Speedy algorithms for lattice fields: status and unfinished (?) developments

Ulli Wolff

Humboldt University, Berlin

.....contain a brief review of the status of the topic and also devote a portion of time to speculation.....

my interpretation:

- a very subjective subset is reviewed
- some off-the-mainstream developments described
- some bias toward cluster methods

outline

- Speed of MC algorithms
- Scalar fields
- Gauge fields
- Fermions
- a page with references at the end

Speed of MC algorithms in equilibrium

equilibrate $\longrightarrow$ $(a_1, b_1, ...)$ update $(a_2, b_{2,...})$ update ..... update $(a_N, b_N, ...)$

estimates $\{a_i, b_i, ...\}$ of observables $A, B, ...$ separated by "update"

$$\langle a_i \rangle = A, ...$$

autocorrelation function:

$$\Gamma_{AA}(t) = \langle (a_i - A)(a_{i+t} - A) \rangle, ..., \quad \Gamma_{AB}(t) = \langle (a_i - A)(b_{i+t} - B) \rangle$$

- $\langle .... \rangle$ dynamical average ($\leftrightarrow$ ensemble of MCs)
- coincides with statistical ('static') mean at equal MC time

Standard estimator:

$$\overline{a} = \frac{1}{N} \sum_{i=1}^{N} a_i \,, \quad \left\langle (\overline{a} - A)^2 \right\rangle = \frac{V_A}{N/2\tau_{\mathrm{int},A}}$$

variance (static, independent of "update"):

$$V_A = \Gamma_{AA}(0) = \left\langle (a_i - A)^2 \right\rangle$$

integrated autocorrelation time (dynamical):

$$2\tau_{\mathrm{int},A} = \sum_{t=-\infty}^{t=+\infty} \frac{\Gamma_{AA}(t)}{\Gamma_{AA}(0)}$$

remarks:

- several definitions of $\tau_{\mathrm{int},A}$ (agreeing when they are large)
- τ_{exp}: slowest mode for all observables (extremely hard to determine)
- $\Gamma, \tau_{\mathrm{int}}$ have to be estimated numerically (construct estimators)
- error of τ_{int}: $\leftrightarrow$ error of the error of observables
- often estimate $f(A, B)$ by $f(\overline{a}\,, \overline{b}\,)$, errors from $\Gamma_{AA}, \Gamma_{AB}, \Gamma_{BB}$

Critical slowing down

- $\tau_{\rm int}$ (usually) diverge in the continuum limit $\xi = (am_{\rm Phys})^{-1} \to \infty$

$$\tau_{\rm exp} \propto \xi^z, \quad \tau_{{\rm int},A} \propto \xi^{z_A}$$

- z, z_A close to 2 for standard local heatbath, Metropolis
  [referring to 'sweeps', i.e. cost("update") $\propto$ # of lattice sites]
- QCD (Ukawa, lat2001): cost $\propto a^{-7}$, $z \sim 3$, large prefactor
- exponents $\leftrightarrow$ dynamical RG, universality classes of algorithms (?), not as well developed as the usual RG

practical upshot:

$\Longrightarrow$ minimize $\tau_{{\rm int},A}$/(exec. time of "update") for optimally estimating A

Scalar fields and cluster algorithms

- 'classical' cluster algorithms for Ising, Potts and O(n) σ-models
- for continuous O(n) spins one practically or truely achieves $z_A = 0$
- $\xi \sim 100$ possible, $\tau_{\rm int} \sim 1$, O(#sites) ops. for independent config.

How?

$$Z = \int D s(x)\; e^{-\beta H(s)}, \quad s \in \mathbb{R}^n,\ |s| = 1$$

- for a given $\{s(x)\}$ parameterize a 'target set' $\{\tilde{s} = s_\perp + \sigma |s_\parallel|\}$ with respect to a random direction and an Ising field $\sigma(x) = \pm 1$
- $\tilde{H}(\sigma|s) = -\sum_{\langle x\,y\rangle} J_{x\,y}\sigma(x)\sigma(y), \quad J_{x\,y} = |s_\parallel(x) s_\parallel(y)|$
- $J_{x\,y}$: no frustration, random bond ferromagnet
- update from initial $\sigma \equiv {\rm sign}(s_\parallel)$ to new $\sigma(x)$ with any algo for $\tilde{H}$
- powerful with Swendsen-Wang or Single cluster algorithm for σ

