Lattice QCD Computations in a Fixed Gauge

Attilio Cucchieri

(in collaboration with Tereza Mendes)

Instituto de Física de São Carlos – USP (Brazil)

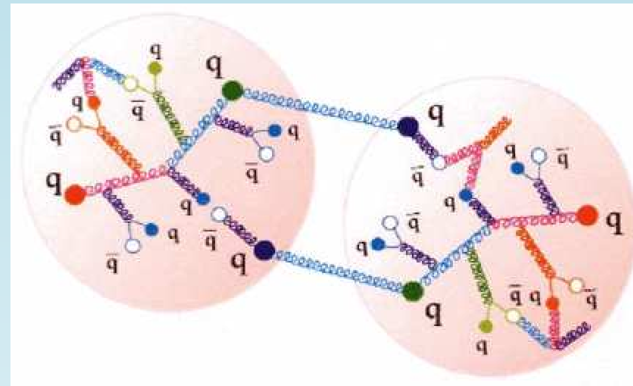
Color Confinement

Millennium Prize Problems by the Clay Mathematics Institute (US$1,000,000): Yang-Mills Existence and Mass Gap: Prove that for any compact simple gauge group G, a non-trivial quantum Yang-Mills theory exists on $\mathbb{R}^4$ and has a mass gap $\Delta > 0$.

Lattice simulations can solve QCD exactly (in discretized Euclidean space-time), allowing quantitative predictions for the physics of hadrons. But they can also help reveal the principles behind a central phenomenon of QCD: confinement. In fact, we can try to understand the QCD vacuum (the “battle for nonperturbative QCD”, E.V. Shuryak, *The QCD vacuum, hadrons and the superdense matter*) by using inputs from lattice simulations and by testing numerically the approximations introduced in analytic approaches (Dyson-Schwinger equations, Bethe-Salpeter equations, Pomeron dynamics, QCD-inspired models, etc).

Pathways to Confinement

- How does linearly rising potential (seen in lattice QCD) come about?
- Green's functions carry all information of a QFT's physical and mathematical structure.



- Confinement given by behavior at large distances (small momenta) $\Rightarrow$ nonperturbative study of IR propagators and vertices $\longrightarrow$ it requires to fix a gauge and very large lattice volumes.
- Gluon propagator (two-point function) as the most basic quantity of QCD. Proposal by Mandelstam (1979) linking linear potential to infrared behavior of gluon propagator as $1/p^4$.

Quantization and Gribov Copies

The invariance of the Lagrangian under local gauge transformations implies that, given a configuration $\{A(x), \psi_f(x)\}$, there are infinitely many gauge-equivalent configurations $\{A^g(x), \psi_f^g(x)\}$ (gauge orbits). In the path integral approach we integrate over all possible configurations

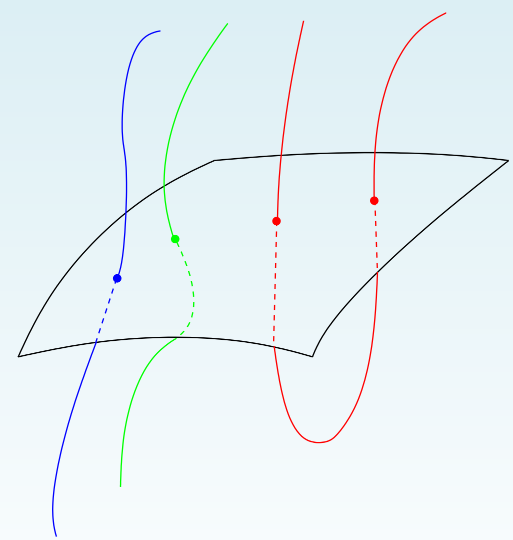
$$Z = \int DA \, \exp\left[- \int d^4x \, \mathcal{L}(x) \right] .$$

There is an infinite factor coming from gauge invariance: $\int DA = \int D\overline{A}^g \, Dg$ and $\int Dg = \infty$.

