Conjugate Directions in Landau and Coulomb Lattice Gauge Fixing

Renwick James Hudspith

Department of Physics and Astronomy, York University, Toronto, Ontario, M3J 1P3, Canada
RBC and UKQCD collaborations

June, 2014

Overview

- Matching lattice Green's functions to continuum perturbation theory, non perturbative renormalisation in Landau gauge.
- Coulomb gauge fixed wall source propagators, static potential measurements.
- Two schools

1. Los Alamos
   - Overrelaxation
2. Cornell
   - Fourier acceleration

Links, Fields and Notation

$$U_\mu\left(x + a\frac{\hat{\mu}}{2}\right) = e^{iag_0 A_\mu\left(x + a\frac{\hat{\mu}}{2}\right)}.$$

Generally for an SU(3) matrix $U = e^{iA} = f_0 + f_1 A + f_2 A^2$.

$$A \approx \frac{1}{2i}\left\{\left[U - U^\dagger\right] - \frac{1}{3}\mathrm{Tr}\left[U - U^\dagger\right]\cdot I_{3\times 3}\right\}.$$

$$A = \frac{f_2^* U - f_2 U^\dagger - 2i\Im(f_0 f_2^*)\cdot I_{3\times 3}}{2i\Im(f_1 f_2^*)}.$$

Links, Fields and Notation ... Cont'd

$$a\Delta_\mu A_\mu(x) = \sum_\mu \left(A_\mu \left(x + a\frac{\hat{\mu}}{2} \right) - A_\mu \left(x - a\frac{\hat{\mu}}{2} \right) \right).$$

$$\Theta^{(n)} = \frac{1}{N_c V} \sum_x \mathrm{Tr} \left[\left(a\Delta_\mu a g_0 A_\mu^{(n)}(x) \right)^2 \right].$$

Fixing on a lattice ... Steepest descent (SD)

Satisfying the Landau condition is the same as minimising a cost functional F(U).

$$F(U) = \frac{1}{N_d N_c V} \sum_{x,\mu} \mathrm{Tr}\left[\left(a g_0 A_\mu\left(x + a\frac{\hat{\mu}}{2}\right)\right)^2\right].$$

Fixing on a lattice ... Steepest descent (SD)

Satisfying the Landau condition is the same as minimising a cost functional F(U).

$$F(U) \approx 1 - \frac{1}{N_d N_c V} \sum_{x,\mu} \Re \left(\mathrm{Tr} \left[U_\mu \left(x + a\frac{\hat{\mu}}{2} \right) \right] \right).$$

Fixing on a lattice ... Steepest descent (SD)

Satisfying the Landau condition is the same as minimising a cost functional F(U).

$$F(U) \approx 1 - \frac{1}{N_d N_c V} \sum_{x,\mu} \Re \left(\mathrm{Tr} \left[U_\mu \left(x + a\frac{\hat{\mu}}{2} \right) \right] \right).$$

$$x^{(n+1)} = x^{(n)} - \alpha f'(x^{(n)}).$$

$0 < \alpha < 1$ is a small tuning parameter.

Fixing on a lattice ... Steepest descent (SD)

Satisfying the Landau condition is the same as minimising a cost functional F(U).

$$F(U) \approx 1 - \frac{1}{N_d N_c V} \sum_{x,\mu} \Re \left(\mathrm{Tr} \left[U_\mu \left(x + a\frac{\hat{\mu}}{2} \right) \right] \right).$$

$$g(x) = e^{-i\alpha a \Delta_\mu a g_0 A_\mu^{(n)}(x)},$$

$$U_\mu^{(n+1)} \left(x + a\frac{\hat{\mu}}{2} \right) = g(x) U_\mu^{(n)} \left(x + a\frac{\hat{\mu}}{2} \right) g(x + a\hat{\mu})^\dagger.$$

$0 < \alpha < 1$ is a small tuning parameter.

