Numerical methods for lattice field theory

Mike Peardon

Trinity College Dublin

August 13, 2007

Hybrid Monte Carlo

Hybrid Monte Carlo

1. Draw new conjugate momenta, π from the normal distribution
2. Store the current field, ϕ and compute $\mathcal{H}(\pi, \phi)$
3. Integrate the equations of motion using a reversible, symplectic integrator with step-size h (such as leap-frog) so $(\phi, \pi) \overset{\text{leapfrog}}{\longrightarrow} (\phi', \pi')$
4. Compute $\mathcal{H}(\pi', \phi')$ and the change, $\Delta\mathcal{H}$
5. Accept ϕ' as the new entry in the Markov chain with probability

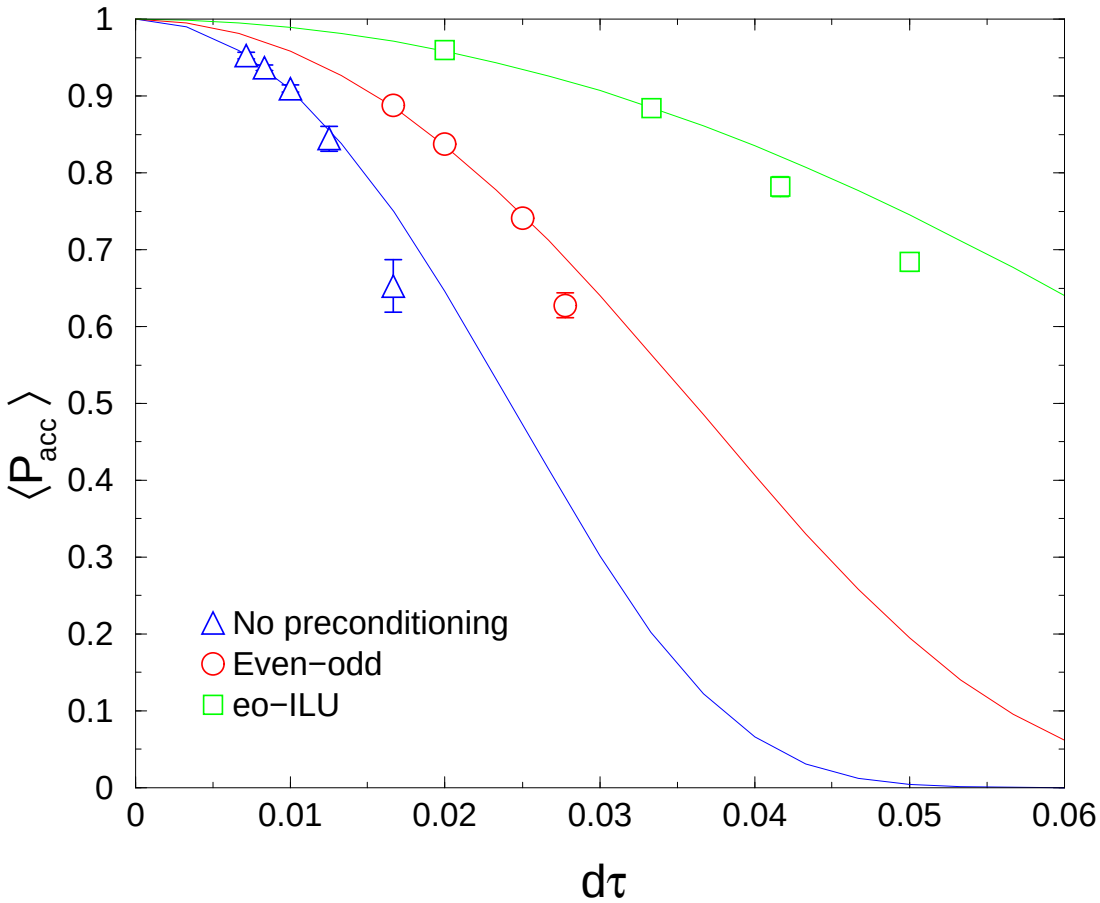
$$\mathcal{P}_{\text{acc}} = \min\left[1, e^{-\Delta\mathcal{H}}\right]$$

if the new configuration is rejected, then make ϕ the new entry.