Trivial identity: on each link write

$$\mathrm{e}^{\beta J\sigma\sigma'} = \mathrm{e}^{-\beta J} \sum_{b=\mathrm{off,on}} [\delta_{b,\mathrm{off}} + (\mathrm{e}^{2\beta J} - 1)\delta_{b,\mathrm{on}}\,\delta_{\sigma,\sigma'}]$$

proof: consider $\sigma\sigma' = +$ and $\sigma\sigma' = -$

- additional ('redundant') bond-variables $\{b_{xy}\}$
- throw new bonds for given spins (trivial)
- determine (bond-)percolation clusters (clever)
- throw new spins for given bonds (trivial)

Why is it so efficient? improved estimator identity:

$$\langle s(x)\cdot s(y)\rangle = \langle(\mathrm{smooth})\times\Theta(x,y)\rangle_{b,s}, \quad \Theta(x,y) = \begin{cases} 1 & \text{if } x,y \in \text{one cluster} \\ 0 & \text{else} \end{cases}$$

- cluster size automatically chained to the correlation length

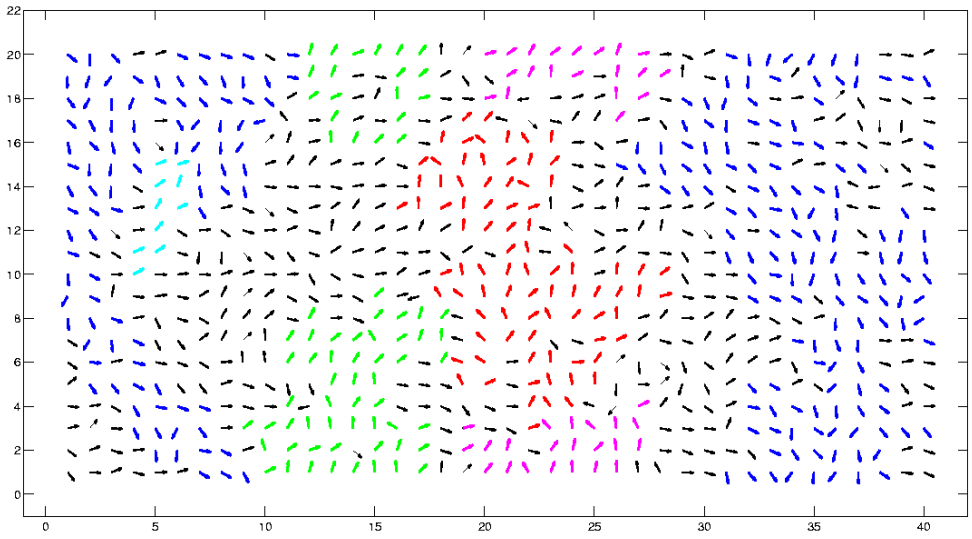
Disappointment of the 90's: no analogous successful algorithms for gauge theories, not even for other σ-models: $\mathrm{RP}(n)$, $\mathrm{CP}(n)$, $\mathrm{SU}(n)\times\mathrm{SU}(n)$

Why is this so?

- Ising embeddings possible, but J_{xy} in $\tilde{H}$ found frustrated
- cluster update still possible, but clusters too large
  $\rightarrow$ cluster flips $\approx$ global symmetry move, not dynamically relevant
- no improved estimator identity

Sokal et al. (heuristic) codimension-one 'no-go' argument:

- think of smooth classical $s(x)$
- clusters break, where $\beta J_{xy} \ll 1 \Rightarrow s_{\parallel}(x) \approx 0$ (s unchanged by flip)
- this is 1 real condition for $x \rightarrow$ clusters isolated by $d-1$ dimensional hypersurfaces
- other embeddings: >1 conditions, no isolation of clusters
- no-go theorem: embedded Ising spins $\leftrightarrow$ involutoric isometry of spin manifold; exists with $\dim = d-1$ fixed point only for sphere $\leftrightarrow$ $\mathrm{O}(n)$



O(2) model, $\xi = 40$, $\| =$ vertical, black=smaller clusters (picture: Tomasz Korzec)

Interesting way out: make (extreme!) use of the freedom in formulation/discretization in view of available algorithms

Cluster updates for 'quantum' models

- d-dimensional Hamiltonian operator, Trotter-Suzuki formula, $H \to D = d+1$ imaginary time path integral ($\sim$ euclidean FT)
- prototype: quantum Heisenberg model, $d = 1$

$$H = J \sum_x \vec{S}_x \cdot \vec{S}_{x+a}\,, \quad [S_x^k, S_y^l] = \delta_{xy}\, i\epsilon^{klm}\, S_x^m$$

