Langevin type dynamics for continuous and discrete systems

Lukas Kades

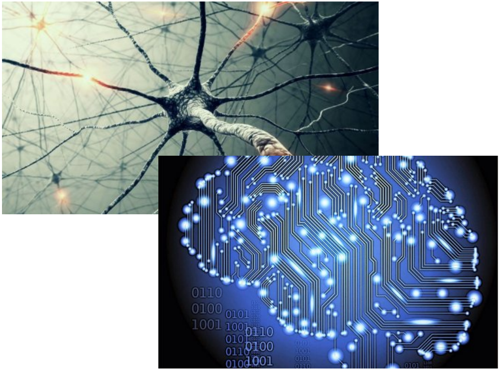
Cold Quantum Coffee

- Heidelberg University -

17 April 2018

Structure

- Motivation
- Langevin dynamics for the 2D Ising model
- Demystification
- Results
- Applications



http://www.kdnuggets.com

Motivation

Langevin Dynamics

Langevin equation:

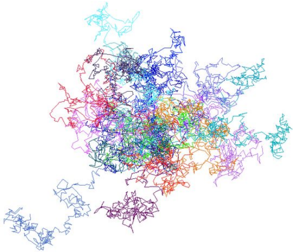
$$\frac{\partial}{\partial \tau}\phi_x(\tau) = -\frac{\delta S}{\delta \phi_x(\tau)} + \eta_x(\tau)$$

with Gaussian noise:

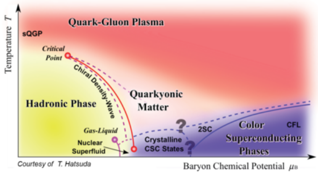
$$\langle \eta_x(\tau), \eta_{x'}(\tau') \rangle_\eta = 2\delta(x - x')\delta(\tau - \tau')$$

$$\langle \eta_x(\tau) \rangle_\eta = 0$$

$\Rightarrow$ Aim: Application of this formalism on neuromorphic hardware



http://mathsissmart.tumblr.com

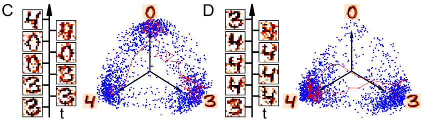


http://www.fair-center.eu

Motivation

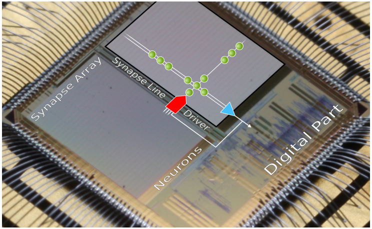
Human Brain Project/BrainScaleS - Introduction

- Neuromorphic computing system
- 1.6 million neurons
- 0.4 billion dynamic synapses
- 10000 times faster than their biological archetypes



Petrovici M.A., PhD Thesis, 2016


Electronic Vision(s) Group, Heidelberg University



Electronic Vision(s) Group, Heidelberg University

Motivation

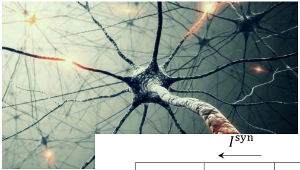
Human Brain Project/BrainScaleS - Stochastic interference

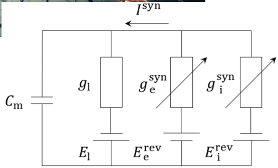
Ornstein-Uhlenbeck process:

$$\frac{du_{\text{eff}}(t)}{dt} = \Theta\left[\mu - u_{\text{eff}}(t)\right] + \sigma\eta(t)$$

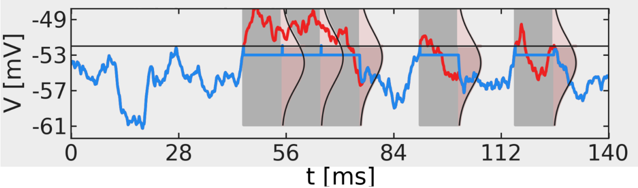
with:

$$\Theta = \frac{1}{\tau_{\text{syn}}}, \qquad \mu = u_{\text{leak}} + \sum_{\text{syn}\, i} \sum_{\text{spks}\, s} \kappa(t, t_{s,i})$$





Free membrane potential $u_{\text{eff}}(t)$:



http://www.kdnuggets.com and Petrovici M.A., PhD Thesis, 2016

Motivation

Comparison

Langevin dynamics:

$$\frac{\partial}{\partial \tau}\phi_x(\tau) = -\frac{\delta S}{\delta \phi_x(\tau)} + \eta_x(\tau)$$

- Continuous system
- Gaussian noise contribution
- Coupled system

$\Leftrightarrow$

BrainScaleS:

$$\frac{du_{\text{eff}}(t)}{dt} = \Theta\left[\mu - u_{\text{eff}}(t)\right] + \sigma\eta(t)$$

- Effective two-state system
- Gaussian noise contribution
- Coupled system

Langevin equation for a discrete two-state system?