- As $h \to 0, \Delta\mathcal{H} \to 0$, so $\mathcal{P}_{\text{acc}} \to 1$
- $E[e^{-\Delta\mathcal{H}}] = 1$ and $E[\Delta\mathcal{H}] = \frac{1}{2}E[(\Delta\mathcal{H})^2]$ for small $\Delta\mathcal{H}$

Metropolis-Hastings Acceptance probabilities

- The algorithm needs some tuning; make h small and the computer cost rises, make h too big and all proposals are rejected as $\Delta\mathcal{H}$ is large.
- $P_{acc} \propto \mathrm{erfc}(h^2/h_0^2)$. Example below from the $N_f = 2$ Schwinger model (2d QED)



Molecular dynamics (4)

- For gauge theories, our degrees of freedom are constrained, so we need to define hamiltonian dynamics on curved manifolds.
- A Lie group G (all continuous gauge groups) has particularly helpful properties. At all points, there is a well-defined, tangent space (the Lie algebra at the identity element) in which conjugate momenta naturally live.
- A useful definition for a momentum variable p conjugate to a group element U is

$$p = p_a T_a \text{ so } p \in \mathcal{L}(G)$$

with T_a the (hermitian) generators of the group and define the equation of motion for U to be

$$\dot{U} = ipU$$

- The extra appearance of U shows we need to rotate the Lie algebra (where p lives) to be tangent to U. The left multiplication is a convention (right works just as well).

Molecular dynamics (5)

- The (group invariant) kinetic term is

$$T = \mathrm{Tr}\ p^2 = \frac{1}{2}\sum_a p_a^2$$

  and so the co-ordinates p_a are still normally distributed.
- A toy example:

$$S[U] = \mathrm{ReTr}\ U\Sigma \ \text{ so } \ \mathcal{H}[U,p] = \mathrm{Tr}\ p^2 + \mathrm{ReTr}\ U\Sigma$$

  with $U \in SU(N)$ and $\Sigma \in GL(N)$ a constant background.
- A simple way to find the equations of motion is to demand the hamiltonian is conserved and use the definition of $\dot{p}$, so:

$$\dot{\mathcal{H}} = 2\mathrm{Tr}\ \dot{p}p + \mathrm{ReTr}\ \dot{U}\Sigma = 0$$

  substitute $\dot{U} = ipU$ and we get

$$\dot{p} = \frac{i}{4}\left\{U\Sigma - \Sigma^\dagger U^\dagger - \frac{1}{N}\mathrm{Tr}\ \left(U\Sigma - \Sigma^\dagger U^\dagger\right)\right\}$$

Molecular dynamics (6)

- Now QCD:

$$S = \beta \sum_x \mathrm{ReTr}\,(1 - U_\Box(x)) + \phi^* \left[M^\dagger M\right]^{-1} \phi$$

  the ϕ fields are (usually) held fixed in the integration stage, and are drawn at the start of the trajectory from a heat-bath (since they are normally distributed).
- The force term from the gauge action is the staple sum - the same object found in the Gibbs sampler methods. For the pseudofermions

$$\frac{d}{dt}\left\{\phi^* \left[M^\dagger M\right]^{-1} \phi\right\} = -\phi^* \left[M^\dagger M\right]^{-1} \frac{d}{dt}\left(M^\dagger M\right) \left[M^\dagger M\right]^{-1} \phi$$

- Define $X = \left[M^\dagger M\right]^{-1} \phi$ and $Y = MX$, we get

$$\frac{d}{dt}\left\{\phi^* \left[M^\dagger M\right]^{-1} \phi\right\} = -Y^* \frac{dM}{dt} X - X^* \frac{dM^\dagger}{dt} Y$$

Molecular dynamics (7)

- The sparse structure of M generates a few terms, bilinear in the derived fields, X, Y; an example (for the Wilson fermion matrix) would yield a similar expression for Σ in

$$\dot{p}_\mu(x) \propto i(U_\mu(x)\Sigma_\mu(x) - \Sigma^\dagger U_\mu^\dagger(x) - \frac{2}{N}\mathrm{Im}\mathrm{Tr}\ U_\mu(x)\Sigma_\mu(x))$$

$$\Sigma_\mu(x) = (1-\gamma_\mu)^{\alpha\beta} X^\alpha(x+\hat{\mu}) Y^{*\beta}(x) + \ldots$$

  (α, β spin components)
- The computationally intensive part is computing X and Y; this requires sparse matrix inversion.
- Fortunately, for each (pair of) inverses computed, **all** links are updated.

