Rearranged Stochastic Heat Equation

William Hammersley

The 9th International Colloquium on BSDEs and Mean Field Systems
27th June 2022
Annecy, France

Joint Work with François Delarue

Both at: Laboratoire J.A. Dieudonné, Université Côte d'Azur
3IA - Interdisciplinary Instute for Artificial Intelligence

Motivation

Aim to *regularise by noise* the gradient flows in the space of probability measures that arise from mean field models.

$$dX_t^{i,N} = f(X_t^{i,N}, \mu_t^N)dt \overset{N\to\infty}{\to} \quad dX_t = f(X_t, \mu_t)dt$$

$$\mu_t^N := \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{i,N}} \qquad \to \qquad \mu_t := \mathbb{P} \circ X_t^{-1} =: \mathscr{L}(X_t)$$

The particular desire is that this noise is *intrinsic*.

Related to the study of diffusions in the space of measures, eg: Wasserstein Diffusion [7, 11], Fleming-Viot Process [12].

Motivation

Two perspectives:

- Introducing noise at the level of the particle system.

$$
\begin{aligned}
dX_t^{i,N} =& f(X_t^{i,N}, \mu_t^N)dt + \sigma dM_t^{i,N} \overset{N\to\infty}{\to} dX_t = f(X_t, \mu_t)dt + \sigma dM_t \\
\mu_t^N :=& \frac{1}{N}\sum_{i=1}^{N} \delta_{X_t^{i,N}} \qquad\qquad \to \qquad \mu_t := \mathscr{L}(X_t | \mathcal{G}_t)
\end{aligned}
$$

  The noise introduced must be exchangeable to respect the mean field symmetry.

- Alternatively, one may consider the formal mean-field limit dynamics for the original particle system and ask what regularisation one would like to see. Then, one may hope to reverse-engineer the particle systems with this regularised limit.

Motivation

Consider the particles live in $\mathbb{R}$ (dimension$= 1$!) and that one expects solutions to have square integrable marginals. Instead of thinking to directly approach randomising the FPK, we lift to the Hilbert-space L^2 and randomise in the following way:

We fix the underlying probability space of the McKean-Vlasov distribution dependent SDE and randomise the resulting PDE. Consider the unit circle equipped with Lebesgue measure $(\Omega, \mathbb{P}) := (\mathbb{S}, \mathrm{Leb})$,

$$dX(\omega)_t = f(X(\omega)_t, \mu_t)dt \overset{\text{rewrite}}{\to} dX(x)_t = f(X(x)_t, \mathrm{Leb} \circ X_t^{-1}(x))dt$$
$$\mu_t := \mathbb{P} \circ X_t^{-1}$$

Regularisation by Noise

$$dX(x)_t = F(X(x)_t, Leb \circ X_t^{-1})dt + \Delta_x X(x)_t dt + dW(x)_t$$

with periodic boundary conditions.

$$W(t,x) := B_t^0 \cdot 1 + \sum_{m \in \mathbb{N}} \sqrt{2}\left(B_t^{m,+} cos(2\pi mx) + B_t^{m,-} sin(2\pi mx)\right) =: \sum_{m \in \mathbb{Z}} B_t^m e_m(x)$$

Expecting solutions to the above dynamics to be $L^2(\mathbb{S})$ valued; the Laplacian enables this.

Regularisation by Noise

$$dX(x)_t = F(X(x)_t, Leb \circ X_t^{-1})dt + \Delta_x X(x)_t dt + dW(x)_t$$

with periodic boundary conditions.

$$W(t,x) := B_t^0 \cdot 1 + \sum_{m \in \mathbb{N}} \sqrt{2}\left(B_t^{m,+} cos(2\pi mx) + B_t^{m,-} sin(2\pi mx)\right) =: \sum_{m \in \mathbb{Z}} B_t^m e_m(x)$$

Expecting solutions to the above dynamics to be $L^2(\mathbb{S})$ valued; the Laplacian enables this.

Dynamics may change depending on the choice of representation of the initial distribution!