Fourier Acceleration (FA)

Davies et al. (1988) pointed out that the SD suffers from critical slowing down. Ameliorated by Fourier acceleration

$$g(x) = e^{-i\alpha a \Delta_\mu a g_0 A_\mu^{(n)}(x)}.$$

Fourier Acceleration (FA)

Davies et al. (1988) pointed out that the SD suffers from critical slowing down. Ameliorated by Fourier acceleration

$$g(x) = e^{-i\alpha\,\tilde{F}\,\frac{p^2_{\mathrm{Max}}}{Vp^2}\,F\,a\Delta_\mu a g_0 A^{(n)}_\mu(x)}.$$

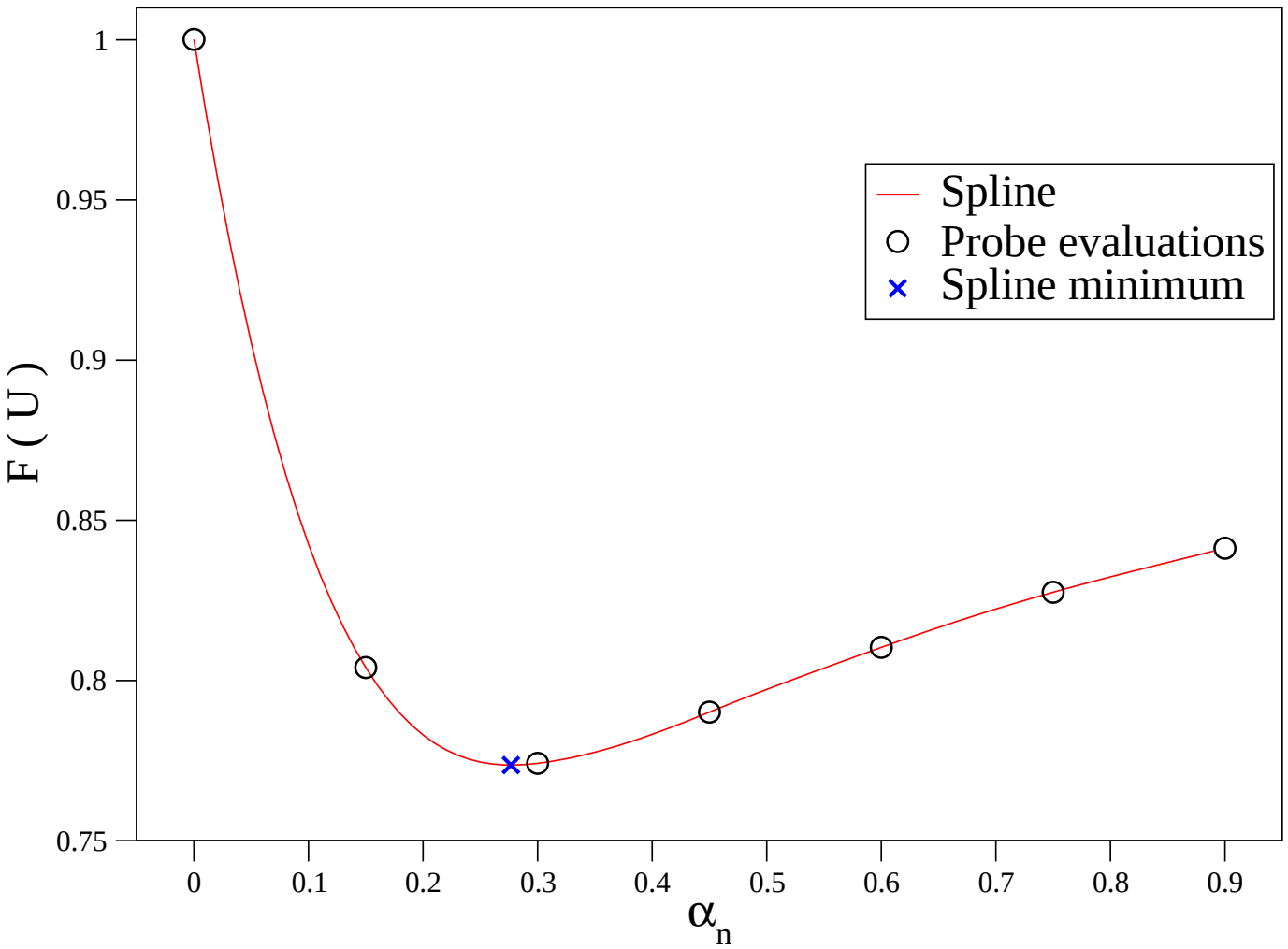
$$p^2 = 2\left(N_d - \sum_\mu \cos\left(\frac{2\pi n_\mu}{L_\mu}\right)\right).$$

We call this the FASD method.