Symplectic integrators

- Classical dynamics is a smooth flow in a phase space, $\Lambda = (\phi_i, \pi_i)$. Phase space has a geometric structure.

Symplectic integrators

- An integrator is a function $A : \Lambda \to \Lambda$ so $(\phi, \pi) \xrightarrow{A} (\phi', \pi')$
- Define the block Jacobian $K_A = \begin{pmatrix} \frac{\partial \phi'}{\partial \phi} & \frac{\partial \pi'}{\partial \phi} \\ \frac{\partial \phi'}{\partial \pi} & \frac{\partial \pi'}{\partial \pi} \end{pmatrix}$ and $J = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$
- Then integrator A is called **symplectic** if

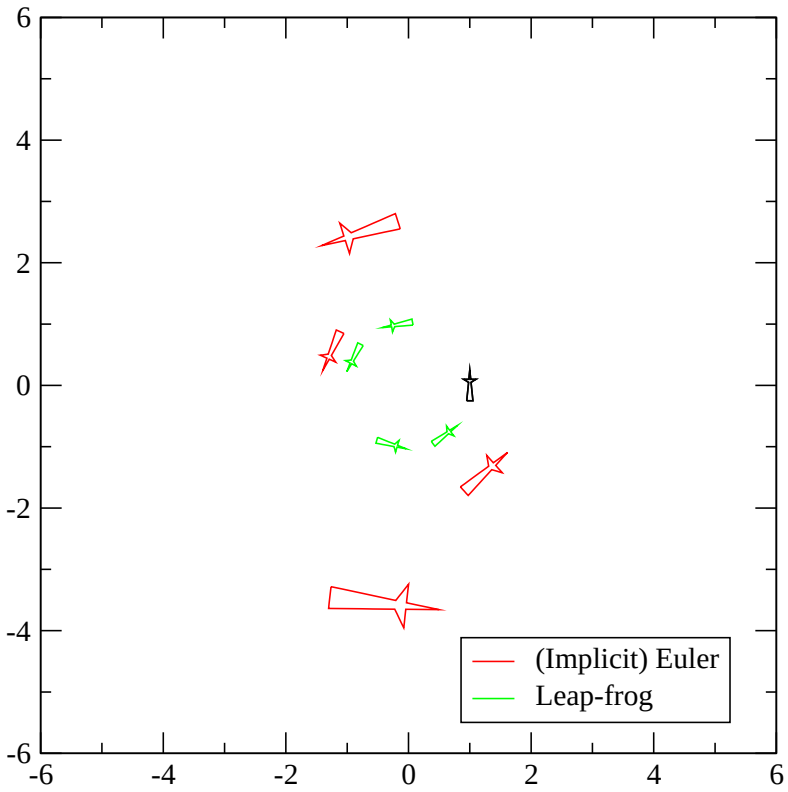
$$K_A^T J K_A = J$$

- This structure means symplectic integrators behave like a Lie group with the Poisson bracket acting as the Lie bracket.

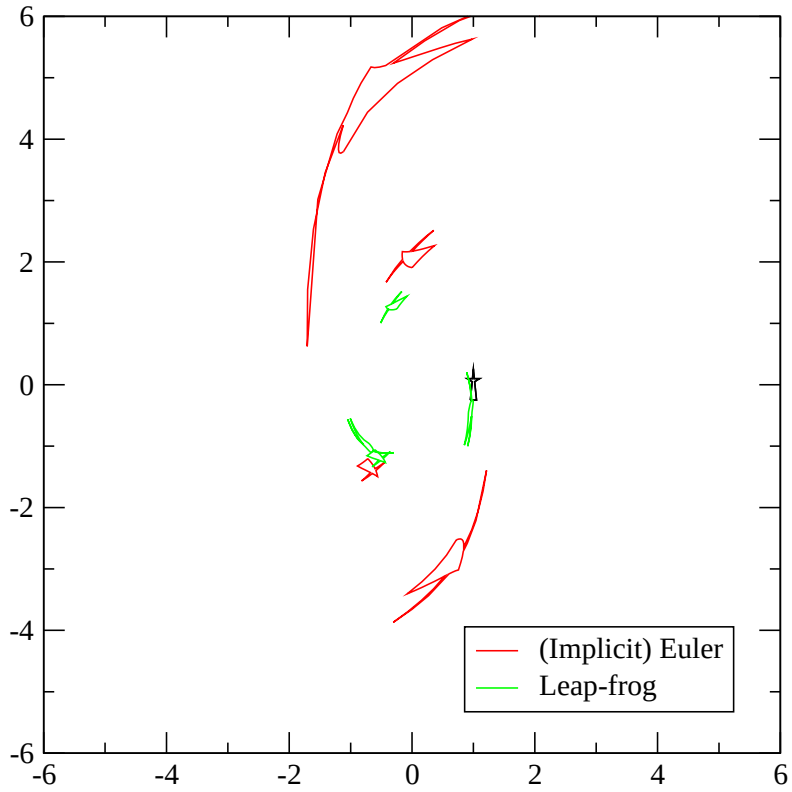
$$\{f, g\} = \sum_i \frac{\partial f}{\partial \phi_i} \frac{\partial g}{\partial \pi_i} - \frac{\partial f}{\partial \pi_i} \frac{\partial g}{\partial \phi_i}$$