How could a Langevin equation look like for the Ising model?

Langevin Dynamics for the 2D Ising Model

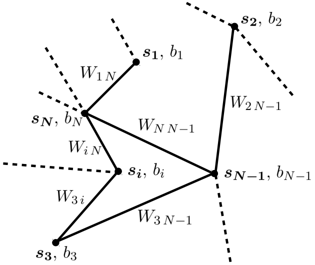
- N^2 states $s_i \in \{-1, 1\} = \{\downarrow, \uparrow\}$ on a square lattice
- Hamiltonian:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i$$

- Second order phase transition

$\Rightarrow$ Can be mapped easily onto a Boltzmann machine with $s_i \in \{0, 1\}$:

$$H = -\frac{1}{2} \sum_{ij} W_{ij} s_i s_j - \sum_i b_i s_i$$



Langevin Dynamics for the 2D Ising Model

Heuristic Approach

Langevin equation:

$$\frac{\partial}{\partial \tau}\phi_x(\tau) = -\frac{\delta S}{\delta \phi_x(\tau)} + \eta_x(\tau)$$

$$\Downarrow$$

Discrete Langevin equation ($\phi_x := \phi_x(\tau)$, $\phi'_x := \phi_x(\tau + \epsilon)$):

$$\phi'_x = \phi_x - \epsilon \frac{\delta S}{\delta \phi_x} + \sqrt{\epsilon}\eta_x \,,$$

$$\Downarrow$$

Identifications $S = \beta H$, $\phi_x = s_x$:

$$s'_i = s_i - \epsilon\beta \frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\eta_i$$

Langevin Dynamics for the 2D Ising Model

Heuristic Approach

Hamiltonian with $s_i \in \{-1, 1\} = \{\downarrow, \uparrow\}$:

$$H = -J \sum_{\langle i,j \rangle} s_i s_j - h \sum_i s_i$$

Langevin equation:

$$s_i' = \text{sign}\left(s_i - \epsilon \beta \frac{\partial H}{\partial s_i} + \sqrt{\epsilon} \eta_i \right)$$

Langevin Dynamics for the 2D Ising Model

Numerical Results

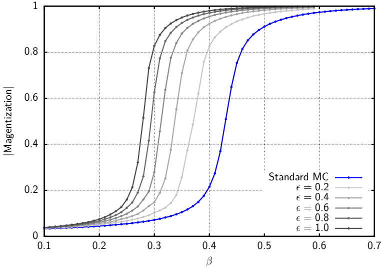
Langevin equation:

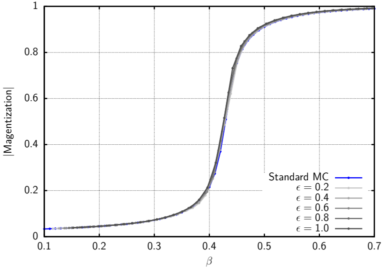
$$s_i' = \text{sign}\left(s_i - \epsilon\beta\frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\eta_i\right)$$

$$\Downarrow$$

Improved Langevin equation with adapted noise $\langle \tilde{\eta}_i, \tilde{\eta}_j' \rangle = \delta(j-i)\delta(t'-t)$:

$$s_i' = \text{sign}\left(s_i - \epsilon\frac{\beta}{\lambda(\epsilon)}\frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\tilde{\eta}_i\right)$$





I. Is this a Langevin equation for the Ising model?

$$s_i' = \text{sign}\left(s_i - \epsilon \frac{\beta}{\lambda(\epsilon)} \frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\tilde{\eta}_i\right)$$

II. Why does this work?

III. Is this a MCMC algorithm with the Boltzmann distribution $P(s) \propto \exp(-\beta H(s))$ as equilibrium distribution?

To III.: Markov Chain Monte Carlo Algorithm?

