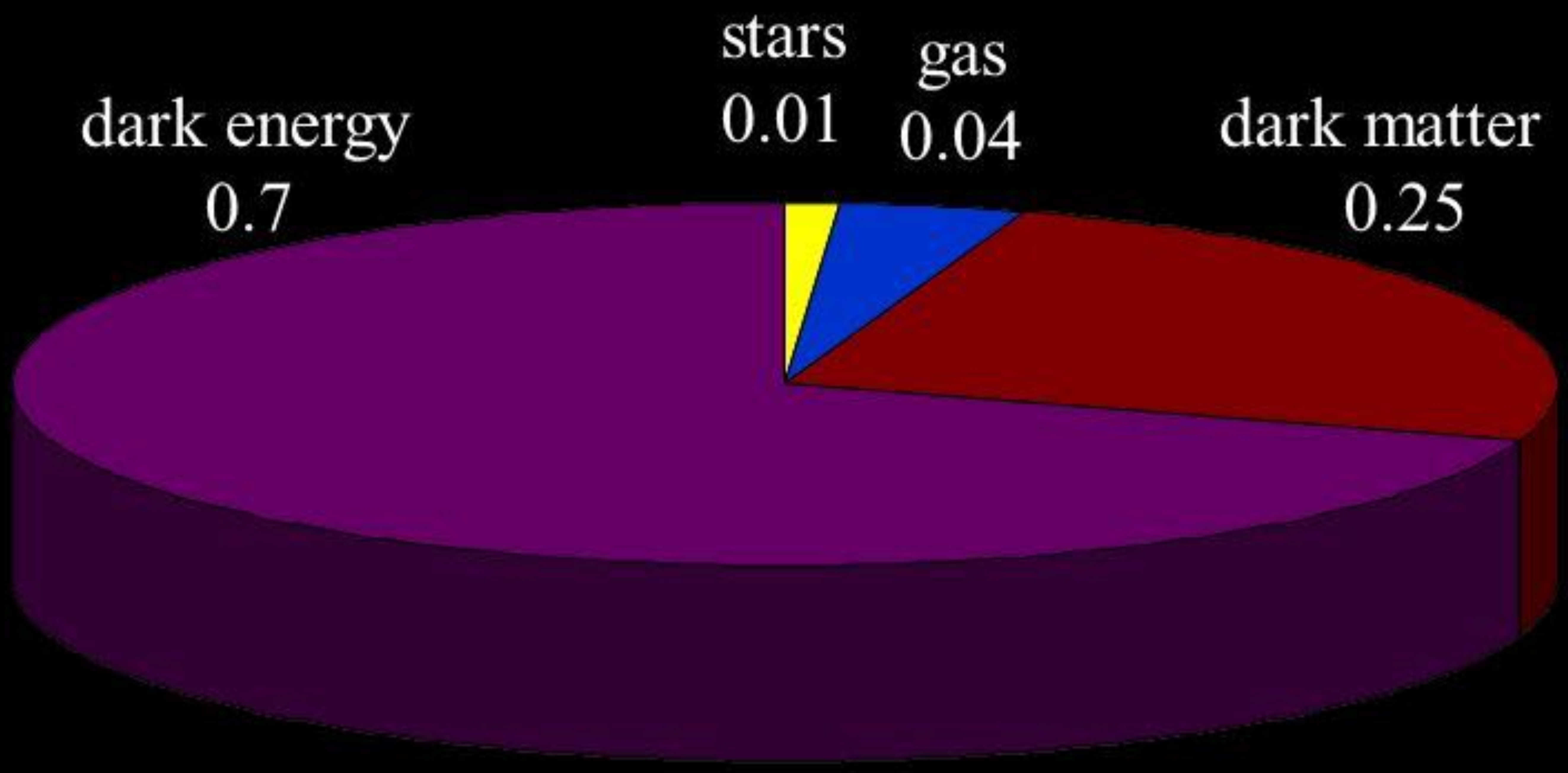
These calculations are made possible by

Lawrence Livermore National Laboratory
U.S. DEPARTMENT OF ENERGY INCITE LEADERSHIP COMPUTING
OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY
SciDAC Scientific Discovery through Advanced Computing
DOE Topical Collaboration Double Beta Decay
NNSA National Nuclear Security Administration
U.S. DEPARTMENT OF ENERGY | Office of Science

Science Drivers

- Can we understand the properties of matter (protons, neutrons, atomic nuclei) directly from the Standard Model of particle physics?
- Can we understand the evolution of stars from their solar fusion cycle through supernovae explosions and collapse to neutron stars and black holes (directly from the Standard Model)?
  - See Bronson Messer's talk just prior - we are hoping to start a new effort connecting our research
- Why does the universe contain more matter than anti-matter?
- …

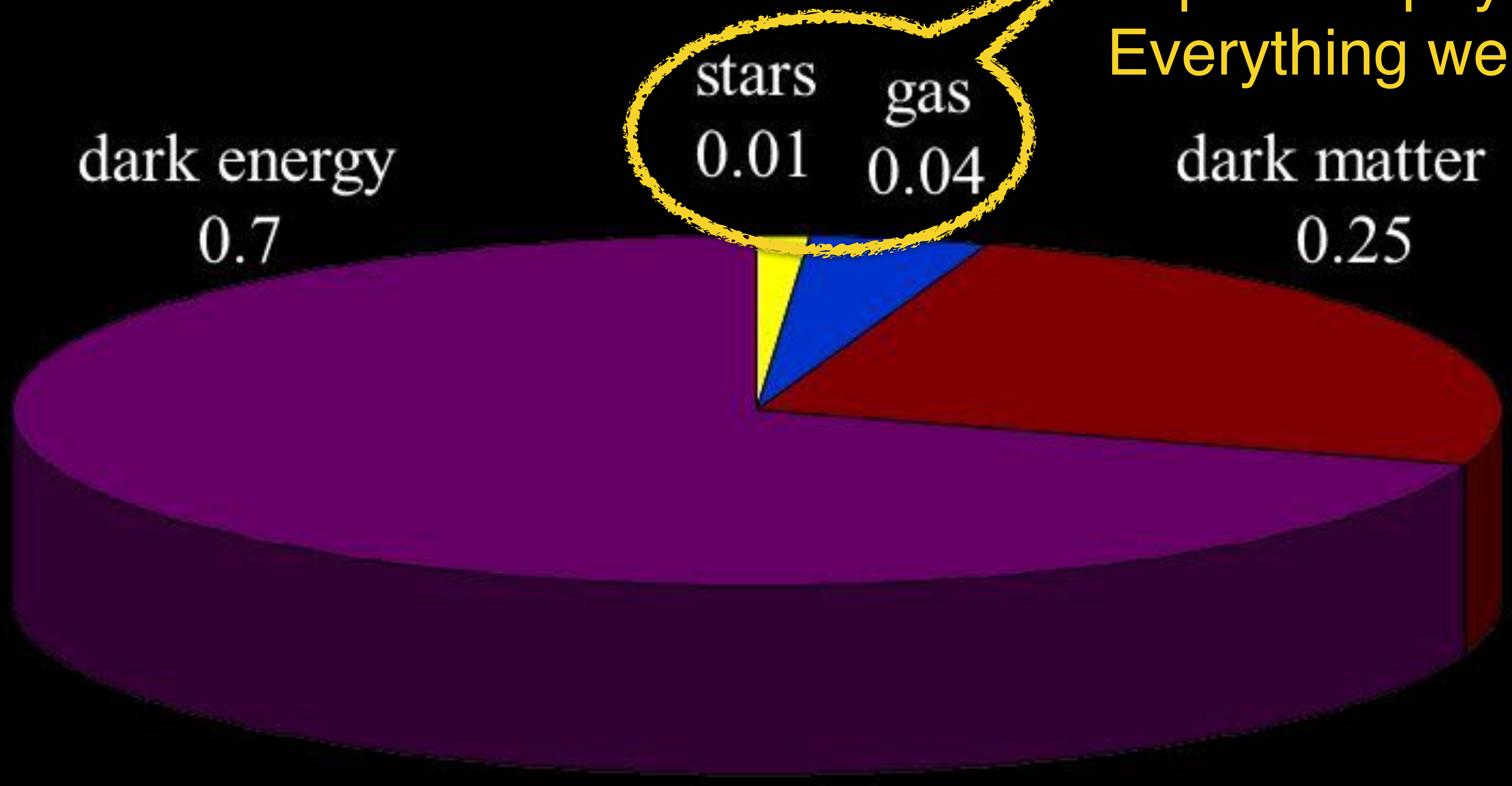
cosmic mass/energy budget



$$\Omega_{mass+energy} = 1 \text{ (or very close)}$$

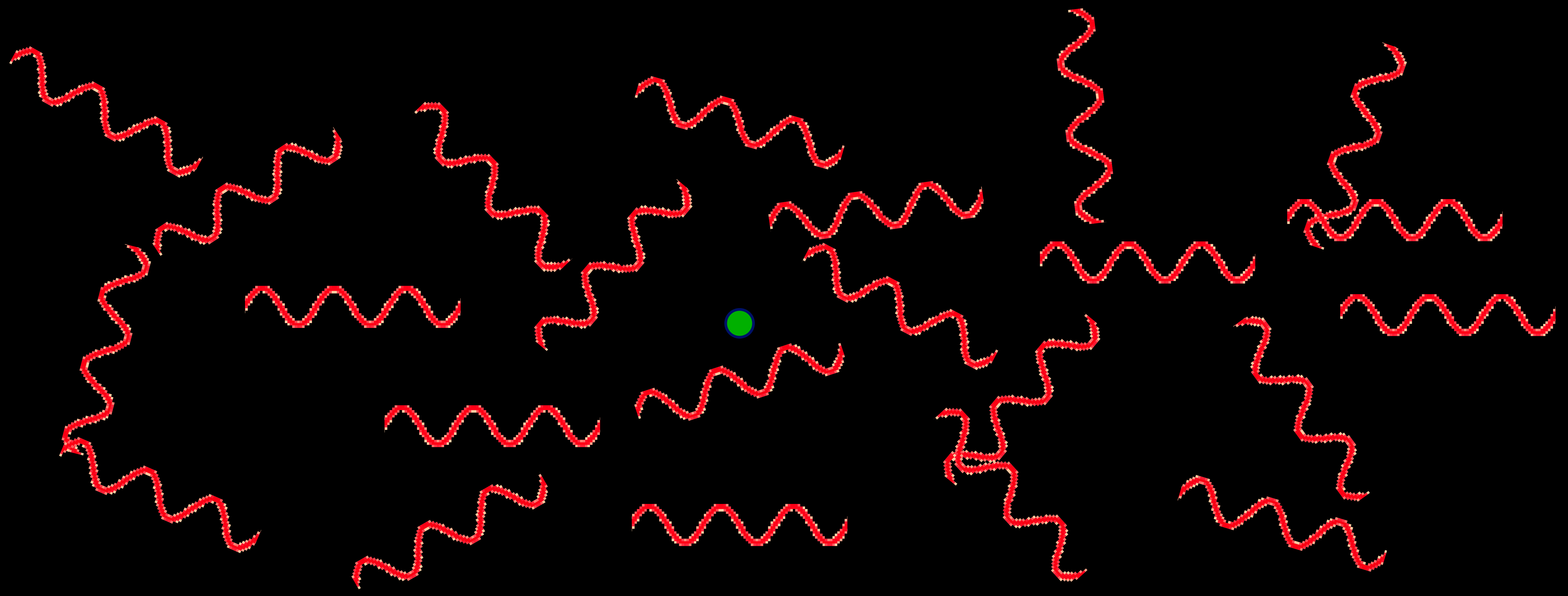
cosmic mass/energy budget

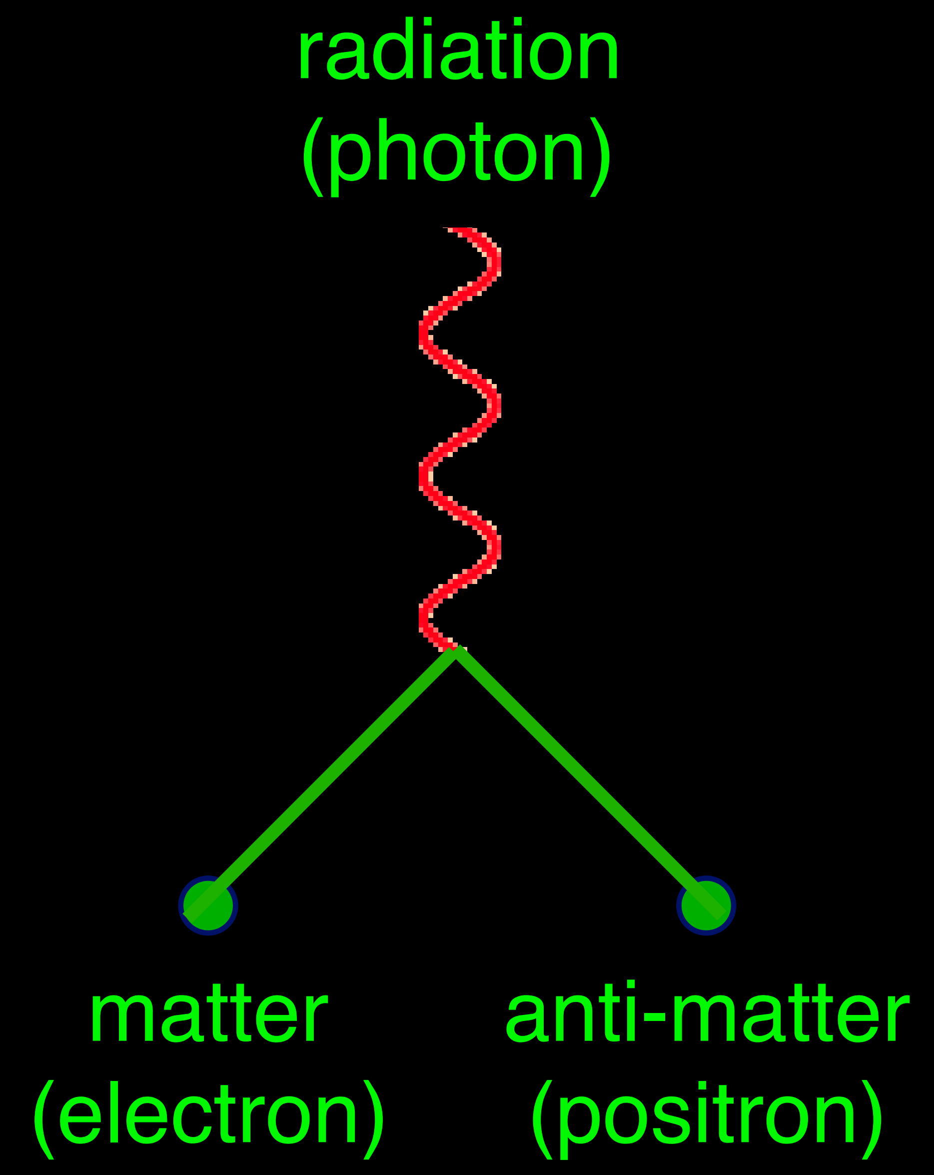
Standard Model
of particle physics
Everything we know…



$$\Omega_{\text{mass+energy}} = 1 \text{ (or very close)}$$

To the best of our knowledge, the Standard Model matter in the Universe is comprised entirely of matter and not anti-matter

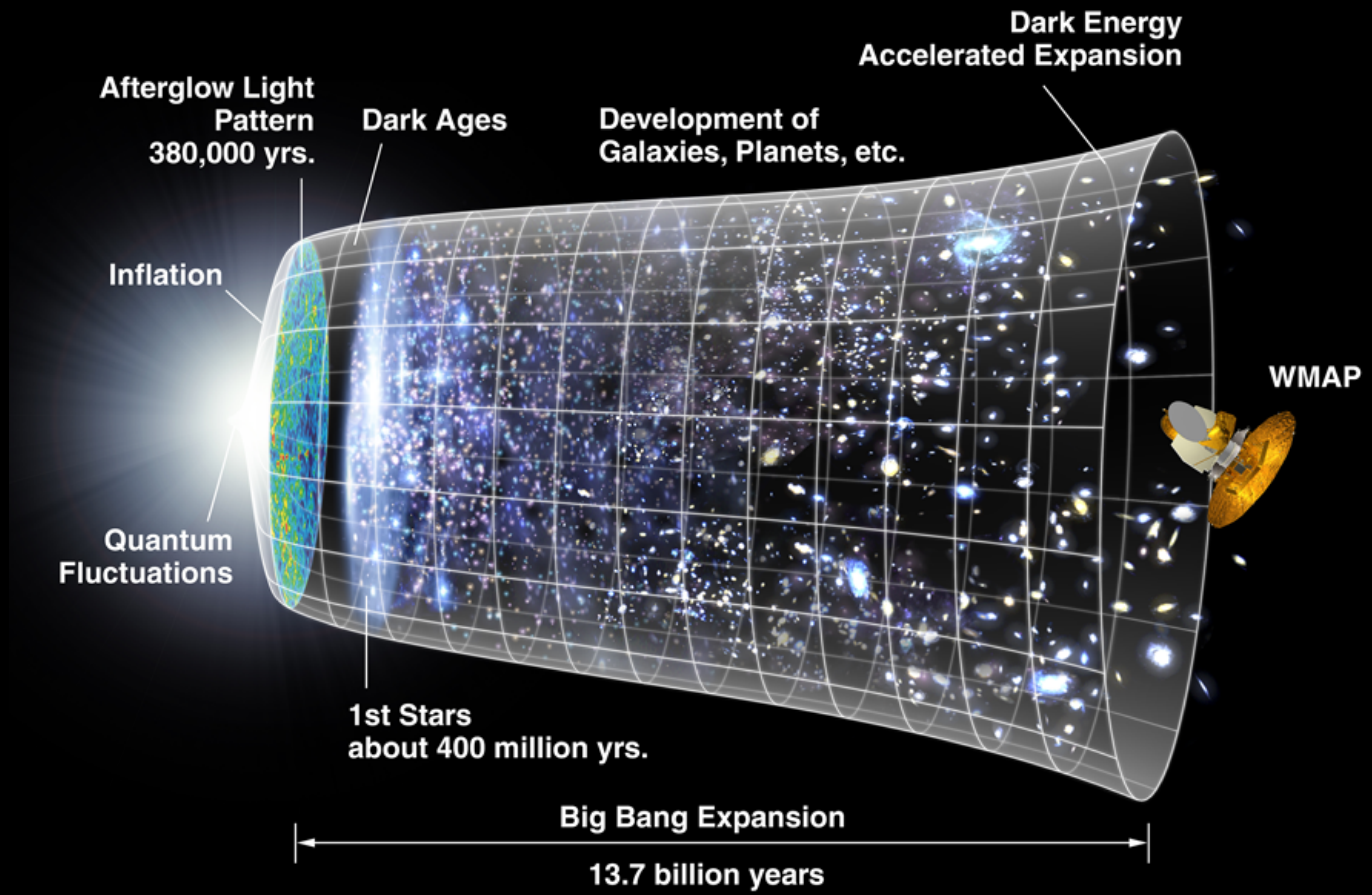




We observe there are more protons than anti-protons in the universe by an amount of roughly

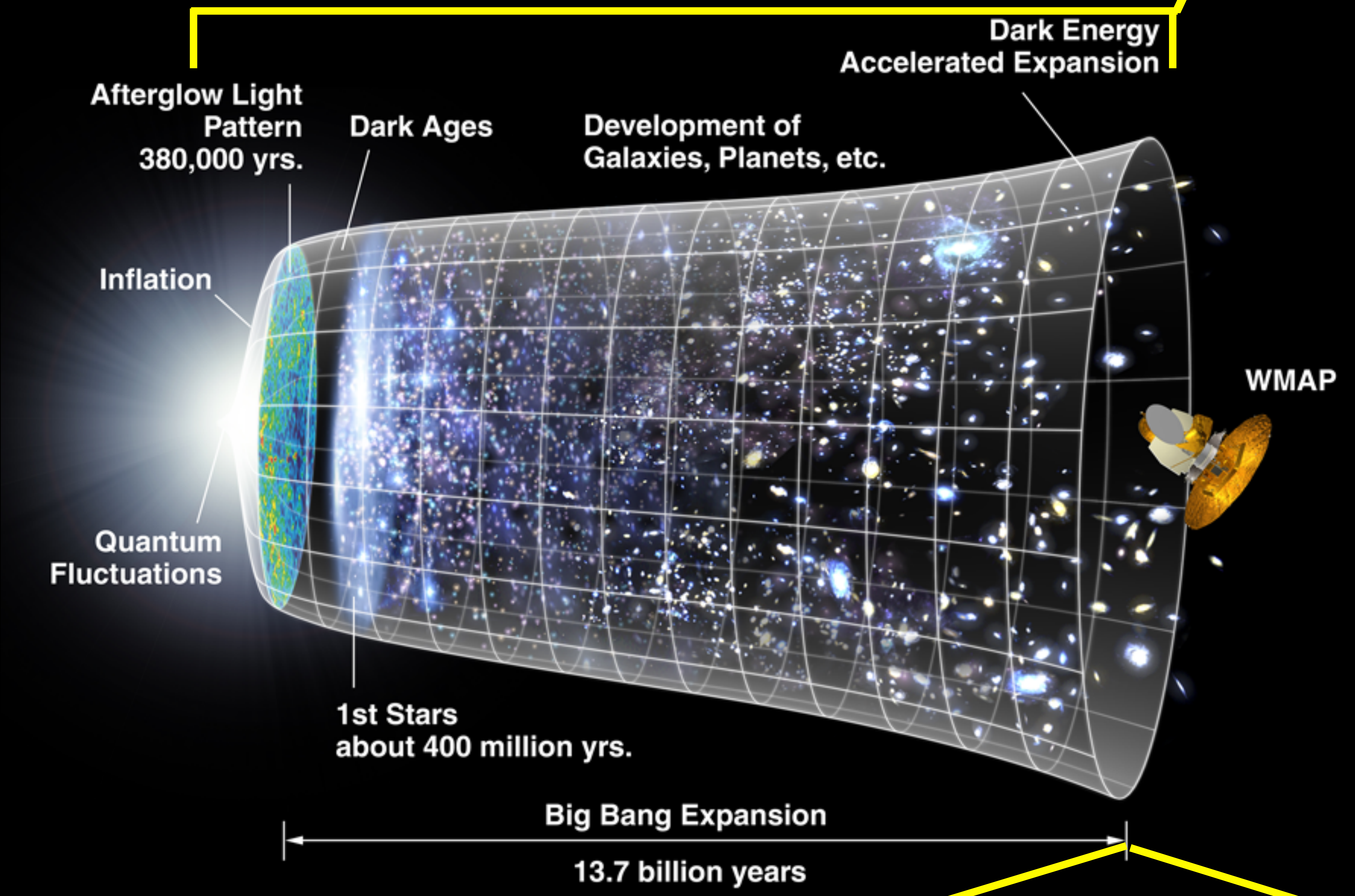
$$\frac{N_{proton} - N_{anti-proton}}{N_{photon}} \simeq 10^{-9}$$

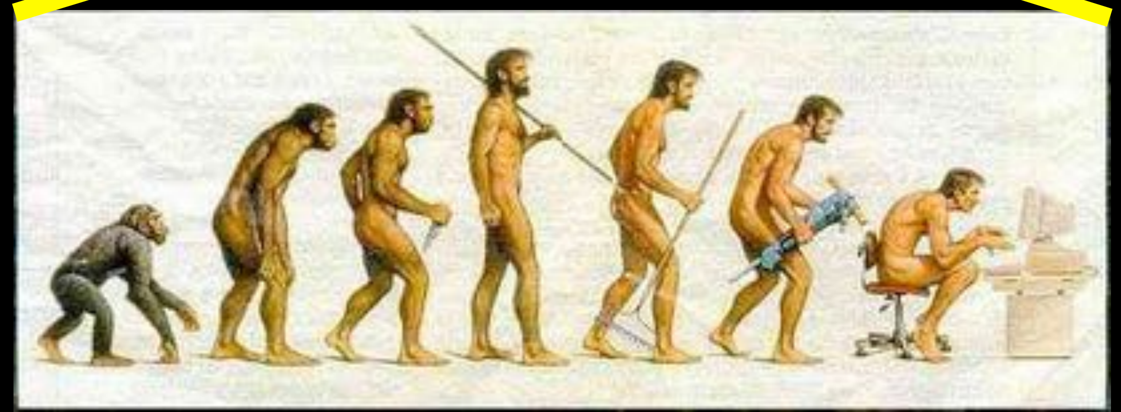
While tiny, this is still **10,000** times or more greater than we would predict with the Standard Model - why is there so much matter in the universe?

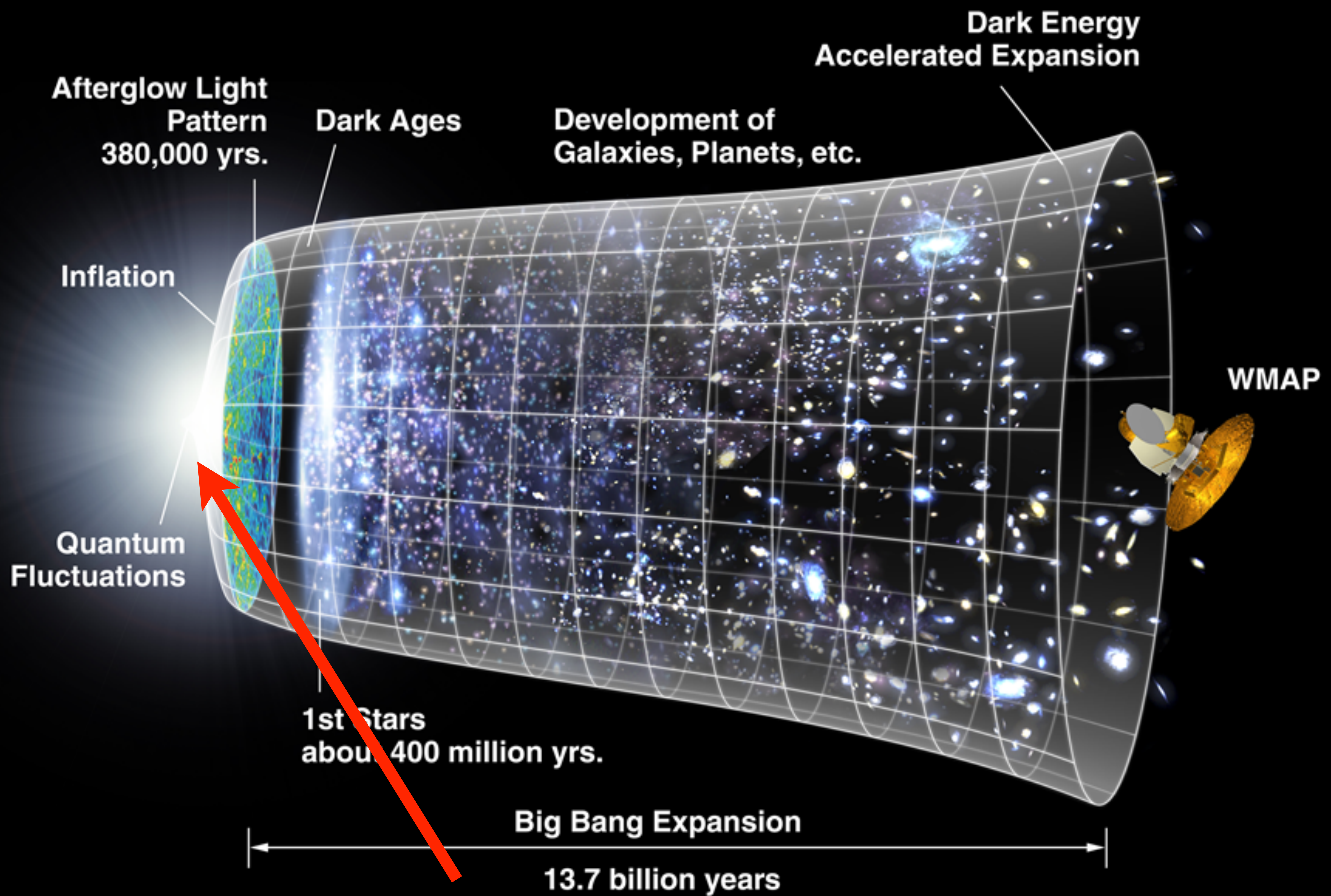


Dark Energy
Accelerated Expansion
Afterglow Light
Pattern
380,000 yrs.
Dark Ages
Development of
Galaxies, Planets, etc.
Inflation
WMAP
Quantum
Fluctuations
1st Stars
about 400 million yrs.
Big Bang Expansion
13.7 billion years

Standard Model of Particle Physics





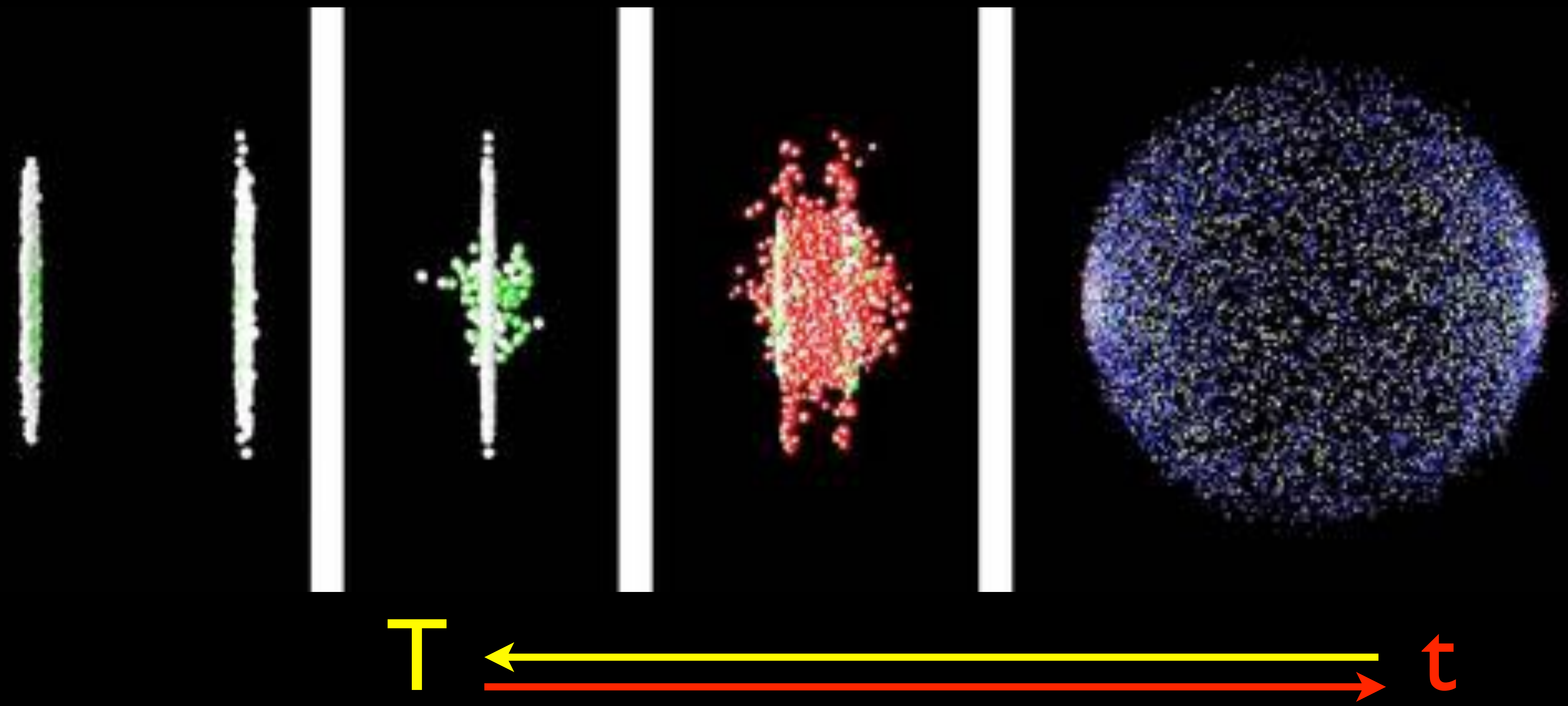


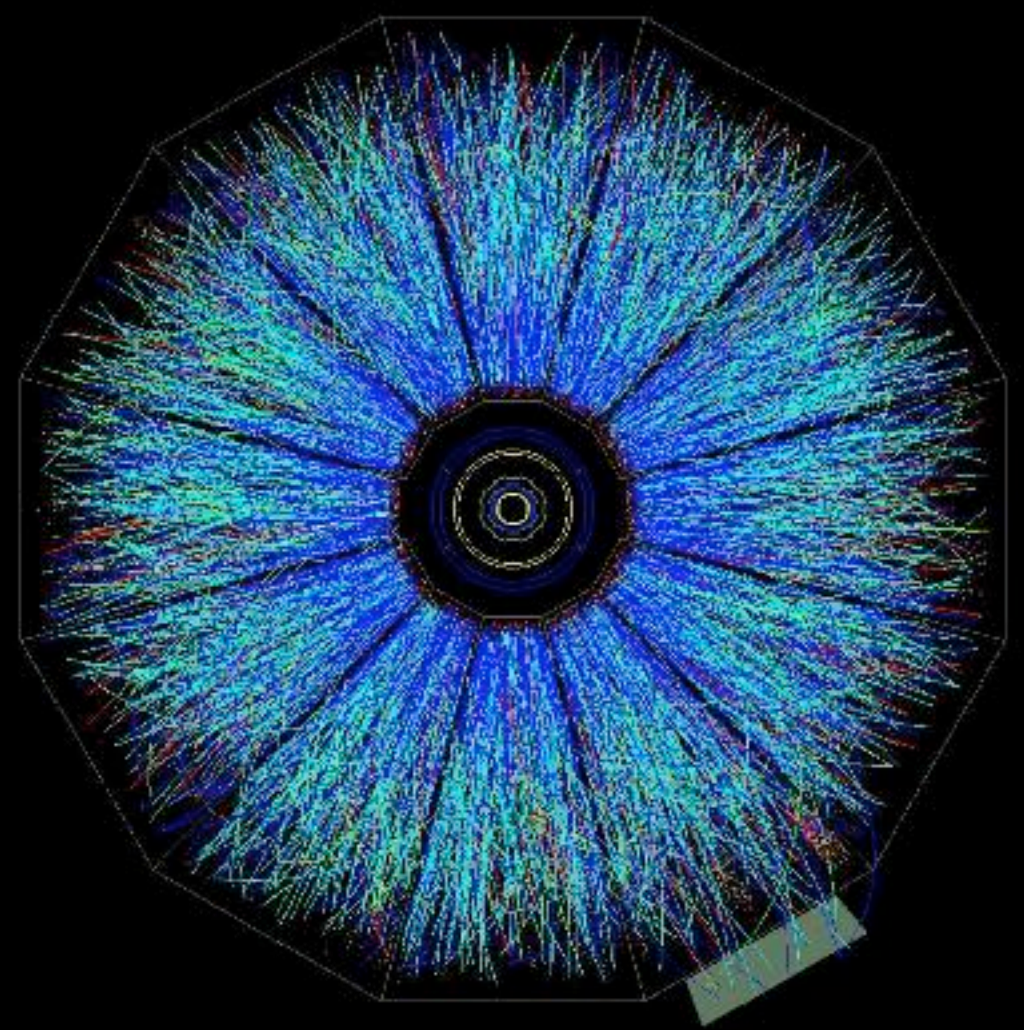
Formation of Matter

T ≃ 1 trillion K (10^{12} K)

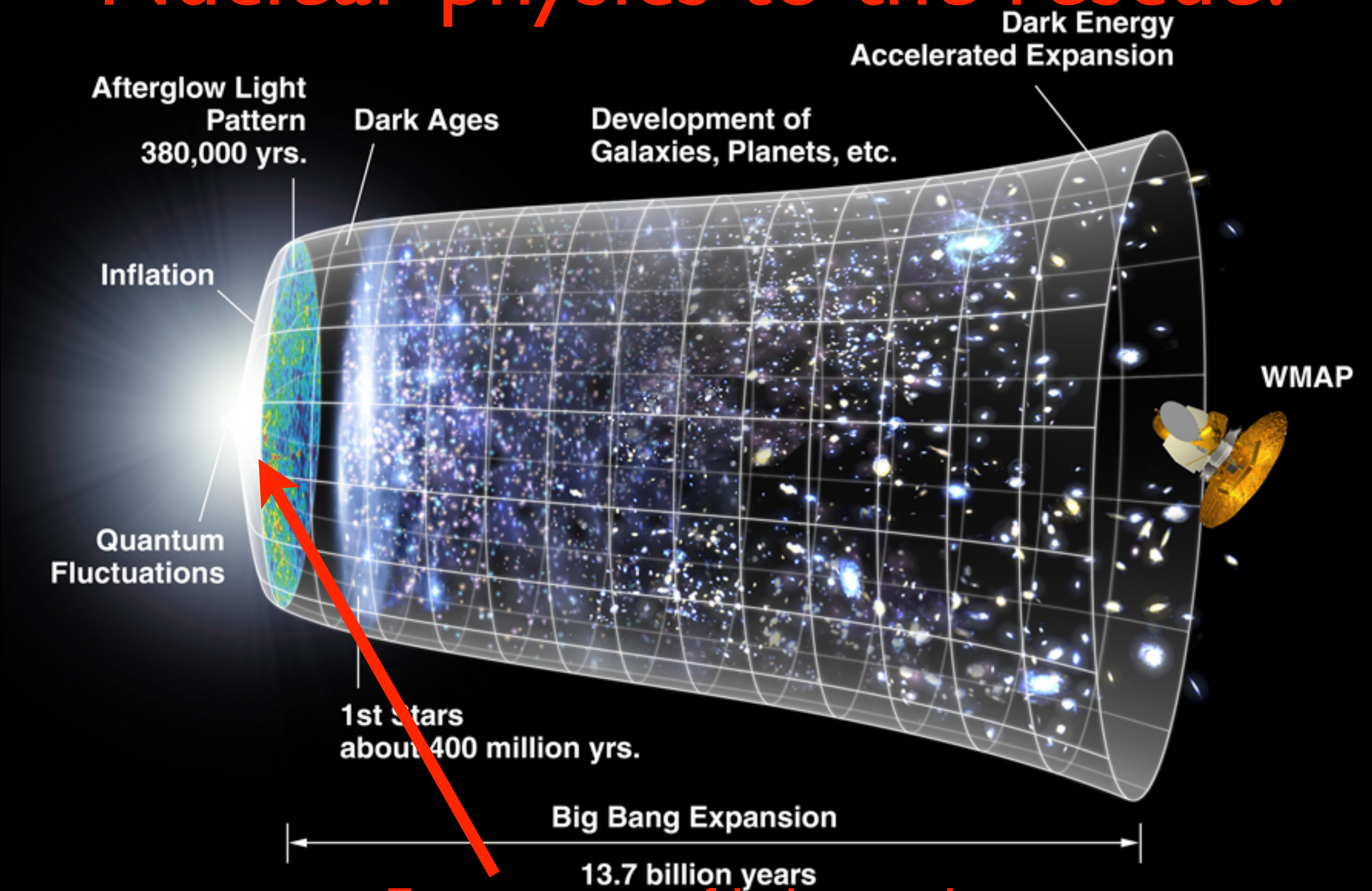
t ≃ 30 micro seconds (3×10^{-5} s)


Sextant 10/11
12:00 o'clock
Sextant 12/1
PHOBOS
10:00 o'clock
BRAHMS
2:00 o'clock
RHIC
4:00 o'clock
Sextant 6/7
STAR
Sextant 4/5
(First sextant to be tested)
BOOSTER
G-2
LINAC
AGS
TANDEMS


T
t



Nuclear physics to the rescue!