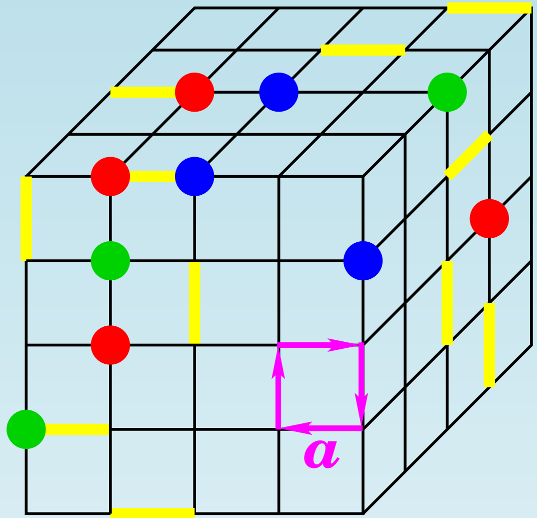
To solve this problem we can choose a representative $\overline{A}$ on each gauge orbit (gauge fixing) using a gauge-fixing condition $f(\overline{A}) = 0$. The change of variable $A \to \overline{A}$ introduces a Jacobian in the measure.

Question: does the gauge-fixing condition select one and only one representative on each gauge orbit?

Answer: in general this is not true (Gribov copies).



Lattice QCD



In the lattice formulation the theory is rewritten in discrete space-time (subtle$\rightarrow$ must preserve the gauge symmetry) and the fields may be thought of as elements of a thermodynamical system at a fixed temperature.

- The only first-principles, nonperturbative way to study QCD with no uncontrolled approximations.
- Lattice is an UV cutoff which goes to zero in the continuum (= physical) limit: rigorous formulation of quantum field theory.
- Application of statistical-mechanics techniques — such as Monte Carlo simulation, study of critical phenomena — to quantum field theories.
- Grand Challenge Problem (report to the US President, 2005, "Computational Science: ensuring America's competitiveness").

Gauge-Related Lattice Features

- Gauge action written in terms of oriented plaquettes formed by the link variables $U_\mu(x) \in SU(N_c)$.
- Closed loops are gauge invariant under gauge transformations $U_\mu(x) \to g(x) U_\mu(x) g^\dagger(x + \hat{e}_\mu)$, where $g(x) \in SU(N_c)$.
- Integration volume is finite: no need for gauge-fixing.
- When gauge fixing procedure is incorporated in the simulation, no need to consider ghost fields or Faddeev-Popov matrix.
- Lattice momenta given by $\hat{p}_\mu = 2 \sin(\pi k_\mu / N)$ with $k_\mu = 0, 1, \ldots, N/2 \;\Leftrightarrow\; p_{min} \sim 2\pi/(a\,N) = 2\pi/L$, $p_{max} = 4/a$ in physical units.

3-Step Code

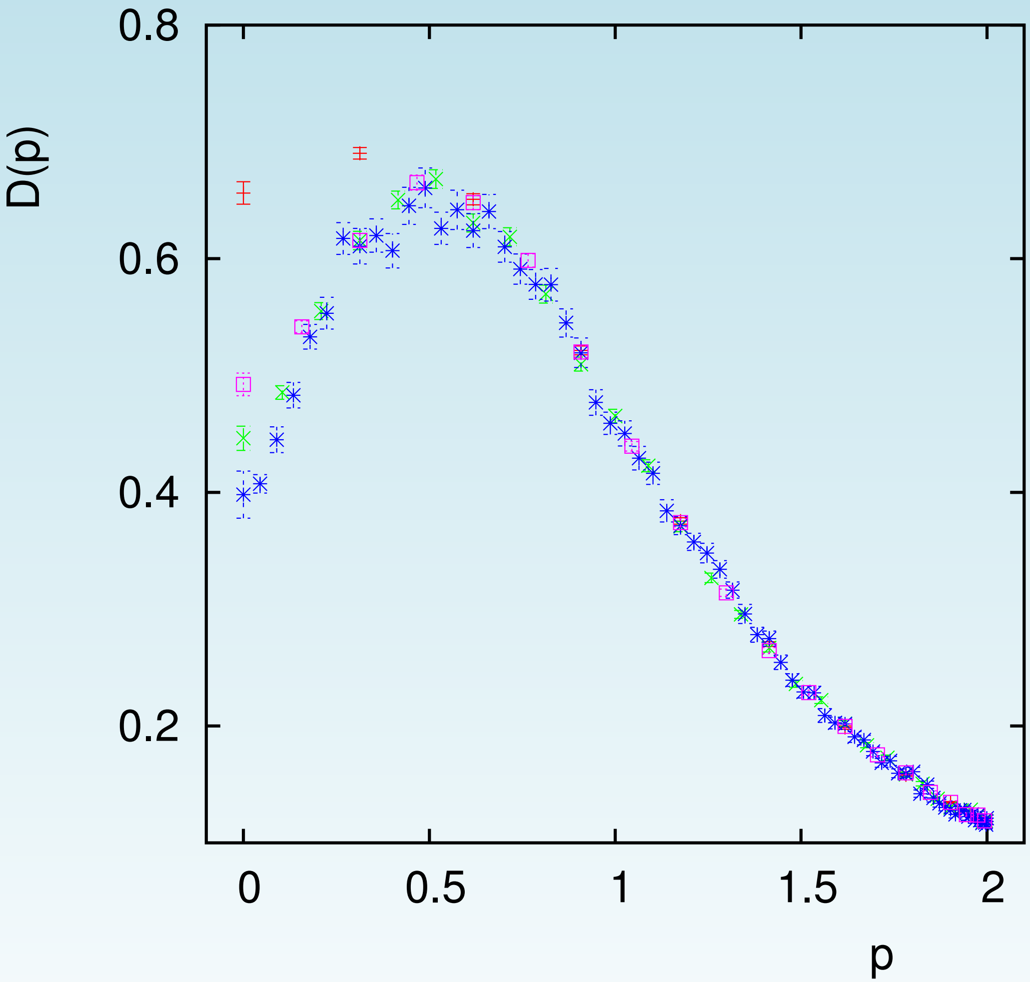
```
main() {
/* set parameters: beta, number of configurations NC,
                   number of thermalization sweeps NT */
     read_parameters();
/* {U} is the link configuration */
     set_initial_configuration(U);
/* cycle over NC configurations */
     for (int c=0; c < NC; c++) {
          thermalize(U,beta,NT);
          gauge_fix(U,g);
          evaluate_propagators(U[g]);
     }
}
```

