Gaussian Approximation of Transition Paths

Andrew Stuart

Mathematics Institute, University of Warwick

COSMOS, Paris, February 2\textsuperscript{nd} – 5\textsuperscript{th} 2016

Collaboration with Yulong Lu and Hendrik Weber
Also Frank Pinski, Gideon Simpson, Florian Theil

Funded by EPSRC, ERC and ONR
Outline

1. Chemical Reactions
2. Transition Paths Overview
3. Best Gaussian Approximation
4. Low Temperature Limit
5. Conclusions
Outline

1. Chemical Reactions
2. Transition Paths Overview
3. Best Gaussian Approximation
4. Low Temperature Limit
5. Conclusions
Brownian Dynamics

- $V : \mathbb{R}^d \rightarrow \mathbb{R}^+$.  
- $\mathcal{E} = \{ x \in \mathbb{R}^d : \nabla V(x) = 0 \}$.  
- $x^\pm \in \mathcal{E}$.  
- $\varepsilon > 0$ temperature.  
- $T > 0$ transition time.

$$\frac{dx}{dt} = -\nabla V(x) + \sqrt{2\varepsilon} \frac{dW}{dt},$$

$x(0) = x^-$,  
$x(T) = x^+$.  

Mathematical Statement
Example I: Vacancy Diffusion

Example II: Two Routes

[2] F. Pinski and A.M. Stuart,
Transition paths in molecules at finite temperature,
Outline

1. Chemical Reactions
2. Transition Paths Overview
3. Best Gaussian Approximation
4. Low Temperature Limit
5. Conclusions
Probabilistic Formulation

\[
\frac{dx}{dt} = -\alpha \nabla V(x) + \sqrt{2\epsilon} \frac{dW}{dt},
\]
\[x(0) = x^-, \quad x(T) = x^+.\]

Path Space Measures

- \(\mu_\alpha\) the induced probability measure on \(C([0, T]; \mathbb{R}^d)\).
- \(\mu_0\) is Brownian bridge.
- \(\mu := \mu_1\) is measure of interest.
## Orientation

### Key Questions

- What is the **most likely path** connecting $x_-$ and $x_+$ at finite temperature $\varepsilon > 0$?
- In the **low temperature limit** $\varepsilon \to 0$ what happens to the most likely path?

### Our Approach

- Characterize **best Gaussian approximation** to $\mu$.
- Use $\Gamma$– **convergence** to study the small temperature limit.
Large Deviations

- Let $B_{\delta}(\varphi)$ be a ball of radius $\delta$ in $C([0, T]; \mathbb{R}^n)$ centred at $\varphi$.
- For fixed $T$:

$$\mathbb{P}^\mu(B_{\delta}(\varphi)) \approx \exp \left( -\frac{1}{2\varepsilon} \mathcal{S}_T(\varphi) \right).$$

- The **action functional** is

$$\mathcal{S}_T(\varphi) := \frac{1}{2} \int_0^T (\varphi'(t) + \nabla V(\varphi))^2 \, dt.$$ 

- The most likely paths are thus minimizers of

$$S_T(\varphi) := \frac{1}{2} \int_0^T |\varphi'(t)|^2 + |\nabla V(\varphi(t))|^2 \, dt.$$
Onsager-Machlup Theory [3]

Maximizing Small Ball Probabilities

- Let \( B_\delta(\varphi) \) be a ball of radius \( \delta \) in \( C([0, T]; \mathbb{R}^n) \) centred at \( \varphi \).
- Then \( \lim_{\delta \to 0} \frac{\mathbb{P}(B_\delta(\varphi_1))}{\mathbb{P}(B_\delta(\varphi_2))} = \exp(I_{\varepsilon, T}(\varphi_2) - I_{\varepsilon, T}(\varphi_1)) \).
- Here \( I_{\varepsilon, T} \) is the Onsager-Machlup functional defined by
  \[
  I_{\varepsilon, T}(\varphi) := S_T(\varphi) - \int_0^T \varepsilon \Delta V(\varphi(t)) \, dt.
  \]
- Probability of the small ball is maximized when centred at minimizers of \( I_{\varepsilon, T} \) (MAP estimator in statistics).
- The Itô correction \( \varepsilon \Delta V \) may produce non-physical transition paths, in contrast to large deviation approach.
Example I – Vacancy Diffusion

Onsager-Machlup Minimizers [2]

- Left: zero temperature.
- Right: finite temperature.
- Saddle is preferred.
Example II – Two Routes

Onsager-Machlup Minimizers [2]

- Left: two routes.
- Middle: straight route.
- Right: curved route.
- \( \varepsilon = 0 \) (black), \( \varepsilon = 10^{-2} \) (blue), \( \varepsilon = 10^{-1} \) (red).
- Low entropy route preferred.
Formulation With $T = \varepsilon^{-1}$, $t \mapsto \varepsilon t$ [3]

\[
\frac{dx}{dt} = -\frac{\alpha}{\varepsilon} \nabla V(x) + \sqrt{2} \frac{dW}{dt},
\]

\[x(0) = x^-, \quad x(1) = x^+.\]