Markov Property

Update rule:

$$s_i' = \operatorname{sign}\left(s_i - \epsilon \frac{\beta}{\lambda(\epsilon)} \frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\tilde{\eta}_i\right)$$

Transition probabilities ($\Phi(x) = \int_{-\infty}^{x} dt \frac{1}{\sqrt{2\pi}} \exp(-t^2/2)$):

$$W(\downarrow\rightarrow\uparrow) = W(\downarrow \mid \uparrow) = \Phi\left(-\frac{1}{\sqrt{\epsilon}} - \frac{\sqrt{\epsilon}\beta}{\lambda(\epsilon)} \frac{\partial H}{\partial s_i}\right)$$

$$W(\uparrow\rightarrow\downarrow) = W(\uparrow \mid \downarrow) = \Phi\left(-\frac{1}{\sqrt{\epsilon}} + \frac{\sqrt{\epsilon}\beta}{\lambda(\epsilon)} \frac{\partial H}{\partial s_i}\right)$$

$\Rightarrow$ Markov property is fulfilled.

To III.: Markov Chain Monte Carlo Algorithm?

Ergodicity and Detailed Balance

- Ergodicity $\Rightarrow$ Yes!

- Detailed balance equation:

$$W(\uparrow\to\downarrow)P(\uparrow) = W(\downarrow\to\uparrow)P(\downarrow)$$

$$\Rightarrow \quad \frac{W(\downarrow\to\uparrow)}{W(\uparrow\to\downarrow)} \stackrel{!}{=} \frac{P(\uparrow)}{P(\downarrow)} = \exp\left[-\beta(E(\uparrow) - E(\downarrow))\right]$$

$\Rightarrow$ Detailed balance equation seems to be satisfied.

To II.: Why does this work?

- Limit between the cumulative Gaussian distribution and the exponential function:

$$\Phi\left(-\frac{1}{\sqrt{\epsilon}} - \frac{\sqrt{\epsilon}}{\lambda(\epsilon)}x\right) \propto \exp(-x)$$

- Symmetry properties of the Ising model ($H = -J\sum_{\langle i,j\rangle} s_i s_j - h\sum_i s_i$):

$$E(\uparrow) = -E(\downarrow) = \frac{\partial H}{\partial s_i}$$

Detailed balance equation:

$$\Rightarrow \quad \frac{W(\downarrow\rightarrow\uparrow)}{W(\uparrow\rightarrow\downarrow)} = \frac{\Phi\left(-\frac{1}{\sqrt{\epsilon}} - \frac{\sqrt{\epsilon}\beta}{\lambda(\epsilon)}\frac{\partial H}{\partial s_i}\right)}{\Phi\left(-\frac{1}{\sqrt{\epsilon}} + \frac{\sqrt{\epsilon}\beta}{\lambda(\epsilon)}\frac{\partial H}{\partial s_i}\right)} = \exp\left[-\beta(E(\uparrow) - E(\downarrow))\right]$$

To I.: Is this is a Langevin equation for the Ising model?

$$s_i' = \text{sign}\left(s_i - \epsilon \frac{\beta}{\lambda(\epsilon)} \frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\tilde{\eta}_i\right)$$

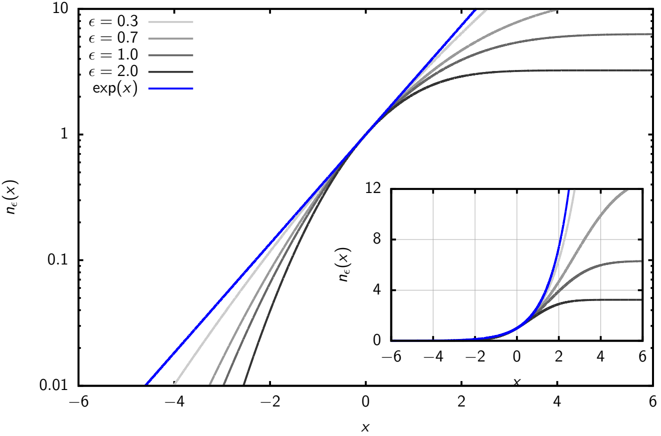
$\Rightarrow$ No!

$\Rightarrow$ It is a Langevin like MCMC algorithm with Gaussian noise input.

Did we learn something?