Algorithms: Heat-Bath and Micro-canonical (thermalization), overrelaxation and simulated annealing (gauge fixing), conjugate gradient and Fourier transform (propagators, etc.).

Parallelization

- We need a parallelized code in order to simulate at very large lattice volumes V.
- Communication is required in each of the three steps.
- Each node gets a contiguous block of $v = V/N$ lattice sites (local lattice).
- Communication is required only for sites on the boundary of the local lattice.
- 3D/4D simulations $\rightarrow$ high granularity due to the surface/volume effect.

Gluon Propagator: Infinite-Volume Limit



Gluon propagator as a function of the lattice momentum p for lattice volumes $V = 20^3$, 40^3, 60^3 and 140^3 at $\beta = 3.0$ About 100 days using a 13 Gflops PC cluster (2003).

Lattice Landau Gauge (I)

In the continuum: $\partial_\mu A_\mu(x) = 0$. On the lattice the Landau gauge is imposed by minimizing the functional

$$\mathcal{E}[U;g] = -\sum_{x,\mu} Tr\, U_\mu^{(g)}(x) \,,$$

where $g(x) \in SU(N_c)$ and $U_\mu^{(g)}(x) = g(x)\, U_\mu(x)\, g^\dagger(x+\hat{e}_\mu)$ is the lattice gauge transformation.

By considering the relations $U_\mu(x) = e^{i\, A_\mu(x)}$ and $g(x) = e^{i\,\tau\,\gamma(x)}$, we can expand $\mathcal{E}[U;g]$ (for small τ):

$$\mathcal{E}[U;g] = \mathcal{E}[U;\mathbb{1}] + \tau\, \mathcal{E}^{'}[U;\mathbb{1}](b,x)\,\gamma^b(x)$$

$$+ \frac{\tau^2}{2}\,\gamma^b(x)\,\mathcal{E}^{''}[U;\mathbb{1}](b,x;c,y)\,\gamma^c(y) + \ldots \,,$$

where $\mathcal{E}^{''}[U;\mathbb{1}](b,x;c,y) = \mathcal{M}(b,x;c,y)[A]$ is a lattice discretization of the Faddeev-Popov operator $-D\cdot\partial$.

Lattice Landau Gauge (II)

At any local minimum (stationary solution)

$$\mathcal{E}'(0) = 0 \quad \forall \quad \{\gamma^b(x)\} \quad \Rightarrow \quad [(\nabla \cdot A)(x)]^b = 0 \quad \forall \quad x,b \ ,$$

where

$$A_\mu(\vec{x}) = \frac{1}{2\,i}\left[U_\mu(\vec{x}) - U_\mu^\dagger(\vec{x})\right]_{\text{traceless}}$$

is the gauge field and

$$\left(\nabla \cdot A^b\right)(\vec{x}) = \sum_{\mu=1}^{d} A_\mu^b(\vec{x}) - A_\mu^b(\vec{x} - \hat{e}_\mu)$$

is the (minimal) Landau gauge condition on the lattice.