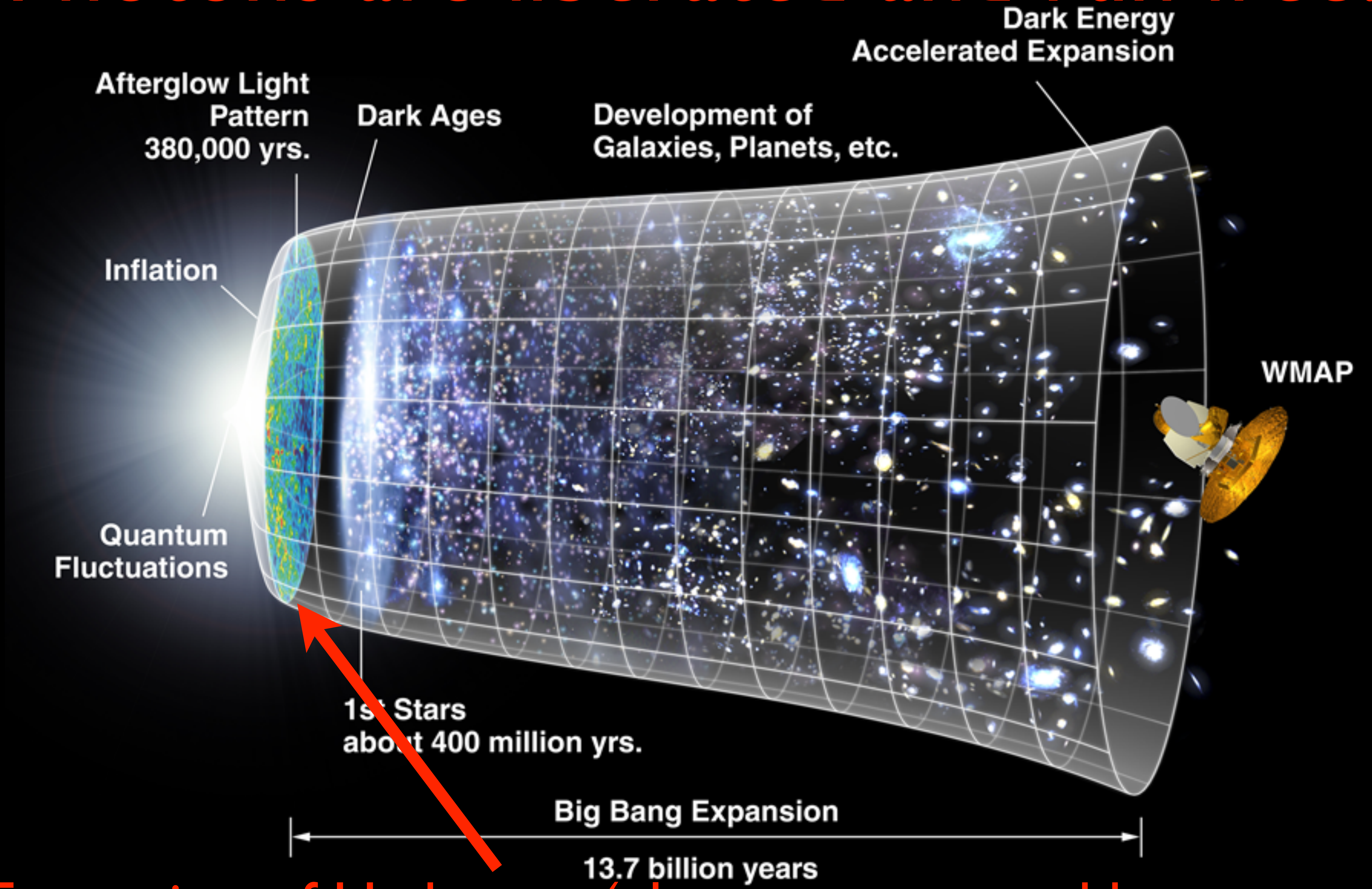


Formation of light nuclei

$T \simeq 1$ billion K (10^9 K)

$t \simeq 3$ minutes

Photons are liberated and run free!

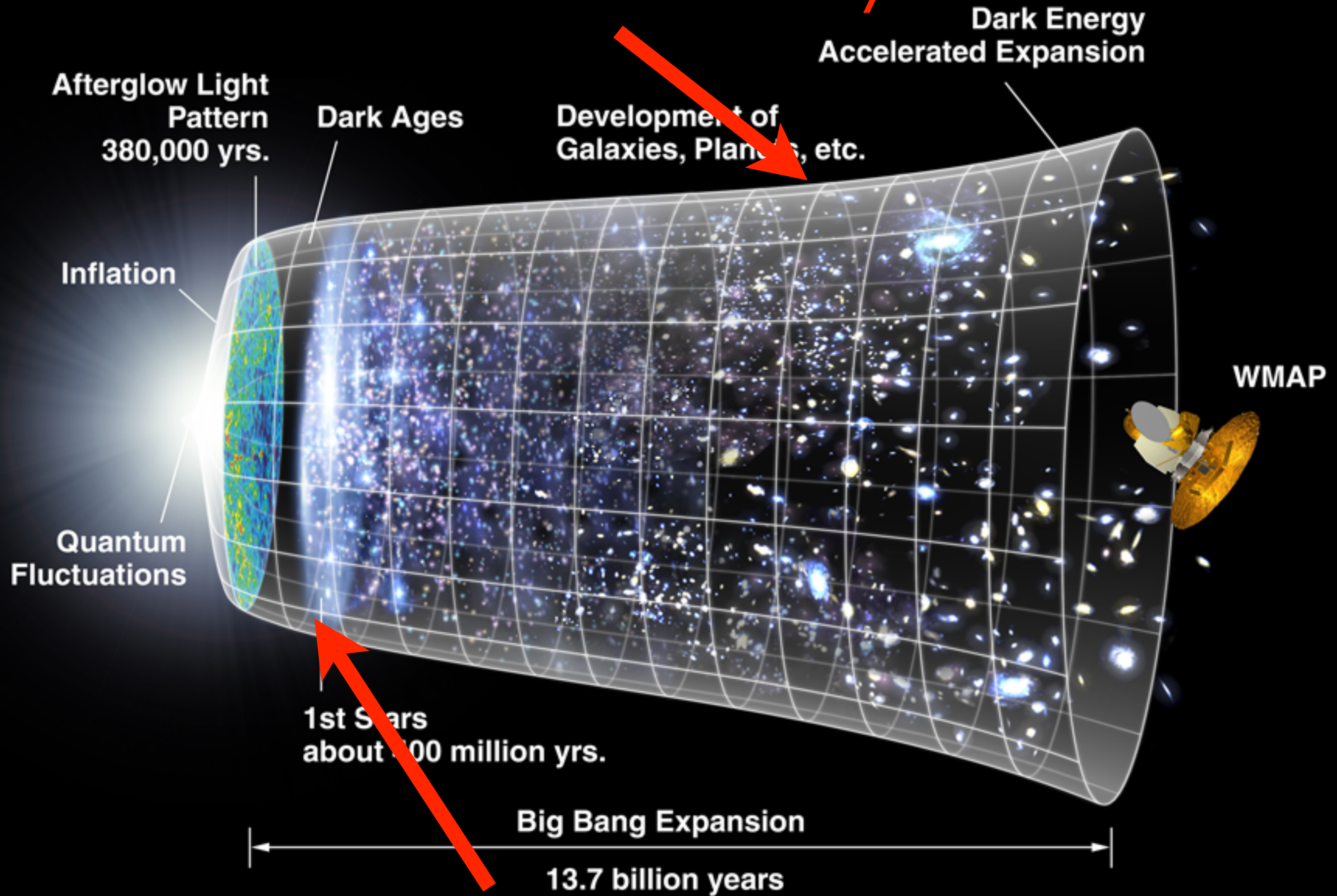


Formation of Hydrogen (electrons captured by protons)

$T \simeq 4{,}000$ K

$t \simeq 380{,}000$ years

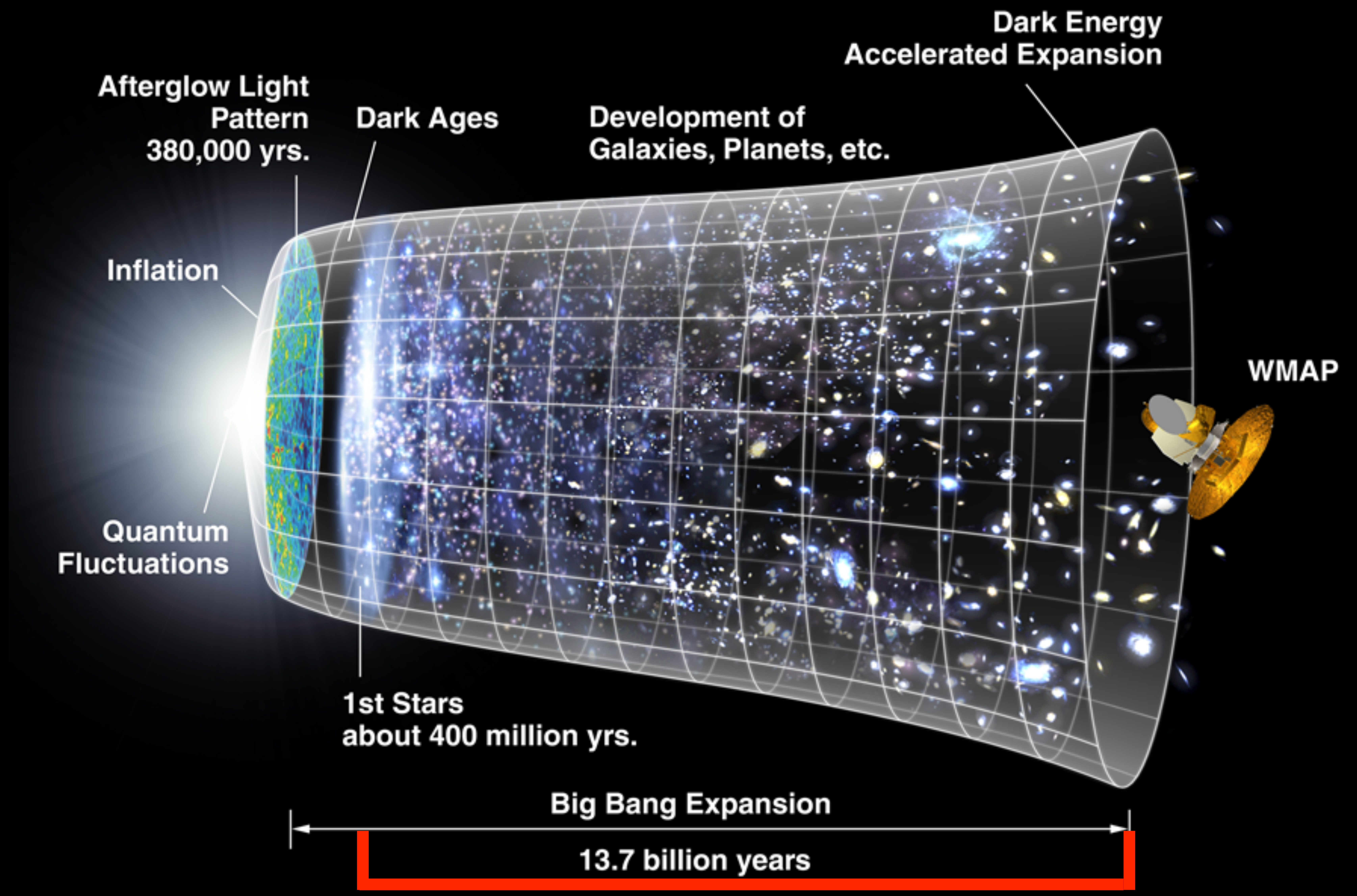
Formation of our solar system



Formation of first stars

$T \simeq 20$ K

$t \simeq 200$ Million years



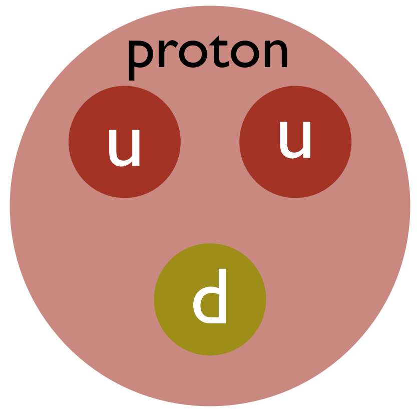
Death of stars, creation of heavy nuclei and *life*, creation of new, *ultradense* states of nuclear matter

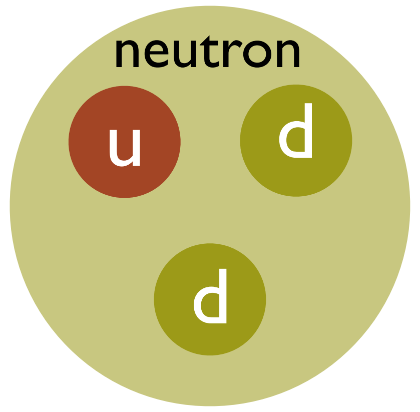
From Quarks to Protons and Neutrons

$$T \simeq 1 \text{ trillion K} \;\; (10^{12} \text{ K})$$

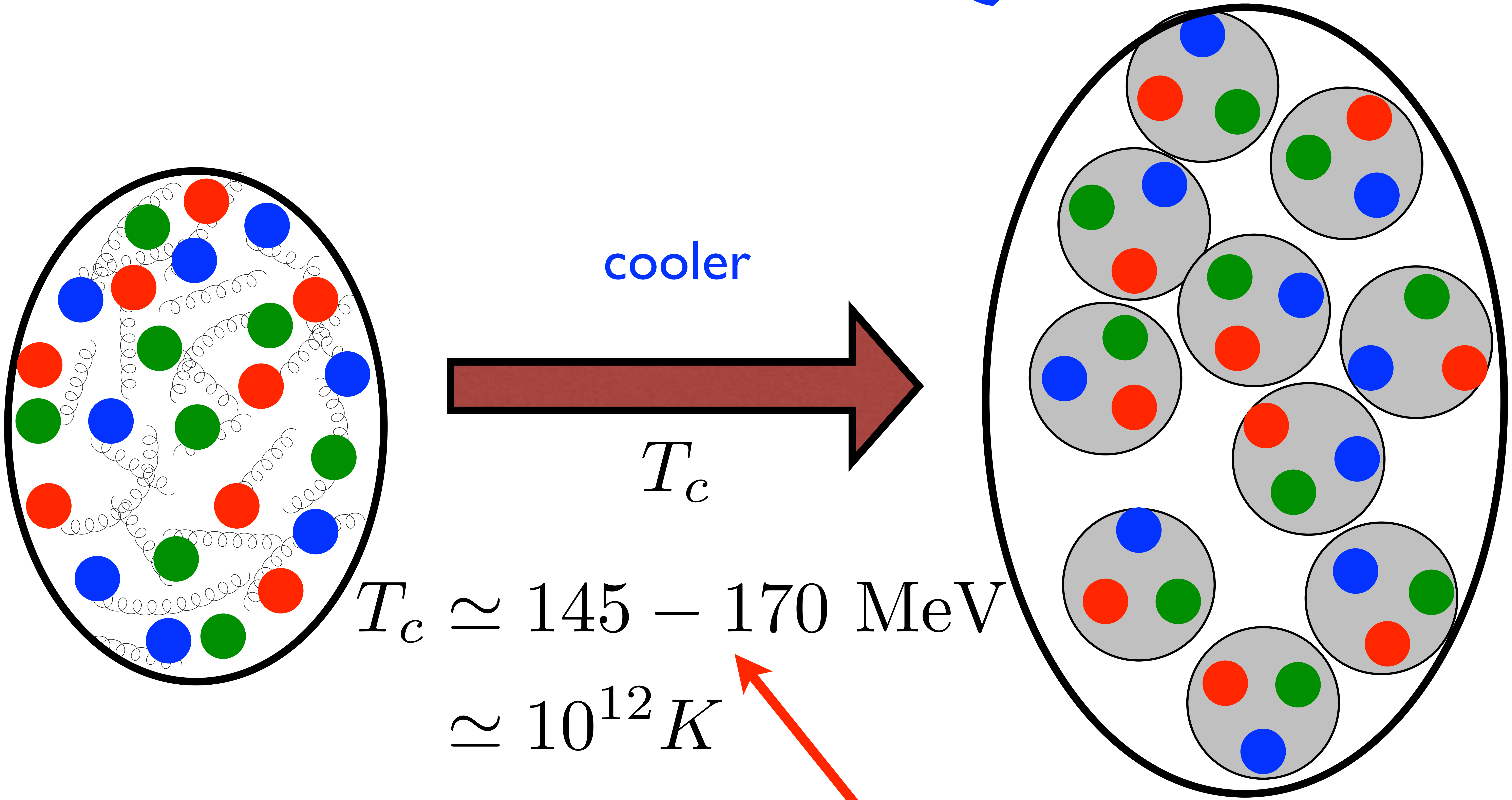
$$t \simeq 30 \text{ micro seconds} \, (3.0 \times 10^{-5} s)$$

- protons and neutrons are not fundamental - but they are composite states of *quarks* and *gluons*





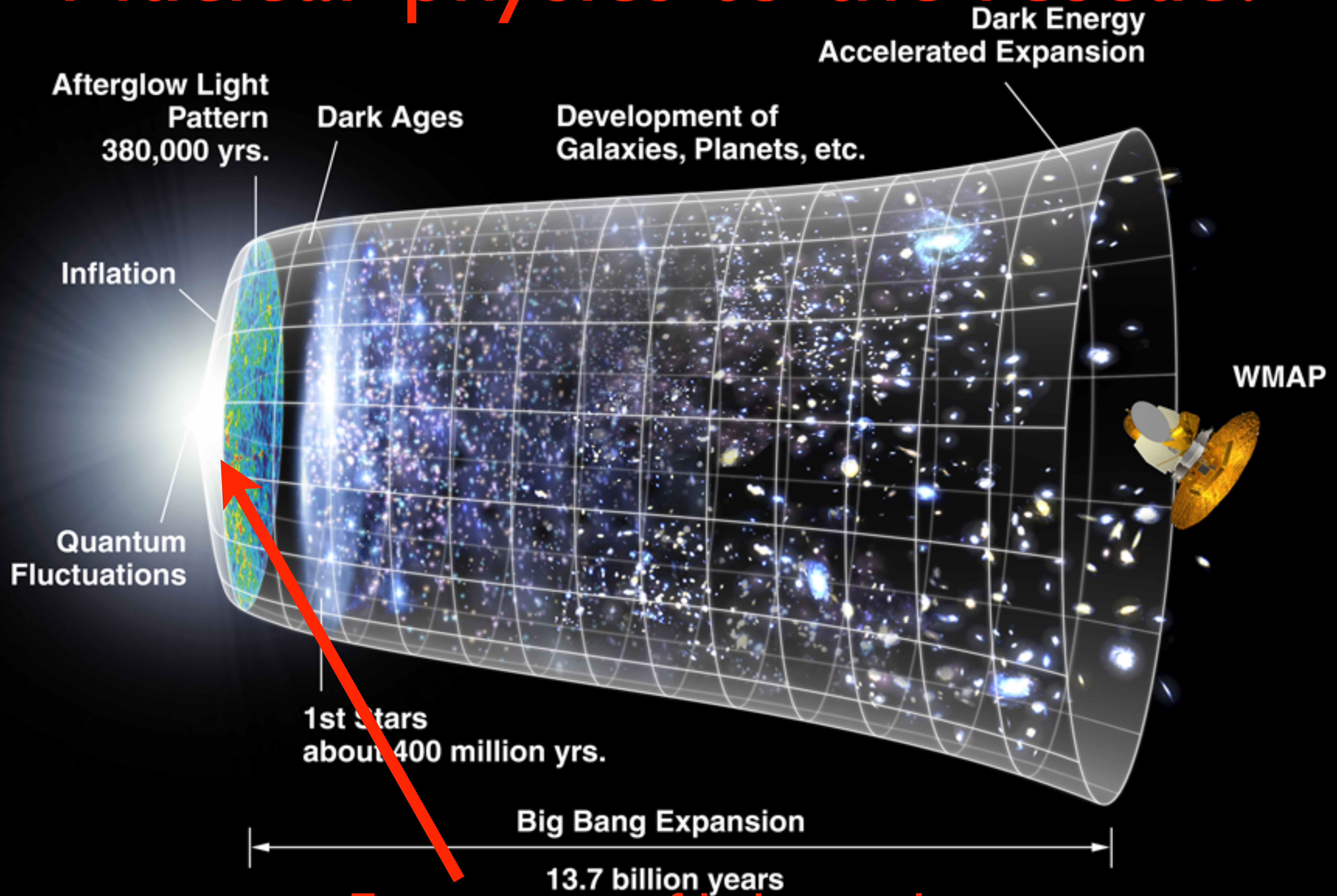
Confinement of Quarks



$$T_c \simeq 145 - 170 \text{ MeV}$$
$$\simeq 10^{12} K$$

computed by hot-QCD and Budapest-Wuppertal Lattice Collaborations with previous generation supercomputers

Nuclear physics to the rescue!



Formation of light nuclei

$T \simeq 1$ billion K (10^9 K)

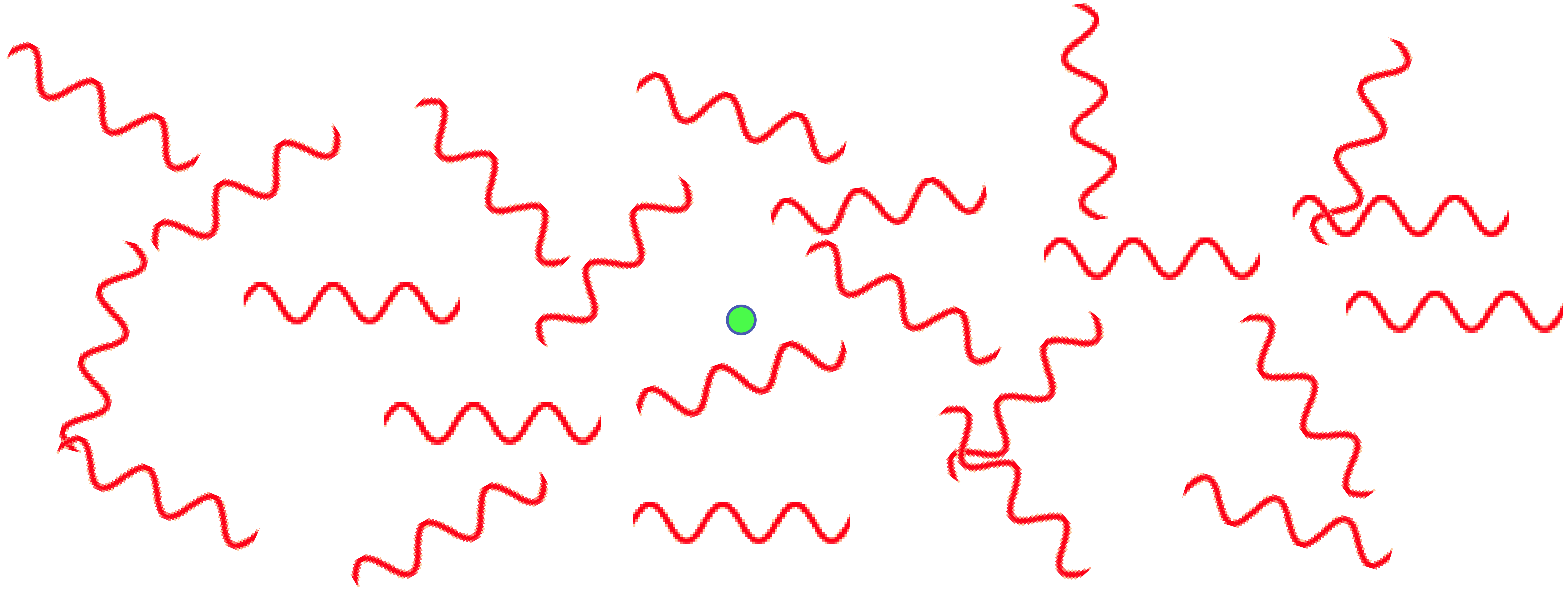
$t \simeq 3$ minutes

Big Bang Nucleosynthesis

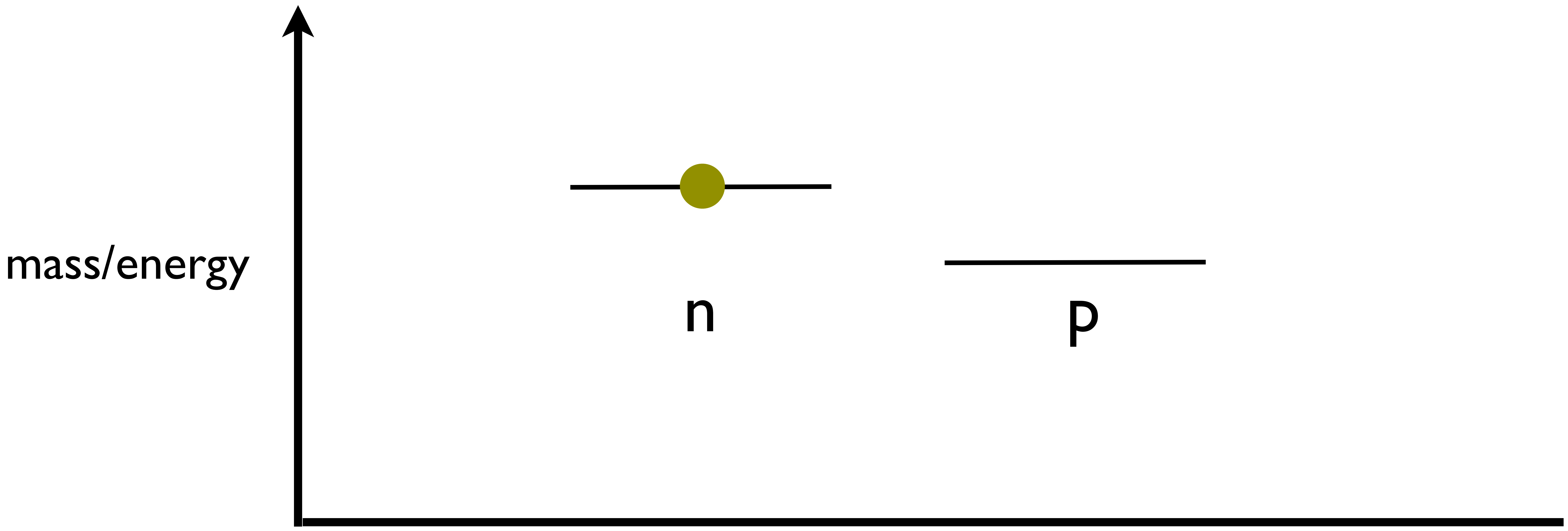
$$T \simeq 1 \text{ trillion K} \to 1 \text{ billion K}$$

$$t \simeq 3 \times 10^{-5} s \to 3 \text{ min}$$

Our initial condition is a soup of radiation plus a small excess amount of matter, in the form of protons, neutrons, electrons and photons



when systems cool, they settle into the lowest energy state



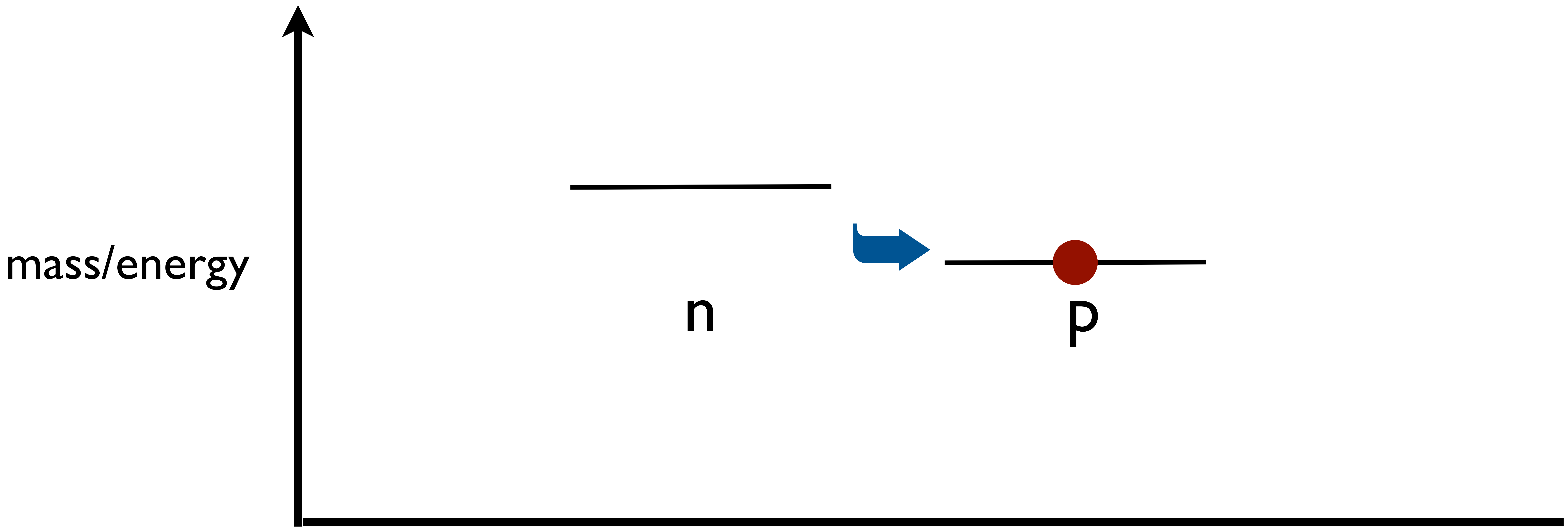
$$M_n - M_p = 1.29333217(42)\ \text{MeV}$$

$$\frac{M_n + M_p}{2} = 938.9187473(58)\ \text{MeV}$$

$$E_{\text{Hydrogen}} = 13.6\ \text{eV}$$

$$1\ \text{MeV} = 10^6\ \text{eV}$$

when systems cool, they settle into the lowest energy state

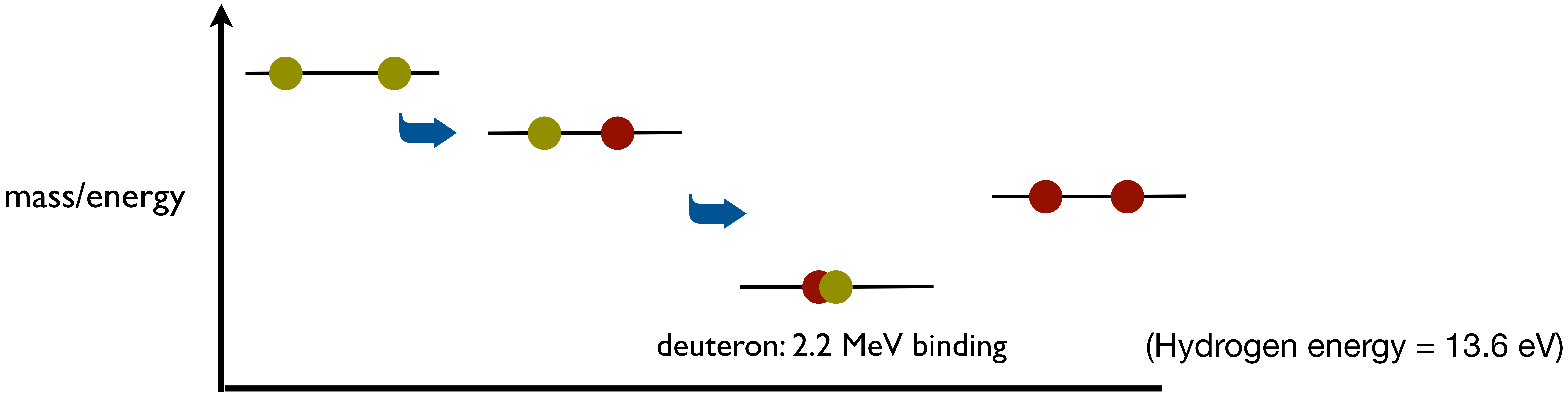


$\tau_n \sim 15 \text{ min}$

what prevented this from destroying all the neutrons?

if nothing else were to happen in the next few minutes,
our universe would be full of only Hydrogen

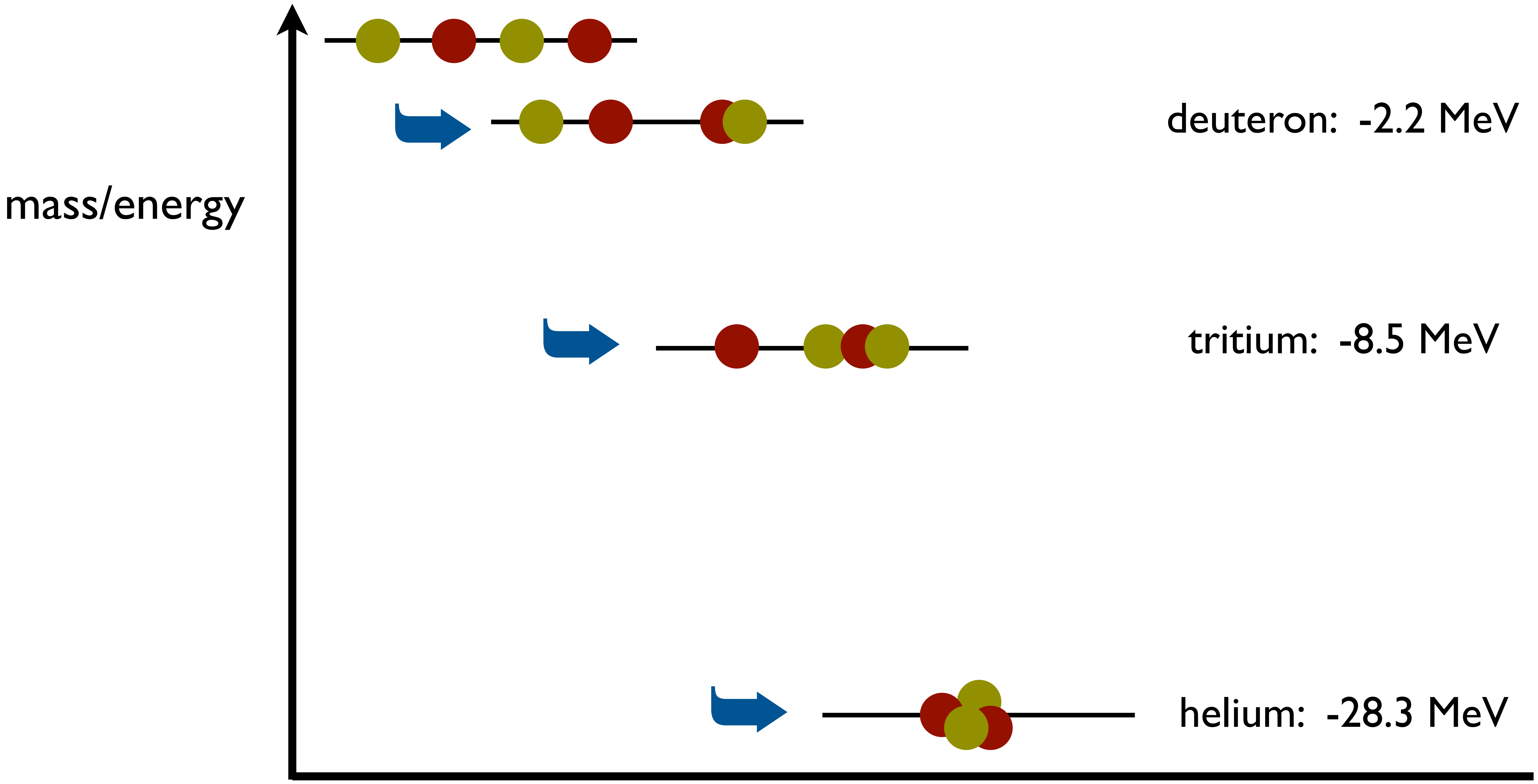
when systems cool, they settle into the lowest energy state



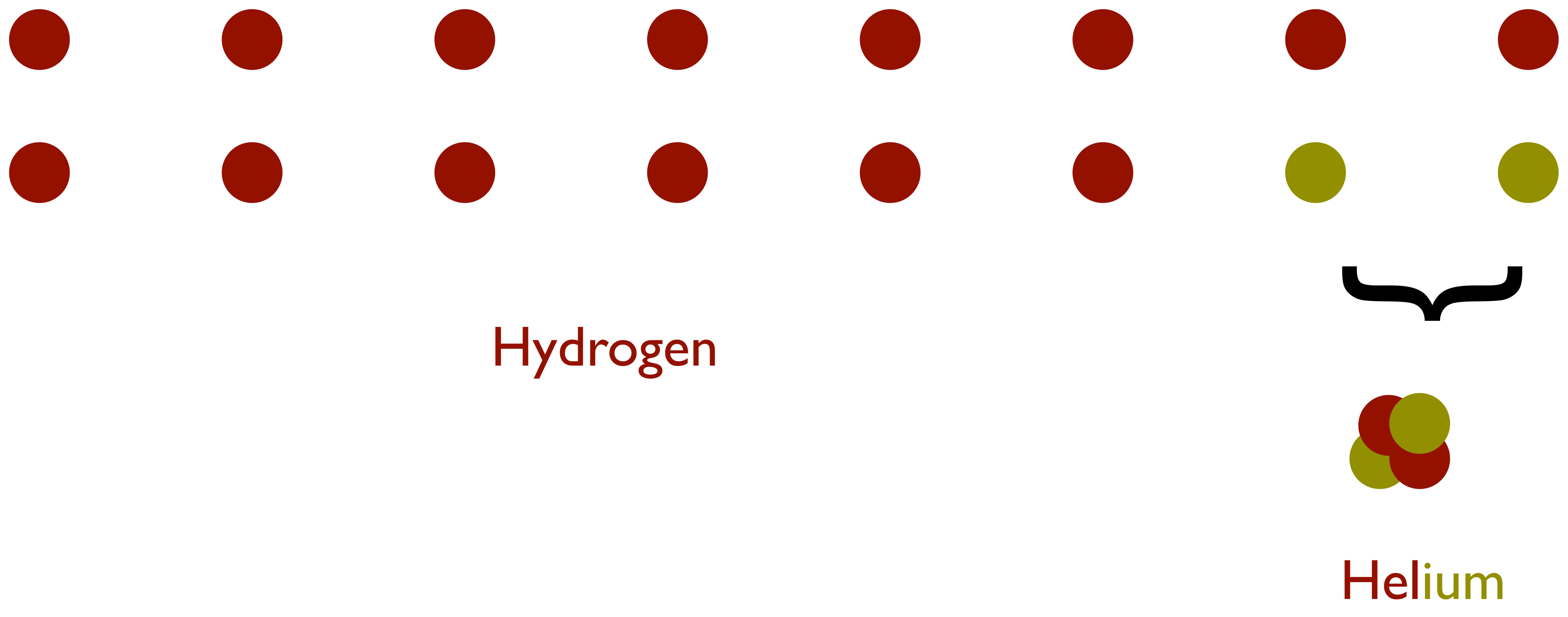
Answer: formation of nuclei

a system with protons and neutrons can collapse to a compact bound state, the deuteron: the attractive binding of a neutron and proton allows neutrons to survive when embedded in nuclei

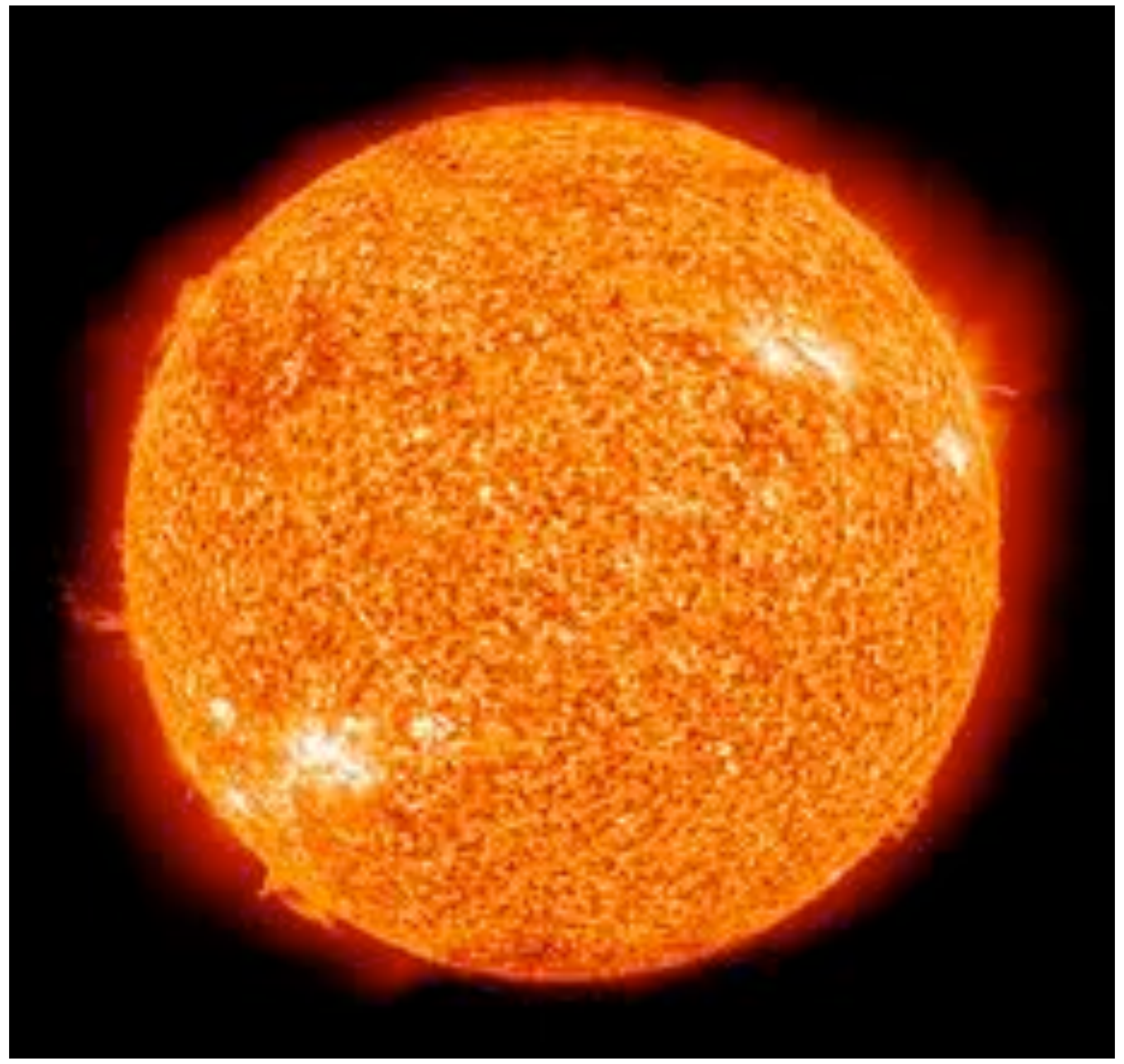
The deuterium "bottleneck" is broken, neutrons flow into He



He stability: ↑,↓ protons and ↑,↓ neutrons can be packed together



The early universe contains 75% H and 25% ^{4}He by mass fraction
("all" deuterium converted to ^{4}He)



The evolution of the universe is very sensitive to
- what is the lifetime of a neutron?
- what is the excess of matter over anti-matter?

this picture very sensitive to binding energy of deuterium which is finely tuned (most nuclei have ~8 MeV binding per nucleon)!

$$B_d = 2.22 \texttt{ MeV}$$

What if

$B_d \ll 2.22$ MeV
more finely tuned
all neutrons decay - no helium
mostly hydrogen stars?