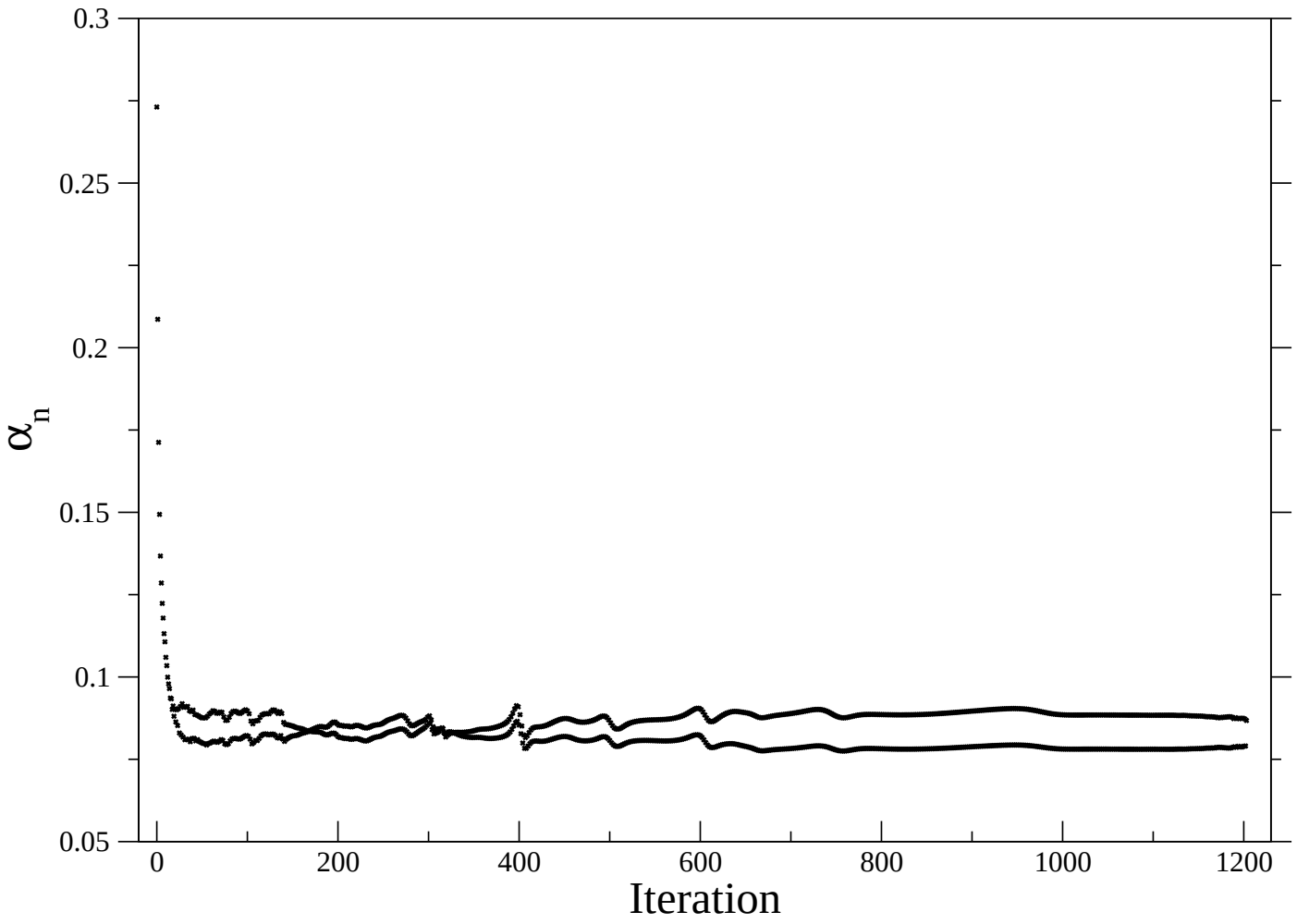
Optimal SD is far from optimal

- What about $\alpha = \alpha_n$, tuned at each step?
- Evaluate "N" probe α_n''s and choose the one that minimises $F(U)$ best
- Use cubic spline interpolation and exact solve for minimum
- Cost is $\approx N\times$ that of fixed α.