Numerical Gauge Fixing: Convergence

Consider an iterative minimization algorithm for $\mathcal{E}[U;g]$, starting from $g(x) = \mathbb{1}$, all x. Check the convergence of the algorithm

$$\Delta\mathcal{E} \;=\; \mathcal{E}[U;g](t) \,-\, \mathcal{E}[U;g](t-1)$$

$$(\nabla A)^2 \;=\; \frac{1}{(N_c^2-1)\,V} \sum_{b=1}^{N_c^2-1} \sum_{\vec{x}\in\Lambda_x} \Big[\left(\nabla\cdot A^b\right)(\vec{x})\Big]^2$$

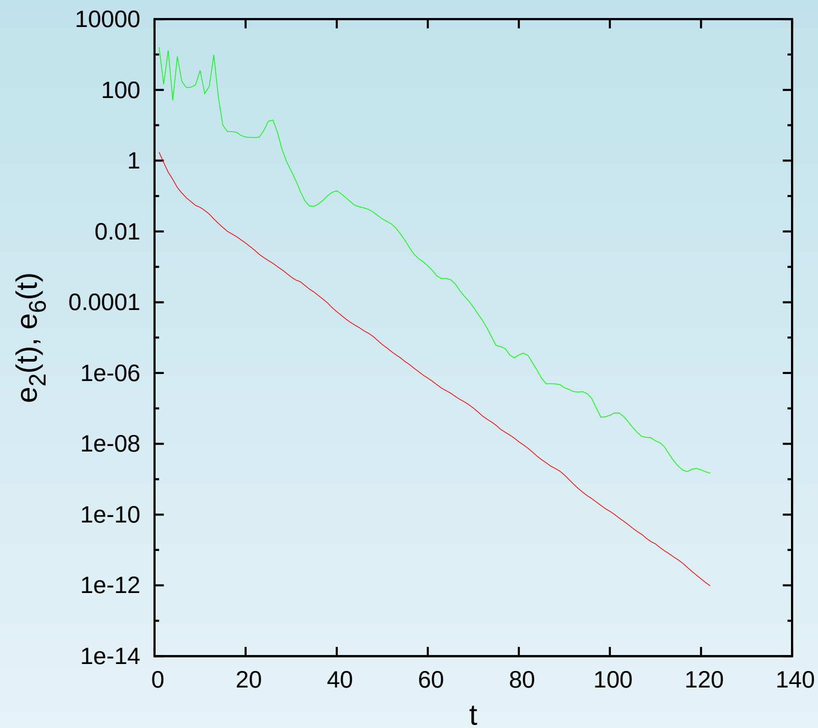
$$\Sigma_Q \;=\; \frac{1}{d\,(N_c^2-1)\,N} \sum_{\mu=1}^{d} \sum_{b=1}^{N_c^2-1} \sum_{x_\mu=1}^{N} \Big[Q_\mu^b(x_\mu) - \widehat{Q}_\mu^b\Big]^2 \Big/ \Big[\widehat{Q}_\mu^b\Big]^2 \,,$$

where

$$Q_\nu(x_\nu) \equiv \sum_{\mu\neq\nu} \sum_{x_\mu} A_\nu(\boldsymbol{x})$$

should be constant.