Regularisation by Noise

$$dX(x)_t = F(X(x)_t, Leb \circ X_t^{-1})dt + \Delta_x X(x)_t dt + dW(x)_t$$

with periodic boundary conditions.

$$W(t,x) := B_t^0 \cdot 1 + \sum_{m \in \mathbb{N}} \sqrt{2}\Big(B_t^{m,+} cos(2\pi mx) + B_t^{m,-} sin(2\pi mx)\Big) =: \sum_{m \in \mathbb{Z}} B_t^m e_m(x)$$

Expecting solutions to the above dynamics to be $L^2(\mathbb{S})$ valued; the Laplacian enables this.

Dynamics may change depending on the choice of representation of the initial distribution!

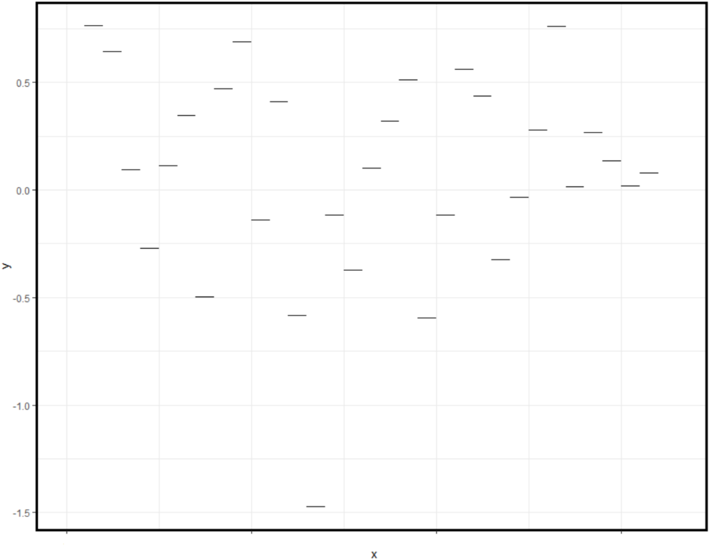
The idea is to constrain the above SPDE to live within a particular set of functions that correspond to the space of probability measures.

Quantile Function Representation

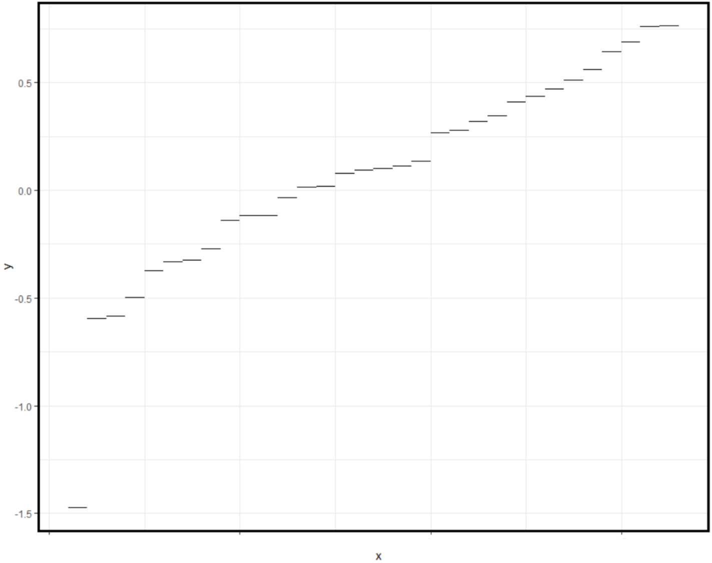
One could choose to represent the starting measure through its quantile, since there exists a unique non-decreasing function Q such that $\mu_0 = \mathrm{Leb}_{\mathbb{S}} \circ Q^{-1}$.

$$Q(u) := \inf \Big\{ x \in \mathbb{R} : u \leq \mu_0\big((-\infty, x]\big) \Big\}$$

