Complex Langevin dynamics: an overview of recent developments

Frank James
Gert Aarts\textsuperscript{1}, Erhard Seiler\textsuperscript{2}, Ion-Olimpiu Stamatescu\textsuperscript{3}

12\textsuperscript{th} March 2012

\textsuperscript{1}Swansea University
\textsuperscript{2}MPI Munich
\textsuperscript{3}Uni. Heidelberg
Introduction to complex Langevin dynamics
Formal arguments for correct convergence
Criteria for correctness
An improved integration algorithm
Study of SU(3) spin model with chemical potential
Introduction

- Goal: lattice simulations to determine QCD phase structure
- Problem: chemical potential makes action complex
- Complex weight can’t be interpreted as a probability

\[ Z = \int D\phi \left| e^{-S(\phi)} \right| e^{i\varphi} \]

- Standard methods based on importance sampling break down
- Origin of the sign problem
Reweighting

- Can shift complex phase from weight to observable
- Simulate with respect to a real and positive weight, phase quenched theory

\[
\langle O \rangle = \frac{\int D\phi \, O(\phi)e^{i\phi}|e^{-S(\phi)}|}{\int D\phi \, e^{i\phi}|e^{-S(\phi)}|} = \frac{\langle Oe^{i\phi} \rangle_{pq}}{\langle e^{i\phi} \rangle_{pq}}
\]

- Problem: the phase vanishes exponentially as volume \( \Omega \to \infty \):

\[
\langle e^{i\phi} \rangle_{pq} = \frac{Z}{Z_{pq}} \sim e^{-\Omega\Delta f}
\]

- Required simulation time grows exponentially
- Can measure \( \langle e^{i\phi} \rangle_{pq} \) to quantify “severeness” of sign problem
Langevin dynamics does not rely on importance sampling

Add fictitious time-like parameter $\vartheta$ (Langevin time), $\phi \rightarrow \phi(\vartheta)$

Equation of motion with noise term, Langevin equation

$$\frac{\partial \phi}{\partial \vartheta} = -\frac{\delta S(\phi)}{\delta \phi} + \eta$$

Fluctuations from Gaussian noise

$$\langle \eta(\vartheta) \eta(\vartheta') \rangle = 2\delta(\vartheta - \vartheta'), \quad \langle \eta(\vartheta) \rangle = 0$$
Expectation values

- Expectation values taken as noise averages
- Equal to quantum expectation values in limit of large times

$$\lim_{\vartheta \to \infty} \langle O(\vartheta) \rangle_\eta = \langle O \rangle$$

- When action is real, can be shown that the stationary solution generates configurations distributed $e^{-S}$
Complex Langevin dynamics

- With action complex, can still write down the (complex) Langevin equation
- Complex drift term forces all degrees of freedom into complex plane
- Need to complexify degrees of freedom $\phi \rightarrow \phi^R + i\phi^I,$

$$\frac{\partial \phi^R}{\partial \vartheta} = K^R + \eta, \quad \frac{\partial \phi^I}{\partial \vartheta} = K^I$$

- Drift terms given by

$$K^R = -\text{Re} \left. \frac{\delta S}{\delta \phi} \right|_{\phi \rightarrow \phi^R + i\phi^I}, \quad K^I = -\text{Im} \left. \frac{\delta S}{\delta \phi} \right|_{\phi \rightarrow \phi^R + i\phi^I}$$
Discretised equations

- Discretise time: $\vartheta = \epsilon n$
- Standard (Euler) integration:
  \[
  \phi^R(n + 1) = \phi^R(n) + \epsilon K^R(n) + \sqrt{\epsilon} \eta(n)
  \]
  \[
  \phi^I(n + 1) = \phi^I(n) + \epsilon K^I(n)
  \]
- Introduces $O(\epsilon)$ stepsize corrections
- Discrete process generates configurations distributed with effective action
  \[
  \tilde{S} = S_0 + \epsilon S_1 + \ldots
  \]
- Correct results obtained by extrapolation to $\epsilon \to 0$
A simple example