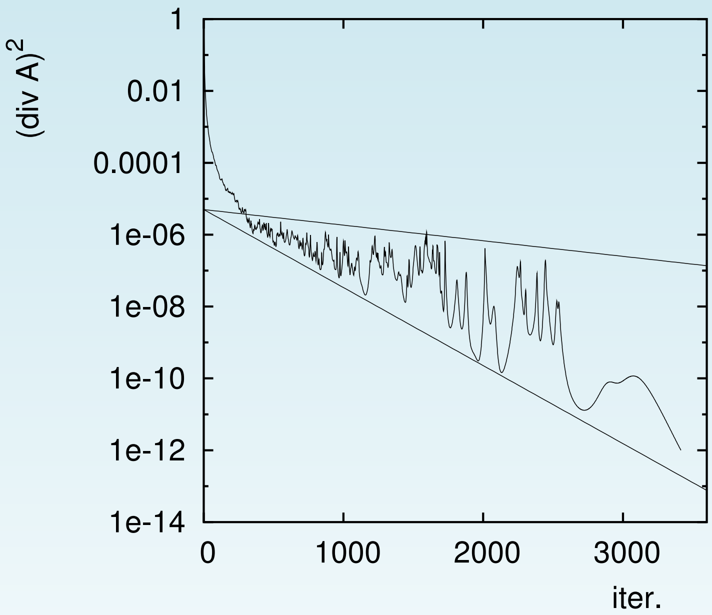
Convergence of the Algorithms



Plot of $e_2(t) = (\nabla A)^2(t)$ and of $e_6(t) = \Sigma_Q(t)$ as a function of the gauge-fixing iterations t. Here $V = 16^2$ and $\beta = 8.0$ (corresponding to $a \approx 0.196$ fermi) for the Cornell method.

Convergence and Landscape of $\mathcal{E}[U;g]$

For very large lattice side (N) the quantities that should go to zero are not only affected by the long-wavelength modes (critical-slowing down), but they show very large fluctuations, due to the complicated landscape of $\mathcal{E}[U;g]$.



Plot of $(\nabla A)^2$ as a function of the gauge-fixing iterations t. Here $V = 200^3$ and $\beta = 4.2$ (corresponding to $a \approx 0.174$ fermi) for the stochastic overrelaxation algorithm.

This result does not depend on the choice of the gauge-fixing algorithm.

Algorithms

Single-site update: $\mathcal{E}[U; g(y)] = constant - Tr[g(y)\, h(y)]$ where

$$h(y) = \sum_{\mu=1}^{d} \Big[U_\mu(y) g^\dagger(y + \hat{e}_\mu) + U_\mu^\dagger(y - \hat{e}_\mu) g^\dagger(y - \hat{e}_\mu) \Big]$$

is the single-site effective magnetic field.

In the SU(2) case: with the parametrization $U = u_0 \mathbb{1} + i\vec{u}\cdot\vec{\sigma}$, where $\vec{\sigma}$ are Pauli matrices $u_0 \in \mathrm{R}$, $\vec{u} \in \mathrm{R}^3$ and $u_0^2 + \vec{u}\cdot\vec{u} = 1$, $h(y)$ is proportional to an $SU(2)$ matrix:

$$h(y) = \sqrt{\det h(y)}\; \widetilde{h}(y) \; .$$

Define:

$$\mathcal{N}(y) = \sqrt{\det h(y)}$$

and

$$\mathcal{T}(y) = Tr[g(y)\, h(y)]$$

Update: $g^{(new)}(y) = R^{(update)}(y)\, g(y)$

Los Alamos Method: $R^{(update)}(y) = \widetilde{h}^{\dagger}(y)\, g^{\dagger}(y)$.

Overrelaxation: $R^{(update)}(y) = [\widetilde{h}^{\dagger}(y)\, g^{\dagger}(y)]^{\omega}$ with $\omega \in (1,2)$.

Stochastic Overrelaxation:

$$R^{(update)}(y) \;=\; \begin{cases} \left[\widetilde{h}^{\dagger}(y)\, g^{\dagger}(y)\right]^2 & \text{with probability } p \\[2em] \widetilde{h}^{\dagger}(y)\, g^{\dagger}(y) & \text{with probability } 1-p \,. \end{cases}$$

Cornell Method: $R^{(update)}(y) \propto \left[\mathbb{1} - \alpha \left(\nabla \cdot A^{(g)}\right)(y)\right]$.

Fourier Acceleration:

$$R^{(update)}(y) \;\propto\; \mathbb{1} - \widehat{F}^{-1} \left\{ \frac{p^2_{max}}{p^2(k)}\, \widehat{F} \left[\alpha \left(\nabla \cdot A^{(g)} \right) \right] \right\} (y) \,,$$

where $\widehat{F}$ is the Fourier transform and $p^2(k)$ is the magnitude square of the lattice momenta. Note: we are inverting the usual lattice Laplacian $\Rightarrow$ we can use other methods, e.g. conjugate gradient and multi-grid, with reduced accuracy.