Thus measure of interest $\mu$ has density with respect to the Brownian bridge $\mu_0$ which is given by:

\[
\mu(dx) = Z^{-1} \exp\left(-\frac{1}{2\varepsilon^2} \Phi_\varepsilon(x)\right) \mu_0(dx),
\]

\[Z = \mathbb{E}^{\mu_0} \exp\left(-\frac{1}{2\varepsilon^2} \Phi_\varepsilon(x)\right),\]

\[
\Phi_\varepsilon(x) = \int_0^1 \left(\frac{1}{2} |\nabla V(x(t))|^2 - \varepsilon \Delta V(x(t))\right) dt.
\]
Kullback-Leibler Approximation

For $\nu, \mu$ probability measures define the K-L divergence

\[
D_{KL}(\nu \| \mu) = \mathbb{E}^\nu \log \left( \frac{d\nu}{d\mu} \right) \text{ if } \nu \ll \mu
\]

\[= \infty \text{ otherwise.}\]

Theorem [4]

- Let $\mu \ll \mu_0 = N(m, C)$.
- Let $\mathcal{A}$ denote the set of all Gaussians equivalent to $\mu_0$.
- Then there is $\nu \in \mathcal{A}$ that minimizes $D_{KL}(\nu \| \mu)$.

Approximation by Inhomogeneous OU Processes [6]

\[ dz(t) = -\varepsilon^{-1} A(t) z(t) + \sqrt{2} dW(t), \]
\[ z(0) = z(1) = 0. \]

Approximation Class (in dimension \( d = 1 \) for exposition)

- If \( C = 2(-\Delta + B_{\varepsilon})^{-1} \), where \( C^{-1} \) has domain \( H^2 \cap H^1_0 \) and \( B_{\varepsilon} = \varepsilon^{-2} A^2 - \varepsilon^{-1} A' \), then \( z \sim N(0, C) \).
- Here we assume that, for some \( a > 0 \),

\[ m \in H^1_\pm(0, 1) := \{ x \in H^1(0, 1) : x(0) = x^-, x(1) = x^+ \}, \]
\[ A \in H^1_a := \{ u \in H^1(0, 1) : u \geq a \text{ a.e.} \}. \]

- Define \( \mathcal{A} := \{ x = m + z, m \in H^1_\pm(0, 1), z \sim N(0, C) \} \).
- We aim to find \( \nu \in \mathcal{A} \) to minimize \( D_{KL}(\nu \| \mu) \).
Calculation of KL Divergence

Let $\bar{A}(t) = \int_{t}^{1} A(s)ds$. Then, in dimension $d = 1$ for exposition,

\[
D_{KL}(\nu \| \mu) = \frac{1}{2\epsilon} F_{\epsilon}(m, A) + \text{const}
\]

\[
F_{\epsilon}(m, A) = \frac{\epsilon}{2} \int_{0}^{1} |m'(t)|^2 dt + \frac{1}{\epsilon} \mathbb{E}^\nu \Phi_{\epsilon}(m + z)
\]

\[
- \frac{1}{4} \mathbb{E}^\nu \int_{0}^{1} B_{\epsilon}(t)(z(t))^2 dt + \frac{1}{2\epsilon} \int_{0}^{1} A(t) dt
\]

\[
+ \frac{1}{2} \log \left( \int_{0}^{1} e^{-2\bar{A}(t)/\epsilon} dt \right)
\]

$F_{\epsilon}$ is an entropically fattened large deviations rate function.
Let $\bar{A}(t) = \int_t^1 A(s) \, ds$. Then, in dimension $d = 1$ for exposition,

$$D_{KL}(\nu \| \mu) = \frac{1}{2\epsilon} F_\epsilon(m, A) + \text{const}$$

$$F_\epsilon(m, A) = \frac{\epsilon}{2} \int_0^1 |m'(t)|^2 \, dt + \frac{1}{\epsilon} \mathbb{E}^{\nu} \Phi_\epsilon(m + z)$$

$$- \frac{1}{4} \mathbb{E}^{\nu} \int_0^1 B_\epsilon(t)(z(t))^2 \, dt + \frac{1}{2\epsilon} \int_0^1 A(t) \, dt$$

$$+ \frac{1}{2} \log \left( \int_0^1 e^{-2\bar{A}(t)/\epsilon} \, dt \right)$$

$F_\epsilon$ is an entropically fattened large deviations rate function.
Calculus of Variations

**Theorem (Regularized $D_{KL}$ Minimization) [6]**

For $m \in H^1_\pm, A \in H^1_a$, and for any $\gamma > 0$, define

$$J_\varepsilon(m, A) = F_\varepsilon(m, A) + \varepsilon^\gamma \|A\|_{H^1}^2.$$ 

Assume that $\inf_{\nu \in A} D_{KL}(\nu \| \mu) < \infty$, then there exists $m_\varepsilon \in H^1_\pm, A_\varepsilon \in H^1_a$ that minimizes $J_\varepsilon(m, A)$.