- Single degree of freedom:
  \[ S = \frac{1}{2} \sigma x^2, \quad \sigma = A + iB \]

- Complexify \( x \rightarrow x + iy \) and get Langevin equations
  \[ \dot{x} = K_x + \eta, \quad \dot{y} = K_y \]

- Force terms
  \[ K_x = -Ax + By, \quad K_y = -Ay - Bx \]

- Can solve the equation of motion directly, taking initial conditions \( x(0) = y(0) = 0 \):
  \[ x(\vartheta) = \int_0^{\vartheta} e^{-A(\vartheta - s)} \cos[B(\vartheta - s)]\eta(s)ds \]
  \[ y(\vartheta) = -\int_0^{\vartheta} e^{-A(\vartheta - s)} \sin[B(\vartheta - s)]\eta(s)ds \]
A simple example

- Expectation values in limit $\nu \to \infty$
  
  \[
  \langle x^2 \rangle = \frac{1}{2A} \frac{2A^2 + B^2}{A^2 + B^2} \\
  \langle y^2 \rangle = \frac{1}{2A} \frac{B^2}{A^2 + B^2} \\
  \langle xy \rangle = -\frac{1}{2} \frac{B}{A^2 + B^2}
  \]

- The correct (holomorphic) combination is recovered
  
  \[
  \langle x^2 \rangle \to \langle x^2 - y^2 + 2i xy \rangle = \frac{A - iB}{A^2 + B^2} = \frac{1}{A + iB} = \frac{1}{\sigma}
  \]
What is the problem?

- Looks good: simple idea, easy to implement and known about since 1980s
- Some problems:
  - New degree of freedom $\phi^I$ is unbounded
  - Simulations can be unstable and follow runaway trajectories in direction $\phi^I$
  - No proof of convergence to correct distribution (or at all)
  - Simulations can converge to a well defined distribution, but results turn out to be wrong

- Instabilities cured by careful integration with an adaptive stepsize

  Aarts, FJ, Seiler, Stamatescu, 2010
Formal arguments

- Aim: understand conditions for correct convergence of complex Langevin process
- For simplicity, consider a single degree of freedom, $x$
- Replace original measure with the equilibrium distribution $P$ of the complex Langevin process
  1. Complex measure $e^{-S}dx$, which suffers from a sign problem
  2. Real and positive measure $Pdx dy$, complex Langevin solution
- Expectation values of holomorphic functions should agree
Fokker-Planck equation

- Fokker-Planck equation dual to complex Langevin process

\[ \frac{\partial}{\partial \vartheta} P(x, y; \vartheta) = L^T P(x, y; \vartheta) \]

- Fokker-Planck operator given by

\[ L^T = \nabla_x[\nabla_x - K_x] - \nabla_y K_y \]

- \( P(x, y; \vartheta) \) is a real distribution
- Probability density of at time \( \vartheta \) for the complexified variables \( x, y \)
Also consider the complex density \( \rho(x; \vartheta) \) with \( x \) real

\[
\frac{\partial}{\partial \vartheta} \rho(x; \vartheta) = L_0^T \rho(x; \vartheta)
\]

Complex Fokker-Planck operator

\[
L_0^T = \nabla_x [\nabla_x + (\nabla_x S(x))] 
\]

Complex density represents original description, suffers from sign problem

Correct stationary solution for \( \rho \) exists

\[
\rho(x; \infty) \propto e^{-S(x)}
\]
Evolution of the densities

• Define expectation values

\[ \langle O \rangle_{P(\vartheta)} = \frac{\int O(x + iy)P(x, y; \vartheta)dxdy}{\int P(x, y; \vartheta)dxdy} \]

\[ \langle O \rangle_{\rho(\vartheta)} = \frac{\int O(x)\rho(x; \vartheta)dx}{\int \rho(x; \vartheta)dx} \]

• Need to show that expectation values match

\[ \langle O \rangle_{P(\vartheta)} = \langle O \rangle_{\rho(\vartheta)} \]