Result I: Generalised Relations

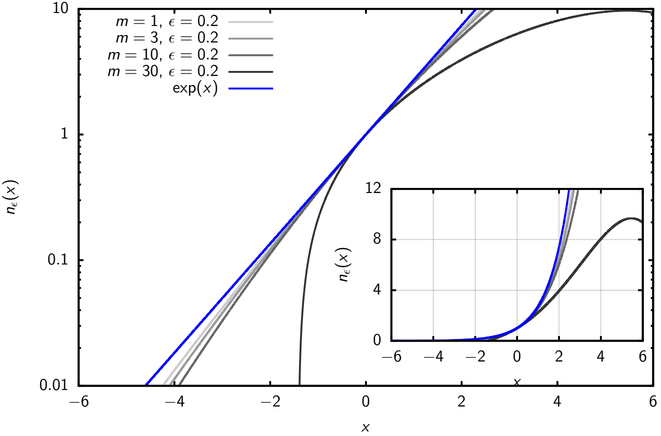
Cumulative Gaussian Distribution



$$\lim_{\epsilon \to 0} \frac{\Phi\left(-\frac{1}{\sqrt{\epsilon}} + \sqrt{\epsilon}\frac{x}{\lambda_\epsilon}\right)}{\Phi(-\frac{1}{\sqrt{\epsilon}})} = \exp(x) + \mathcal{O}(\epsilon x^2)\,, \quad \text{with:} \quad \lambda_\epsilon = \frac{\sqrt{\epsilon}\varphi\left(-\frac{1}{\sqrt{\epsilon}}\right)}{\Phi\left(-\frac{1}{\sqrt{\epsilon}}\right)}$$

Result I: Generalised Relations

Derivatives of the Cumulative Gaussian Distribution



$$\lim_{\epsilon \to 0} \frac{\frac{\partial^m}{\partial t^m} \Phi\left(-\frac{1}{\sqrt{\epsilon}} + \sqrt{\epsilon} t\right)\Big|_{t = -\frac{\mathrm{He}_m(-1/\sqrt{\epsilon})}{\sqrt{\epsilon}\mathrm{He}_{m-1}(-1/\sqrt{\epsilon})} x}}{\frac{\partial^m}{\partial t^m} \Phi(-\frac{1}{\sqrt{\epsilon}} + \sqrt{\epsilon} t)\Big|_{t=0}} = \exp(x) + \mathcal{O}(\epsilon x^2)$$

Result II: MCMC Algorithm based on Gaussian Noise

Generalised update rule:

$$s_i' = s_i + (\nu - s_i)\Theta\left[-1 - \frac{\epsilon\beta}{2\lambda(\epsilon)}\Delta E(\nu, s_i) + \sqrt{\epsilon}\eta_i^T\right]$$

with a proposal state ν.

For the Ising model this is equivalent to:

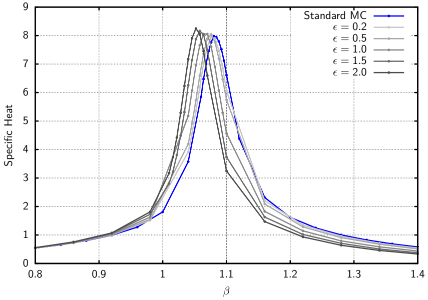
$$\Leftrightarrow \quad s_i' = \text{sign}\left(s_i - \epsilon\frac{\beta}{\lambda(\epsilon)}\frac{\partial H}{\partial s_i} + \sqrt{\epsilon}\tilde{\eta}_i\right)$$

Application I: Numerical Results for Other Models

q-Potts model:

$$H_p = -J_p \sum_{\langle i,j \rangle} \delta_{s_i, s_j} \ ,$$

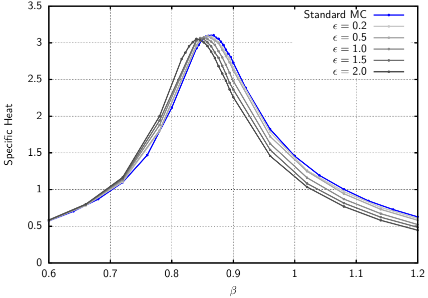
with $s_i \in \{1, 2, \dots, q\}$.



Clock model:

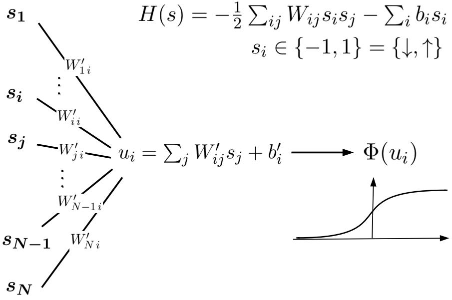
$$H_c = -J_c \sum_{\langle i,j \rangle} \cos\left(\theta_i - \theta_j\right) \ ,$$

with $\theta_i = \frac{2\pi n}{q})$ and $n \in \{1, 2, \dots, q\}$.