**Remarks**

- **Regularization** is applied to obtain compactness w.r.t. $A$.
- Optimal transition paths characterized by a **Gaussian tube** centered at $m$ with variance of order $\varepsilon$. 

Outline

1. Chemical Reactions
2. Transition Paths Overview
3. Best Gaussian Approximation
4. Low Temperature Limit
5. Conclusions
Questions

- When $\varepsilon \to 0$: what is limit $(m, A)$ of $(m_\varepsilon, A_\varepsilon)$?
- what is the limiting functional $F$ of $F_\varepsilon$?
- does $(m, A)$ minimize $F$?

Answer

- Study convergence of functionals by $\Gamma$-convergence.
- Compactness + $\Gamma$-convergence implies convergence of minima.
- Compactness means

\[
\limsup_{\varepsilon \to 0} F_\varepsilon(u_\varepsilon) < \infty \quad \text{implies} \quad \exists u \text{ such that } u_\varepsilon \to u.
\]
Theorem [6]

Under conditions on $V$ and if $\gamma \in (0, \frac{1}{2})$, then the $\Gamma$-limit of $F_{\varepsilon}$ on $L^1(0, 1) \times L^1_a(0, 1)$ is

$$F(m, A) := E(m) + \int_0^1 \frac{1}{4A(t)} (V''(m(t)) - A(t))^2 \, dt$$

where

$$E(m) := \begin{cases} \sum_{\tau \in J(m)} \psi(m(\tau-), m(\tau+)) & \text{if } m \in BV_\pm (0, 1; \mathcal{E}), \\ +\infty & \text{otherwise in } L^1(0, 1). \end{cases}$$

Here $J(m)$ is the jump set of $m$ and $\psi$ is the energy cost for each jump, namely for $x_1, x_2 \in \mathcal{E},$

$$\psi(x_1, x_2) = |V(x_2) - V(x_1)|.$$
Sketch Proof

Expanding in $\varepsilon$ and using $\mathbb{E}z^2(t) = \mathbb{E}[x(t) - m(t)]^2 \approx \varepsilon/A(t)$:

$$F_\varepsilon(m_\varepsilon, A_\varepsilon) \approx \frac{\varepsilon}{4} \int_0^1 m'_\varepsilon(t)^2 \, dt + \frac{1}{4\varepsilon} \int_0^1 V'(m_\varepsilon(t))^2 \, dt$$

$$+ \int_0^1 \frac{1}{4A_\varepsilon(t)} \left( (V''(m_\varepsilon(t)) - A_\varepsilon(t))^2 + V'(m_\varepsilon(t)) V^{(3)}(m_\varepsilon(t)) \right) \, dt.$$  

Note that:

$$F_\varepsilon(m_\varepsilon, A_\varepsilon) \approx \frac{1}{2} S_{\varepsilon^{-1}}(m_\varepsilon) + \int_0^1 \frac{1}{4A_\varepsilon(t)} (V''(m_\varepsilon(t)) - A_\varepsilon(t))^2$$

$$+ V'(m_\varepsilon(t)) V^{(3)}(m_\varepsilon(t)) \, dt.$$
Sketch Proof

\[ F_\varepsilon(m_\varepsilon, A_\varepsilon) \approx \frac{1}{2} S_{\varepsilon^{-1}}(m_\varepsilon) + \int_0^1 \frac{1}{4A_\varepsilon(t)} \left( V''(m_\varepsilon(t)) - A_\varepsilon(t) \right)^2 \]
\[ + V'(m_\varepsilon(t)) V^{(3)}(m_\varepsilon(t)) dt. \]

Remarks

- The first term reflects large deviations, depends only on the mean, and minimizers are well understood, using \( \Gamma \)-convergence, in the small temperature limit.

- The Itô correction term can be killed by setting \( A_\varepsilon = V''(m_\varepsilon) \): fluctuation determined by linearization of the original dynamics around mean path.

- \( V'(m_\varepsilon) V^{(3)}(m_\varepsilon) \) vanishes in the limit since the large deviations energy functional forces \( V'(m_\varepsilon) \approx 0. \)
Outline

1. Chemical Reactions
2. Transition Paths Overview
3. Best Gaussian Approximation
4. Low Temperature Limit
5. Conclusions
Conclusions

- Large deviations predicts **physical** transition paths.
- Onsager-Machlup theory (MAP estimator) produces **non-physical** transition paths.
- Entropy is key to understanding this dichotomy.
- “Best Gaussian” approximation w.r.t. Kullback-Leibler.
- Gaussian approximation explicitly incorporates **entropy**.
- Characterize most likely transition paths as optimal Gaussian tubes around large deviation mean.
- Gaussian tube is defined by **OU fluctuations** involving linearization at critical points.
- In the low temperature limit, Kullback-Leibler approximation approach removes undesirable Itô correction in Onsager-Machlup and recovers large deviations.