• Initial conditions match requires

\[ P(x, y; 0) = \rho(x; 0)\delta(y) \]
Shifting time dependence

- Shift time dependence from densities to observable
- On holomorphic observables, may act with Langevin operator

\[ \tilde{L} = \left[ \nabla_x - (\nabla_x S(x)) \right] \nabla_x \]

- Action of \( \tilde{L} \) on holomorphic functions agrees with that of \( L \)
- Evolution of observables given by

\[
\frac{\partial}{\partial \vartheta} O(x; \vartheta) = \tilde{L} O(x; \vartheta)
\]

- Formally solved by

\[
O(x; \vartheta) = \exp(\vartheta \tilde{L}) O(x)
\]
Conditions for correct results

Consider

\[ F(\vartheta, \vartheta') = \int P(x, y; \vartheta - \vartheta') O(x + iy; \vartheta') dx dy \]

It interpolates between the two expectation values

\[ F(\vartheta, 0) = \int P(x, y; \vartheta) O(x + iy; 0) dx dy = \langle O \rangle_{P(\vartheta)} \]

\[ F(\vartheta, \vartheta) = \int P(x, y; 0) O(x + iy; \vartheta) dx dy \]

\[ = \int \rho(x; 0) \left( e^{\vartheta L_0} O \right) (x; 0) dx \]

\[ = \int O(x; 0) \left( e^{{\vartheta}^T L_0} \rho \right) (x; 0) dx \]

\[ = \langle O \rangle_{\rho(\vartheta)} \]
Integration by parts

- Expectation values match if $F(\vartheta, \vartheta')$ independent of $\vartheta$:

$$
\frac{\partial}{\partial \vartheta'} F(\vartheta, \vartheta') = - \int (L^T P(x, y; \vartheta - \vartheta')) O(x + iy; \vartheta') \, dx \, dy + \int P(x, y; \vartheta - \vartheta') L O(x + iy; \vartheta') \, dx \, dy
$$

- Integration by parts gives required cancellation

$$
\int P(x, y; \vartheta - \vartheta') L O(x + iy; \vartheta') \, dx \, dy \rightarrow \int L^T P(x, y; \vartheta - \vartheta') O(x + iy; \vartheta') \, dx \, dy
$$

- Needs boundary terms to vanish for $\langle O \rangle_\rho = \langle O \rangle_P$
Vanishing boundary terms requires decay of distribution to be sufficiently fast.

Products of observable and distribution (and derivatives)

\[ P(x, y; \vartheta - \vartheta') O(x + iy; \vartheta') \]

Real direction \( x \) will be either compact or distribution rapidly decaying.

Need distribution "narrow" and fast decay in imaginary direction \( y \).
Criteria for correctness

- Take slightly weaker condition \( \vartheta' = 0 \) and \( \vartheta \to \infty \)
- Condition now becomes

\[
\left. \frac{\partial}{\partial \vartheta'} F(\infty, \vartheta') \right|_{\vartheta' = 0} = - \int (L^T P(x, y, \infty)) O(x + iy, 0) dx dy + \\
\int P(x, y, \infty) LO(x + iy, 0) dx dy
\]

- First term vanishes automatically due to \( L^T P(x, y; \infty) = 0 \)
- Therefore \( \vartheta' \)-independence requires

\[
\langle LO \rangle = \int P(x, y; \infty) LO(x + iy; 0) dx dy = 0
\]

- Can be checked for any given observable
- Strong statement: should be true for all observables
SU(3) spin model

- Effective dimensionally reduced polyakov loop model for QCD
- Studied using complex Langevin dynamics in 1980s

- Recently developed method using flux formalism to circumvent sign problem in an alternate way

- Action given by \( S = S_B + S_F \),

\[
S_B = -\beta \sum_x \sum_{\nu=1}^3 \text{Tr} \ U_x \text{Tr} \ U_{x+\hat{\nu}}^\dagger + \text{Tr} \ U_{x+\hat{\nu}} \text{Tr} \ U_x^\dagger
\]

\[
S_F = -h \sum_x e^{\mu \text{Tr} \ U_x} + e^{-\mu \text{Tr} \ U_x^\dagger}
\]