$B_d \gg 2.22$ MeV
natural scenario
all neutrons captured in deuterium and helium - no hydrogen
no stars like ours!

Can we understand the emergence of B_d and how fine-tuned it is in terms of the fundamental theory?

this picture very sensitive to binding energy of deuterium which is finely tuned (most nuclei have ~8 MeV binding per nucleon)!

$$B_d = 2.22 \text{ MeV}$$

What if

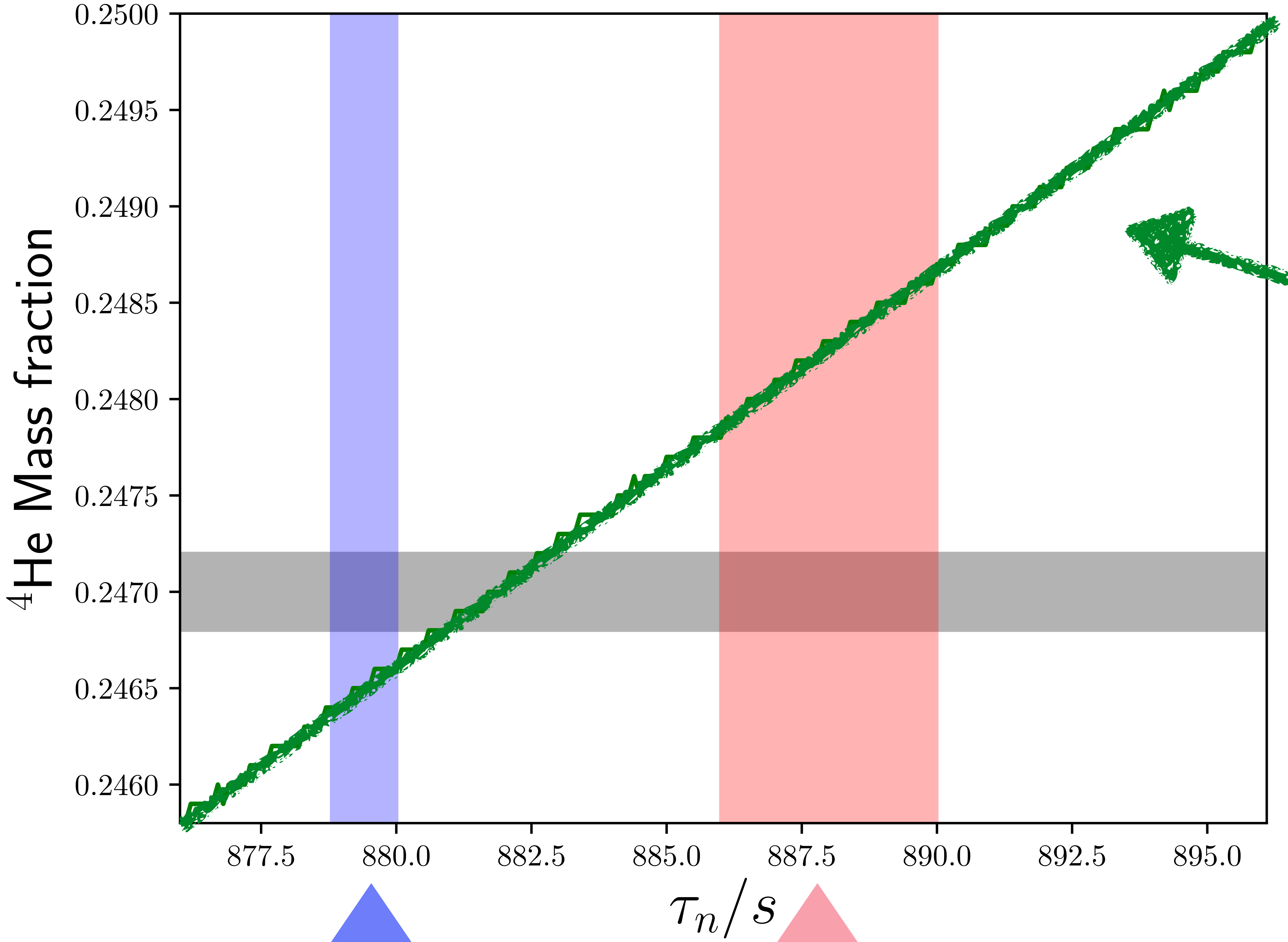
$B_d \ll 2.22 \text{ MeV}$ more finely tuned
all neutrons decay - no helium
mostly hydrogen stars?

$B_d \gg 2.22 \text{ MeV}$ natural scenario
all neutrons captured in deuterium and helium - no hydrogen
no stars like ours!

Can we understand the emergence of B_d and how fine-tuned it is in terms of the fundamental theory?

Yes!
With summit!

This picture is also very sensitive to the lifetime of the neutron

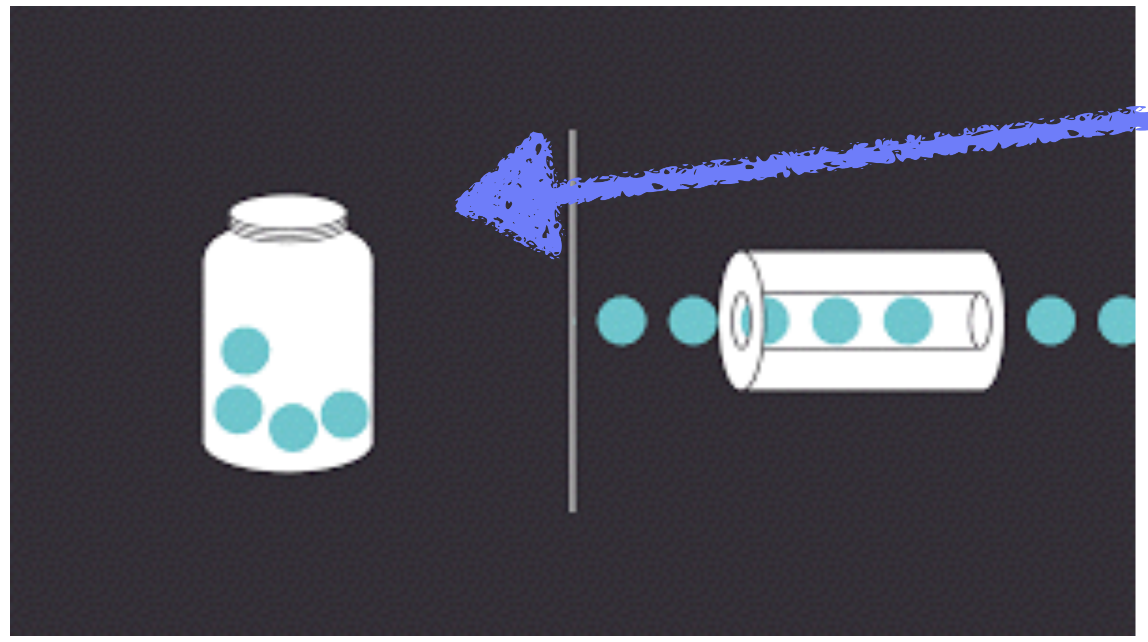


$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1 + 3g_A^2}$$

Czarnecki, Marciano, Sirlin

Predicted primordial ^{4}He mass fraction of the universe as a function of the neutron lifetime

Observed primordial ^{4}He mass fraction of the universe

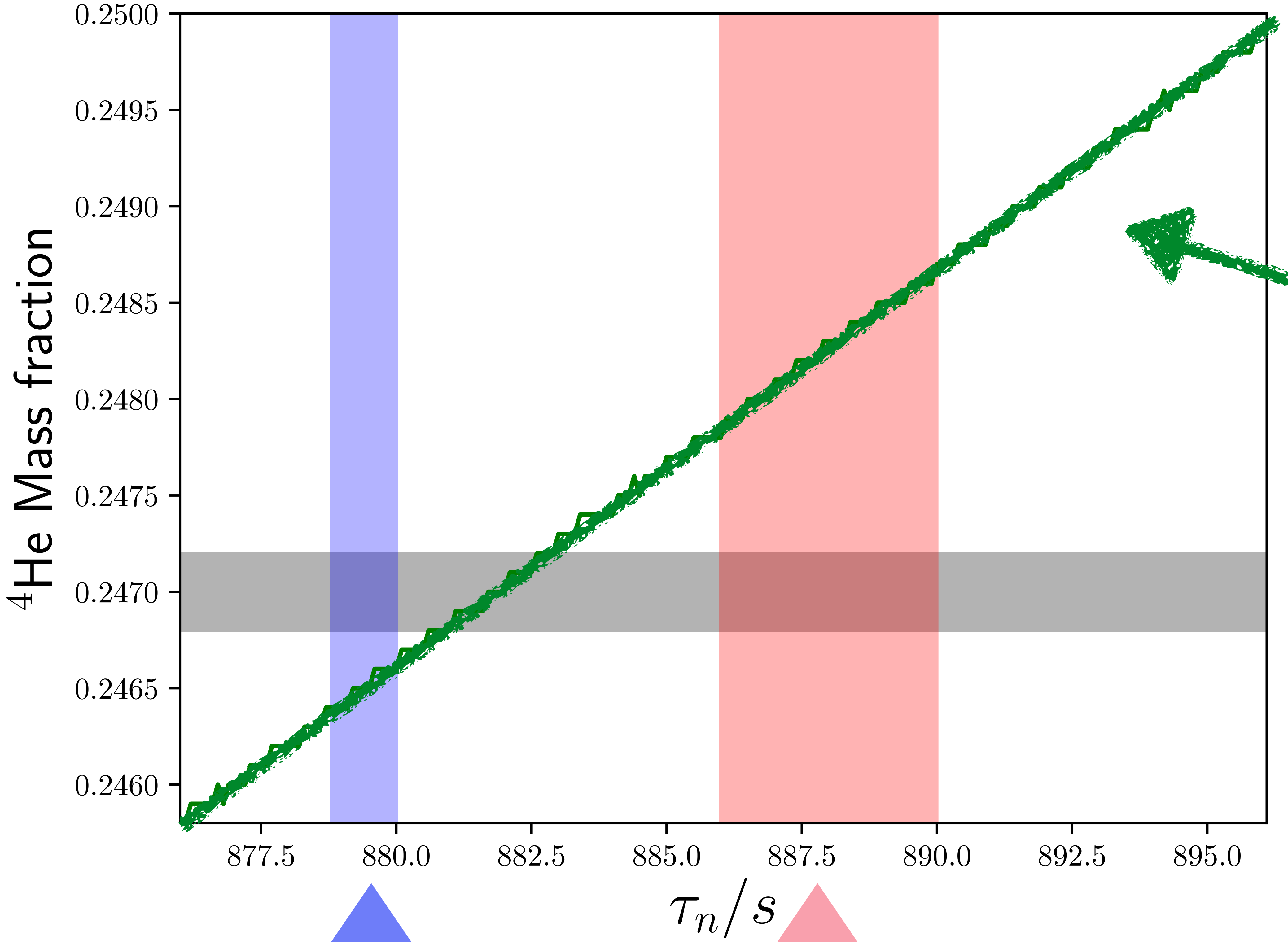


Two different methods of measuring the neutron lifetime disagree at the 99% level

bottle and beam

Is one of the experiments wrong? Or is there new physics hiding here?

This picture is also very sensitive to the lifetime of the neutron



$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1 + 3g_A^2}$$

Czarnecki, Marciano, Sirlin

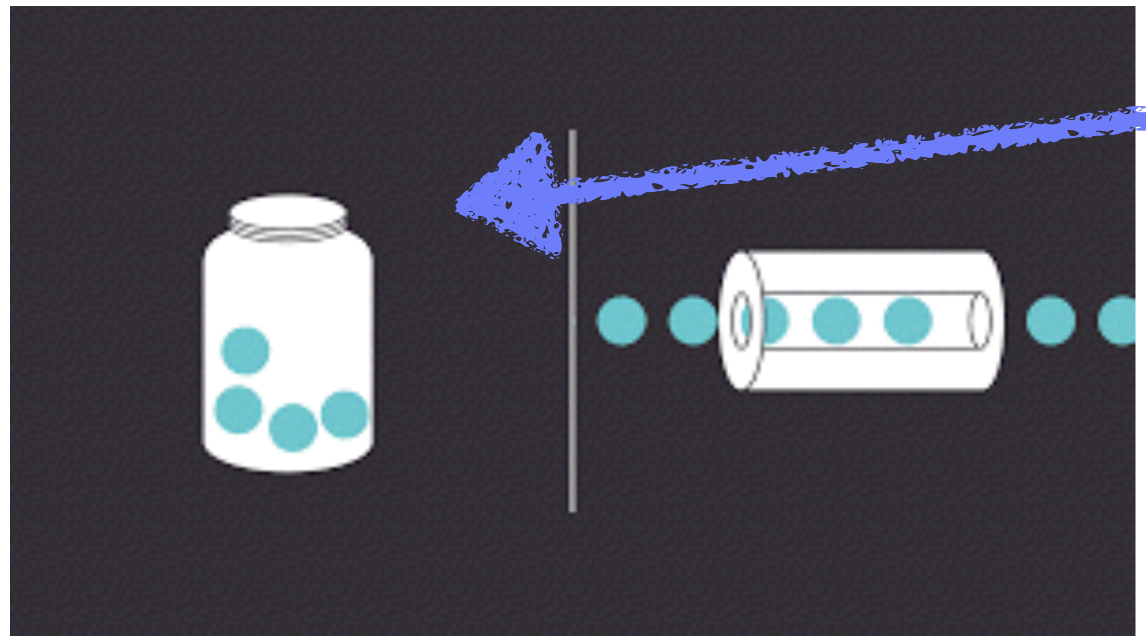
In order to predict this number, we need to use massive super-computers

Predicted primordial ^{4}H fraction of the universe of the neutron lifetime

Observed primordial ^{4}He mass fraction of the universe

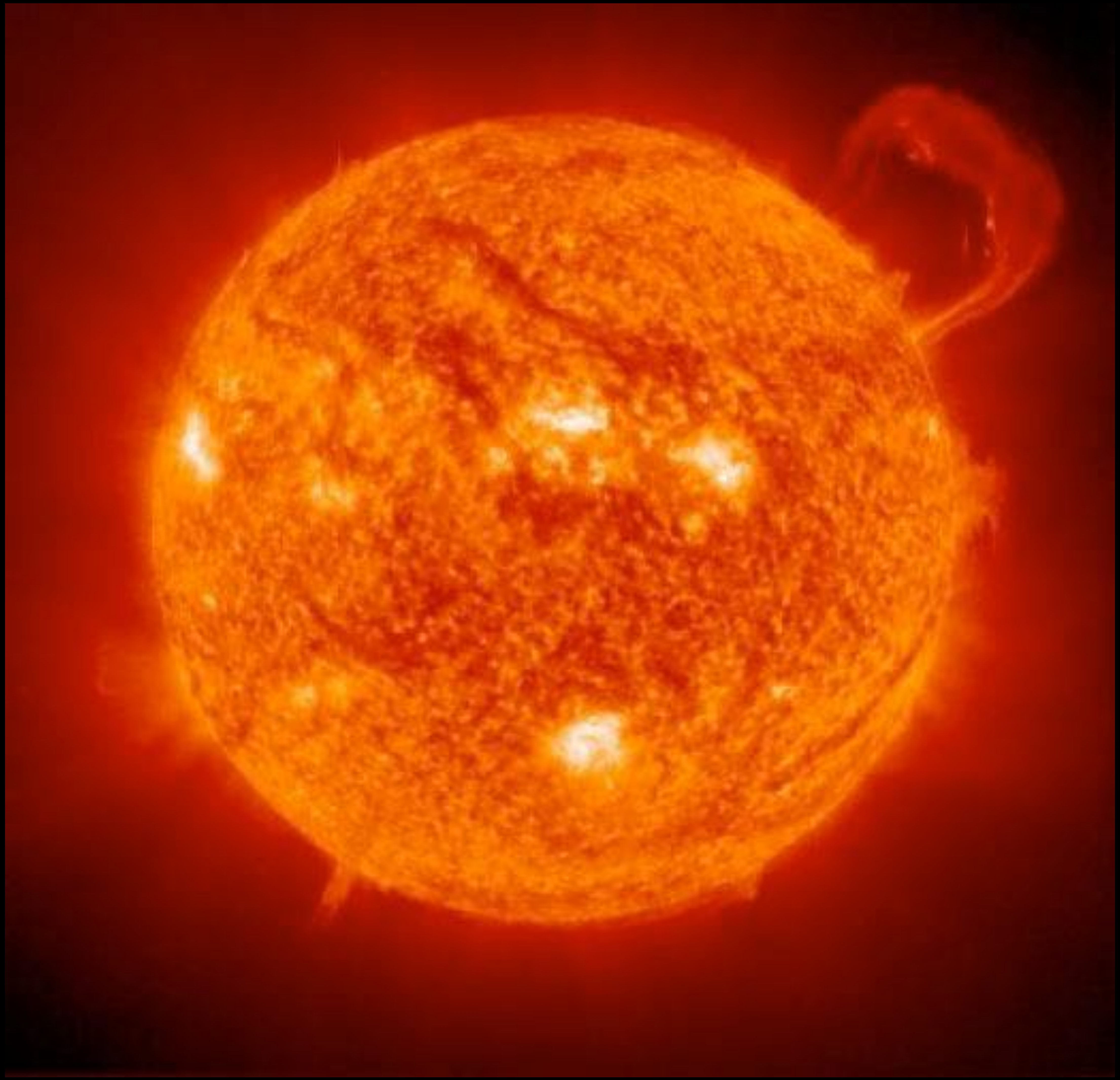
Two different methods of measuring the neutron lifetime disagree at the 99% level

bottle and beam



Is one of the experiments wrong? Or is there new physics hiding here?

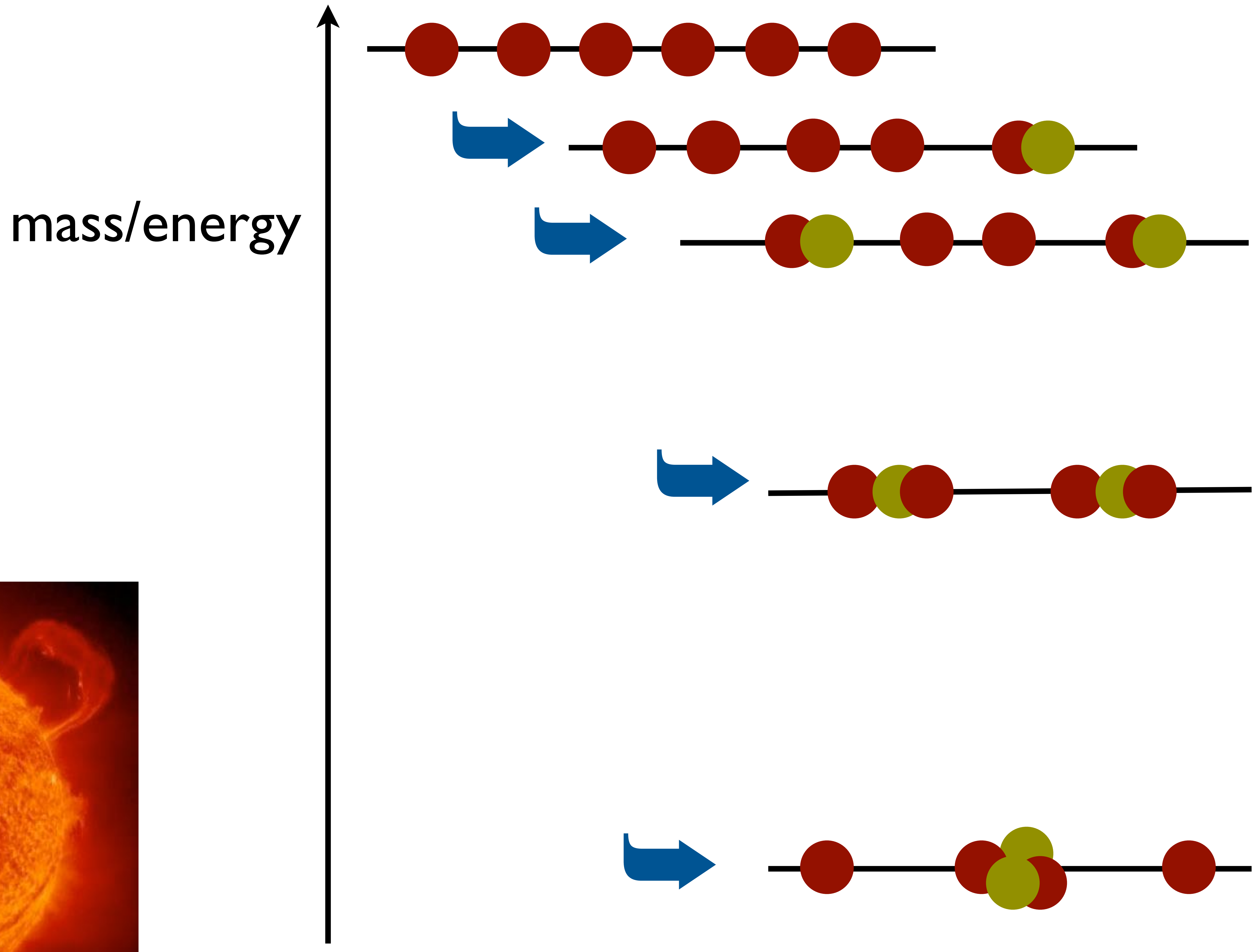
Solar Fusion



$$T \simeq 20 \text{ K}$$
$$t \simeq 200 \text{ Million years}$$

One needs neutrons and protons to make new nuclei.

Small stars burn protons only, manufacturing the needed neutrons



$$p + p \rightarrow d + \nu_e$$

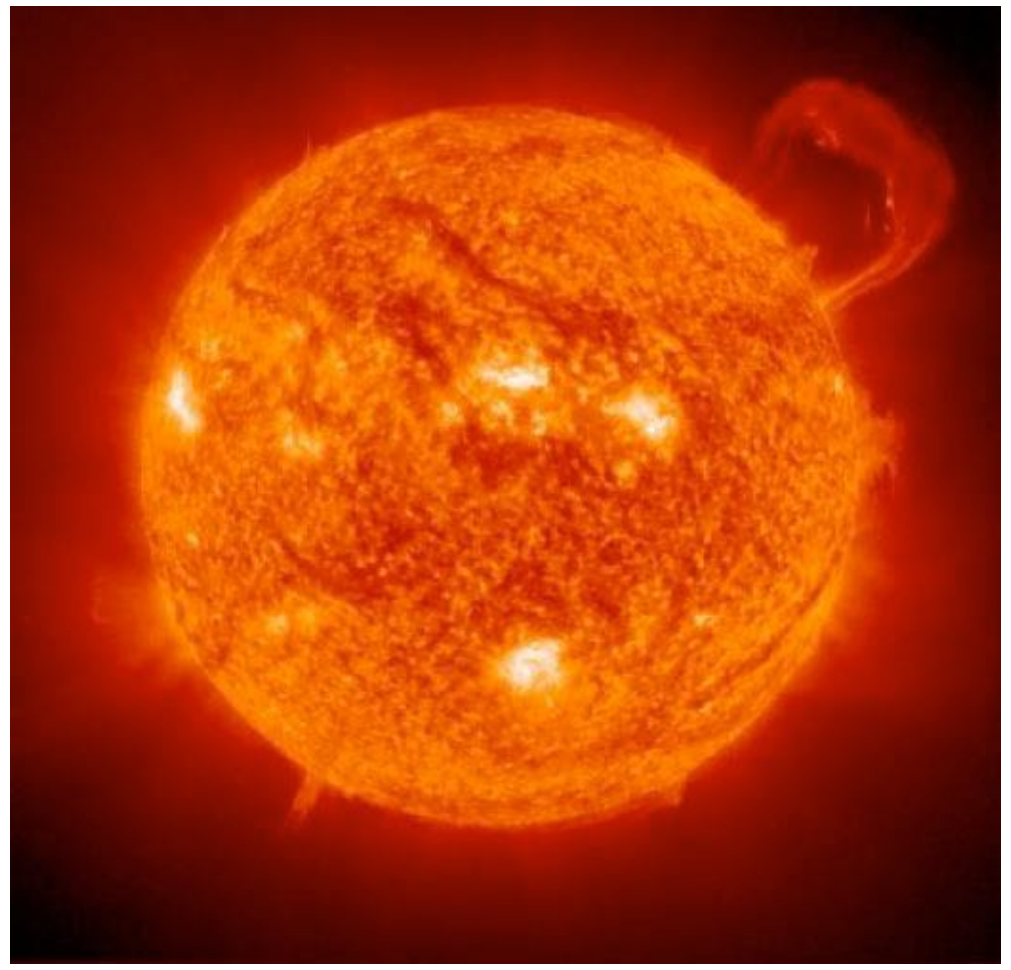
another bottleneck

deuteron: -1.2 MeV

2 deuterons: -2.4 MeV

2 ^{3}He: -12.4 MeV

helium: -28.3 MeV



pp chain

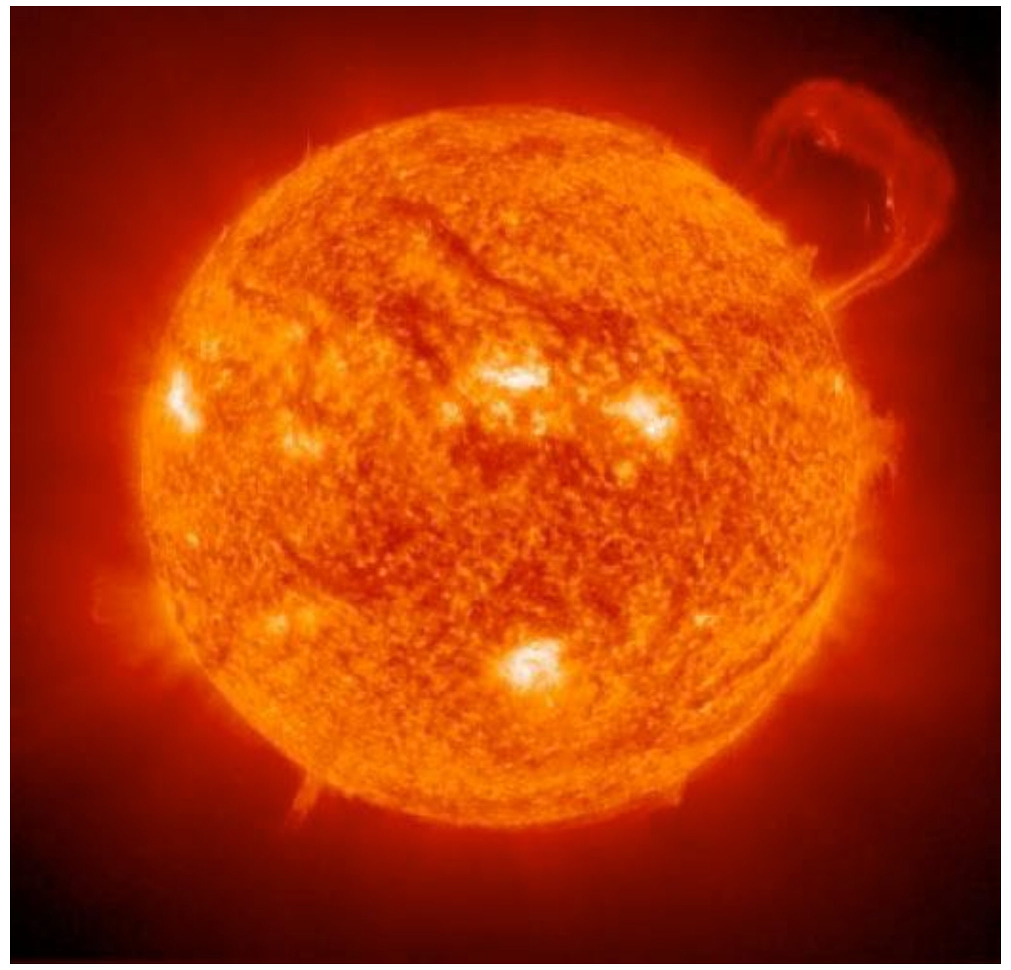
thermal energy to support star

This is how our Sun generates its energy

80% of all stars generate their energy by hydrogen burning

At its very center the Sun generates 275 watts/m^3 - similar to the energy generated by a compost (garbage) heap (of the same size)!

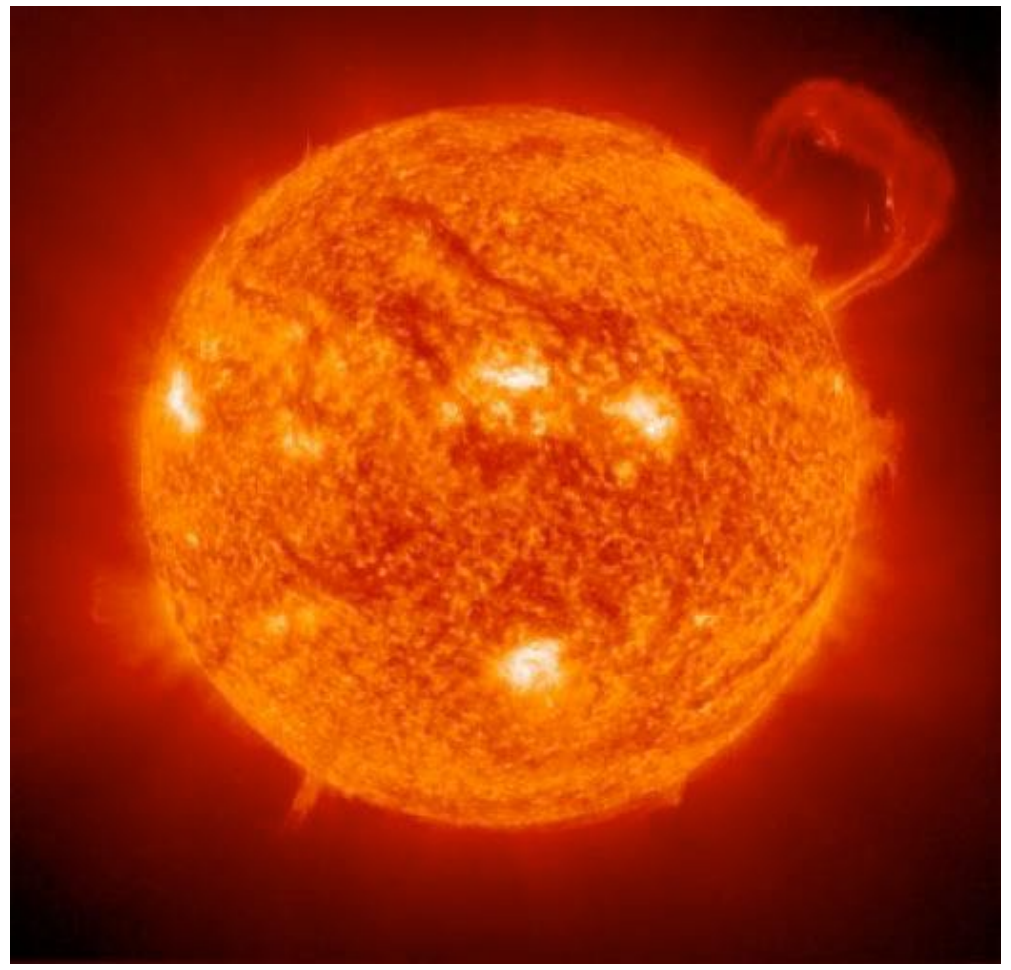
And this is why the Sun has burned for 4.6 b.y., and will burn for 5 b.y. more, fortunately -- a very big, very slow reactor



$$p + p \rightarrow d + \nu_e + e^+$$

This fundamental reaction can not be measured!
(Coulomb Repulsion)

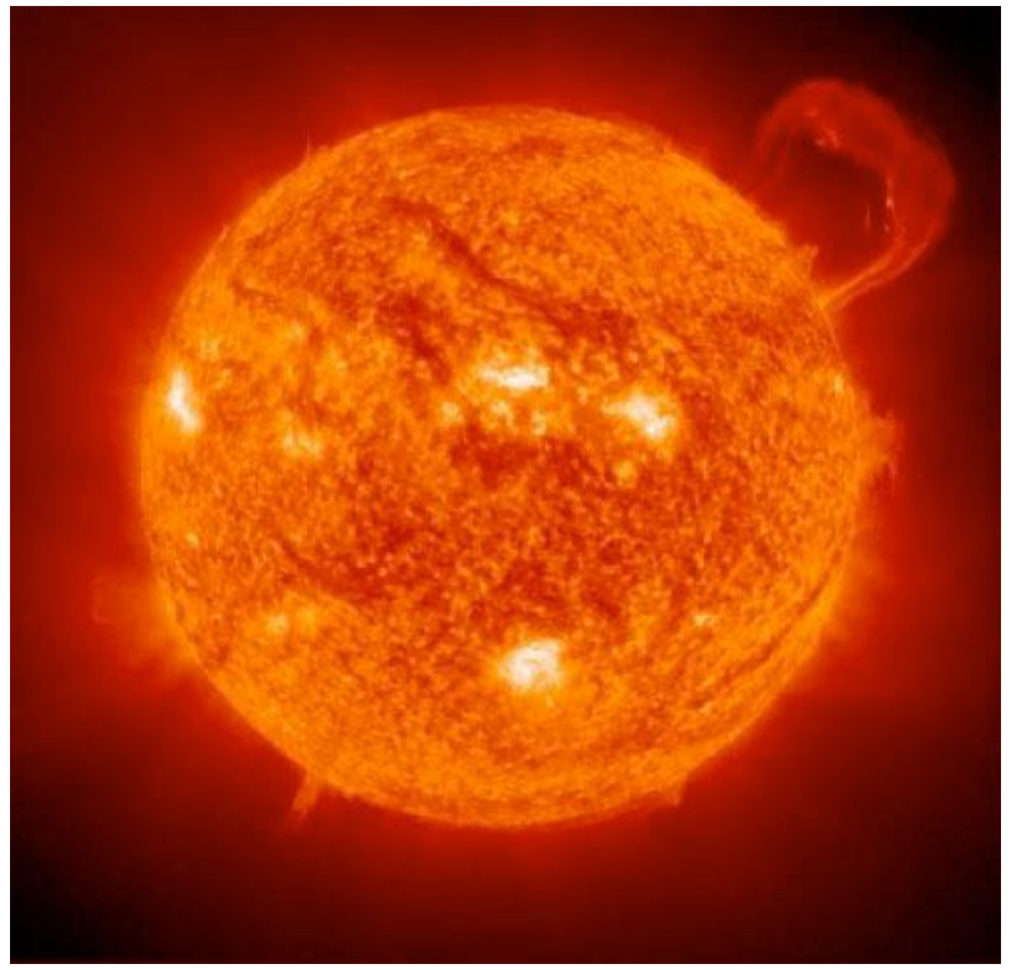
We believe we know the rate, but we have not been able to predict it directly from the fundamental theory, can we?



$$p + p \rightarrow d + \nu_e + e^+$$

This fundamental reaction can not be measured!
(Coulomb Repulsion)

We believe we know the rate, but we have not been able to predict it directly from the fundamental theory, can we?



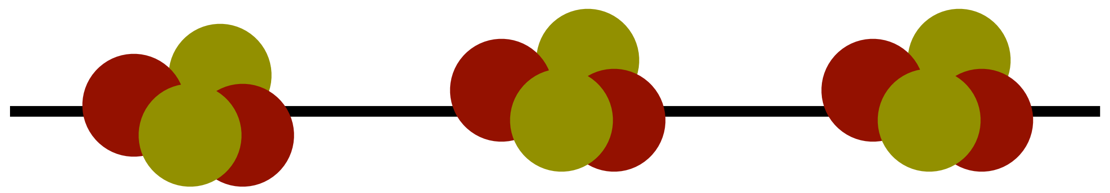
Yes!
With summit!

Large stars use He and neutrons to build new nuclei.

Higher temperatures and higher densities are needed.