Splines



Spline history



The general Non-linear CG procedure

Compute the gradient direction $f'(x_0)$
Perform a line search for α_0 s.t $\min(f(x_0 - \alpha_0 f'(x_0)))$
Perform the update $x_1 = x_0 - \alpha_0 f'(x_0)$
Set $s_0 = -f'(x_0)$
$n = 1$
while $|f'(x_n)|^2 >$ Tolerance **do**
    Compute the gradient $f'(x_n)$
    Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n)-f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$
    Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$
    Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
    Update $x_{n+1} = x_n + \alpha_n s_n$
    $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$

Perform a line search for α_0 s.t $\min(f(x_0 - \alpha_0 f'(x_0)))$

Perform the update $x_1 = x_0 - \alpha_0 f'(x_0)$

Set $s_0 = -f'(x_0)$

$n = 1$

while $|f'(x_n)|^2 >$ Tolerance **do**

  Compute the gradient $f'(x_n)$

  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n)-f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$

  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$

  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$

  Update $x_{n+1} = x_n + \alpha_n s_n$

  $n = n + 1$

end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$

$g(x) \leftarrow e^{-i\alpha_0\Gamma_0(x)}$

Perform the update $x_1 = x_0 - \alpha_0 f'(x_0)$

Set $s_0 = -f'(x_0)$

$n = 1$

while $|f'(x_n)|^2 >$ Tolerance **do**

  Compute the gradient $f'(x_n)$

  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n)-f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$

  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$

  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$

  Update $x_{n+1} = x_n + \alpha_n s_n$

  $n = n + 1$

end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$

$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$

$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$

Set $s_0 = -f'(x_0)$

$n = 1$

while $|f'(x_n)|^2 >$ Tolerance **do**

  Compute the gradient $f'(x_n)$

  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n) - f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$

  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$

  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$

  Update $x_{n+1} = x_n + \alpha_n s_n$

  $n = n + 1$

end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $|f'(x_n)|^2 >$ Tolerance **do**
  Compute the gradient $f'(x_n)$
  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n) - f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$
  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$
  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} > \text{Tolerance}$ **do**
  Compute the gradient $f'(x_n)$
  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n) - f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$
  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$
  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} >$ Tolerance **do**
  $\Gamma_n(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(n)}(x)$
  Compute $\beta_n = \max\left[0, \frac{f'(x_n)^T(f'(x_n) - f'(x_{n-1}))}{f'(x_{n-1})^T f'(x_{n-1})}\right]$
  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$
  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} >$ Tolerance **do**
  $\Gamma_n(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(n)}(x)$
  $\beta_n = \max\left[0, \frac{\sum_x \mathrm{Tr}[\Gamma_n(x)(\Gamma_n(x) - \Gamma_{n-1}(x))]}{\sum_x \mathrm{Tr}[\Gamma_{n-1}(x)^2]}\right]$
  Conjugate direction $s_n = -f'(x_n) + \beta_n s_{n-1}$
  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} >$ Tolerance **do**
  $\Gamma_n(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(n)}(x)$
  $\beta_n = \max\left[0, \frac{\sum_x \mathrm{Tr}[\Gamma_n(x)(\Gamma_n(x) - \Gamma_{n-1}(x))]}{\sum_x \mathrm{Tr}[\Gamma_{n-1}(x)^2]}\right]$
  $s_n(x) \leftarrow \Gamma_n(x) + \beta_n s_{n-1}(x)$
  Perform a line search for α_n s.t $\min(f(x_n + \alpha_n s_n))$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} >$ Tolerance **do**
  $\Gamma_n(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(n)}(x)$
  $\beta_n = \max\left[0, \frac{\sum_x \mathrm{Tr}[\Gamma_n(x)(\Gamma_n(x) - \Gamma_{n-1}(x))]}{\sum_x \mathrm{Tr}[\Gamma_{n-1}(x)^2]}\right]$
  $s_n(x) \leftarrow \Gamma_n(x) + \beta_n s_{n-1}(x)$
  $\alpha_n = \min(F(U); \alpha' s_n)$
  Update $x_{n+1} = x_n + \alpha_n s_n$
  $n = n + 1$
end while