- Contribution \( S_F \) makes action complex when \( \mu \neq 0 \), sign problem
Phase transition in region of small $h$
Disordered (confined) phase for lower $\beta$ values
Ordered (deconfined) phase for higher $\beta$ values
Phases separated by a first-order transition
Increasing chemical potential weakens the transition and becomes a crossover at a critical point
At larger $h$ there is a crossover only
Phase structure with small $h$
Langevin equations

- Can diagonalise $U_x$, write in terms of angles

$$\text{Tr } U_x = e^{i\phi_1 x} + e^{i\phi_2 x} + e^{-i(\phi_1 x + \phi_2 x)}$$

- Must include reduced Haar measure

$$S_H = -\sum_x \ln \left[ \sin^2 \left( \frac{\phi_1 x - \phi_2 x}{2} \right) \sin^2 \left( \frac{2\phi_1 x + \phi_2 x}{2} \right) \sin^2 \left( \frac{\phi_1 x + 2\phi_2 x}{2} \right) \right]$$

- Effective action $S_{\text{eff}} = S_B + S_F + S_H$

- Langevin dynamics then given by

$$\frac{\partial}{\partial \vartheta} \phi_{ax} = K_{ax} + \eta_{ax}, \quad K_{ax} = -\frac{\partial S_{\text{eff}}}{\partial \phi_{ax}}$$
Phase transition at $\mu = 0$

\[
\mu=0, \ h=0.02, \ 10^3
\]

\[
\langle \text{Tr}(U+U^{-1})/2 \rangle
\]

\[
\beta
\]

\[
0.12 \ 0.125 \ 0.13 \ 0.135 \ 0.14
\]
Imaginary chemical potential

- With imaginary chemical potential the action is real, no sign problem
- Complex Langevin results should be continuous across $\mu^2 = 0$ from the imaginary chemical potential results
- Non-analyticity is a sign of convergence to wrong limit
- XY model is an example of non-analyticity and incorrect convergence, where CL failed in part of the phase diagram

Choose observable even in $\mu$, $\langle \text{Tr} (U + U^{-1}) \rangle / 2$

Aarts and FJ, 2010
Analyticity in $\mu^2$

\[ \langle \frac{\text{Tr}(U+U^{-1})}{2} \rangle \]

$\beta = 0.135, 0.134, 0.132, 0.130, 0.128, 0.126, 0.124, 0.120$

$h = 0.02, 10^3$

\[ -1 -0.5 0 0.5 1 \]

\[ 0 0.5 1 1.5 2 \]
Taylor series expansion

- Can perform a Taylor series expansion in $\mu$
- Simulations at $\mu = 0$ used to extrapolate to $\mu > 0$
- Provides test for correct results at $\mu \neq 0$ for small chemical potentials
- Free energies in full and phase quenched theories:

$$f(\mu) = f(0) - (c_1 + c_2 h) h \mu^2 + O(\mu^4)$$
$$f_{pq}(\mu) = f(0) - c_1 h \mu^2 + O(\mu^4)$$

- With

$$c_1 = \frac{1}{\Omega} \sum_x \langle \text{Tr } U_x \rangle_{\mu=0} = 0.1146(21),$$
$$c_2 = \frac{1}{2\Omega} \sum_{xy} \langle \text{Tr } (U_x - U_x^\dagger) \text{Tr } (U_y - U_y^\dagger) \rangle_{\mu=0} = -3.534(72)$$
$\beta=0.125$, $h=0.02$, phase quenched
Density and Silver Blaze problem