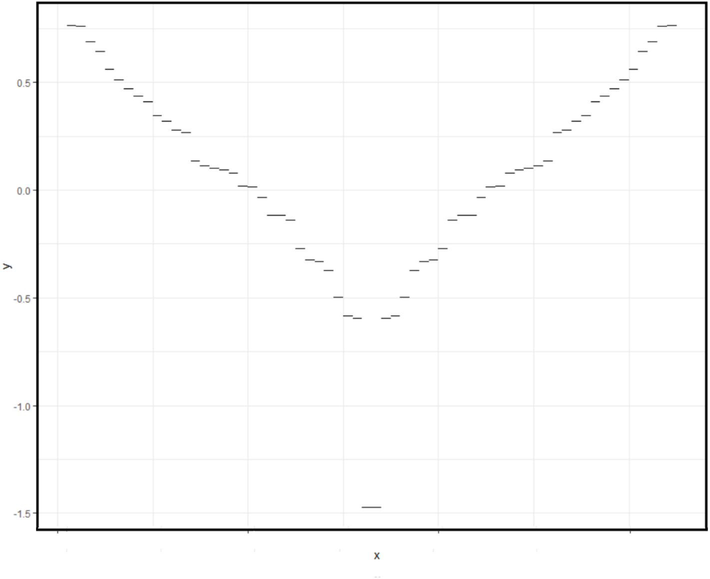
A Simple Random Variable



Quantile Function



Symmetric Decreasing Rearrangement



Scheme

We will work with the subset of $L^2(\mathbb{S})$ comprised of symmetric non-increasing functions, denoted $L^2_{sym}(\mathbb{S})$.

Consider the following scheme to construct a noise on L^2_{sym}. (The rearrangement of f is denoted f^*).

1. Initialise with the symmetric decreasing representative of the initial distribution:
$$X_0 = X_0^* \sim \mu_0.$$
2. Solve on some small time interval the stochastic heat equation:
$$\hat{X}_{\delta t} = SHE(X_0, \delta t).$$
3. Rearrange the terminal state:
$$X_{\delta t} = \hat{X}^*_{\delta t}.$$
4. Interpolate linearly between iterates.

Play two-step illustration

Play the discrete time scheme

The Scheme

For intervals of length h, define the sequences $\{X_n^h\}_{n\in\mathbb{N}}$ by,

$$X_{n+1}^h = \left(e^{h\Delta} X_n^h + \int_0^h e^{(h-s)\Delta} d\tilde{W}_s^{n+1}\right)^*$$
$$X_0^h = X_0 \equiv X_0^*,$$

The noise we construct is the limiting process as the time mesh becomes ever finer.

We are able to prove uniform estimates and tightness, but are required to penalise the amount of noise we add into the system. This makes sense since the operator preserves the L^2 norm, yet moves weight towards lower Fourier modes.

Coloured Noise

Instead of the cyclindrical Brownian motion,

$$W(t,u) = \sum_{m\in\mathbb{Z}} B_t^m e_m(u)$$

we take the coloured noise:

$$\tilde{W} := B^0 e_0 + \sum_{m\in\mathbb{N},\, n\neq 0} m^{-\lambda} B^m e_m \equiv \sum_{m\in\mathbb{N}_0} \lambda_m B^m e_m.$$

It is assumed that λ_m defines a square-summable sequence, i.e. $\lambda > \frac{1}{2}$.

Whilst the noise is now in L^2, retain the Laplacian for its smoothing effect.

Rearrangement Theorem

We use fundamental results from the theory of rearrangements.

Preservation of L_p norms

$$\|f^*\|_p = \|f\|_p$$

Re-arrangement doesn't harm our estimates, however one needs to unravel the scheme through each interval in the time-mesh.

Non Expansive Property

$$\|f^* - g^*\|_p \leq \|f - g\|_p$$

NB: Non-expansion also holds for projections onto convex sets!

Rearrangement Theorem

Riesz Rearrangement Inequality

$$\iint f(x)g(x-y)h(y)dxdy \leq \iint f^*(x)g^*(x-y)h^*(y)dxdy$$

Here, we see an instance where using heat is convenient since the periodic heat kernel is indeed symmetric non-increasing

Pólya–Szegő Inequality

$$\|\nabla f^*\|_p \leq \|\nabla f\|_p$$

These theorems are usually stated over $\mathbb{R}^d$ for non-negative functions. Fortunately, this restriction is not required in the case of the circle.