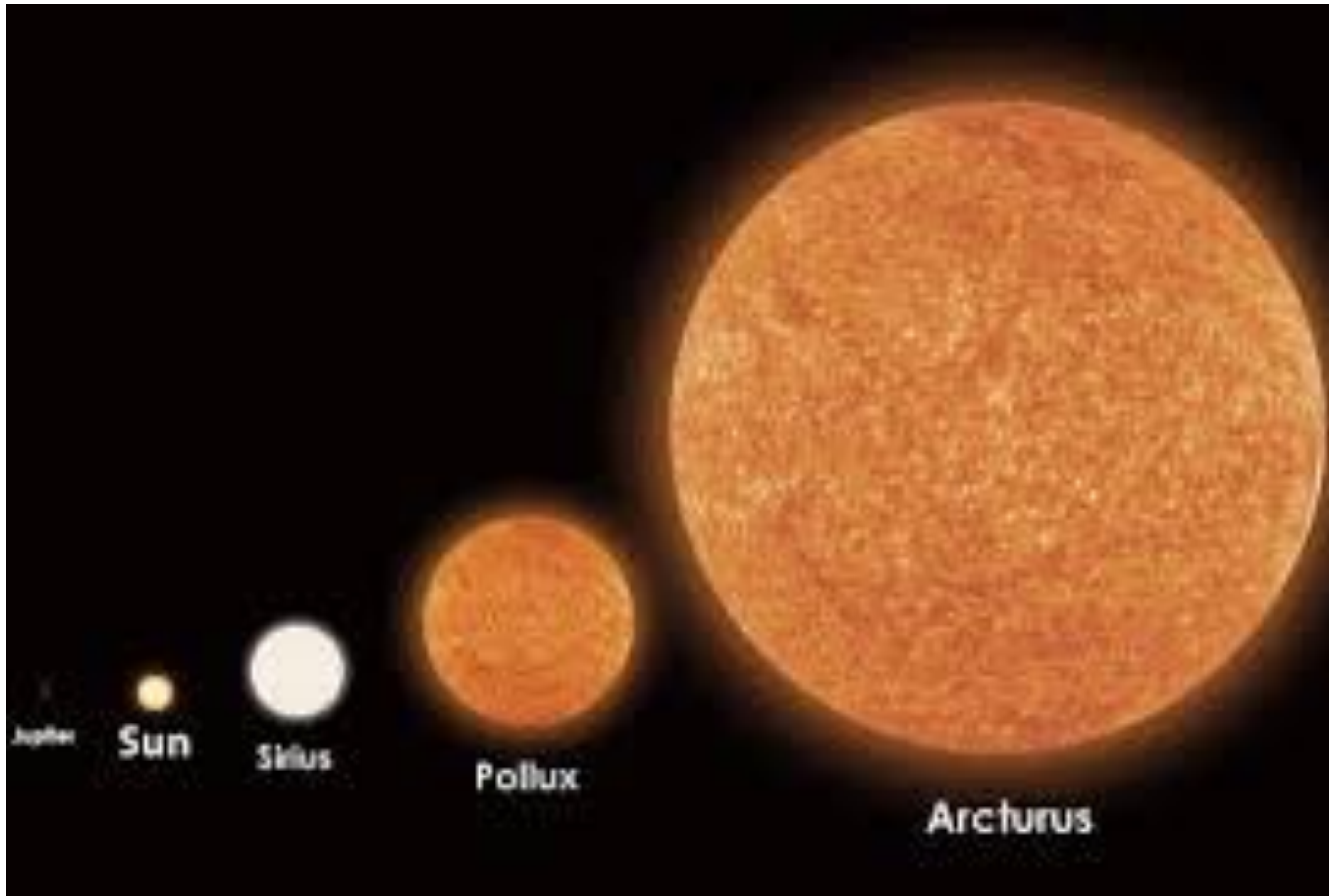
The Big Bang could not do this because the density was too low.

mass/energy

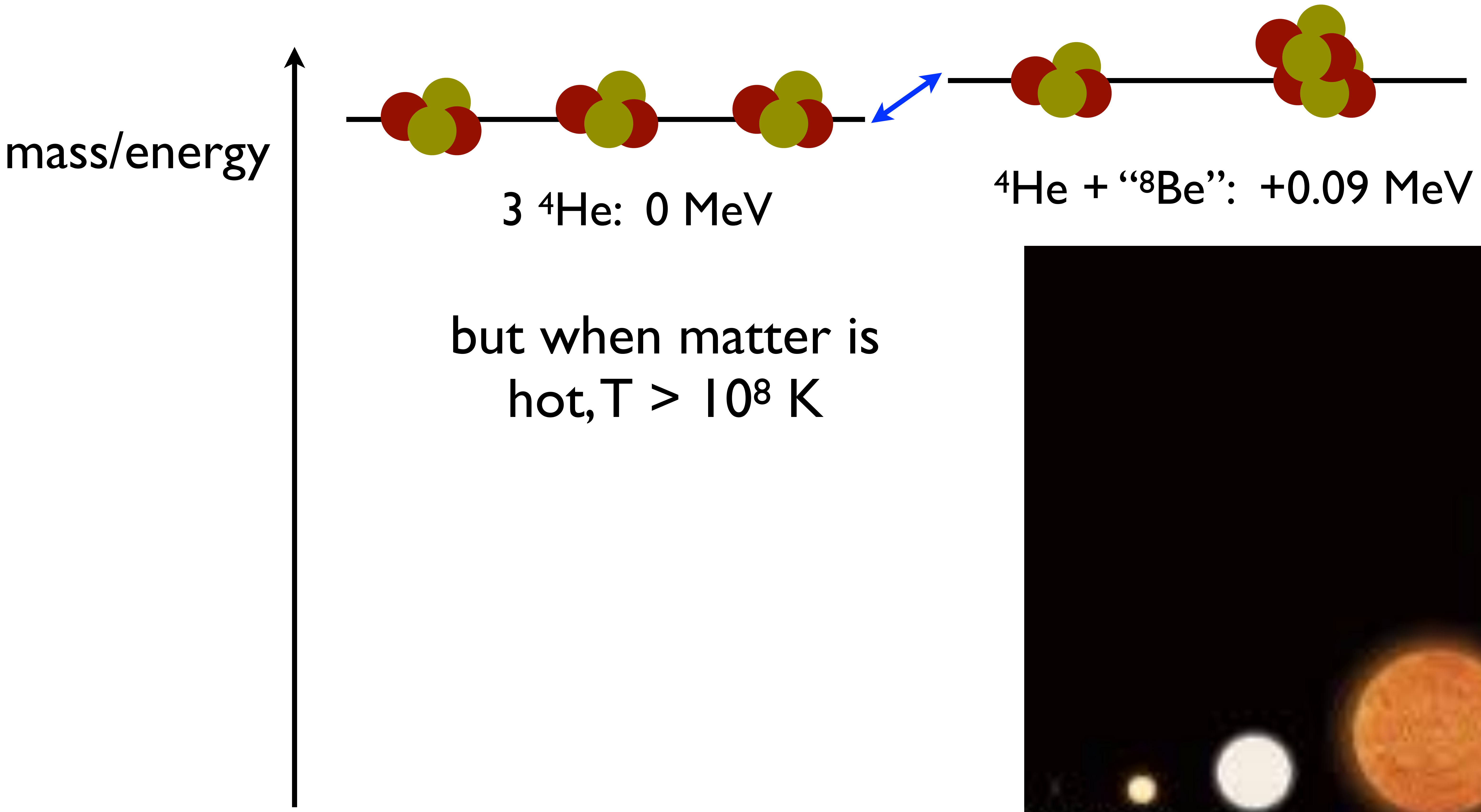


3 ^{4}He: 0 MeV

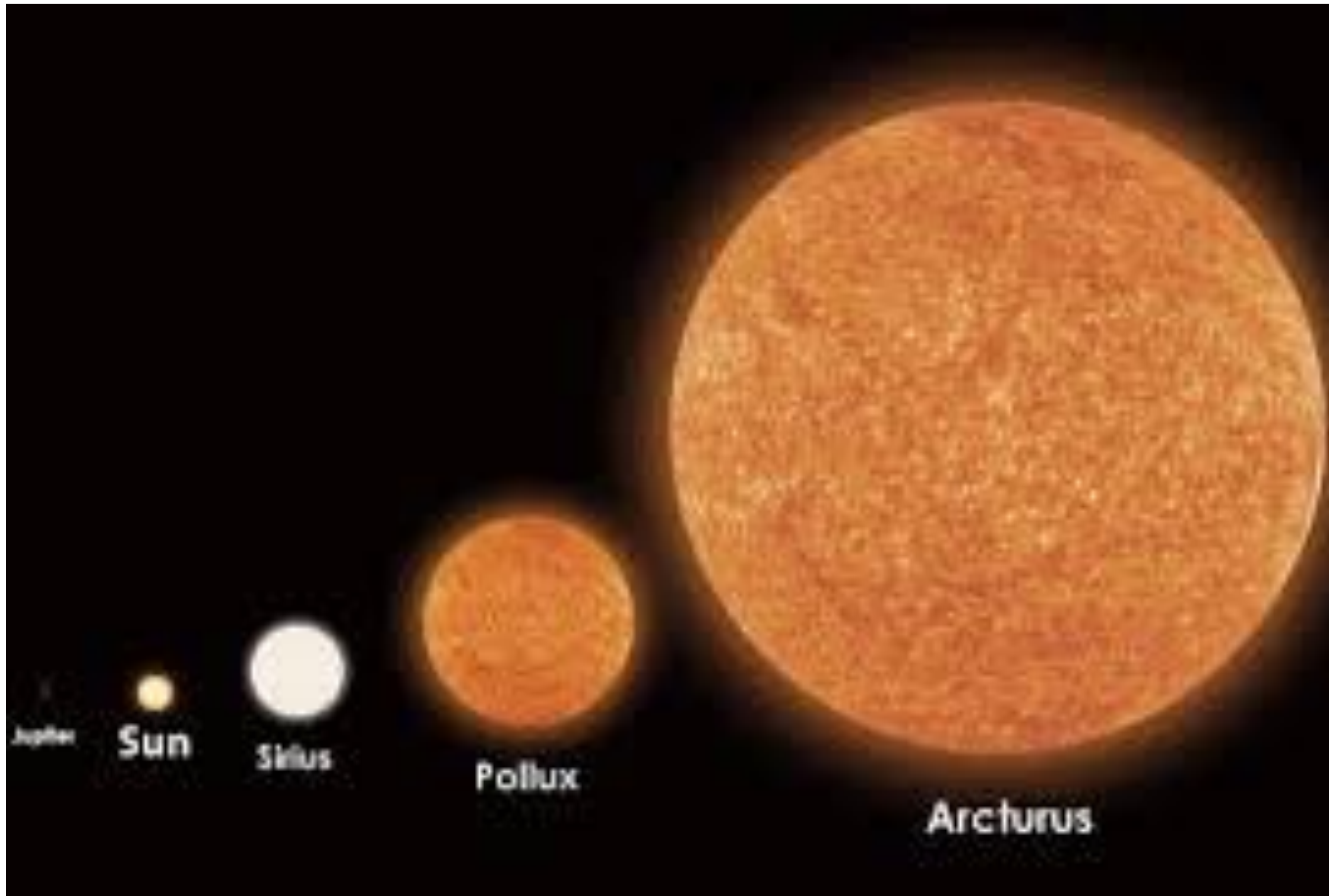
when matter is cold



even more finely tuned

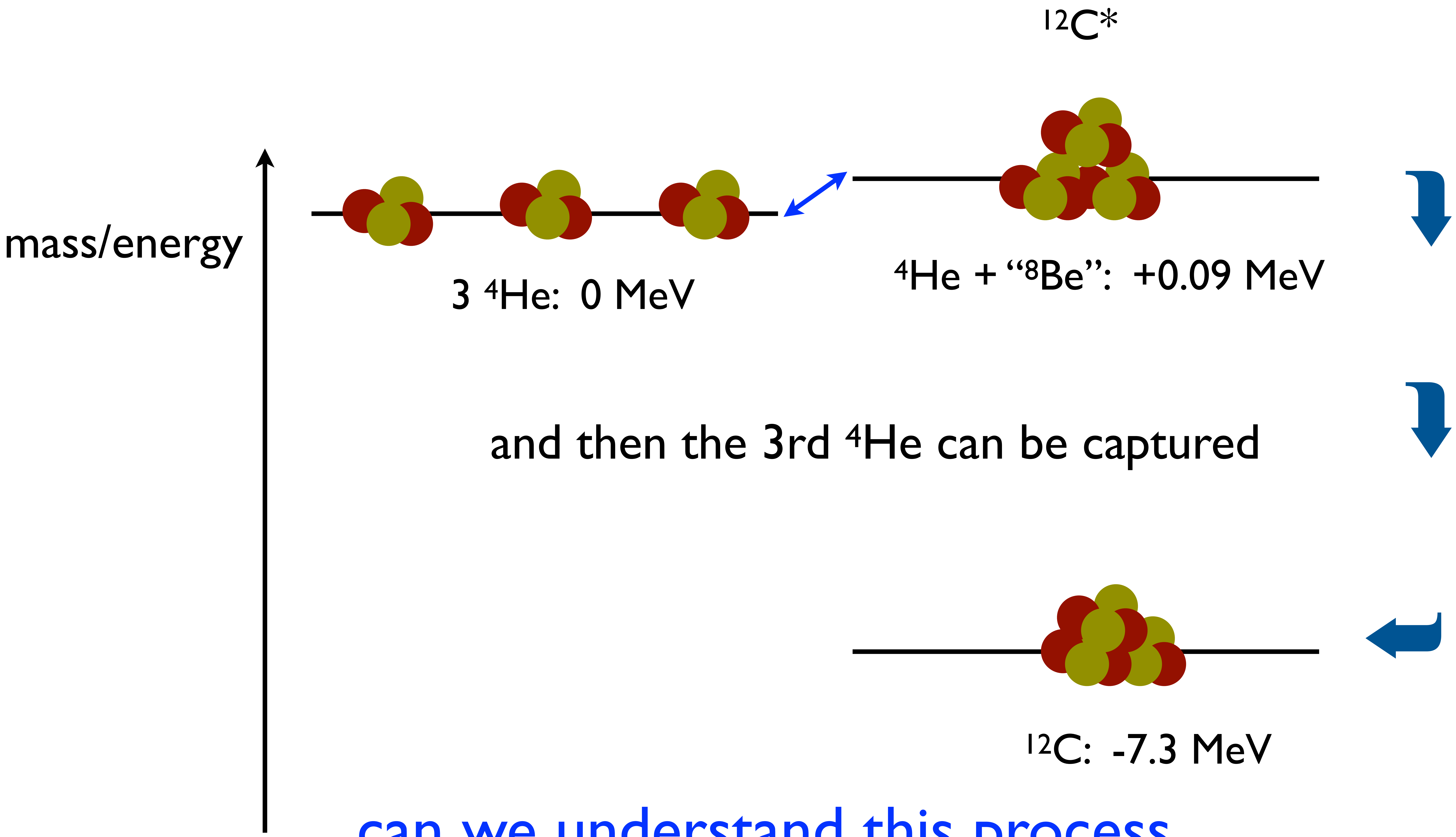


but when matter is hot, T > 10^8 K



even more finely tuned - source of complex life

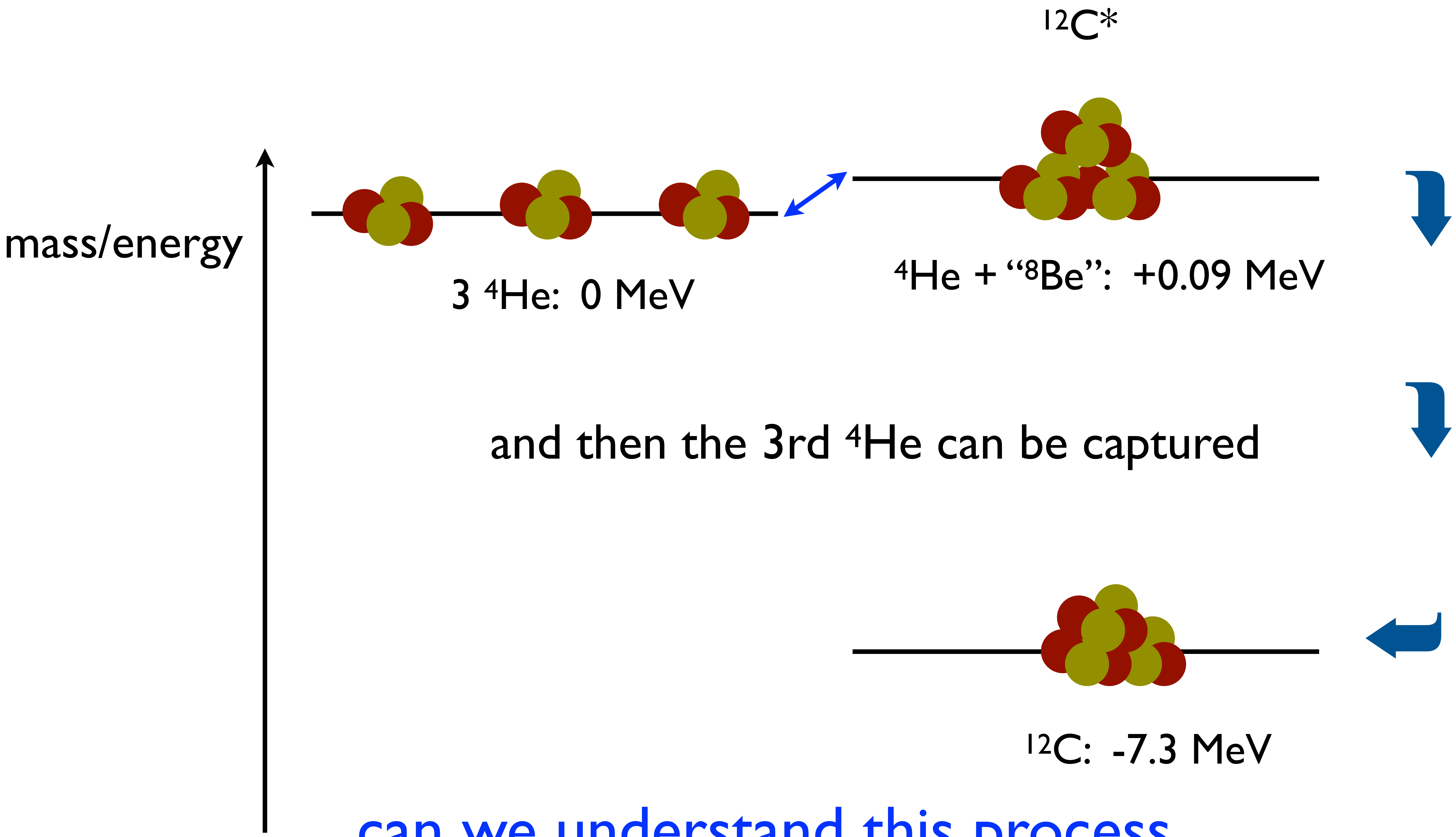
the triple-α process Hoyle State



can we understand this process
from the fundamental theory?

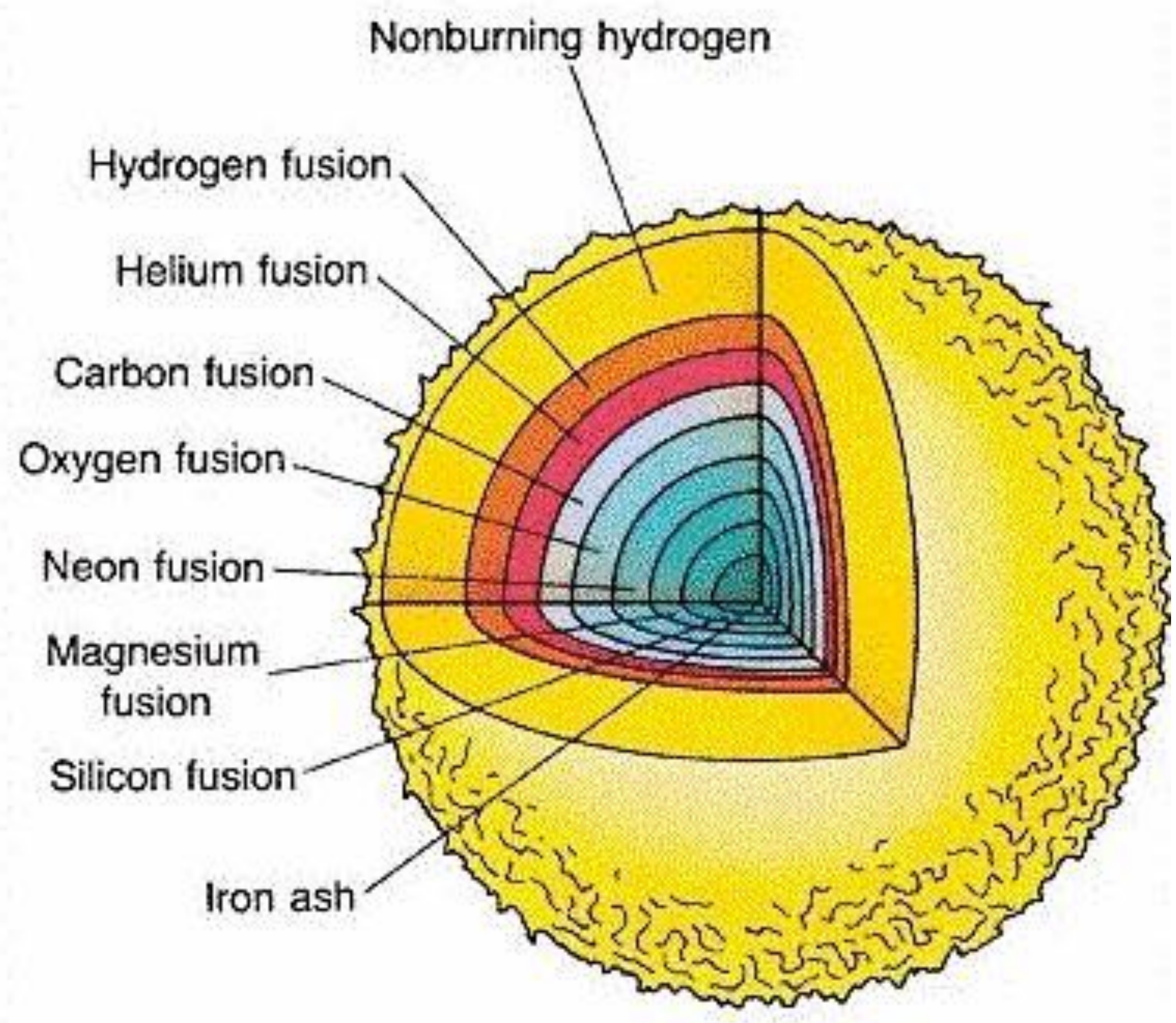
even more finely tuned - source of complex life

the triple-α process Hoyle State

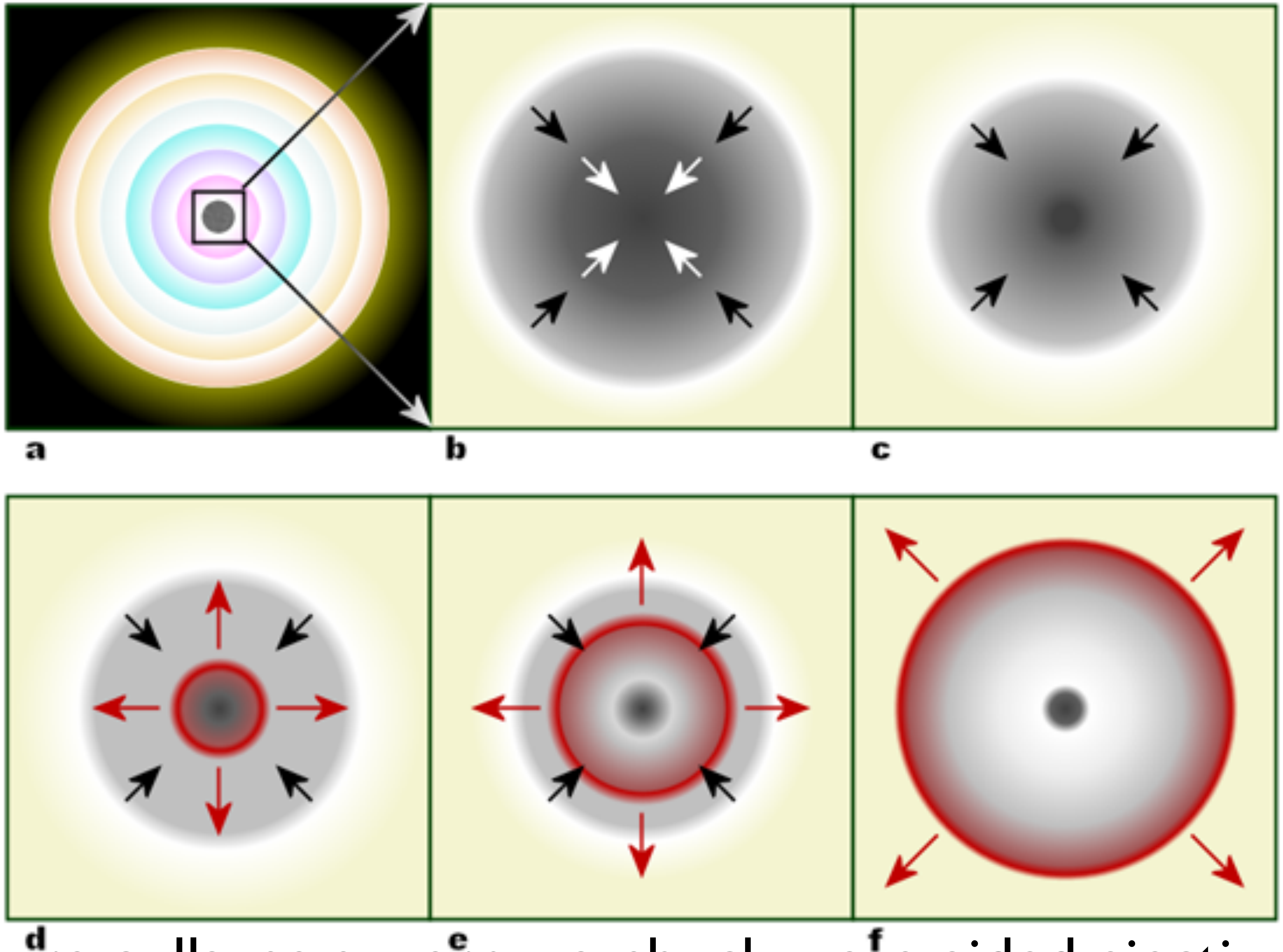


can we understand this process
from the fundamental theory?

Yes!
maybe with summit!



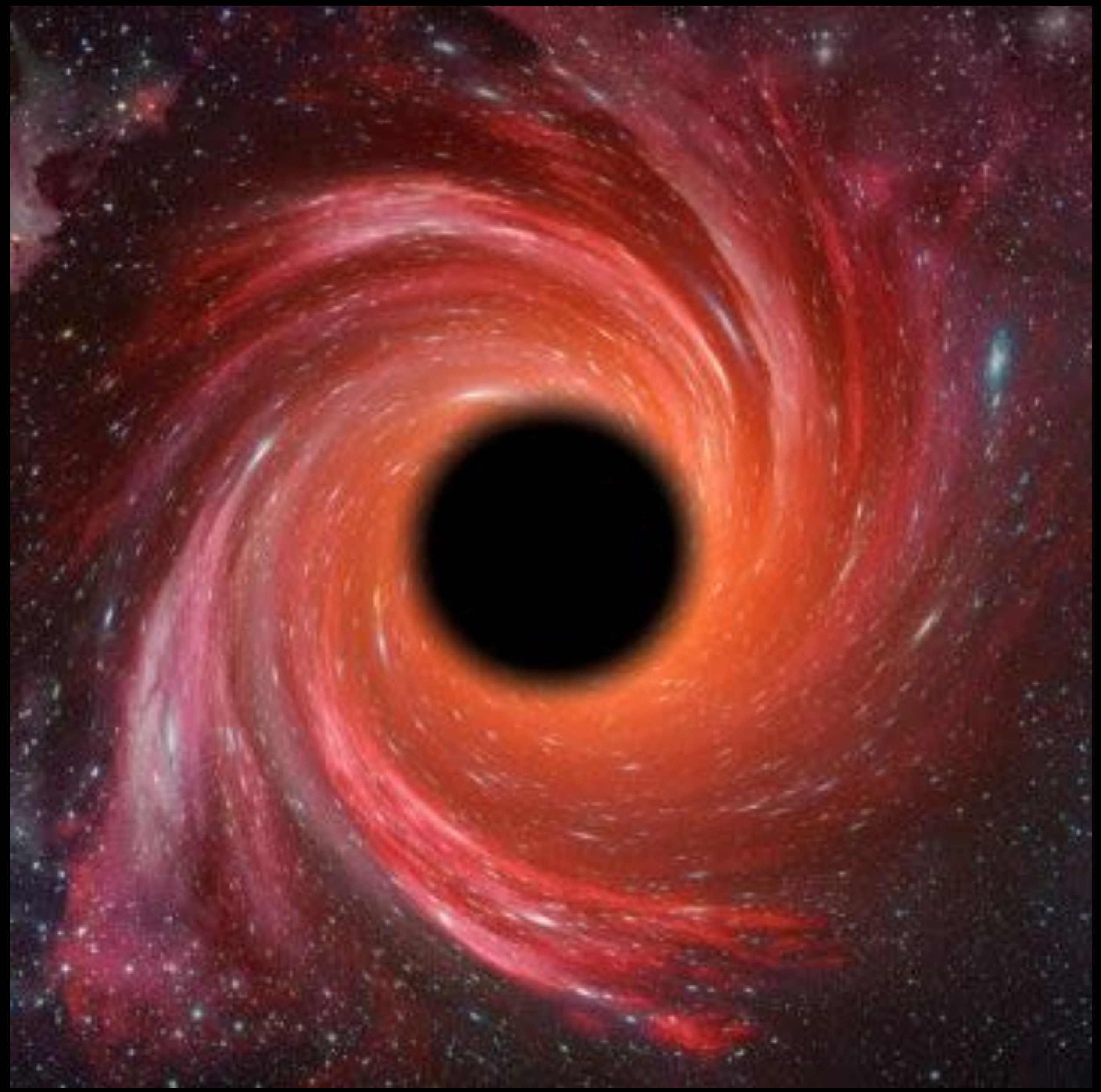
He, C, O, ... Si burning produces energy until Iron (Fe)



core collapse supernova, shock-wave-aided ejection of mantel

Supernova

neutron star



black hole

"We are all made of star stuff"

much coming from the ejecta of supernova

Can we understand properties of neutron stars directly from the fundamental theory?

Can we understand properties of neutron stars directly from the fundamental theory?

Yes! With FRONTIER

Can we understand properties of neutron stars directly from the fundamental theory?

Yes! With FRONTIER

Will it have GPUs?
I hope so!

Need for HPC

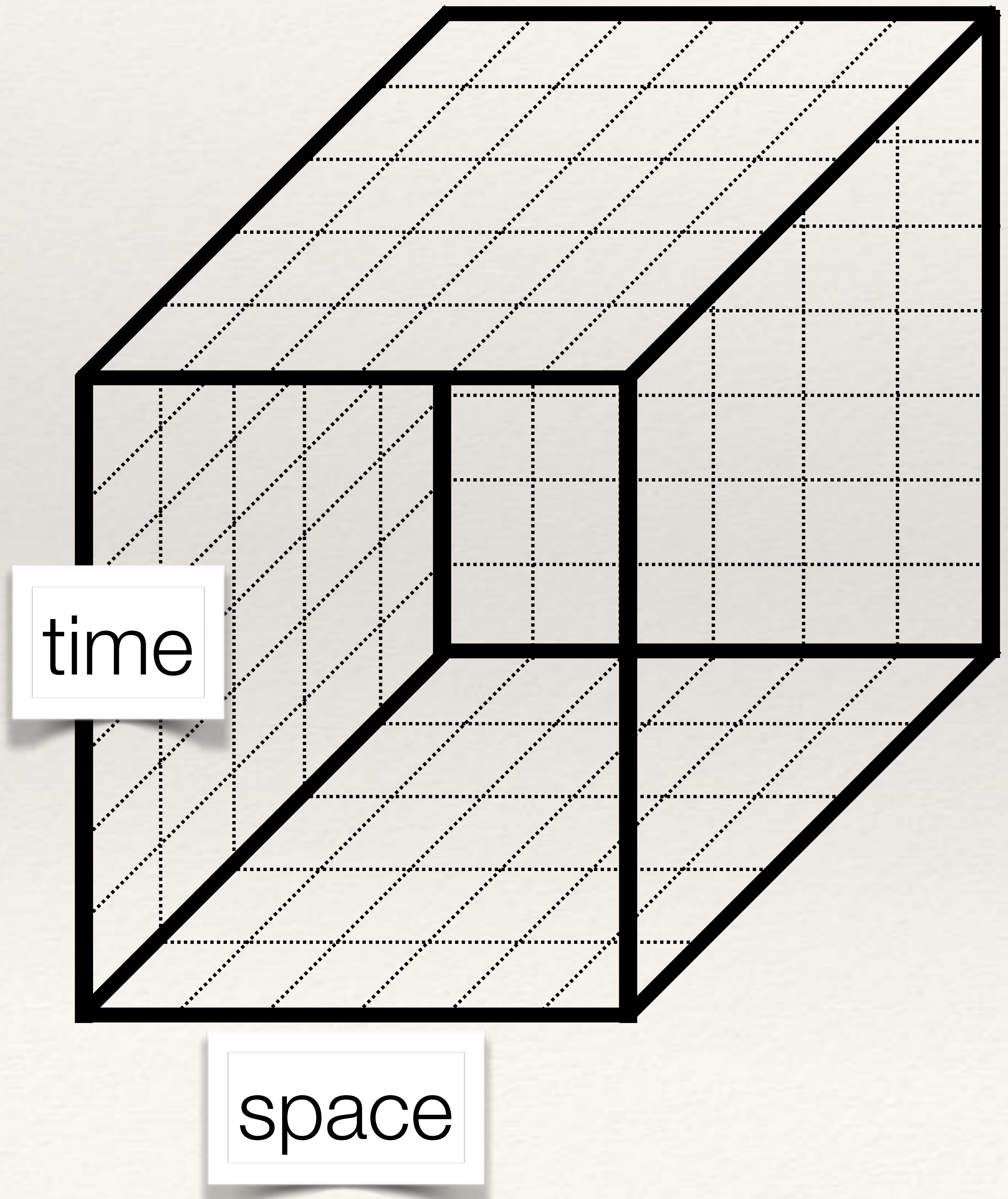
- The inherent challenge in understanding properties of matter (nuclear physics) directly from the fundamental theory, the Standard Model of Particle Physics - is that the nuclear interactions are very strong - so perturbative methods fail
  - Analogy: $1/(1-x) = 1 + x + x^2 + x^3 + \ldots$
    If $|x|<1$, then each successive contribution is smaller.
    If $|x|>1$, then contributions grow.

- The fundamental theory of nuclear strong interactions is Quantum Chromodynamics
  - a relativistic, quantum mechanical theory of quarks and gluons
  - inherently a 4-dimensional theory

Need for HPC

- ultimately, we have to perform this integral of integrals

$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

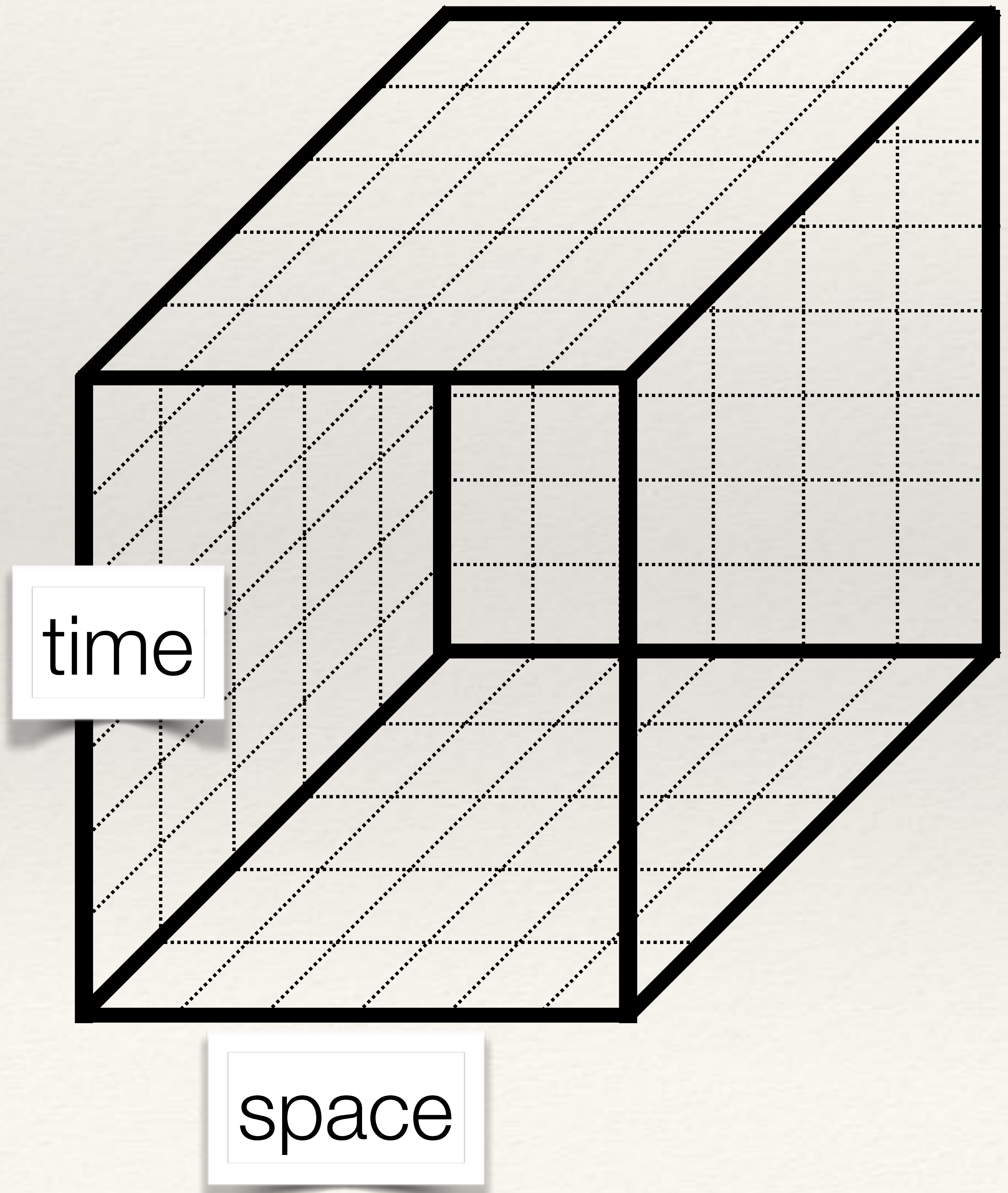


Need for HPC

- ultimately, we have to perform this integral of integrals

$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability



Need for HPC

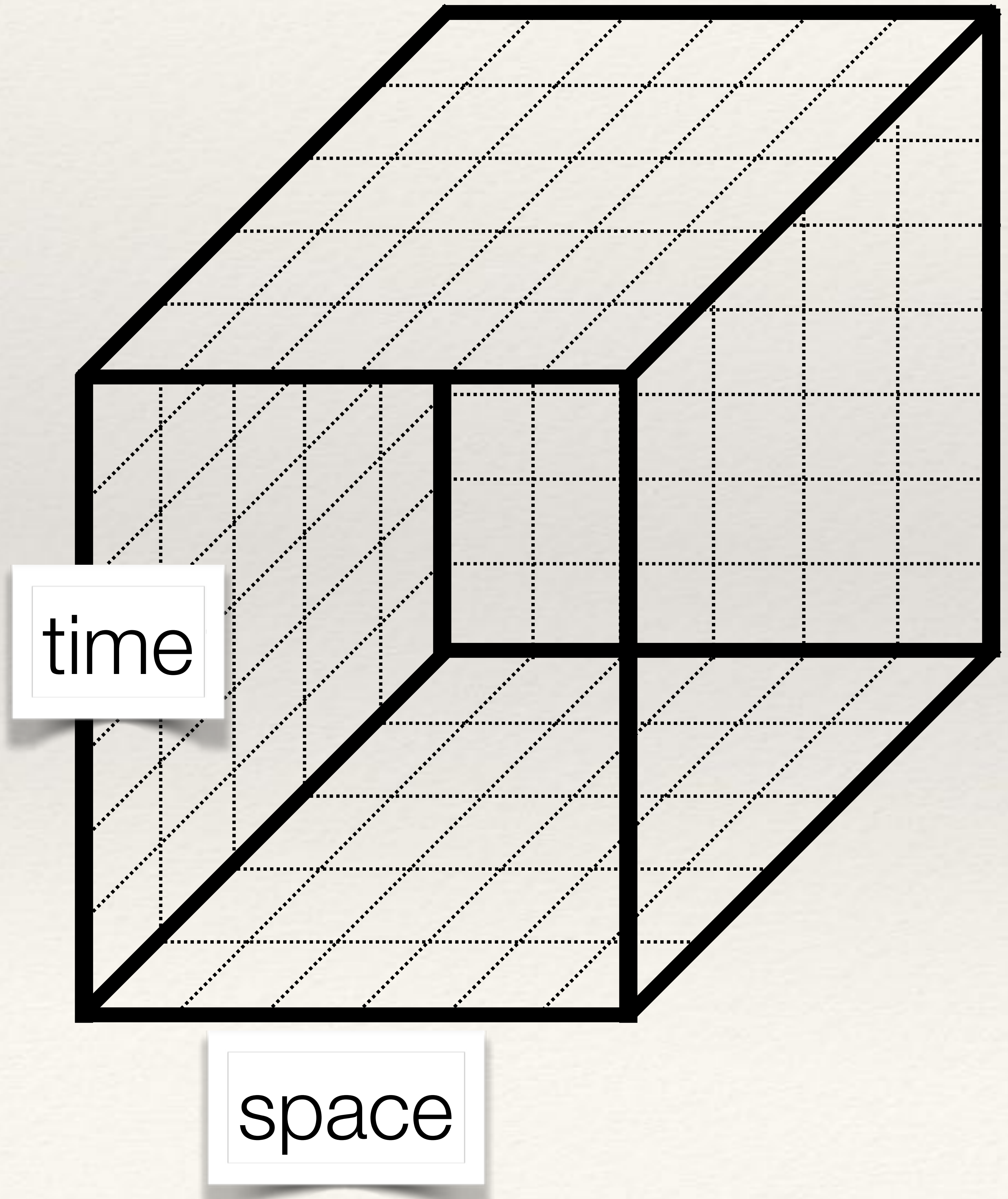
- ultimately, we have to perform this integral of integrals

$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability

$$\{U_1, U_2, U_3, \dots, U_N\}$$

Markov Chain Monte Carlo

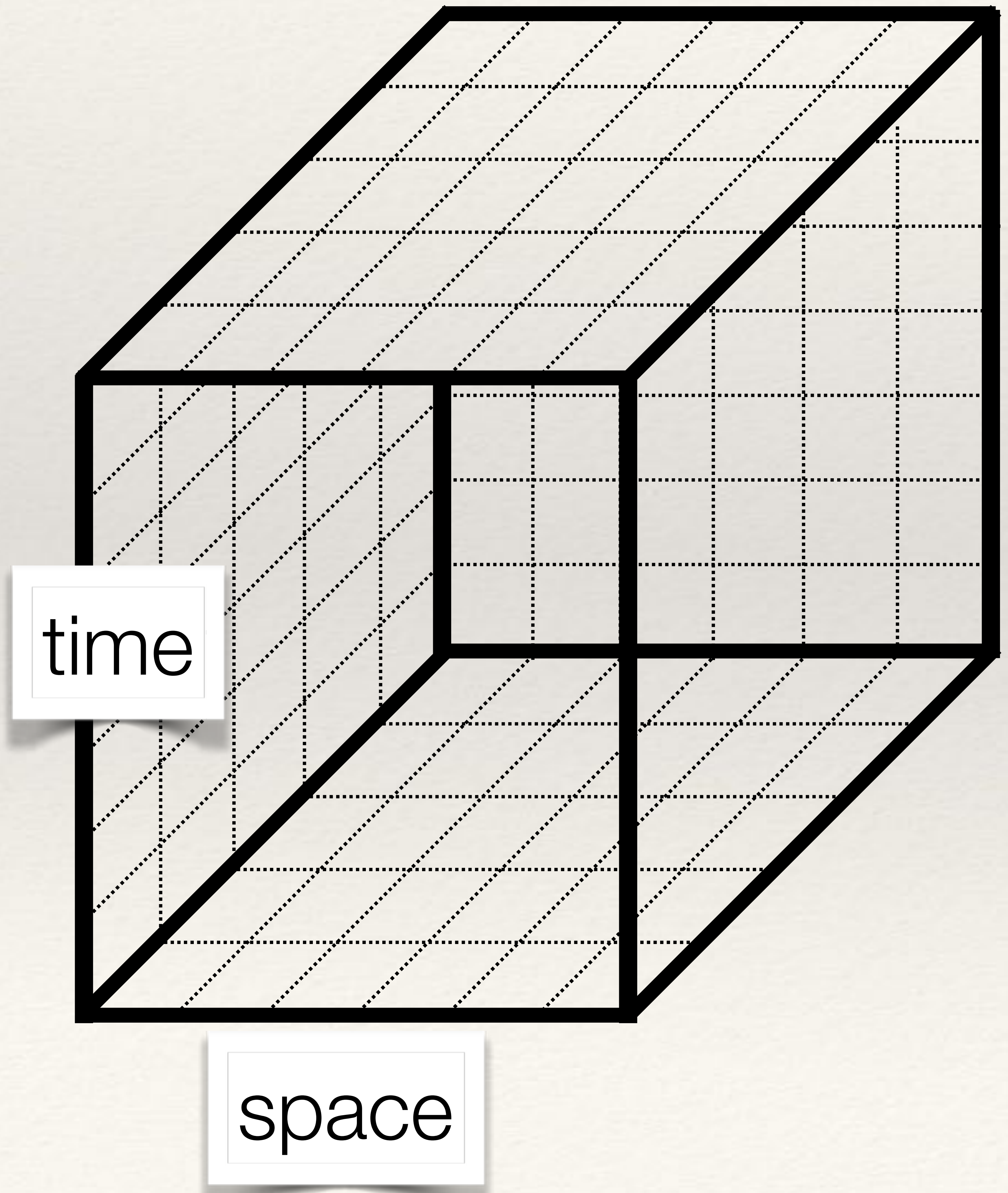


Need for HPC

- ultimately, we have to perform this integral of integrals

$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability



$$\{U_1, U_2, U_3, \ldots, U_N\}$$

Markov Chain Monte Carlo

$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(t)\mathcal{O}^\dagger(0)[U_i]$$