The general Non-linear CG procedure

$\Gamma_0(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(0)}(x)$
$g(x) \leftarrow e^{-i\alpha_0 \Gamma_0(x)}$
$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
$s_0(x) \leftarrow \Gamma_0(x)$
$n = 1$
while $\Theta^{(n)} >$ Tolerance **do**
  $\Gamma_n(x) \leftarrow a\Delta_\mu a g_0 A_\mu^{(n)}(x)$
  $\beta_n = \max\left[0, \frac{\sum_x \mathrm{Tr}[\Gamma_n(x)(\Gamma_n(x) - \Gamma_{n-1}(x))]}{\sum_x \mathrm{Tr}[\Gamma_{n-1}(x)^2]}\right]$
  $s_n(x) \leftarrow \Gamma_n(x) + \beta_n s_{n-1}(x)$
  $\alpha_n = \min(F(U); \alpha' s_n)$
  $g(x) \leftarrow e^{-i\alpha_n s_n(x)}$
  $U_\mu^{(n+1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x) U_\mu^{(n)}\left(x + a\frac{\hat{\mu}}{2}\right) g(x + a\hat{\mu})^\dagger$
  $n = n + 1$
end while

The general Non-linear CG procedure

$$\Gamma_0(x) \leftarrow \tilde{F}\,\frac{p_{\text{Max}}^2}{Vp^2}\,F a\Delta_\mu a g_0 A_\mu^{(0)}(x)$$

$$g(x) \leftarrow e^{-i\alpha_0\Gamma_0(x)}$$

$$U_\mu^{(1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x)U_\mu^{(0)}\left(x + a\frac{\hat{\mu}}{2}\right)g(x + a\hat{\mu})^\dagger$$

$$s_0(x) \leftarrow \Gamma_0(x)$$

$$n = 1$$

while $\Theta^{(n)} >$ Tolerance **do**

$$\Gamma_n(x) \leftarrow \tilde{F}\,\frac{p_{\text{Max}}^2}{Vp^2}\,F a\Delta_\mu a g_0 A_\mu^{(n)}(x)$$

$$\beta_n = \max\left[0, \frac{\sum_x \mathrm{Tr}[\Gamma_n(x)(\Gamma_n(x)-\Gamma_{n-1}(x))]}{\sum_x \mathrm{Tr}[\Gamma_{n-1}(x)^2]}\right]$$