Baernstein: Symmetrization in Analysis [2].

Limiting Dynamics

We expect that the limit process X should satisfy a reflected equation of the type considered in Röckner, Zhu and Zhu [10],

$$dX_t = \Delta X_t + \tilde{W}_t + \eta_t.$$

Such an equation was studied for the stochastic heat equation constrained to live above some fixed/static boundary in a series of papers by Donati-Martin, Nualart, Pardoux and Zambotti, [6, 9, 13].

For us, η is a forcing term that reflects the process X into the cone of symmetric decreasing functions in $L^2(\mathbb{T})$. However, in the absence of a corresponding integration by parts formula, the dynamics will be studied via a smaller class of test functions, sufficient to demonstrate the well-posedness and regularising effect of the limiting dynamics and ergodicity for a class of 'drifts'.

Limiting Dynamics

Taking inspiration from works of Brenier [3], we are able to identify the following equations for the limit process:

for $\varphi \in C^2$, $s, t \in \mathbb{R}_+$,

$$\langle X_t - X_s, \varphi \rangle = \int_s^t \langle X_r, \Delta\varphi \rangle dr + \langle \eta_t - \eta_s, \varphi \rangle + \langle \tilde{W}_t - \tilde{W}_s, \varphi \rangle, \quad \mathbb{P}\text{-a.s.}$$

Note that for symmetric decreasing φ, $\langle \eta_t, \varphi \rangle$ is non-decreasing.

Since we want to study, say for uniqueness:

$$\|X_t - Y_t\|_2^2 = \sum_{m\in\mathbb{Z}} \langle X_t - Y_t, e_m \rangle^2 = \sum_{m\in\mathbb{Z}} \left(\widehat{X_t - Y_t}^m \right)^2$$

From the Itô-formula, we need to define in some manner,

$$\sum_{m\in\mathbb{Z}} \int_s^t \langle X_r - Y_r, e_m \rangle d \langle \eta_r^X - \eta_r^Y, e_m \rangle$$

Limiting Dynamics

However, the integrals are only established, thus far, against symmetric decreasing test functions. Nevertheless, one can decompose each element of the Fourier basis into a difference of two symmetric decreasing functions.

Indeed, for $\{e_m\}_{m\in\mathbb{Z}}$, the Fourier basis, each may be written as the difference of two symmetric non-increasing functions:

$$e_m^+(x) := e_m(0) + \int_0^1 \mathbb{1}_{0,x}(y)\left[-\mathbb{1}_{0,1/2}(De_m)_- + \mathbb{1}_{1/2,1}(De_m)_+\right](y)dy,$$
$$e_m^-(x) := \int_0^1 \mathbb{1}_{0,x}(y)\left[-\mathbb{1}_{0,1/2}(De_m)_+ + \mathbb{1}_{1/2,1}(De_m)_-\right](y)dy.$$

Indeed, the functions e_m^+ and e_m^- are symmetric decreasing (courtesy of the symmetry properties of e_m) and $e_m = e_m^+ - e_m^-$.

Limiting Dynamics

Therefore, one may write the Riemann-Stieltjes integral,

$$\int_s^t \langle Z_r, e_m\rangle \cdot d\langle \eta_r, e_m\rangle := \int_s^t \langle Z_r, e_m\rangle \cdot d\langle \eta_r, e_m^+\rangle - \int_s^t \langle Z_r, e_m\rangle \cdot d\langle \eta_r, e_m^-\rangle.$$

Ultimately, we are able to define

$$\int_s^t Z_r \cdot d\eta_r := \lim_{M\to\infty} \sum_{m^2 \leq M} \int_s^t \langle Z_r, e_m\rangle \cdot d\langle \eta_r, e_m\rangle \geq 0$$

for sufficiently regular processes Z, valued in $L^2_{sym}(\mathbb{S})$.