Need for HPC

- ultimately, we have to perform this integral of integrals

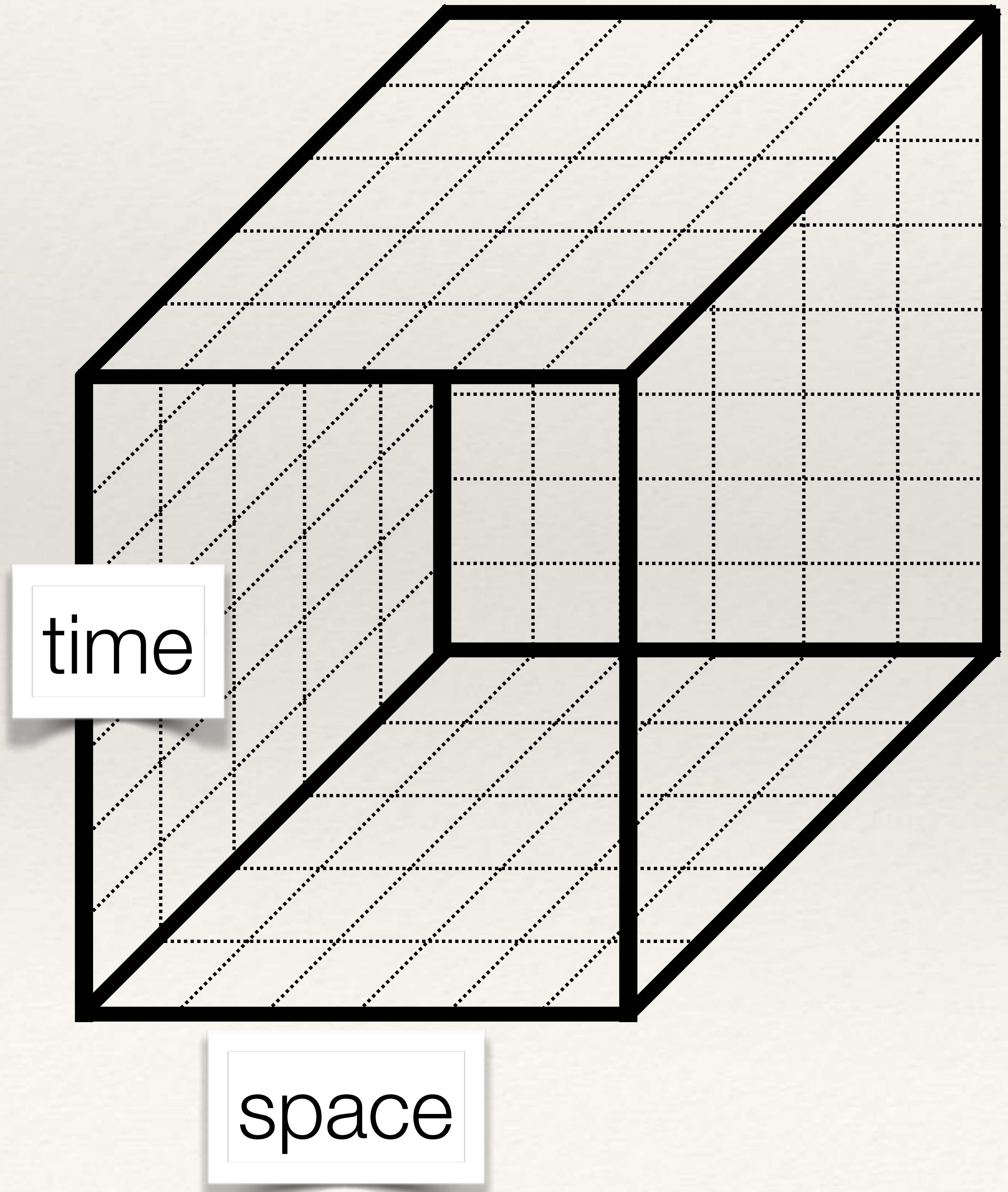
$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t)O[U]^\dagger(0)$$

Probability

$$\{U_1, U_2, U_3, \ldots, U_N\}$$

Markov Chain Monte Carlo

$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(t)\mathcal{O}^\dagger(0)[U_i] + O\left(\frac{1}{\sqrt{N}}\right)$$



Need for HPC

- ultimately, we have to perform this integral of integrals

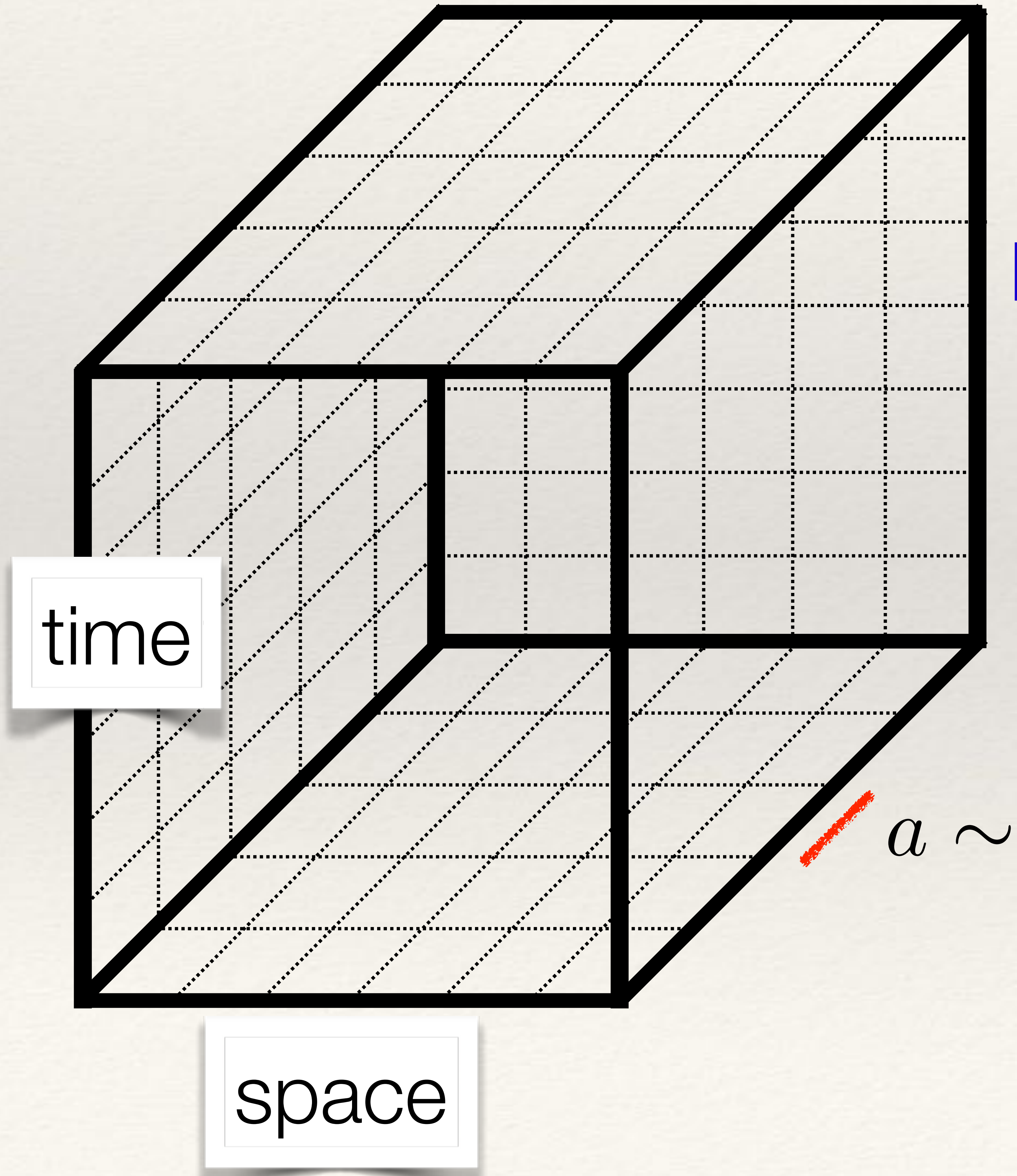
$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability

$$\{U_1, U_2, U_3, \ldots, U_N\}$$

Markov Chain Monte Carlo

$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(t)\mathcal{O}^\dagger(0)[U_i] + O\left(\frac{1}{\sqrt{N}}\right)$$



$$a \sim 10^{-16} \text{ meters} \sim \frac{1}{10} \text{size proton}$$

Need for HPC

- ultimately, we have to perform this integral of integrals

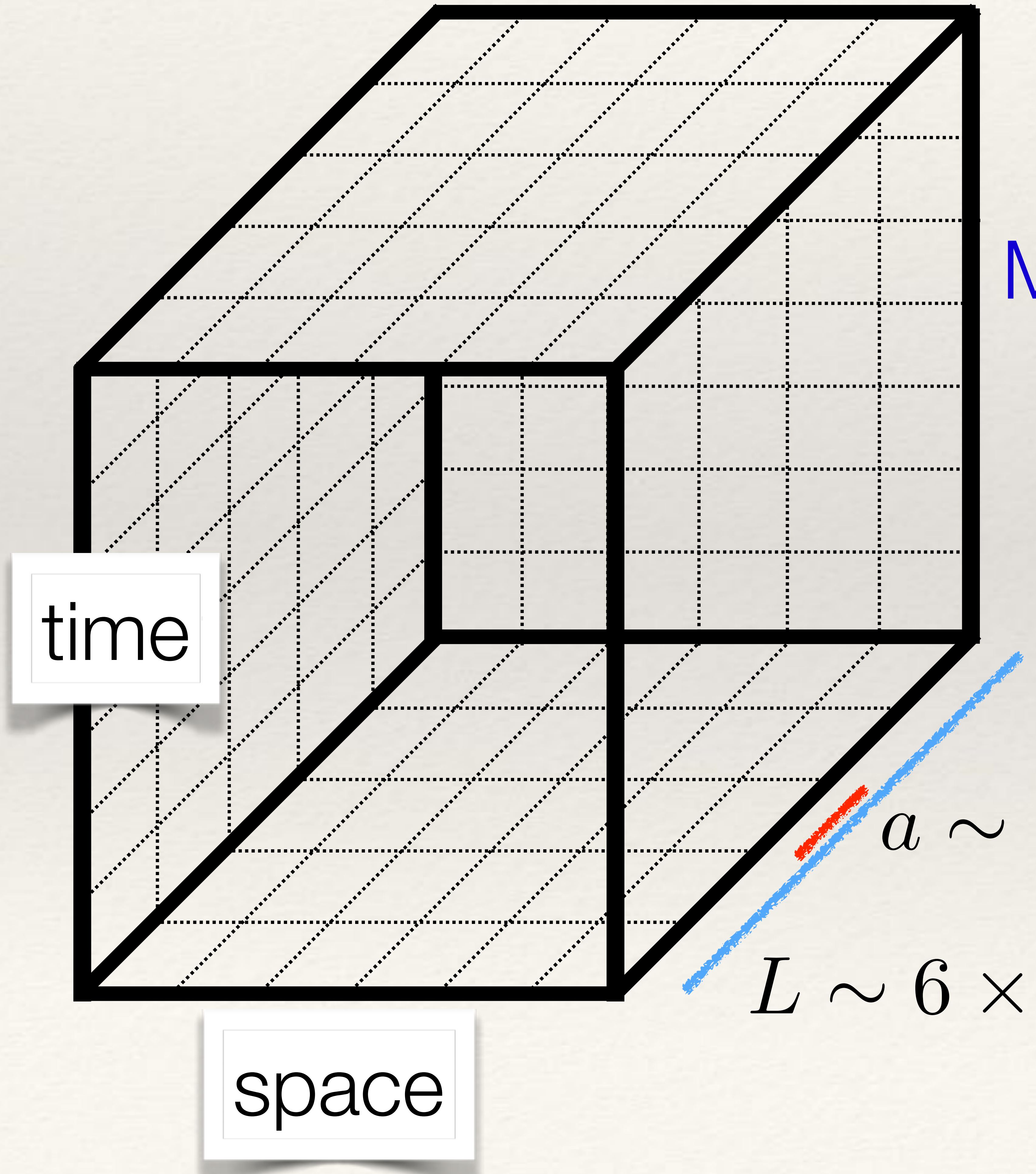
$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability

$$\{U_1, U_2, U_3, \ldots, U_N\}$$

Markov Chain Monte Carlo

$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(t)\mathcal{O}^\dagger(0)[U_i] + O\left(\frac{1}{\sqrt{N}}\right)$$



$$a \sim 10^{-16} \text{ meters} \sim \frac{1}{10} \text{size proton}$$

$$L \sim 6 \times 10^{-15} \text{ meters} \sim 6 \times \text{ size proton}$$

Need for HPC

- ultimately, we have to perform this integral of integrals

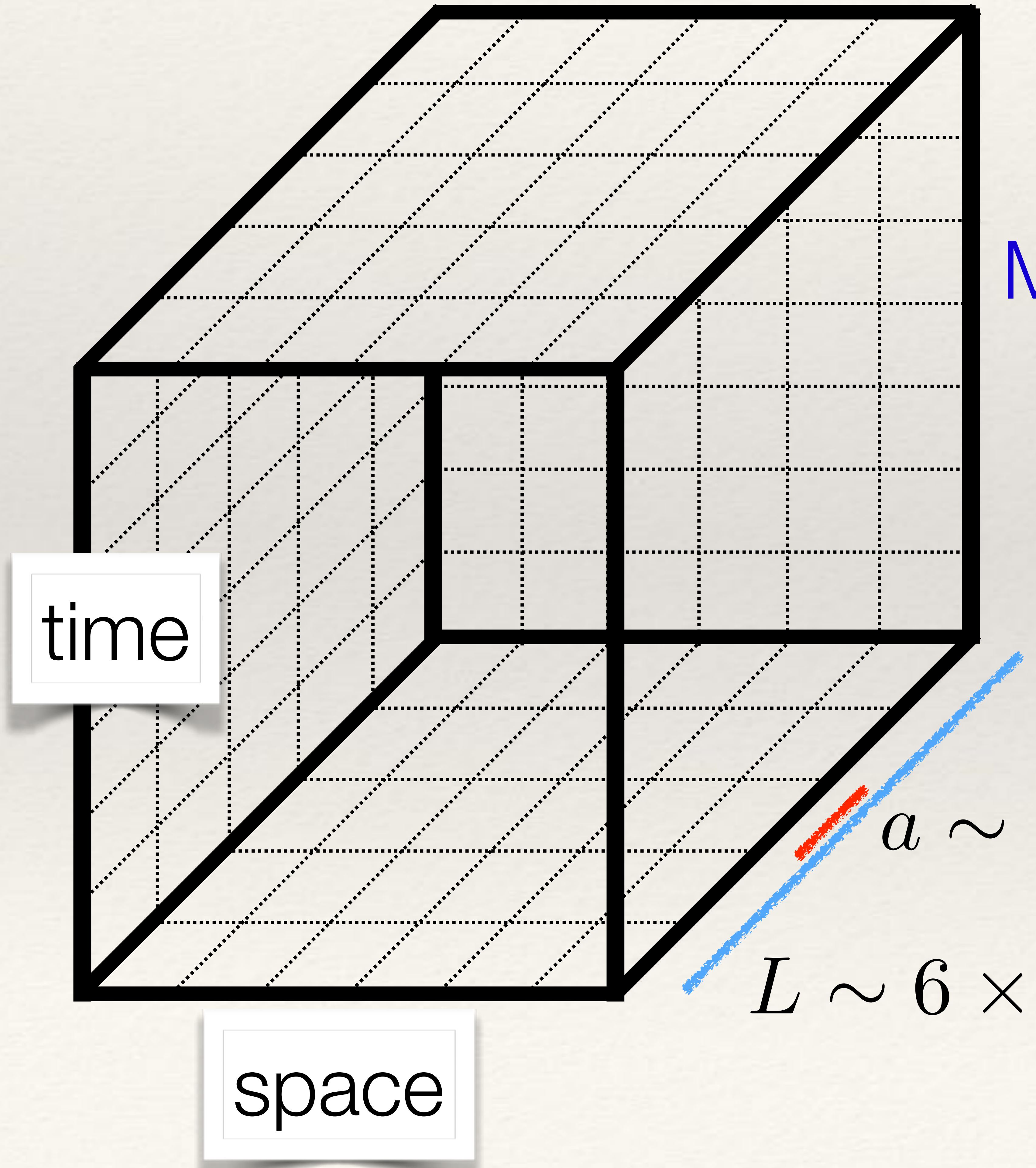
$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$

Probability

$$\{U_1, U_2, U_3, \ldots, U_N\}$$

Markov Chain Monte Carlo

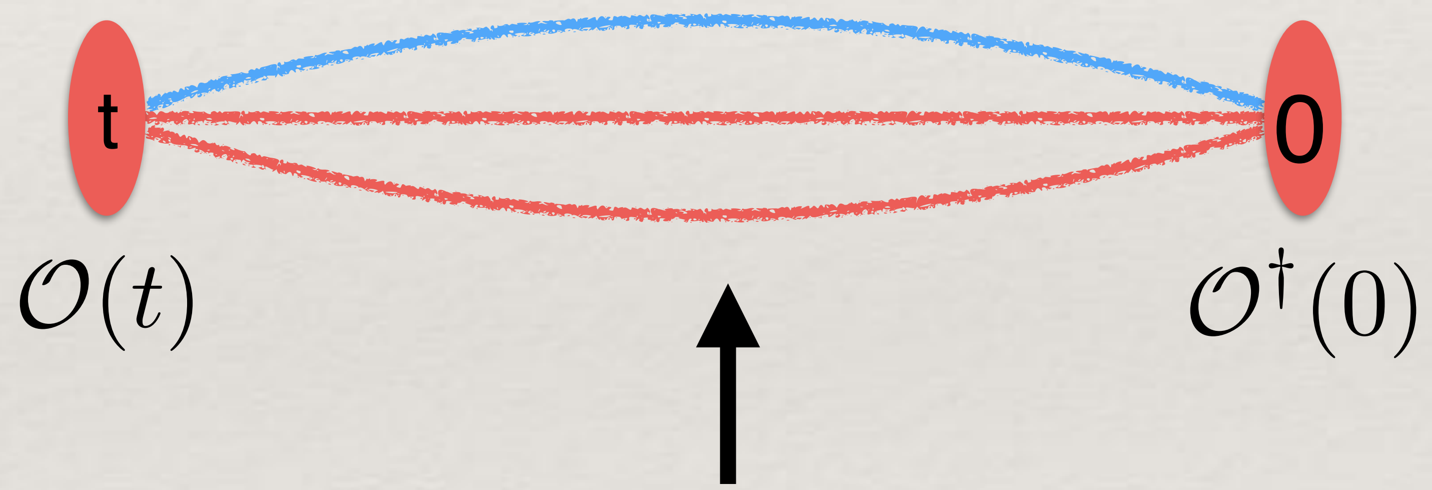
$$\approx \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(t)\mathcal{O}^\dagger(0)[U_i] + O\left(\frac{1}{\sqrt{N}}\right)$$



$$a \sim 10^{-16} \text{ meters} \sim \frac{1}{10} \text{size proton}$$

$$L \sim 6 \times 10^{-15} \text{ meters} \sim 6 \times \text{ size proton}$$

Suppose $\mathcal{O}^\dagger(0)$ is an initial proton and $\mathcal{O}(t)$ is a final state proton

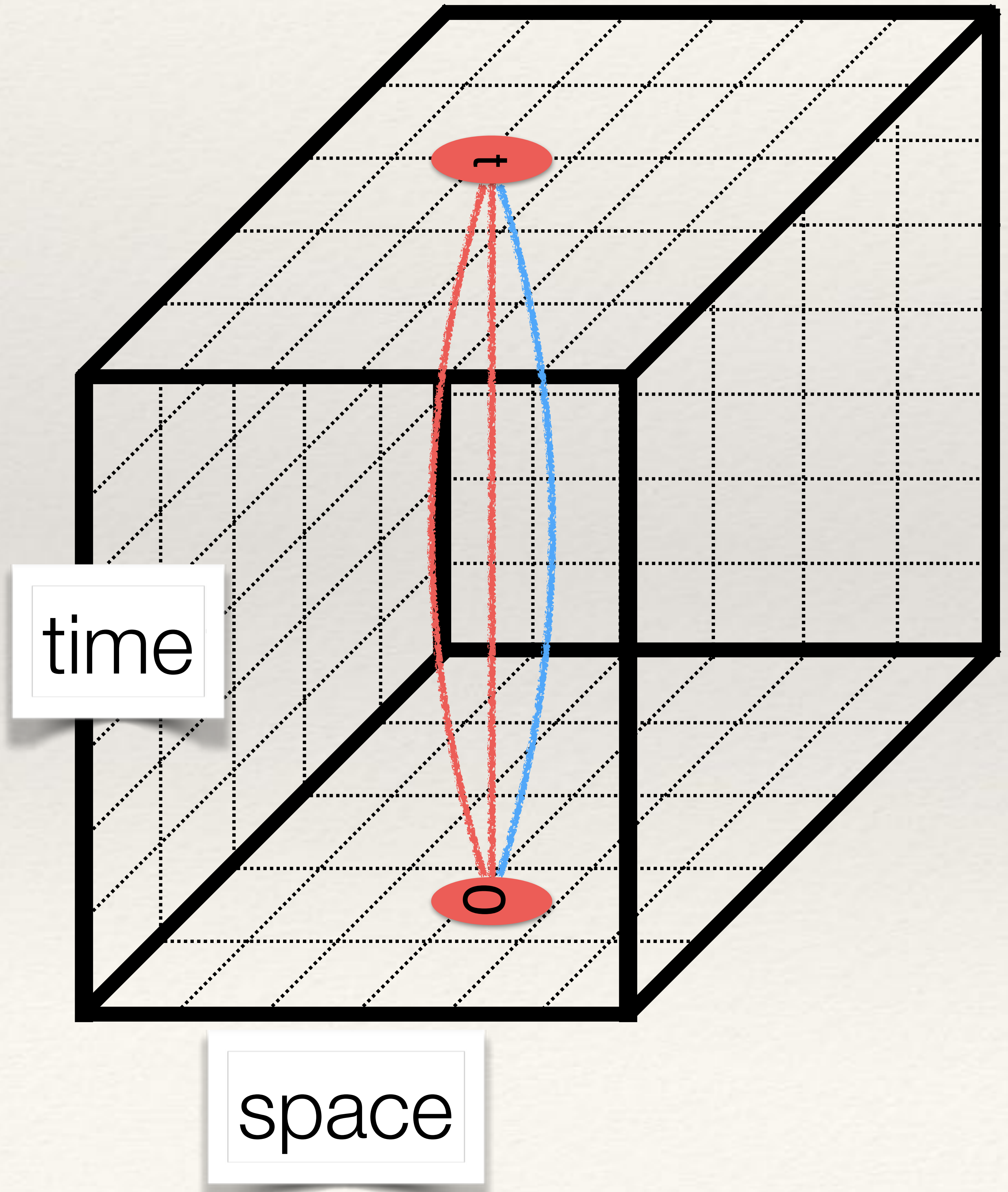


These objects are *quark propagators*

Need for HPC

- This integral describes the probability of finding the final state proton given the initial proton

$$C(t) = \langle \hat{O}(t)\hat{O}^\dagger(0)\rangle = \frac{1}{\mathcal{Z}}\int \mathcal{D}U \det(D+M) e^{-\int d^4x \mathcal{L}[U(t,x,y,z)]} O[U](t) O[U]^\dagger(0)$$



"$S_Q(t,0)$ = Quark Propagator"

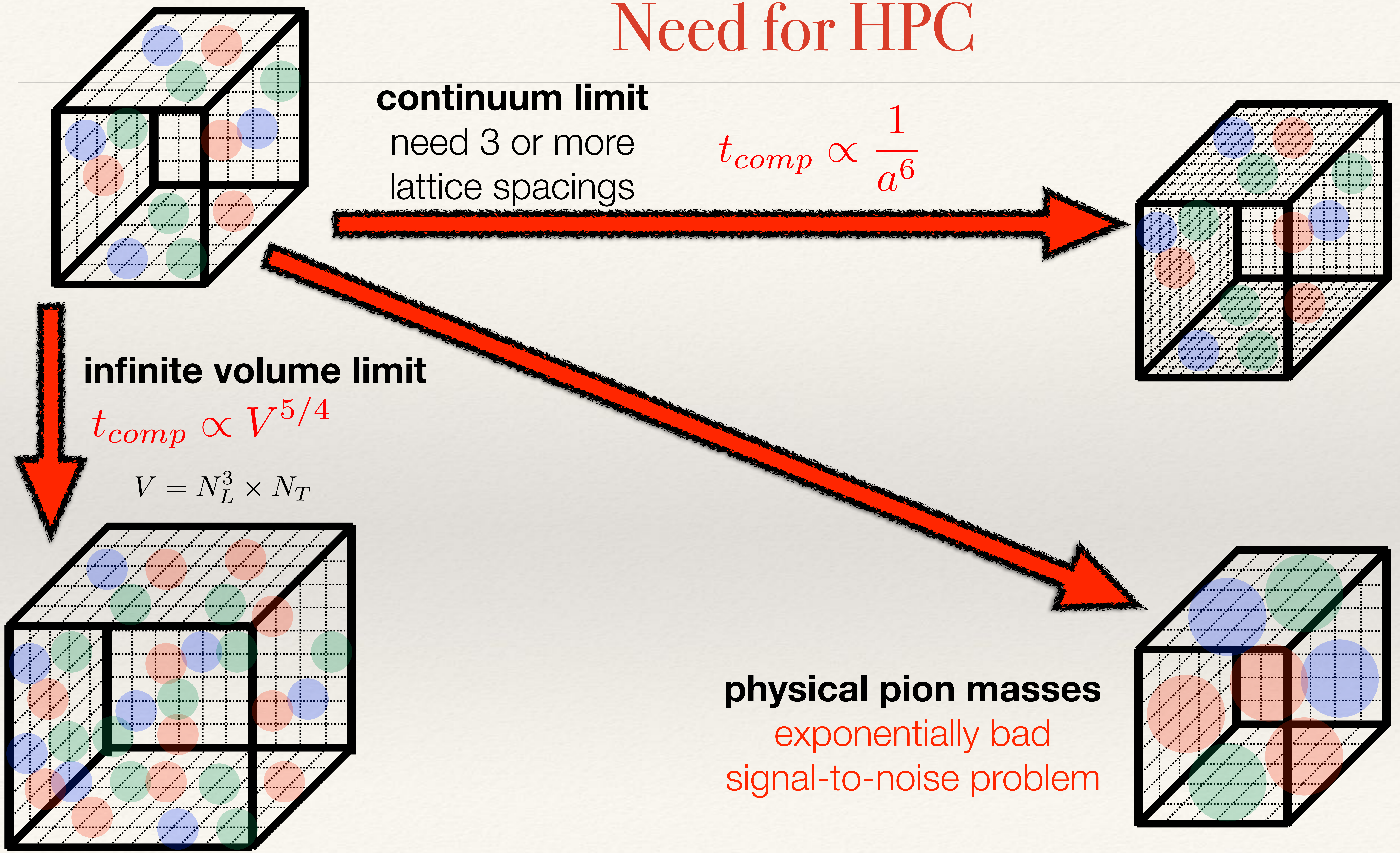
$$[D+M]_{y,z} S_{z,x} = \delta_{y,x} = \begin{cases} \text{if } x=y & 1 \\ \text{else} & 0 \end{cases}$$

Known Matrix
Sparse, Large :
~400,000,000

Solve the Quark Propagators with Conjugate Gradient

- **sparse Matrix-vector**
- GPUs are what allow us to do the calculations very efficiently
  - lots of memory bandwidth
  - memory access patters are structured and predictable

Need for HPC


continuum limit
need 3 or more
lattice spacings
$t_{comp} \propto \frac{1}{a^6}$
infinite volume limit
$t_{comp} \propto V^{5/4}$
$V = N_L^3 \times N_T$
physical pion masses
exponentially bad
signal-to-noise problem

Efficient Use of Sierra Summit

Efficient Use of Sierra/Summit

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

#1 on Top 500 list

O($200M) machine

Important to get as much science per flop as possible

Nodes	4608
Processor	IBM Power9 (2/node)
GPUs	27,648 NVIDIA V100 (6/node)
Node Performance	42 TeraFlops
Memory/node	512GB DDR4 + 96GB HBM2
NV-LINK Memory	1600GB/node
Interconnect	Mellanox 100G InfiniBand, Non-blocking Fat Tree
Peak Power	13 MegaWatts

Efficient Use of Sierra/Summit

OAK RIDGE National Laboratory | LEADERSHIP COMPUTING FACILITY

#1 on Top 500 list

O($200M) machine

Important to get as much science per flop as possible

- ❑ Need optimized use of NVIDIA GPUs
- ❑ Need scalability (communication avoiding algorithms)
- ❑ Need to take advantage of the heterogeneous architecture
- ❑ Need to efficiently manage hundreds of thousands of tasks

TEN YEARS OF QUDA

in use as GPU backend for BQCD, Chroma, CPS, MILC, TIFR, etc.

Solvers for all major fermionic discretizations

Routines needed for gauge-field generation

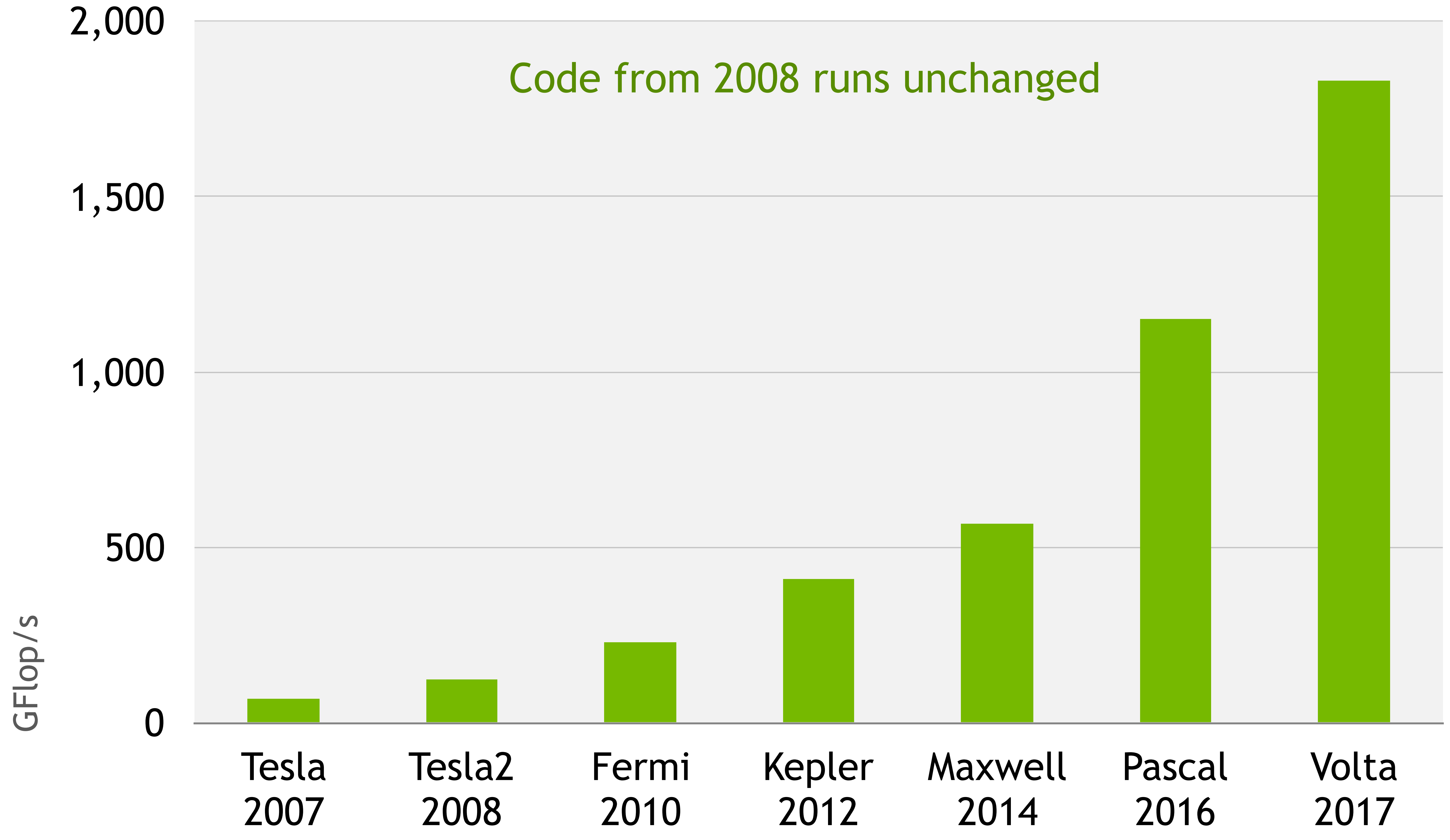
Maximize performance

- **Exploit symmetries to minimize memory traffic**
- **Mixed-precision methods (16 bit / 8 bit)**
- **Domain-decomposed (Schwarz) preconditioners for strong scaling**
- Eigenvector and deflated solvers (Lanczos, EigCG, GMRES-DR)
- Multi-source solvers
- Multigrid solvers for optimal convergence

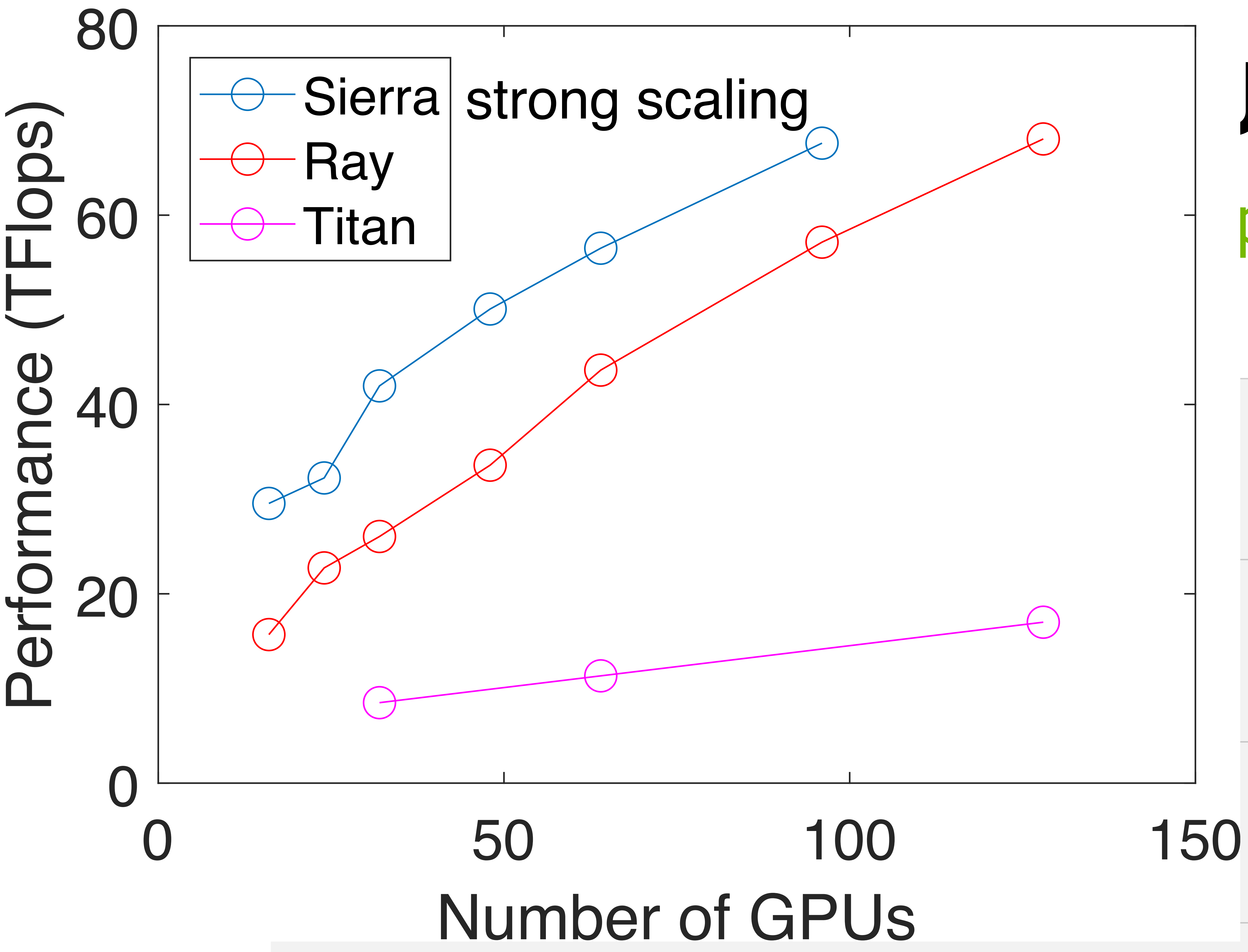
(thanks Kate and Mathias for slide)

RECOMPILE AND RUN

Autotuning provides performance portability

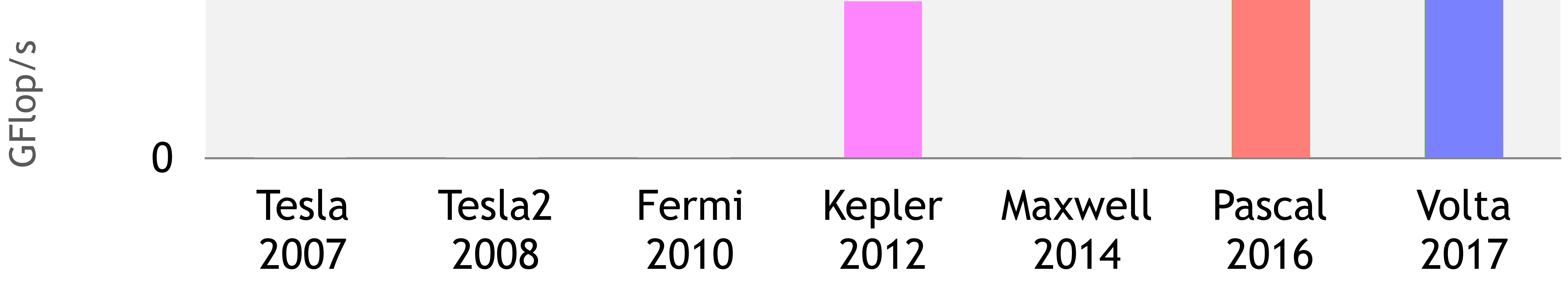


(thanks Mathias and Kate for slide)



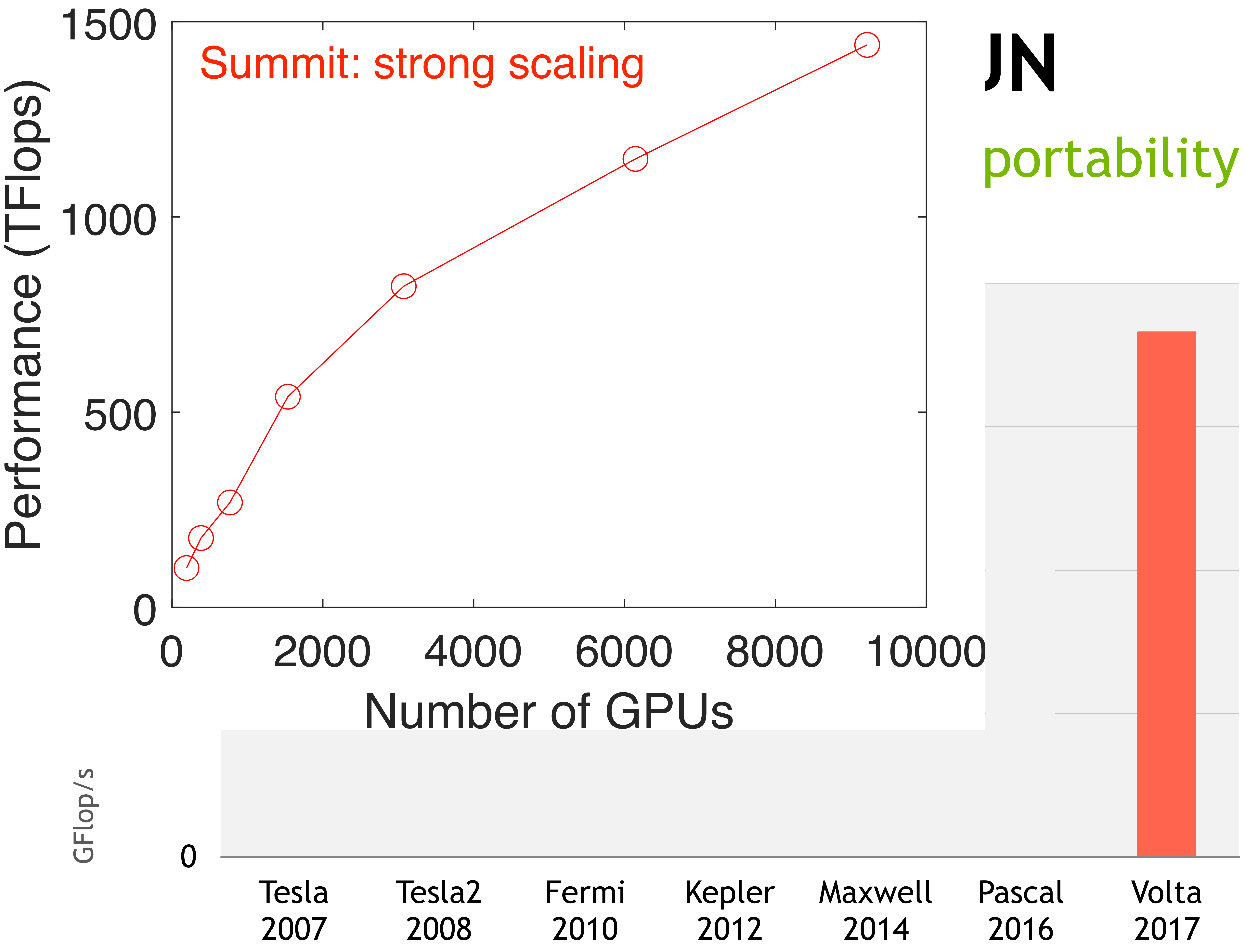
JN

portability



(thanks Mathias and Kate for slide)

NVIDIA



JN

portability

(thanks Mathias and Kate for slide)

Efficient Use of Sierra/Summit

- For our research problems of today, the majority of computing requirements reside in the large, sparse, linear system solves (quark propagators)

- We use the QUDA library, which is highly optimized for NVIDIA GPUs, written using CUDA C++ - large open BSD developed code
  https://github.com/lattice/quda

- See the talks

  S9708 - Strong Scaling HPC Applications: Best Practices with a Lattice QCD case study
  Kate Clark (one of our collaborators) and Mathias Wagner
  Thursday, 11am

  S9330 - Lattice QCD with Tensor Cores
  Jiqun Tu
  Thursday, 10am

  for a detailed discussion of QUDA and new uses of the tensor cores, respectively

Our work flow: managing millions of tasks

TASK GPU or CPU	RESOURCE NEEDS
Create a “snapshots” of the QCD vacuum via Monte-Carlo sampling	O(6 - 100) Summit nodes serial walk through Monte-Carlo ~1000 snapshots per “ensemble”
Prepare source for linear solve	~1-8 Summit nodes for ~1 minute ~8-32 sources per snapshot ~1000 snapshots ~20 ensembles
Perform linear system solve	for each: 1-8 Summit nodes for 5min - 2 hours depending on the ensemble for all sources/snapshots/ensembles
Dense Matrix multiplication with solve — subsequent solves based on first solve	for each: 1-8 Summit nodes for 5min - 30 min; 1-8 Summit nodes for 5min - 2 hours, ~100 subsequent solves for each solve
Dense Matrix multiplication with solve and secondary solve	for each: 1-8 Summit nodes for 5min - 30 min;

Our work flow: managing millions of tasks

TASK GPU or CPU	RESOURCE NEEDS
Create a “snapshots” of the QCD	O(6 - 100) Summit nodes

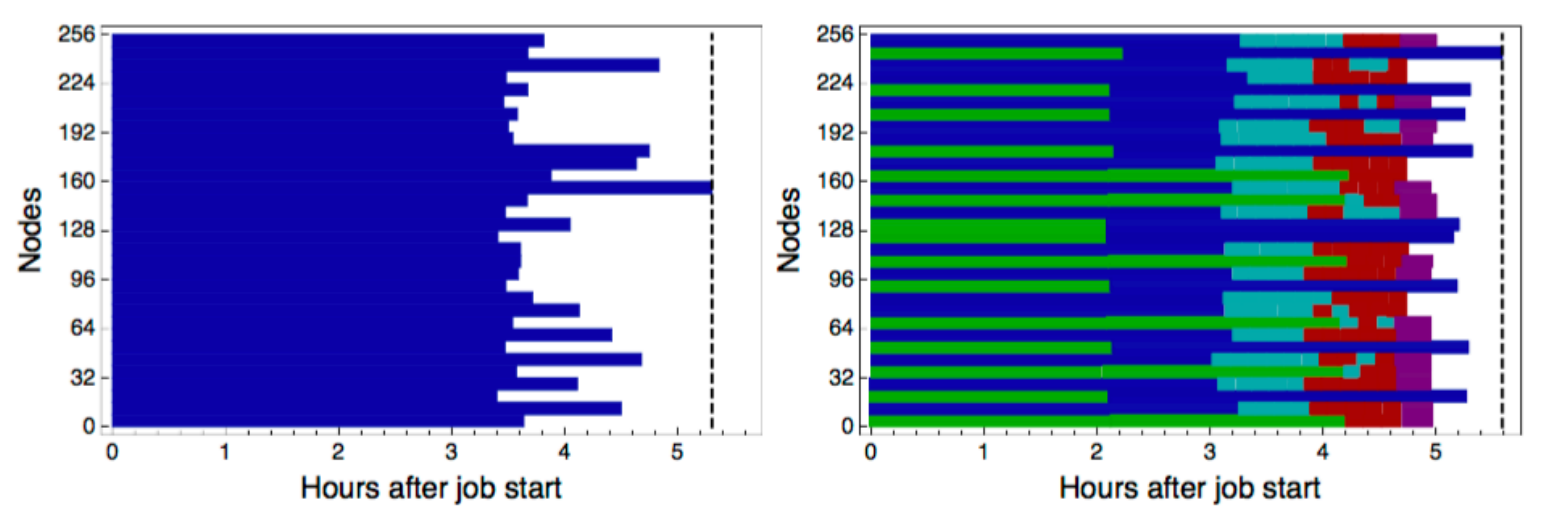
- We need to run hundreds of thousands to millions of independent, small node tasks
  - Running each job separately taxes the launch nodes on these supercomputers
- If we stack GPU tasks after CPU and so on, much of the wall-clock time will be wasted in the sense that the GPUs will not be in use - and the CPU wall-clock time is approaching a significant fraction of the GPU wall-clock time
- Job bundling wastes significant amounts of wall-clock time as performance of each task can vary substantially if nodes are close together or far apart
- We need a light-weight task manager capable of efficiently using all node resources and scheduling the tasks

Dense Matrix multiplication with solve and secondary solve	for each: 1-8 Summit nodes for 5min - 30 min;

First Generation: MetaQ

- Bash script system to improve efficiency of running many jobs in a single large allocation
  https://github.com/evanberkowitz/metaq

example from Titan



arXiv:1702.06122

- Naïve system resulted in waste of ~30% of time.