Critical Slowing-Down

We expect to observe

$$e_i(t) \propto \exp\left(-t/\tau_i\right) \quad \text{and} \quad \tau_i = c_i N_i^z$$

for the three quantities considered, i.e. $\Delta\mathcal{E}$, $(\nabla A)^2$ and Σ_Q. The z_i are the dynamic critical exponents. The quantities e_i have the same τ for each given algorithm.

In the $2d$ case:

algorithm	N_{min}	z	c	χ^2/DF
Los Alamos	12	1.99 ± 0.04	0.22 ± 0.03	0.89 (49 %)
Cornell	16	0.83 ± 0.09	0.6 ± 0.2	0.24 (91 %)
overrelax.	16	1.12 ± 0.07	0.22 ± 0.05	0.27 (90 %)
stoch. overr.	12	1.09 ± 0.05	0.30 ± 0.05	0.27 (93 %)
Fourier acc.	8	0.04 ± 0.06	2.9 ± 0.6	0.38 (69 %)

Weighted least-squares fit for $\tau = c\,N^z$ at $N^2/\beta = 32$ [$N/\xi \approx 7$ and $a \in (0.049, 0.392)$ fermi] using lattice sizes $N \geq N_{min}$.

The Case $\beta = \infty$

The same results for the dynamic critical exponents can be obtained analytically (!) by considering the case $\beta = \infty$, i.e. $U_\mu(x) = \mathbb{1}$. Then,

$$
\begin{aligned}
\mathcal{E}[\mathbb{1}; g] &= -\sum_{x,\mu} Tr\; g(x)\; g^\dagger(x + \hat{e}_\mu) \\
&= constant - \sum_{x,\mu} \frac{Tr}{2}\; [\, g(x)\; - \; g(x + \hat{e}_\mu)\,]\; [\, g(x)\; - \; g(x + \hat{e}_\mu)\,]^\dagger \; .
\end{aligned}
$$

When we are close to the minimum we can write

$$
g(x) \, = \, \mathbb{1} - i\, \epsilon\, \vec{\sigma} \cdot \vec{f}(x) \, + \, \mathcal{O}(\epsilon^2)
$$

and we obtain $g(x)\, g^\dagger(x) \, = \, \mathbb{1} \, + \, \mathcal{O}(\epsilon^2)$ so that

$$
\mathcal{E}[\mathbb{1}; g] \, \approx \, -\sum_{x,\mu} \frac{\epsilon^2\, Tr}{2} \, \| \, \vec{f}(\boldsymbol{x}) \, - \, \vec{f}(\boldsymbol{x} + a\boldsymbol{e}_\mu) \, \|^2
$$

and we can use standard analytic results (for z, τ and the tuning) in the minimization of this quadratic form.

The Case $\beta = 0$

These results are confirmed also in the limit case $\beta = 0$, i.e. link variables $U_\mu(x)$ are completely random, except for the Cornell method ($z = 2$) and the Fourier method ($z = 1$). In the SU(2) case we have:

Overrelaxation:

$$g^{(new)}(y) \propto (1 - \omega)\, g(y) + \omega\, \widetilde{h}^\dagger(y) \; .$$

Cornell Method (for large t):

$$g^{(Corn)}(y) \propto [1 - \alpha\, \mathcal{N}(y)]\, g(y) + \alpha\, \mathcal{N}(y)\, \widetilde{h}^\dagger(y) \; .$$

If $\alpha\, \mathcal{N}(y)$ is larger than $2 \Rightarrow$ maximizing(!).

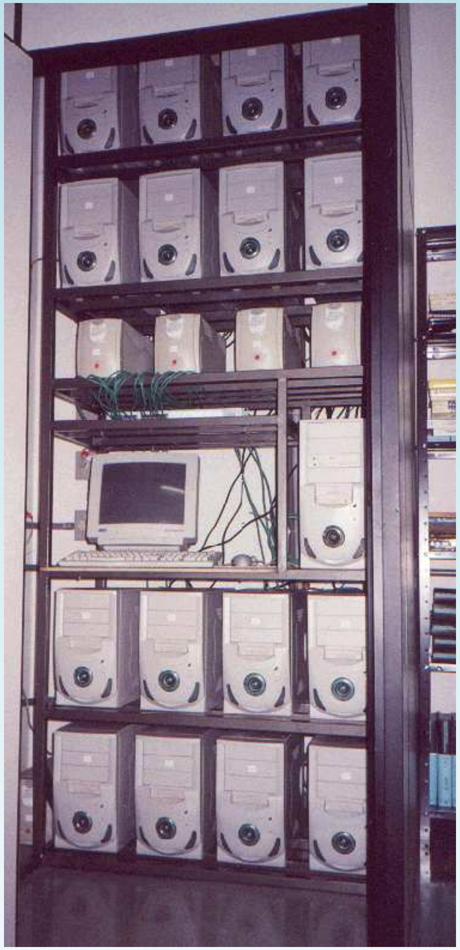
Solution: new definition for Cornell with $z = 1$:

$$g^{(Corn)}(y) \propto (1 - \tilde{\alpha}(y))\, g(y) + \tilde{\alpha}(y)\, \widetilde{h}^\dagger(y) \; ,$$

where $\tilde{\alpha}(y) \equiv \min\left(\, \alpha\, \mathcal{N}(y), \, 2 \,\right)$.

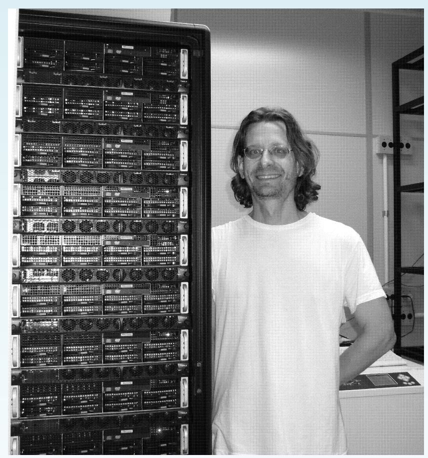
Not easy to improve the Fourier method and difficult tuning.

Old Computer Facilities



2001

2007



2010

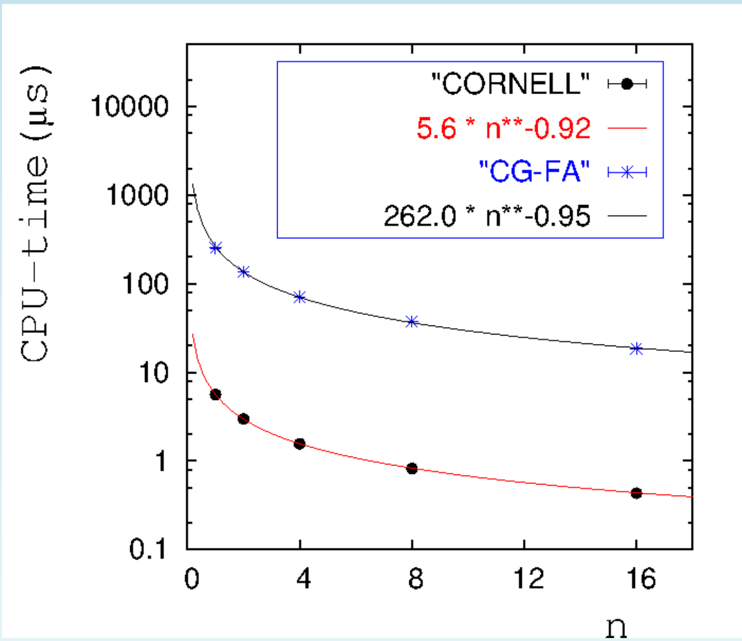
2012

From a Pentium III PC-cluster, to the IBM Mamute, an Intel Xeon cluster and the Blue Gene P.

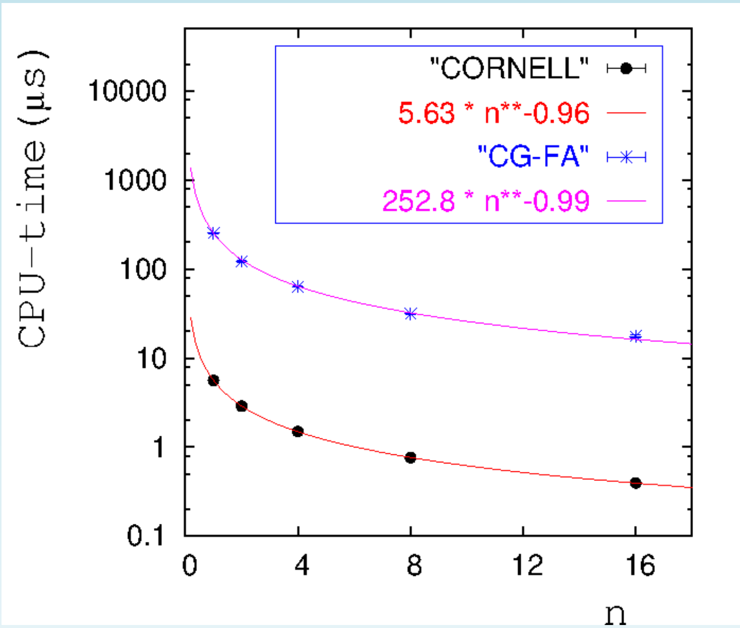
☺ The New Toy: BG/Q

The Rice BlueGene/P was upgraded to BlueGene/Q around March 2015.