The Seattle Phase Space Needle

- Integrating a simple one-dimensional system illustrates the difference between symplectic and non-symplectic (the Euler integrator)



Simple harmonic oscillator



$S = \frac{1}{2}\phi^2 + 0.375\phi^4$

Symplectic integrators (3)

- A useful linear operator on the space of functions on Λ is $\Delta_X = \{\cdot, X\}$ with X another operator on Λ then time-evolution for any function, f becomes

$$\frac{df}{dt} = \Delta_{\mathcal{H}} f = \{f, \mathcal{H}\} \quad \text{so} \quad f(t) = e^{t\Delta_{\mathcal{H}}} f(0)$$

- $e^{t\Delta_{\mathcal{H}}}$ is hard (impossible) to construct in practise. Is there a recipe for constructing useful symplectic integrators?
- $\mathcal{H} = T + S$ so consider the action of Δ_T and Δ_S. For example,

$$\Delta_T f = \{f, T\} = \sum_i \frac{\partial f}{\partial \phi_i} \pi_i \quad \text{so} \quad \Delta_T \phi_i = \pi_i, \Delta_T \pi_i = 0$$

this then leads to

$$e^{h\Delta_T} f(\phi(t), \pi(t)) = f(\phi(t) + h\pi(t), \pi(t))$$

and so $e^{h\Delta_T}$ is just the symplectic operator that adds $h\pi$ to ϕ

Symplectic integrators (4)

- It is easy to apply the symplectic operators $e^{h\Delta_T}$ and $e^{h\Delta_S}$; they are just adding momenta and the force to π and ϕ respectively.
- The two simplest symmetric symplectic integrator are then

$$e^{\frac{h}{2}\Delta_T} e^{h\Delta_S} e^{\frac{h}{2}\Delta_T} \text{ and } e^{\frac{h}{2}\Delta_S} e^{h\Delta_T} e^{\frac{h}{2}\Delta_S}$$

- and applying these n times forms the n-step leap-frog integrator.
- Since the poisson bracket behaves like a Lie algebra, we can use the Baker-Campbell-Hausdorff, so: $e^{\frac{h}{2}\Delta_T} e^{h\Delta_S} e^{\frac{h}{2}\Delta_T} = e^{h\Delta_{\mathcal{H}} + h^3\Delta'}$ with

$$\Delta' = \frac{1}{12}[\Delta_S, [\Delta_S, \Delta_T]] + \frac{1}{24}[\Delta_T, [\Delta_S, \Delta_T]]$$

- The Lie-algebraic properties also imply any compound of symplectic integrators can be written as $e^{h\Delta_{\mathcal{H}'}}$ with $\mathcal{H}'$ some Hamiltonian. This implies there is a different energy function that is exactly conserved by every symplectic integrator. For leap-frog,

$$\mathcal{H}' = \mathcal{H} + \frac{h^2}{12}\left(\pi_i \frac{\partial^2 S}{\partial\phi_i\phi_j}\pi_j - \frac{1}{2}\frac{\partial S}{\partial\phi_i}\frac{\partial S}{\partial\phi_i}\right)$$

Symplectic integrators (5)