- MetaQ backfills tasks, reducing wasted (idle) cycles to about 5%

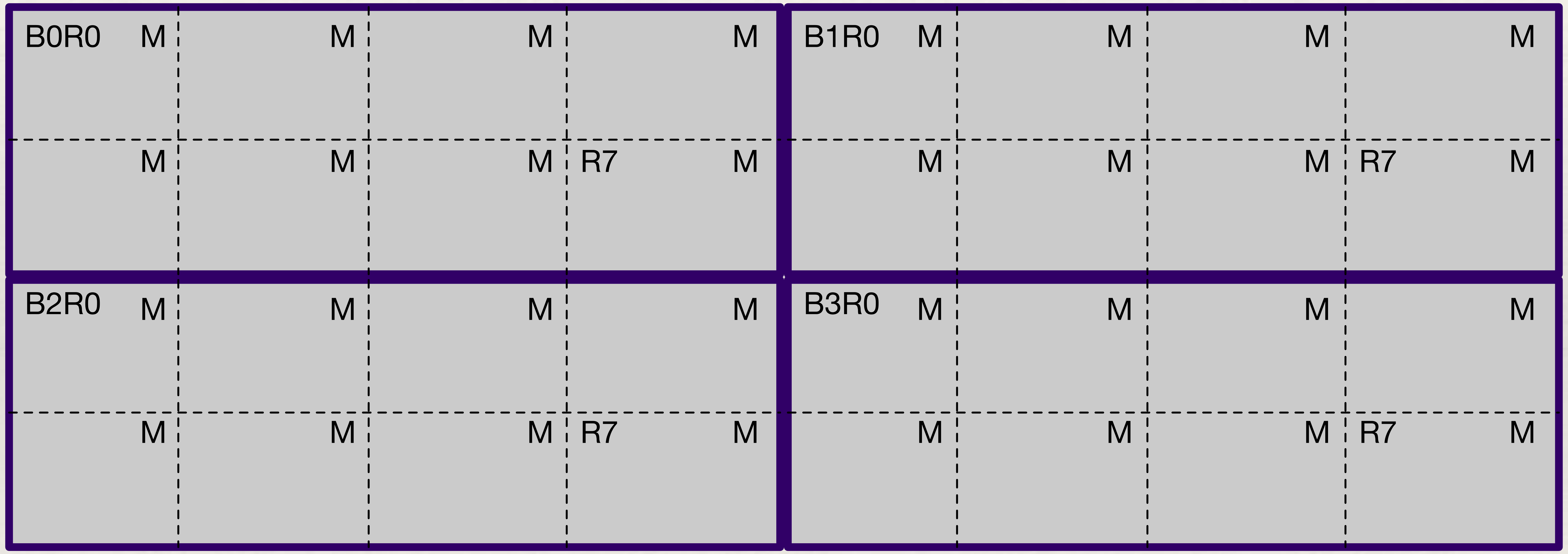
MPI_JM: MPI Job Manager

- Efficiently run a large set of tasks of bounded size in a large allocation of nodes.
- Low overhead on service node.
- Guarantee that communications are always physically local.
- Support overlay of jobs using distinct resources on the same nodes, i.e. GPU vs CPU jobs.
- Customizable collection/generation of workload with python based front-end.
    - Support pre and post actions that can be used to chain computations.
- Fault tolerance for startup on new super computers.

MPI_JM: https://github.com/kenmcelvain/mpi_jm (temporarily private)

Start by placing jm_master on each node in allocation.

mpirun -n 32 -map-by node jm_master jobs



Block sizes, memory, and slot description come from a configuration file. Each block gets a sub-communicator.

MPI_JM

Controlling Interconnect Performance

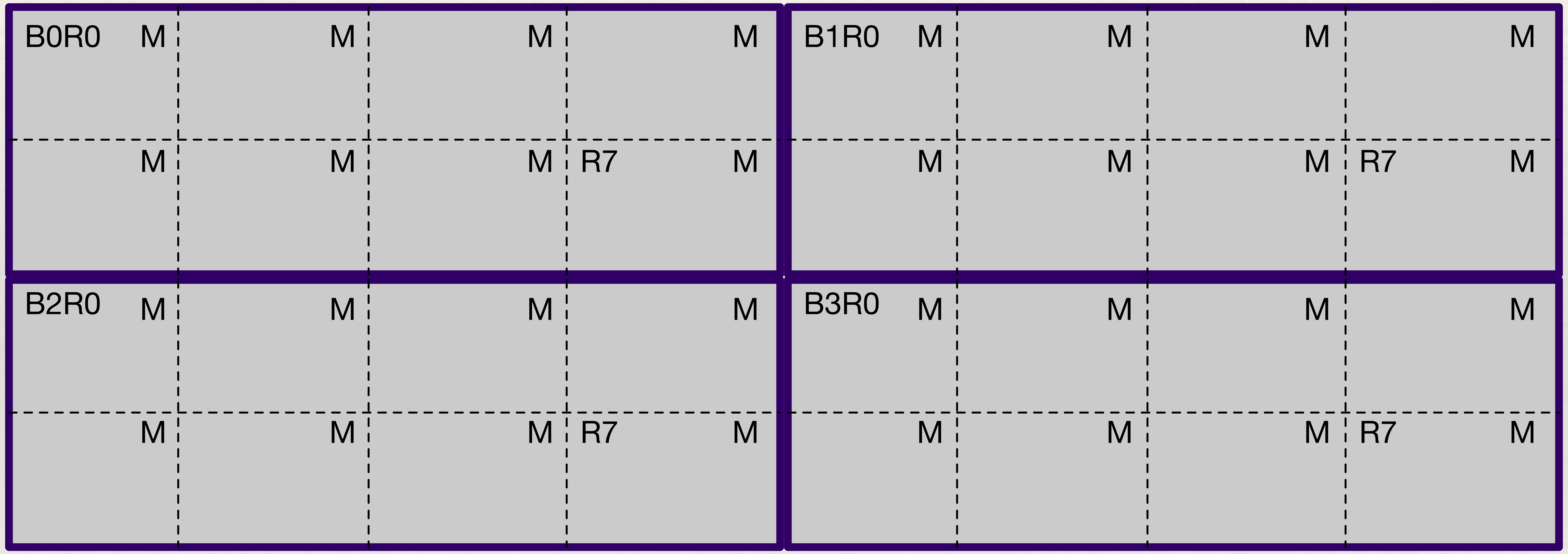
Maximize blocks with node numbers in sequence.

B0R0 M 807	M 1130	M 2101	M 3355	B1R0 M 1122	M 1123	M 1124	M 1125
M 3356	M 3516	M 3517	R7 M 3518	M 1126	M 1127	M 1128	R7 M 1129
B2R0 M 3340	M 3341	M 3342	M 3343	B3R0 M 3348	M 3349	M 3349	M 3350
M 3344	M 3345	M 3346	R7 M 3347	M 3351	M 3352	M 3353	R7 M 3354

As tasks finish and new ones are started, they will get consistent interconnect performance. We rely on individual tasks having bounded size.

MPI_JM

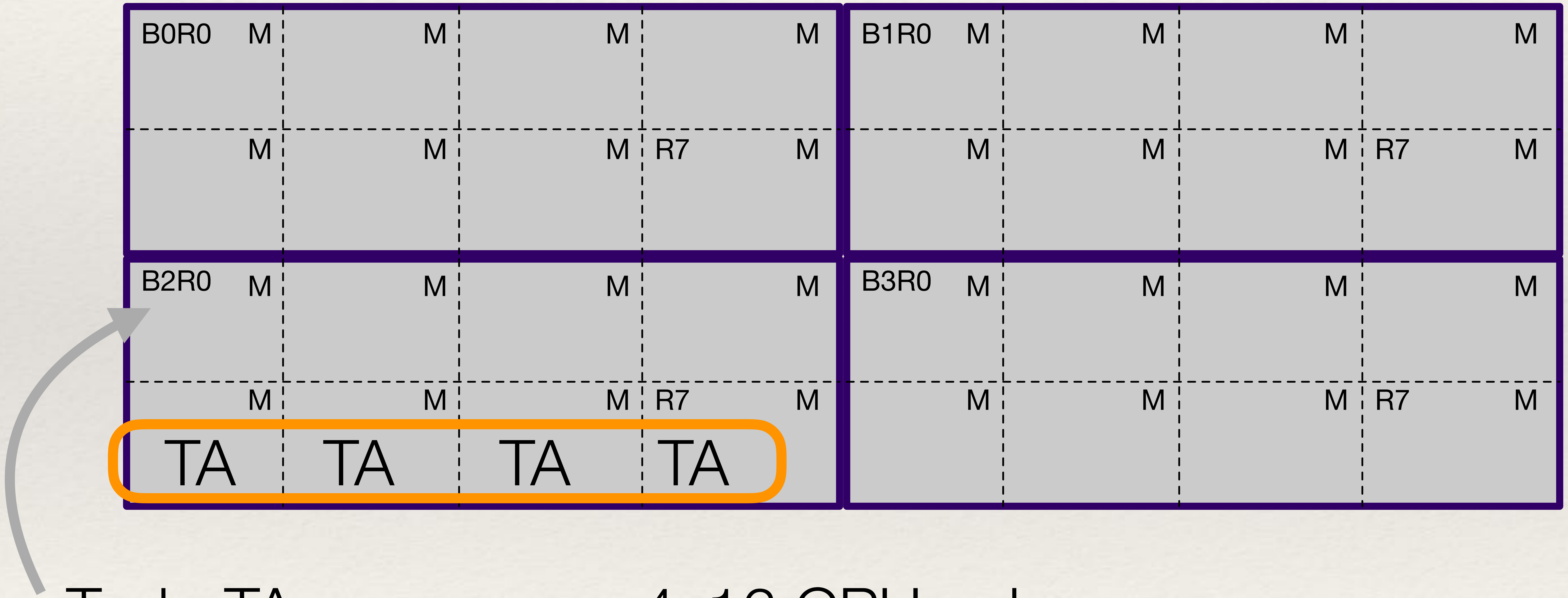
- The scheduler matches a task to block resources and sends task parameters to rank 0 of a block.
- The task “disconnects” until it completes.



Task	TA:	src,	4x16 CPU only
Task	TB:	solve,	8x1 CPU/GPU
Task	TC:	contract,	8x14 CPU only

MPI_JM

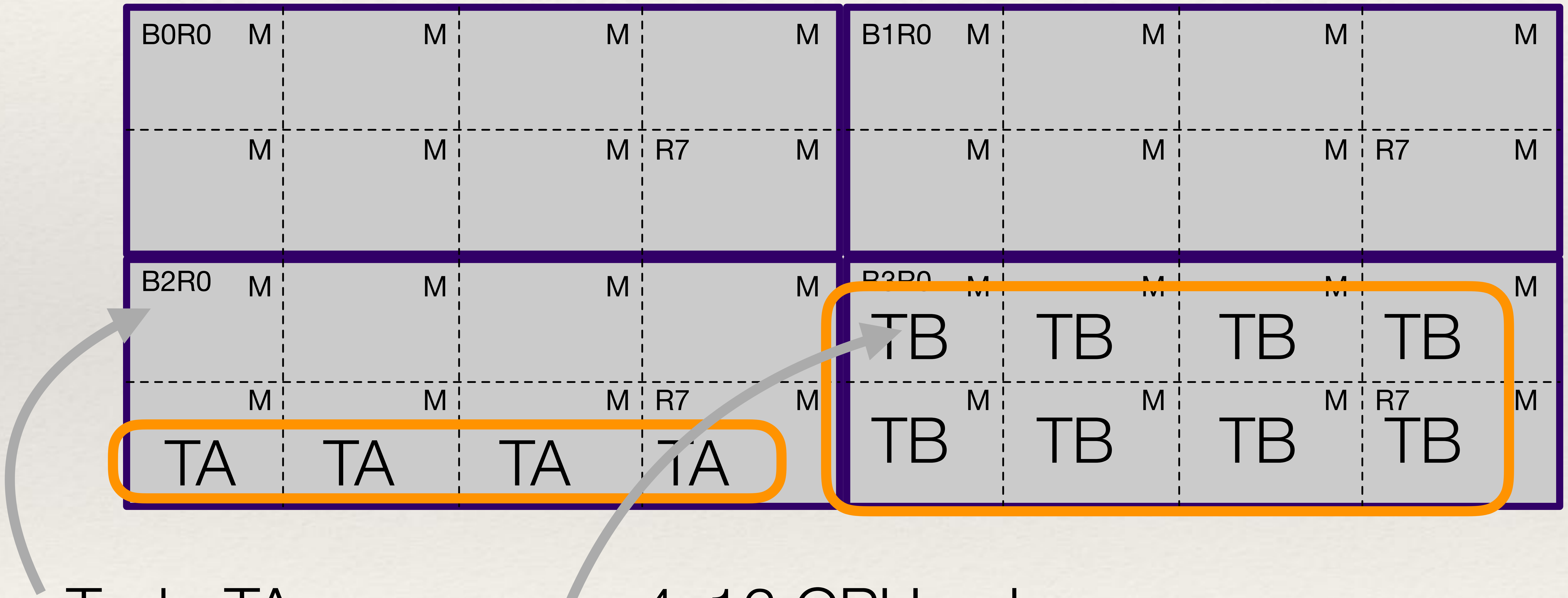
- The scheduler matches a task to block resources and sends task parameters to rank 0 of a block.
- The task "disconnects" until it completes.



Task TA: src, 4x16 CPU only
Task TB: solve, 8x1 CPU/GPU
Task TC: contract, 8x14 CPU only

MPI_JM

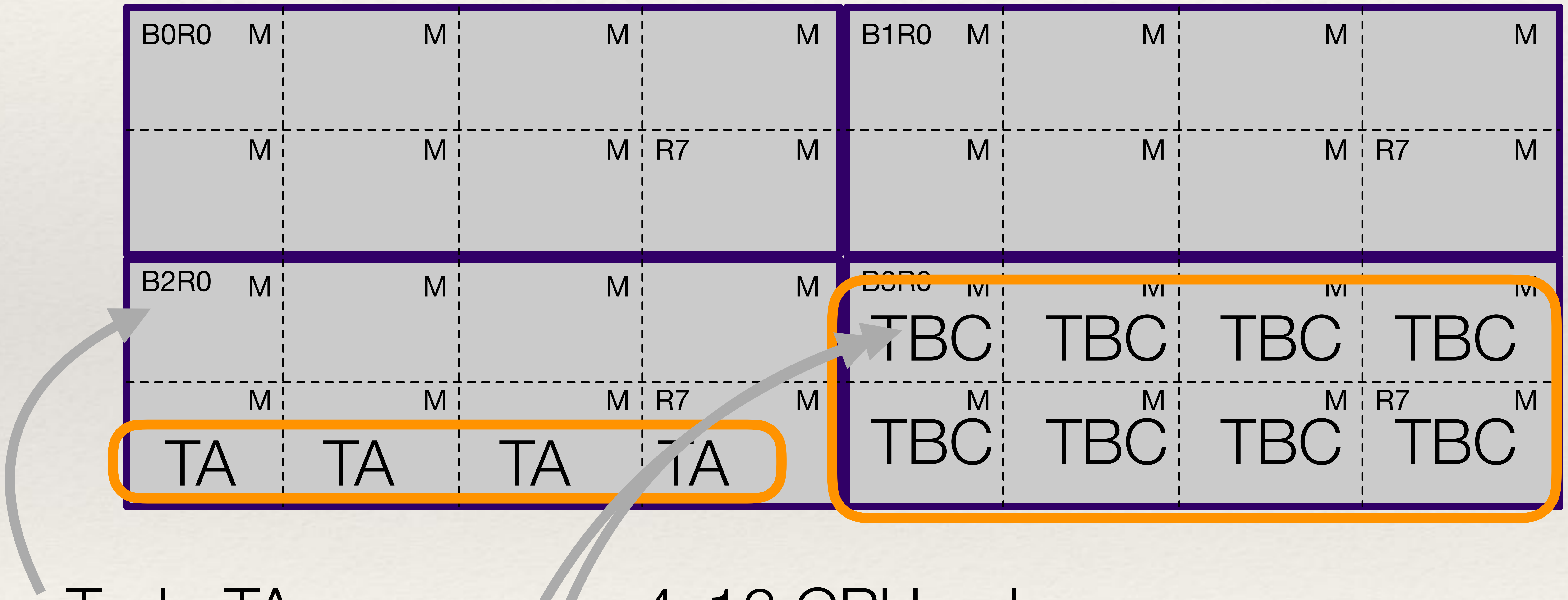
- The scheduler matches a task to block resources and sends task parameters to rank 0 of a block.
- The task “disconnects” until it completes.



Task TA: src, 4x16 CPU only
Task TB: solve, 8x1 CPU/GPU
Task TC: contract, 8x14 CPU only

MPI_JM

- The scheduler matches a task to block resources and sends task parameters to rank 0 of a block.
- The task "disconnects" until it completes.



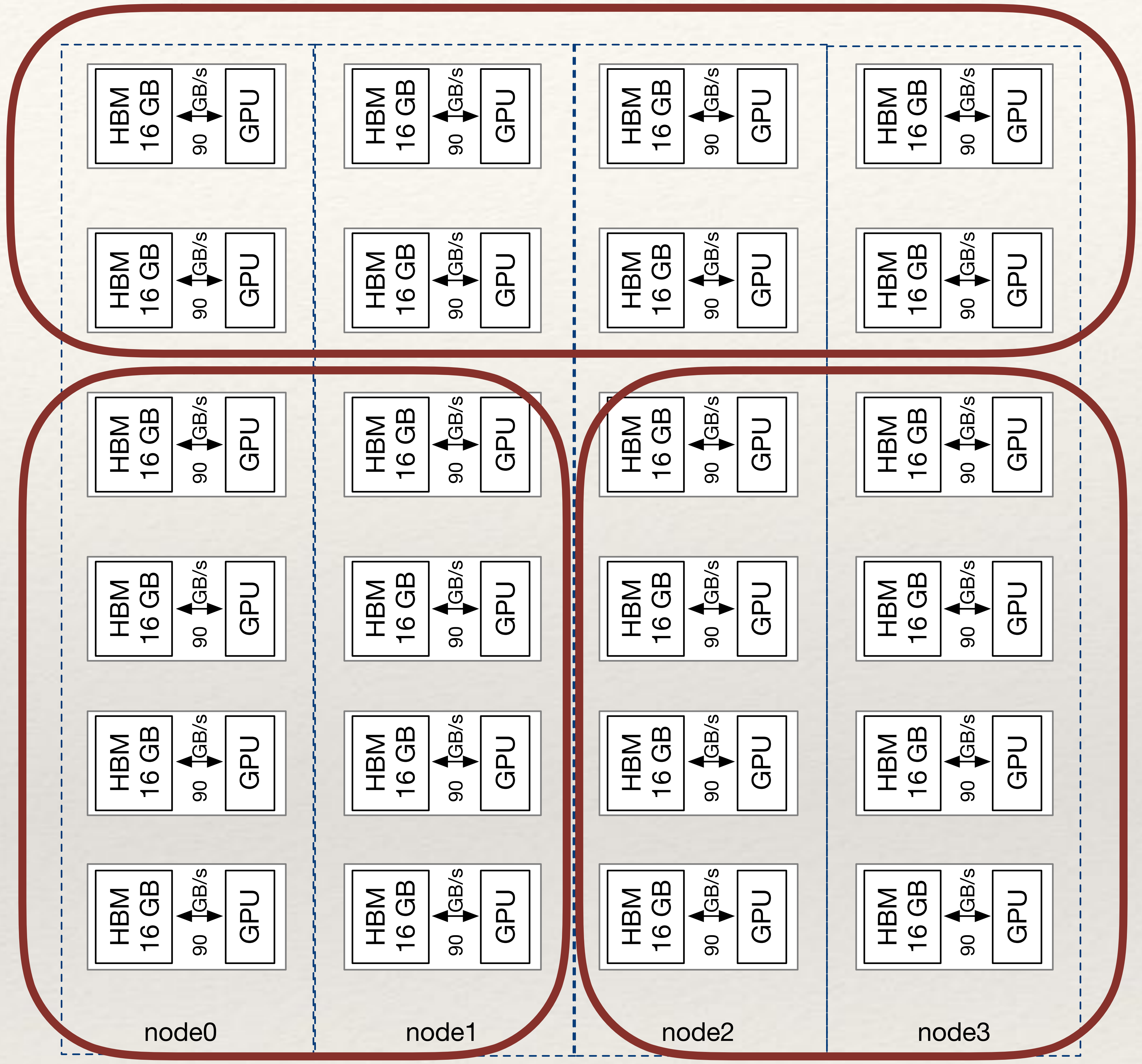
Task TA: src, 4x16 CPU only
Task TB: solve, 8x1 CPU/GPU
Task TC: contract, 8x14 CPU only

MPI_JM

Not all problems have factors of 3 in them (Summit nodes have 6 GPUs/node)

MPI_JM provides micro-control of rank placement

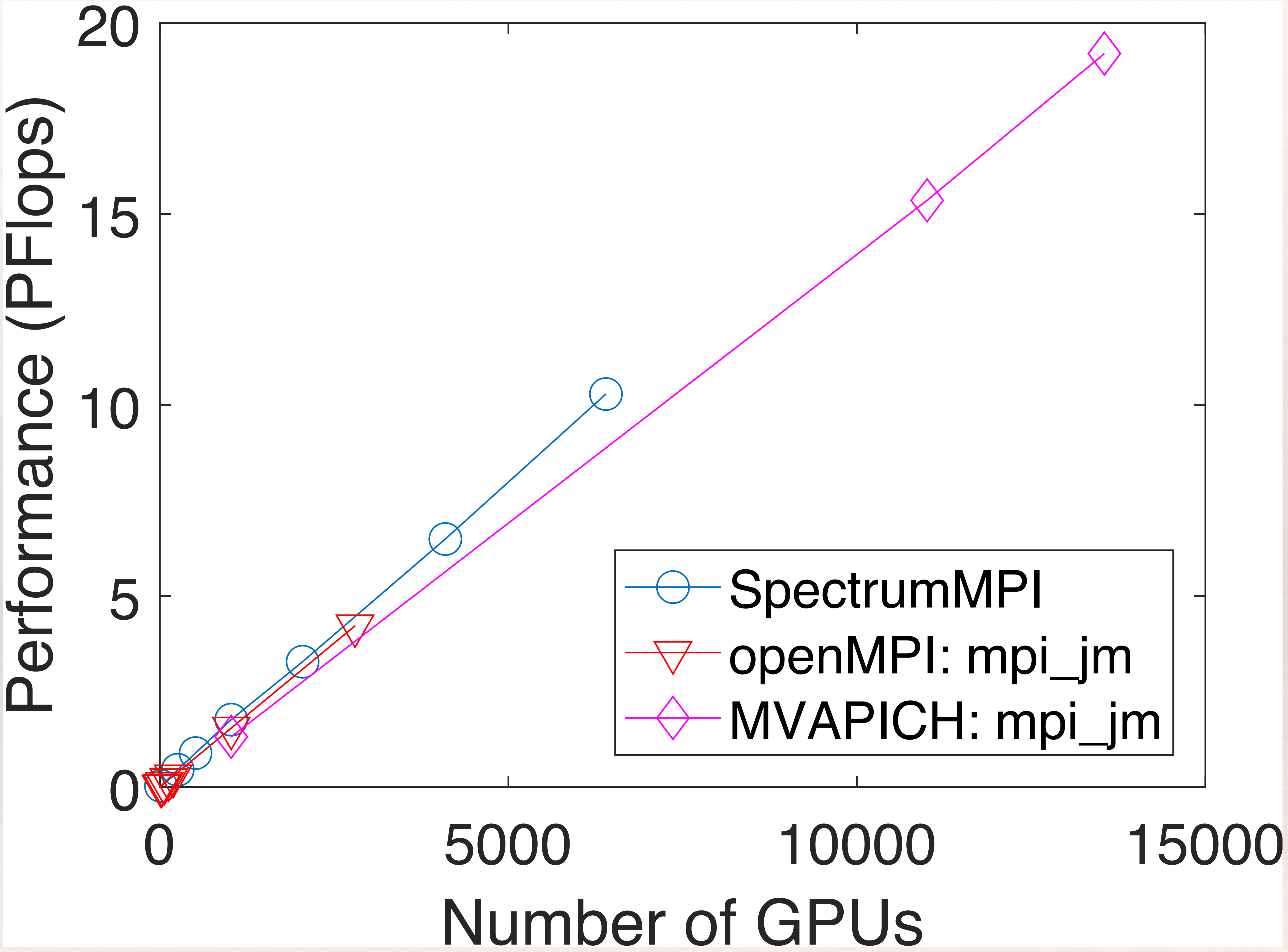
- Three 8 GPU jobs distributed over 4 Summit nodes.
- QUDA Autotuner configures the top job differently because interconnect is different.
- Most CPU cores remain open for CPU only jobs.
- Try to keep all parts of nodes doing productive work!



MPI_JM: Scaling in Sierra

- The SpectrumMPI run was done with 400 separate jobs.
- MPI_JM on OpenMPI was run with 7 independent jobs of 100 nodes due to reliability issues.
- Under **MVAPICH** full DPM support was available and separate lumps of nodes were started independently. Successful lumps connect afterwards and are managed together! ~4 min startup for all 4000 nodes!
- The combination of GPU performance and our management software got us selected as a 2018 Gordon Bell Finalist

MVAPICH runs done with single job



MPI_JM: Scaling in Sierra

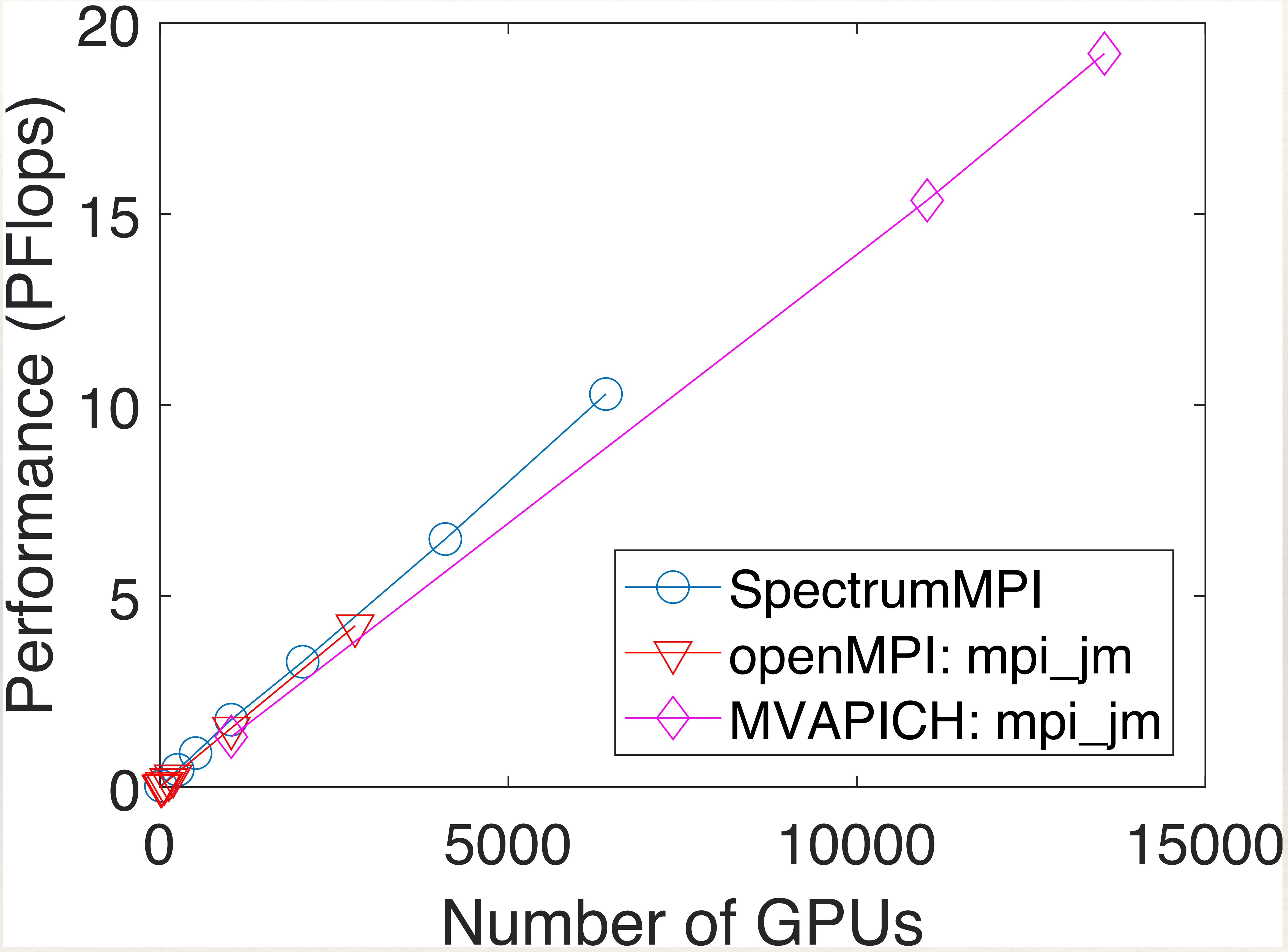
- Compared to other applications, our peak performance does not seem impressive

- Machine-to-machine, compared to Titan @ ORNL (Titan is 18,688 nodes 16-core AMD + 1 K20/node)

  For our research
  Sierra (~4300) is **~10 times** faster than Titan
  Summit (~4600) is **~15 times** faster than Titan

  These machines are **disruptively** faster than previous computers

MVAPICH runs done with single job



MPI_JM: Scaling in Sierra

- Compared to other applications, our peak performance does not seem impressive

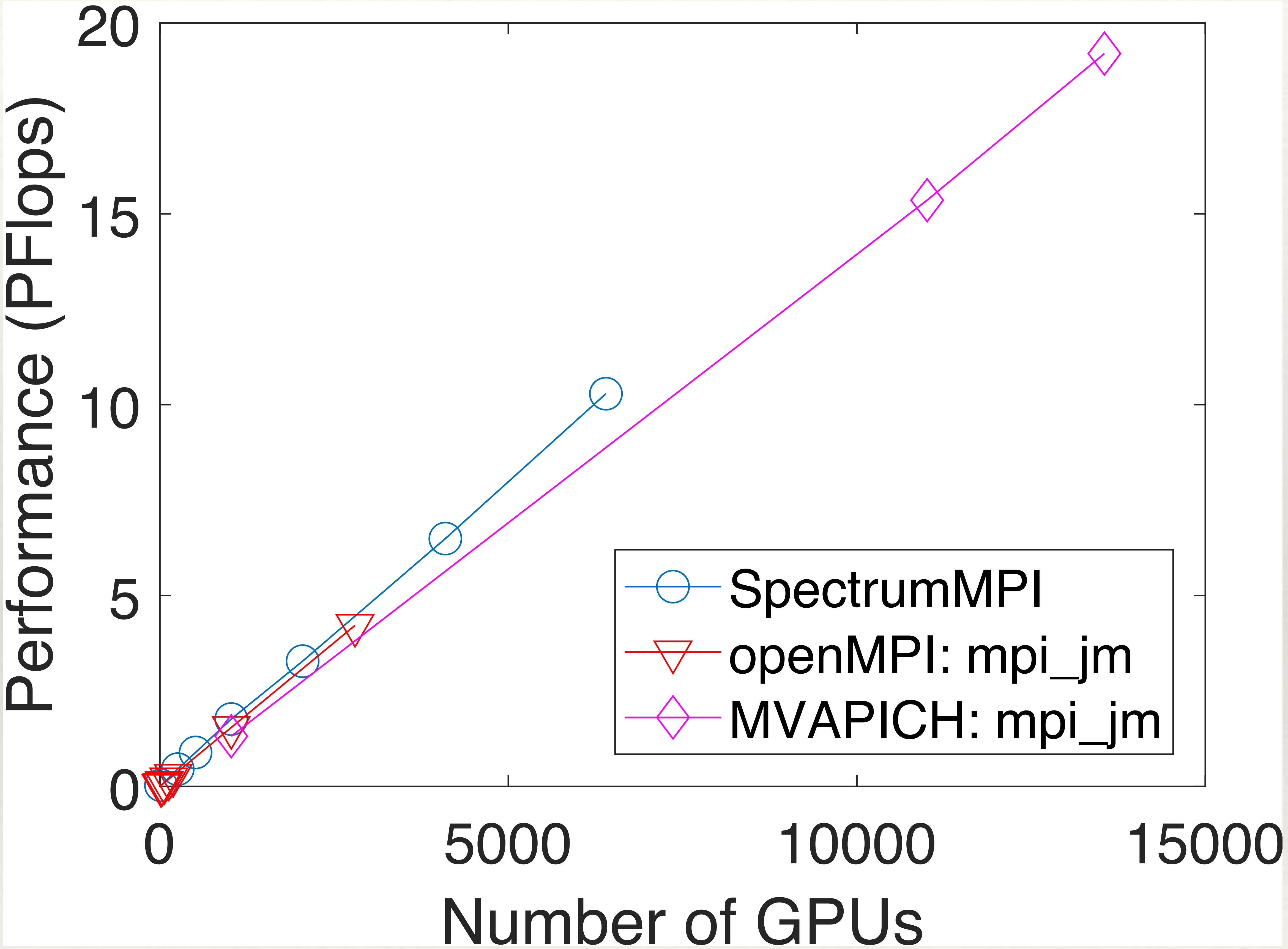
- Machine-to-machine, compared to Titan @ ORNL (Titan is 18,688 nodes 16-core AMD + 1 K20/node)

  For our research
  Sierra (~4300) is **~10 times** faster than Titan
  Summit (~4600) is **~15 times** faster than Titan

  These machines are **disruptively** faster than previous computers

- In 2.5 weekends on Sierra - we accomplished 5x more than in 1 year on Titan

MVAPICH runs done with single job



What did this get us?

- ~3 months Early Science Time on Sierra @ LLNL
  Sep. - Dec. 2018

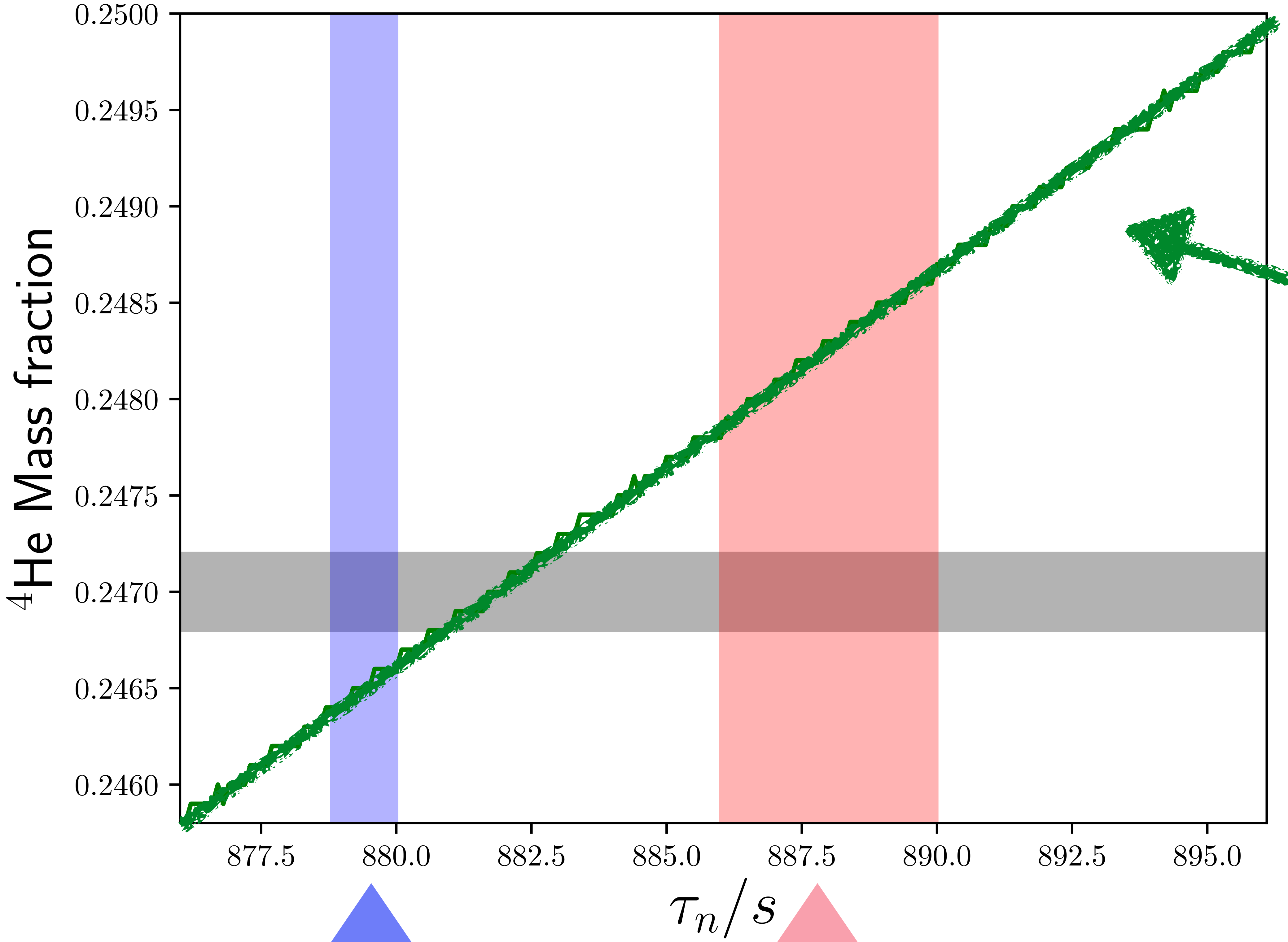
- Early Science Time on Summit @ ORNL
  Jan. +

- 2019 DOE INCITE Allocation on Summit
  Jan. +

Why is understanding the neutron lifetime interesting?