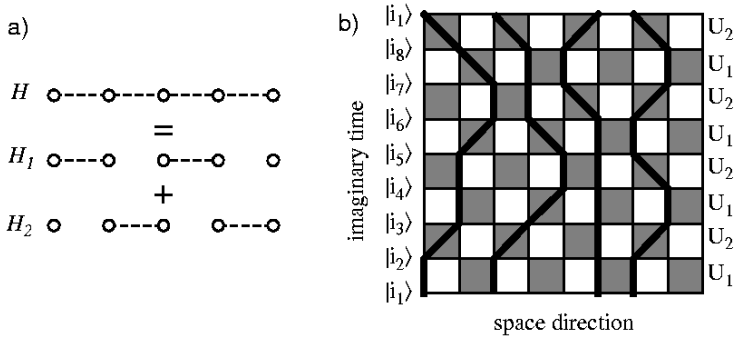
$$Z = \mathrm{Tr}\, \mathrm{e}^{-\beta H} \approx \mathrm{Tr}\, (\mathrm{e}^{-\varepsilon H_2}\, \mathrm{e}^{-\varepsilon H_1})^N$$

with $\varepsilon = \beta/N, H = H_1 + H_2,\ 1,2 \leftrightarrow$ even/odd links

- intermediate states: $S_x^3\,|s\rangle = s_x|s\rangle$, $s_x = \pm\frac{1}{2}$

$$Z = \sum_{\{s(t,x)=\pm\frac{1}{2}\}} A[s]$$

- A made of many factors $\langle s'|\mathrm{e}^{-\varepsilon H_1}|s\rangle$ and $\langle s'|\mathrm{e}^{-\varepsilon H_2}|s\rangle$
- classical form, $D=2$, here $A\geqslant 0$ ($=0$ for many configs)
- connecting sites with $s(t,x)=+\frac{1}{2} \leftrightarrow$ 'worldlines'



picture: Troyer et al, physics/0306128

- hard for local algorithms, but:
- nonlocal loop-cluster algorithm (Evertz, Lana, Marcu) works extremely well, gobal changes of worldlines
  - additional bond-type variables stochastically chosen
  - closed loops identified ($\leftrightarrow$ cluster search)
  - flip s-variables along loops
- high precision results for (anti)ferromagnet and other models

particular application (Wiese et al.):
cluster simulations of $D=2$ CP(n) models:

- devise a $d=2$ quantum spin system with SSB of SU(n) to U($n-1$) with Goldstone bosons $\simeq \mathrm{SU}(n)/\mathrm{U}(n-1) = \mathrm{CP}(n-1)$
- on a strip geometry: no real SSB but very small mass $\rightarrow$ effectively $d=1$ or $D=2$ for the classical model (dimensional reduction)

- CP($n-1$) model emerges as effective theory (χPT)
- universality check: same finite size scaling as conventional discretization, also the O(3) has been simulated in this way

to some degree, the 'quantum model' approach extends to fermions:
see below
construction of gauge theories from 'quantum link models' also investigated

Pure gauge systems

standard (local) procedure Hybrid OverRelaxation
we discuss SU(2), $\leftarrow$ SU(n) by Cabibbo-Marinari embedded SU(2)
take actions, for which the single link dependence is

$$P(U) \propto \exp[\mathrm{Re\,tr}\,(UM^\dagger)] = \exp[u^a m^a], \quad U = u^0 + u^k i\tau^k, \quad u^a u^a = 1$$

- includes many actions (counterexample: adjoint, $\mathrm{tr}(U^2_{\mathrm{plaq}})$)
- heatbath: new u with $P(U)$
- microcanonical/'overrelaxation': $u = u_\parallel + u_\perp \to u_\parallel - u_\perp$ $(\parallel, \perp \leftrightarrow m)$
- HOR: 1 HB sweep followed by N_{OR} OR sweeps
- $z \sim 1$ achievable, requires $N_{\mathrm{OR}} \propto \xi$ (similar $\omega = 2 - \frac{c}{\xi}$ for 'true' OR)
- link order can be important, [loop-μ [loop-x .....]] found ++

$\Rightarrow$ update time often negligible in quenched simulations

cluster algorithms for Z(2) gauge theories:

- throw bond on plaquettes

$$\mathrm{e}^{\beta\sigma_1\sigma_2\sigma_3\sigma_4} = \mathrm{e}^{-\beta} \sum_{b=\mathrm{off,on}} [\delta_{b,\mathrm{off}} + (\mathrm{e}^{2\beta}-1)\delta_{b,\mathrm{on}}\,\delta_{\sigma_1\sigma_2\sigma_3\sigma_4,1}]$$