- Better integrators can be constructed, by building longer (symmetric) compounds. The Omelyan integrator is

$$e^{h\mathcal{H}+h^3\Delta_O} = e^{\lambda h\Delta_T} e^{\frac{h}{2}\Delta_S} e^{(1-2\lambda)h\Delta_T} e^{\frac{h}{2}\Delta_S} e^{\lambda h\Delta_T}$$

  and this gives

$$\Delta_O = \alpha(\lambda)[\Delta_T,[\Delta_S,\Delta_T]] + \beta(\lambda)[\Delta_S,[\Delta_S,\Delta_T]]$$

  Minimising $\alpha^2+\beta^2$ gives $\lambda \approx 0.193$. This integrator works well, giving a speed-up of about 1.5-2.
- Getting rid of $\mathcal{O}(h^2)$ errors altogether using a compound of $e^{h\Delta_T}$ and $e^{h\Delta_S}$ requires seven terms, and is not useful in practice, since the integrators go unstable at smaller step-size.

Symplectic integrators (6)

- The action, S can often be split into a sum of terms, $S = S_1 + S_2 + \ldots$ each with its own force, $\frac{\partial S_1}{\partial \phi_i}, \frac{\partial S_2}{\partial \phi_i}, \ldots$ and symplectic integrator $e^{h\Delta_{S_1}}, e^{h\Delta_{S_2}}, \ldots$
- If the forces have very different sizes, then this splitting can be used to build better integrators; write

$$e^{h\Delta_{\mathcal{H}'_1}} = \prod_{i=1}^{m} e^{\frac{h}{2m}\Delta_T} e^{\frac{h}{m}\Delta_{S_1}} e^{\frac{h}{2m}\Delta_T}$$

  then a modified leapfrog integrator is

$$e^{\frac{h}{2}\Delta_{S_2}} e^{h\mathcal{H}'_1} e^{\frac{h}{2}\Delta_{S_2}}$$

  J.Sexton and D.Weingarten, Nucl. Phys. B380 (1992), 665
- This integrator has two time-scales, h and $\frac{h}{m}$ and tuning these scales leads to a faster algorithm, provided the force that is computationally cheap dominates.

Extensions to HMC (1)

- Odd-flavour simulations can be performed in HMC. Now the required importance sampling measure is $|\det M|$ and this is converted (using γ_5-hermiticity) to $\det \sqrt{M^\dagger M}$. This is then bosonised as before.
- Computationally efficient ways of representing $\sqrt{M^\dagger M}$ needed.
- **Polynomial approximations: PHMC**
  K. Jansen and R. Frezzotti, Phys. Lett. B402 (1997) 328
  If $\mathcal{P}(x) \approx \frac{1}{\sqrt{x}}$, then the action on the pseudofermions becomes

$$S_\phi = \phi^* \mathcal{P}(M^\dagger M)\phi = \phi^* \left(\prod_{i=1}^{N} (M^\dagger M - z_i) \right) \phi, \text{ with } z_i \text{ the roots of } \mathcal{P}$$

$$= \phi^* \left(\prod_{i=1}^{N} (\gamma_5 M - \sqrt{z_i})(\gamma_5 M - \sqrt{z_i}^*) \right) \phi$$

- Differentiating with respect to molecular dynamics time gives

$$\frac{dS_\phi}{dt} = \sum_{i=1}^{N} Y_i^* \frac{dM}{dt} X_i + X_i^* \frac{dM^\dagger}{dt} Y_i$$

Extensions to HMC (2)

- **Rational approximations: RHMC**
  M.Clark and A.Kennedy, Nucl.Phys.B (Proc. Suppl.) 129 (2004) 850
  If $\mathcal{R}(x) = \frac{A(x)}{B(x)} \approx \frac{1}{\sqrt{x}}$, then the action on the pseudofermions becomes

$$S_\phi = \phi^* \mathcal{R}(M^\dagger M)\phi$$

- A rational approximation (with A, B suitably chosen) can be written

$$R(x) = c_0 + \sum_{i=1}^{n} \frac{c_i}{x + d_i}$$

  For example, an optimised rational approximation (A. Kennedy, `hep-lat/0504038`) to $\frac{1}{\sqrt{x}}$ with $x \in [0.003, 1]$ is $1/\sqrt{x} \approx$

$$0.390460391 \frac{(x + 2.3475661045)(x + 0.1058344600)(x + 0.0073063814)}{(x + 0.4105999719)(x + 0.0286165446)(x + 0.0012779193)}$$

$$= 0.3904603901 + \frac{0.0511093775}{x + 0.0012779193} + \frac{0.1408286237}{x + 0.0286165446} + \frac{0.5964845033}{x + 0.4105999719}$$