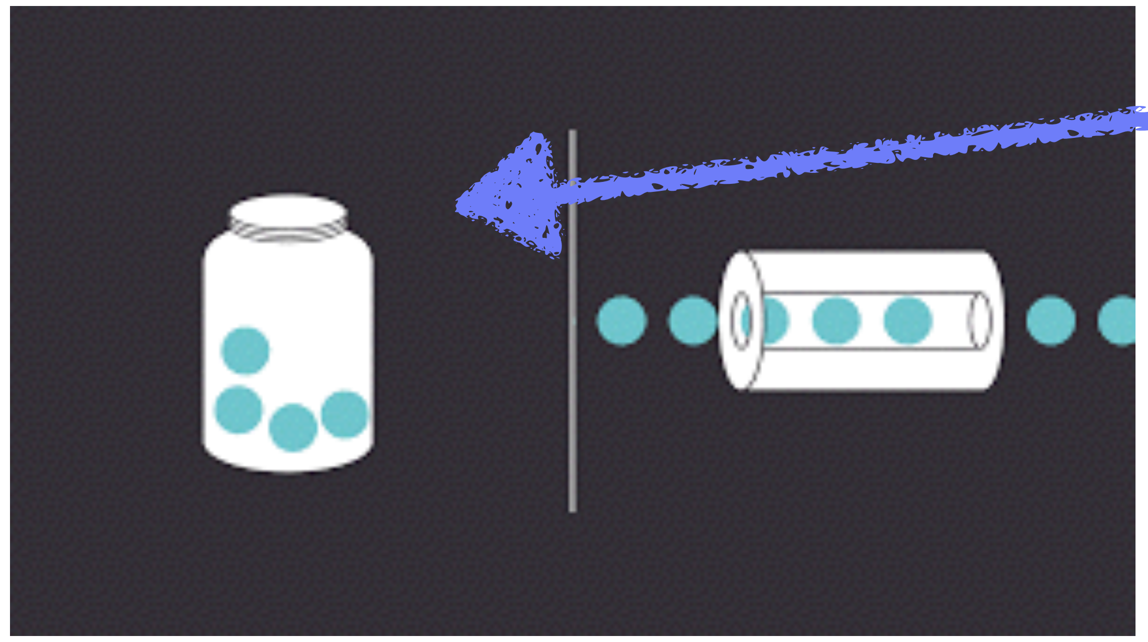
$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1 + 3g_A^2}$$

Czarnecki, Marciano, Sirlin



Predicted primordial ^{4}He mass fraction of the universe as a function of the neutron lifetime

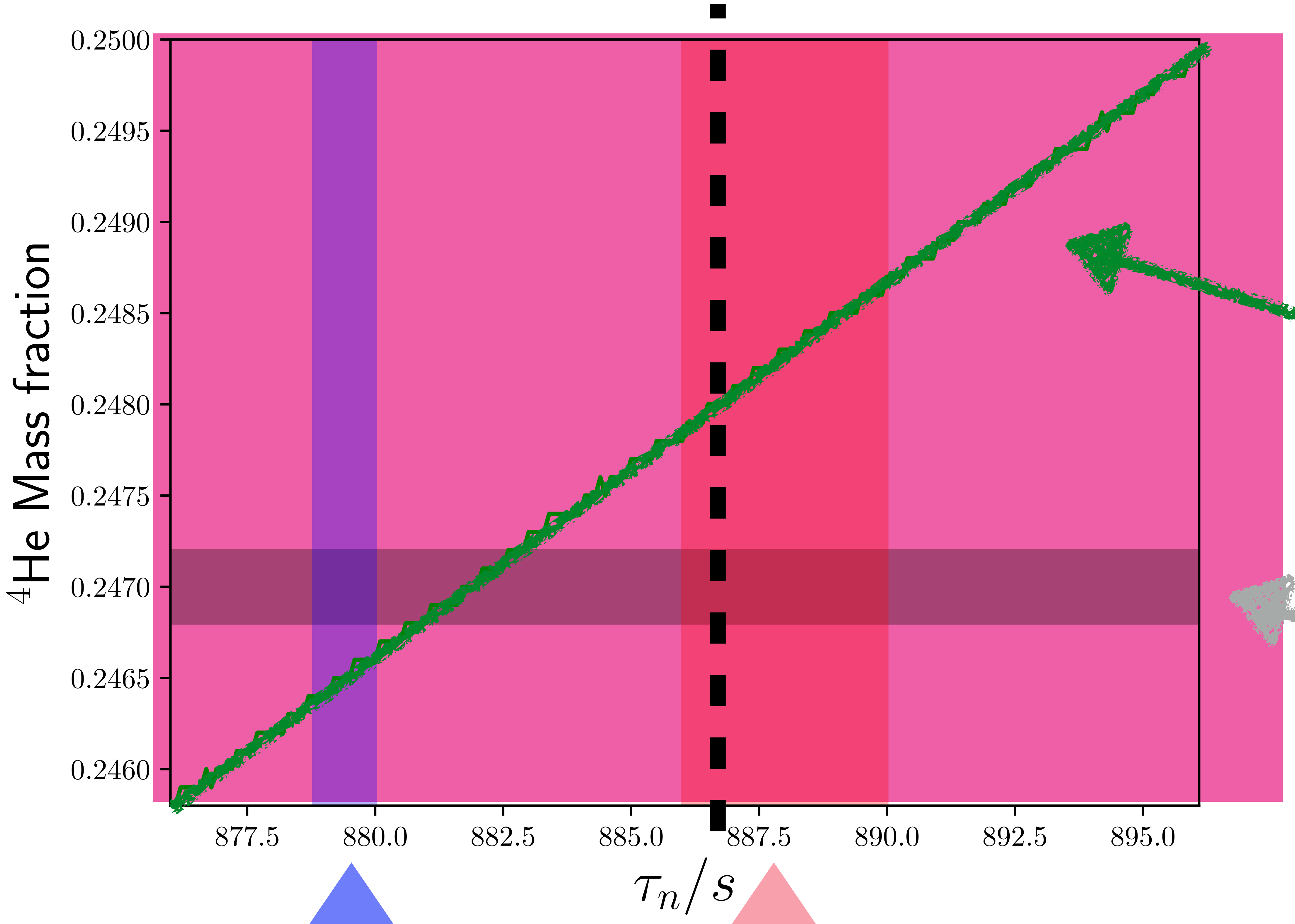
Observed primordial ^{4}He mass fraction of the universe



Two different methods of measuring the neutron lifetime disagree at the 99% level
bottle and beam

Is one of the experiments wrong? Or is there new physics hiding here?

Why is understanding the neutron lifetime interesting?

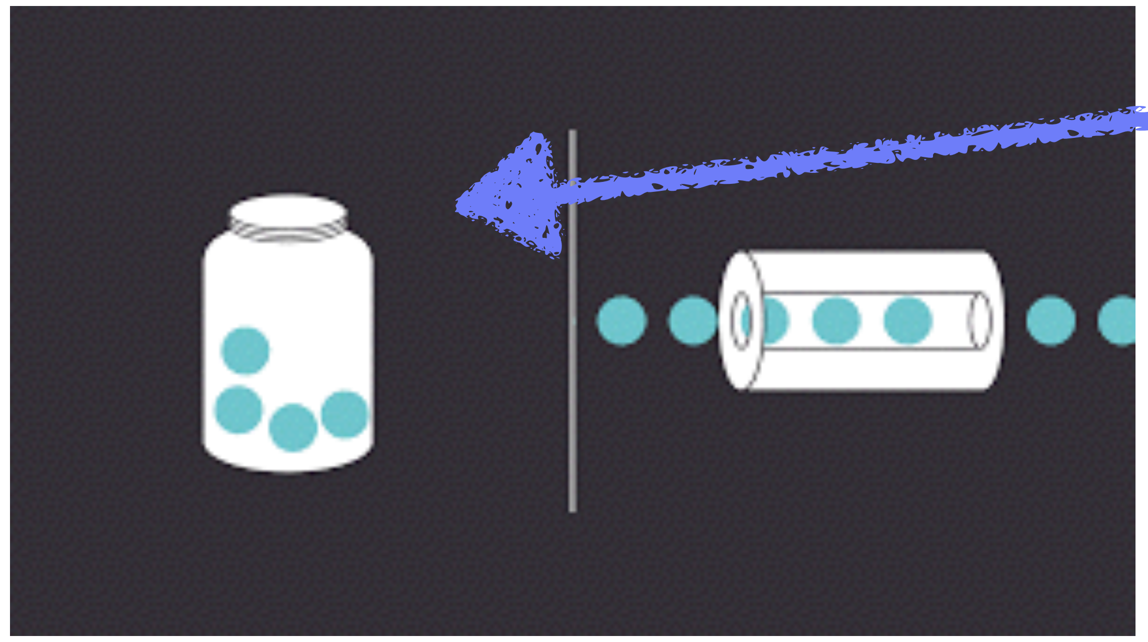


$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1 + 3g_A^2}$$

Czarnecki, Marciano, Sirlin

Predicted primordial ^{4}He mass fraction of the universe as a function of the neutron lifetime

Observed primordial ^{4}He mass fraction of the universe

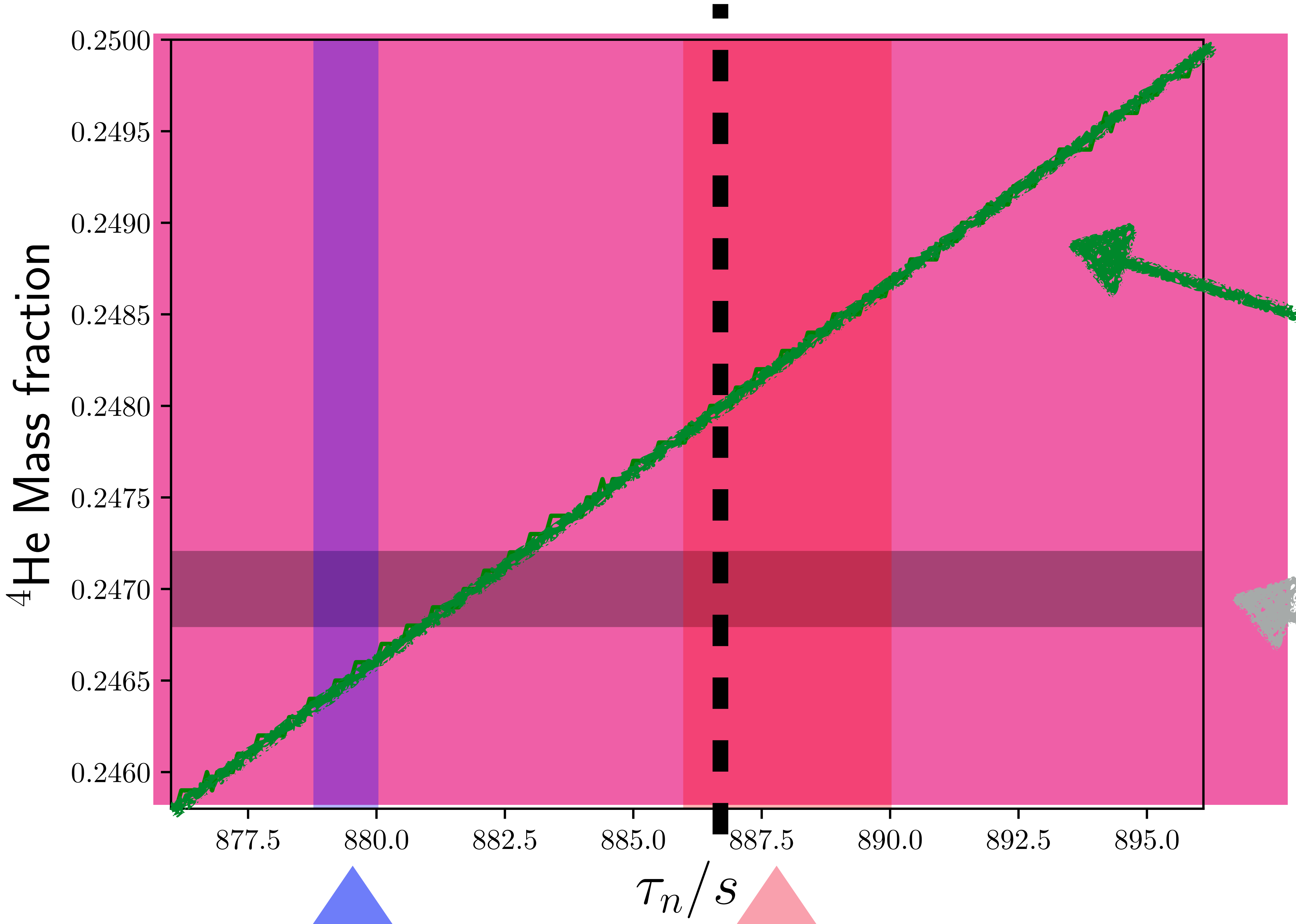


Two different methods of measuring the neutron lifetime disagree at the 99% level

bottle and beam

Is one of the experiments wrong? Or is there new physics hiding here?

Why is understanding the neutron lifetime interesting?

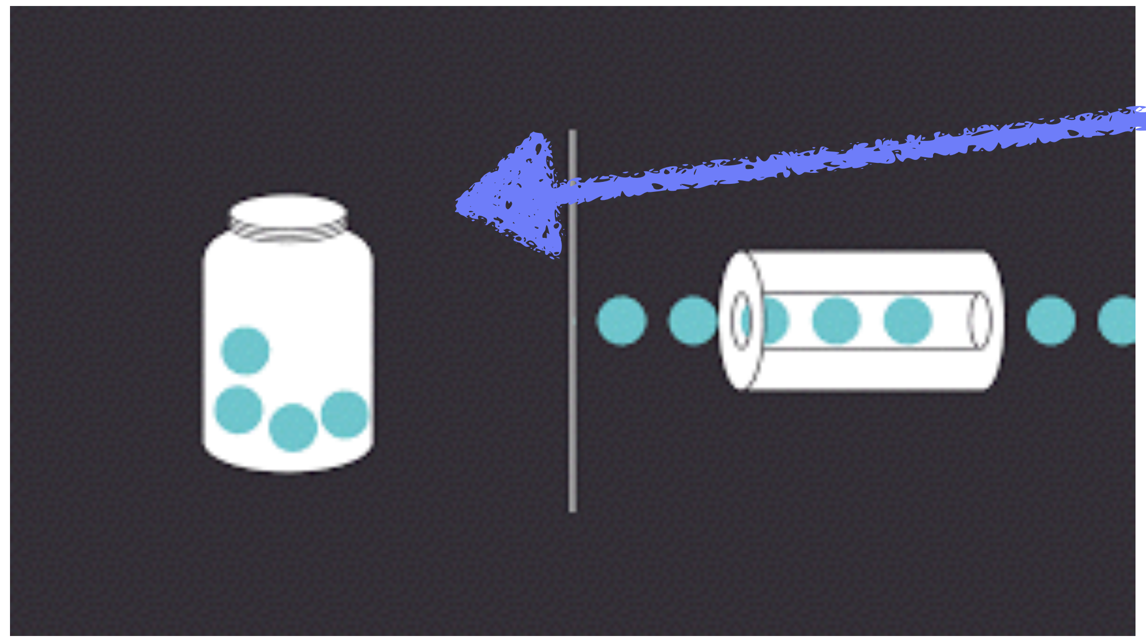


$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1 + 3g_A^2}$$

Czarnecki, Marciano, Sirlin

Predicted primordial ^{4}He mass fraction of the universe as a function of the neutron lifetime

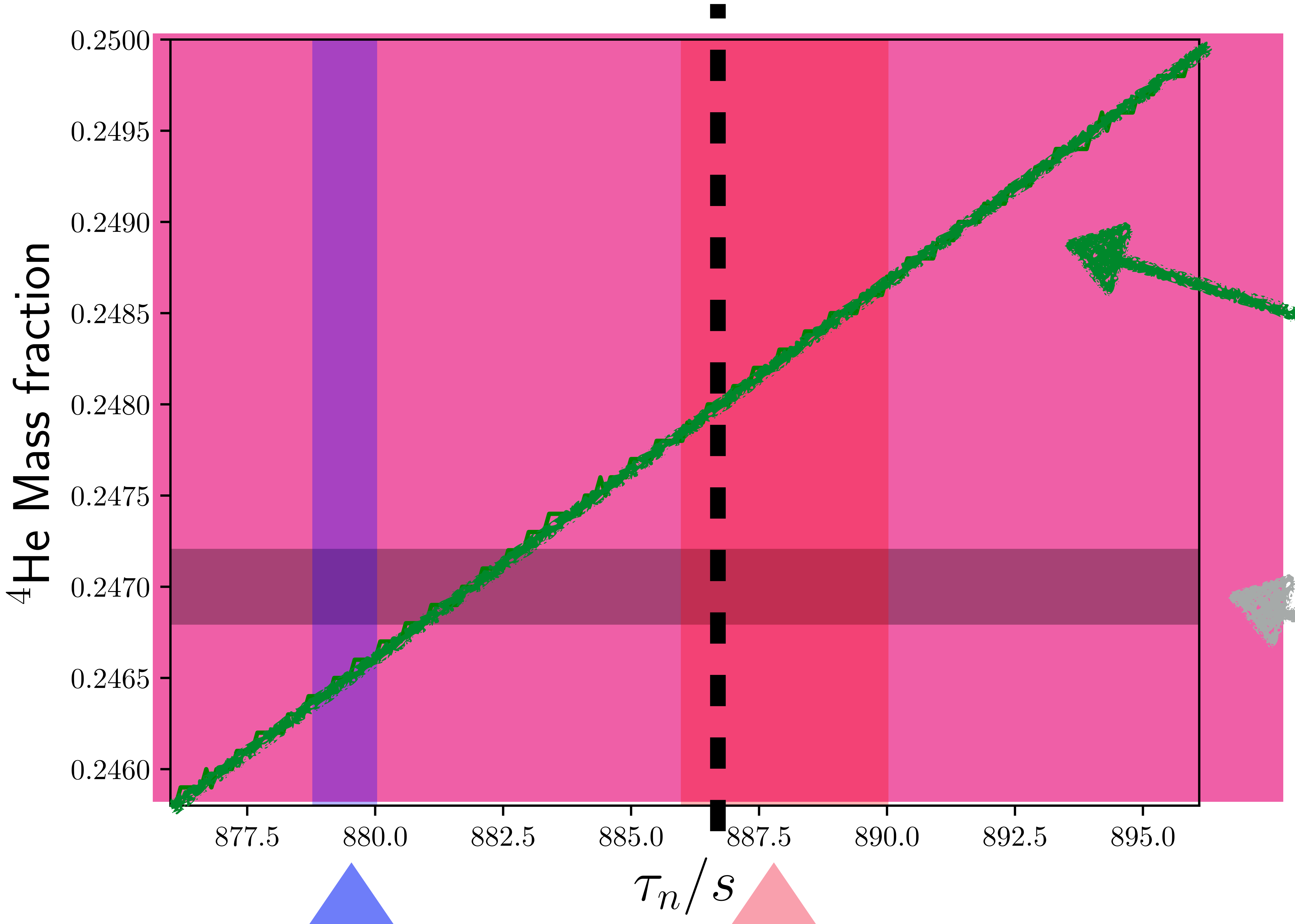
Can we reduce the uncertainty to a discriminating level?



Two different methods of measuring the neutron lifetime disagree at the 99% level
bottle and beam

Is one of the experiments wrong? Or is there new physics hiding here?

Why is understanding the neutron lifetime interesting?



$$\tau_n = \frac{(5172.0 \pm 1.1)\ \text{seconds}}{1+3g_A^2}$$

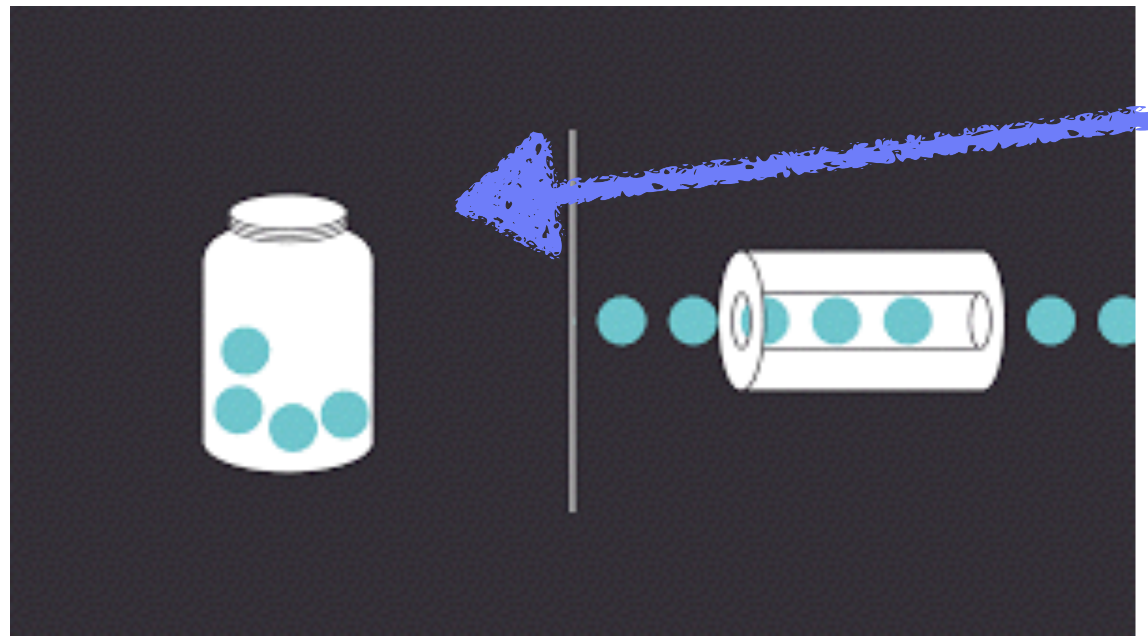
Czarnecki, Marciano, Sirlin

Predicted primordial ^{4}He mass fraction of the universe as a function of the neutron lifetime

Can we reduce the uncertainty to a discriminating level?

Yes!

With summit!



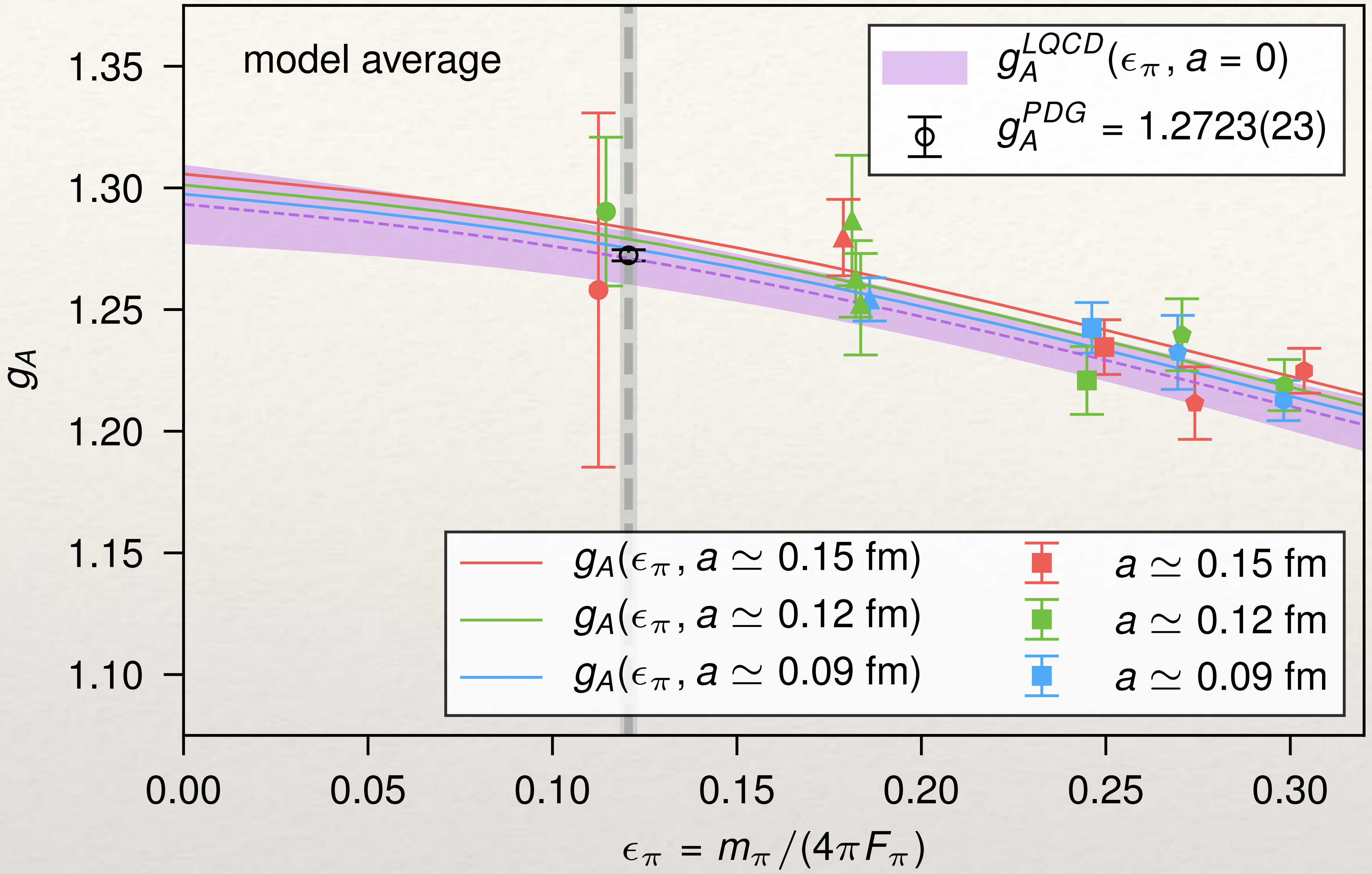
Two different methods of measuring the neutron lifetime disagree at the 99% level

bottle and beam

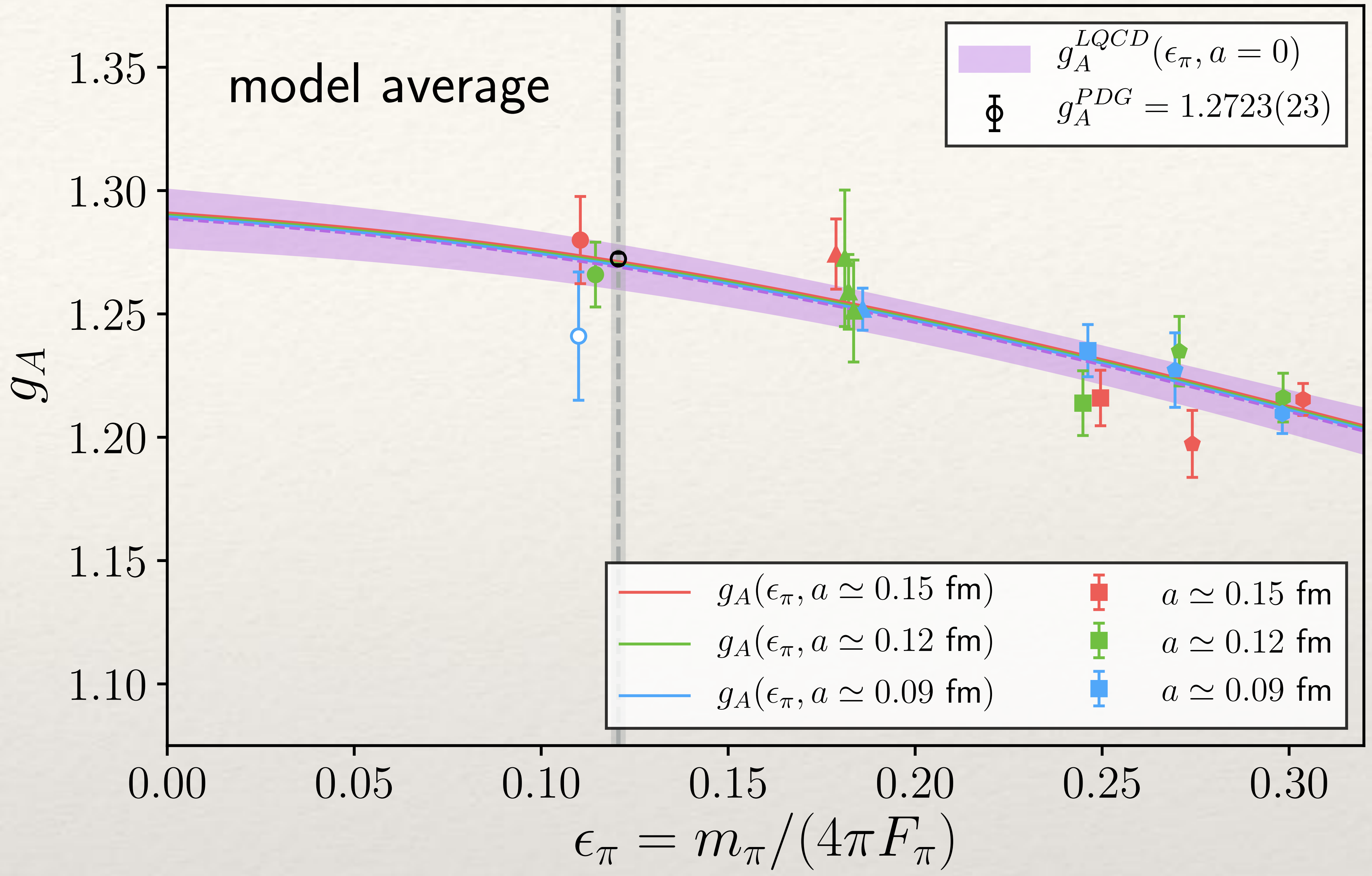
Is one of the experiments wrong? Or is there new physics hiding here?

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



Nature 558 (2018) no. 7708, 91-94

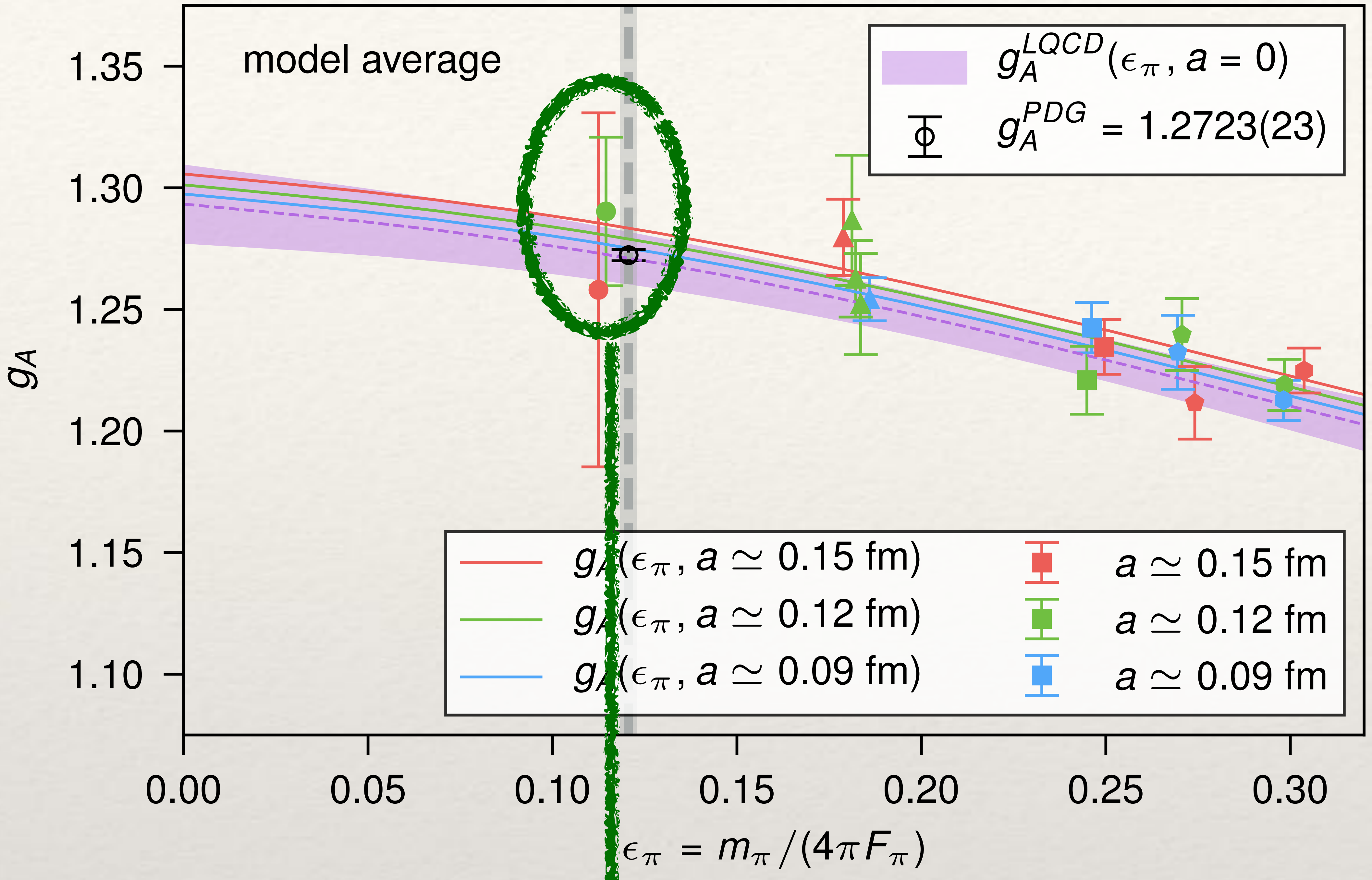


Sierra Early Science

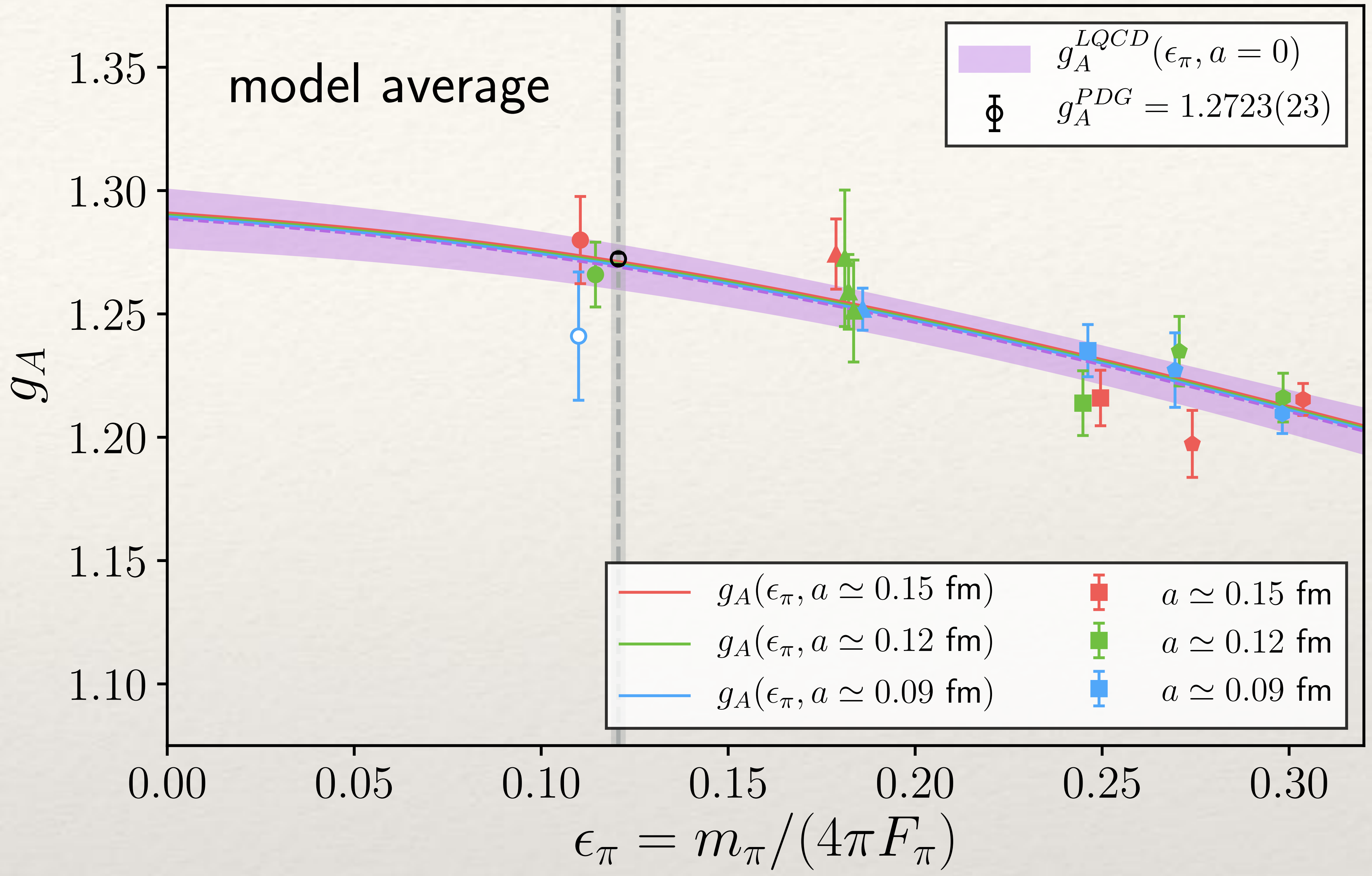
- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has **10x** more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