$$s_n(x) \leftarrow \Gamma_n(x) + \beta_n s_{n-1}(x)$$

$$\alpha_n = \min(F(U); \alpha' s_n)$$

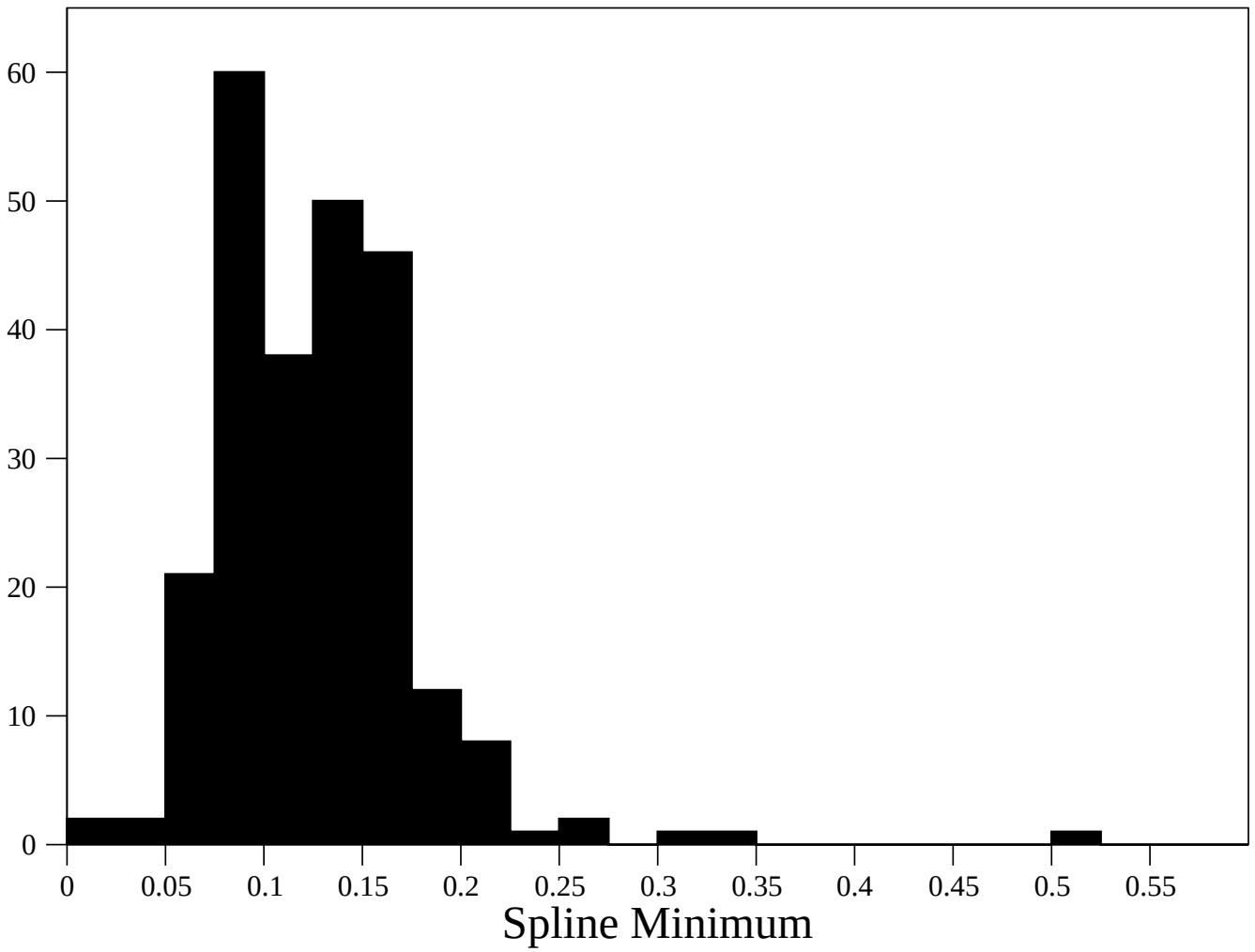
$$g(x) \leftarrow e^{-i\alpha_n s_n(x)}$$

$$U_\mu^{(n+1)}\left(x + a\frac{\hat{\mu}}{2}\right) \leftarrow g(x)U_\mu^{(n)}\left(x + a\frac{\hat{\mu}}{2}\right)g(x + a\hat{\mu})^\dagger$$