Limiting Dynamics

We identified three conditions satisfied by our limit processes:

$$\langle X_t - X_s, e^{\varepsilon\Delta} e_m \rangle = \int_s^t \langle X_r, \Delta e^{\varepsilon\Delta} e_m \rangle dr + \langle \eta_t - \eta_s, e^{\varepsilon\Delta} e_m \rangle + \langle \tilde{W}_t - \tilde{W}_s, e^{\varepsilon\Delta} e_m \rangle,$$

$$\int_s^t e^{\varepsilon\Delta} Z_r \cdot d\eta_r \geq 0,$$

and

$$\liminf_{\varepsilon \searrow 0} \int_0^t e^{\varepsilon\Delta} X_r \cdot d\eta_r = 0.$$

Limiting Dynamics - Uniqueness

Considering two candidate limit processes, labelled Y and Z.

$$
\begin{aligned}
&d\langle e^{\varepsilon\Delta}(Y_t - Z_t), e_m\rangle^2 \\
&\quad = 2\langle e^{\varepsilon\Delta}(Y_t - Z_t), e_m\rangle \left[\frac{1}{2}\langle \Delta e^{\varepsilon\Delta}(Y_t - Z_t), e_m\rangle dt + d\langle \eta_t^Y - \eta_t^Z, e^{\varepsilon\Delta} e_m\rangle\right]
\end{aligned}
$$

Writing the above in integral form and summing over the Fourier modes m, one obtains

$$
\begin{aligned}
&\|e^{\varepsilon\Delta}(Y_t - Z_t)\|_2^2 - \|e^{\varepsilon\Delta}(Y_0 - Z_0)\|_2^2 \\
&\quad = \int_0^t \langle e^{\varepsilon\Delta}(Y_r - Z_r), \Delta e^{\varepsilon\Delta}(Y_r - Z_r)\rangle dr + 2\int_0^t e^{2\varepsilon\Delta}(Y_r - Z_r)\cdot d(\eta_r^Y - \eta_r^Z) \\
&\quad \leq 2\left(\int_0^t e^{2\varepsilon\Delta} Y_r \cdot d\eta_r^Y + \int_0^t e^{2\varepsilon\Delta} Z_r \cdot d\eta_r^Z\right)
\end{aligned}
$$

Applying expectation and setting $\varepsilon \to 0$, one obtains that $Y_t = Z_t$ $\mathbb{P} - a.s.$ for any $t \in I$, uniqueness follows from the continuity of the processes. Furthermore, there is 1-Lipschitz dependence on initial conditions.

Regularisation

Due to the presence of the term η, one does not expect to easily obtain the differentiability of the solution in the initial condition, see Deuschel and Zambotti [5]. Nonetheless, it is still possible to estimate the directional derivatives of the semigroup flow.

By adapting to this setting a method of Norris [8], we are able to estimate the directional derivatives of the semigroup $\{P_t\}_{t\in I}$ defined by $P_t f(x) := \mathbb{E}[f(X_t^x)]$ for $f \in B_b(L^2)$.

This method uses Girsanov theorem and the previous established Lipschitz dependence on initial conditions. We truncate the Fourier modes of the initial values in order to apply the Rademacher theorem on almost everywhere differentiability of Lipschitz functions.

Regularisation

Considering the difference, for $u, v \in L^2_{sym}$,

$$P_T f(u + \delta v) - P_T f(u) = \mathbb{E}[f(X_T^{u+\delta v})] - \mathbb{E}[f(X_T^u)].$$

We reproduce the law of the process started from $u + \delta v$ by the law of the *shifted* process $X_t^{u^M + \delta \frac{T-t}{T} v^M}$ under a change of measure. We do not have a well defined change of measure without the Laplacian!

Ultimately, we are able to show that the semigroup maps bounded functions f into Lipschitz functions with constant

$$c_{\|f\|_\infty} \cdot T^{-(\frac{1}{2} + \frac{\lambda}{2})}$$

Notably for $\lambda < 1$ this is integrable.

Ergodicity

At this stage we have the drift-less form of the rearranged stochastic heat equation.

One may introduce a drift via a Girsanov tranformation, bearing in mind that this drift must respect the regularity of the solution in order to be well defined. This may be seen also as coming from the colouring of the noise.