Nature 558 (2018) no. 7708, 91-94

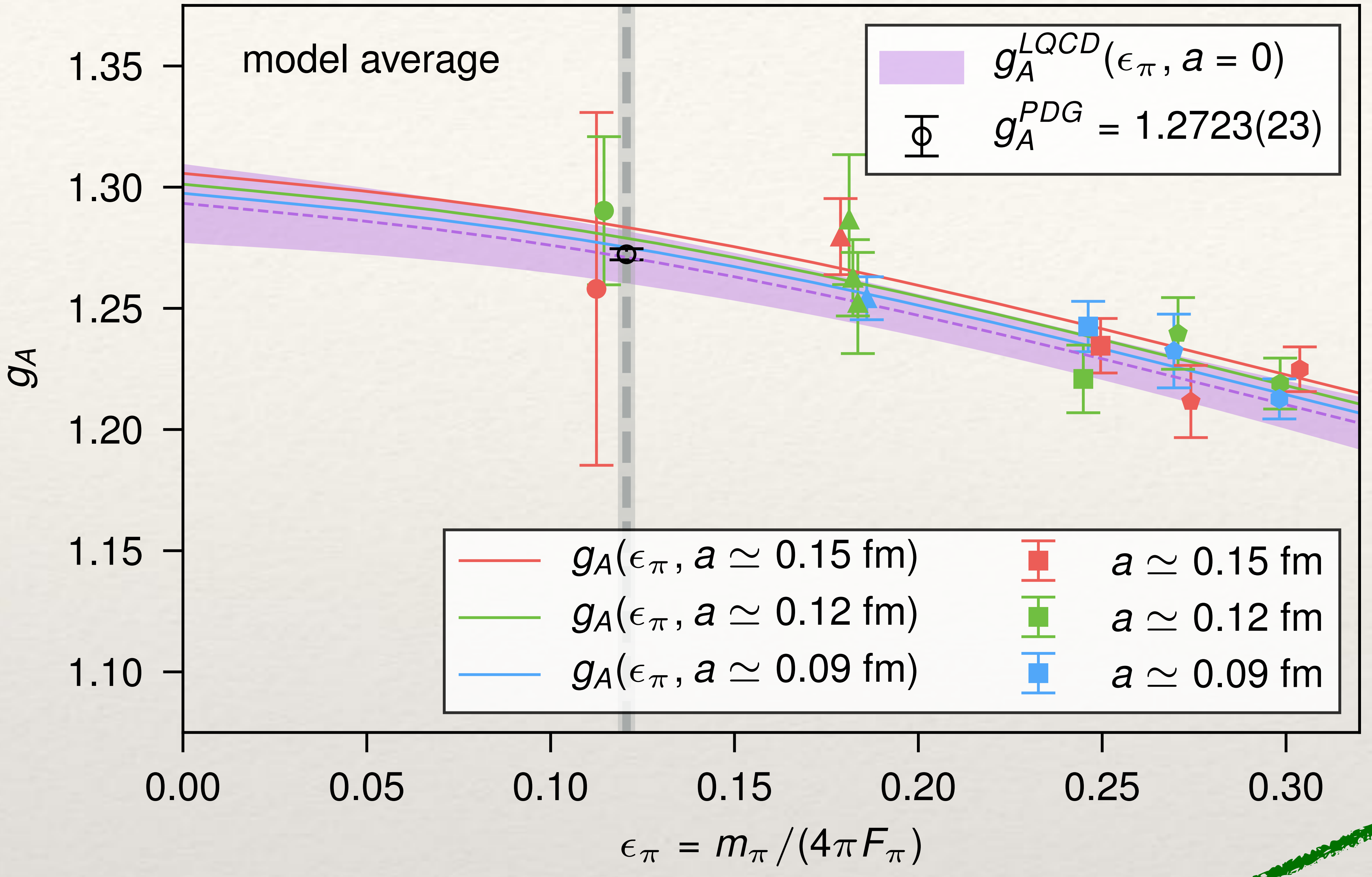


Sierra Early Science

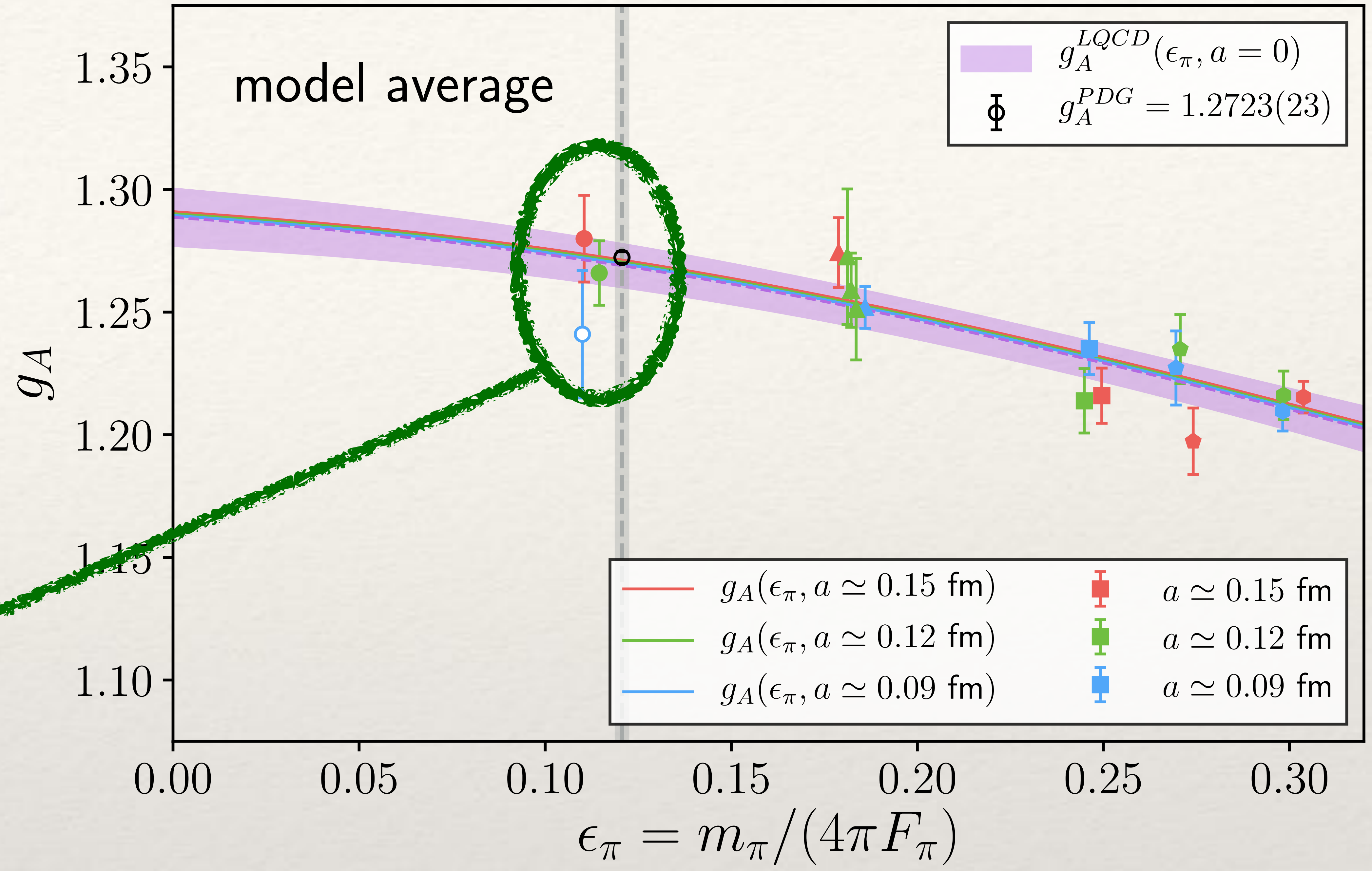
- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has 10x more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



Nature 558 (2018) no. 7708, 91-94

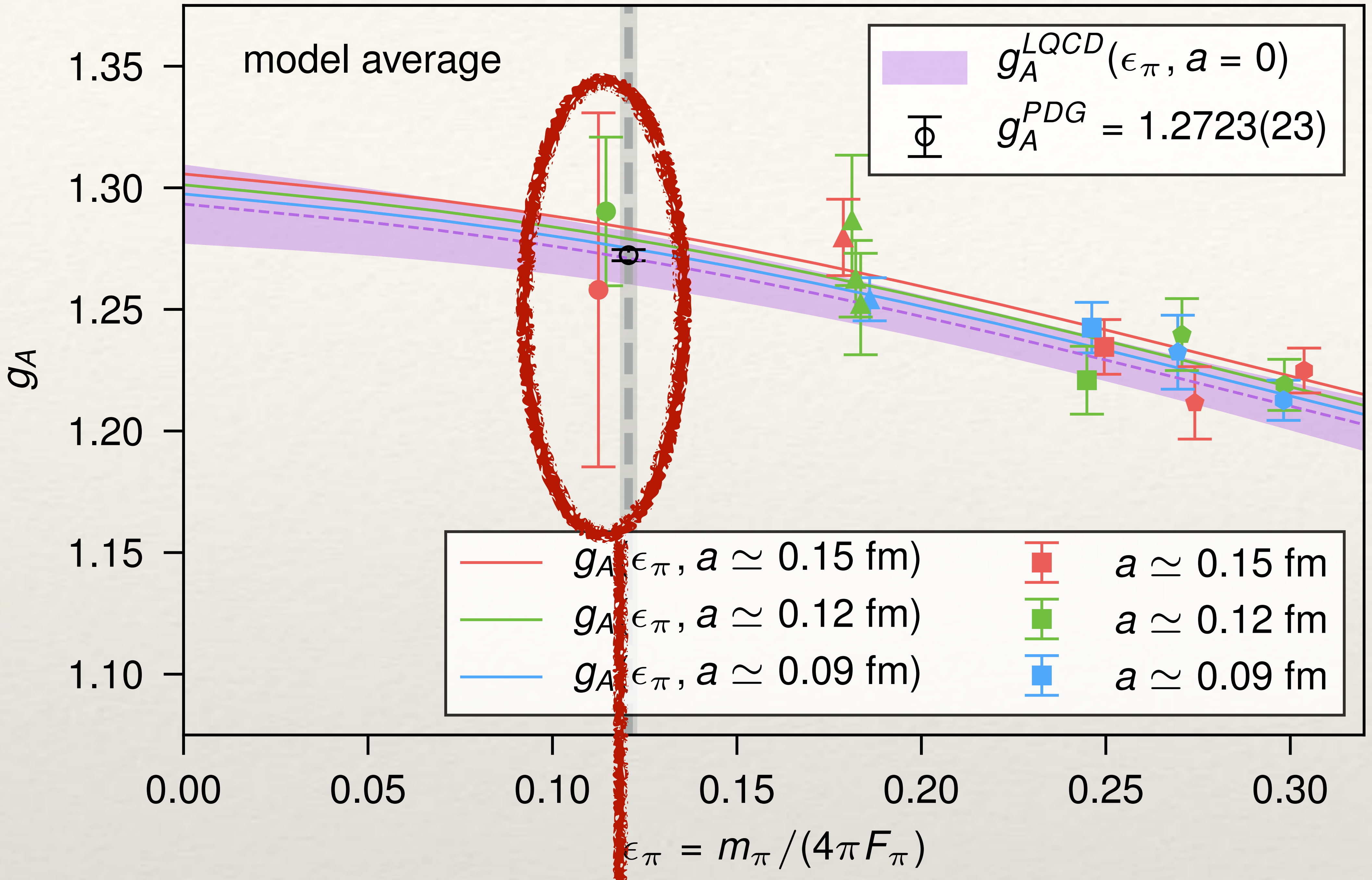


Sierra Early Science

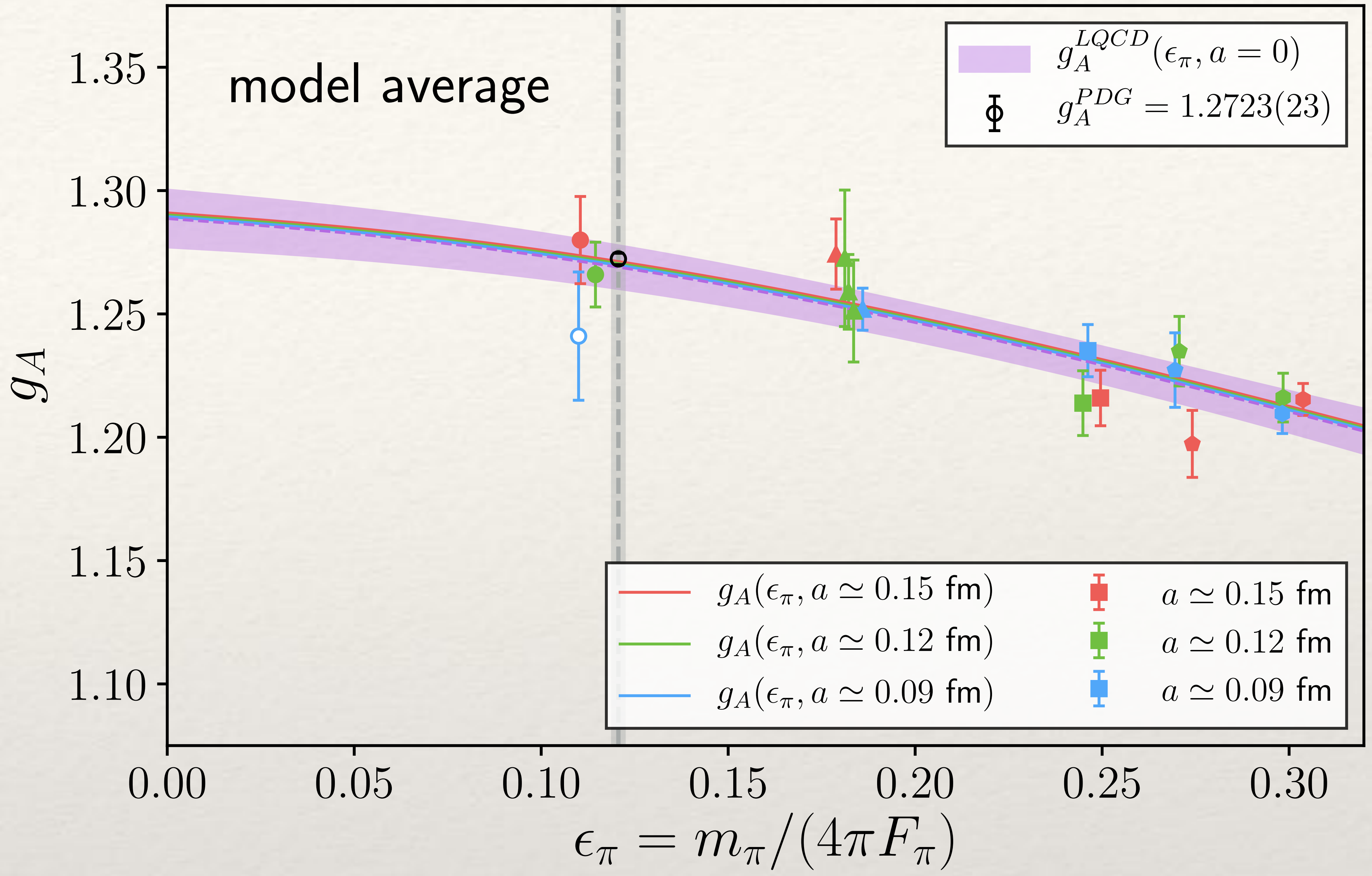
- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has **10x** more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



Nature 558 (2018) no. 7708, 91-94

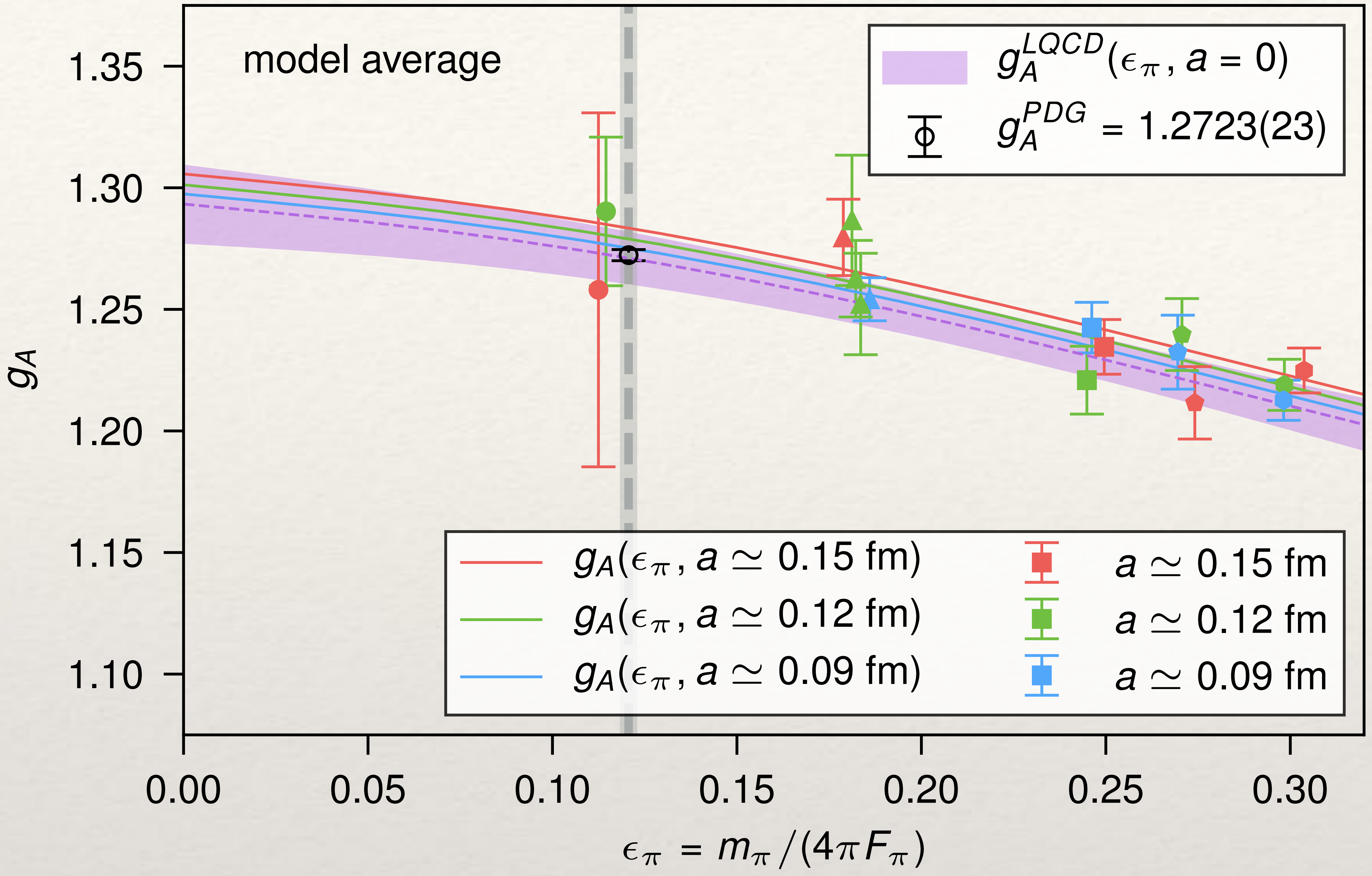


Sierra Early Science

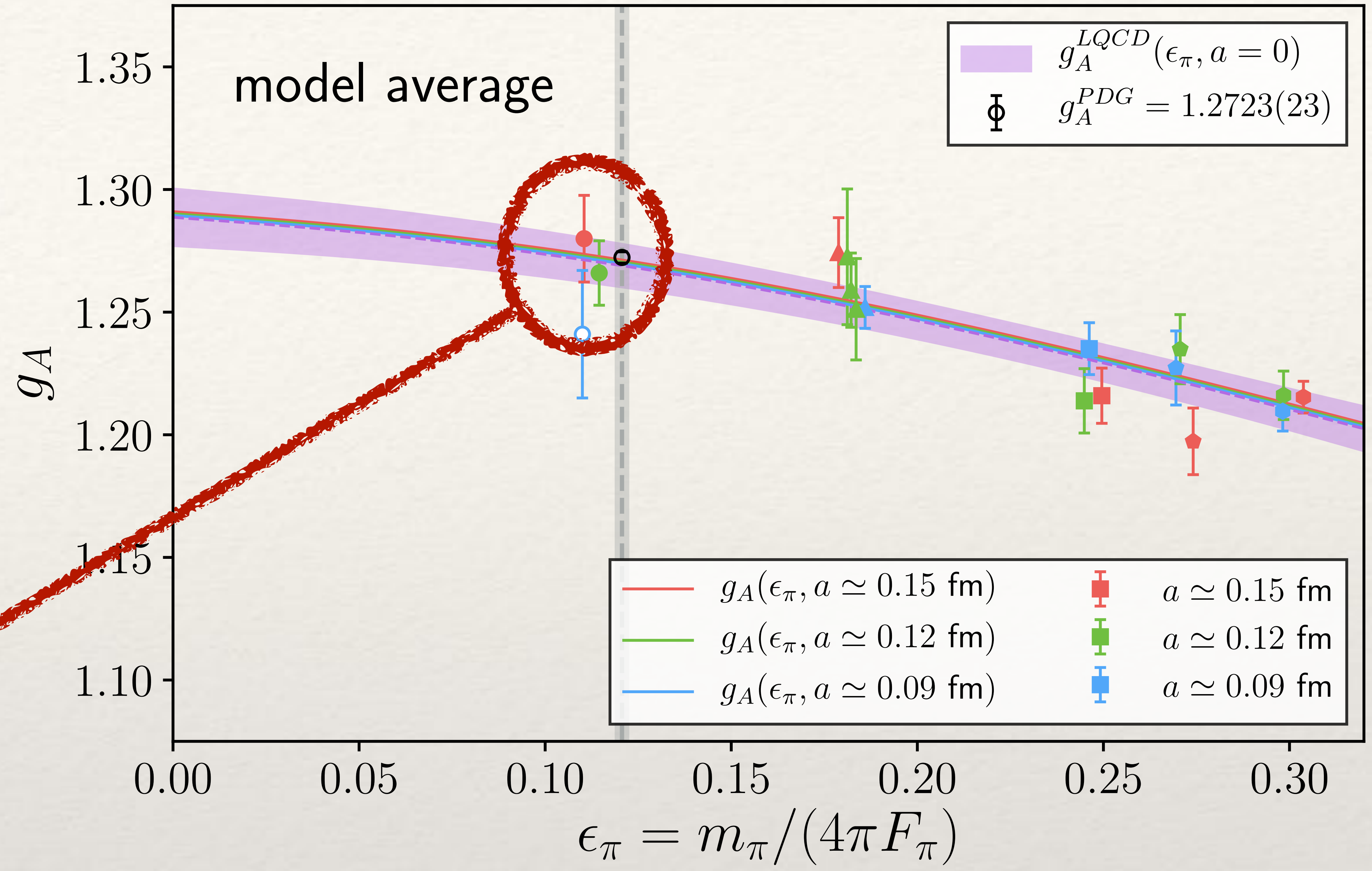
- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has **10x** more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



Nature 558 (2018) no. 7708, 91-94

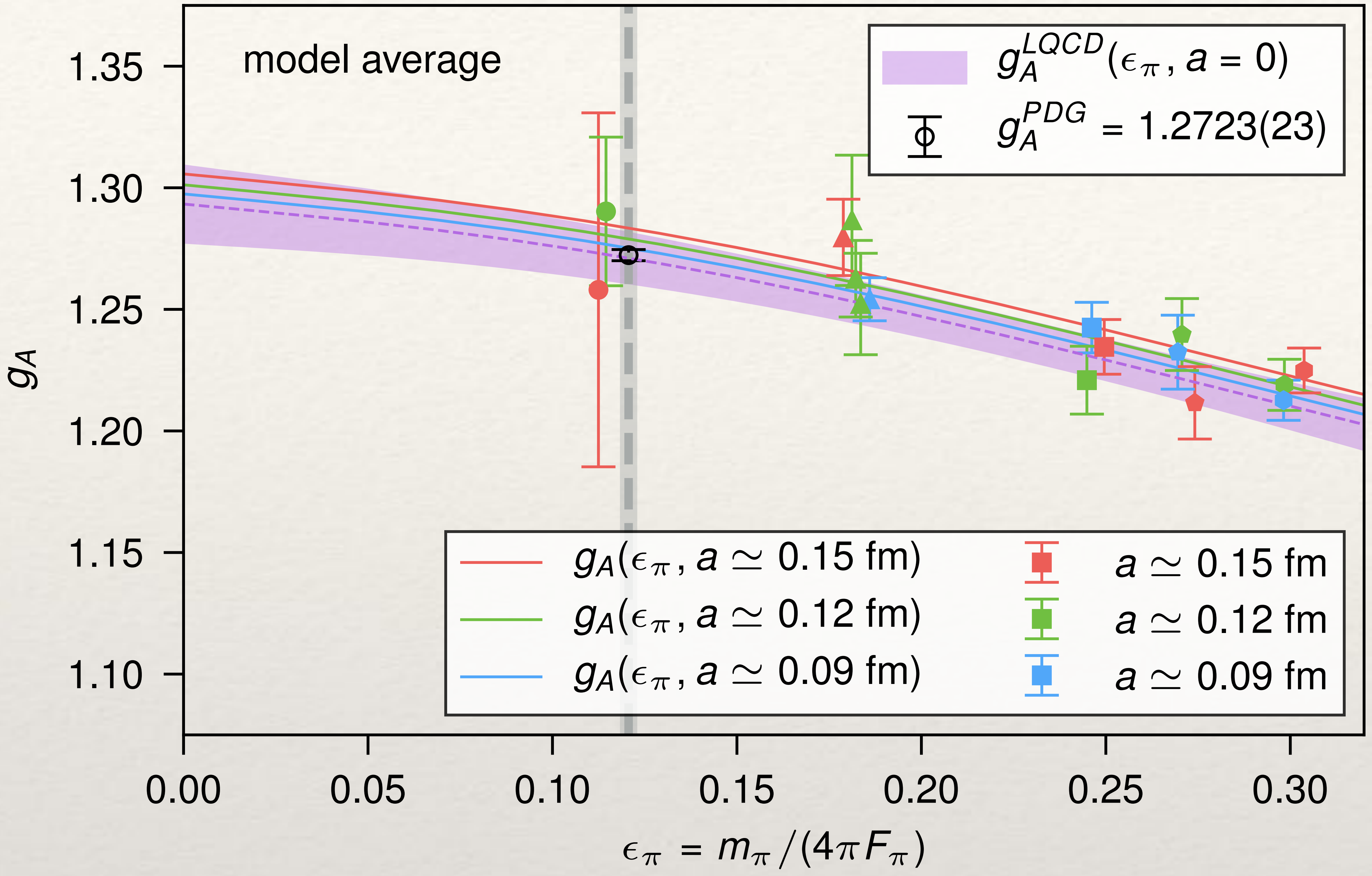


Sierra Early Science

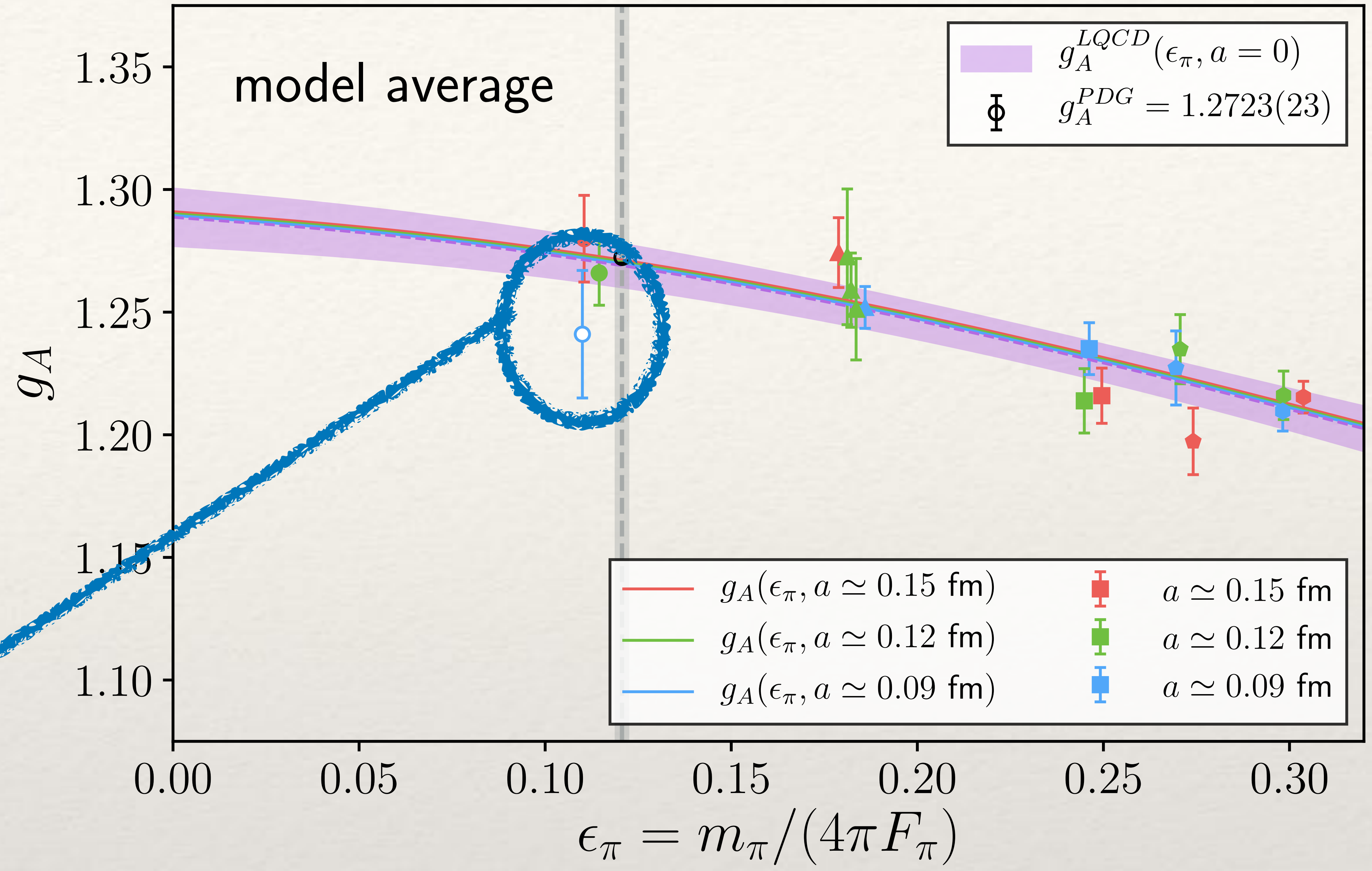
- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has **10x** more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

The Neutron Lifetime on Sierra Early Science

1 year on Titan (ORNL) + 2 years on GPU machines at LLNL



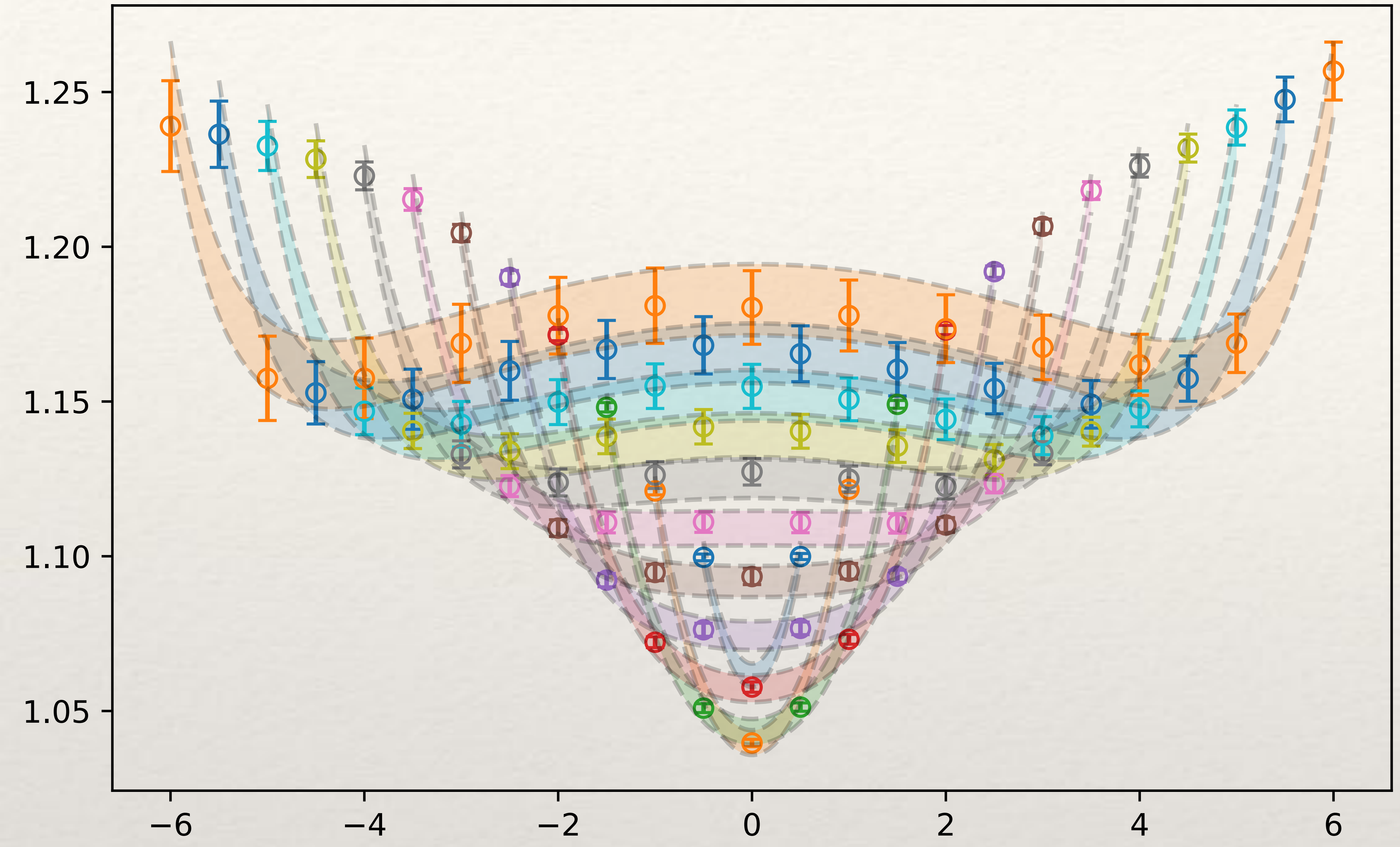
Nature 558 (2018) no. 7708, 91-94

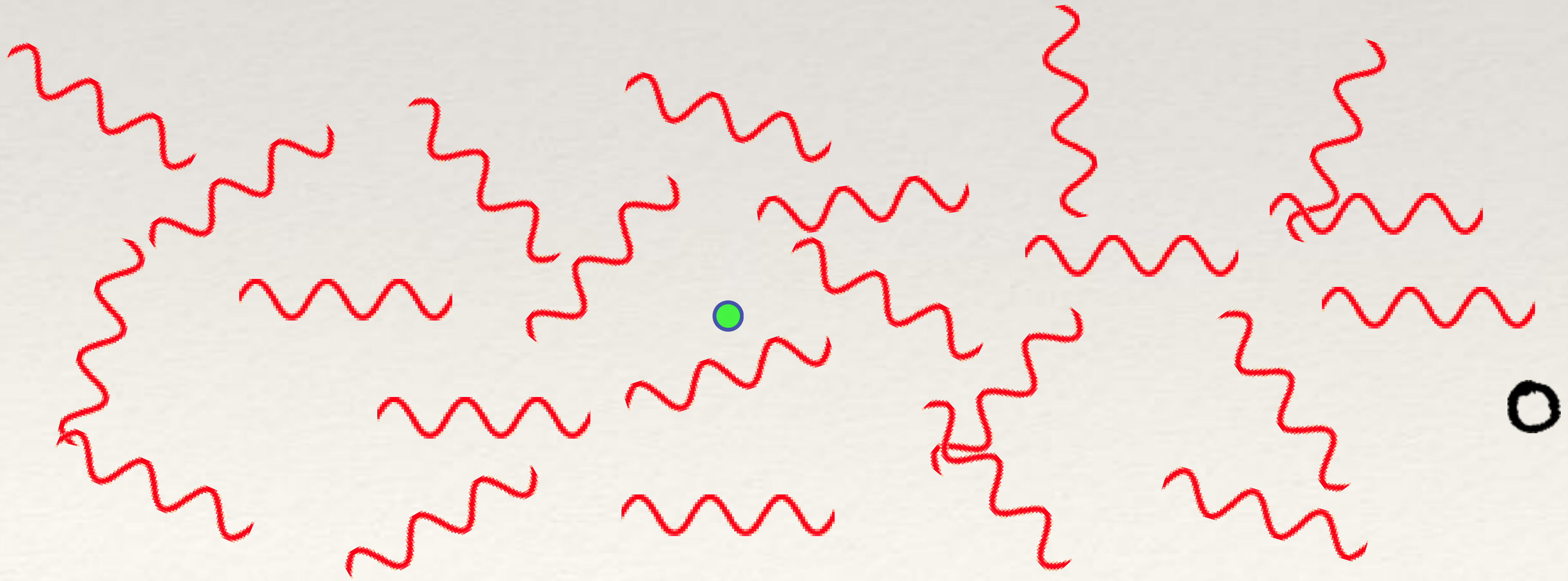


Sierra Early Science

- The vertical gray band denotes the physical value of pion mass - these points are significantly more expensive than the rest to compute - but the most valuable for the final predictions
- The green point in our publication cost as much computing time as all the other points combined
- The green point from Sierra has 10x more statistics than our publication
- The red point from our publication was not useful
- The red point from Sierra came from an entirely new calculation and is now very useful
- The blue point from Sierra was entirely unattainable from previous computers (it still needs more statistics to be useful)

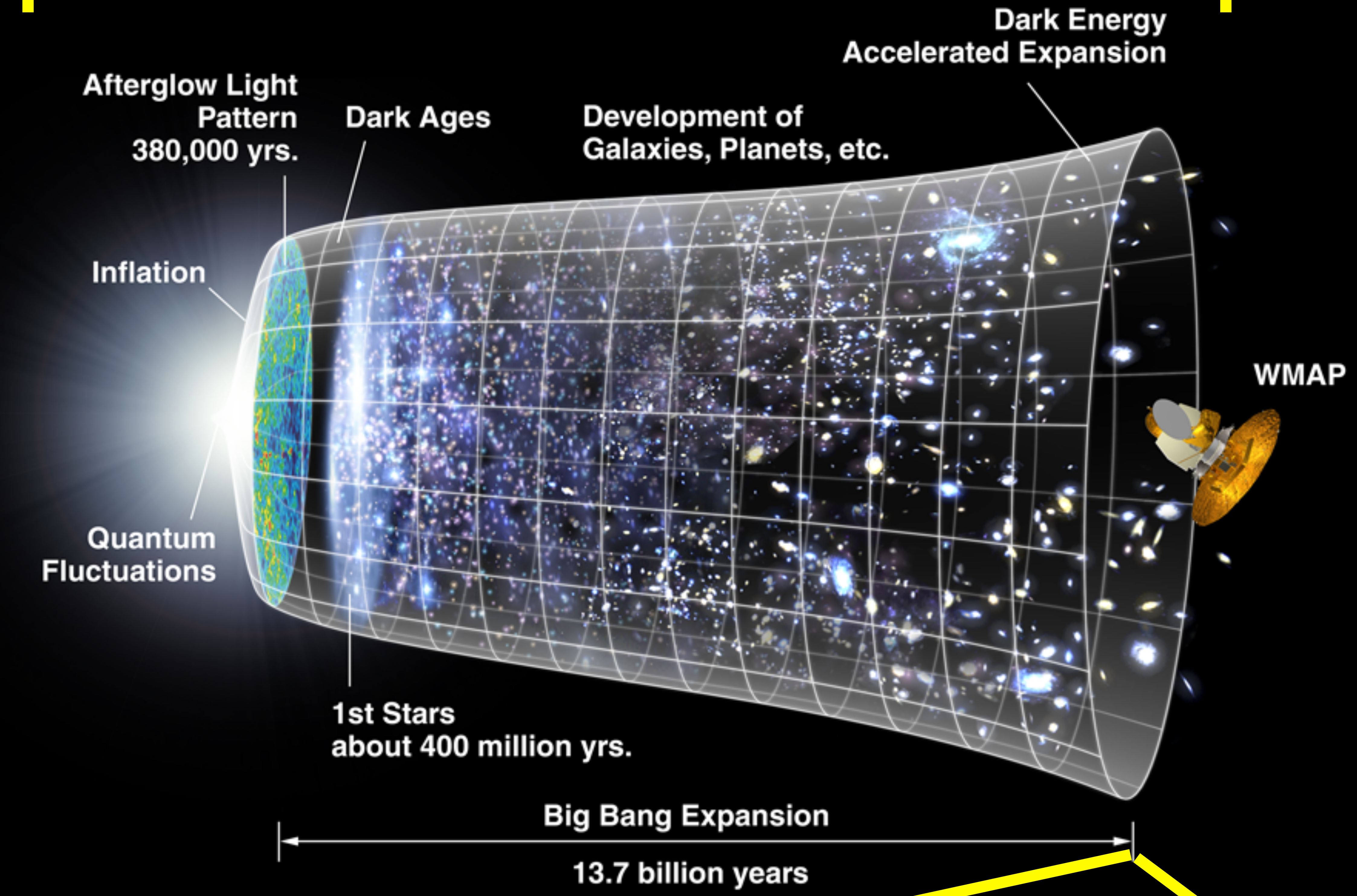
Sneak peak of Summit Results

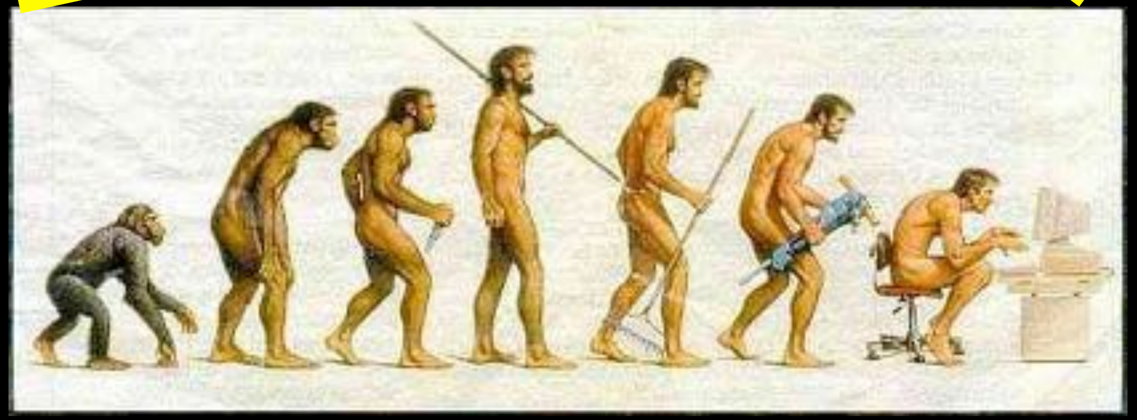




- We recently finished a “tuning” run on Summit
  - not yet processed into a nice result
- This data set is about **5x** larger than ever created for such a calculation (it also temporarily occupies about 1 PetaByte of scratch space)
- Our theoretical understanding (fit bands) is in remarkable agreement with the numerical results
  - This will provide us with an unprecedented ability extract the physics of interest
- The results we obtain this year will help in the quest to understand the origin of matter over anti-matter in the universe

Standard Model of Particle Physics





- This is a very exciting time for basic science research with high-performance-computing
- Titan (K20) brought us to the edge of making predictive contact with basic nuclear physics quantities relevant to our understanding of the universe
- Summit is **disruptively** faster than Titan
  - we are still learning to expand our vision of what can be accomplished
- Making full use of the new heterogeneous architecture requires
  - Optimized GPU libraries, QUDA
  - the development of sophisticated, and light-weight job managers
MetaQ, MPI_JM