Extensions to HMC (3)

- The matrix approximation is then

$$(M^\dagger M)^{-1/2} \approx \mathcal{R}(M^\dagger M) = c_0 + \sum_{i=1}^{n} c_i (M^\dagger M + d_i)^{-1}$$

- Now evaluating $\mathcal{R}$ requires matrix inversion again, but the advantage is that machine precision can be reached for fairly low orders of polynomials, A, B in contrast to polynomial approximation.
- For rational approximations to $x^{-1/2}$ the coefficients are +ve.
- The solution of $(M^\dagger M + d_i)\chi_i = \phi$ for many different values of d_i can be achieved with "multi-mass" solver. Convergence determined by the least-well-conditioned problem (i.e. smallest d_i).
- Further modifications to the way the fermion determinant is represented are possible. Two of the most popular current methods are the Hasenbusch mass preconditioner and Lüscher's Schur alternating approach.
  M.Hasenbusch - Phys.Lett.B519 (2001) 177
  M.Lüscher - Comput.Phys.Commun 165 (2005) 199.

Krylov Space methods (1)

- For almost all these methods, we need an efficient way of solving $A[U]\psi = \eta$ for a sparse matrix A.
- The most common way to solve these problems is to use a Krylov space method. The n-dimensional Krylov space, $\mathcal{K}_n(A, v_0)$ is the vector space

$$\text{span}\left\{v_0, Av_0, A^2v_0, \dots\right\}$$

- When $n = N$, the rank of A the solution must lie in $\mathcal{K}_n(A, v_0)$ (Cayley-Hamilton). N is large, these methods are considered as iterative, with "good" convergence properties (usually exponential).
- New methods are emerging that are not Krylov space methods (such as deflation methods).

Krylov Space methods (2)

- For A positive-definite hermitian, then the best method is (usually) conjugate gradient:

The conjugate gradient algorithm

$$r_0 = \eta - A\psi_0, p_0 = r_0$$

Until convergence ($|r_k|$ small), repeat for $k = 0, 1, 2, \ldots$

$$\alpha_k = \frac{r_k^* \cdot r_k}{p_k^* \cdot A p_k}$$

$$\psi_{k+1} = \psi_k + \alpha_k p_k$$

$$r_{k+1} = \psi_k - \alpha_k A p_k$$

$$\beta_k = \frac{r_{k+1}^* \cdot r_{k+1}}{r_k^* \cdot r_k}$$

$$p_{k+1} = r_{k+1} + \beta_k p_k$$

- Simple iterations and low storage requirements

Krylov Space methods (3)

- For A non-hermitian, the method is generalised to the BiCG algorithm, which has some improved versions, BiCGStab,
- Convergence of these methods is accelerated through **preconditioning**. In lattice QCD with Wilson-like quarks, the most commonly used form is even-odd (or red-black) preconditioning. Write $M\psi = \eta$ as

$$\begin{pmatrix} M_{ee} & M_{eo} \\ M_{oe} & M_{oo} \end{pmatrix} \begin{pmatrix} \psi_e \\ \psi_o \end{pmatrix} = \begin{pmatrix} \eta_e \\ \eta_o \end{pmatrix}$$

and then solve the equivalent problem:

$$\begin{pmatrix} M_{ee} & 0 \\ 0 & M_{oo} - M_{oe} M_{ee}^{-1} M_{eo} \end{pmatrix} \begin{pmatrix} \psi_e + M_{ee}^{-1} M_{eo} \psi_o \\ \psi_o \end{pmatrix} = \begin{pmatrix} \eta_e \\ \eta_o - M_{oe} M_{ee}^{-1} \eta_e \end{pmatrix}$$

- Recently, interest has been in methods that use "deflation". Here, as the Krylov space is constructed, an set of approximate low eigenvectors for A is built and stored. These can be used in to accelerate inversion. See e.g.
  W. Wilcox, presentation at Lattice 2007.