Assumption

The drift $B(x) := \bar{B}(x) - \omega\langle x, e_0\rangle e_0$*, where* $\omega > 0$ *and* $\bar{B}$ *satisfies*

$$\|\bar{B}\|_0 := \sup_{x \in L^2_{sym}(\mathbb{S})} \sum_{m \in \mathbb{Z}} \lambda_m^{-2} \langle \bar{B}(x), e_m\rangle^2 < \infty.$$

There exists a constant c_B *such that,*

$$\|\bar{B}(x) - \bar{B}(y)\|_2 \leq c_B \|x - y\|_2$$

$$
\begin{aligned}
\langle X_t - X_s, \varphi \rangle = \int_s^t \langle B(X_r), \varphi \rangle + \langle X_r, \Delta\varphi \rangle dr \\
+ \langle \tilde{W}_t - \tilde{W}_s - \int_s^t \langle B(X_r), \varphi \rangle dr, \varphi \rangle + \langle \eta_t - \eta_s, \varphi \rangle,
\end{aligned}
$$

$$
\int_s^t e^{\varepsilon\Delta} Z_r \cdot d\eta_r \geq 0,
$$

$$
\liminf_{\varepsilon \searrow 0} \int_0^t e^{\varepsilon\Delta} X_r \cdot d\eta_r = 0.
$$

Under the conditions stated previously, we have well-posedness of the dynamics and exponential convergence to the unique invariant measure. The exponential convergence adapts a coupling argument of Debussche, Ying and Tessitore [4].

Comments

- Corresponding particle system and McKean-Vlasov drift.
- Other bases/operators/quantile representations.
- Higher dimension.
- Application to MFG.

Thank you for listening!

Selected References I

[1] A. Baernstein II. “Convolution and rearrangement on circle”. In: *Complex Variables, Theory and Application: An International Journal* 12.1-4 (1989), pp. 33–37.

[2] A. Baernstein II. *Symmetrization in Analysis*. New Mathematical Monographs. Cambridge University Press, 2019.

[3] Y. Brenier. “L2 Formulation of Multidimensional Scalar Conservation Laws”. In: *Archive for Rational Mechanics and Analysis* 193 (2009), 1–19.

[4] A. Debussche, Y. Hu, and G. Tessitore. “Ergodic BSDEs under weak dissipative assumptions”. In: *Stochastic Processes and their Applications* 121.3 (2011), pp. 407–426.

[5] J.-D. Deuschel and L. Zambotti. “Bismut–Elworthy’s formula and random walk representation for SDEs with reflection”. In: *Stochastic Processes and their Applications* 115.6 (2005), pp. 907–925.

Selected References II

[6] C. Donati-Martin and E. Pardoux. "White noise driven SPDEs with reflection". In: *Probability Theory and Related Fields*. 95, 1-24 (1993).

[7] V. Marx. "A Bismut–Elworthy inequality for a Wasserstein diffusion on the circle". In: *Stochastics and Partial Differential Equations: Analysis and Computations* (2021).

[8] J. R. Norris. "Simplified Malliavin calculus". en. In: *Séminaire de probabilités de Strasbourg* 20 (1986), pp. 101–130.

[9] D. Nualart and E. Pardoux. "White noise driven quasilinear SPDEs with reflection.". In: *Probab. Th. Rel. Fields*. 93, 77–89 (1992).

[10] M. Röckner, R.-C. Zhu, and X.-C. Zhu. "The stochastic reflection problem on an infinite dimensional convex set and BV functions in a Gelfand triple". In: *The Annals of Probability* 40.4 (2012), pp. 1759 –1794.

[11] M.-K. von Renesse and K.-T. Sturm. "Entropic measure and Wasserstein diffusion". In: *The Annals of Probability* 37.3 (2009), pp. 1114 –1191.

Selected References III

[12] W. Stannat. "Long-time behaviour and regularity properties of transition semigroups of Fleming–Viot processes". In: *Probability Theory and Related Fields* 122 (2002), pp. 431–469.

[13] L. Zambotti. "A Reflected Stochastic Heat Equation as Symmetric Dynamics with Respect to the 3-d Bessel Bridge". In: *Journal of Functional Analysis* 180 (2001), pp. 195–209.