- Silver Blaze problem: $\mu \neq 0$ but observables $\mu$-independent
  - Requires precise cancellations in numerical simulations
  - Density given by
    \[
    \langle n \rangle = \frac{1}{\Omega} \frac{\partial \ln Z}{\partial \mu} = \langle h e^{\mu} \text{Tr} \, U_x - h e^{-\mu} \text{Tr} \, U^\dagger_x \rangle
    \]
  - When $\mu \neq 0$ there is a difference between $\langle \text{Tr} \, U \rangle$ and $\langle \text{Tr} \, U^\dagger \rangle$
  - Silver Blaze effect requires $\mu$-independence: $\langle \text{Tr} \, U \rangle = \langle \text{Tr} \, U^\dagger \rangle$
  - Not possible to satisfy both requirements: no Silver Blaze here
  - Phase quenched:
    \[
    \langle n \rangle_{pq} = h \sinh \mu \langle \text{Tr} \, U_x + \text{Tr} \, U^\dagger_x \rangle_{pq}
    \]
- $\langle n \rangle_{pq} \neq 0$ immediately once $\mu \neq 0$
Density Taylor expansion

\[ \beta = 0.125, \ h = 0.02, \ 10^3 \]

\[ \langle n \rangle \]

- full
- phase quenched
Improved algorithm

- Need to extrapolate to vanishing stepsize $\epsilon \to 0$
- Standard algorithm (Euler integration) has $O(\epsilon)$ corrections
- Simple mid-point scheme with improved drift terms but noise unchanged does not improve corrections
- Must also modify noise terms
- An improved algorithm proposed for real Langevin dynamics

Chien-Cheng Chang, 1987

- Reduces corrections to $O(\epsilon^2)$ for free theories and $O(\epsilon^{3/2})$ for coupled systems
Improved algorithm

- Add intermediate steps $\psi, \tilde{\psi}$ and modify noise terms

\[
\psi_{ax}(n) = \phi_{ax}(n) + \frac{1}{2} \epsilon K[\phi_{ax}(n)] + k \sqrt{\epsilon} \tilde{\alpha}_{ax}(n),
\]
\[
\tilde{\psi}_{ax}(n) = \phi_{ax}(n) + \frac{1}{2} \epsilon K[\phi_{ax}(n)] + l \sqrt{\epsilon} \tilde{\alpha}_{ax}(n),
\]
\[
\phi_{ax}(n + 1) = \phi_{ax}(n) + \epsilon \left( aK[\psi_{ax}(n)] + bK[\tilde{\psi}_{ax}(n)] \right) + \sqrt{\epsilon} \alpha_{ax}(n)
\]

- Coefficients chosen to cancel $O(\epsilon)$ contributions:

\[
a = \frac{1}{3}, \quad b = \frac{2}{3}, \quad k = 0, \quad l = \frac{3}{2}
\]

- Random variable $\tilde{\alpha}_{ax}(n) = \frac{1}{2} \alpha_{ax}(n) + \frac{\sqrt{3}}{6} \xi_{ax}(n)$

- Gaussian noise terms $\alpha, \xi$
Stepsize corrections: observables

\[ \langle \text{Tr} U \rangle \]

\[ \beta=0.125, \mu=3, h=0.02, 10^3 \]

\[ \langle \text{Tr} U^{-1} \rangle \]

- Lowest order
- Improved

\[ \varepsilon \]

\[ 0 \quad 0.00025 \quad 0.0005 \quad 0.00075 \quad 0.001 \]
Stepsize corrections: criteria

$\beta=0.125$, $\mu=3$, $h=0.02$, $10^3$

- $\langle LTrU \rangle$ - lowest order
- $\langle LTrU \rangle$ - improved
- $\langle LTrU^{-1} \rangle$ - lowest order
- $\langle LTrU^{-1} \rangle$ - improved
Criteria for correctness

- Clear linear stepsize correction with standard algorithm
- Stepsize corrections much smaller with improved algorithm
- Find that $\langle LO \rangle$ vanish in limit of $\epsilon \to 0$
- Note: condition must be satisfied even with real Langevin dynamics
- Condition that $\langle LO \rangle = 0$ therefore quantifies stepsize corrections
Distribution of observables

- Compute histogram of observables during Langevin evolution
- Gives a distribution of values sampled by the process
- Complexified space should be explored
- Need distribution sufficiently localised for criteria to be satisfied
- Look at distributions of $\text{Re Tr } U$, $\text{Im Tr } U$
Distributions: $\text{Tr } U$