CPU-Time vs. Number of Nodes n



Fixed Volume ($V = 64^3$).

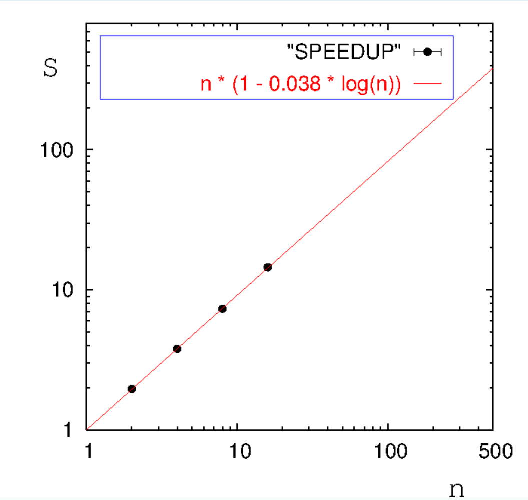


Variable Volume ($V = 64^3 n$).

Speedup

	Fixed volume		Variable volume	
	$S = T_1/T_n$	$E = S/n$	$S = T_1/T_n$	$E = S/n$
$1 \to 2$	1.87(2)	0.937(8)	1.96(2)	0.981(8)
$1 \to 4$	3.61(1)	0.904(3)	3.79(1)	0.948(3)
$1 \to 8$	6.91(2)	0.863(2)	7.31(8)	0.91(1)
$1 \to 16$	13.38(6)	0.836(3)	14.5(3)	0.88(1)

Average speedup S and efficiency E at fixed and at variable volume.



Weak and Strong Scaling at BG/P

V	Nodes	Site up.	HB/link	CG it.
8^4	4	13.6	17.3	0.0022
16^4	64	13.9	17.6	0.0026
32^4	1024	14.0	18.7	0.0030
16^4	32	13.2	16.6	0.0035
16^4	64	13.9	17.6	0.0026
16^4	128	15.0	19.5	0.0025
16^4	256	16.7	22.7	0.0018
16^4	512	18.1	26.4	0.0014

Weak (with 3 different lattice volumes) and strong (with 5 different volumes) scaling: time (in microseconds) for 2 different updates of local variables and the time (in seconds) for one conjugate gradient iteration. Link and site variables are SU(2) matrices.

Weak and Strong Scaling at BG/Q

V	Nodes	HB	Micro	Gfix	GluonProp	CG
$64^2 \times 32^2$	32	494.9	54.7	0.0044	0.041	0.0081
$64^3 \times 32$	64	496.3	62.1	0.0049	0.041	0.0088
64^4	128	496.8	59.2	0.0047	0.050	0.0084
$64^3 \times 128$	256	499.4	63.0	0.0050	0.041	0.0090
$64^2 \times 128^2$	512	499.7	56.4	0.0046	0.042	0.0083
64^4	128	496.8	59.2	0.0047	0.0050	0.0084
64^4	256	256.3	37.9	0.0029	0.0028	0.0055
64^4	512	134.6	27.3	0.0020	0.0018	0.0040
64^4	1024	74.4	22.5	0.0016	0.0012	0.0035
64^4	512	2943.6	218.5	0.0171	0.0179	0.0239

Weak (with 5 different lattice volumes) and strong (with 4 different volumes) scaling: time (in seconds) for 3 different updates of local variables and for the evaluation of the gluon propagator and the time (in seconds) for one conjugate gradient iteration. Link and site variables are SU(2) matrices. The last row is for the BG/P.

Conclusions

Numerical gauge-fixing for minimal Landau gauge is an interesting case where it is possible to

- study the algorithms numerically and analytically
- understand how an algorithm reduces critical-slowing down
- find prescriptions for tuning the algorithms

and, therefore, test possible new ideas.

Challenge: how to deal with Gribov copies.

MERCI!