A Uniform Estimate

Lemma (Uniform $2p^{th}$ moment of L_2 norm of scheme)

For $p \geq 1$,

$$\mathbb{E}[\|X_n^h\|_2^{2p}] \leq (1 + c_p h)^n \left(\mathbb{E}\left[\|X_0^h\|_2^{2p}\right] + c_p nh\right).$$

Proof Sketch

$$\mathbb{E}\left[\|X_n^h\|_2^{2p}\right] \equiv \mathbb{E}\left[\left\|\left(e^{h\Delta}X_{n-1}^h + \int_0^h e^{(h-s)\Delta} d\tilde{W}_s^n\right)^*\right\|_2^{2p}\right]$$

preservation of L_p norms

$$= \mathbb{E}\left[\left\|e^{h\Delta}X_{n-1}^h + \int_0^h e^{(h-s)\Delta} d\tilde{W}_s^n\right\|_2^{2p}\right]$$

multinomial theorem, Hölder inequality, Young's inequality

$$\leq (1 + c_p h)\mathbb{E}\left[\left\|e^{h\Delta}X_{n-1}^h\right\|_2^{2p}\right] + c_p h^{1-p}\mathbb{E}\left[\left\|\int_0^h e^{(h-s)\Delta} d\tilde{W}_s^n\right\|_2^{2p}\right]$$

We require then, that

$$\mathbb{E}\left[\left\|\int_0^h e^{(h-s)\Delta} d\tilde{W}_s^n\right\|_2^{2p}\right] \leq c_p h^p$$

Tightness

The linear interpolation schemes are tight within the space of $L^2_{sym}(\mathbb{S})$ valued continuous paths.

Let $N_t \in \{\lceil \frac{t}{h} \rceil, \lfloor \frac{t}{h} \rfloor\}$ and $N_s \in \{\lceil \frac{s}{h} \rceil, \lfloor \frac{s}{h} \rfloor\}$.

$$\begin{aligned}
&\mathbb{E}\left[\|X^h_{N_t} - X^h_{N_s}\|_2^{2p}\right] \\
\leq &2^{2p-1}\mathbb{E}\left[\left\|X^h_{N_t} - e^{(N_t - N_s)h\Delta}X^h_{N_s}\right\|_2^{2p} + \left\|e^{(N_t - N_s)h\Delta}X^h_{N_s} - X^h_{N_s}\right\|_2^{2p}\right]
\end{aligned}$$

Non Expansive Property: $\|f^* - g^*\|_p \leq \|f - g\|_p$

$$\mathbb{E}\left[\left\|e^{h\Delta}X^h_{N_t-1} - e^{(N_t - N_s)h\Delta}X^h_{N_s}\right\|_2^{2p}\right] \leq \mathbb{E}\left[\left\|X^h_{N_t-1} - e^{(N_t - N_s - 1)h\Delta}X^h_{N_s}\right\|_2^{2p}\right]$$

Tightness

To estimate

$$\mathbb{E}\left[\left\|e^{(N_t-N_s)h\Delta}X^h_{N_s}-X^h_{N_s}\right\|_2^{2p}\right]$$

One needs to control the norm of the derivative since:

$$\left\|e^{(N_t-N_s)h\Delta}X^h_{N_s}-X^h_{N_s}\right\|_2 \leq 2(N_t-N_s)\sup_{n\in[N_t,N_s]}\left\|DX^h_n\right\|_2$$

However, we want to be able to initialise with any distribution in $\mathcal{P}_2(\mathbb{S})$. This requires careful analysis towards $t=0$.

With the following Lemma, we are able to retain some smoothing effect of the heat kernel when initial function is replaced by something in the pre-image of the rearrangement.

Lemma

Let U be uniformly distributed on $[0,1]$. Then,

$$\mathbb{E}[\|De^{hU\Delta}u^*\|_2^2] \leq \mathbb{E}[\|De^{hU\Delta}u\|_2^2].$$

Proof: Uses the aforementioned rearrangement inequalities.