- find random new gauge field under constraint: $\sigma_{\mathrm{plaqu}} = 1$ where $b = \mathrm{on}$
  - this step is hard
  - efficient graph theoretical solution in $D = 3$ (Ben-Av et al.)
  - linear algebra solution (Bunk), also for gauge-Higgs
  - $z \approx 0.5$

no embedding scheme for continuous variables, not even $3D$ U(1) model

- possible recent progress: 'worm' algorithms (Alet et al.), changing global flux loops

Dynamical fermions

chiral invariant quarks: $\rightarrow$ talks by Tony Kennedy & Kostas Orginos

$$Z = \int DU\, D\bar{\psi}\, D\psi\, \mathrm{e}^{-S_{\mathrm{G}}(U) - a^4 \sum_x \bar{\psi}(\not{D}+m)\psi}$$

determinant methods (N_{f} degenerate flavors) use

$$\int D\bar{\psi}\, D\psi\, \mathrm{e}^{-a^4 \sum_x \bar{\psi}(\not{D}+m)\psi} = \det(\not{D}+m)^{N_{\mathrm{f}}}$$

pseudofermionic methods [take N_{f}=2, assume $\not{D}^\dagger = \gamma_5 \not{D} \gamma_5$, set $Q = \gamma_5(\not{D}+m)$] use

$$\det(\not{D}+m)^2 = \det(Q)^2 = \int D\varphi^\dagger\, D\varphi\, \mathrm{e}^{-a^4 \sum_x \varphi^\dagger Q^{-2} \varphi}$$

standard workhorse: HMC ($a \equiv 1$)

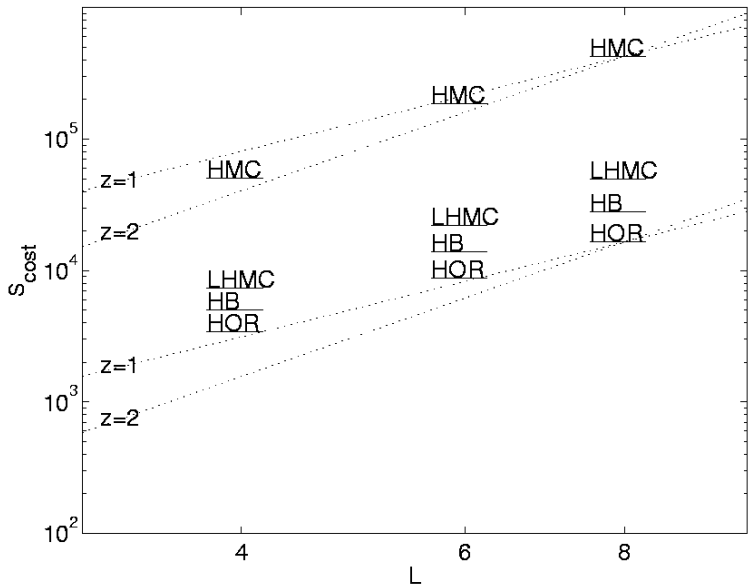
$$\mathcal{H} = \frac{1}{2}\sum_{x\mu} \mathrm{tr}\left[\Pi_\mu^2\right] + S(U), \quad S(U) = S_{\mathrm{G}}(U) + \sum_x \varphi^\dagger Q^{-2}\varphi$$

- update φ by global heatbath
- update (Π, U) 'almost microcanonically' $\leftrightarrow$ discretized Hamilton eqs.
- accept step

general remarks:

- HMC is an exact algorithm including fermions
- HMC is a relatively slow algorithm in the pure gauge limit compared to HOR

plot: B. Gehrmann



cost of algorithms at constant physics, $L \equiv L/a$

recent improvements
A: momenta need not be exponentiated exactly (Lüscher)
B(essential): more than 1 pseudofermion (Hasenbusch)

- small $m \to$ large condition number $Q^{-2} \to$ large force $\to$ only small $\delta\tau$
- factorization: $\det(Q)^2 = \det(\tilde{Q}\tilde{Q}^\dagger) \times \det\Big(Q(\tilde{Q}\tilde{Q}^\dagger)^{-1} Q\Big)$
- one pseudofermion for each det $\to$ smaller total force, larger $\delta\tau$

ALPHA realization of this: $Q \leftrightarrow$ Wilson/clover (even/odd), $\tilde{Q} = Q - i\rho$

$$\Rightarrow S_\varphi = \sum_x \varphi_1^\dagger (Q^2+\rho^2)^{-1}\varphi_1 + \varphi_2^\dagger (Q^{-2}+\rho^{-2})\varphi_2$$

so far: $\rho \leftrightarrow \approx$ equal condition numbers, main cost: Q^{-2}
better: different step size for φ_1, φ_2 (QCDSF), smaller ρ
gains reported range form 2...6, only minor modification of code