Histograms for different values of $\mu$ and $\beta = 0.125$, $h = 0.02$:
- $\mu = 0$, $8^3$
- $\mu = 0$, $12^3$
- $\mu = 1$, $8^3$
- $\mu = 1$, $12^3$
- $\mu = 3$, $8^3$
- $\mu = 3$, $12^3$
Distributions: $\text{Tr } U$

\[ \beta=0.125, \ h=0.02 \]

- $\mu = 0, 8^3$
- $\mu = 0, 12^3$
- $\mu = 1, 8^3$
- $\mu = 1, 12^3$
- $\mu = 3, 8^3$
- $\mu = 3, 12^3$
Distributions: $\phi^I$

$\beta=0.125$, $h=0.02$

$\mu=0$

$\mu=1$, $8^3$

$\mu=1$, $12^3$

$\mu=3$, $8^3$

$\mu=3$, $12^3$

$\sim \exp(-35|\phi^I|)$

$\sim \exp(-45|\phi^I|)$
Decay of $\phi^I$

- Decay of $P(\phi^I) \sim e^{-a|\phi^I|}$, with $a \sim 40$
- Rapid decay enough for correct convergence of $O = \text{Tr} \ U$
- Problem for observables of high powers like $\text{Tr} [U^k]$ for $k \gtrsim 40$
- Contributions are like $e^{-k\phi^I} \cos(k\phi^R)$
- These should vanish due to oscillating sign?
- Compare with U(1) one-link model where $a \sim 2$ and complex Langevin failed

Aarts, FJ, Seiler, Stamatescu, 2011
Complex neighbours

• Focus on single lattice site $x$

$$S = - \sum_x \text{Tr} \, U_x \left( \beta \sum_\nu \left[ \text{Tr} \, U^\dagger_{x+\hat{\nu}} + \text{Tr} \, U^\dagger_{x-\hat{\nu}} \right] + h e^\mu \right) +$$

$$\text{Tr} \, U^\dagger_x \left( \beta \sum_\nu \left[ \text{Tr} \, U_{x+\hat{\nu}} + \text{Tr} \, U_{x-\hat{\nu}} \right] + h e^{-\mu} \right)$$

• Combine neighbours

$$u_x = \frac{1}{6} \sum_{\nu=1}^3 \text{Tr} \, U^\dagger_{x+\hat{\nu}} + \text{Tr} \, U^\dagger_{x-\hat{\nu}}$$

• Action

$$S = - \sum_x (6 \beta u_x + h e^\mu) \text{Tr} \, U_x + (6 \beta u^*_x + h e^{-\mu}) \text{Tr} \, U^\dagger$$
Effective 1-link model

- Action
  \[ S = -\beta_1 \text{Tr} \, U - \beta_2 \text{Tr} \, U^\dagger \]

- Complex parameters
  \[ \beta_1 = \beta_{\text{eff}} e^{i\gamma} + h e^\mu, \quad \beta_2 = \beta_{\text{eff}} e^{-i\gamma} + h e^{-\mu} \]

- Couplings related to full model \( \beta_{\text{eff}} e^{i\gamma} = 6\beta u \)

- Write with angles, gain reduced Haar measure
  \[ S_H = -\log \left[ \sin^2 \left( \frac{\phi_1 - \phi_2}{2} \right) \sin^2 \left( \frac{2\phi_1 + \phi_2}{2} \right) \sin^2 \left( \frac{\phi_1 + 2\phi_2}{2} \right) \right] \]
Results

\[ \text{Re} \langle \text{Tr} U \rangle \]

\( \beta=0.5, \mu=1, h=0.02 \)
Results

\[
\text{Re} \langle \text{Tr} \ U \rangle
\]

\[
\gamma
\]

\[
\beta=2.0, \mu=1, h=0.02
\]
Conclusions

- Complex Langevin dynamics is still a candidate for circumventing sign problem
- Criteria provide necessary conditions for correct results
- Improved algorithm eliminates leading order stepsize corrections
- Criteria also provide general method for quantifying stepsize corrections
- SU(3) spin model passes the tests for correct results on both sides of phase diagram
- Effective 1-link model works for all complex parameters
- No dependence on severeness of sign problem and performance of complex Langevin dynamics