Application II: Langevin Machine

Scheme:



Identifications:

- $W'_{ii} = \frac{1}{\sqrt{\epsilon}}$
- $b'_i = \frac{\sqrt{\epsilon}}{\lambda(\epsilon)} b_i$
- $W'_{ij} = \frac{\sqrt{\epsilon}}{\lambda(\epsilon)} W_{ij}$

Implicit update rule:

$$s'_i = \text{sign}(u_i + \tilde{\eta}_i)$$

Activation function: $p_i(\uparrow) = \dfrac{1}{1 + \exp(2H_i(\uparrow))}$

$\Rightarrow$ Alternative implementation of the Boltzmann machine with a different update dynamic and self interaction

Does this help for a computation on the neuromorphic hardware of the BrainScaleS project?

$\Rightarrow$ No! Still different dynamics.

Application III: Modified Ornstein-Uhlenbeck Process

Ornstein-Uhlenbeck process (free membrane potential):

$$\frac{ds_i}{dt} = \theta(\mu - s_i) + \sigma\eta(t)$$

Resulting activation function for $\theta = 1$ and $\sigma = 1$:

$$p(\uparrow) = p(s_i \geq 0) = \int_0^\infty p(s_i) ds_i = \Phi(\mu)$$

Desired activation function:

$$p(\uparrow) = p(s_i \geq 0) = \frac{1}{1 + \exp(-2\mu)}$$

Application III: Modified Ornstein-Uhlenbeck Process

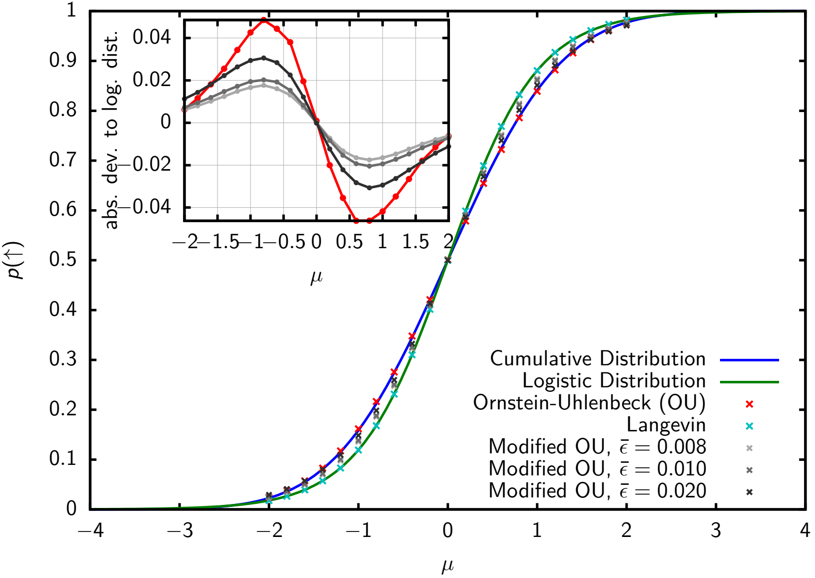
Modified Ornstein-Uhlenbeck process (free membrane potential):

$$\frac{ds_i}{dt} = \theta \left(\frac{\sqrt{\bar{\epsilon}}}{\lambda(\bar{\epsilon})} \mu - s_i + \frac{\mathrm{sign}(s_i)}{\sqrt{\bar{\epsilon}}} \right) + \sigma \eta(t)$$

Resulting activation function for $\theta = 1$ and $\sigma = 1$:

$$\lim_{\bar{\epsilon} \to 0} p(\uparrow) = \lim_{\bar{\epsilon} \to 0} p(s_i \geq 0) = \frac{1}{1 + \exp(-2\mu)}$$

Application III: Modified Ornstein-Uhlenbeck Process



Conclusion

Results:

- Limit between the cumulative Gaussian distribution and the exponential function
- A Langevin like MCMC algorithm based on Gaussian noise for discrete systems
- Langevin machine
- Modified Ornstein-Uhlenbeck process

Future work:

- Find useful applications (other neuromorphic systems)
- Investigate interactions and further aspects of the neuromorphic hardware