No pseudofermions

- want: $\lim_{N_f \to 0}$ (myalgo) = HOR (not HMC)
- true for: (detailed bal.) proposal with HOR + accept with $|\det Q|^2$
- improve: proposal need not take just $S_G \to$ so-called UV-filtering
- problem: no fermion-guided proposal $\to$ acceptance ok?
- probably in general no, but SF in (phys.) small volume?
- successful in 2-D Schwinger model (Knechtli, Wolff)
- any stochasticity in the Metropolis step for $|\det Q|^2$ costs acceptance, exact inequality, a few exact generalized eigenvalues can help

$$Q^2(U_{\rm new})\psi = \lambda Q^2(U_{\rm old})\psi$$

No determinant

- Grassmann integral: finite sum of nonzero terms on any finite lattice
- can this be sampled directly? beware: minus-sign problem
- early example with recent progress: LQCD$_{\beta=0}$

staggered fermions, gauge group U(n):

$$Z = \int D\overline{\psi} D\psi DU \, \mathrm{e}^{\sum_x \left(m\overline{\psi}\psi + \sum_\mu \frac{\eta_\mu(x)}{2} \left[\overline{\psi}(x) U_\mu(x) \psi(x+\hat{\mu}) - \overline{\psi}(x+\hat{\mu}) U_\mu^\dagger(x) \psi(x) \right] \right)}$$

integrate U link by link $\Rightarrow$ local model in $\overline{\psi}\psi(x)$ ('ultralocal confinement')
do Grassman integral site by site $\Rightarrow$ discrete monomer-dimer representation:

$$Z = \sum_{\{k_\mu(x)=0,\ldots,n\}} \prod_{x\mu} \alpha_{k_\mu(x)} \prod_x \beta_{\sigma(x)} m^{\sigma(x)}$$

α_i, β_i known weights ($\geqslant 0$), $k_\mu(x)$ dimers on link $x\mu$
$\sigma(x) = n - \sum_\mu k_\mu(x) + k_\mu(x-\hat{\mu})$ monomers at site x,

extensions to SU(n) include polymers/baryons in addition

- local updates of $k_\mu(x)$ only possible for $m > 0$
- Adams, Chandrasekharan: much more efficient nonlocal cluster algorithm, also at $m = 0$ (constraint $\sigma(x) \equiv 0$)
- precision study for chiral SSB for $m \to 0$, finite T transition
- precision matching to χPT
- still: $\beta = 0$ only :-(

alternative quantum model method:

- fermionic d-dim. Hamiltonian, Trotter-Suzuki with occupation number intermediate states $\to D = d + 1$ dim. 'euclidean' world line theory
- generically $\to$ minus sign problem $\leftrightarrow$ topology of world lines
- #configs $\propto \exp(cV)$ needed

- overcome for a class of Hubbard-like models by the meron cluster technique (Chandrasekharan, Wiese)
    - combine analytically configurations $\rightarrow$ weights $\geqslant 0$
    - modified importance sampling, reweighting
- only certain models, clever construction needed case by case
- warning: the generic 'sign problem' may NP-hard (Troyer, Wiese, cond-mat/0408370)

Conclusions

- classical cluster algorithms restricted to O(n) and discrete models
- progress via quantum models $\rightarrow$ think of non-standard discretizations
- steady progress for the 'standard' fermion method (via det, pseudofermions, HMC): optimized solvers, integration schemes, preconditioning, 2 (or more) pseudofermions, domain decomposition (Lüscher),...
- break-through type gains???
- maybe we should keep looking for a representation of relativistic fermions as a weighted sum over configurations (other than the standard nonlocal 'bosonization') in connection with cluster-type algorithms
- clearly a long way to go for QCD

Some references

incomplete, reviews where available, add 'and further references found there' to each entry....
errors, autocorrelations: U. Wolff, hep-lat/0306017
review on classical cluster algorithms: F. Niedermayer, hep-lat/9704009
codim. one no-go theorem: S. Caracciolo et al., hep-lat/9205005
loop cluster algorithm: H.G. Evertz, cond-mat/9707221
CP(n) cluster simulation: B.B.Beard, U. Wiese, hep-lat/0406040
meron cluster algorithm: S. Chandrasekharan et al., cond-mat/0201360
quantum spin and fermion models: M. Troyer et al, physics/0306128,
2 pseudofermions: M. Hasenbusch, hep-lat/0107019, hep-lat/0310029
dimer models: D.H. Adams, S. Chandrasekharan, hep-lat/0303003