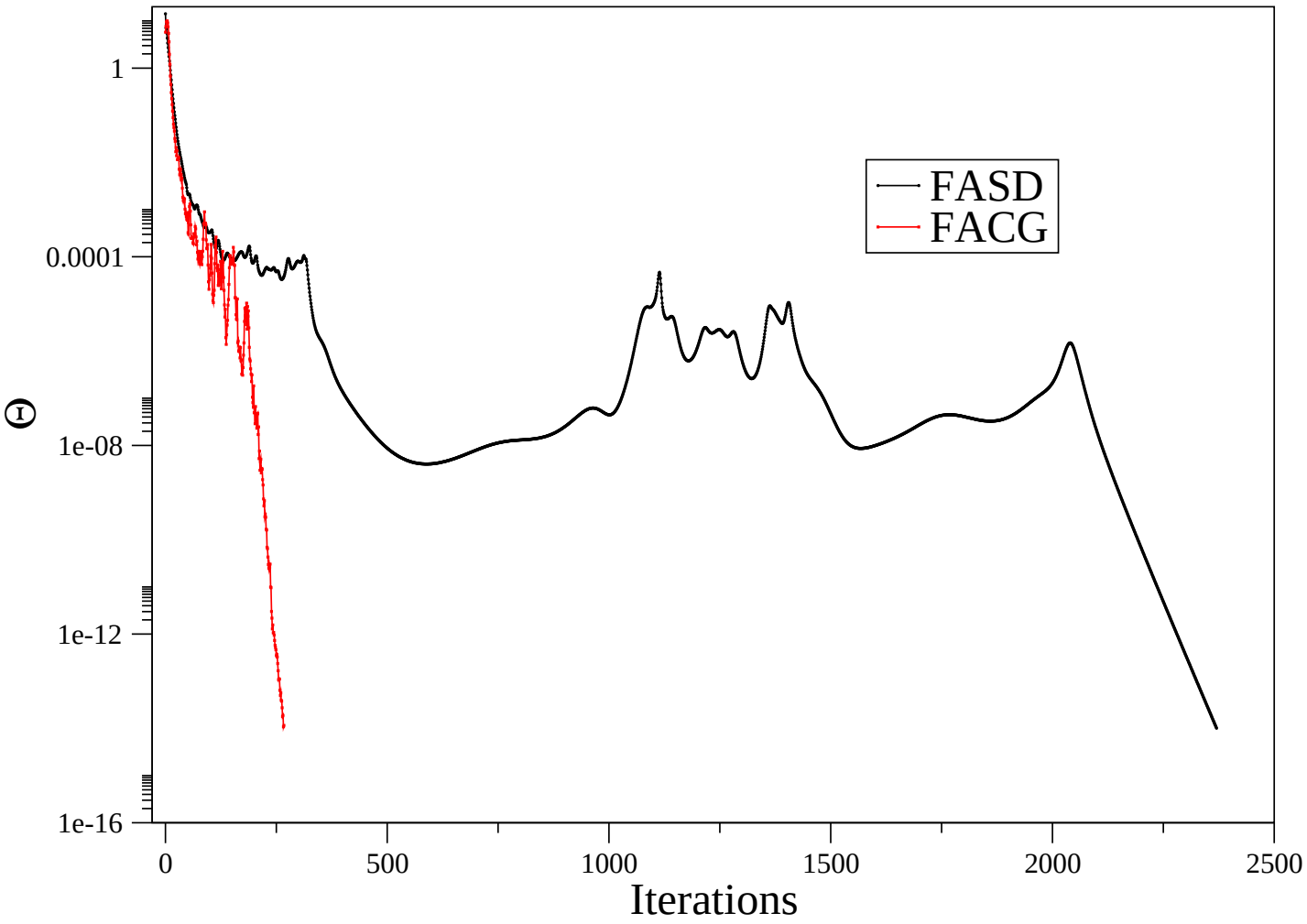
$$n = n + 1$$

end while

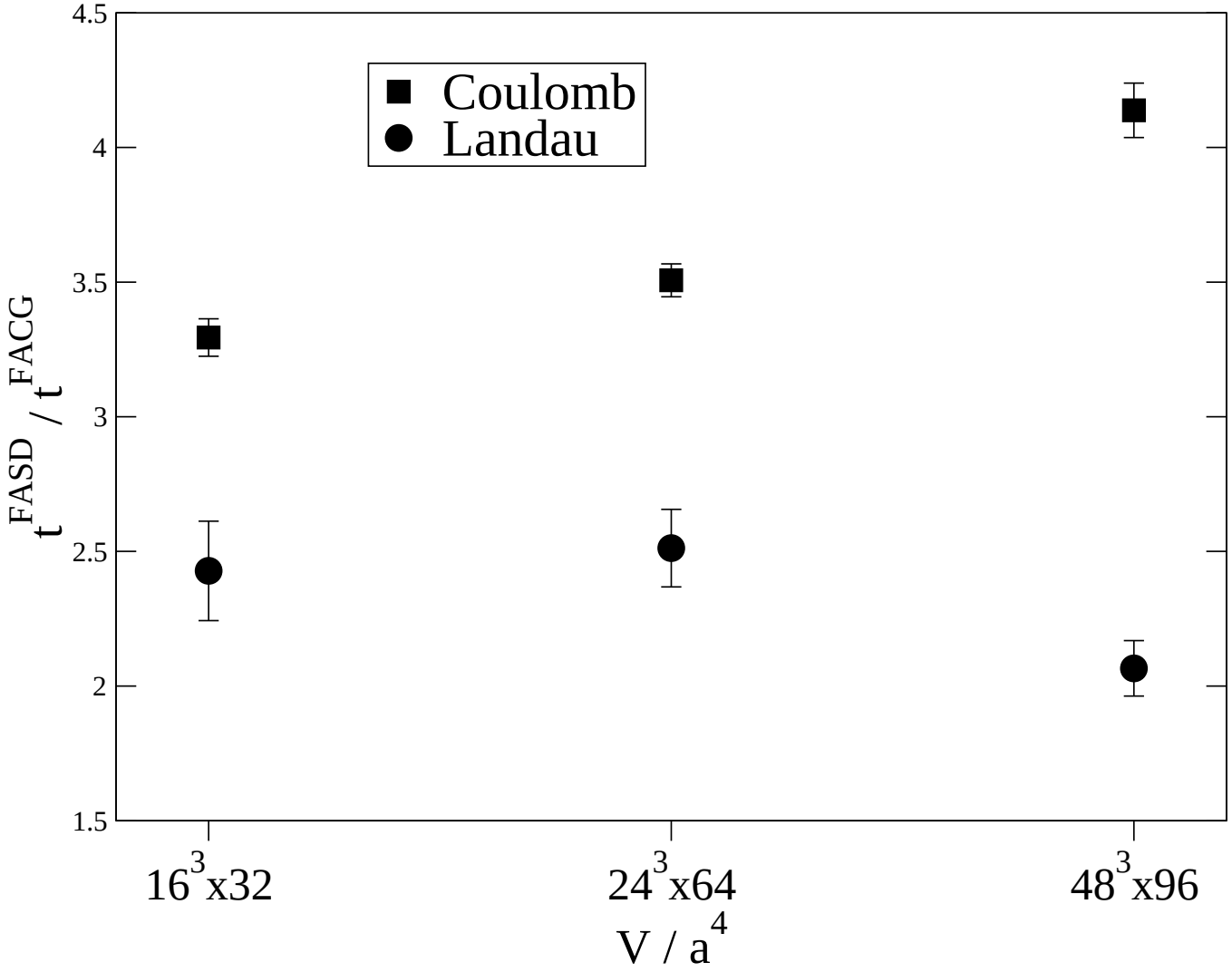
FACG spline history



FASD vs. FACG



Effective speed up



Timings

	Landau		Coulomb	
Volume	FASD (s)	FACG (s)	FASD (s)	FACG (s)
$16^3 \times 32$	25(1)	10.5(5)	27.4(6)	8.3(1)
$24^3 \times 64$	306(18)	122(7)	194(3)	55.4(4)
$48^3 \times 96$	12995(857)	6292(404)	10882(180)	2629(39)

Table: The time in seconds taken to achieve an accuracy of $\Theta^{(n)} < 10^{-14}$ for 25 randomly transformed copies of the same, well-thermalised configuration. Each configuration has the same lattice spacing. This measurement was performed using four 8-core AMD Opteron 6128 processors. All operations were performed in double precision.

Conclusions

- This method relies on fast functional evaluations and a line search
- At least a $2\times$ speed up from this procedure for all volumes considered
- Fewer FFTs compared to the FASD
- Method could be faster
- Check out www.github.com/RJHudspith/GLU/
- And my paper http://arxiv.org/abs/1405.5812